Reliability: Theory and Applications

ELECTRONIC JOURNAL OF INTERNATIONAL GROUP ON RELIABILITY
JOURNAL IS REGISTERED IN THE LIBRARY OF THE U.S.CONGRESS

2nd International Scientific and Technical Conference 'Infocommunication Systems and Artificial Intelligence Technologies' (ICSAIT 2024)

December 4-5, 2024, AzTU, Baku

SELECTED ARTICLES

Special Issue 7 (83)
Volume 20

ISSN 1932-2321
MAY 2025

Edited by
Prof. Vilayat Valiyev
Prof. Subhan Namazov
Prof. Nizami Yusubov
Prof. Sylvio Simon

GNEDENKO FORUM
INTERNATIONAL GROUP ON RELIABILITY

2nd International Scientific and Technical Conference 'Infocommunication Systems and Artificial Intelligence Technologies' (ICSAIT 2024) December 4-5, 2024, AzTU, Baku

Selected peer-reviewed, extended articles based on abstracts presented at the 2nd International Scientific and Technical Conference “Infocommunication Systems and Artificial Intelligence Technologies” (ICSAIT 2024)

Edited by
Prof. Vilayat Valiyev
Prof. Subhan Namazov
Prof. Nizami Yusubov
Prof. Sylvio Simon

Preface

The 2nd International Scientific-Technical Conference on “Infocommunication Systems and Artificial Intelligence Technologies” (ICSAIT) was held on December 4–5, 2024, at the Azerbaijan Technical University (AzTU) in Baku, Azerbaijan. The conference was organized by AzTU (Azerbaijan); the Institute of Information Technology of the Ministry of Science and Education of the Republic of Azerbaijan (Azerbaijan); the Institute of Control Systems of the Ministry of Science and Education of the Republic of Azerbaijan (Azerbaijan); the National Defense University (Azerbaijan); the National Aerospace Agency (Azerbaijan); Istanbul Technical University (Turkey); the National Technical University “Kharkiv Polytechnic Institute” (Ukraine); and Gazi University (Turkey).

The activities of the conference were aimed at addressing the architectural concepts of new and next-generation communication networks and the development prospects of telecommunication, control systems, information security, radio-electronic, and aerospace systems based on artificial intelligence technologies. The discussions focused on modern scientific and technical solutions; the establishment of strategically effective service spheres in the current environment; the formalization of a unified information infrastructure in the industry-oriented communication sector; and the transformation model for the digital economy. Additionally, the conference emphasized the synthesis and advancement of the fundamental principles for transitioning to the Fourth Industrial Revolution, comprehensive analyses related to the efficient integration of information and communication technologies, and fostering collaboration between higher education institutions and telecommunication firms, as well as building connections with prominent scientific centers and industry leaders.

The participants of the conference represented various countries, including Azerbaijan, Turkey, Uzbekistan, the USA, India, Germany, Russia, Belarus, Poland, Ukraine, and Latvia. A total of 140 papers were selected for the ICSAIT 2024 conference from the submissions made through the freely accessible conference system created on the AzTU website. Employees from 44 institutions across different countries took part in the conference.

We extend our heartfelt gratitude to the co-chairs of ICSAIT 2024, the members of the International Program Committee and the Organizing Committee, the invited speakers, reviewers, and all participants for their valuable contributions to the successful organization of ICSAIT 2024. We would also like to express our special thanks to the staff of the Gnedenko Forum, publisher of the *Electronic Journal Reliability: Theory & Applications* (ISSN 1932-2321), for their significant support in publishing the proceedings of ICSAIT 2024.

With best regards,

Prof. Vilayat Valiyev, Chair of ICSAIT 2024

Prof. Nizami Yusubov, Executive Secretary of ICSAIT 2024

Committees

Co-Chairs of the Conference

Prof. Vilayat Valiyev

(Rector of the Azerbaijan Technical University, Baku, Azerbaijan)

Prof. Rasim Aliguliyev

(Full member of ANAS, Director General of the Institute of Information Technology of the Ministry of Science and Education of the Republic of Azerbaijan)

Prof. Ali Abbasov

(Full member of ANAS, Director General of the Institute of Control Systems of the Ministry of Science and Education of the Republic of Azerbaijan)

Prof. İsmail Koyuncu (Rector of the Istanbul Technical University, Turkey)

International Program Committee

Prof. Nurali Yusifbayli – (Vice-rector for educational affairs AzTU, Azerbaijan)
Prof. Subhan Namazov - (Vice-rector for science and innovation, AzTU, Azerbaijan)
Prof. Ramiz Aliguliyev – (Corresponding Member of ANAS, Institute of Information Technologies, Azerbaijan)
Prof. Natig Javadov – (National Aerospace Agency, Azerbaijan)
Gunduz Abdulov – (National Defence University, Azerbaijan)
Prof. Arif Hasanov – (Military Scientific Research Institute of the National Defence University)
Ph.D. Elkhan Sabziyev – (Institute of Control Systems, Azerbaijan)
Ph.D. Rashid Alekperov – (Institute of Information Technologies, Azerbaijan)
Prof. Kamil Ayda-zade – (Corresponding Member of ANAS, Institute of Control Systems, Azerbaijan)
Prof. Agasi Malikov – (Corresponding Member of ANAS, Baku Engineering University, Azerbaijan)
Prof. Mehman Hasanov – (Azerbaijan Technical University, Azerbaijan)
Prof. Andrey Zenevich – (Belarusian State Academy of Communications, Belarus)
Prof. Bayram Ibrahimov – (Azerbaijan Technical University, Azerbaijan)
Prof. Yadigar Imamverdiyev – (Azerbaijan Technical University, Azerbaijan)
Prof. Famil Mammadov – (Azerbaijan Technical University, Azerbaijan)
Prof. Elshad Ismibeyli – (Azerbaijan Technical University, Azerbaijan)
Prof. Ramiz Humbatov - Baku Engineering University, Azerbaijan
Prof. Tofik Mansurov - Azerbaijan Technical University, Azerbaijan
Prof. Sylvio Simon – (Brandenburg Technical University, Cottbus-Zenftenberg, Germany)
Prof. Viktor Guzeev – (South Ural State University, Russia)
Prof. Nizami Yusubov – (Azerbaijan Technical University, Azerbaijan)
Prof. Vagif Movla-zadeh – (Azerbaijan Technical University, Azerbaijan)

Prof. Wilko Flügge – (Fraunhofer Institute for Large Structures in Manufacturing Engineering, Germany)
Prof. Nariman Rasulov – (Azerbaijan Technical University, Azerbaijan)
Prof. Alakbar Huseynov – (Azerbaijan Technical University, Azerbaijan)
Prof. Konstantin Samuylov – (Peoples' Friendship University of Russia, Russia)
Prof. Oleg Varlamov – (Moscow Technical University of Communications and Informatics, Russia)
Prof. Mustafa Guden – (Izmir University of Technology, Turkey)
Prof. Fakhraddin Aliyev – (National Aerospace Agency, Azerbaijan)
Prof. Kochetkova Irina – (Peoples' Friendship University of Russia, Russia)

Organizing Committee

Ph.D. Fariz Mammadov – (Azerbaijan Technical University, Azerbaijan)
Ph.D. Anar Eminov – (Azerbaijan Technical University, Azerbaijan)
Ph.D. Bakhtiyar Badalov – (Azerbaijan Technical University, Azerbaijan)
Prof. Vagif Maharramov - (Azerbaijan Technical University, Azerbaijan)
Ph.D. Rashad Tahirov – (National Defence University, Azerbaijan)
Ph.D. Ali Tagiyev – (Azerbaijan Technical University, Azerbaijan)
Ph.D. Mehman Binnetov – (Azerbaijan Technical University, Azerbaijan)
Ph.D. Zafar Ismayilov – (Azerbaijan Technical University, Azerbaijan)
Prof. Jamal Yildiz –(Gazi University, Ankara, Turkey)
Prof. Dmitrii Ardashev – (South Ural State University, Russia)
Prof. Igor Savin – (Kazan Technical University, Russia)
Prof. Ibrokhimali Normatov – (National University of Uzbekistan, Uzbekistan)
Prof. Yulya Gaydamaka – (Peoples' Friendship University of Russia, Russia)
Prof. Efendi Nasiboglu – (Dokuz Eylül University, Turkey)
Ph.D. Gulchin Abdullayeva – (Institute of Control Systems, Azerbaijan)
Ph.D. Almaz Aliyeva – (Mingachevir State University, Azerbaijan)
Prof. Georgiy Kuchuk – (Kharkov Polytechnic Institute of the National Technical University, Ukraine)
Prof. Asad Rustamov – (National Defence University, Azerbaijan)

Conference Executive Secretary

Prof. Nizami Yusubov – Azerbaijan Technical University, Azerbaijan

Table of Contents

AXIOMATIC MODEL FOR ENSURING STABILITY OF COMPLEX INTELLIGENT SYSTEMS

Mehman Hasanov[1], Vitaly Evdokimov[2], Anatoliy Navrotskiy[2], Hasan Najafov[3], Ilham Suleymanov[3]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan
[2]Belarus State University, Minsk, Belarus
[3]Nakhchivan State University, Nakhchivan, Azerbaijan
mehman.hasanov@aztu.edu.az, vigandvdk@gmail.com, navrotsky@bsuir.by, hasan.nacafov@ndu.edu.az, ilhamsuleymanov@ndu.edu.az

Abstract

This paper presents a review analysis of an axiomatic model aimed at ensuring stability, integration, and adaptability of complex intelligent systems within the context of Industry 4.0 and Industry 5.0 technologies. The model encompasses 15 axioms that describe key principles for the development and synthesis of independent subsystems into a unified ecosystem. Special attention is paid to the issue of balanced interaction between adaptability, resilience, resource optimization, and the capacity for evolution and scaling, which is crucial for stable functioning and long-term development of smart city infrastructure and Industry 4.0 and 5.0 systems in general. The article is intended for researchers and specialists in the fields of systems analysis, telecommunications, and control, as well as those involved in the design and implementation of complex adaptive intelligent systems.

Keywords: axiomatic model, stability, adaptability, adaptive intelligent systems, Industry 4.0, smart cities, telecommunications, communications

I. Introduction

Industry 4.0 and 5.0 technologies, such as smart cities, intelligent farms, and industrial IoT systems [1], are becoming increasingly significant in the context of global climate change, urbanization, and demographic and macroeconomic shifts. These technologies offer substantial opportunities for resource optimization, energy efficiency improvement, and enhanced management of urban and industrial infrastructure. For instance, in the context of smart cities, management systems integrate data from various subsystems to coordinate traffic flows, power supply, and water distribution, which helps reduce operational costs [2] and minimize environmental impact [3], while ensuring uninterrupted operation of critical infrastructure facilities. Nevertheless, the development and implementation of such systems face several significant challenges. In most cases, solutions are developed to meet local needs [4], which complicates the creation of holistic and synchronized ecosystems. This leads to interoperability issues and reduced efficiency at the system level.

To overcome these challenges, standardized approaches are required to unify processes and ensure coordinated functioning of all system components.

II. Review of Existing Systems, Approaches, and Concepts

Let us examine the key challenges that specialists face when implementing Industry 4.0 and 5.0 technologies (Figure 1):

Software aspect

-Lack of unified standards and protocols

One of the key problems is the absence of generally accepted and standardized interaction protocols between various components of cyber-physical systems (CPS) and the Internet of Things (IoT). This is particularly relevant for industrial systems, where different management levels (for example, APCS, MES, ERP) must interact to achieve a common goal.

-Lack of dynamic optimal algorithm selection

Modern software solutions in CPS and IoT are often developed with fixed algorithms, which limits their ability to adapt to changing conditions. This is particularly crucial for systems with high real-time requirements, such as neural network solutions or production process control systems [5].

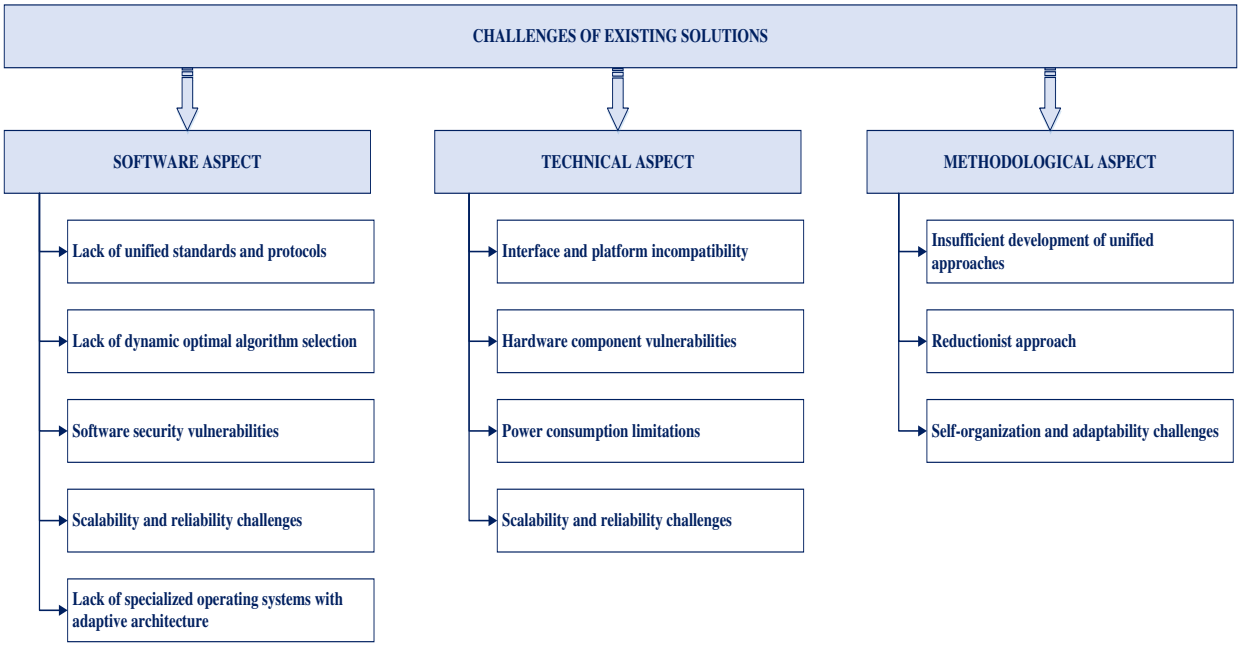


Figure 1: *Challenges of existing solutions*

Software security vulnerabilities

Software systems, particularly in distributed IoT networks, are vulnerable to multiple cyber threats. This necessitates the implementation of more robust protection mechanisms, including multi-layer security systems capable of preventing unauthorized access and attacks.

Scalability and reliability challenges

Software systems face challenges in maintaining reliability and scalability as the number of connected devices or data volume increases. Scalability is particularly critical in large ecosystems, such as smart cities, where the number of connected devices can grow exponentially.

-Lack of specialized operating systems with adaptive architecture

Most operating systems used in CPS and IoT lack built-in adaptation and flexibility mechanisms. This limits systems' ability to modify their structure or architecture in response to changing environmental conditions. Specialized operating systems are needed that could dynamically alter their configuration to ensure maximum efficiency and resilience [6].

Technical aspect

-Interface and platform incompatibility

Hardware components of modern CPS and IoT often utilize proprietary interfaces and protocols, which complicates their integration with other systems. The lack of standards and compatible interfaces creates barriers to effective interaction between systems and platforms [7].

-Hardware component vulnerabilities

Physical and information security of hardware components, especially in IoT devices, represents a significant challenge.Hardware component vulnerability necessitates the development of solutions that provide both physical protection and data security at the device level.

-Power consumption limitations

Power consumption constraints represent a critical challenge in the context of IoT devices, particularly under resource-limited conditions. Modern technologies such as 5G and NB-IoT partially address this challenge; however, further work is required to improve energy efficiency and enhance the scalability of hardware solutions.

Scalability and reliability challenges

Scalability and reliability issues concern not only software but also hardware components. Difficulties arise both with increasing numbers of connected devices and growing volumes of transmitted data. System reliability can be compromised due to insufficient coordination between various components, potentially leading to failures in global networks [8].

Methodological aspect

-Insufficient development of unified approaches

The lack of unified methodologies for the development and integration of CPS, IoT, and intelligent systems represents one of the key challenges. Standardized approaches and methods are needed to integrate systems of different levels and platforms with minimal effort [9].

Reductionist approach

Modern systems are frequently developed using a reductionist approach, where each system is viewed as a separate unit rather than a part of a global ecosystem. In the context of global networks, a transition to a holistic approach is necessary, one that considers the interaction of all ecosystem components and their capacity for adaptation [10].

-Self-organization and adaptability challenges

Current design methodologies do not provide sufficient mechanisms for real-time system self-organization and adaptability. This results in systems being unable to autonomously adapt to changing environmental conditions beyond input signal modification, requiring external intervention and manual adjustment. Implementation of methodologies that enable systems to automatically modify their architecture and algorithms could significantly enhance their resilience and efficiency [11].

III. Conclusions

The presented analysis of systemic challenges in technological transformation reveals a fundamental need for developing a principally new approach to creating integrated ecosystems.

A qualitative transition is required from the traditional paradigm of system integration to a holistic perception of the ecosystem as a unified organism, where technological, social, and economic components form an inseparable unity. In this context, the development of an axiomatic approach appears promising, providing a formal basis for describing systemic patterns and constructing a comprehensive theory of integrated ecosystems.

The axiomatic model can serve as the conceptual foundation that enables the transition from intuitive understanding of systemic interactions to rigorous scientific description of the principles governing the formation and development of integrated ecosystems. This opens the path to creating theoretically grounded methods for designing and managing new-generation complex sociotechnical systems.

IV. Axiomatic Model for Complex Adaptive Intelligent Systems

The accumulated experience in implementing Industry 4.0 and 5.0 solutions demonstrates the necessity of revising the fundamental principles of complex technological systems development. In response to this challenge, an axiomatic model is proposed, consisting of fifteen basic postulates that formalize key requirements for the design and development of modern technological solutions.

1.ASHR (Axiom of Systemic Homeorrhesis)

Ensures ecosystem stability through continuous monitoring and maintenance of balance among all components. Includes mechanisms for preventing and compensating deviations, protecting system integrity.

$$\Delta S(t) = \int \left[\sum(K_i * H_i) + \sum\left(R_j * E_j\right) + \sum(W_n * F_n)\right] dt * (1 + D) * (1 - C) \quad (1)$$

where $\Delta S(t)$ – is the change in system state over time, t–is time, K_i –is the time-dependent homeostatic regulation coefficient, H_i–is the internal environment stability characteristic, R_j–is the degree of control action influence, E_j–is the energy state characteristic, W_n–is the functional connection weight coefficient, F_n–is the functional activity characteristic, D–is the supplementation coefficient, C– is the compensation coefficient.

2.ATIS (Axiom of Topological Invariance of Systems)

Defines invariant basic principles and critical parameters of ecosystem functioning. rotects fundamental interaction mechanisms between components.

$$\Delta T(t) = \mu T * \left[\sum(N_i * L_i * M_i) + \sum(B_i * P_i * V_i)\right] * \left(1 - \frac{|T - T_{opt}|}{T_{max}}\right) * \exp\left(\frac{-t}{\tau c}\right) * (1 + D) * (1 - C) \quad (2)$$

where $\Delta T(t)$ – is the rate of change in topological structure over time, μT – is the topological stability coefficient, N_i – is the number of nodal system elements B_i – is the system elements connectivity, P_i – is the connection conductivity, V_i – is the volume of topological space, T – the current topological structure, T_{opt} – is the optimal value of topological structure, T_{max} – is the maximum allowable topological complexity, τc – is the characteristic time of topological changes.

3.AMAS (Axiom of Morpho-Adaptive Self-organization)

Ensures the system's ability to flexibly respond to external and internal changes without loss of functionality. Supports evolutionary system development through continuous adjustment to new conditions.

$$M(a) = \oint \left[a(t) * \frac{dF}{dt} + \beta(t) * \frac{dS}{dt}\right] * \exp(-\gamma * \tau) * (1 + D) * (1 - C) \quad (3)$$

where $M(a)$ – is the morphoadaptive system function, $a(t)$ – is the dynamic coefficient of functional adaptation, $\beta(t)$ – is the structural plasticity coefficient, F – is the functional state of the system, S – is the structural state of the system, γ – is the damping coefficient of adaptive changes.

4.AMSB (Axiom of Minimal System Basis)

Defines and maintains the critical minimum of resources and functions necessary for system survival. Creates the foundation for restoration of full functionality under favorable conditions.

$$B(min) = \sum(V_i * W_i) * (1 + D) * (1 - C) \geq B_{crit} * (1 + \varepsilon) \quad (4)$$

where $B(min)$ – is the minimum system basis, V_i – is the volume of the i-th basis element, W_i –

is the weight of the i-th basis element, B_{crit} – is the critical value of the system basis, $(1+\varepsilon)$ – is the stability margin coefficient.

5.AESO (Axiom of Entropic Self-Organization)

Ensures efficient distribution and utilization of all types of resources within the system. Optimizes resource flows for maximum efficiency of the entire ecosystem.

$$E = -k * \sum(p_i * \ln(p_i)) * \eta(R,t) * (1+D) * (1-C) \quad (5)$$

where E – is system entropy, k – is the normalization constant, p_i – is the probability of the i-th system state, $\eta(R,t)$ – is the self-organization efficiency coefficient, R – is system resources.

6. ASSE (Axiom of Systemic Synergy and Emergence)

Facilitates the emergence of new systemic properties and capabilities through component interaction. Integrates local innovations into system-wide improvements.

$$Syn(t) = \sum(E_i * I_j) * (1 + \mu * \sum(U_k * F_k)) * (1+D) * (1-C) \quad (6)$$

where $Syn(t)$ –is the system's synergetic effect over time, E_i –is the energy potential of the i-th element, I_j–is the information potential of the j-th element, μ–is the mutual enhancement coefficient, U_k–is the connectivity coefficient of the k-th interaction, F_k–is the dynamic factor of the k-th interaction.

7.ASBR (Axiom of Self-organizing Bifurcation and Development)

Creates safe mechanisms for testing and implementing system changes. Controls risks during innovation implementation while protecting ecosystem stability.

$$B(r) = \pm\sqrt{[\sum(\xi_i * p_i)]} * \exp(-\sigma * t) * \Psi(s) * (1+D) * (1-C) \quad (7)$$

where $B(r)$ – is the bifurcation function of system development, ξ_i – is the instability parameter of the i-th state, p_i – is the probability of the i-th bifurcation transition, σ – is the damping coefficient of bifurcation oscillations, $\Psi(s)$ – is the system stability function.

8.ADSB (Axiom of Dynamic System Balance)

Maintains optimal balance between system development and preservation of its stability. Manages the speed and scale of changes to maintain system stability.

$$D(b) = \int[a * \frac{dI}{dt} - \beta * \frac{dS}{dt} + \gamma * \frac{d^2R}{dt^2}] * \Omega(t) * (1+D) * (1-C) \quad (8)$$

where $D(b)$ – is the dynamic balance function, a – is the information dynamics coefficient, β – is the structural dynamics coefficient, γ – is the resource dynamics coefficient, I – is the information state, S–is the structural state, R – is the resource state, Ω(t) – is the temporal stabilization function.

9.AASR (Axiom of Autonomous System Regulation)

Ensures the ability of system components to independently respond to changes and threats. Enhances system resilience through autonomy of its elements.

$$A(r) = \sum[K_i * (1 - e^{-\varphi_i * t})] * \theta(R) * (1+D) * (1-C) \quad (9)$$

where$A(r)$–is the autonomous regulation function, K_i–is the coefficient of the i-th regulation loop, φ_i–is the establishment rate of i-th regulation loop, θ(R)–is the resource availability function.

10.ACSR (Axiom of Cognitive Self-Reference)

Creates mechanisms for experience accumulation and utilization at all system levels.Develops system's capacity for self-improvement through analysis of its own experience.

$$K(t) = \sum[M_i * L_i * \exp(V_i * t)] * \Phi(I) * (1 + D) * (1 - C) \quad (10)$$

where $K(t)$ – is the cognitive reference function, M_i – is the complexity measure of the i-th cognitive pattern, L_i – is the connectivity level of the i-th cognitive pattern, V_i – is the development rate of the i-th cognitive pattern, $\Phi(I)$ – is the information completeness function.

11.APSM (Axiom of Predictive System Modification)

Ensures the system's ability to anticipate changes and adapt development plans. Coordinates planning at all levels to achieve common goals. Creates mechanisms for flexible response to predicted changes.

$$P(m) = \sum[W_i * H_i * \exp(-\tau_i * \Delta t)] * \Psi(F) * (1 + D) * (1 - C) \quad (11)$$

where $P(m)$ – is the predictive modification function, W_i– is the weight of the i-th predictive factor, H_i – is the forecasting horizon of the i-th factor, τ_i – is the temporal decay coefficient of the i-th factor, Δt – is the forecasting time interval, $\Psi(F)$ – is the factor significance function.

12.ADSF (Axiom of Determined System Functionality)

Ensures stability and predictability of all system components' functioning. Guarantees predictable system behavior under various conditions.

$$F(d) = \sum[Q_i * Z_i * e^{-\eta_i * t}] * \Theta(S) * (1 + D) * (1 - C) \quad (12)$$

where $F(d)$ – is the determined functionality function, Q_i – is the quality of the i-th functional component, Z_i – is the significance of the i-th functional component, η_i – is the establishment rate of the i-th functionality, $\Theta(S)$ – is the system coherence function.

13.ASSD (Axiom of Sustainable System Development)

Ensures balanced development of the entire ecosystem in the long term perspective. Coordinates growth of various components to maintain overall sustainability.

$$U(r) = \sum[E_i * V_i * (1 + \mu_i * t)] * \Omega(R) * (1 + D) * (1 - C) \quad (13)$$

where $U(r)$ – is the sustainable development function, E_i – is the efficiency of the i-th development component, V_i – is the contribution of the i-th component to sustainability, μ_i – is the growth coefficient of the i-th component, $\Omega(R)$ – is the resource availability function for development.

14.ATSS (Axiom of Temporal System Synchronization)

Ensures balance between current tasks and strategic system development. Coordinates objectives of various components to achieve common results. Creates mechanisms for coherent achievement of short-term and long-term goals.

$$T(s) = \sum[S_i * P_i * exp(-\sigma_i * \Delta t)] * \Gamma(Q) * (1 + D) * (1 - C) \quad (14)$$

where $T(s)$ – is the temporal synchronization function, S_i – is the synchronization degree of the i-th process, P_i – is the priority of the i-th process, σ_i – is the temporal correlation coefficient of the i-th process, $\Gamma(Q)$ – is the system coherence function.

15.APVS (Axiom of Polymorphic Variability of Systems)

Maintains necessary diversity of components and connections in the system to ensure its adaptability. Ensures resilience through maintaining variability of solutions and approaches.

$$V(p) = \sum[B_i * G_i * (1 + p_i * t)] * \Lambda(M) * (1 + D) * (1 - C) \quad (15)$$

where $V(p)$ – is the polymorphic variability function, B_i – is the basic variability of the i-th morphotype, G_i – is the flexibility of the i-th morphotype, p_i – is the development coefficient of the i-th morphotype, $\Lambda(M)$ – is the morphological compatibility function.

After comparing the parameters and generalizing the functionality of the axiomatic model, we obtain the following formula:

$$\Psi(S,t) = \int[\sum(a_i F_i + \beta_i R_i + \gamma_i M_i)] * \exp(-\lambda t) * \Phi(S) * (1+D) * (1-C)dt \quad (16)$$

where $\Psi(S,t)$ –is the generalized system state function, a_i, β_i, γ_i–are influence coefficients, F_i–are functional parameters, R_i–are resource parameters, M_i –are morphological parameters, λ–is the generalized damping coefficient, $\Phi(S)$–is the system coherence function.

Based on this axiomatic model, a Python-based software package has been developed that implements mathematical modeling of adaptive intelligent system behavior. The modeling results obtained under critical perturbations are presented in Figure 2.

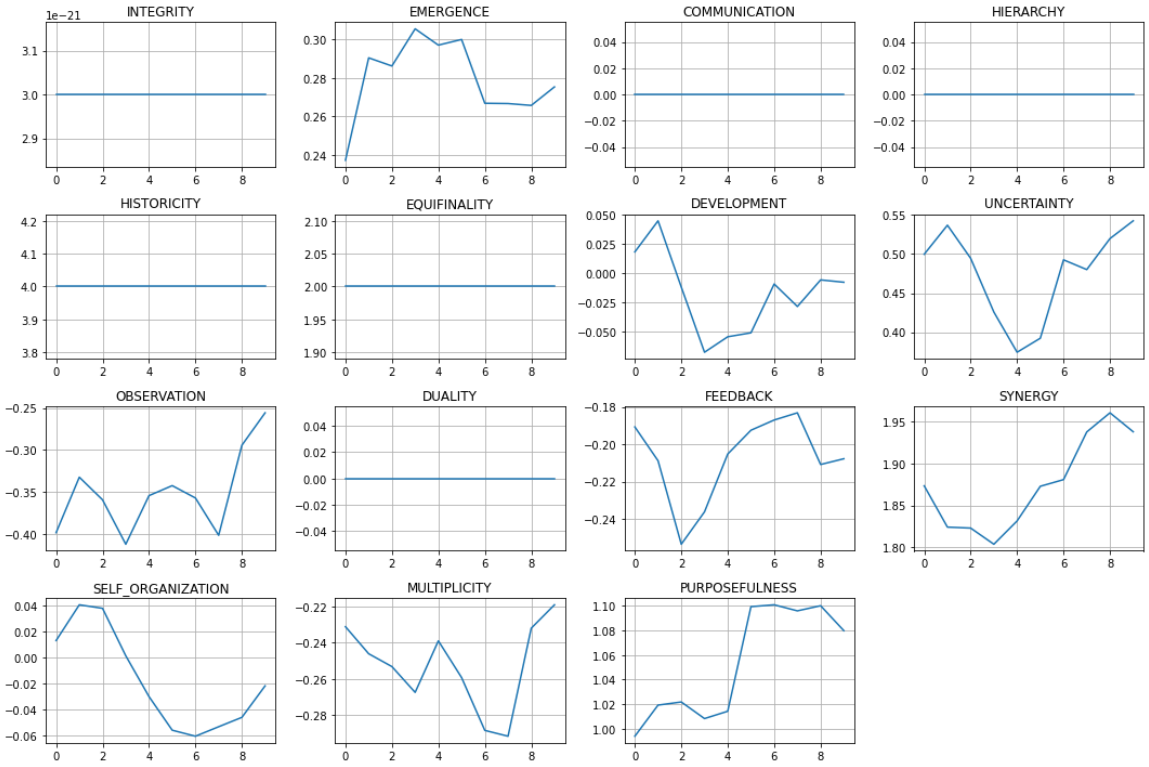


Figure 2: *Simulation Results*

The presented graphs demonstrate the results of numerical modeling of system parameter dynamics, reflecting the fundamental characteristics postulated in the axioms. Analysis of the obtained data allows evaluating system performance efficiency and verifying the theoretical provisions of the axiomatic model.

Analysis of the observed system dynamics processes reveals three key patterns:

– The structural core of the system (INTEGRITY, HISTORICITY, EQUIFINALITY) demonstrates high stability, indicating the robustness of basic system characteristics.

– The adaptive circuit (EMERGENCE, UNCERTAINTY, SYNERGY) shows directed positive dynamics with cyclic fluctuations, indicating active processes of system self-organization.

– The functional complex (SELF-ORGANIZATION, MULTIPLICITY, PURPOSEFULNESS) is characterized by wave dynamics with a tendency toward recovery, reflecting the current phase of system reconfiguration.

– The observed processes indicate a transitional state of the system while maintaining basic stability during active transformation of adaptive mechanisms. The overall development trend is positive, despite local instability of individual parameters.

V. Conclusion

The proposed axiomatic model represents an attempt to formalize key principles underlying evolving technological ecosystems. Fifteen fundamental axioms forming the core of the model create a holistic theoretical framework that enables comprehension and structuring of the multidimensional requirements space for modern technological solutions.The uniqueness of this approach lies in its abstraction from specific technological implementations, focusing instead on universal principles of self-organization, adaptation, and system evolution.

The presented axioms form a conceptual basis for understanding fundamental patterns in complex system development, which is critical in the context of increasing technology convergence and blurring boundaries between physical and digital worlds.

It is important to note that this model serves not only as a theoretical construct but also as a practical tool that enables:

– forming a holistic vision of complex system architecture;

– defining criteria for evaluating their effectiveness and sustainability;

– predicting the evolution directions of technological solutions;

– developing specific design and implementation methodologies;

– creating standards and protocols for component interaction.

Further development of the proposed approach requires:

– detailed elaboration of methodological apparatus for practical implementation of the formulated principles;

– creation of formal methods for verifying system compliance with axiomatic requirements;

– development of tools for evaluating the effectiveness of implemented solutions;

– formation of standards and recommendations for model application in various domains;

– advancement of theoretical foundation for describing self-organization processes in technological systems.

Of particular significance is the possibility of using this model as a universal language for describing and analyzing complex systems, which is critically important in the context of growing interdisciplinarity and technology convergence. This creates prerequisites for forming a unified conceptual space necessary for effective integration of various technological solutions and creation of truly holistic ecosystems of the future.

References

[1] Lu Y., Xu X. Internet of Things (IoT) Cybersecurity Research: A Review of Current Research Topics //IEEE Internet of Things Journal. 2019.

[2] Theorin A., Bengtsson K., Provost J. An event-driven manufacturing information system architecture for Industry 4.0 //International Journal of Production Research. 2017.

[3] Khan M., Wu X., Xu X., Dou W. Big data challenges and opportunities in the hype of Industry 4.0 //IEEE International Conference on Communications (ICC).2017. P. 1-6.

[4] Winfield A.F., Jirotka M. Ethical governance is essential to building trust in robotics and artificial intelligence systems //Philosophical Transactions of the Royal Society A.2018.

[5] Chen B., Wan J., Shu L. Smart Factory of Industry 4.0: Key Technologies, Application Case, and Challenges //IEEE Access. 2018. Vol. 6.

[6] Qu Y., Ming X.G., Liu Z.W., Zhang X. Smart manufacturing systems: state of the art and future trends //International Journal of Advanced Manufacturing Technology. 2019.

[7] Thoben K.D., Wiesner S., Wuest T. "Industry 4.0" and Smart Manufacturing–A Review of Research Issues and Application Examples/International Journal of Automation Technology. 2017.

[8] Xu L.D., Xu E.L., Li L. Industry 4.0: State of the art and future trends //International Journal of Production Research. 2018.

[9] Kang H.S.,Lee J.Y.,Kim S.S. Smart Manufacturing: Past resent Findings and Future Directions.International Journal of Precision Engineering and Manufacturing-Green Tech.2016

[10] Monostori L., Kádár B., Bauernhansl T. Cyber-physical systems in manufacturing //CIRP Annals - Manufacturing Technology. 2016. P.621-641.

[11] Frank A.G., Dalenogare L.S., Ayala N.F. Industry 4.0 technologies: Implementation patterns in manufacturing companies // International Journal of Production Economics. 2019.

MODELS OF CONTROL SYSTEMS WITH PARAMETRIC AND ADDITIVE-PARAMETRIC FEEDBACK

Sergey Yablochnikov[1], Tofig Mansurov[2],
Irina Yablochnikova[3], Rahman Mammadov[2]

•

[1]Plekhanov Russian University of Economics, Moscow, Russia
[2]Azerbaijan Technical University, Baku, Azerbaijan
[3]Moscow Technical University of Communications and Informatics, Moscow, Russia
vvkfek@mail.ru, tofiq-mansurov@rambler.ru, irayablochnikova@mail.ru, r.s.mamedov@mail.ru

Abstract

For processes implemented in object management systems of various nature, quite often the input signal can be formalized as one of a limited set of some, so to speak, standard functions. In particular, such a function may correspond to an increasing exponential. The functioning of control systems for these processes should be ensured in such a way that the reaction of the system itself to such highly probable signals is adequate and successfully formed in accordance with the criterion of minimum error. One of the ways to solve this problem is to form a parametric or additive–parametric feedback circuit in the control system, as well as integrate components with parameters changing according to a certain law into the structure of the system. At the same time, the error should not depend on the magnitude of the parameters noted above and, in fact, is minimized.
The authors synthesized and substantiated models and algorithms for the functioning of control systems with parametric feedback by transmission coefficient, which make it possible to eliminate control errors and ensure the noise immunity of the system. It is also proposed to use other means that ensure the high efficiency of such systems that implement the management of objects and processes of various types.

Keywords: mathematical modeling, control systems, feedback, features of the functioning of systems, parametric feedback, optimization of systems.

I. Introduction

The fundamental foundations of the modern theory of optimal control were formed by a large body of researchers in the second half of the twentieth century. First of all, the efforts of researchers from various countries in this field were focused on the successful solution of problems correlating with the synthesis of control systems for objects of various nature, which in the process of their functioning in a more or less stable external environment did not undergo significant qualitative and quantitative changes [1-3].

Mathematical models reflecting the functioning of such control systems, as a rule, were synthesized by most authors on the basis of systems of differential equations. Over time, researchers began to pay attention to the development of theoretical foundations and practical recommendations for the design of control systems that would be able to adapt to significant changes in their operating conditions, including by correcting the structure of the system and the

parameters of its individual components).

One of the basic concepts of the above-mentioned theory of control of dynamic systems is the concept of "feedback", which is used in the formalization of the description of the functioning of the so-called closed-type control systems. The main advantage of such systems over alternative systems of the "open" type is their ability to independently form a reaction (usually a compensation type reaction) to external disturbances. At the same time, effective algorithms for the functioning of feedback systems are either based on the principle of actually isolating the system from external disturbances, or on the principle of adaptation to dynamic conditions of an aggressive external environment.

The feedback principle is successfully applied by researchers and designers of systems and processes of various nature (technical, technological, informational, chemical-biological, economic, etc.) for their analysis and synthesis. In the most general case, the functioning of a feedback system can be understood as the formation of some impact of the results of the vital activity of the system on the nature of the implementation of this vital activity. The type of feedback (additive or parametric) determines a set of specific tasks that can be successfully solved in practice due to its formation in the structure of the system, as well as approaches to the synthesis and optimization of mathematical models for an adequate formal description of the functioning of the system. The synthesis of optimal control system models is still a very urgent task today [3-5].

II. Methodology and purpose of the study

The purpose of this article is the analysis, synthesis and scientific substantiation of mathematical models, as well as the analysis of the set of features of the functioning of control systems with parametric and additive–parametric feedback. Such systems are widely used in practice to ensure effective management of objects and processes of various nature. In particular, the authors pay appropriate attention to the issues of the accuracy of the functioning of such systems and their noise immunity.

The methodology of this scientific research is based on the theoretical provisions of technical cybernetics as the science of managing complex, hierarchical, multidimensional and adaptive systems, in particular, modern ideas in the field of automatic control theory.

III. Modeling of control systems with parametric feedback

Undoubtedly, a rather important point for the successful development and design of effective control systems for various objects and processes is to ensure the necessary level of accuracy of functioning, which, as a rule, determines their quality and, accordingly, the effectiveness of solving practical problems. In the current conditions, when any control system implements its functioning surrounded by a huge number of other technical and technological systems that exchange large flows of information with each other, its noise immunity is no less important. In turn, the noise immunity of the system determines its reliability.

As it was shown in previously published articles, the condition for accurate reproduction of exponential input effects in automated control systems is the equality of exposure time constants and an unstable aperiodic link, which is intentionally included in the feedback circuit. Since the value of the time constant of the exponential function corresponding to the input effect on the system may change during the operation of the system as a whole, it becomes necessary to adapt the value of the time constant of the unstable aperiodic link in such a way that the above equality condition is ensured. It is the fulfillment of such a condition that will minimize the magnitude of the dynamic error ε, in fact, reducing it to zero [6-7].

Thus, a peculiar parameter, depending on the value of which, in fact, optimization of the functioning of the control system as a whole, as well as the aperiodic link, in particular, is the gradient of the dynamic error.

In previously published works [8-9], the authors of this article substantiated the thesis that the aperiodic link using a parametric feedback circuit is characterized by parameters that functionally depend on the amplitude and sign of the input signal, in particular, in our case, the signal corresponding to the error $\pm\varepsilon$. In particular, the formal functioning of such a circuit can be represented by a set of the following equations.

For the direct circuit of the control system block diagram (1):

$$T_1 \frac{du}{dt} - u = k_n u_1 \tag{1}$$

where $u(t) -$ is the output signal for this component of the control system; $u_1(t) -$ is the input signal for the aperiodic component of the control system; $k_n -$ is the transmission coefficient of the so–called "short circuit link"; $T_1 -$ is the time constant of the aperiodic component of the control system.

For the closure link, which, in the case of additive feedback, is a comparator (2):

$$u_1 = (k_m + k_1 x_1)\varepsilon \tag{2}$$

where $k_m -$ is the transmission coefficient of the closed control system; $k_1 -$ is the transmission coefficient of the open control system; $x_1 -$ is the output signal of the parametric feedback circuit; $\varepsilon -$ is the error signal.

For the feedback circuit of the control system (3):

$$x_1 = k_0 u \tag{3}$$

Then the following equation (4) will be characteristic of the control system as a whole, which describes its functioning from a formal point of view

$$T_{ekv} \frac{du}{dt} - u = k_{ekv}\varepsilon \tag{4}$$

where $T_{ekv} = T_1/(1 + k_1 k_0 k_n \varepsilon,$

$$k_{ekv} = k_m k_n/(1 + k_1 k_0 k_n \varepsilon \tag{5}$$

For $\varepsilon = 0$, the equation of the transfer function of the quasi-static link of the control system with parametric feedback, when the condition of adaptation of the time constant is fulfilled, must correspond to (4). In order to ensure the fulfillment of this condition, according to the authors, it is necessary to accept $k_m = 1$.

Depending on the magnitude and sign of the error ε the T_{ekv} parameter will also change. By changing the parameters k_1 and k_0, it is possible to ensure an optimal ratio between the variations of ε and T_{ekv}. Simultaneously with the change in the time constant T_{ekv}, a proportional change in the transmission coefficient k_{kv} is realized, the value of which (5) was defined by us somewhat higher in the text as a parameter of equation (4). In turn, the fulfillment of such a condition will ensure very small changes in the values of the roots of the characteristic equation during the adaptation process of the functioning of the entire system. Consequently, the quality indicators of the transition process implemented in such a management system will actually be constant. Moreover, the effective signal bandwidth will also be unchanged, and the characteristics of the control system as a whole will be optimal.

The implementation of an approach to optimizing the functioning of the control system by adapting the parameters of a quasi-static component forcibly integrated into the system allows, in

fact, to transform it from a dynamic to a static system, since additional information about the values of the parameters can be obtained only if the condition $\varepsilon \neq 0$ is met. And in order to implement astatic adaptation, it is additionally necessary to have an integrating link in the structure of the quasi-static component.

In any case, the adaptation of self-adjusting automatic control systems, the operation of which ensures constant quality indicators of transients, regardless of the type and nature of parametric effects, is realized due to the presence of a parametric feedback circuit in its structure. In some cases, in addition to the parametric feedback circuit, an additional circuit is used that provides a direct parametric connection of the input signal with the parameters of the direct circuit. In general, systems have proven themselves well, in the structure of which there is a combination of the following components: a means of determining process quality parameters; a means of forming an adaptation algorithm (self-tuning); some specific executive device.

Parametric feedback is also used in the design of control systems with the implementation of the identification of control objects. But, in any case, the structure of adaptive (self-adjusting) control systems contains elements that provide combined additive-parametric feedback. At the same time, the main circuit is providing additive feedback, and the auxiliary one is providing parametric feedback.

If the parameters of the control object vary quite slowly compared to the variations of the processes in the system itself, then the analysis of the additive and parametric feedback circuits can be implemented quite simply by separating them. If the speeds of the above processes are sufficiently close to each other, in this case the self-adjusting control system is a complex nonlinear system, the analysis of the functioning of which requires the use of some specific methods, which were described by researchers in a number of publications at the time [7-9].

IV. Adaptive properties of parametric feedback systems in circuits providing self-tuning

Next, we will consider a number of problems that may arise in control systems with a large variation in the transmission coefficient of the control object $k(t)$. In this case, the equation of the control object takes the following form (6).

$$\sum_{i=0}^{n} a_i y^i = k(t) \sum_{j=0}^{m} b_j u^j \tag{6}$$

Further, let's assume that there is a possibility of some compensation for variations in the transmission coefficient $k(t)$ due to the corresponding variations in the transmission coefficient of the regulator $k_r(t)$ itself. If $u = \varepsilon k_r(t)$, and $\varepsilon = x - y$, then the equation of the feedback control system, taking into account (6), will take the following form (7):

$$\sum_{i=0}^{n} a_i y^i = k(t) \sum_{j=0}^{m} b_j \big(k_r(t)(x-y)\big)^j \tag{7}$$

By revealing the right side of equation (7) according to the Leibniz formula and simultaneously implementing the rearrangement of the components of the formula, the following fact can be established. The coefficients on the left side of this equation (7) for derivatives up to the order of m will functionally depend on the transmission coefficient of the regulator $k_r(t)$ as well as on its derivatives.

Therefore, only with sufficiently slow changes in the values of this coefficient $k_r(t)$ and, accordingly, sufficiently slow changes in the values of the coefficient $k(t)$, it is possible to transform

(7) into (8):

$$\sum_{i=m+1}^{n} a_i y^i + \sum_{j=0}^{m} (a_i + k(t)k_r b_j) y^j = k(t)\, k_r(t) \sum_{j=0}^{m} b_j x^j \tag{8}$$

It follows from equation (8) that in order to ensure the condition of independence of the properties of the control system from the values of the transmission coefficient $k(t)$, it is necessary that, if possible, the following equality (9) be fulfilled:

$$k(t) \cdot k_r(t) = const \tag{9}$$

And, accordingly, the fulfillment of equality (9) can be ensured only if there is a parametric feedback circuit in the control system. However, the more precisely the ratio (9) is fulfilled, the more unacceptable are the conditions for ensuring effective transmission of information in the system as a whole. Therefore, as a rule, control subsystems and adaptation subsystems are separated in such systems.

A number of problems that arise in this case can be stopped by exciting some natural oscillations in the main circuit of the system, used as a signal for the implementation of self-adjustment (self-adaptation). But, at the same time, such fluctuations have practically no effect on the implementation of the basic management regime. As some of the features, it is necessary to note a number of specific requirements for the parameters of self-oscillations in the main circuit of the system.

The first requirement is the constancy of the oscillation frequency, which makes it possible to stabilize other self-oscillation parameters. The second requirement is to exceed the self-oscillation frequency of the maximum frequency of the spectrum of the input signal acting on the system. This condition is determined by the need to ensure a sufficient level of noise immunity of the system and to prevent distortion of the input signal.

The following types of signals can also be used to configure control systems with parametric feedback by monitoring the passage of certain testing influences: various types of harmonic signals, the so-called "white noise", as well as periodically repeating sequences of pulses of various shapes (rectangular, triangular, trapezoidal, etc.). To form such "test" signals, it is necessary to have an appropriate external source is available, the main requirement for the practical implementation of which is the stability of the parameters of the signal generated by it. In some cases, it is very problematic to ensure such stability.

If, for the implementation of testing of the system, its own oscillations are used, excited in the main circuit of the control system, as we indicated above, then such a problem is automatically solved. In this case, with respect to the control signal x(t), the control system behaves as if there is no parametric feedback circuit in its structure. This is due to the fact that the control signal does not really penetrate the self-tuning (adaptation) circuit of the system. It is stopped by a special filter. And the transition process in such a system differs from the transition process in a system in which there is no self-adaptation chain, only by the presence of an additional harmonic signal of very small amplitude.

An essential feature of self-adjusting systems with a so-called limit cycle and parametric or additive parametric feedback is their significant performance, structural simplicity and high level of reliability of operation. This allows, in particular, minimizing hardware and software tools to ensure successful management of a wide range of facilities, including those for which the operator's direct participation in the implementation of management actions is very difficult.

V. Conclusions

The material presented in this article is based on a set of fundamental approaches within the framework of technical cybernetics, and, therefore, can be used for the successful synthesis of control systems for complex technical and technological objects and processes. However, the practical application of the above ideas is not limited to technical or technological fields. This is due to the fact that the feedback principle, in this case parametric feedback, is a powerful and universal tool for the successful implementation of effective management on a scientific basis of a set of processes in any other industries (economics, biology, etc.) In all the above cases, the objects of management are complex, multidimensional, hierarchical systems with the property of adaptation [9-10].

Successful achievement of the set of goals for the functioning of such multidimensional systems is impossible without the formation and implementation of numerous feedback loops (both positive and negative, both additive and parametric). Today, this is an indisputable fact. In this article, we have considered individual cases of the implementation of adaptive management, focused on ensuring compliance with the criterion of minimum errors that occur during the operation of complex objects and systems. The use of additional components that, in a certain sense, stop errors that occur during the operation of systems, in practice turns out to be very productive. In the future, in a number of their publications, the authors of this article plan to specify a set of practical tools for implementing such control using systems with parametric and additive-parametric feedback.

References

[1] L. Lilburne, M. Robson-Williams and N. Norton, Improving Understanding and Management of Uncertainty in Science-Informed Collaborative Policy Processes, *Sustainability 2022*, vol. 14, pp. 6041, [online] Available: https://doi.org/10.3390/su14106041.

[2] I. A. Shcherbatov, Intellectual management of robotics systems in the conditions of uncertainty, *Vestn. Astrakhan State Technical Univ. Ser. Management Computer Sciences and Informatics*, vol. 1, pp. 73-77, 2010.

[3] T. V. Sibalija and V. D. Majstorovic, An integrated approach to optimise parameter design of multi-response processes based on Taguchi method and artificial intelligence, *Journal of Intelligent Manufacturing*, vol. 23, no.5, pp. 1511-1528, 2012.

[4] N. V. Mokrova, S. L. Yablochnikov, A. B. Semenov and I. K. Kuchieva, Intensification of Intelligent Automated Control Systems, 2023 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, Russian Federation, 2023, pp. 1-6, doi: 10.1109/IEEECONF56737.2023.10091978.

[5] V. S. Artemyev, M. N. Makhiboroda, S. L. Yablochnikov, V. A. Kurbatov and Z. E. Tarkhanova, Implementation of Adaptive Control with Parametric Uncertainty,*2022 Intelligent Technologies and Electronic Devices in Vehicle and Road Transport Complex (TIRVED)*, Moscow, Russian Federation, 2022, pp. 1-7, doi: 10.1109/TIRVED56496.2022.9965505.

[6] Kuptsov, M.I., Minaev, V.A., Faddeev, A.O. et al. On the Stability of Integral Manifolds of a System of Ordinary Differential Equations in the Critical Case. J Math Sci 262, 825-834 (2022). https://doi.org/10.1007/s10958-022-05861-5.

[7] I.O. Yablochnikova, V.B. Dzobelova, S.L. Yablochnikov, S.D. Savostin, Ensuring a Given Level of Control Accuracy for Systems of Various Nature, 2024 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO), Vyborg, Russian Federation, 2024, pp. 1-5, doi: 10.1109/SYNCHROINFO61835.2024.10617751.

[8] S. L. Yablochnikov, I. O. Yablochnikova, V. B. Dzobelova and L. V. Medvedeva, Aspects of Assessing the Absolute Stability of Control Systems Containing Nonlinear Components, 2024 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO), Vyborg, Russian Federation, 2024, pp. 1-6,

doi: 10.1109/SYNCHROINFO61835.2024.10617666.

[9] S.L. Yablochnikov, I.O. Yablochnikova, M.I. Kuptsov, M.N. Machiboroda and I. V. Bondarenko, On the Issue of Synchronization of Real and Virtual Processes in the Context of the Evolution of the Modern Information Society, 2021 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO, 2021, pp. 1-6, doi: 10.1109/SYNCHROINFO51390.2021.9488411.

[10] S.L. Yablochnikov, I.O. Yablochnikova, S.V. Vidov, M.I. Kuptsov and A.V. Olisaeva, The Aspects of Modeling Information Processes Realized in Complex Telecommunication Systems, 2018 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), St. Petersburg, 2018, pp. 1-5, doi: 10.1109/WECONF.2018.8604360.

FORMATION OF AN INFORMATION LEAKAGE CHANNEL FROM AN OPTICAL FIBER BY THERMAL ACTION

Andrey Zenevich[1], Tatiana Matkovskaia[1], Nurana Camalzadeh[2], Javid Namazov[2]

[1]Belarusian State Academy of Communications, Minsk, Republic of Belarus
[2]Azerbaijan Technical University, Baku, Azerbaijan
a.zenevich@bsac.by, t.matkovskaya@bsac.by, nurana002@gmail.com,
namazovjavid@gmail.com

Abstract

The article evaluates the possibility of forming a channel for information leakage from a defect in optical fiber created by thermal action. The properties of optical fiber inhomogeneities caused by such action are currently practically unstudied, which determines the relevance of the research. The paper shows that local temperature action makes it possible to form a defect in optical fiber that allows part of the optical radiation to be removed beyond this fiber, that is, to create a channel for unauthorized data retrieval. It is shown that with an increase in the wavelength of optical radiation propagating along the fiber, the loss of radiation power on the defect formed by thermal action on the optical fiber increases. It is found that with the same loss of power on the defect formed by thermal action, the optical radiation power removed from such a defect has the greatest value when using G.652 optical fiber, and the least when using G.657 fiber. The results presented in the article can be used in designing systems for protecting information transmitted over fiber-optic communication lines.

Keywords: optical fiber; optical fiber defect; thermal effect; information leakage channel.

I. Introduction

Today, fiber-optic communication lines (FOCL) are widely used to transmit information. Optical fibers included in FOCL have advantages in speed and throughput compared to copper cores [1-3].

Data transmitted via optical fibers are more protected from unauthorized access to the transmitted information, but it is possible to form an information leakage channel by diverting part of the optical radiation from the fiber without breaking it [4]. Methods for implementing an information leakage channel include the formation of a macrobend and microbend of an optical fiber [5-6], as well as methods implemented on the basis of optical tunneling and compression of an optical fiber [7].

To connect devices implementing these methods, access to a section of optical fiber of a certain length is usually required, but it is not always possible to obtain it. At the same time, it is possible to create a connection zone for a device that provides unauthorized information retrieval by thermal action on the fiber. Such action leads to the appearance of non-uniformity of the optical fiber, from which part of the optical radiation power can escape beyond the fiber.

To detect thermally induced optical fiber inhomogeneities, it is necessary to know the characteristics of these inhomogeneities. The most important characteristics of inhomogeneities are the radiation power losses they introduce in the optical fiber and the proportion of optical radiation power diverted through the inhomogeneity formed outside the optical fiber. The first of these properties determines the possibility of detecting the presence of an inhomogeneity. The second allows us to estimate the possibility of using the inhomogeneity for data retrieval. However, the properties of thermally induced optical fiber inhomogeneities have not been studied to date. Therefore, the purpose of this work is to determine the characteristics of thermally induced optical fiber inhomogeneities.

II. Experimental setup and measurement technique

The objects of the study were standard single-mode optical fibers G.652, G.655 and G.657, which are quite often used in optical cables. The properties of optical fiber inhomogeneities caused by thermal effects were studied using an experimental setup, the structural diagram of which is shown in Fig. 1. The following designations are used in the diagram: RS – radiation source, OV – optical fiber, PM – power meter, D - diaphragm, FD – photodetector, G – burner, A – ammeter, V – voltmeter, PS – power supply, R_l – load resistor.

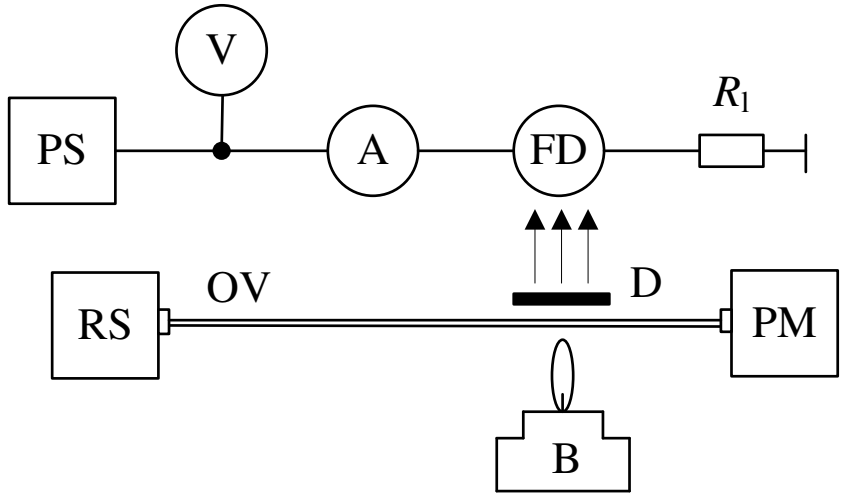


Figure 1: *Block Diagram of the Experimental Setup*

The principle of operation of the experimental setup: from the source RS, optical radiation enters the fiber OV, and from it – by measuring the intensity IM. From the constant voltage source PS, the supply voltage Us is supplied to the photodetector PD. The voltmeter V is needed to control the measurement of the reverse bias voltage. The ammeter A is used to measure the turn-off current flowing through the photodetector FD. A load resistor Rl with a nominal value of 1 kOhm is connected in series with the photodetector, which corresponds to the limitations of the current flowing through the photodetector.

The wavelength of the optical signal RS can change during the measurements and the values 1310, 1490, 1550 and 1650 nm are accepted, which corresponds to the "transparency windows" of single-mode optical fibers. During the studies, the optical signal power losses in the fiber were not taken into account, since its length was only 1 m. It should be noted that for all applications of optical fibers at constant lengths, a change in the wave intensity does not lead to an increase of 0.4 dB / km.

To form a defect in the optical fiber OV, a small part of this fiber was placed in the flame of the

burner B. In this case, the diaphragm D was closed to limit the penetration of light and heat from the burner B onto the photodetector FD. During the studies, the photocurrent I_{ph} of the photodetector FD was calculated:

$$I_{ph} = I - I_d \quad (1)$$

where I_d is the dark electric current measured with the diaphragm D closed; I is the electric current measured with the diaphragm open.

The photocurrent value was used to determine the power of optical radiation arriving at the photodetector FD from the inhomogeneity of the optical fiber caused by the temperature effect: $P_b = I_{ph}/S$, where S is the sensitivity of the photodetector.

Note that different areas of the flame have different temperatures [8]. Thus, the temperature of the upper region of the flame is the highest (1500 K), and the temperature of the region located near the wick is the lowest (800 K) [9]. Therefore, the optical fiber was placed in the upper part of the flame. This allowed the fiber to be exposed to thermal action with a constant temperature of 1500K.

When performing measurements, the loss of radiation power D_l introduced by the defect caused by the temperature effect was determined by the following formula:

$$D_l = 10\lg\left(\frac{P}{P_{PM}}\right) \quad (2)$$

where P is the power of the radiation source RS, P_{PM} is the power of optical radiation arriving at the power meter PM.

The time of thermal action varied from 1 to 10 s. The longer the action on the optical fiber, the greater the insertion loss of radiation power on the created defect.

Note that with a thermal action time on the optical fiber of less than 1 s, it was not possible to form a defect with significant insertion loss of radiation power D_l. With a thermal action time on the optical fiber of more than 10 s, the insertion loss of radiation power Dl on the defect exceeded 20 dB. Such a value of D_l for zonal and trunk fiber-optic communication lines leads to the cessation of data transmission. Therefore, it was considered inappropriate to consider effects introducing such losses of radiation power.

The length of the section where the defect caused by the thermal effect occurred was about 1 cm. When the optical fiber was thermally affected, the fiber's paint coating, if any, burned in the area of this effect. After the thermal effect, the fiber became brittle in the area of the formed defect and broke even with a slight bend. In the absence of a paint coating, the appearance of the area where the thermal effect was carried out remained the same as before the effect.

To determine the radiation power D_b, diverted beyond the optical fiber from the inhomogeneity caused by the thermal effect, the following expression was used:

$$D_b = 10lg\left(\frac{Pb}{P}\right) \quad (3)$$

During the studies, the characteristics of the optical fiber inhomogeneities caused by thermal exposure were also estimated using the reflectometric method on an experimental setup, the structure of which is given in [10]. In this case, the inhomogeneity caused by thermal exposure was formed in the middle of an optical fiber 1.5 km long. During the measurements, reflectograms of signals in the fiber with such inhomogeneity were recorded, for which a reflectometer was connected to the input of the optical fiber. The duration of the reflectometer's optical pulse was 3 ns. With such a duration of optical pulses, the length of the dead zone for attenuation of the reflectometer used was minimal, which made it possible to determine with the greatest accuracy the locations of the

thermal exposure and the amount of radiation power loss on the formed defect.

Measurements of the values of D_l and D_b, as well as reflectograms, were performed at room temperature T = 293 K.

III. Results of measurements and their discussion

During the research, a defect in the optical fiber was created by thermal action on the optical fiber with a burner flame (Figure 1). In this case, the magnitude of the power loss in the defect depended on the time of action on the optical fiber of the flame exposed to the flame.

During thermal action, diffusion of impurities introduced into the fiber during its production can occur in the optical fiber. Such diffusion leads to a change in the absolute values of the refractive indices of the core and cladding of the optical fiber at the point of thermal action, and, consequently, their difference. This leads to a violation of the condition for the existence of one mode in a single-mode fiber and the appearance of additional modes. Redistribution of energy between modes leads to a loss of power of the transmitted optical signal and the release of energy of additional modes beyond the fiber in the area of its local heating. Removal of the paint coating also contributes to the release of energy of optical radiation beyond the fiber.

The results of measuring the optical radiation power loss for the same defect caused by the thermal effect on the optical fiber for different wavelengths of optical radiation are presented in Table 1. As follows from the obtained results, the longer the wavelength, the greater the optical radiation power loss for all the studied optical fibers. Table 1 also displays information on the optical radiation power removal from the side surface of the optical fiber in the defect area. As can be seen from the obtained data, with increasing wavelength, more optical radiation power is removed from the side surface of the optical fiber in the defect area. The above trends are observed for all the studied optical fibers. Note that the defects of G.652 optical fibers, which have a lower optical radiation power loss than the other studied optical fibers, had higher values of the optical radiation power removed from the side surface of the optical fiber in the defect area. This is due to the different internal structure of the studied fibers.

Table 1: *The results of measurements of the power loss of optical radiation in the area of the defect caused by thermal effects on the optical fiber*

Type of optical fiber	λ, [nm]	Optical radiation power loss, [dB]	Optical power diversion, [dB]
G.652	1310	0,40	-48,70
	1490	1,23	-48,10
	1550	1,29	-47,40
	1625	1,51	-46,90
G.655	1310	0,97	-56,60
	1490	2,16	-56,10
	1550	2,79	-55,60
	1625	3,52	-55,10
G.657	1310	5,01	-57,80
	1490	5,27	-57,20
	1550	5,49	-56,30
	1625	6,10	-54,30

An analysis of the reflectogram section of the G.652 optical fiber containing a defect shows that the location of such a defect is characterized by the presence of a power drop in the form of a “step”

(Figure 2). The magnitude of such a power drop increases with increasing wavelength, i.e. for these defects, a dependence of the loss of optical radiation power on the wavelength of optical radiation is observed. Note that the reflectograms for these defects and their behavior with increasing wavelength of optical radiation are similar to the reflectograms characteristic of macrobends in optical fiber [11].

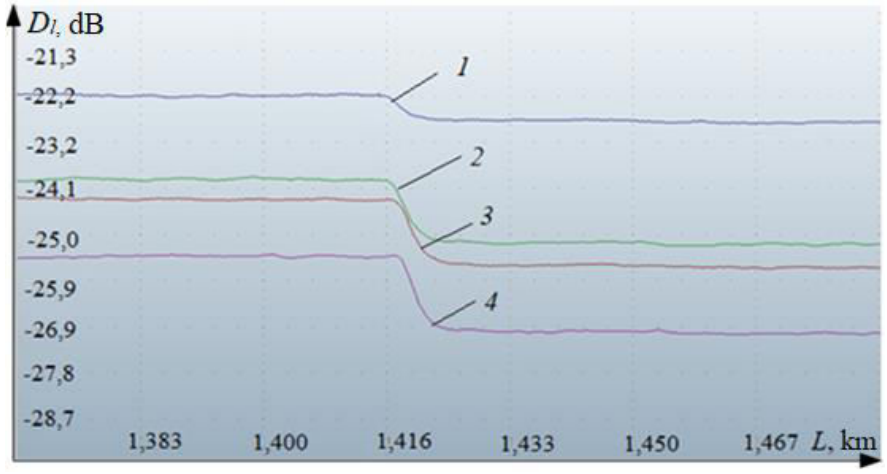


Figure 2: *The section of the reflectogram of the optical fiber G652 containing the defect, for wavelengths: 1 – 1310 nm, 2 – 1490 nm, 3 – 1550 nm, 4 – 1625 nm*

Reflectograms are given for the optical fiber G.652. Reflectograms for other optical fibers are identical. It can be seen from the reflectograms that an increase in the loss of optical radiation power on a defect leads to an increase in the value of the optical radiation power diverted from the defect. With the same value of the power loss on a defect, different values of the optical radiation power diverted from the defect were observed for different wavelengths of optical radiation.

Figure 3 shows typical dependences of the value of the radiation power Pref diverted from a defect formed as a result of thermal exposure on the value of the loss of radiation power that occurred due to this defect for different wavelengths. As follows from the obtained dependences, with an increase in the loss of radiation power, the value of the radiation power diverted from the defect Pref increases. The obtained dependences were nonlinear. This indicates that the power losses on such a defect are caused not only by the radiation that goes beyond the optical fiber in the area of this defect.

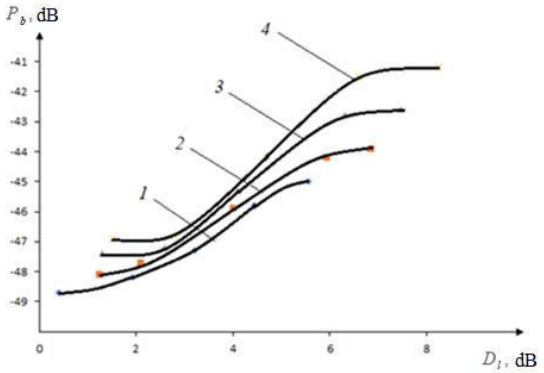


Figure 3: *The dependence of the removal of optical radiation power from a defect formed as a result of thermal exposure on the magnitude of the power loss caused by this effect for wavelengths: 1 – 1310 nm, 2 – 1490 nm, 3 – 1550 nm, 4 – 1625 nm*

In the course of the conducted study, a comparison of the optical radiation power removal from the defect was performed for different optical fibers. During the comparison, defects were created in each of the studied optical fibers that introduced the same power loss at the same optical radiation wavelength. The optical radiation power loss was measured for the same optical radiation wavelength. Information on the obtained results is presented in Table 2. Based on the data presented in Table 2, the highest power removal from the defect formed as a result of thermal exposure is observed for the G.652 optical fiber, and the lowest – for G.657. This is observed for all studied radiation wavelengths. Such a difference in the values of power removal from the defect is associated with a different internal structure of the studied optical fibers. Note that the studied fibers have a difference in the geometric dimensions of the core and cladding of the fiber [12-14].

Table 2: *The results of measurements of the removal of optical radiation power from the lateral surface of the optical fiber in the area of the defect formed by thermal action*

Type of optical fiber	Wavelength, [nm]	Loss of power, [dB]	Power diversion from the side surface of the fiber in the defect area, [dB]
G.652	1310	3,5	-47,1
G.655			-52,8
G.657			-58,6
G.652	1490	4,5	-46,0
G.655			-52,0
G.657			-58,0
G.652	1550	5,0	-44,1
G.655			-51,0
G.657			-56,5
G.652	1625	6,0	-42,0
G.655			-50,7
G.657			-54,4

IV. Conclusion

It has been shown that by means of local temperature action it is possible to form a defect in an optical fiber, allowing part of the optical radiation to be removed beyond the fiber.

It has been determined that with an increase in wavelength, the power of optical radiation removed from a defect formed by thermal action on an optical fiber increases.

The studies have shown that with the same power loss on a defect formed by thermal action on an optical fiber, the removal of optical radiation power from such a defect has the greatest value when using G.652 optical fiber, and the least when using G.657.

References

[1] Sklyarov O. K. Fiber-optic networks and communication systems. St. Petersburg: Lan, 2021. 268 p. (in Russ.)

[2] Ionov A. D. Fiber-optic transmission lines: Textbook. Novosibirsk: SibSUTI, 2003. 152 p. (in Russ.)

[3] Govind P. Agrawal Fiber-Optic Communication Systems. New York: Wiley-Interscience, 2002. 563 p.

[4] Zenevich A. O. Optical fiber information leak detectors. Minsk: Belarusian State Academy of Communications, 2017. 142 p. (in Russ.)

[5] Gulakov I., Zenevich A., Kochergina O., Matkovskaia T. Investigation of an Information Leakage Channel in the Area Optical Fiber Bending. Proc. of Telecom. Universities. 2022;8(3):44-49. (in Russ.) DOI:10.31854/1813-324X-2022-8-36-44-49

[6] Gulakov I., Zenevich A., Matkovskaya T., Novikov E. Investigations of Single-Mode Optical Fiber Microbending Properties. Proc. of Telecom. Universities. 2023;9(4):15-20. (in Russ.) DOI:10.31854/1813-324X-2023-9-4-15-20

[7] Shubin V. V. Information security of fiber-optic systems. – Sarov: RFNC-VNIIEF, 2015. 257 p. (in Russ.)

[8] Devisilov V.A., Drozdova T.I., Timofeeva S.S. Theory of combustion and explosion. M.: FORUM, 2012. 352 p. (in Russ.)

[9] Starikov A. N. Fundamentals of the Theory of Combustion and Explosion. Vladimir: VlSU Publishing House, 2019. 148 p. (in Russ.)

[10] Listvin A. V., Listvin V. N. Reflectometry of optical fibers. Moscow: LESARart, 2005. 208 p. (in Russ.)

[11] Zenevich A.O. Detection of optical fiber bends near welded and mechanical joints / A.O. Zenevich, E.V. Novikov, T.A. Matkovskaya, O.Yu. Gorbaday, G.V. Vasilevsky // Problems of infocommunications. 2022; 2(16):32-38.

[12] Recommendation ITU-T G.652. 11/2016. Characteristics of a single-mode optical fibre and cable.

[13] Recommendation ITU-T G.655. 11/2009. Characteristics of a non-zero dispersion-shifted single-mode optical fibre and cable.

[14] Recommendation ITU-T G.657. 11/2016. Characteristics of a bending-loss insensitive single-mode optical fibre and cable.

COMPARATIVE ANALYSIS OF LANET AND FE-NET IN IMAGE SEGMENTATION

Zhao Di[1], Alevtina Gourinovitch[1], Shadiye Sultanova[2],

Baloghlan Najafov[2], Gizilgul Israfilova[2]

•

[1]Belarussian State University, Minsk, Belarus
[2]Azerbaijan Technical University, Baku, Azerbaijan
3189124246@qq.com, alevt5555@gmail.com, shadiye.sultanova@aztu.edu.az,
baloglan.necefov@aztu.edu.az, qizilgul.israfilova@aztu.edu.az

Abstract

Image segmentation in infocommunications helps analyze infrastructure, enhance video quality, manage networks to improve efficiency and service quality. Biomedical image segmentation is very important for the medical imaging. It requires precise delineation of anatomical structures and pathological regions for various diagnostic tasks (for example, tumor detection and treatment planning) in improving clinical outcomes. In recent decades, the automatic medical segmentation methods which are based on deep learning (DL) models, such as convolutional neural network (CNN) architectures, have facilitated the automatic delineation of organ and lesion boundaries. This changed traditional manual segmentation approaches and improved efficiency and accuracy in clinical practice. This paper presents a comparative analysis of two DL-based models: LANet and FE-Net. LANet employs an Efficient Fusion Attention (EFA) module and an Adaptive Feature Fusion Decoder (AFF) module to improve segmentation efficiency and precision. In contrast, FE-Net integrates a Feature Awareness Module (FAM) for enhances its features to capture multi-scale and process segmentation details. The advantages of both methods in handling different scales, details, and edges were investigated and experimentally evaluated on various public datasets, and their performance in specific scenarios was assessed. The experiment showed that LANet is superior in computational efficiency and feature refinement. FE-Net shows superior performance in handling complex variations and edge details. The source code of LANet and FE-Net can be found on GitHub at [https://github.com/tyjcbzd/LANet] and [https://github.com/tyjcbzd/FE-Net].

Keywords: Image segmentation, service quality, deep learning, FE-Net, LANet.

I. Introduction

Biomedical image segmentation is crucial for various medical applications, including clinical diagnosis, treatment, and quantitative analysis, being indispensable for both clinical analysis and surgical procedures. This method involves partitioning images into multiple regions or Regions of Interest (ROIs) based on specific characteristics such as color, texture, and shape, and subsequently extracting these areas for further analysis. Accurate segmentation is vital for ensuring the quality of clinical diagnoses and treatment plans. Traditional segmentation methods, including threshold-

based approaches and region-growing algorithms, often depend on manual or semi-automatic processes [1]. These methods are not only time-consuming and labor-intensive but also susceptible to the operator's experience and subjective judgment, which limits their widespread applicability and accuracy in clinical settings. However, with the advent of deep learning technologies, the field of biomedical image segmentation has experienced a transformative shift from conventional manual methods to automated, algorithm-driven approaches that deliver higher precision and efficiency.

In recent years, the rapid advancement of deep learning techniques, particularly Convolutional Neural Networks (CNNs) with their robust feature representation capabilities, has significantly improved performance in biomedical image segmentation. U-Net, a prominent algorithm for biomedical image segmentation based on CNNs, introduced skip connections to merge multi-level features, achieving exceptional segmentation results. Numerous U-Net-based variant architectures, such as Res-UNet[2], U-Net++[3], and TransUNet [4], have been proposed for biomedical image segmentation. These models can automatically delineate organ or lesion contours, effectively overcoming the limitations of manual segmentation. The rapid development of computer hardware, coupled with advancements in deep learning, has enabled these models to demonstrate remarkable performance in the automatic segmentation of objects of interest. However, biomedical images frequently affected by noise, exhibit complex backgrounds, and display similar appearances across different tissues, which complicates the extraction of valuable information. The traditional CNN methods may face challenges in effectively balancing the capture of relevant information with the elimination of redundant features, which can lead to poor segmentation performance and inefficient feature representation. Furthermore, the fusion of low-level and high-level features poses a significant challenge for biomedical image segmentation. This kind of fusion process requires adaptability and context awareness to ensure the seamless fusion of complementary information to improve the network's ability to accurately analyze and represent the input features.

In this context, our previous efforts created two state-of-the-art (SOTA) biomedical image segmentation methods: LANet [5] and FE-Net[6]. LANet consists of an Efficient Fusion Attention (EFA) module and an Adaptive Feature Fusion (AFF) decoding block, aiming to improve the abilities of feature extraction by effectively capturing task-specific information and minimizing redundancy in channel and spatial dimensions [5]. Conversely, FE-Net introduces a Feature-Aware Module (FAM) that emphasizes important features while suppressing irrelevant ones, utilizing an encoder-decoder architecture to address multi-scale challenges in biomedical images [6]. In this paper evaluates the performance of these two models using three public datasets: Kvasir-SEG, low-grade glioma (LGG), and the 2018 Data Science Bowl, applying standard evaluation metrics including precision, Dice coefficient, mean Intersection over Union (mIoU) and recall.

II. Methods Overview

Data Description and Preprocessing

The three public datasets were used in this article: Kvasir-SEG [7], LGG [8], and the 2018 Data Science Bow l[9]. These datasets contain a variety of medical images and multiple pathological conditions. They can provide a comprehensive testing environment to evaluate the generalization abilities and robustness of the algorithms. Below is the detailed description of each dataset:

Kvasir-SEG: The dataset includes 1,000 gastrointestinal polyp images of diverse resolutions, the features segmentation masks of the database is annotated by knowledgeable gastroenterologists. These notations contain regular tissue, polyps, ulcers, and various other lesions, offering a valuable resource for segmentation purposes.

LGG: This collection comprises 1,310 RGB medical images, each having a resolution of 256×256 pixels. Acquired from the Cancer Imaging Archive at the National Cancer Institute, this compilation highlights LGG studies and showcases FLAIR mode photographs from individuals diagnosed with

TCGA. Neuroradiology experts have manually labeled irregular sections in these scans to aid in the development of automated segmentation techniques.

2018 Data Science Bowl: This collection comprises 670 cellular images, each featuring a 256×256 pixels clarity, showcasing a varied spectrum of cellular information, encompassing both cancerous and healthy cells.

The three public datasets were divided into training, validation, and testing segments by employing random selection in an 8:1:1 ratio. The goal of this segmentation approach is to ensure the model receives enough samples in training and to assess its ability for generalization and predictive precision through validation and test datasets. Comprehensive details on each dataset are provided in Table 1.

Table 1: *The information of the dataset*

Dataset	Images	Size	Train	Validation	Test	Application
Kvasir-SEG	1000	Variables	800	100	100	Colonoscopy
LGG	1310	256 x 256	1048	131	131	Brain
2018 Data Science Bowl	670	256 x 256	530	67	67	Nuclei

In the phase of data preprocessing, as shown in Figure 1, the article standardized all images to maintain data consistency. The process entailed adjusting the image scale to fit the model's input needs and conducting normalization to reduce discrepancies due to various devices and imaging scenarios. In order to enhance the model's robust and improve its adaptability to unseen samples, the article implemented multiple techniques for data augmentation applied to the training data. These methods utilized included random cropping and rotation, which help to introduce variability and enhance training set's diversity, thus improving the model to generalize better to new data.

However, in order to maintain the integrity and fairness of the test results, the article refrained from applying any data augmentation to the test set. This approach ensures that the evaluation metrics reflect the model's true performance on unaltered and real-world data.

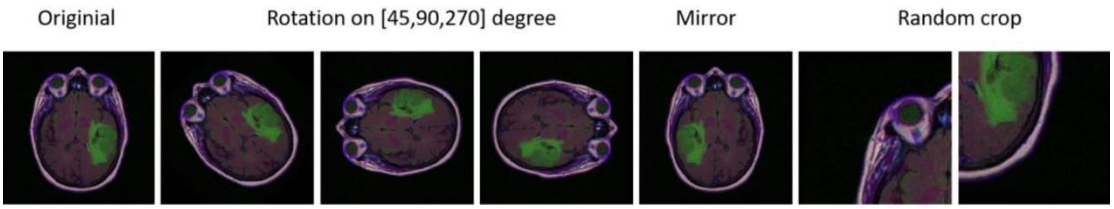
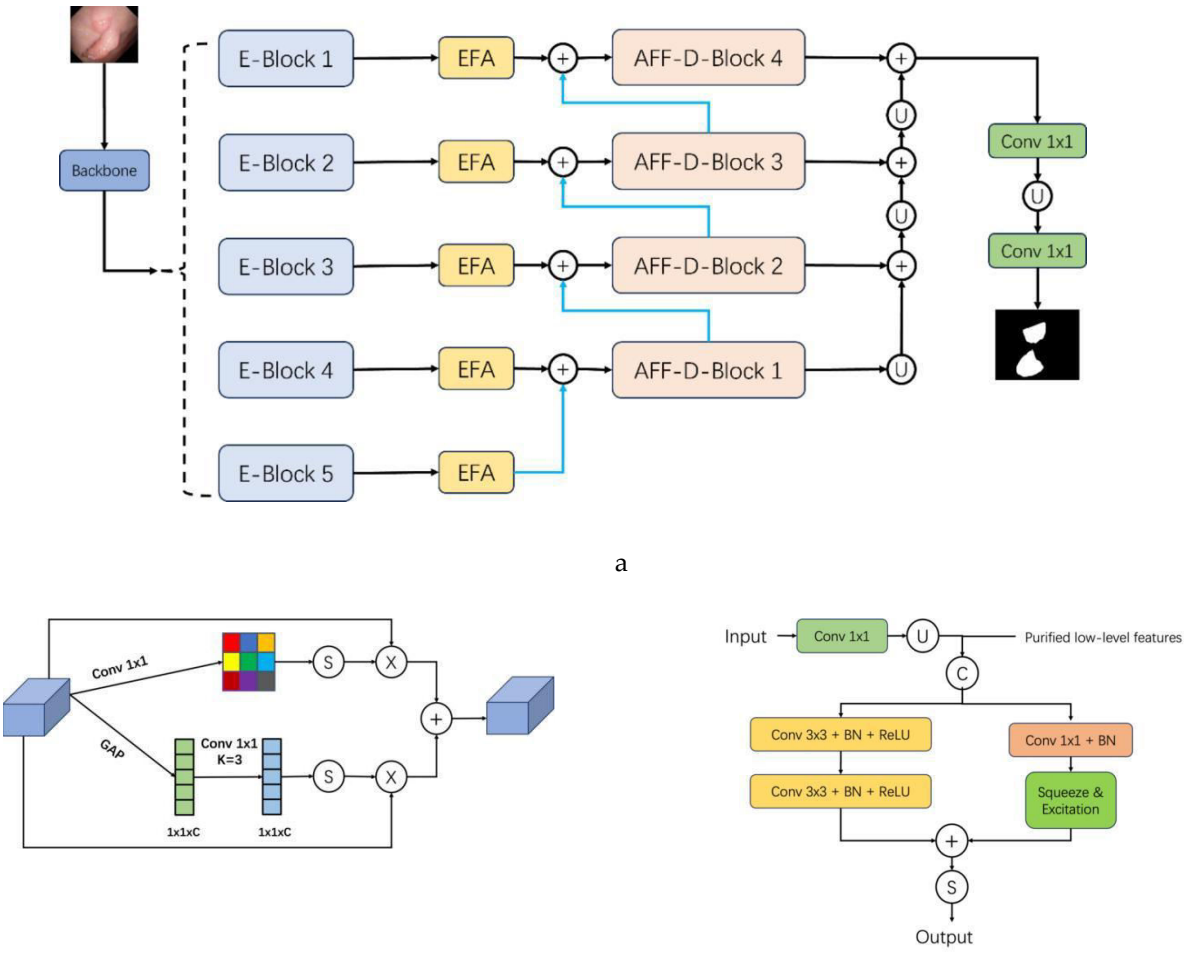


Figure 1: *Data augmentation techniques applied on the LGG dataset*

III. Model Architecture

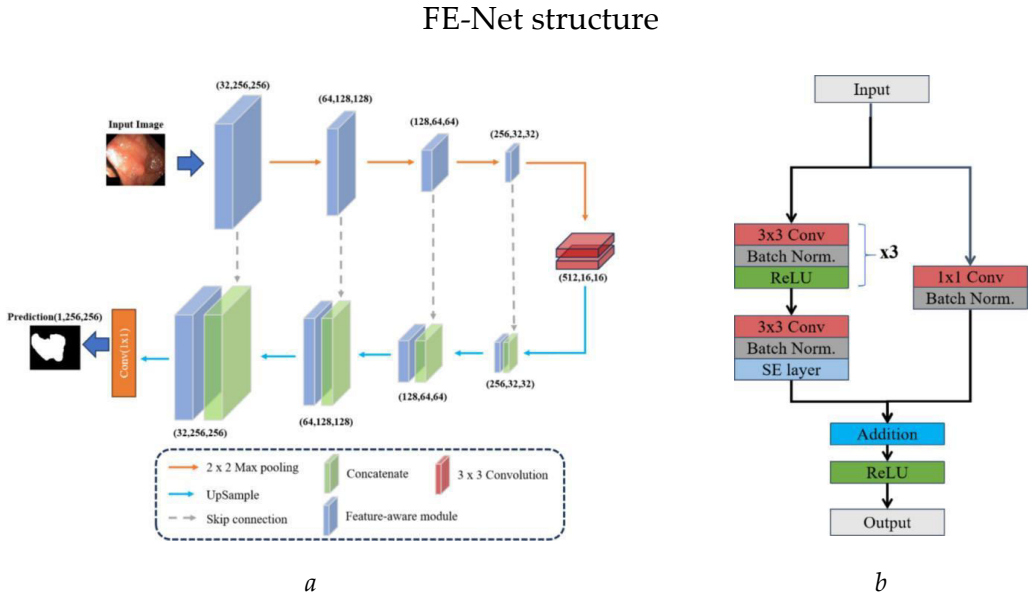
LANet structure

As shown in Figure 2, LANet is based on a lightweight MobileViT backbone network, incorporating the EFA module and the AFF decoding module. The EFA module extracts key task-relevant features and reduces redundancy through channel and spatial attention. Meanwhile, the AFF module enhances feature representation by combining low-level and high-level features from the encoding path during the decoding process. The lightweight design of LANet makes it suitable for resource-constrained environments, such as real-time clinical devices. The LANet source code is hosted on GitHub at [https://github.com/tyjcbzd/LANet].



a b c

Figure 2: *The structure of LANet*
a- LANet architecture overview; b- EFA Block structure; c- AFF-D Block structure.



a *b*

Figure 3: *The architecture of the proposed network and module:*
a - FE-Net architecture; b - feature-aware module flowchart

As illustrated in Fig.3, FE-Net utilizes a traditional encoder-decoder architecture, augmented by the incorporation of a FAM aimed at improving detail segmentation. The FAM module utilizes attention mechanisms to highlight important features and reduce noise during feature extraction, and it also uses skip connections to facilitate the fusion of low-level and high-level features. In contrast to other segmentation models, FE-Net uses bilinear interpolation instead of transposed convolution during the decoding phase. This strategic choice not only reduces computational complexity but also preserves the smoothness of the resulting images. The FE-Net source code is hosted on GitHub at [https://github.com/tyjcbzd/FE-Net].

IV. Evaluation Metrics

To assess the effectiveness of LANet and FE-Net, this study employs a range of standard evaluation metrics, including precision, the Dice coefficient (commonly referred to as the F1 score), recall, and mIoU. The Dice coefficient serves as a measure of the accuracy of the segmentation results, while the Intersection over Union (IoU) quantifies the degree of overlap between the model's predictions and the ground truth annotations. Precision reflects the model's accuracy in identifying target objects, whereas recall indicates the model's effectiveness in detecting all relevant targets. All evaluation metrics are derived from true positives (TP) and false positives (FP).

1. Precision: Precision measures the proportion of TP in the predicted segmentation results, in particular the ratio of actual positives among the samples predicted as positive. A high precision indicates that there are fewer FP in the model's predictions. Precision is calculated using the following formula:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

2. Dice Coefficient (F1 Score): The Dice score measures the accuracy by comparing the overlap region between the predicted segmentation results and the ground truth annotations. It quantifies the model's ability to match the true segmentation in biomedical image analysis tasks. The score reflects the extent of overlap, with scores between 0 and 1. The higher scores indicate that a closer match between the model's output and the actual labels. The Dice coefficient is computed with the following formula:

$$\text{Dice} = \frac{2TP}{2TP + FP + FN} \tag{2}$$

3. Recall: Recall measures the ratio of true positive samples that the model successfully detected, specifically the ratio of correctly detected true positives among all true positive samples. A high recall rate indicates that the model can identify more true positive samples, thereby reducing the number of false negatives (FN). Recall can be calculated using the following formula:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

4. MIoU: MIoU is a metric used to evaluate the performance of image segmentation tasks.It evaluate the degree of overlap between the predicted results from a model and the true labels by calculating the ratio of the intersection to the union for each category and then averaging this ratio across all categories. MIoU is commonly used in multi-class segmentation tasks as a comprehensive performance metric. Assuming there is a total of k categories, the calculation formula is as follows:

$$\mathrm{m}IOU = \frac{1}{\mathrm{k}} \sum_{i=0}^{k} \frac{TP}{TP + FP + FN} \tag{4}$$

V. Experiment results overview
Quantitative comparison

Table 2: *Comparison of experimental results on three different medical datasets namely LGG, Kvasir-SEG and 2018 Data Science Bowl. The best results are in bold.*

Dataset	Module	mIoU	Dice	Precision	Recall
LGG	LANet	0.854	0.906	0.926	0.866
	FE-Net	0.914	0.963	0.988	0.962
Kvasir-SEG	LANet	0.851	0.911	0.949	0.903
	FE-Net	0.830	0.889	0.915	0.901
2018 Data Science Bowl	LANet	0.871	0.930	0.946	0.918
	FE-Net	0.864	0.924	0.986	0.939

1. Performance on the Kvasir-SEG Dataset: On the Kvasir-SEG dataset, the performance difference between LANet and FE-Net decreased. LANet achieved a mIoU of 0.851, while FE-Net scored 0.830, indicating that both of the two models performed comparably, with LANet having a slight edge. However, in terms of precision, LANet outperformed FE-Net with a value of 0.949, which underscores higher accuracy of LANet in identifying positive samples. In terms of recall, FE-Net slightly surpassed LANet with a value of 0.903 compared to the value of 0.901 of LANet. It is clear that FE-Net had a slight advantage in the recall of detecting positive samples.

2. Performance on the LGG Dataset: LANet achieved a mIoU value of 0.854 On this dataset, while FE-Net exceeded this with an mIoU value of 0.914. This result indicates that FE-Net is more effective in identifying and segmenting lesion areas. Regarding Dice coefficient, precision, and recall, FE-Net also outperformed LANet, achieving values of 0.963, 0.988, and 0.962, respectively, compared to the values of 0.906, 0.926, and 0.866 of LANet. The numerous benefits of FE-Net across various metrics further confirm its dominance in precise segmentation.

3. Performance on the 2018 Data Science Bowl Dataset: On the dataset, the performance of LANet and FE-Net diverged from the trends observed in the previous datasets. LANet achieved a mIoU of 0.871, while FE-Net scored 0.864,which can indicate a minimal variance. However, regarding the Dice coefficient, accuracy, and recall, FE-Net outperformed LANet with values of 0.924, 0.986, and 0.939. In comparison, LANet scored 0.930, 0.946, and 0.918. These results demonstrate that FE-Net's strengths in these metrics reflect its adeptness in controlling segmentation details and boundaries.

VI. Qualitative Analysis

Figure 4 shows a comparison of the segmentation results of LANet and FE-Net on different datasets. The unique strengths of each model are evident. LANet is particularly effective at segmenting larger targets. FE-Net excels at handling details and edges. For example, in the segmentation of gastrointestinal polyp images from the Kvasir-SEG dataset, LANet achieved impressive boundary alignment, resulting in a precision value of 0.949. On the other hand, FE-Net demonstrated a greater ability to capture finer details, achieving a recall value of 0.915, which highlights its sensitivity to subtle features. When processing cell images with complex boundaries, FE-Net was able to segment these boundaries more accurately, while LANet showed higher overall

accuracy in identifying larger targets. LANet excelled in detecting larger tumors, in contrast to FE-Net, who was more adept at detecting smaller or initial stage tumors.

LANet and FE-Net demonstrated good adaptability in various imaging techniques. However, FE-Net outperformed LANet in processing high-contrast images. Additionally, LANet exhibited robust in dealing with noise, particularly with noise patterns commonly found in MRI images. Despite FE-Net is more sensitive to noise in some situations, its feature perception module can reduce noise impact by learning contextual information from the medical images.

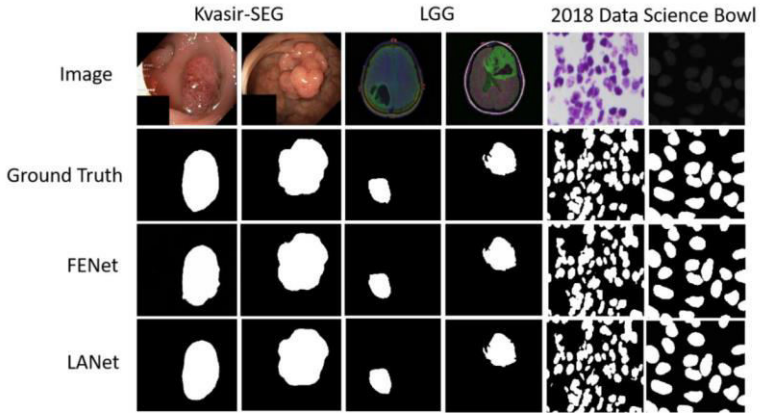


Figure 3: *The comparison of the segmentation results of LANet and FE-Net*

VII. Conclusion

The study was conducted jointly with the Department of Magnetic Resonance Imaging of the N. N. Alexandrov National Oncology Center.

It provides a comparative analysis of LANet and FE-Net in biomedical image segmentation. This study highlights joint efforts to improve biomedical image analysis capabilities. Both models exhibit unique strengths and limitations. This significantly impacts their performance and applicability in practice. LANet and FE-Net showed excellent performance and unique strengths in biomedical image segmentation. However, they also demonstrated limitations that impact their performance and applicability.

LANet's lightweight architecture, augmented with EFA and AFF modules, optimizes computational efficiency and processing speed, making it suitable for resource-constrained settings such as mobile devices and real-time clinical applications. Accuracy may not be achieved in high-resolution segmentation tasks and datasets with significant variability and noise.

FE-Net with FAM excels in capturing multi-scale features and segmentation detail. This is especially true for images with complex boundaries and fine structures. The inclusion of bilinear interpolation in FE-Net reduces the computational burden while maintaining image smoothness, which is an important factor for real-time applications. However, the complexity of FAM increases the computational requirements, potentially hindering its deployment in resource-constrained environments. Moreover, FE-Net's sensitivity to noise poses challenges in noisy imaging modalities such as low-dose CT or MRI. Its detailed feature extraction increases the risk of overfitting when trained on limited datasets.

Both methods show significant promise for the task of biomedical image segmentation. Future research can aim to integrate their respective strengths to obtain more efficient and accurate segmentation results.

References

[1] X. Liu, Z. Deng, Y. Yang, Recent progress in semantic image segmentation, Artif. Intell. Rev. 52 (2019) 1089-1106.

[2] X. Xiao, S. Lian, Z. Luo, et al., Weighted Res-UNet for high-quality retina vessel segmentation, 2018 9th International Conference on Information Technology in Medicine and Education (ITME), IEEE, (2018) 327-331.

[3] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh,, & J. Liang, Unet++: A nested u-net architecture for medical image segmentation , Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20,2018, Proceedings 4.Springer International Publishing,(2018) 3-11.

[4] J. Chen, Y. Lu, Q. Yu, et al., TransUNet: Transformers make strong encoders for medical image segmentation, arXiv preprint arXiv:2102.04306, (2021).

[5] Y. Tang, D. Pertsau, D. Zhao, D. Kupryianava, M. Tatur, LANet: Lightweight Attention Network for Medical Image Segmentation, Information Technologies and Their Applications. ITTA 2024. Communications in Computer and Information Science, Springer, Cham.(2025) 213-227.https://doi.org/10.1007/978-3-031-73420-5_18

[6] D. Zhao., Y. Tang, A.B. Gourinovitch, Effective Algorithm for Biomedical Image Segmentation. Doklady BGUIR. (2024) 84-92. https://doi.org/10.35596/1729-7648-2024-22-3-84-92.

[7] D. Jha, P.H. Smedsrud, M.A. Riegler, et al., Kvasir-SEG: A segmented polyp dataset, Multi Media Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020, Proceedings, Part II 26, Springer International Publishing, (2020) 451-462.

[8] M. Buda, A. Saha, M.A. Mazurowski, Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm, Comput. Biol. Med. 109 (2019) 218–225.

[9] J. C. Caicedo, A. Goodman, K. W. Karhohs, B. A. Cimini, J. Ackerman, M. Haghighi, A. E. Carpenter, Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. Nature methods,(2019) 1247-1253.

TWO-LAYER CHARCOAL-CONTAINING MICROWAVE ABSORBERS WITH A RELIEF SURFACE FOR SERVER EQUIPMENT PROTECTION FROM INTERFERENCE

Olga Boiprav[1], Vadim Bogush[1], Mehman Hasanov[2],
Vyacheslav Mokerov[1], Elena Belousova[1]

•

[1]Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus
[2]Azerbaijan Technical University, Baku, Azerbaijan
smu@bsuir.by, bogush@bsuir.by, mehman.hasanov@aztu.edu.az, vyacheslav.mokerov@mail.ru, belousova@bsuir.b

Abstract

The article presents the electromagnetic radiation reflection and transmission characteristics in the frequency range of 0.7–17.0 GHz of multilayer microwave absorbers. These absorbers consist of modules with a relief surface made of a mixture of powdered activated charcoal and a binder (polyvinyl acetate dispersion aqueous solution or polyurethane mastic). According to the presented characteristics, electromagnetic radiation reflection and transmission coefficients values in the frequency range of 0.7–17.0 GHz of the specified absorbers vary, respectively, within the limits from –2.0 to –18.0 dB and from –10.0 to –40.0 dB. The studied absorbers can be used for wall cladding or creating internal partitions in server rooms. Using such absorbers, it is possible to solve the practical problem of ensuring electromagnetic compatibility of server equipment and other information processing equipment.

Keywords: charcoal, interference, microwave absorber, reflection coefficient, transmission coefficient.

I. Introduction

Some of the tasks currently being solved when creating server rooms are:

– ensuring electromagnetic compatibility of equipment located inside these rooms;

– protection of equipment located inside these rooms from external electromagnetic interference.

As a rule, the solution to this problem is associated with the use of materials that absorb electromagnetic radiation (EMR) energy, i.e. microwave absorbers. Most of the microwave absorbers currently being developed and studied are characterized by a relief surface. This is due to the fact that such absorbers, compared to absorbers with a smooth surface, are characterized by lower the EMR reflection coefficient values due to the fact that:

– the relief elements of the surfaces of such absorbers ensure the scattering of electromagnetic waves interacting with them [1];

– the relief elements of the surfaces of such absorbers are a set of conditional resonators that ensure the absorption of EMR energy at certain frequencies. In addition, when electromagnetic waves interact with absorbers with a relief surface, as a rule, there is no formation of standing electromagnetic waves and / or passive electromagnetic interference [2]. The designated waves and

interference can affect the performance of the equipment.

As shown in [3–6], the scattering of electromagnetic waves by the relief elements of the surfaces of the microwave absorbers is ensured both in the case when the size of these elements is comparable with the wavelengths and in the case when it is significantly smaller than these lengths. In this regard, microwave absorbers with a relief surface are divided into the following types.

1. Perforated absorbers, i.e. absorbers with regularly or irregularly placed through or blind holes. The presence of holes in materials causes a decrease in their impedance, which in turn causes a decrease in the value of their EMR reflection coefficient.

2. Absorbers, the surface of which is characterized by the presence of nano- and / or microroughnesses, distributed uniformly or unevenly. The surface area of such absorbers exceeds the surface area of the absorbers with a relatively smooth surface, due to which the value of the EMR reflection coefficient value of the first of the designated absorbers is lower than EMR reflection coefficient value of the second of the designated absorbers.

3. Absorbers with a profiled surface, which is a set of identical or unequal bulges and / or depressions, the size of which is from units to tens of centimeters.

The following technologies are used to manufacture microwave absorbers with a relief surface of the listed types.

1. Perforation (for the manufacture of the absorbers of the first of the above types).

2. Electrochemical treatment of material surfaces. The procedure for implementing this technology is as follows:

– the surface of the material is immersed in an electrolytic solution;

– electric current is applied to the material, which leads to the formation of a relief surface.

This technology is used to manufacture absorbers of the second of the above types.

3. Photolithography. The procedure for implementing this technology is as follows:

– applying a photosensitive layer to the surface of the material;

– exposure of the applied layer to ultraviolet radiation through a mask that creates the desired pattern;

– treatment of the surface of the material with a solution that helps remove those areas of the photosensitive layer that were not exposed to ultraviolet radiation.

This technology is used to manufacture absorbers of the second of the above types.

4. Laser cutting of materials (used to manufacture absorbers of the third of the above types).

It should be noted that at present, microwave absorbers with a relief surface of the third of the above types are most often developed, researched and used, and accordingly, the fourth of the above technologies is used for their manufacture. It should be noted that the main disadvantage of such absorbers, as a rule, is their low mechanical strength, due to the fact that porous carbon-containing materials are usually used for their manufacture [7].

The paper [8] presents the results of the development and study of the microwave absorbers with a relief surface, which are not characterized by the indicated disadvantage. This is due to the fact that such absorbers are modules in the form of solid forms with hemispherical depressions filled with a mixture of powdered activated charcoal and a binder (polyvinyl acetate dispersion aqueous solution, polyurethane mastic, gypsum binder).

The advantage of powdered activated charcoal over other carbon-containing materials currently used to manufacture microwave absorbers is its low cost and high availability. The low cost of powdered activated charcoal is due to the fact that it is usually made from industrial or agricultural waste [9–11].

The choice of the type of solid forms is due to the absence of acute-angled vertices in them, which reduces the time and financial costs of manufacturing such forms. In addition, such forms provide higher mechanical strength of microwave absorbers, in the structure of which they are included. The height and width of the recesses in these forms meet the requirements for the

dimensions of the relief elements of the surface of microwave absorbers [12].

The presented study is a continuation of the study, the results of which are presented in [8]. Its purpose was to establish the patterns of change in the EMR reflection and transmission coefficients values in the frequency range of 0.7–17.0 GHz of two-layer microwave absorbers in the form of sets of the above-mentioned modules depending on:

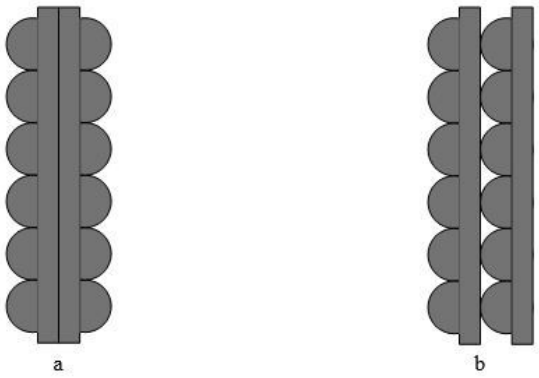
– the type of binder included in the composition of such absorbers;

– the mutual arrangement of the modules of which such absorbers consist.

Taking into account the designated patterns, it is possible to establish how EMR reflection and transmission coefficients values of the absorbers presented in work [8] change as a result of the inclusion of an additional layer in their structure.

II. Research methodology

Four groups of experimental samples were prepared for the study. Figure 1 shows a schematic representation of these samples.



a – samples of groups 1, 3; b – samples of groups 2, 4

Figure 1: *Schematic representation of the developed samples*

The appearance of one of the modules used to manufacture the experimental samples is shown in Figure 2.

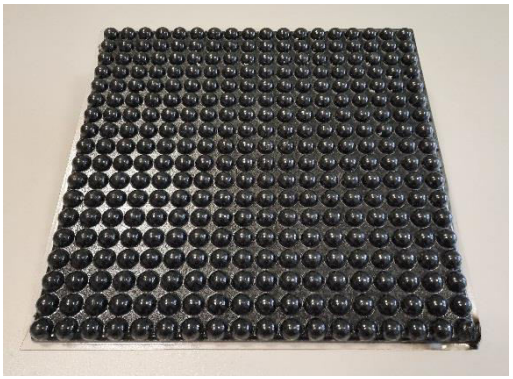


Figure 2: *The appearance of one of the modules used to manufacture the experimental samples*

Table 1 presents the characteristics of the samples of each group.

Table 1: *The characteristics of the samples of each group*

Conventional designation of the sample	Composition of the mixture used to prepare the sample	Weight of 1 m2 sample, kg
Sample of group 1	Polyvinyl acetate dispersion aqueous solution, powdered activated charcoal	10.0
Sample of group 2		
Sample of group 3	Polyurethane mastic, powdered activated charcoal	14.5
Sample of group 4		

The measurements of EMR reflection and transmission coefficients values (S11 and S21, respectively) of the manufactured samples were carried out in the frequency range of 0.7–17.0 GHz using a setup that included a personal computer, a panoramic meter of transmission and reflection coefficients SNA 0.01–18, and two horn antennas. The measurements were carried out in accordance with the methodology presented in [13].

The frequency dependences of the EMR reflection coefficient in the range of 0.7–17.0 GHz of the manufactured samples are shown in Figure 3.

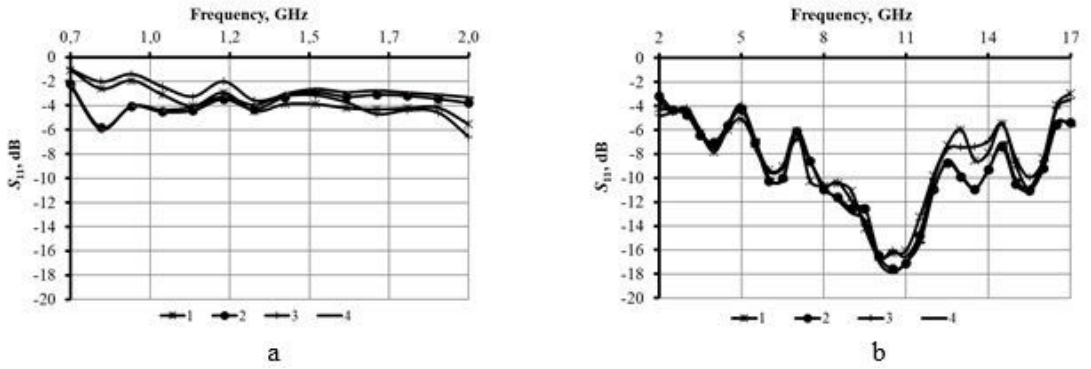


Figure 3: *The frequency dependences of the EMR reflection coefficient in the range 0.7–2.0 GHz (a) and 2.0–17.0 GHz (b) for samples of groups 1, 2, 3 and 4 (curves 1, 2, 3 and 4 respectively)*

It is evident from Figure 3 that EMR reflection coefficient values in the frequency range of 0.7–2.0 GHz for all the manufactured samples vary within the range from –2.0 to –6.0 dB, and in the frequency range of 2.0–17.0 GHz – from –4.0 to –18.0 dB. It follows from Figure 3 that EMR reflection coefficient values in the specified frequency range for the manufactured samples don't depend on either the type of binder included in their composition or the relative position of the modules used to manufacture them. Consequently, most of the electromagnetic waves characterized by frequencies whose values belong to the frequency range specified above are scattered by the relief elements of the outer layer of the manufactured samples and are scattered or absorbed by the second layer of these samples.

The frequency dependences of the EMR transmission coefficient in the range of 0.7–17.0 GHz for the manufactured samples are shown in Figure 4.

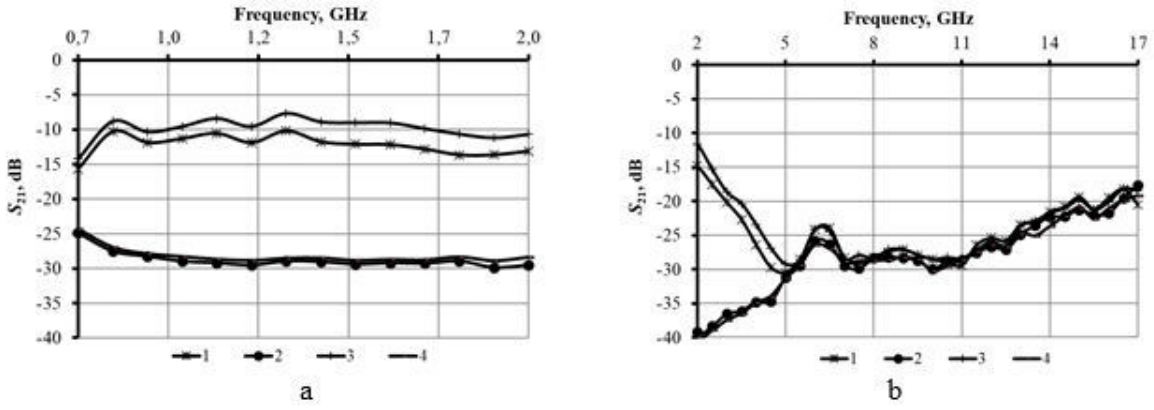
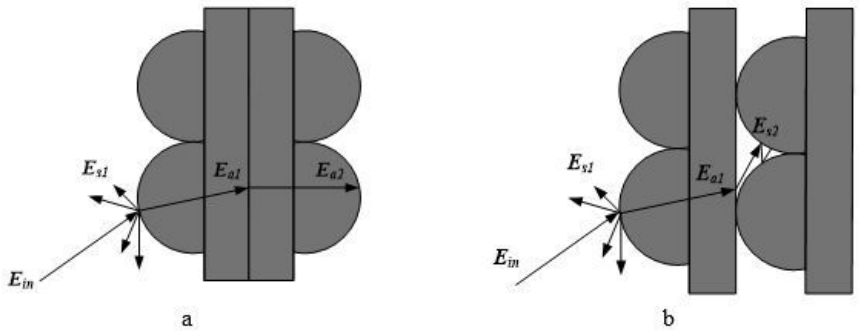


Figure 4: *The frequency dependences of the EMR transmission coefficient in the range 0.7–2.0 GHz (a) and 2.0–17.0 GHz (b)for samples of groups 1, 2, 3 and 4 (curves 1, 2, 3 and 4 respectively)*

It is evident from Figure 4 that EMR transmission coefficient values in the frequency range of 0.7–2.0 GHz for samples 1 and 3 vary, respectively, within the range of –10.0 to –15.0 dB and from –8.0 to –15.0 dB, and for samples 2 and 4 – from –25.0 to –30.0 dB. EMR transmission coefficient values in the frequency range of 2.0–17.0 GHz for samples 1 and 3 vary, respectively, within the ranges from –16.0 to –30.0 dB, from –12.0 to –30.0 dB, and for samples of groups 2 and 4 – from –20.0 to –40.0 dB. Sample of group 3 is characterized by higher EMR transmission coefficient values compared to sample of group 1 due to the fact that its wave impedance is lower than the wave impedance of sample of group 1. This feature is due to the fact that the relative permittivity value of the polyurethane mastic, using which sample of group 3 was manufactured, is higher than the relative permittivity value of the polyvinyl acetate dispersion aqueous solution, using which sample of group 1 was manufactured [9]. Lower EMR transmission coefficient values in the frequency range of 0.7–17.0 GHz of samples of groups 2, 4 compared to samples of groups 1, 3 may be due to the fact that the energy of electromagnetic waves that can be scattered by the relief elements of the surface of the second layer of samples of groups 2, 4 is higher than the energy of electromagnetic waves that can be absorbed in the thickness of the second layer of samples of groups 1, 3 (Figure 5).



Ein – incident wave; Es1 – wave scattered by the first layer; Ea1 – wave absorbed by the first layer; Es2 – wave scattered by the second layer; Ea2 – wave absorbed by the second layer

Figure 5: *Schematic representation of the mechanisms of electromagnetic waves interaction with samples of groups 1, 3 (a) and samples of groups 2, 4 (b)*

It should be noted that EMR transmission coefficient values in the frequency range of 0.7–17.0 GHz of the manufactured absorbers samples are lower by 5.0–25.0 dB, than EMR transmission coefficient values of the absorbers samples, the results of the study of which are

presented in the paper [8] (Figure 6). EMR reflection coefficient values in the frequency range of 0.7–17.0 GHz of the first of the designated samples are practically equivalent to EMR reflection coefficient values in the specified frequency range of the second of the designated samples. Thus, the inclusion of an additional layer in the structure of the absorbers presented in the work [8] seems advisable.

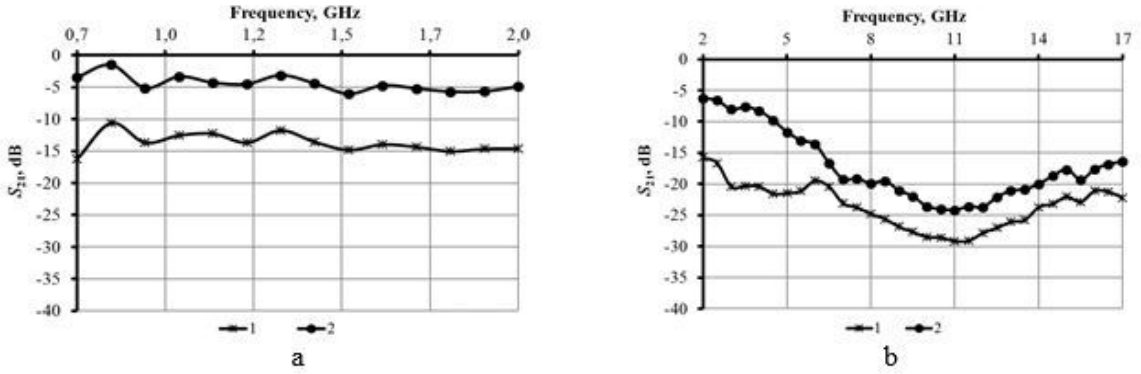


Figure 6: *The frequency dependences of the EMR transmission coefficient in the range 0.7–2.0 GHz (a) and 2.0–17.0 GHz (b) for of the absorbers samples presented in the paper [8] and made from polyvinyl acetate dispersion aqueous solution and polyurethane mastic (curves 1 and 2 respectively)*

III. Conclusion

The studied microwave absorbers are characterized by the following advantages over their analogs.

1. A wider operating frequency range (~ 9.0 GHz) due to the fact that:

– the studied absorbers are multilayered;

– electromagnetic waves interacting with the studied absorbers are scattered both by the relief elements of their surfaces (i.e. hemispheres) and by particles of powdered activated charcoal, the size of which is 1.0–5.0 mm (i.e. comparable with the lengths of electromagnetic waves in the frequency range of 0.7–17.0 GHz).

2. Lower cost due to the fact that powdered activated charcoal, which is part of the studied absorbers, is characterized by a lower cost compared to other carbon-containing powdered materials used to manufacture their analogs.

3. Increased mechanical strength due to the fact that the structure of the studied absorbers includes solid forms made of radio-transparent polymeric material and without acute-angled vertices. The studied microwave absorbers can be used for wall cladding or creating internal partitions in server rooms. With the help of such absorbers, it is possible to solve the practical problem of ensuring electromagnetic compatibility of server equipment and other means of information processing.

References

[1] E.A. Shorokhova, A.V. Kashin, Some features of electromagnetic wave scattering from statistically rough land covers in the millimeter-wave range of wavelengths, Radiophysics and Quantum Electronics. (2005) 426-434.

[2] A. Mildred, The scattering of a standing electromagnetic wave, Journal of the Franklin Institute, (1929) 19-45.

[3] N.N. Grinchik, Electrodynamics of inhomogeneous (laminated, angular) structures, Journal of Electromagnetic Analysis and Applications. (2014) 57-105.

[4] A. Potapov, Multiple scattering of waves in fractal discrete randomly-inhomogeneous media from the point of view of radiolocation of the self-similar multiple targets, Radioelectronics. Nanosystems. Information Technologies. (2018) 3-22.

[5] M.A. Aliseyko, O.V. Boiprav, N.N. Grinchik, A.V. Tarasevich, Modeling the interaction of solitlike pulse signals with electromagnetic shields in the form of heterogeneous media, Edelweiss Chemical Science Journal. (2020) 1-5.

[6] O. Boiprav, M. Hasanov, V. Bogush, L. Lynkou, Flexible double-layered microwave absorbers based on foiled materials with mechanically treated surface, New Materials, Compounds and Applications. (2023) 100-110.

[7] J.H. Kim, A.Y. Jo, Y.J. Choi, K.B. Lee, J.S. Im, B.C. Bai, Improving the mechanical strength of carbon–carbon composites by oxidative stabilization, Journal of Materials Research and Technology. (2020) 16513-16521.

[8] O.V. Boiprav, E.S. Belousova, N.V. Bogush, S.E. Savanovich, M.M. Kasperovich, F.V. Gusinsky, I.A. Zakharov, Charcoal-containing microwave electromagnetic radiation absorbers with relief surface. Vestsi Natsyyanal'nai akademii navuk Belarusi. Seryya fizika-tekhnichnykh navuk. Proceedings of the National Academy of Sciences of Belarus.

Physical-technical series. (2024) 17-27 (in Russian).

[9] W.T. Gore, Energy and charcoal production from solid waste generated in the south African pulp and paper industry, South African Forestry Journal. (1986) 10.1080/00382167.1986.9629630.

[10] J.M. Dias, M.C.M. Alvim-Ferraz, M.F. Almeida, J. Rivera-Utrilla, M. Sanchez-Polo, Waste materials for activated carbon preparation and its use in aqueous-phase treatment: a review, J. Environ. Manage. (2007), 10.1016/j.jenvman.2007.07.031.

[11] W. Bogale, Preparation of charcoal using agricultural wastes, Ethiopian journal of education and sciences, Ethiopian Journal of Education and Sciences. (2009).

[12] E.S. Belousova, O.V. Boiprav, S.E. Savanovich, Charcoal-containing absorbers of electromagnetic radiation with hemispherical geometric inhomogeneities, Electromagnetic waves and electronic system. (2024) 22-29 (in Russian).

[13] O. Boiprav, H. Ayad, S.A. Abdaljlil, L. Lynkou, M. Abdulmawlay, Charcoal- and foil-containing materials for radio electronic control systems protection from electromagnetic interferences, 2022 IEEE 21st international Ccnference on Sciences and Techniques of Automatic Control and Computer Engineering. (2022) 299-304.

MODELING THE TOPOLOGY OF FSON USING DIJKSTRA'S ALGORITHM

Agil Movsumov[1], Shadiya Sultanova[1], Emin Payizov[1], Turana Rasullu[1], Nurana Camalzadeh[1], Javid Namazov[1], Mala Dutta[2]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan
[2]Assam down town University, Assam, India
agil.movsumov@aztu.edu.az, shadiye.sultanova@aztu.edu.az,
emin.payizov.sh@student.aztu.edu.az, turana.rasullu.k@student.aztu.edu.az,
nurana002@gmail.com, namazovjavid@gmail.com, Mala.dutta@adtu.in

Abstract

This paper focuses on modeling the topology of Free Space Optical Networks (FSON) using Dijkstra's algorithm, demonstrating its potential for efficient and reliable data transmission in distributed communication systems. The work highlights the integration of Li-Fi technology and graph theory to optimize routing and minimize time costs, ensuring adaptability and scalability for real-world applications. By addressinFg challenges like interference and line-of-sight constraints, the proposed methodology enhances network performance for smart cities, IoT systems, and space communications. The results confirm the effectiveness of the approach in improving bandwidth efficiency, reducing delays, and dynamically adapting to changing network conditions.

Keywords: Free Space Optical Networks, Dijkstra's Algorithm, Network Topology, Li-Fi Technology, Wireless Communication

I. Introduction

In the context of the rapid growth of data transmission volumes and the increasing number of connected devices, the demands on wireless communication technologies have risen significantly. Traditional radio frequency channels, despite their widespread use, face several limitations, including spectrum congestion, high energy consumption, and scalability challenges in dense urban environments. In response to these challenges, optical data transmission technologies, such as Li-Fi (Light Fidelity) and Free Space Optical Networks (FSON), have garnered particular interest [1-8].

Li-Fi utilizes visible light, ultraviolet, or infrared waves for data transmission, offering unique advantages such as high-speed communication, low latency, and reliability. This technology is actively being implemented in smart city systems, industrial automation, and even household devices. It is particularly useful for indoor data transmission, such as in offices and residential buildings, as well as for creating high-speed links between mobile devices.

On the other hand, FSON focuses on data transmission through open space using laser beams. These systems can provide high-speed connectivity over long distances, making them a promising solution for space communications, inter-building networks in densely populated metropolises, and

satellite connectivity. A key advantage of FSON is its lack of dependence on physical infrastructure, which reduces costs and simplifies network deployment. However, these systems are sensitive to external factors, such as weather conditions and the need for line-of-sight communication.

Thus, Li-Fi and FSON technologies represent two complementary approaches to addressing modern communication challenges. Their use not only expands data transmission capabilities but also allows adaptation to various scenarios, including integration into existing systems, deployment in remote areas, and even applications in space.

Figure 1 illustrates the topology diagram developed based on graph theory. As shown in the figure, the FSON network establishes connections between the monitoring center, optical locator, low-Earth orbit satellites, aircraft, drones, and buildings within line-of-sight range. To ensure reliable communication among all system elements, various wavelengths are employed, minimizing interference and enhancing network capacity.

Specifically, the wavelengths used for communication between the optical locator and satellites are denoted as $\lambda_1, \lambda_2, \dots, \lambda_n, \lambda_d, \lambda_p$; between satellites as $\lambda'_1, \dots, \lambda'_n$; between satellites and buildings as $\lambda''_1, \dots, \lambda''_n$; between buildings as $\lambda^B_1, \dots, \lambda^B_n$; between the optical locator and buildings as $\lambda'''_1, \dots, \lambda'''_n$; and between the monitoring center and the optical locator as λ'_m. These wavelengths enable independent data transmission between nodes, improving the overall efficiency of the network.

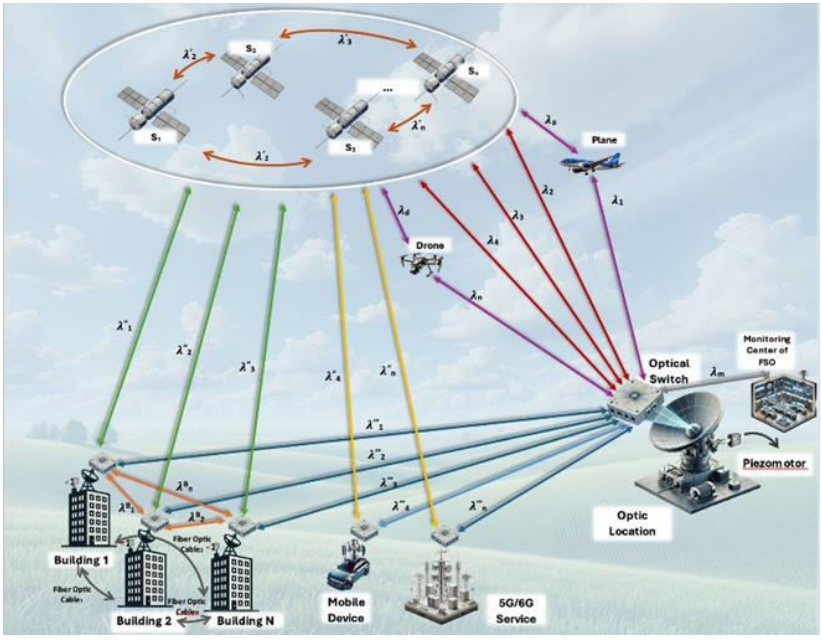


Figure 1: *Data Transmission Scheme in a Network Using Li-Fi and Satellites*

II. Network Topology and Routing Parameters

Modern data transmission systems impose stringent requirements on network topology and routing principles [1-6]. In distributed systems such as Free Space Optical Networks (FSON), efficient organization of communication between nodes is critically important. This chapter describes the network topology, based on graph theory, and the parameters that determine optimal data transmission routes.

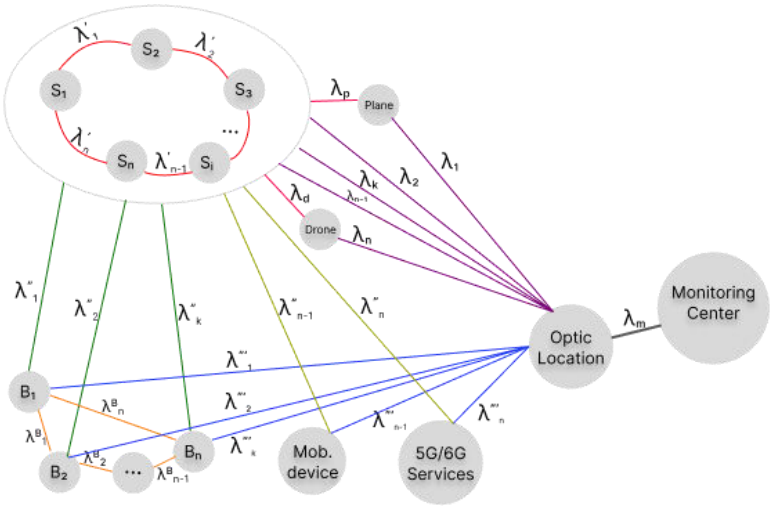


Figure 2: *FSON Network Topology with Nodes and Wavelengths*

Representing a network as a graph is one of the most effective ways to model complex interconnections [9]. In this topology, each node corresponds to an active network element, and the edges represent data transmission lines between the nodes. Fig. 2 illustrates a network structure in which communication is organized between:

- Buildings ($B_1, B_2, \dots, B_n$),
- Low-Earth orbit satellites ($S_1, S_2, \dots, S_n$),
- Mobile devices,
- Drones and aircraft,
- The monitoring center via an optical locator.

Each edge of the graph is characterized by the wavelength used for data transmission, minimizing interference between communication lines.

For instance:

- Wavelengths between satellites are denoted as $\lambda'_1, \lambda'_2, \dots, \lambda'_n$,
- Wavelengths between satellites and buildings as $\lambda''_1, \dots, \lambda''_n$,
- Wavelengths between a building and the locator as $\lambda'''_1, \dots, \lambda'''_n$.

One of the primary constraints of such a network is the requirement for line-of-sight communication between nodes. This imposes limitations on node placement and interaction but enable s significantly higher data transmission speeds by using laser beams and minimizing signal loss.

The use of graph-based topology provides several key advantages:

- Flexibility: The ability to scale the network by adding new nodes without altering the existing structure,
- Efficiency: Minimization of data transmission delays through optimized routing,
- Reliability: The availability of alternative routes for data transmission reduces the risk of connection disruptions in case of individual node failures.

The system supports a high degree of adaptation to external conditions, such as weather changes, by utilizing alternative routes based on Dijkstra's algorithm (described in the next chapter).

To ensure efficient network operation, the following parameters are used to define data transmission routes between nodes:

- d_i: The distance between node i and $i + 1$, which directly impacts the data transmission speed,
- S_i: The data transmission speed for a given network segment (e.g., via Li-Fi),
- δ_i: The data processing delay at node i, depending on the type of equipment and the number

of operations at the node.

To calculate the total data transmission time, the following formula is used:

$$T_{total} = \sum_{i=1}^{n} \left(\frac{d_i}{S_i} + \delta_i \right) \quad (1)$$

Where T_total – total time required for data transmission between two nodes. This formula accounts for all route parameters, enabling precise calculation of the time required for data transmission.

Impact of Parameters:

• Distance (d_i): The larger the distance between nodes, the greater the time required. Therefore, minimizing route length is one of the optimization goals,

• Data Transmission Speed (S_i): Depends on the technology used in the network (e.g., Li-Fi, laser communication lines). High-speed technologies significantly reduce T_{total},

• Delays (δ_i): Includes signal processing time at the node. Minimizing delays is achieved through optimization of hardware and software.

To describe the network, a weight matrix is used (Table 1), where each node is connected to others with corresponding parameters (distance, data transmission speed, delays). This matrix allows for efficient modeling of temporal costs between nodes:

• Values along the diagonal are zero, as a node has no cost to communicate with itself

•Represent the temporal costs of data transmission between nodes ($\lambda_1, \lambda_2, \ldots, \lambda_n$) through the respective communication channels.

Table 1: *Weight Matrix Example*

From/To	Building 1	Building 2	…	Mobile device (Building n-1)	Cell tower (Building n)	Satellite 1	Satellite 2	…	Satellite n	Optic locator	Drone	Aircraft	Monitoring center
Building 1	0	λ^{b1}	∞	∞	$\lambda^{b}{}_{n}$	λ''^{1}	∞	∞	∞	λ'''^{1}	∞	∞	∞
Building 2	λ^{b1}	0	λ^{b2}	∞	∞	∞	λ''^{2}	∞	∞	λ'''^{2}	∞	∞	∞
…	∞	λ^{b2}	0	…	∞	∞	∞	…	∞	∞	∞	∞	∞
Mobile device (Building n-1)	∞	∞	…	0	$\lambda^{b}{}_{n}{}^{-1}$	∞	∞	∞	$\lambda''{}_{n}{}^{-1}$	$\lambda'''{}_{n}{}^{-1}$	∞	∞	∞
Cell tower (Building n)	$\lambda^{b}{}_{n}$	∞	∞	$\lambda^{b}{}_{n}{}^{-1}$	0	∞	∞	∞	$\lambda''{}_{n}$	$\lambda'''{}_{n}$	∞	∞	∞
Satellite 1	λ''^{1}	∞	∞	∞	∞	0	λ'^{1}	∞	$\lambda'{}_{n}$	$(\lambda^{1} + \lambda_{p})$	∞	λ_{p}	∞
Satellite 2	∞	λ''^{2}	∞	∞	∞	λ'^{1}	0	λ'^{2}	∞	λ^{2}	∞	∞	∞
…	∞	∞	…	∞	∞	∞	λ'^{2}	0	…	∞	∞	∞	∞
Satellite n	∞	∞	∞	$\lambda''{}_{n}{}^{-1}$	$\lambda''{}_{n}$	$\lambda'{}_{n}$	∞	…	0	$(\lambda_{n} + \lambda_{d})$	λ_{n}	∞	∞
Optic locator	λ'''^{1}	λ'''^{2}	∞	$\lambda'''{}_{n}{}^{-1}$	$\lambda'''{}_{n}$	$(\lambda^{1} + \lambda_{p})$	λ^{2}	∞	$(\lambda_{n} + \lambda_{d})$	0	λ_{n}	λ^{1}	λ_{m}
Drone	∞	∞	∞	∞	∞	∞	∞	∞	λ_{d}	λ_{n}	0	λ_{a}	∞
Aircraft	∞	∞	∞	∞	∞	λ_{p}	∞	∞	∞	λ^{1}	λ_{a}	0	∞
Monitoring center	∞	∞	∞	∞	∞	∞	∞	∞	∞	λ_{m}	∞	∞	0

III. Route Optimization and Results Analysis

Efficient data transmission in distributed communication systems is impossible without route optimization. Using Dijkstra's algorithm and the time parameters described in the previous chapter, the fastest and most stable paths between network nodes can be identified. This chapter describes the process of route optimization based on input data, the application of Dijkstra's method, and the analysis of results.

The input data for routing is based on a weight matrix reflecting the time costs of data transmission between nodes. The weight matrix considers:

• *Direct Line-of-Sight Communication:* Communication is possible only if a direct line of sight exists between nodes, a critical constraint in FSON networks.

• *Time Costs:* For each pair of nodes, time costs are calculated based on distance (d_i), transmission speed (S_i) and processing delay (δ_i). If direct communication between nodes is not possible, the cost is set to ∞.

• *Communication Channel Characteristics:* Different wavelengths ($\lambda_1, \lambda_2, \dots, \lambda_n$), are used to minimize interference and improve efficiency.

Example Weight Matrix: Table 1 demonstrates the temporal costs between nodes, where each element represents the cost of transmitting data between the respective nodes.

Dijkstra's Algorithm for Path Optimization — is a classical method for finding the shortest path in graphs. In the context of this system, the algorithm identifies routes with the minimal time cost. The main steps are as follows:

1.Initialization:

• All nodes are assigned initial cost values (∞), except for the starting node (0).

• A list of unprocessed nodes is created.

2.Node Selection with Minimal Cost:

• From the list of unprocessed nodes, the node with the lowest cost is selected.

3.Updating Neighboring Nodes:

• For each neighboring node, the cost is recalculated based on the current node.

• If the new route offers a lower cost than previously calculated, the values are updated.

4.Repeat:

• Steps 2 and 3 are repeated until all nodes are processed.

5.Route Formation:

•After completing the algorithm, the optimal path from the starting node to each other node is determined.

Table 2: *Final Table of Minimal Weights Between Network Nodes*

From/To	Building 1	Building 2	…	Mobile device Building n-1	Cell tower Building n	Satellite 1	Satellite 2	…	Satellite n	Optic locator	Drone	Aircraft	Monitoring center
Building 1		$x_{1(2)}$	…	$x_{1(n-1)}$	$x_{1(n)}$	$x_{1(n+1)}$	$x_{1(n+2)}$	…	$x_{1(2n)}$	$x_{1(k)}$	$x_{1(d)}$	$x_{1(p)}$	$x_{1(m)}$
Building 2	$x_{2(1)}$		…	$x_{2(n-1)}$	$x_{2(n)}$	$x_{2(n+1)}$	$x_{2(n+2)}$	…	$x_{2(2n)}$	$x_{2(k)}$	$x_{2(d)}$	$x_{2(p)}$	$x_{2(m)}$
…	…	…	…	…	…	…	…	…	…	…	…	…	…
Mobile device (Building n-	$x_{n-1(1)}$	$x_{n-1(2)}$	…		$x_{n-1(n)}$	$x_{n-1(n+1)}$	$x_{n-1(n+2)}$	…	$x_{n-1(2n)}$	$x_{n-1(k)}$	$x_{n-1(d)}$	$x_{n-1(p)}$	$x_{n-1(m)}$

1)													
Cell tower (Building n)	$x_{n(1)}$	$x_{n(2)}$	…	$x_{n(n-1)}$		$x_{n(n+1)}$	$x_{n(n+2)}$	…	$x_{n(2n)}$	$x_{n(k)}$	$x_{n(d)}$	$x_{n(p)}$	$x_{n(m)}$
Satellite 1	$x_{n+1(1)}$	$x_{n+1(2)}$	…	$x_{n+1(n-1)}$	$x_{n+1(n)}$		$x_{n+1(n+2)}$	…	$x_{n+1(2n)}$	$x_{n+1(k)}$	$x_{n+1(d)}$	$x_{n+1(p)}$	$x_{n+1(m)}$
Satellite 2	$x_{n+2(1)}$	$x_{n+2(2)}$	…	$x_{n+2(n-1)}$	$x_{n+2(n)}$	$x_{n+2(n+1)}$		…	$x_{n+2(2n)}$	$x_{n+2(k)}$	$x_{n+2(d)}$	$x_{n+2(p)}$	$x_{n+2(m)}$
…	…	…	…	…	…	…	…	…	…	…	…	…	…
Satellite n	$x_{2n(1)}$	$x_{2n(2)}$	…	$x_{2n(n-1)}$	$x_{2n(n)}$	$x_{2n(n+1)}$	$x_{2n(n+2)}$	…		$x_{2n(k)}$	$x_{2n(d)}$	$x_{2n(p)}$	$x_{2n(m)}$
Optic locator	$x_{k(1)}$	$x_{k(2)}$	…	$x_{k(n-1)}$	$x_{k(n)}$	$x_{k(n+1)}$	$x_{k(n+2)}$	…	$x_{k(2n)}$		$x_{k(d)}$	$x_{k(p)}$	$x_{k(m)}$
Drone	$x_{d(1)}$	$x_{d(2)}$	…	$x_{d(n-1)}$	$x_{d(n)}$	$x_{d(n+1)}$	$x_{d(n+2)}$	…	$x_{d(2n)}$	$x_{d(k)}$		$x_{d(p)}$	$x_{d(m)}$
Aircraft	$x_{p(1)}$	$x_{p(2)}$	…	$x_{p(n-1)}$	$x_{p(n)}$	$x_{p(n+1)}$	$x_{p(n+2)}$	…	$x_{p(2n)}$	$x_{p(k)}$	$x_{p(d)}$		$x_{p(m)}$
Monitoring center	$x_{m(1)}$	$x_{m(2)}$	…	$x_{m(n-1)}$	$x_{m(n)}$	$x_{m(n+1)}$	$x_{m(n+2)}$	…	$x_{m(2n)}$	$x_{m(k)}$	$x_{m(d)}$	$x_{m(p)}$	

The execution of Dijkstra's algorithm produces a routing table that provides the minimal time costs for each pair of nodes. Table 2 highlights the following:

• Displays the most efficient paths and their associated transmission times.

• Identifies key nodes with the most efficient connections, indicating their central role in the network.

• Highlights nodes where optimization is required, such as improving transmission speeds or reducing delays.

Advantages of Using Dijkstra's Algorithm:

1. Dynamic Routing: Routes are updated in real time based on current network conditions, ensuring adaptability to environmental factors or system changes.

2. Efficient Resource Utilization: The algorithm selects routes with minimal costs, maximizing the network's overall throughput.

3. Minimized Delays: By optimizing time costs, the network ensures fast and stable data transmission.

4. Scalability: The system easily adapts to the addition of new nodes or connections without losing its effectiveness.

IV. Conclusion

Data transmission route optimization is a fundamental element for the efficient operation of distributed networks. The system considered, based on Dijkstra's algorithm and Li-Fi technologies, demonstrates a high level of adaptability, reliability, and performance.

The primary advantages of such a network lie in the efficient utilization of resources by organizing communication between nodes with direct line-of-sight, minimizing signal loss, and increasing data transmission speeds. The use of multiple wavelengths significantly reduces interference and enhances bandwidth, making the network scalable and flexible for adding new nodes without requiring a complete infrastructure overhaul.

Dijkstra's algorithm enables dynamic routing by finding the shortest paths with minimal time costs, which is crucial for systems requiring high-speed and reliable data transmission. Analysis results confirm that nodes with minimal time costs become key points in the network, serving as primary data hubs, while high-cost segments highlight areas needing further optimization, such as upgrading hardware or increasing transmission speeds.

This system offers broad application prospects, including smart cities, space communications, Internet of Things (IoT) systems, and critical emergency or traffic management systems. An important area for future development is the integration with other communication technologies, consideration of external factors such as weather conditions, and the creation of automated routing systems capable of adapting to network changes in real time.

In conclusion, optimizing routes in distributed systems using Li-Fi technologies and Dijkstra's algorithm allows for the creation of a flexible, reliable, and efficient network. This network can tackle complex challenges and ensure stable data transmission even under constraints and dynamic conditions.

References

[1] Khalighi, M. A., Uysal, M. Survey on Free Space Optical Communication: A Communication Theory Perspective / M. A. Khalighi, M. Uysal // IEEE Communications Surveys & Tutorials. 2014.

Vol. 16, No. 4.pp. 2231–2258.

[2] H. Hemmati, A. Biswas, and I. B. Djordjevic, Deep-space optical communications: Future perspectives and applications, Proc. IEEE, vol. 99, no. 11, 2011, pp. 2020–2039.

[3] Himani Kaushal,Georges Kaddoum. Optical Communication in Space: Challenges and Mitigation Techniques. IEEE Communications Surveys & Tutorials, 2017, pp. 57 – 96.

[4] M.H Hasanov, NA Atayev. Early Conceptual Model of Nanosatellite with Laser Beam Control and Active Transponder System.Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO) IEEE. 2022. pp.1-4.

[5] M.H. Hasanov, N.A. Atayev. Algorithm design nanosatellite based on radio frequency and optical communication. problems of information technology (2022), vol. 13, no. 2, pp. 61-68.

[6] Carrasco-Casado, A. & Mata-Calvo, R. (2020). Free-space optical links for space communication networks. Japan. p. 66.

[7] CubeSat 101 Basic Concepts and Processes for First-Time CubeSat Developers, California, p.22.

[8] K.A.Balaji.K.Prabu.Performance evaluation of FSO system using wavelength and time diversity over Malaga turbulence channel with pointing errors.Optics Communications Volume 410, 2018, pp. 643-651.

[9] Dijkstra, E. W. A Note on Two Problems in Connexion with Graphs / E. W. Dijkstra // Numerische Mathematik.1959. Vol. 1. pp. 269–271.

COMPREHENSIVE EVALUATION AND PERFORMANCE ANALYSIS OF A DEEP LEARNING MODEL WITH HYPERPARAMETER TUNING FOR LUMPY SKIN DISEASE CLASSIFICATION IN DAIRY COWS

Gunikhan Sonowal[1], Soraisam Gobinkumar Singh[1], Prasanta Bairagi[1], Utpal Barman[1], Dulumani Das[1], Mammadov Iltimas[2], Gulnar Gurbanova[2]

•

[1]Assam down town University, Assam, India
[2]Azerbaijan Technical University, Baku, Azerbaijan
gunikhan.sonowal@gmail.com, soraisam@gmail.com, bairagi@gmail.com, utpal@gmail.com, dulumani@gmail.com, iltimas.memmedov@aztu.edu.az, gulnar.qurbanova@aztu.edu.az

Abstract

This work attempts to classify lumpy skin conditions using CNN and hyperparameter tuning. This model is comprised of many procedures, including selecting a pre-trained model, altering the architecture, and training the model on a specific dataset. During tweaking, the proposed model attained a validation accuracy of 89.73 percent. The model's generalisation performance was confirmed with an accuracy of 80.68% in the final test set evaluation. It significantly increased the timeliness of LSD identification, making it a valuable tool for farmers and veterinarians. Furthermore, a Receiver Operating Characteristic (ROC) curve with an Area Under the Curve (AUC) of 0.88 indicates that our binary classifier performed satisfactorily.

Keywords: LSD; CNN; hyperparameter tuning; Lumpy Skin Disease; Cows

I. Introduction

Lumpy Skin Disease (LSD) is a viral disease affecting cattle, characterized by nodules on the skin, fever, and other systemic symptoms. The disease can lead to severe economic losses due to decreased milk production, weight loss, and increased mortality. According to the FAOSTAT1 production data, India is the leading milk producer globally, holding the top position with a 24% share of world milk production in 2021-22. Over the past eight years, from 2014-15 to 2021-22, India's milk production has surged by 51%, reaching a total of 22 crore tonnes in 2021-22.

Traditional diagnostic methods, such as physical examination and laboratory tests, are often time-consuming and may not be feasible for large herds[5]. In recent years, deep learning techniques, particularly Convolutional Neural Networks (CNNs), have demonstrated considerable potential in the field of medical imaging and disease diagnosis [8, 7]. CNNs, a class of deep learning algorithms designed to process and analyze visual data, have shown remarkable performance in tasks such as image classification, object detection, and disease recognition. These models leverage hierarchical feature extraction and learning capabilities, allowing for improved accuracy and efficiency in detecting and classifying diseases from images.

The application of CNNs to LSD diagnosis represents a promising advancement. CNNs can automate and expedite the diagnostic process by analyzing images of cattle with high precision, potentially addressing the limitations of traditional methods. However, the effectiveness of CNNs in diagnosing LSD is contingent upon several factors, including the quality of the dataset, the architecture of the neural network, and the optimization of hyperparameters. Hyperparameter tuning is a critical step in training deep learning models, as it involves adjusting various parameters to enhance model performance and generalization capabilities. This study aims to develop and evaluate a CNN model for the early diagnosis of LSD in dairy cows. The specific contributions include:

- To preparing a comprehensive dataset of images.
- To designing and implementing a CNN architecture.
- To training and optimizing the CNN model.
- To evaluating the model based on accuracy, sensitivity, specificity, and other relevant metrics.

II. Literature Review

Lumpy Skin Disease (LSD) is a significant viral infection affecting cattle, caused by the Lumpy Skin Disease Virus (LSDV), a member of the Capripoxvirus genus. Therefore, many researchers provided a wide number of models to detect these diseases which are discussed below:

Rai et. al., [7] developed an architecture utilizing machine learning techniques for disease diagnosis and detection. This framework employs tools such as VGG-16, VGG-19, and Inception-v3 for feature extraction. The work was tested on a proprietary dataset and compared with other advanced methodologies, including kNN, SVM, NB, ANN, and LR. The results demonstrated considerable performance in feature extraction.

Girma et al., [3] developed a model for detecting and classifying Lumpy Skin Disease (LSD) in animals categorizing skin conditions into Severe, Mild, and Normal . The dataset was sourced from the Oromia region, specifically Bale Zone's Medawelabu Wereda and Arsi Zone's Chole Wereda Livestock Production Offices, as well as from an external image repository on the internet. Experimental results indicate that the Support Vector Machine (SVM) classifier out- performs both the Random Forest (RF) and Softmax classifiers. The SVM classifier achieved an overall accuracy of 95.7%, whereas the RF classifier reached 87.4%, and the Softmax classifier achieved 94.8%.

Genemo et al., [2] proposed a model for the segmentation and classification of cattle's lumpy skin disease . The framework incorporates a deep learning-based segmentation method and CNN feature optimization. The proposed method was evaluated on well-known datasets for cattle's lumpy skin disease, and the results indicate promising performance. The best classification result achieved in this work is with the ELM classifier, which attained an accuracy of 0.9012. ELM was found to have the overall best performance on the dataset. However, one constraint of our work is the computational time, which will be addressed in future research. Additionally, in future studies, we aim to enhance our segmentation technique to prevent our deep models from training on irrelevant visual features.

Ujjwal et al., [9] aimed to predict the likelihood of cattle contracting lumpy skin disease in a specific geographic region, either in the present or the future, to facilitate timely preventive actions. We applied multiple machine learning algorithms to a dataset containing 18,603 instances and 16 features, with the target column indicating whether the disease occurred (0) or not (1). Among all the algorithms tested, Random Forest achieved the highest accuracy at 97.7%, outperforming other methods in predicting the occurrence of lumpy skin disease.

Patel et al., [6] mentioned that veterinary doctors typically detect Lumpy Skin Disease through manual observation, but it is not possible to detect the disease in its early stages using these methods.

In such cases, AI-based methods can achieve higher accuracy in disease prediction. In this work, a Random Forest-based machine learning model is used to detect Lumpy Skin Disease, utilizing data from Kaggle for training and validation.

Kukreja et al., [4] developed a robust machine learning model for the accurate identification of a variety of skin diseases in cattle. Precision values, ranging from 88.12% to 97.57%, indicate the model's proficiency in distinguishing between different disease classes. With an overall accuracy of 91.56%, the model demonstrates high reliability, crucial for its real-world application in veterinary contexts. By leveraging a comprehensive dataset and advanced machine learning techniques, the model enables timely interventions and treatments, offering veterinarians a valuable tool for managing cattle skin diseases.

This work highlights the potential for future advancements in disease detection strategies, improving animal health and treatment outcomes. This research paper aims to provide a comprehensive evaluation and performance analysis of a deep learning model, specifically a CNN, for LSD classification in dairy cows. Hyperparameter tuning is a critical step in training deep learning models, as it involves adjusting various parameters to enhance model performance and generalization capabilities.

III. Methodology

The following diagram (Figure 1) provides an overview of the model's workflow, from data collection to predicting the presence of cow lump disease. It outlines the key stages of preprocessing, CNN implementation, hyperparameter tuning, and model evaluation. CNNs are particularly effective for image classification tasks, as they capture spatial hierarchies in images, which are crucial for distinguishing between healthy and diseased cows. 3.1.

IV. Data Preprocessing

Dataset Split: The dataset was split into 80% for training and 20% for testing, following a common practice to ensure robust model evaluation. The training set was used to train the CNN model, while the test set was reserved for final model evaluation. Image Normalization: Prior to training, all images were normalized to ensure consistent pixel intensity distributions across the dataset. This step helps to improve convergence during model training by preventing large gradients.

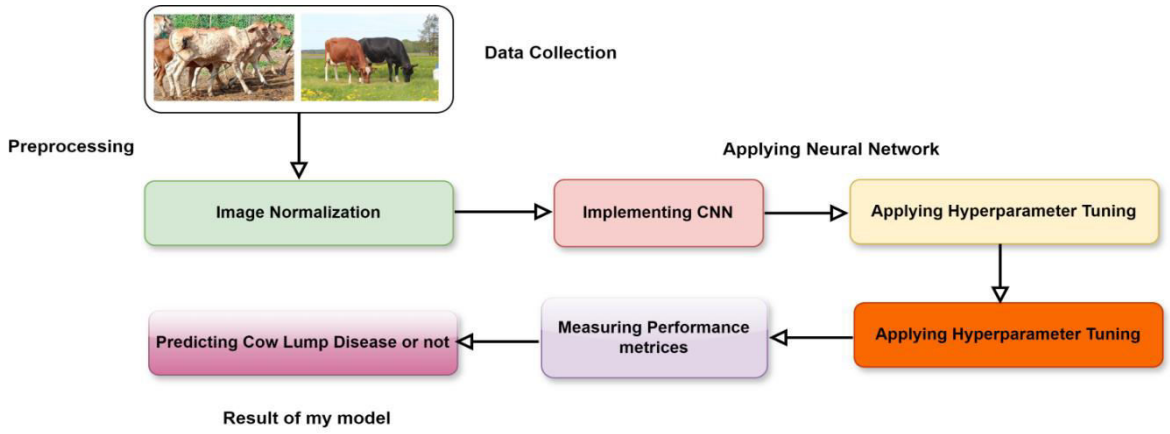


Figure 1: *CNN model*

V. Model Selection and Initialization

For the classification task, a pre-trained CNN model such as VGG16, ResNet50, or InceptionV3 is selected. These models have been trained on large datasets like ImageNet and provide a strong starting point. In this methodology, VGG16 is used. The pre-trained model is loaded without the top fully connected layers, allowing for the addition of custom layers suited to the specific task of LSD classification. The base model is initially frozen, meaning its weights are not updated during the initial training phase.

We opted for a custom CNN rather than a pre-trained model like InceptionV3, ResNet50, or ResNet152V2. This decision was based on the size of the dataset and the need for fine control over the number of filters, layers, and dropout rates. A lighter, customized architecture is often more efficient for smaller datasets where overfitting is a concern.

Pre-trained models like ResNet152V2 are more suitable for large-scale datasets. However, given the nature of the cow disease classification dataset, using a large pre-trained model would have added unnecessary complexity and could lead to overfitting due to the limited data. Instead, we focused on a simpler architecture, using hyperparameter tuning to find the optimal configuration for our specific dataset. 3.3.

VI. Model Architecture

The Convolutional Neural Network (CNN) was designed as follows:

- Input Layer: The model accepts input images of size 256 × 256 × 3 (RGB images).
- Convolutional Layers: The CNN consists of multiple convolutional blocks (1 or 2). Detects local features such as edges and textures with filter sizes ranging from 16 to 64. The activation function used is ReLU.
- MaxPooling Layer: Reduces the spatial dimensions while retaining important features.
- Dropout Layer: Applied to reduce overfitting with a rate tuned between 0 to 0.5.
- Flatten Layer: Flattens the 2D feature maps into a 1D vector to be fed into fully connected layers.
- Dense Layer: A fully connected layer with 64 to 128 units, activated using ReLU, processes the extracted features. Output Layer: A single neuron with a sigmoid activation function is used for binary classification to predict whether the cow has a lump disease or not.

VII. Hyperparameter Tuning

Keras Tuner's Hyperband was used for optimizing key hyperparameters. It tuned the number of convolutional blocks, filter sizes, dropout rates, and units in the dense layer. The search objective was to maximize the validation accuracy. The optimal hyperparameters found during tuning include using 1 convolutional block, 32 filters, 128 dense units, and dropout rates of 0.1 for the convolutional block and 0.4 for the dense layer.

VIII. Result Analysis

To classify Lumpy Skin Disease (LSD) in dairy cows, the first step involves collecting a comprehensive dataset of images. These images should include both affected and unaffected cows. Table 1 shows that 421 affected cows and 515 unaffected cows were collected for experimenting with the proposed model. The Lumpy disease cow is shown in Figure 2a and The Healthy cows are shown in Figure 2b.

Table 1: *Dataset*

Affected Cows	Unaffected Cows	Total Cows
421	515	936

Data preprocessing is crucial for ensuring the quality and consistency of the dataset. All images should be resized to a consistent resolution and format to ensure uniformity. In this model, the images are resized into height 256, and width 256. The batch size parameter is set to 16, which means each batch will contain 16 images. Suppose we have 1,000 images in our dataset. With a batch size of 16, our dataset will be divided into 1000/16 = 62.5 batches. Since we can't have half a batch, there will be 62 full batches of 16 images and one final batch containing the remaining 8 images.

Our next step is to split a dataset into training, validation, and test sets, we allocate a certain percentage of the data to each set. In this model 70% of the total dataset is allocated to the training set, 20% of the total dataset is allocated to the validation set, and 10% of the total dataset is allocated to the test set. We aim to ensure that the data is properly divided into the specified proportions for training, validation, and testing. Each subset can then be used independently for model training, validation, and evaluation.

Figure 2: *Sample Images of Dataset*

Scaling data is an essential step in preprocessing, especially when working with image data in neural networks. The provided code snippet scales the image pixel values from the range [0, 255] to [0, 1], which is a common practice to improve the performance of deep learning models.

To define a model-building function for hyperparameter tuning using Keras Tuner, we can use the build model function we provided. This function allows for the tuning of several hyperparameters, such as the number of filters in each convolutional layer, dropout rates, the number of convolutional blocks, and the number of units in the dense layer. This workflow sets up hyperparameter tuning using Keras Tuner with a customizable model. The buildmodel function defines the model architecture with tunable hyperparameters. The RandomSearch tuner searches for the best hyperparameter configuration based on validation accuracy. After finding the best model, we can evaluate it on the test set to get its final performance metrics.

Our next step is to search for the best hyperparameters for our model using the Hyperband tuner, which balances the exploration of a wide range of hyperparameters and the exploitation of promising configurations. The Keras Tuner's Hyperband tuner is used in our model. It is an effective method for hyperparameter optimization. It intelligently allocates resources to different configurations, allowing for an efficient search process.

Once the best hyperparameters are retrieved, the model is built using these hyperparameters. The model is trained with the optimal hyperparameters for a specified number of epochs (80 in this case). The trained model is evaluated on the test set to determine its accuracy. By executing these

steps, we evaluated the performance of our trained model on unseen test data and visualized the training and validation metrics over epochs, providing insights into the model's learning process and performance. The performance of the model is shown in Figure 3.

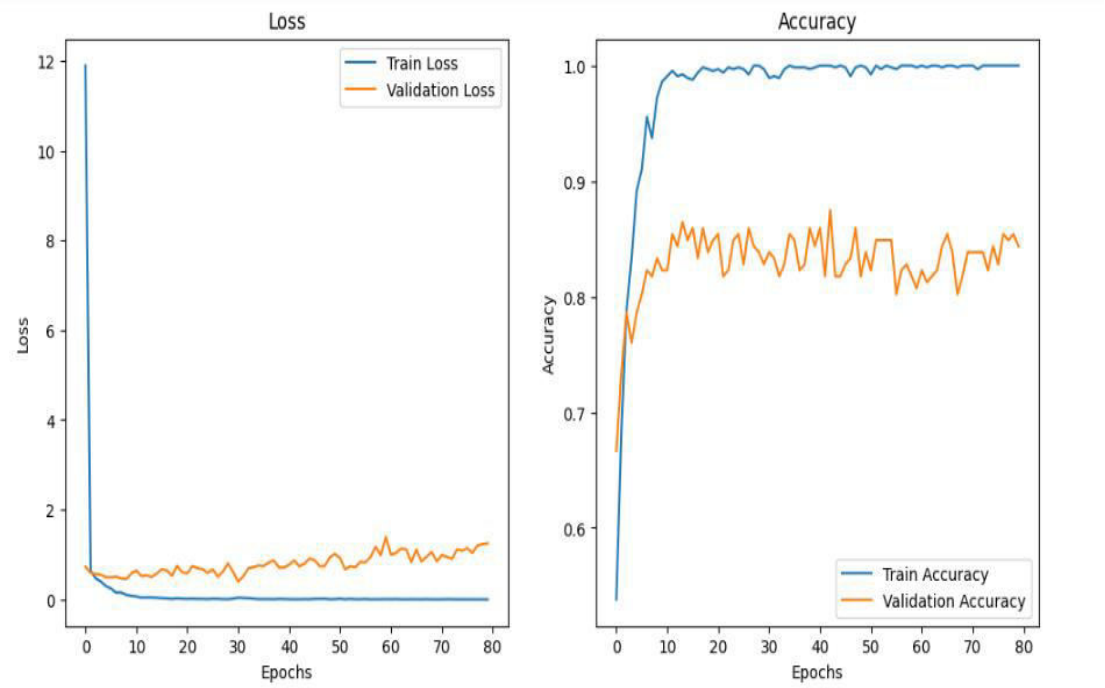


Figure 3: *Test Accuracy and Test Loss*

From Figure 3, it has been seen that the model achieved a test accuracy of approximately 80.68% and a test loss of 1.01. Test Accuracy represents the percentage of correctly classified images in the test dataset and Test Loss value represents the average loss (or error) per sample in our test dataset. Our next step is to evaluate the confusion matrix which is shown in Figure 4.

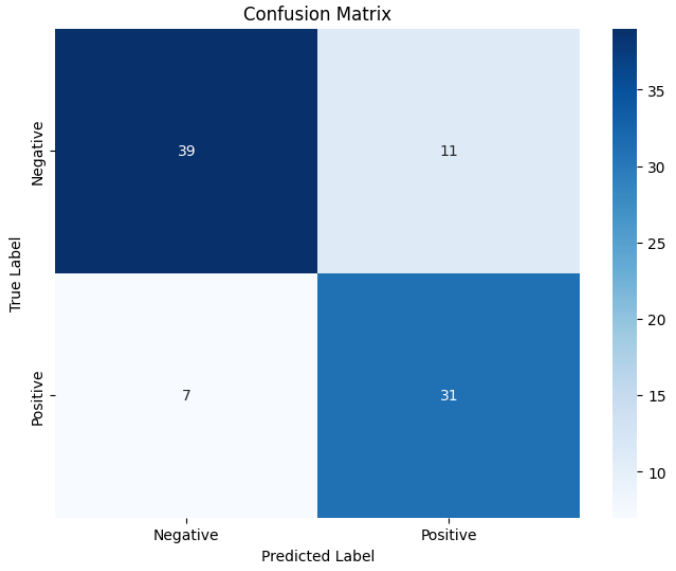


Figure 4: *Confusion matrix*

The confusion matrix provides a detailed breakdown of how well the model performed in classifying instances into positive and negative categories. It shows that out of all instances where the actual class was positive, the model correctly predicted 31 as positive (True Positives). However, there were 7 instances where the model incorrectly predicted them as negative when they were positive (False Negatives). On the other hand, when the actual class was negative, the model correctly predicted 39 instances as negative (True Negatives). There were 11 instances where the model incorrectly predicted them as positive when they were actually negative (False Positives). These metrics, derived from the confusion matrix, offer insights into the model's accuracy, precision, and recall for both positive and negative classes, indicating areas where the model performs well and where improvements may be needed.

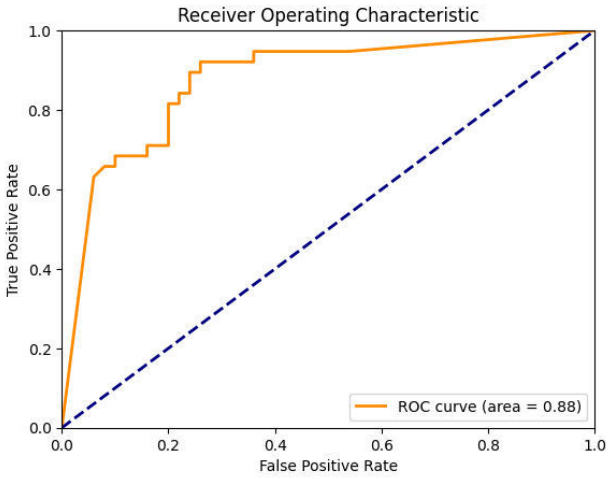


Figure 5: *ROC-AUC*

The final step is to evaluate the ROC-AUC which is shown in Fig 5. A Receiver Operating Characteristic (ROC) curve with an Area Under the Curve (AUC) of 0.88 indicates good performance of our binary classifier. Interpreting AUC of 0.88, AUC of 0.88 suggests that our model has good discriminatory power. It correctly ranks a randomly chosen positive instance higher than a randomly chosen negative instance approximately 88% The higher the AUC, the better the model's ability to distinguish between positive and negative classes. AUC values above 0.5 indicate better-than-random performance, where 0.5 is equivalent to random guessing. In order to justify the models, a comparative analysis is conducted with other existing models as shown in Table 2.

Table 2: *Comparative analysis of other models*

Authors	Accuracy (%)
Workee et al., [10]	80
Ansari et al., [1]	80
The proposed model	80.68.

IX. Discussion

In earlier studies on animal disease detection, various traditional machine learning approaches such as SVM, kNN, and Random Forest have been applied, which rely heavily on feature engineering and may struggle with complex image data. More recent studies have adopted Convolutional Neural Networks (CNNs) due to their ability to automatically learn hierarchical features from images, significantly outperforming manual feature extraction methods.

In addition, Traditional machine learning models rely on hand-crafted features, which can be inadequate for image data with complex patterns. These models are often sensitive to noise and irrelevant features. Some studies apply pre- trained CNN models, which may not be fully optimized for specific tasks like cow disease detection unless fine-tuned appropriately. Although pre-trained models are useful, our approach aims to develop a model optimized specifically for this dataset using custom architecture and hyperparameter tuning, providing a more focused solution.

Unlike traditional methods, CNNs do not require manual feature extraction, which reduces biases and enhances accuracy. CNN architectures can be tailored to various datasets by adjusting layers, filters, and dropout rates, making them more adaptable than static traditional models.

X. Conclusion

This methodology outlines a comprehensive approach to classifying Lumpy Skin Disease in dairy cows using fine- tuning of a pre-trained CNN model. Our study demonstrates that the developed deep learning model shows promise in the automated detection of Lumpy Skin Disease in dairy cows. The combination of accurate classification metrics and strong discriminatory power (AUC of 0.88) supports its potential application in veterinary diagnostics, contributing to early disease detection and proactive management strategies in dairy farming.

References

[1] Ansari, A., Singh, A., Singh, M., Kukreja, V., 2024. Enhancing skin disease classification: A hybrid cnn-svm model approach, in: 2024 International Conference on Automation and Computation (AUTOCOM), IEEE. pp. 29–32.

[2] Genemo, M., 2023. Detecting high-risk area for lumpy skin disease in cattle using deep learning feature. Advances in Artificial Intelligence Research 3, 27–35.

[3] Girma, E., 2021. Identify animal lumpy skin disease using image processing and machine learning. Ph.D. thesis. St. Mary's University.

[4] Kukreja, V., Srivastava, P., Garg, A., Hariharan, S., et al., 2024. Synergizing cnn and random forest for accurate cattle disease identification,

in: 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS), IEEE. pp. 1–6.

[5] Madhanakumar, S., et al., 2024. Lumpy skin disease prediction using machine learning, in: 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), IEEE. pp. 887–895.

[6] Patel, S., Thakkar, V., Swain, D., Bhilare, A., 2023. An early lumpy skin disease detection system using machine learning, in: International Conference on Data Science, Computation and Security, Springer. pp. 51–59.

[7] Rai, G., Naveen, Hussain, A., Kumar, A., Ansari, A., Khanduja, N., 2021. A deep learning approach to detect lumpy skin disease in cows, in: Computer Networks, Big Data and IoT: Proceedings of ICCBI 2020, Springer. pp. 369–377.

[8] Saha, D.K., 2024. An extensive investigation of convolutional neural network designs for the diagnosis of lumpy skin disease in dairy cows. Heliyon .

[9] Ujjwal, N., Singh, A., Jain, A.K., Tiwari, R.G., 2022. Exploiting machine learning for lumpy skin disease occurrence detection, in: 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), IEEE. pp. 1–6.

[10] WORKEE, G., 2021. Cattle skin diseases identification model using machine learning approach. Ph.D. thesis

DEVELOPMENT PROSPECTS AND MATHEMATICAL SOLUTION METHODS FOR INTEGRATING BEACON SYSTEMS INTO UAVS

Elshan Hashimov[1], Elkhan Sabziev[2], Samad Muradov[3]

[1]Azerbaijan Technical University, Baku, Azerbaijan Republic
[2]Institute of Control Systems, Baku, Azerbaijan Republic
[3]Military Scientific Research Institute, Baku, Azerbaijan Republic
hasimovel@gmail.com, elkhan.sabziev@gmail.com, semedmuradov@yahoo.com

Abstract

The paper describes radio beacon systems and presents a mathematical solution for locating an unmanned aerial vehicle (UAV) equipped with a direction-finding device. This system provides continuous flight and allows you to determine the exact coordinates of targets regardless of satellite signals. It enables the successful execution of combat missions in adverse weather conditions and when using radio electronic warfare systems based on signals received from radio beacons. Thus, due to the recent development of unmanned aerial vehicles, many countries are improving their radio electronic warfare systems. Because in a real war, destroying cheap UAVs with expensive Air Defensive Systems missiles does not benefit any country financially. Radio electronic warfare systems are being developed rapidly because they are more effective in this respect. Taking this into account, it is possible to increase the resistance against radio electronic warfare systems by developing the issue of integration of the beacon systems into the UAV proposed in the article. Thus, by setting up beacon systems, it is possible to perform a UAV flight during radio electronic warfare application in any conditions.

Keywords: unmanned aerial vehicle (UAV); direction finder; beacon systems; radio beacons; navigation systems; air defense; radio electronic warfare (REW).

I. Introduction

The analysis of the development directions of the forms and methods of modern military operations shows that unmanned aerial vehicles are now viewed as highly effective tools capable of solving a wide range of combat tasks. It is believed that in the near future, unmanned aerial vehicles will play a significant role in determining the location of air defenses, silencing and destroying them, obtaining the exact coordinates of fortified enemy positions, as well as launching missiles and bombs at detected objects. Already, the course and outcome of military operations, the degree of army readiness for combat, and the ability to perform assigned tasks have begun to depend significantly on unmanned aerial vehicles [1,2]. Additionally, unmanned aerial vehicles have become a powerful factor for commanders when deciding on the initiation of combat operations. They are subject to constant development and improvement, necessitating a careful and detailed analysis of all aspects of their application [3,4].

The article discusses the implementation of a stable navigation system in unmanned aerial

vehicles, enabling them to successfully perform tasks in challenging weather conditions, closed spaces, and even during the application of radio-electronic combat systems. This advanced system allows uninterrupted flight and precise target coordinate determination, regardless of satellite signals. To achieve this, unmanned aerial vehicles are equipped with a direction-finding device, which offers a mathematical solution to determine their location based on signals received from ground-based beacons.

II. Understanding how satellite navigation systems work

Currently, beacon systems integrated with UAVs change in a completely different form. They are mainly mounted on UAVs and are used for various purposes. The most common of them is the GPS beacon. They are a small device that initially determines its location based on the signals it receives from the GPS and transmits this information to a receiver [5]. They are mainly used to track the location of UAVs. They can usually be attached to UAVs, aircraft, other vehicles, and even people. They use signals from satellites to transmit location information in real time, allowing the movement of an object or person to be tracked. GPS beacons receive signals from satellites and use them to calculate their location. They transmit this information via radio signals to a radio receiver.

One possible solution to finding UAVs in an emergency could be the use of "emergency locator transmitters" used on manned aircraft. These transmitters operate at a frequency of 406 MHz to signal distress and communicate with the so-called Cospas-Sarsat System. This is an international satellite system for search and rescue operated by 43 countries and organizations. Its mission is to provide accurate, rapid, and reliable distress signals and location information for search and rescue operations to aircraft, ships, and other equipment in distress. Currently, more than 1.5 million Cospas-Sarsat emergency beacons are in operation. The price of 406 MHz radio beacons varies from $500 to $1,500, depending on their technical characteristics. Both the size and weight depend on the model of the radio beacon. However, in general, optimized hazardous situations for a manned aircraft, the transmitters are approximately 20 cm x 10 cm x 10 cm and weigh 1 kg [6]. The Cospas-Sarsat System can be used for important technological and life-saving tasks. It should not be made to unnecessary demands and its resources should not be wasted. When using a radio beacon, some characteristics of the UAV, such as weight, size and flight range, should be taken into account. In addition, the designed beacon should be as cheap as possible.

The most widely used field of lighthouses until modern times has been maritime. From ancient times to the present, they have been widely used for ship navigation and tracking systems to prevent collisions with coastal or other ships. They are designed to provide identification and position information to both ships and coastal stations. Currently, these systems are the most important means of navigational safety for every sailor after radar. This is a digital position information system operating in a very high frequency range and designed for the marine environment. Its purpose is to identify ships, assist in target tracking, search and rescue operations, simplify information exchange, and obtain real-time information about the current situation. This system was previously developed as a means for merchant ships to see each other more clearly in simple and complex conditions and to avoid collisions by providing detailed information about the environment to the sailor. Lighthouses can organize such safety rules by continuously transmitting the identification code, position, speed, and course of ships to other ships along with relevant information. The transmission range of beacon signals can vary from 20 nautical miles (37 km) to 350 nautical miles (648 km), depending on the power of the transmitting and receiving antenna, atmospheric conditions, and other conditions [7].

When installing radio beacons on the ground, the following requirements should be considered:

- Distance: The first consideration is how far the UAV can be from the beacon. If it is required

to send a signal to the UAV over a longer distance, the power of the transmitter must be increased. In this case, the beacon system may be required to have a larger and higher transmission power.

- Environmental (meteorological) factors: If the beacon system will be used in extremely humid, cold or other harsh environments, it must be designed to withstand these conditions. This may require a more robust and weatherproof design.

- Power source: The power source for the beacon system will affect its size and design. If the system is battery-powered, the size of the battery must be considered when designing the system. There are many different types of batteries, each with their own characteristics and advantages. The most common types of batteries for beacon systems are alkaline, lithium-ion and lithium polymer. Of these, alkaline batteries are widespread and relatively inexpensive, but their lifespan is shorter than other types of batteries. Lithium-ion and lithium polymer batteries have a longer lifespan and are often used in high-performance devices, but they can be more expensive. If the battery is rechargeable, the design of the beacon system should also include a charging system. If the battery cannot be recharged, the system should be designed to allow for easy battery replacement. In general, the design of a battery-powered beacon system should depend on the size, capacity, power consumption, and whether the battery is rechargeable or replaceable. Careful consideration of these factors will help ensure that the beacon system provides reliable performance and meets the needs of the intended application.

- Type of UAV: The size of the UAV into which the beacon system is integrated should not affect the size of the beacon system. The UAV can be small or large. The most important consideration here is that the UAV must be equipped with a receiver capable of receiving the signal sent by the beacon. The receiver must be integrated into the UAV's on-board computer system, which can then use the information received from the beacon system to perform navigation, tracking, or other tasks. These beacon systems can be activated for any type of UAV if their transmission and reception systems are properly coded in advance.

The development of beacon systems and their integration into UAVs can be applied for the following purposes:

1. Increasing navigation capabilities in the operational area. If REW is applied in the operational area, pre-installed beacons will send their location to the UAV through stronger radio signals or laser signals, helping it to accurately calculate its coordinates and apply correct navigation. There were many such problems during the Second Karabakh War. Also, during the Russian-Ukrainian war, the REW systems used by Ukraine caused Iranian UAVs to go to the wrong coordinates and miss the target, and the antiaircraft systems could easily detect and destroy them.

2. Safe takeoff and landing in the runway when REW is applied. Since the takeoff and landing of large UAVs are directly related to GPS, major problems arise when REW is applied in the runway. The reason for this is that the INS, which is considered the second type of navigation in this type of UAVs, takes its origin from GPS. If GPS is obstructed, this means that it will not be able to take its origin. For this reason, the system does not allow takeoff. If it takes off, the risk of the UAV crashing is very high because it does not have a correct starting point and is close to the ground. If the operator waits for the GPS signal to arrive, the takeoff will be delayed for hours, which will also delay the operation. On the other hand, if the GPS is not correct during landing, the UAV will not be able to determine its location and the risk of leaving the runway and crashing is very high. To prevent these problems, if we place a beacon in an area close to the runway, the UAV will determine its exact location based on the signals from the beacon in the event of a GPS obstruction. In this way, we will have prevented the accident situation [8].

The main navigation system used in modern UAVs and airplanes is satellite navigation, which includes GPS, GLONASS, GALILEO, BeiDou, and others. Satellites are positioned approximately 20,000 km above the Earth's surface and continuously transmit their location and real-time data to ground receivers. UAVs use a satellite receiver to capture these signals, and the distance to the GPS

satellites is calculated based on the time it takes for the signal to travel.

To determine its position accurately in 3 dimensions (x, y, and z), the GPS receiver on the UAV must receive signals from at least 4 satellites.

Recently, the effectiveness of satellite navigation has been compromised due to the blocking of incoming signals caused by radio-electronic combat systems employed to thwart UAVs used for illicit purposes such as pillaging. When satellite signals are obstructed, determining the UAV's position becomes challenging.

To overcome this problem, a "beacon system" can be applied to UAVs, similar to the systems used for determining the correct position and direction of ships. In the absence of GPS signals, the UAV will automatically rely on signals from beacons to report its position and maintain navigation.

During the landing, if the GPS is not correct, the UAV cannot determine its location, and the risk of an accident after leaving the runway is greatly increased. To avoid these problems, if we place a beacon in an area close to the runway, the UAV will determine its exact location based on the signals from the beacon during GPS interference. Thus, we will have prevented the accident situation [9].

According to Fig. 1, GPS signals are intercepted by radio-electronic warfare means. Since GPS signals are in a single frequency range, it is possible to easily block those frequencies [10]. But we can control the frequencies of the beacon signals as we want, within the secret range.

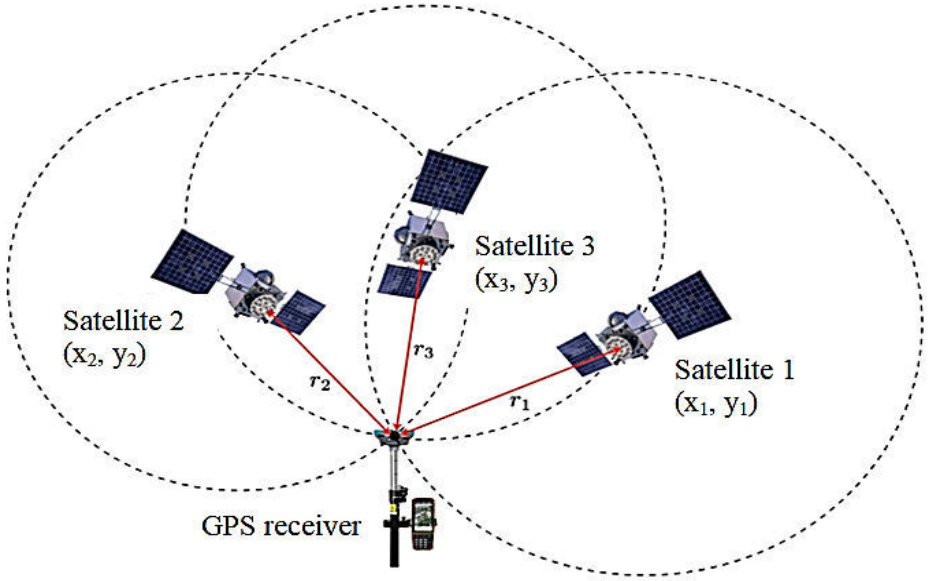


Figure 1: *GPS principle of operation*

Proper installation of beacon systems will significantly increase the stability and effectiveness of UAVs during radio-electronic warfare deployment. For this, a highly accurate and stable technology using different algorithms should be developed. Environmental factors, topography, and other geological factors should be taken into account when installing beacons in the area.

Currently, almost all types of UAVs use GPS navigation. However, beacon systems can be used in non-GPS areas during radio-electronic warfare application.

In the scientific literature, it is impossible to find information about radio beacons specially designed for UAVs.

In order to implement this system, a considerable number of fixed beacons need to be strategically placed across the operational area or at various strategic locations throughout the

country. These beacons will be spaced approximately 20-30 km apart and will continuously emit circular signals on pre-coded frequencies. Similar to GPS, these signals will contain precise coordinates of the beacon and the time of transmission.

The UAV's receiver, installed on board, will be capable of determining its exact coordinates, altitude, and speed by receiving signals from at least four beacons. To ensure comprehensive coverage, it is recommended that radio beacons be positioned in proximity to the operational area in such a way that each point in the area can receive signals from at least three radio beacons via the UAV's bearing.

Once the necessary information is obtained, the UAV will be able to effortlessly and accurately follow the operator's commands for navigation. Furthermore, all the collected data will be transmitted to the Ground Control Station (GCS) through a dedicated communication channel (datalink).

III. Mathematical formalization and solution of the problem

Taking into account that the area over which the UAV flies for the purpose of monitoring is quite limited, the curvature of the Earth's surface in the observed area can be ignored. Therefore, let's enter a rectangular positive Oxy coordinate system with respect to the Earth in order to locate the beacons and the UAV.

It is assumed that the UAV regularly carries out the bearing of its surroundings at a full angle, and at this time it can stably receive the signal of at least $n \geq 3$ radio beacons. Let us denote the full bearing period by T. It is assumed that at a certain time t, the UAV was at the point $\tilde{A}$, whose coordinates relative to the Oxy system are known $(\tilde{x}, \tilde{y})$.

For simplicity, let's number the receiving stations as $k = 1, 2, 3, \dots, n$ clockwise according to the sequence of UAV coverage. It is considered that the coordinates of the k th beacon related to the Oxy system are known and are accordingly (x_k, y_k).

The angle between the direction beams of the UAV bearing towards the 1st, 2nd, 3rd, ..., n-th station at the instant $t + T$ is $\varphi_{1,2}, \varphi_{2,3}, \dots, \varphi_{n*1,n}$.

Let us denote the coordinates of the UAV at the moment $t + T$ related to the Oxy system as $(\tilde{x}^T,\ \tilde{y}^T)$. Then, the issue of determining the location of the UAV equipped with a direction finder device on the basis of radio beacons can be expressed as follows:

It is necessary to find the coordinates $(\tilde{x}^T,\ \tilde{y}^T)$ of the UAV corresponding to the moment of time $t + T$ such that the remaining angle between the beams directed from that point to the k and $(k + 1)$-th beacon is $\varphi_{k,k+1}$. The remaining angle between the beams directed towards the n-th and 1st beacons be $\varphi_{n,1} = \left(2\pi - \sum_{k=1}^{n-2} \varphi_{k,k+1}\right)$.

To solve the problem, let's first determine the set of points that ensure that the angle $\varphi_{k,k+1}$ remains between the rays directed towards the k and $(k + 1)$-th beacon. Essentially, this set is a circle passing through the points $A_k(x_k, y_k)$, $A_{k+1}(x_{k+1}, y_{k+1})$ və $\tilde{A}^T(\tilde{x}^T, \tilde{y}^T)$ consists of (Fig. 2). To write the equation of that circle, find the coordinates of the point $A_k^D(x_k^D, y_k^D)$ located on it, so that it is on the perpendicular raised from the middle of the straight line segment connecting the points A_k, A_{k+1} and $A_k A_k^D A_{k+1}$. Let the angle $A_k A_k^D A_{k+1}$be equal to $\varphi_{k,k+1}$.

Let $A_k^M(x_k^M, y_k^M)$ be the middle point of the straight line segment connecting the points $A_k(x_k, y_k)$ və $A_{k+1}(x_{k+1}, y_{k+1})$. It is obvious that,

$$x_k^M = \frac{x_k + x_{k+1}}{2}, \qquad y_k^M = \frac{y_k + y_{k+1}}{2} \tag{1}$$

Let's write the equation of the straight line passing through the points $A_k^M(x_k^M, y_k^M)$ və $A_k^D(x_k^D, y_k^D)$

as follows:

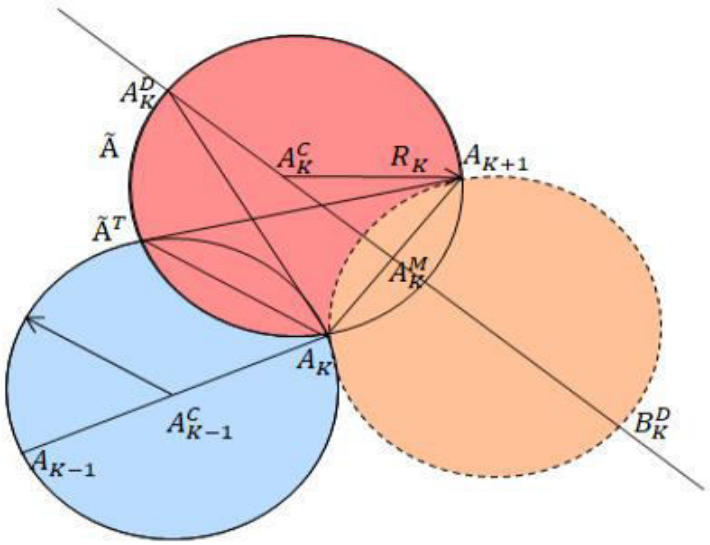


Figure 2: *Determining the remaining angle between the beams*

$$(y_{k+1} - y_k)(y - y_k^M) + (x_{k+1} - x_k)(x - x_k^M) = 0 \tag{2}$$

$A_k A_k^D A_{k+1}$ the condition that the angle is $\varphi_{k,k+1}$ can be written as follows:

$$\operatorname{tg}\frac{\varphi_{k,k+1}}{2} = \sqrt{\frac{(x_k^M - x_k)^2 + (x_k^M - y_k)^2}{(x_k^M - x_k^D)^2 + (x_k^M - y_k^D)^2}}$$

or

$$((x_k^M - x_k^D)^2 + (x_k^M - y_k^D)^2)\operatorname{tg}^2\frac{\varphi_{k,k+1}}{2} = (x_k^M - x_k)^2 + (x_k^M - y_k)^2 \tag{3}$$

Equations (2)-(3) are a system of algebraic equations of the second form with respect to the unknowns x_k^D, y_k^D, and it is clear that it has 2 solutions. Those solutions will correspond to the coordinates of points A_k^D və B_k^D, as shown in Fig. 2. According to the essence of the problem, the closest to the point $\tilde{A}(\tilde{x}.\tilde{y})$ should be chosen from the points A_k^D və B_k^D, in other words,

$$(x_k^D, y_k^D) = \arg\min\{|A_k^D - \tilde{A}|, |B_k^D - \tilde{A}|\} \tag{4}$$

Suppose that the center of the circle passing through the points A_k, A_{k+1} və $\tilde{A}$ is located at the point $A_k^C(x_k^C, y_k^C)$. The coordinates of this point must satisfy the equation (2) and also satisfy the following equation, which expresses the equality of the radii $A_k A_k^C$ və $A_k^D A_k^C$:

$$(x_k^C - x_k^D)^2 + (y_k^C - y_k^D)^2 = (x_k^C - x_k)^2 + (y_k^C - y_k)^2$$

Thus, the calculation of coordinates x_k^C, y_k^C leads to the solution of the following system of linear algebraic equations:

$$\begin{cases} (x_{k+1} - x_k)x_k^C + (y_{k+1} - y_k)y_k^C = (x_{k+1} - x_k)x_k^M + (y_{k+1} - y_k)y_k^M \\ 2(x_k^D - x_k)x_k^C + 2(y_k^D - y_k)y_k^C = (x_k^D)^2 + (y_k^D)^2 - x_k^2 - y_k^2 \end{cases} \tag{5}$$

Using the solution of system (5), we can write the equation of the circle with its center at the

point A_k^C and passing through the points A_k, A_{k+1}:

$$(x - x_k^C)^2 + (y - y_k^C)^2 = R_k^2 \tag{6}$$

here

$$R_k = \sqrt{(x_k - x_k^C)^2 + (y_k - y_k^C)^2}$$

is the radius of the circle.

It is clear that the coordinates of the UAV corresponding to the moment $T + t$ must satisfy equations (6) with a certain accuracy for each k. This means that the new coordinates of the UAV can be calculated as $(\tilde{x}^T . \tilde{y}^T)$ giving a minimum to the following functional:

$$J(\tilde{x}^T . \tilde{y}^T) = \sum_{k=1,2,\dots,n} \{(\tilde{x}^T - x_k^C)^2 + (\tilde{y}^T - y_k^C)^2 - R^2\}^2 \tag{7}$$

(7) Numerical methods can be applied to find the minimum of the functional [5].

IV. Conclusion

In order for UAVs to successfully perform their combat tasks in all operational conditions, it is possible to create stable navigation systems that ensure its uninterrupted flight and determine the exact coordinates of targets, regardless of satellite signals, in closed space and during radioelectron warfare application. Currently, new research is being conducted to detect, identify and neutralize UAVs, anti-GPS signal jamming systems are being developed, and their false spatial deviation is being applied. Considering these methods, it is necessary to create new stable navigation systems for UAVs to successfully perform reconnaissance tasks without satellite signals. The biggest research effort against UAVs is to disable them by affecting their navigation systems and thereby preventing their deployment. The development of radio-beacon systems and their integration into UAVs will greatly expand navigation capabilities.

Thus, a mathematical solution to the issue of determining the position of UAVs provided with a direction finder device using the equations shown in expressions (6) - (7) is provided. Using these equations, it is possible to use radio beacons as an additional navigation tool in an environment where there is no GPS.

References

[1] Hashimov E., Muradov S. Analysis of the application of UAVs in modern wars // -Baku: Military knowledge, – 2023, No. 1. - pp. 7-16

[2] Hashimov , E. ., & Khaligov , G. . (2024). The issue of training of the neural network for drone detection . Advanced Information Systems, 8(3), 53–58. https://doi.org/10.20998/2522-9052.2024.3.06

[3] Hashimov, E.G., Sabziev, E.N., Huseynov, B.S., Huseynov, M.A. Mathematical aspects of determining the motion parameters of a target by UAV // -Kharkov: Advanced Information Systems, - 2023. 7(1), 18–22. https://doi.org/10.20998/2522-9052.2023.1.03

[4] Bayramov A. A. et al. Unmanned Aerial Vehicle Applications for Military GIS Task Solutions. In I. Management Association (Ed.), Research Anthology on Reliability and Safety in Aviation Systems, Spacecraft, and Air Transport (pp. 1092-1115). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-7998-5357-2.ch044

[5] V. Grinyak, A. Shurygin, A. Devyatisilnyi. Accuracy of indoor navigation with Bluetooth beacons: [Electronic resource] / researchgate. – January 29, 2020. URL: https://www.researchgate.net/publication/338074979_Accuracy_of_Indoor_Navigation_with_Bluet

ooth_Beacons?enrichId=rgreq-eb0355122574a2f4267c6eb8ecc76cfa-XXX&enrichSource=Y292ZXJQYWdlOzMzODA3NDk3OTtBUzo4NTI1MzM2ODE2MDI1NjBAMTU4MDI3MTI3NzExMA%3D%3D&el=1_x_2&_esc=publicationCoverPdf

[6] Juana M, Martínez-Heredia, Francisco Colodro, Jose Luis Mora-Jimenez, Alejandro Remujo, Joaquín Soriano, Sergio Esteban. Development of an emergency radio beacon for small unmanned aerial vehicles: [Electronic resource] / researchgate. –april 13, 2018. URL: https://www.researchgate.net/publication/324514029_Development_of_an_Emergency_Radio_Beacon_for_Small_Unmanned_Aerial_Vehicles

[7] AIS (Automatic Identification System) overview: [Electronic resource] / URL: https://shipping.nato.int/nsc/operations/news/2021/ais-automatic-identification-system-overview

[8] Thomas Dautermann, Bernd Korn, Karin Flaig, Maarten Uijt de Haag. GNSS double differences used as beacon landing system for aircraft instrument approach: [Electronic resource] / researchgate. – May 18, 2021. URL: https://www.researchgate.net/publication/352906228_GNSS_Double_Differences_Used_as_Beacon_Landing_System_for_Aircraft_Instrument_Approach

[9] Thomas Dautermann, Bernd Korn, Karin Flaig, Maarten Uijt de Haag. GNSS double differences used as beacon landing system for aircraft instrument approach: [Electronic resource] / researchgate. – May 18, 2021. URL: https://www.researchgate.net/publication/352906228_GNSS_Double_Differences_Used_as_Beacon_Landing_System_for_Aircraft_Instrument_Approach

[10] Reza Nasiri Mahalati. Global Positioning System (GPS): [Electronic resource] –april 03, 2023. URL: https://web.stanford.edu/class/ee259/lectures/ee259_01_gps.pdf

IMPROVING THE PERFORMANCE OF ASSOCIATION RULES HIDING USING HYBRID PARTICLE SWARM OPTIMIZATION ALGORITHM

Eirene Barua[1], Mala Dutta[1], Zafar Jafarov[2]

•

[1]Assam down town University, Assam, India
[2]Azerbaijan Technical University, Baku, Azerbaijan Republic
eirene.adtu@gmail.com, maladuttasid@gmail.com, zafar.cafarov@aztu.edu.az

Abstract

In today's digitized world, data can be taken from many sources like e-marketing sites, social platforms, social networking sites etc. in bulk volumes for usage. Privacy Preserving is a very delicate issue to be looked upon. Hence it becomes necessary to focus on the important privacy preserving parameters. Algorithms for optimization plays an important role in reducing non-sensitive rules in association rule hiding. This paper speaks about a hybrid Particle Swarm Optimization algorithm that requires the properties of all the algorithms which are used for hiding Association Rules and it also highlights the usage of less time.

Keywords: association rules, data privacy-preservation, Particle Swarm Optimization.

I. Introduction

Mining of Association Rules is a very interesting trend which talks about databases on transactions for hiding information which is sensible. There are many types of Rule Hiding Algorithms/Techniques when we talk about sensitive data:

1.Classification Mining Algorithms

2.Decision Region-Based Algorithms

3.Data Perturbation Techniques

Then we have Optimization Algorithms like Particle Swarm Optimization, Cuckoo Optimization Algorithm and Ant Colony Optimization.

This paper describes a hybrid optimization algorithm that gets its characteristics from the mentioned algorithms stated above for hiding association rules, thus producing effective results in less time [1-2].

1.1.Privacy-Preserving Data Mining

Privacy-Preserving Data Mining solely focuses on sensitive information by hiding it based on various rule hiding techniques like association, classification mining, decision region-based and data perturbation techniques.

Here the data is bothered by addition of noise to the set of data used for sensitive data hiding.

There are different models like Sweeney's, Samarati's and Incognito. All the models use multiple secured methods for hiding the data.

1.2. Motivation

At present, there are many applications that are connected with huge amount of data consisting sensible data and information. But when this information is transferred to a third party for extracting the data, there are high chances of the data getting lost. Thus, to solve this problem, data privacy-preservation comes into being.

• Hospital datasets contain a lot of sensible information about patients. Therefore, security should be intact because there are a lot of sensible information related to the patients. Thus, to safeguard this sensible information, there are a lot of privacy-preserving data mining models that can safeguard that sensible information.

• There are cases when robbery takes place in people's houses. Here, the sensible information is the count of the faces appearing in the image, which means however many faces are there in the video has to be saved in order to identify the culprit.

• At present, there exists many companies working on various projects. And that project might contain a lot of information given by the customer. Thus, when the project is sent for processing, there are high chances that the sensible information might be leaked. Thus, for securing such information, privacy-preserving models come into being.

• Also, when talking about bank database, there are a lot of sensible information about the existing customers which should be secured, so that people without proper authorization cannot have access to that information.

1.3. Sweeney's Algorithm:

Sweeney's Algorithm is an algorithm for showing anonymity of Electronic Health Records. Anonymization is accomplished by means of mechanically generalising, substituting, inserting and removing statistics without losing details for research.

1.4. Samarati's Algorithm:

This algorithm scans for the capable k-anonymous explanations by capturing several levels in Domain Generalization Hierarchy. It avails the binary search to get the solution in very less time. Samarati makes the theory that good solutions are the ones where results in a table have minimum generalizations. Thus, her algorithm is considered to look at the concept that determine k-anonymity with minimal suppression. This algorithm fulfils the AGTS model, generalization is applied on column, and suppression is applied on row. MaxSup is the greatest number of tuples that are granted to be suppressed to attain k-anonymity.

1.5. Incognito Algorithm:

Incognito Algorithm produces the set of all possible k-anonymous full domain generalizations of relation T, with an optional tuple suppression threshold. The algorithm consists of iterations of two parts. It begins by examining single-attribute subsets of the quasi-identifier, and afterward repeats scanning k-anonymity with respect to bigger subsets of quasi-identifiers.

II. Literature survey

Using the big itemsets, association rules discover all sets of items with support greater than the minimum support, and then construct the desired rules with confidence greater than the minimal confidence. The lift of a rule is the difference between the actual and predicted support if X and Y were independent. Market basket analysis is a common and widely used application of association rules. Support and confidence are two crucial factors to consider when evaluating association rules. To deal with the sensitivity of association rules, a cyclic technique was adopted (Agrawal et al., 1993; Atallah et al., 1999; Belwal et al., 2013). Similarly, there is confidence reduction (CR), CR2, and generating itemset hiding to improve the hiding process (GIH). For hiding association rules, such type of approaches is being used. (Hahsler et al., 2005; Hong et al., 2011; Kalariya et al., 2015;

Kennedy & Eberhart, 1995; Modi et al., 2010a, 2010b). The work was improved, and algorithm 2b was included to disguise the generating itemset of sensitive rules. (Verykios et al., 2004). An improved algorithm based on Decrease Support and Confidence (DSC) provided better results by hiding predictive association rules. (Wang et al., 2004). Using a genetic algorithm to hide the sensitive rules, privacy-preserving association rule mining over dispersed datasets is possible. (Kesava Murthy & Khan, 2013). Similarly, for the first time, genetic algorithms were used to hide itemsets, and it included a compact pre-large GA-based technique (Goldberg, 1989, 2002) to remove transactions and (Lin et al., 2014) and modified algorithms are introduced to give and insert transactions that are recent.

The most advanced methods are a basic genetic algorithm for transaction deletion and a pre-large genetic algorithm for transaction deletion. For transaction deletion of a fewer variables, and to determine the number of transactions that should be removed in order to reduce negative effects, a modified particle swarm optimization-based algorithm was used.

III. Groundwork for hiding Association rules

The link between the original database D and the sanitised database D is depicted in Figure 1a. The completed cleaned database D is shown in Figure 1h. The coloured area of Figure 1b represents the original database D's frequent itemsets FIs. Figure 1c depicts the sensitive items in D, i.e., itemsets with a Support count of less than Min Support threshold. The non-sensitive data elements are depicted in Figure 1d, where S is in D. Figure 1e–g depicts the three side effects of the sanitization process. The coloured region in Figure 1e represents the hidden failures of sensitive data items that happened. This displays the itemsets that should have been concealed in the sanitised database but were not by the sanitization method. Figure 1f depicts the Sensitive Items that aren't meant to be hidden and must have existed in the sanitization database, but the sanitization algorithm was unable to include all Sis values. The region formed as a result of the production of extra rules that were not available in the original database is seen in Figure 1g.

IV. Provocation for hiding Association rules

Hiding Association Rules poses a number of difficulties, the majority of which are NP-hard problems. This section goes over some of the more common issues [4].

- Failure to Hide: Some sensitive rules were discovered in a sanitised data base that was supposed to be kept concealed. In GA and PSO techniques, this is accomplished through increased computational complexity.
- Rules lost: Some of the non-sensitive frequent item sets are concealed and will not be present in the sanitised database, because the sanitization method failed.
- Artificial rules: In a sanitised database, the effect of sanitization can result in the development of some ghost rules or artificial rules.
- Differences in the database: The ratio between the sanitised database D and the original database must be kept as low as possible, that is, the original database's transactions must be deleted or changed as little as possible. A database similarity ratio of higher than 90 is required.
- Estimated difficulty: The sanitization procedure must be able to give the output database in the shortest amount of time possible, as opposed to other techniques. More efficient the algorithm is, the less complex it is.
- Precision: When data accuracy deteriorates, the apprehension collected from the sanitised database becomes meaningless. The precision is inversely related to the database differences.

V. Defining the problem

In a given a database D, which contains a collection of transactions T1, T2..., Tn, each transaction consists of a set of items I1, I2..., In. Finding the set of rules is the main goal of the association rule mining method, which includes both a priori and FP-Growth algorithms. Based on MST and MCT, these rules are categorised into sensitive and non-sensitive rules. Hiding association rules is mainly associated with hiding association rules which are sensitive. It uses the set of data as input and outputs the set of data which are unrealized. The association rules which are sensitive are hidden, and ghost and lost rules are minimized, thanks to the un-realization dataset used as input to the association rules mining method.

VI. Proposed Hybrid Particle Swarm Optimization Algorithm

This work suggests a hybrid particle swarm optimization approach for proper hiding of association rule that combines the traits of the original particle swarm optimization and other listed optimization methods. Modified C4ARH, Perturbed Dataset, Modified PSO(), ACO() and ARM() are some of the components in this algorithm [3]. The objective function values, as well as their differences and side effect data, are computed by Modified C4ARH. The entire set of data is handled by the Pertubed Dataset() algorithm, which results in the standard set of data S and perturbed set of data P. The values for pbest and gbest are found using the Modified PSO() algorithm. The association rules which are sensitive are defined by gbest, whereas the association rules which are non-sensitive are being defined by pbest.

The association rules were generated using the ARM() technique. Ant Colony Optimization Algorithm is a presumed technique which is used to solve the problems of estimation which can be minimized to finding accurate paths through graphs [4-5].

VII. Hybrid Algorithm

```
Input Original Dataset D, MST, and MCT;
Output A Sanitized Dataset R; begin
do Dataset D
Find item count; call ARM();
end;
begin ARM(D)
Generate association rules
count = number of association rules
Return O, S;
end
begin
for (O, S) apply PSO
for each rule R;
calculate best() for all O,S;
  COA4ARH()
  end
begin
best(R,MST,MCT)
if(R > MST && R > MCT)
```

```
gbest = R;
gbest[]++;
return gbest;
end
begin
COA4ARH() for all O,S
remove gbest;
compute objective function;
R = compare the objective function values of O with S;
R is the dataset with sensitivity;
end
begin
  ACO()
begin pheromone trails;
while (termination condition not satisfied)
do
construct candidate conformations;
perform local search;
update pheromone values;
end
end
```

VIII. Experimental Results

All the experiments are performed on Windows operating system with i3 processor, 16 GB RAM. Corresponding Matlab programs along with necessary datasets are used to test the performance of the association rule hiding algorithm.

There are a variety of well-known techniques for optimization accessible, however Hybrid Particle Swarm Optimization (PSO) and Hybrid Ant Colony Optimization (ACO) are the most used ones. Although we employ a variety of strategies to improve individual advantages, these two optimization approaches outperform them all due to its total effectiveness. In case of best cost function, these two approaches outperform Cuckoo and Generic Optimization techniques. It does not believe in showing a significant change, but it does show a significant decrease in the percentage of optimization that can be viewed using the graphs.

However, in the case of side effect factors, there is a significant difference, which aids in the elimination of side effects to a higher extent than other optimization strategies. This is a big benefit of this technique, and there is also generation of lost rule, which deals with missing values and aids in the synthesis of lost rules; it has a remarkable 98 percent accuracy rate that no other optimization technique can match. This is an important component in deciding whether or not to use Hybrid optimization.

Figure 1 and Figure 2 shows PSO and ACO in terms of cost value.

Figure 3 and 4 shows modified PSO and modified ACO in terms of best cost function.

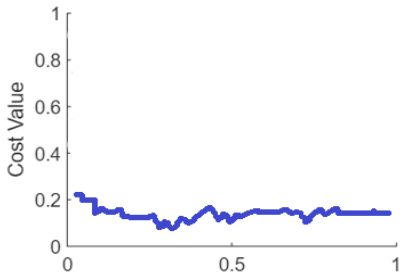


Figure 1: *Experimental result of PSO algorithm*

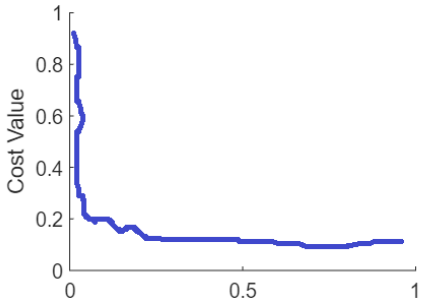


Figure 2: *Experimental result of ACO algorithm*

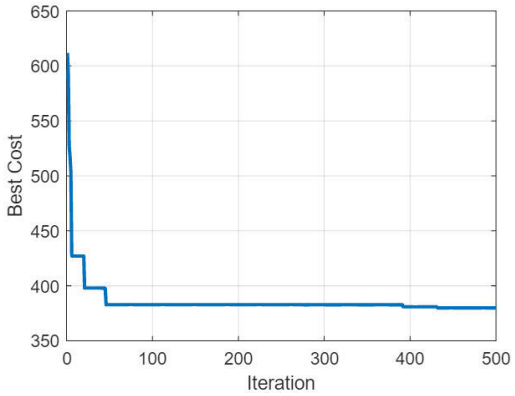


Figure 3: *Experimental result of Modified PSO algorithm*

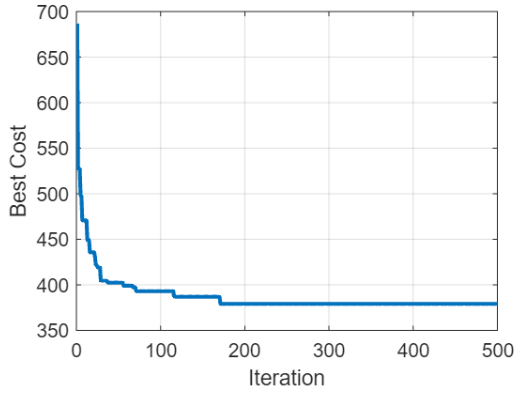


Figure 4: *Experimental result of Modified ACO algorithm*

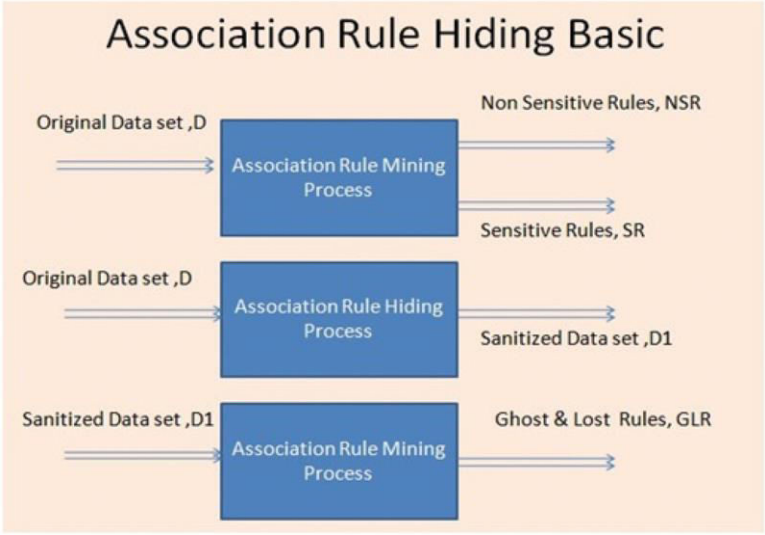


Figure 5: *Framework for Association Rule Hiding*

X. Conclusion

In this article, a hybrid optimization algorithm has been proposed to improve the performance of the mentioned parameters like cost value and best cost function. Experimentally, the result showed better performance than all the other traditional algorithms.

References

[1] Jerry Chun-Wei Lin a,n , Qiankun Liu a , Philippe Fournier-Viger b , Tzung-Pei Hong c , Miroslav Voznak d , Justin Zhan (2021) A sanitization approach for hiding sensitive itemsets based on particle swarm optimization. *Engineering Applications of Artificial Intelligence 53* (2016) 1–18 www.elsevier.com/locate/engappai

[2] Lin, C.-W., Hong, T.-P., & Hsu, H.-C. (2014). Reducing side effects of hiding sensitive item sets in privacy preserving data mining. *The Scientific World Journal*, *2014*, 1–12. https://doi.org/10.1155/2014/235837

[3] Lin, J. C. W., Liu, Q., Fournier-Viger, P., Hong, T. P., Voznak, M., & Zhan, J. (2016). A sanitization approach sfor hiding sensitive itemsets based on particle swarm optimization. *Engineering Applications of Artificial Intelligence*, *53*, 1–18. https://doi.org/10.1016/j. engappai.2016.03.007

[4] Satyanarayana Murthy, T., Gopalan, N. P., & Venkateswarlu, Y. (2018). An efficient method for hiding association rules with additional parameter metrics. *International Electronic Journal of Pure and Applied Mathematics*, *118*(7), 285–290.

[5] Satyanarayana Murthy, T., Gopalan, N. P. (2018). A novel algorithm for association rule hiding. *International Journal of Information Engineering and Electronic Business*, *10*(3), 45–50. https://doi.org/10.5815/ijieeb.2018.03.06.

RESEARCH SPECTRAL EFFICIENCY OF NEW MODULATION FORMATS IN FIBER-OPTIC NETWORKS

Bayram Ibrahimov[1], Asmar Nabieva[1] , Aygun Hamidova[1]

•

[1]Azerbaijan Technical University, Azerbaijan, Baku
bayram.ibrahimov@aztu.edu.az, esmer.nebiyeva@aztu.edu.az, aygun.gamidova69@mail.ru

Abstract

The rapid growth of the use optical technologies requires the development of methods and means to improve the spectral efficiency and noise immunity of fiber-optic networks when using wavelength multiplexing. On the basis of the research a new approach to the construction of the method of calculation spectral efficiency indices of new modulation formats in fiber-optic networks has been developed. The proposed method calculation of indicators takes into account the efficiency indicators fiber-optic transmission systems, the algorithm demodulator synthesis and effective methods modulation formats such as M-ary Quadrature Amplitude Modulation, Differential Phase Shift Keying and M-ary pulse position modulation. On the basis of the calculation method, important analytical expressions evaluating the characteristics of line capacity and noise immunity of optical signal reception are obtained.

Keywords: Spectral efficiency, signal to-noise ratio, fiber optic network, bit rate, line capacity, noise immunity.

I. Introduction

Nowadays, in conditions constant growth of the range infocommunication services and applications provided by optical telecommunication systems based on wavelength multiplexing technology, increasing volume transmitted packet streams, rapid growth requirements to noise immunity of multiservice traffic message reception and the issues effective frequency resource allocation are most acute [1, 2].

The advanced technologies future generation wave multiplexing and 2030 network include both WDM (Wavelength Division Multiplexing), CWDM (Coarse WDM), DWDM (Dense WDM), IoT (Internet of Think), and HDWDM (High DWDM), AI (Artificial Intelligence), 5G-NR-U (New Radio-Unlicen), and 6G (Generation) [2, 3, 4, 5].

In optical telecommunication systems considering the above mentioned modern spectral technologies, one of the important indicators is the line capacity, network performance, frequency efficiency and noise immunity reception optical signals in the provision of infocommunication services and applications.

On the basis of research it was established [6, 7, 8] that one of the important methods of increasing spectral efficiency and noise immunity in fiber-optic networks when using new modulation formats are methods of mathematical modeling. Among them, the spectral efficiency of

modulation formats and optimal methods optical signal reception and efficient methods of signal-code design [9, 10] using modern WDM and DWDM systems with free-length , .

Taking into account the above mentioned, the task of research and the problem of providing the required spectral efficiency and noise immunity of the paths of fiber-optic transmission systems, as one of their most important characteristics, is quite relevant.

In [7, 9, 11] some problems line capacity, bandwidth, and spectral efficiency of telecommunication systems related to the optimization of the demodulator of a noise-tolerant receiver are analyzed. In [5, 12], noise immunity, modulation methods, spectrum expansion, and issues noise-tolerant coding in optical communication systems, which are only defined signal-to-signal-noise ratios.

An optimization problem arises to investigate the spectral efficiency and noise immunity performance fiber-optic networks built according to the latest wavelength multiplexing technology.

The purpose of this work is to develop a new approach to create methods for calculating the characteristics of spectral efficiency and noise immunity of signal reception in fiber-optic networks using efficient modulation formats.

II. General statement of the research problem

Conducted a study of the characteristics fiber-optic networks based on WDM and DWDM technology show [7, 9, 12] that the spectral efficiency and noise immunity of reception characterizes its ability to provide a given quantity and quality of transmitted messages with the least amount time, bandwidth , signal power and guaranteed quality of service.

Considering the constituent components of the vector communication quality, both spectral efficiency, line capacity, and bit error probability reception in terms of the performance of fiber-optic networks based on DWDM technology, which functional dependence is described as follows:

$$S_{EF}(E_b, \lambda_i) = W[SNR(E_b, N_0), V_b(b_i, \lambda_i), E(\lambda_i), \eta_{SE}(\Delta F_k, \lambda_i)] \quad (1)$$

where $SNR(E_b, N_0)-$ function, taking into account the signal-to-noise ratio (Signal to-Noise to Rate) at the demodulator input with consideration of bit signal energy E_b and interference power spectral density N_0, which characterize the complex indices fiber-optic networks based on DWDM system; $\mathrm{V_b}(b_i, \lambda_i)-$ bit rate message with wavelength λ_i and with binary code element b_i and these are binary symbols from the Galois field $F_2(b_i) = GF(2) = \{0,1\} \rightarrow b_i = \{0,1\}; \eta_{SE}(\Delta F_k, \lambda_i)-$ spectral efficiency fiber-optic transmission systems, which characterizes the channel utilization efficiency over the frequency band ΔF_k with wavelength λ_i, (bps/Hz) and is equal to [2]:

$$S_{SE}(\Delta F_k, \lambda_i) = [V_b(\lambda_i)] / \Delta F_k] \geq 1 \quad (2)$$

where $E(\lambda_i)-$ line capacity fiber-optic transmission systems and can be increased by increasing both the bit rate transmission and the number spectral channels N_k, is expressed as follows:

$$E_{DWDM}(\lambda_i) = V_b(b_i, \lambda_i) \cdot N_k \quad (3)$$

where N_k – is the number of spectral channels and is equal to the ratio of the light amplification bandwidth $\Delta\lambda_y$ to the inter-channel spacing width $\Delta\lambda_k$ and is expressed as $N_k = (\Delta\lambda_y / \Delta\lambda_k) > 40,...,96$ channels.

Expression (1), (2) and (3) define the essences of the proposed new approach and characterize the spectral efficiency and fidelity performance of transmission systems.

It is worth noting that the task managing heterogeneous spectral efficiency resources when using modulation formats, researching and evaluating the fidelity of message transmission using the performance criterion fiber optic transmission systems in providing a wide range of multimedia services has not yet been fully explored [5, 9, 11, 12].

Modulation formats for $\mathrm{V_b}(b_i, \lambda_i) \geq 40\,Gbps$ are divided into amplitude-based M-QAM and phase-based DPSK (Differential Phase Shift Keying), which reduces the width of the optical spectrum occupied by the signal, $\Delta\nu_S$ (for binary formats, $\Delta\nu_S = 2\mathrm{V_b}(b_i, \lambda_i)$, increases the spectral efficiency of the modulation $\eta_{SE}(\Delta F_S, \lambda_i) \to \max$, and thus enables high bandwidth fiber optic transmission systems.

Therefore, there is a need to study the DWDM system and create a new approach to build a method for calculating the spectral efficiency and noise immunity using new modulation formats, allowing to optimize the characteristics of fiber-optic networks based on DWDM technology.

III. Structural diagram of the investigated DWDM system link

The surface roughness formed during the cutting of HARDOX-500 steel workpieces varies based on numerous technological, kinematic, structural, and processing environment factors. Consequently, one of the most important technological factors affecting the roughness formed on the surface during abrasive waterjet cutting is the pressure of the water-abrasive mixture applied to the cutting zone. An increase in the pressure of the water-abrasive mixture enhances the cutting capabilities of the water-abrasive, which plays the role of the cutting tool in the steel cutting process, thereby intensifying the cutting of the workpiece. Our research has determined that the study of the surface roughness dimensions in hydroabrasive machining varies widely depending on the processing conditions and regime parameters of the process. Therefore, examining the regularities of roughness changes during the machining of the selected material is one of the important tasks for identifying the advantages of the process. In this case, the roughness formed on the processed surface, depending on the optimal cutting process, has been determined through experimental research to remain within the required limits. In the experiments, a workpiece thickness of 15 mm was taken, with a longitudinal feed rate of S_{long}=26,7 mm/min, an abrasive particle size of 80 μm, and an abrasive consumption of Q=125g/l, while the surface roughness obtained during abrasive waterjet cutting [3] is shown in Figure 1, and the dependence of the obtained experimental and theoretical values on the influence of the water-abrasive jet is presented in Tables 2 and 3.

As mentioned above - constantly emerging types infocommunication services and their new user applications create an increasing load on the backbone optical transport network with increased efficiency. Transportation high-speed traffic flow requires the latest data transmission technology, which, on the one hand, has sufficient spectral efficiency and performance, on the other hand, provides the operator with the ability to scale the network without changing the infrastructure and quality of communication. These requirements are met by the technology of spectral multiplexing WDM and DWDM, which is the main technology for building backbone fiber-optic transmission systems [5, 9].

WDM wave multiplexing is a physical layer technology that is the transmission of multiple optical signals in a single optical fiber at different wavelengths. The first WDM systems were dual channel with transmission at wavelengths $\lambda_i = (1.31,...,1.55)\,\mu m$.

Somewhat later multi-channel solutions appeared: CWDM and DWDM , where the names refer to the density of information optical channels in the optical band. DWDM is a technology of dense spectral multiplexing. Optical channels are located in the range from 1.53 to 1.565 μm with a step of 0.4 nm, 50 GHz or 0.8 nm, 100 GHz [7, 9].

Pic.1 shows the principle of operation of the structural scheme of wavelength multiplexing with optical amplifiers - Erbium Doped Fiber Amplifier (EDFA).

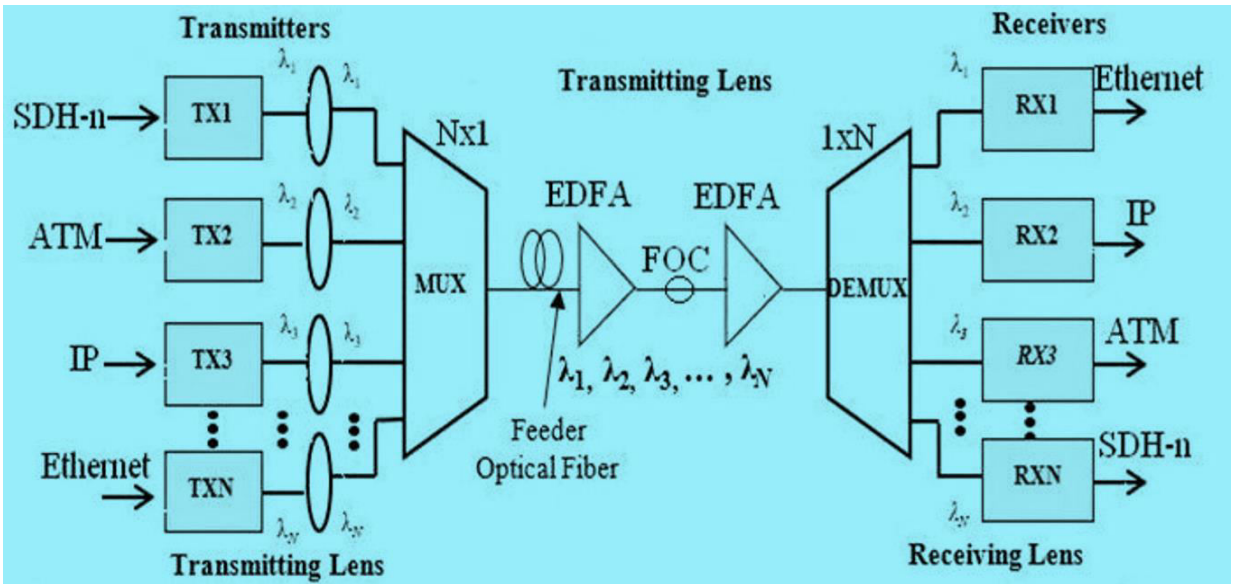


Figure 1: *DWDM system - FOC transmission optical signals on many wavelengths with EDFA optical amplifiers*

From the schematic, it can be seen that figure1 consists of the following important blocks: optical transmitter (TXN), Erbium Doped Fiber Amplifier (EDFA), Fiber Optical Cable (FOC, ITU-T, G.652, SSMF (Standard Single Mode Fiber)) and receiver (RXN). In addition, the main components of a DWDM system are:

- Transponders that generate signals at different wavelengths;
- Multiplexers (MUX), which combine signals from different fibers at different wavelengths on a single fiber, and demultiplexers (DEMUX), which separate multiple signals at different wavelengths from a single fiber over different fibers;
- Amplifiers that amplify a multichannel signal during its transmission over optical fiber, EDFA.

In this scheme, the used erbium-doped fiber amplifier (EDFA) allows uniform amplification of information channels at different wavelengths just in that spectral range of optical fiber where the signal attenuation is minimal, i.e. $\alpha_z(\lambda_i) = A_d \cdot (1/L_y) \to \min$ at the range $C(\lambda_i) = (1.530,...,1.565)\mu m$, (where the $A_d -$ allowable losses in the FOC).

In DWDM system for transmission optical signals primarily use the spectral bands $C(\lambda_i) = (1.530,...,1.565)\mu m$, $S(\lambda_i) = (1.460,...,1.530)\mu m$ and $L(\lambda_i) = (1.565,...,1.625)\mu m$. For the technology of spectral compaction wavelengths in the working bands are used and can be described by a one-dimensional matrix of the form

$$E[L_y(m,k)] = \left|\lambda_1, \lambda_2, ..., \lambda_{i.\max}\right| \;,\; m \neq k \,,\; i = \overline{1,n} \tag{4}$$

where $L_{ó}(\grave{o},k)-$ is the length of the section between network elements in the transmission system; m and $k-$ respectively, indices for the numbers of the initial and final network elements-for the operating bands $C(\lambda_i)$, $S(\lambda_i)$ and $L(\lambda_i)$ when using a frequency grid 50 GHz.

The carrier spacing in DWDM systems for optical networks can be $\Delta F_{DWDM}=(25,...,200)GHz$. However, in modern fiber-optic networks the most commonly used channel grid with a 50 GHz step.

It should be noted that the technology WDM and DWDM transmission with a wide possibility creates a basis for the organization of flexible high-speed intelligent fiber-optic networks, providing transparent transmission ever-growing traffic, including delay- sensitive. Directly related to the performance DWDM systems is the efficiency and noise immunity of fiber optic networks.

IV. Analysis and evaluation of spectral efficiency of multilevel modulation formats

It is known [7, 12], that the spectral efficiency of fiber-optic transmission systems based on DWDM technology is the ratio of the bit rate information transmission ($bps=Bod$) to the bandwidth $\Delta\nu_k$ and is measured in the index (bps/Hz). For M-QAM ($M-$number of signals), modulation formats, spectral efficiency is expressed as:

$$S_{SE}(\Delta F_k,\lambda_i)=[V_b(\lambda_i)]/\Delta\nu_k]\geq 1\ ,\ V_b(\lambda_i)\geq\Delta\nu_k \tag{5}$$

Formula (5) defines the spectral efficiency of multilevel modulation formats, which characterizes how efficiently bandwidth is used in an arbitrary DWDM system.

For M-QAM modulation formats, the spectral efficiency is expressed as:

$$S_{SE}(\Delta F_k,\lambda_i)=\frac{R_k/T_b(\lambda_i)}{\Delta F_k}\cdot\ell og_2 m\geq 1,\quad m=2^M \tag{6}$$

where R_k - is the rate of the interference-free code, $R_k=(3/4)<1$, $T_b(\lambda_i)-$ is the bit cycle in the system and is equal to $T_b(\lambda_i)=1/V_b(b_i,\lambda_i)=\frac{T_S}{M}\cdot 2^M$, M - is the coding level.

Using M-QAM modulation formats ($M\geq 4$) and formulas (5) and (6) at $V_b(\lambda_i)=40Gbps$ and $\Delta F_k=100GHz$, then $S_{SE}(\lambda_i)=0.40bps/\text{Hz}\geq \text{S}_{\text{SE.all.}}(\lambda_i)$ and $\lambda_i=1.55\mu m$, where $\text{S}_{\text{SE..all.}}(\lambda_i)-$ is the acceptable value of spectral efficiency [2, 9].

Now, based on (5) and (6), we can express the DWDM system capacity through the spectral efficiency of modulation formats:

$$E_{DWDM}(\lambda_i)=\text{S}_{\text{SE}}(\Delta F_k,\lambda_i)\cdot\Delta\nu_{ó} \tag{7}$$

In this representation, the capacity optical links when using a DWDM system does not depend explicitly on the bit rate transmission, but is directly proportional to $\Delta\lambda_y$.

The analysis shows [5, 7, 9, 11] that to increase the channel transmission rate and, consequently, the spectral efficiency by a factor of two can be achieved by polarization multiplexing. In this case,

two independent message streams are transmitted at the same wavelength in orthogonal polarization states.

To increase the spectral efficiency and line capacity, we consider the most widely used M-ary pulse position modulation (M-PPM), which reduces the optical signal bandwidth. Taking into account the M-PPM parameters, the spectral efficiency of new modulation formats is expressed as follows [12]

$$S_{SE}(M,\lambda_i) = (1/\grave{I}\ ) \cdot \ell \mathrm{og}_2 \binom{M}{m} \tag{8}$$

Based on (8), when using M-PPM modulation formats, the required bandwidth optical communication networks is determined as follows:

$$C(M,\lambda_i) = M \cdot V_b(b_i,\lambda_i) / \ell og_2 \binom{M}{m} \tag{9}$$

Considering (8) and (9), the line capacities of the WDM and DWDM system can be determined as follows:

$$E_{DWDM}(\lambda_i) = (\Delta \nu_y / M) \cdot \ell og_2 \binom{M}{m} \tag{10}$$

On the basis expression (8), (9) and (10), the package Communications Toolbox in Matlab environment (R 2019b) [10] was used and numerical values were obtained by means of which the graphical dependence was plotted, $S_{SE}(M,\lambda_i) = \mathrm{W[OSNR\,(h),\,M,\,V_b\,(b_i,\lambda_i)]}$.

Figure.2. Shows the graphical dependence of spectral efficiency on the optical signal-to-noise ratio.

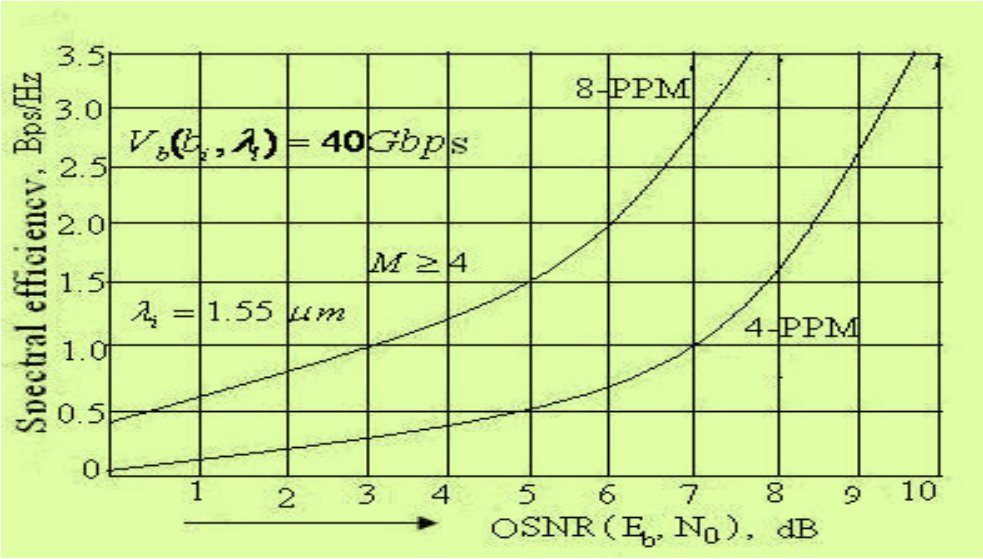


Figure 2: *Graphical dependence spectral efficiency on optical signal to-noise ratio when using the new 4-PPM and 8-PPM modulation formats*

Analysis graphical dependencies $S_{SE}(M,\lambda_i) = \mathrm{W[SNR\,(h),\,M,\,V_b\,(b_i,\lambda_i)]}.$, depicted in figure 2, clearly show the improvement spectral efficiency with increasing signal to-noise ratio when using PPM signals 4-PPM and 8-PPM types for a given bit rate $\mathrm{V_b}(b_i,\lambda_i) = 40\,Gbps$, $\lambda_i = 1.55\,\mu m$ and $M \geq 4$.

V. Investigations of noise immunity optical signals reception in DWDM system

Nowadays, optical telecommunication systems have stringent requirements on bit rates transmission, signal processing methods and noise immunity optical signal reception [3, 11]. To solve this problem, efficient modulation formats M-ary Quadrature Amplitude Modulation (QAM) type were chosen.

For modulation formats such as M-QAM - one of the main criteria for assessing the quality of communication, is the working characteristics noise immunity reception optical signals in DWDM system, showing the dependence of the probability erroneous reception per symbol and the probability of bit errors from the required optical signal to noise ratio (OSNR, Optical Signal to - Noise Ratio):

$$P_{BER}(h) = W[SNR(E_b, \lambda_i, N_0)],\ P_{SER}(h) = W[SNR(E_b, \lambda_i, N_0)] \tag{11}$$

Suppose that in a DWDM system QAM signal is orthogonal in nature. For an orthogonal system $OSNR(E_b, \lambda_i, N_0)$ is expressed as follows:

$$SNR(E_b, \lambda_i, N_0) = \mathrm{SNR}\,(\mathrm{h}, \lambda_{\mathrm{i}}) = \sqrt{P_S T_S / N_0} \tag{12}$$

where $N_0 -$ is the power spectral density of the optical signal under the influence of additive white Gaussian noise (AWGN, Additive White Gaussian Noise).

Taking into account expressions (12) we determine the probability bit errors optical reception signals in DWDM system. Given the noise components of the output signal and the noise variance AWGN or error variance σ_A^2, it is possible to estimate the value of $OSNR(E_b, N_0)$ when receiving M-QAM signals ($M -$ number of QAM modulation levels), which with σ_A^2 is expressed as follows [7, 8]:

$$OSNR(E_b,\ \sigma_A^2) = O\mathrm{SNR}\,(\sigma_{\mathrm{A}}^2) = (E_b^2 / 2\sigma_A^2) \cdot \ell og_2 M \tag{13}$$

It follows from (13) that $OSNR(E_b, \lambda_i, N_0)$ is determined only by the signal energy E_b and error variance σ_A^2, which does not depend on the shape of the optical signal under study.

Given the formats QAM modulation in DWDM system and expressions (11),...,(13) the probability of bit error of reception is calculated as follows [5, 12]:

$$P_{BER}(h) = 1 - \{\frac{1}{2^{m-1}\sqrt{2\pi}} \int_{-\infty}^{\infty} [2 - \mathrm{erfc}(\mathrm{u}/\sqrt{2})]^{\mathrm{m\text{-}1}} \cdot \exp[-0.5[(u - OSNR(E_b, \lambda_i, \sigma_A^2)]^2]^m\, du\}^{\mathrm{m}} \tag{14}$$

Formula (14) defines the dependence of the probability error reception of one *m* subcarriers on $OSNR(E_b, \lambda_i, N_0)$, which characterizes the probability bit errors of optical signal reception.

Note that on the basis calculation methods similarly, it is possible to obtain analytical expressions for bit error probability and $OSNR(E_b, \lambda_i, N_0)$ when using M-QAM type QAM-64 and QAM-256.

On the basis formula (14), the Communications Toolbox package in Matlab(R 2019b, 9.7; 64 bit) environment was used [10]. One of the modules this extension is the graphical environment BERTool with the help which the dependence bit error probability on $OSNR(E_b, \lambda_i, \sigma_A^2)$ was plotted. The following numerical value swore used to plot the graphical dependence:

$OSNR(E_b,\sigma_A^2)=[1,...,24]dB$, $\lambda_i=1.55\,\mu m$ and $V_b(b_i,\lambda_i)=(5,...,10)Gbps$, modulation scheme QAM-8 and QAM-16.

Figure 3 the dependence of the bit error probability on the signal-to-noise ratio per bit, $P_{BER}(h)=\mathrm{W}[\mathrm{OSNR}(\mathrm{E_b},\sigma_\mathrm{A}^2),\mathrm{M},\mathrm{V_b}(\mathrm{b_i},\lambda_\mathrm{i})]$ was presented based on numerical values.

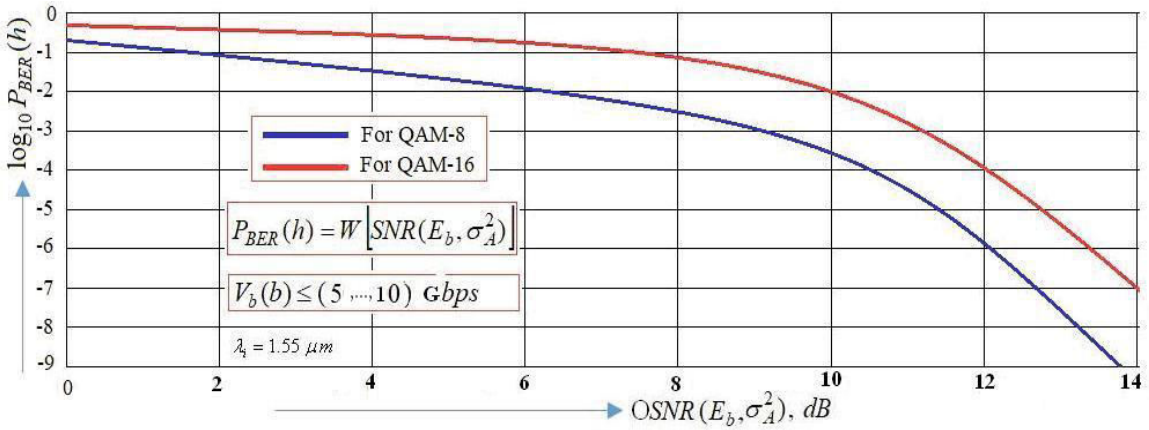


Figure 3: *Graphical dependence bit error probability on OSNR when using QAM-8 and QAM-16*

Graphical dependence, depicted in figure 3, clearly demonstrates the improvement bit error probability level in DWDM system with increasing modulation $M \geq 8$ and $\lambda_i=1.55\,\mu m$ when using new QAM-8 and QAM-16 modulation formats.

The conducted studies have shown that in the DWDM system, when using new modulation formats such as M-PPM, the requirement for optical $SNR(h)$ increases with the growth of positional modulation to achieve the same value of the probability error in receiving P_{BER} an optical signal.

The conducted studies have shown that in a DWDM system, when using new modulation formats such as M-PPM, the requirements for optical SNR increase with the growth positional modulation to achieve the same value of the probability error in receiving an optical signal P_{BER}. Figure 4 shows a family of graphical dependencies $P_{BER}=\mathrm{F}[\mathrm{V_b}(\mathrm{b_i},\lambda_\mathrm{i}),\mathrm{OSNR}(\mathrm{h})]$.

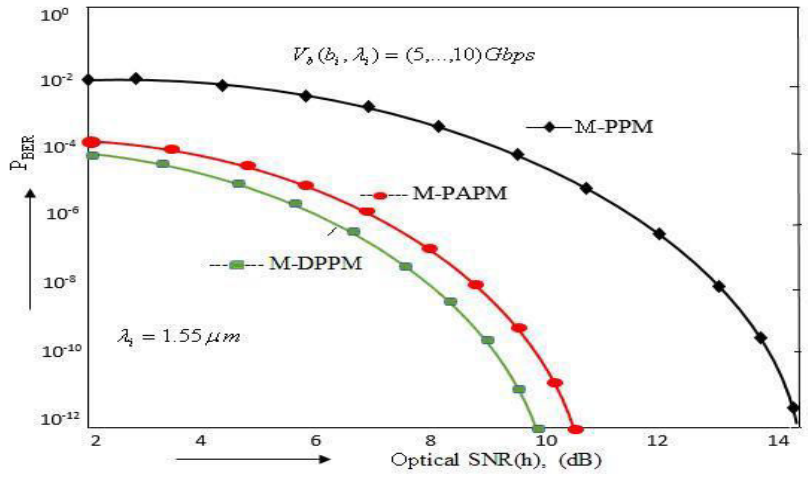


Figure 4: *Graphical dependence of the probability of bit errors P_{BER} on the optical SNR(h)*

From the family graphic dependencies it follows that with increasing optical $SNR(h)$ the probability bit errors is minimized due to the use new formats such as M-ary PPM, M-ary PAPM

(M-ary Pulse Amplitude and Position Modulation) so and M-Digital PPM at a given speed $V_b(b_i, \lambda_i) = (5,...,10)Gbps$.

The above-mentioned efficient and new modulation formats with high noise immunity types can be used in various infocommunication applications, including free-space optical (FSO) lines, hybrid optical fibers, optical wireless communications, next generation FSO, satellite communication systems and atmospheric optical communications.

VI. Conclusions

1. As a result of the study of fiber-optic networks, a new approach to the creation of methods for calculating the spectral efficiency and noise immunity optical signal reception using wavelength multiplexing was proposed.

2. The principle constructing fiber-optic networks using WDM and DWDM technology was analyzed, a structural diagram fiber-optic systems for transmitting optical signals via FOC at many wavelengths with EDFA optical amplifiers based on wavelength multiplexing was proposed.

3. Based on the calculation methods, analytical expressions were obtained that allow estimating the line capacity indicators, spectral efficiency of new modulation formats, optical signal-to-noise ratio and bit error probabilities in a DWDM system.

4. Numerical calculations of spectral efficiency and noise immunity indicators of reception were carried out in the MATLAB environment and graphical dependencies were constructed for comparative analysis using new modulation formats M-QAM and M-DPPM.

References

[1] ITU-T FG NET-2030 Technical Report "Gap Analysis of New Services, Capabilities and Use Cases for the Networks in 2030 and Beyond". Geneva, 2020.

[2] Ibrahimov B. G. (2023) Investigation of noise immunity telecommunication systems according to the criterion energy efficiency. Transport and Telecommunication, 24(4), 375-384.

[3] ITU-T FG NET2030 Technical Report. "Network 2030 – Additional Representative Use Cases and Key Network Requirements for Network 2030". Geneva, 2020.

[4] Proakis J. G. (2007) Digital communications. 5th edition, McGraw-Hill. 1150 p.

[5] Ibrahimov B.G., Hashimov E.Q., Hasanov A.H., Talıbov A.M. Research and analysis indicators fiber-optic communication lines using spectral technologies//Advanced Information Systems. 2022. Vol. 6, No. 1. pp. 61- 64.

[6] Sklar B. (2017) Digital Communications: Fundamentals and Applications. Prentice Hall Communications Engineering and Emerging Technologies Series, 1104 p.

[7] Hien T., Pham T., Ngoc D. Analog network coding aided multiuser visible light communication networks using optical CDMA. OSA Continuum, Vol. 2, No. 9 , 2019 . pp. 2569 – 2580.

[8] Pushkina E. O. (2013) BER performance OFDM-QAM over AWGN channel. Infocommunication technologies, 12 (2), pp. 59-62.

[9] Listvin V.N., Treshchikov V.N. DWDM systems. M.: Technosfera. 2015. 256p.

[10] Shelukhin O. I. (2018) Modeling of information systems. Textbook for universities. M.: Hotline – Telecom, 516 p.

[11] Elsayed, E.E.; Yousif, B.B. Performance enhancement of hybrid diversity for M-ary modified pulse-position modulation and spatial modulation of MIMO-FSO systems under the atmospheric turbulence effects with geometric spreading. Opt. Quantum Electron. 2020, 52.pp.1–18,

[12] Hayal, M.R.., Yousif, B.B., Azim, M.A. Performance Enhancement of DWDM-FSO Optical Fiber Communication Systems Based on Hybrid Modulation Techniques under Atmospheric Turbulence Channel. Photonics, 2021, 8, 464, pp. 1-19.

RESEARCH AND ANALYSIS QOS AND QOE INDICATORS IN MULTISERVICE TELECOMMUNICATION NETWORKS

Vyacheslav Shuvalov[1], Sevinc Ismayilova[2], Samina Rustamova[3]

[1]Siberian State University of Telecommunications and Informatics, Russia
[2]Azerbaijan Technical University, H. Javid ave. 25, AZ1073, Baku, Azerbaijan
[3]Nakhchivan State University, University campus, AZ7012, Azerbaijan
shvp04@mail.ru, sevinc.ismayilova@aztu.edu.az, saminarustamova@ndu.edu.az

Abstract

The rapid evolution multiservice telecommunication networks and the increasing demand for multimedia services have heightened the importance of ensuring Quality of Service (QoS) and Quality of Experience (QoE) for user satisfaction. This study explores QoS and QoE metrics in multiservice networks, leveraging insights from the "Network 2030" project, advanced digital technologies like SDN, NFV, 5G, and emerging 6G frameworks. With a focus on modern, end-to-end digital architectures, the paper proposes a mathematical model for analyzing and optimizing performance indicators, considering parameters such as network throughput, reliability, delay, and packet loss. The model aims to support real-time QoE assessments and enhance service quality by identifying performance thresholds across applications, from voice and video to M2M traffic. Metrics are examined through ITU-T recommendations, incorporating probabilistic-temporal characteristics and subjective user perceptions measured by the MOS (Mean Opinion Score) and R-factor ratings. Findings suggest that dynamic QoE monitoring, aligned with the convergence of NGN and FN architectures, offers significant potential to address rising user expectations in content quality and service reliability, guiding future research toward efficient QoS and QoE assessment in multiservice environments.

Keywords: Bandwidth, QoS, based routing, multiservice network, Quality of Experience, network resources.

I. Introduction

The development research outcomes from the "Network 2030" project, conducted by the ITU-T Focus Group FG NET-2030 on exploring the potential and principles building fixed communication networks, alongside the growth in multimedia traffic volume (voice, data, real-time streams, machine to machine (M2M), and video traffic), highlights the importance ensuring timely delivery. One of the key objectives in multiservice telecommunication networks based on end-to-end digital technologies is to guarantee quality of service (QoS) and quality of experience (QoE) for both user and service traffic when providing infocommunication services and applications [1, 2, 3].

It is worth noting that the analysis of real packet transmission paths conducted in [1, 2] shows that both user and service traffic in multiservice telecommunication networks exhibit structural complexity. This complexity reduces the timeliness traffic processing at nodes in multiservice

networks based on modern end-to-end digital technologies [4, 5].

Advanced end-to-end digital technologies include SDN (Software Defined Networking), NFV (Network Functions Virtualization), IoT (Internet of Things), LTE (Long Term Evolution), IMS (IP Multimedia Subsystem), AI (Artificial Intelligence), ML, WDM & DWDM (Wavelength Division Multiplexing & Dense Wavelength Division Multiplexing), 5G–NR-U (New Radio-Unlicensed), and 6G [5, 6, 7].

A new model for ensuring service quality and experience has become crucial to achieving the service level expected by today's subscribers. Critical infocommunication services require proactive SLA (Service Level Agreements) monitoring and Diffserv in real time to ensure guaranteed service quality and user experience [6, 7, 8].

Given the above and the shift toward next generation (NGN, Next Generation Network) and future networks (FN, Future Network), improving algorithms and protocols has become especially relevant. These networks provide voice, data, and video transmission services and incorporate the convergence of mobile and fixed public networks [3, 8].

In studies [2, 3, 4], certain tasks related to ensuring QoS and QoE, focusing on optimizing probabilistic and temporal performance metrics in multiservice telecommunication networks, were analyzed. In work [5, 6, 7], methods for routing, traffic clustering, and guaranteed service quality in communication systems were examined, identifying only some performance metrics for transmission efficiency and interference resistance in reception.

An optimization task arises to investigate QoS and QoE metrics by selecting probabilistic and temporal characteristics based on the efficiency criterion of multiservice telecommunication networks built according to the architectural concepts NGN (Next Generation Network), FN (Future Network), and the research outcomes ITU-T FG NET-2030 using the latest end-to-end digital technologies [9, 10, 11].

The goal of this work is to develop a new approach for creating a mathematical model to analyze QoS and QoE characteristics in multiservice telecommunication networks using end-to-end digital technologies.

II. General Problem Statement

In recent years, there has been a shift from heterogeneous public telecommunication networks - each designed to provide a wide range infocommunication services -NGN and FN, known as multiservice telecommunication networks.

It is known [2, 3, 5] that designing multiservice telecommunication networks requires considering that the primary load is generated by infocommunication services and applications such as "Triple Play Service" and "Bandwidth on Demand."

As a result, multimedia load is increasingly occupying a larger share in public communication network traffic, corporate networks, and military networks. The primary development focus of these networks is on creating scalable, mobile, reliable, and secure networks with context-aware delivery of essential, additional, and intelligent services.

It should be noted that to study QoS and QoE for multimedia traffic in NGN and FN networks, the following tasks need to be considered [4, 10, 11]:

- The fundamental concepts of QoS and QoE in multiservice telecommunication networks, taking into account ITU-T recommendations E.800 and G.1000, which define the relationships among operational characteristics - performance, reliability, and probabilistic-temporal characteristics of multiservice networks;
- QoS and QoE characteristics in NGN and FN multiservice networks, based on ITU-T Y.1540;

● Network throughput and reliability indicators in public communication networks, including network availability and the readiness of hardware-software systems;

● Probabilistic-temporal characteristics of the multiservice communication network;

● A broad range of QoS and QoE requirements in multiservice networks;

● The sensitivity various applications to multiservice network characteristics, considering traffic types (voice, e-commerce, video conferencing, email) and their parameters (bandwidth, packet loss, delay, and jitter).

Considering the aforementioned QoS and QoE metrics, a new approach is proposed for developing a mathematical model to analyze and synthesize the quality characteristics multiservice telecommunication networks. This approach selects the target function, which is described by the following objective functions:

$$Q_{KF}[(\lambda_i)] = \mathrm{W}\{Arg\max_i B[Q_S(\lambda_i), Q_E(\lambda_i)]\},\ i = \overline{1,k} \quad (1)$$

under the following constraints

$$G_P(\lambda_i) \geq G_{P.all.}(\lambda_i),\ T_{ptx}(\lambda_i) \geq T_{ptx.all.}(\lambda_i),\ I_S(\lambda_i) \leq I_{S.\mathrm{all.}}(\lambda_i),\ i = \overline{1,k} \quad (2)$$

$$C_{.\max}(\lambda_i) \leq C_{\max\ all.}(\lambda_i),\ H_S(\Lambda_i, t) \geq H_{S.all.}(\Lambda_i, t),\quad i = \overline{1,k} \quad (3)$$

where $G_P(\lambda_i)-$ a function that considers the performance criteria of multiservice telecommunication networks, taking into account QoS and QoE metrics, as well as the rate of incoming user and service traffic $\lambda_i,\ i = \overline{1,k}$;

$B[Q_S(\lambda_i), Q_E(\lambda_i)]-$ a function that takes into account QoS and QoE characteristics considering the intensity of incoming user and service traffic, $\lambda_i,\ i = \overline{1,k}$ respectively; $Q_{KF}(\lambda_i)-$ a function that considers the performance quality criteria multiservice telecommunication networks, taking into account QoS and QoE metrics, as well as the rate of incoming user and service traffic $\lambda_i,\ i = \overline{1,k}$;

$T_{ptx}(\lambda_i)-$ a function that considers the probabilistic-temporal characteristics of multiservice telecommunication networks based on packet switching when transmitting the *i*-th packet stream, taking into account the rate incoming user and service traffic $\lambda_i,\ i = \overline{1,k}$;

$C_{\max}(\lambda_i)-$ criteria that consider the maximum throughput metrics of multiservice communication networks, taking into account the rate of incoming traffic λ_i flow when processing the *i*-th traffic packets;

$H_S(\Lambda_i, t)-$ criteria that consider single and composite reliability metrics of the hardware and software systems of communication networks with a failure rate of Λ_i at time t, respectively $i = \overline{1,k}$;

$I_S(\lambda_i)-$ the information security coefficient of the hardware and software systems communication networks, considering the failure rate λ_i when servicing the *i*-th packet stream at time t:
W – operator of the combined transmission of user and service traffic at a network node when evaluating their performance quality as QoS, SLA, and QoE.

$G_{S.all.}(\lambda_i), T_{ptx.all.}(\lambda_i), C_{\max\ all.}(\lambda_i), I_{S.all.}(\lambda_i, t)$ and $H_{S.all.}(\Lambda_i, t)-$ respectively, the allowable values of performance, probabilistic-temporal characteristics, maximum throughput, and resilience of the telecommunication system to information security threats; single and composite reliability metrics of the hardware and software systems of multiservice telecommunication networks, taking into account the intensity $\lambda_i\ and\ \Lambda_i$ when servicing the *i*-th packet stream at time t.

The proposed expressions (1), (2), and (3) define the essence of a new scientific and practical approach through which a mathematical model is built for analyzing and synthesizing QoS and QoE characteristics in multiservice telecommunication networks using modern SDN, NFV, and IMS end-to-end digital technologies that account for comprehensive quality metrics of the communication system's performance.

Additionally, formulas (1), (2), and (3) offer a straightforward analytical expression for the efficiency and interference resistance functions of multiservice network performance based on the next-generation NGN and FN network architecture concepts, projected up to 2030, under the initiative "Network 2030," as reflected in ITU-T recommendations Y.3000 [3, 10, 11, 12].

Thus, a mathematical model is proposed to formalize the problem, aiming to accurately represent the telecommunications processes occurring within the studied segment of the multiservice network when providing infocommunication services and to enable analytical expressions for calculating key QoS and QoE characteristics.

III. Research and Analysis of Quality of Service Metrics in Multiservice Networks

It should be noted that the development of multiservice telecommunication networks is determined by three factors [4, 5, 6]:

- Traffic growth;
- Society's need for new info communication services;
- Advances in end-to-end digital technologies.

However, these factors are not independent; each of them shapes the ideology behind the development of the public telecommunication system. For example, competition among suppliers of terminal, channel, and switching equipment, along with technological advances, has led to a reduction in equipment costs, which, in turn, has stimulated traffic growth, improved QoS and QoE metrics, and the development new core, supplementary, and intelligent services.

One of the most popular data transmission services in recent years is data transmission via the IP protocol. The widespread adoption of the Internet and the services based on it-such as access to Internet resources, Internet service providers, and e-commerce services-has played a key role in this popularization.

However, the rise in the popularity of Internet services does not necessarily mean an increase in the revenue of telecommunications operators providing these services, primarily due to the decline in rates for Internet service provision.

Considering the concept QoS, let us examine the basic concepts in the field of telecommunication service quality. According to the ITU-T recommendation I.112, the entire set of telecommunication services is divided into two types [4, 5, 6, 7]:

- Bearer Service (BS) – traffic transport;
- Teleservice (TS) – communication provision.

The concept of the above-mentioned types of service encompasses [4, 8, 9]:

- Various types of communication – telephony, data transmission, fax, document retrieval;
- Core and supplementary services;
- Information transmission using various switching methods;
- Provision of different transmission mediums – wired, optical (based on WDM and DWDM technologies), and radio communication.

The hierarchy of concepts in the field of telecommunication service quality is shown in Figure 1, as considered in recommendation E.800.

Figure 1 presents the ITU-T model explaining terms related to QoS metrics. As seen in Figure 1, quality of service combines the concepts of effectiveness, security, reliability, and ease of use.

The group of properties such as availability, integrity, and continuity is combined into a single concept – effectiveness, as shown in Figure 1.

Furthermore, from Figure 1, it follows that the performance of a multiservice telecommunication network (NP, Network Performance) characterizes the effectiveness of service for both user and service traffic.

Network characteristics are defined as the ability to ensure communication between users. NP refers to a set parameters that can be calculated and measured according to ITU-T recommendation I.350. Here, the ability to ensure communication is one of the important network resources [5, 10].

Based on the task formulation, the wide range QoS metrics in multiservice networks relates to the resources of network throughput capacity hardware-software systems in the communication system.

The network resources, such as the maximum throughput capacity communication systems $C_{\max}(\lambda_i)$, considering the intensity of the incoming flow λ_i when processing the *i*-th traffic packets, are determined as follows [5]:

$$C_{\max}(\lambda_i) = (\frac{\lambda_i}{N_m}) \cdot E[L_{i.n}] \cdot (\rho_{i.upl.} + \rho_{i.\text{dow.}})^{-1}, \quad i = \overline{1,k} \tag{4}$$

where N_m – the number communication channels established in multiservice communication networks; $E[L_{i.n}]$ – the average length of the transmitted packet when transmitting the *i*-th traffic stream $i = \overline{1,k}$;

$\rho_{i.upl.}$ and $\rho_{i.upl.}$ – respectively, the load coefficients in multiservice communication networks when transmitting messages over the downstream and upstream communication channels.

It should be noted that with the increasing demands for quality of service and quality of experience in multiservice telecommunication networks built on the basis NGN and FN concepts, more and more attention is being paid to routing mechanisms [1, 4].

The reason for this is that the DiffServ functionality, based on priority packet processing at the nodes (routers) of the telecommunication network, can only improve QoS at specific network elements. However, the routing protocol in next generation telecommunication networks must ensure the calculation of one or more paths for packet delivery.

Figure 1 shows the ITU-T model explaining terms in the field QoS and QoE [4, 10]. It should be noted that the terms in the model related to quality of service and quality of experience are defined by ITU-T recommendations E.800, Y.2000, G.628, I.112, and Y.3001 [3, 4, 10, 12].

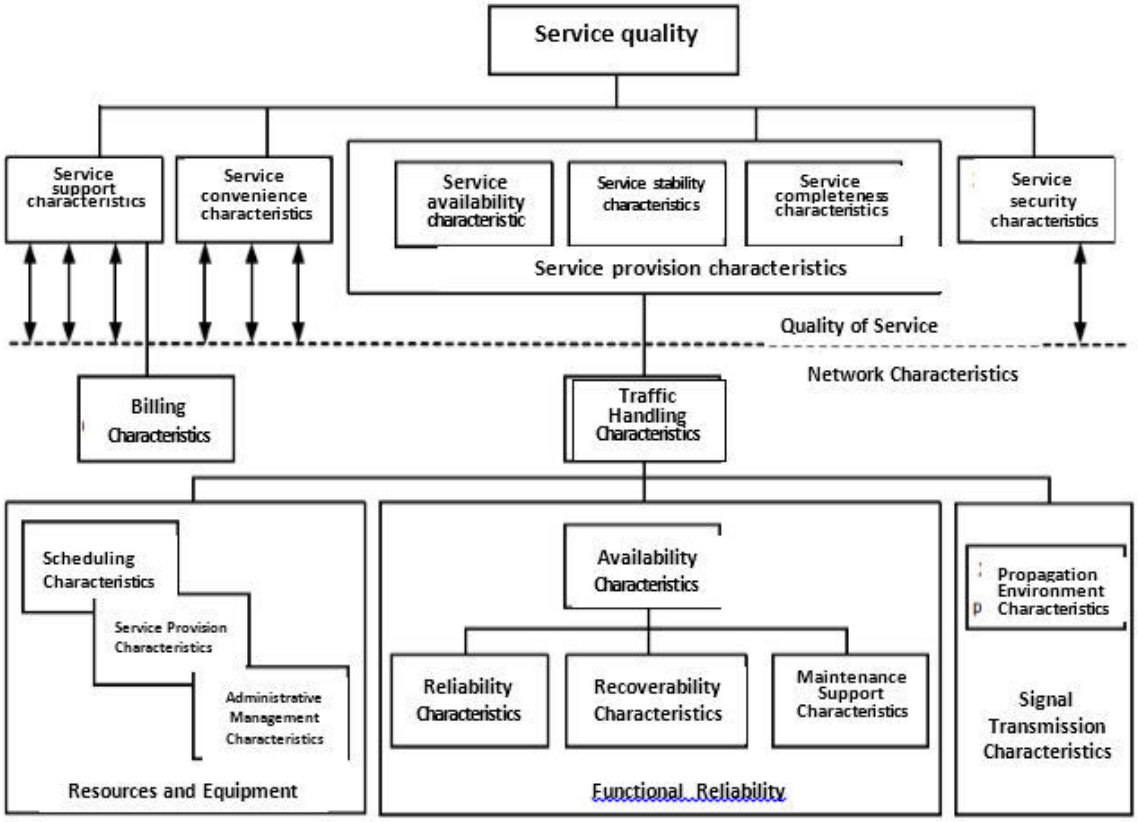


Figure 1: *ITU-T model explaining terms in the field QoS and QoE*

It should be noted that with the increasing demands for QoS and QoE in multiservice telecommunication networks built on the basis of the NGN and FN concepts, more and more attention is being paid to routing mechanisms [1, 7].

The reason for this is that the DiffServ functionality, based on priority packet processing at the nodes (routers) of the telecommunication network, can only improve QoS at specific network elements. However, the routing protocol in next-generation telecommunication networks must ensure the calculation of one or more paths for packet delivery.

Table 1 presents the network performance metrics, taking into account QoS classes [4, 11].

Table 1: *Standards for IP network characteristics with distribution by QoS classes*

Network Characteristics	QoS Classes					
	0	1	2	3	4	5
IP Packet Delivery Delay (IPTD)	100 ms	400 ms	100 ms	400 ms	1 s	H
IP Packet Delay Variation (IPDV)	50 ms	50 ms	H	H	H	H
IP Packet Loss Ratio (IPLR)	$1x10^{-3}$	$1x10^{-3}$	$1x10^{-3}$	$1x10^{-3}$	$1x10^{-3}$	H
IP Packet Error Ratio (IPER)	$1x10^{-4}$	$1x10^{-4}$	$1x10^{-4}$	$1x10^{-4}$	$1x10^{-4}$	H

The analysis Table 1 shows that the network characteristics include the probabilistic and time characteristics multiservice telecommunication networks, which relate to the performance and interference resistance metrics for receiving user and service traffic.

Furthermore, Table 1 presents the standards for IP network characteristics with distribution by QoS classes. In Table 1 H- indicates that the metrics are not standardized.

IV. Research on Quality of Experience indicators in multiservice networks

Currently, the range infocommunication services provided by operators is significantly expanding, with various applications emerging that enhance users' ability to communicate and exchange information. In addition to the wide range existing applications related to data, voice, and video transmission, the capabilities of heterogeneous networks are opening new horizons for providing new services.

It is worth noting that a prominent example of service penetration into all areas of life is the concept of the Internet of Things, which is changing users' perceptions infocommunication services and telecommunication networks as a whole [1, 7]. It is also a natural process to reconsider the factors affecting the quality of the services provided and, as a result, the transformation of QoS and QoE [2] and the expansion indicators characterizing the level of service quality.

A completely new issue in the field of multiservice networks is the issue of "trust" in the information received from terminal devices, sensors, and detectors. How reliable are the data received, or is it a network failure, data distortion.

This leads to a change in approaches to evaluating QoE for both service and user traffic. The components of perception by humans, which are part of the subjective indicators included in QoE, no longer seem so unattainable, unclear, and useless.

It is known [3, 5, 7] that the most widely used subjective quality assessment methodology is described in ITU-T Recommendation P.800 and is known as the MOS (Mean Opinion Score) method. In multiservice networks, for user and service traffic, expert assessments are determined according to the following five-point scale: 5 – excellent, 4 – good, 3 – acceptable, 2 – poor, 1 – unacceptable.

For speech quality, a MOS score of 3.5 and above corresponds to standard and high telephone quality, 3.0…3.5 is acceptable, and 2.5…3.0 represents synthesized sound. For speech transmission with good quality, it is advisable to aim for a MOS score of at least 3.5. At the same time, R-factor values are directly correlated with MOS scores as follows [4, 10, 11]:

$$R = Ro - Is - Id - Ie + A \quad (4)$$

where Ro = 93.2 – the initial value of the R-factor; Is – distortions introduced by codecs;

Id – distortions due to total end-to-end delay in the network;

Ie – distortions introduced by equipment, including packet losses; A – the so-called advantage factor.

From expression (4), it follows that considering the distortions that occur when converting real speech to an electrical signal and back, the theoretical value of the R-factor, without distortions, decreases to a value of 93.2, which corresponds to a MOS score of 4.4.

Table 2 presents the R-factor values and the corresponding MOS scores. From Table 1, it follows that, considering the quality categories and user assessments, the QoS and QoE ratings based on the R-factor and the resulting MOS scores are provided [4,].

Table 2: *R-factor indicators and MOS score values*

R-Factor Value	Quality Category and User Rating	MOS Score Value
90<R<100	Highest	4,34 – 4,50
80<R90	High	4,03 – 4,34
70<R<80	Medium (some users rate quality as unsatisfactory)	3,60 – 4,03

60<R<70	Low (most users rate quality as unsatisfactory)	3,10 – 3,60
50<R<60	Poor (not recommended	2,58 – 3,10

Therefore, the need for a detailed assessment of the quality of service perception, along with the constantly increasing user demands for content parameters, forces the development new approaches for evaluation in multiservice networks. However, dynamic real-time assessment QoE reveals new ways to achieve maximum user satisfaction with services and deepen the emotions they experience.

Thus, timely determination subjective QoE parameters for each user and their correct analysis are promising research directions in the field QoE.

V. Conclusions

1. As a result of the study, a new approach was proposed for creating a mathematical model for analyzing QoS and QoE characteristics in multiservice telecommunication networks and the target function was selected under certain restrictions using end-to-end digital technologies.

2. The proposed mathematical model for analyzing QoS and QoE characteristics takes into account the criteria for the quality of functioning multiservice telecommunication networks and selected as performance parameters, probabilistic-temporal characteristics, throughput, individual and complex indicators of the reliability functioning hardware and software systems, as well as information security of the functioning of multiservice telecommunication networks.

3. The ITU-T model is given, explaining the terms in the field of QoS and QoE, taking into account the ITU-T recommendations, E.800, Y.2000, G.628, I.112 and Y.3001, and the standards for the characteristics IP networks with distribution by quality of service classes and QoE indicators are analyzed, taking into account the R-factor and the MOS assessment value.

References

[1] Network 2030. A Blueprint of Technology, Applications and Market Drivers Towards the Year 2030 and Beyond. FG-NET-2030. Geneva, 2019.

[2] Ositis A.P., Efimushkin V.A. The role of "end-to-end" digital technologies in the development of telecommunications // Telecommunications. No. 1, 2021. pp. 7-11.

[3] Jafarova E.M., Ibrahimov B.G., Ismailova S.R. Analysis complex indicators multiservice telecommunication networks based on FN architectural concepts // Collection scientific articles - IX-International Scientific, Technical and Scientific-Methodological Conference "Actual Problems Infotelecommunications in Science and Education". SPb.: SPbSUT. Vol.1. 2020. pp.399-404.

[4] Telecommunication systems and networks: Textbook / In 3 volumes. Volume 3. – Multiservice networks / V. V. Velichko, E. A. Subbotin, V. P. Shuvalov, A. F. Yaroslavtsev; edited by Professor

V. P. Shuvalov. M.: Hot- Link-Telecom, 2015. – 592 p (in Russian).

[5] Bayram G. Ibrahimov and Almaz A. Alieva. An Approach to Analysis of Useful Quality Service Indicator and Traffic Service with Fuzzy Logic//10th International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions-ICSCCW-2019. Advances in Intelligent Systems and Computing. Vol. 1095. 2019. pp.495-504.

[6] Listopad N.I., Mikhnevich S.Yu.,. Hayder A.A. QoS based routing of information flows in telecommunication networks// Problems of Physics, Mathematics and Technology, No. 2 (27), 2016. pp. 90–96.

[7] Recommendation G.1011. Reference Guide to Quality of Experience Assessment Methodologies. ITU-T, Geneva, May 2013.

[8] Narayanan, A., Rochman, M.I., Hassan, A., Firmansyah, B.S., Sathya, V., Ghosh, M., Qian, F., Zhang, Z.L. A Comparative Measurement Study of Commercial 5G mmWave Deployments. In Proceedings of the IEEE INFOCOM 2022-IEEE Conference on Computer Communications, London, UK, 2–5 May 2022. pp. 800 – 809.

[9] Zwart A. P.Queueing systems with heavy tails/Eindhoven uneversity of technology.2019. 227 p.

[10] Goldstein B.S., Sokolov N.A., Yanovsky G.G. Telecommunication Networks. The textbook for high schools. – BHV. St.Petersburg. 2010. 400 P.

[11] Taehun Jung, Taehoon Kwon, Chan-Byoung Chae. QoE-based transmission strategies for multi-user wireless information and power transfer//ICT Express 1, 2015, pp.116–120.

[12] Choi B. D., Kim B., Wee I. Asymptotic behavior loss probability in GI/M/1/K queue as K-tends to infinity//Queueing Systems. Vol. 36, 2020. pp. 437 – 442.

SECURITY CONTROL SYSTEM FOR INFORMATION EXCHANGE IN TELECOMMUNICATION NETWORKS

Elvin Abaszade[1], Zafar İsmayilov[1], Almaz Mehdiyeva[2], Huseyn Qasimov[3]

[1]Azerbaijan Technical University, Huseyn Javid Ave. 25, Az1073, Baku
[2]Azerbaijan State Oil and Industry University, Azadliq Avenue 20, AZ1010, Baku
[3]Nakhchivan State University, University campus, AZ7012, Azerbaijan
elvinabaszadaphd@gmail.com, zafar.ismayilov@aztu.edu.az, almaz.mehdiyeva@asoiu.edu.az, huseynqasimov@ndu.edu.az

Abstract

The principles of information security and the problems faced were investigated. Cyber security issues are explored. Also, technical requirements for ensuring information security and means ensuring the security of Information systems were studied. Based on Visual Studio and MySQL programs, the issue of web page security and user data protection has been resolved. In this application, Users can access their personal cabinet by typing username and password on the web page. If a user is not registered on this page, then that user will encounter a problem, which means that the database does not have that user's information. The proposed method ensures security.

Keywords: Information security, telecommunication networks, cyber security, MySQL programs, information system, information exchange.

I. Introduction

Information system (IS) is a formal, socio-technical, organizational system designed to collect, process, store and distribute information. From a socio-technical point of view, information systems consist of four components: task, people, structure (or roles), and technology. Information systems can be defined as the integration of components for the collection, storage and processing of data, from which data are used to provide information, contribute to knowledge, as well as digital products that facilitate decision making.

A computer information system is a system consisting of humans and computers that process or interpret information. The term is also sometimes used to refer simply to a computer system on which software is installed.

"Information systems" is also the academic field study of the information-specific systems and complementary networks of computer hardware and software that people and organizations use to collect, filter, process, create, and disseminate information. Emphasis is placed on an information system that has a defined boundary, users, processors, memory, inputs, outputs, and the aforementioned communication networks.

In many organizations, the department or unit responsible for information systems and data processing is known as "information services" [1-5].

Any specific information system aims to support operations, management and decision making. An information system is the information and communication technology (ICT) used by an

organization, as well as how people interact with this technology to support business processes [6, 7].

Some authors make a clear distinction between information systems, computer systems and business processes. Information systems usually include an ICT component, but are not concerned with ICT alone, instead focusing on the end use of information technology. Information systems are also different from business processes. Information systems help control the execution of business processes [8, 9].

Alter advocates the advantages of viewing an information system as a specific business system. A work system is a system in which people or machines perform processes and activities using resources to produce specific products or services for customers. An information system is a business system whose activity is dedicated to capturing, transmitting, storing, retrieving, manipulating and displaying information [10-15].

Thus, information systems interact with information systems on the one hand, and activity systems on the other. An information system is a form of communication system in which information is represented and processed as a form of social memory. An information system can also be considered a semi-formal language that supports human decision-making and actions.

Information security, sometimes abbreviated to InfoSec, is the practice of protecting information by reducing information risks. It is part of information risk management. This typically involves preventing or mitigating unauthorized/inappropriate access or illegal use of information, disclosure, disruption, deletion, corruption, alteration, verification, recording or impairment of information. It also includes measures to reduce the negative effects of such events. Protected information can be in any form, e.g. electronic or physical, tangible (e.g. paperwork) or intangible (e.g. knowledge)) is balanced protection. This is mainly achieved through a structured risk management process that includes:

identification of information and related assets, plus potential threats, vulnerabilities and impacts;

- risk assessment;
- deciding how to address or treat risks ie. avoid, reduce, share or accept them;
- selecting or designing appropriate security controls and implementing them when risk mitigation is required;
- monitoring activities, making corrections, changes and improvements as necessary to solve any problems.

In order to standardize this discipline, scholars and professionals are working on passwords, antivirus software, firewalls, encryption software, legal liability, security awareness and training, etc. collaborate to offer guidelines, policies, and industry standards for the application of any standards and guidelines within may have limited impact.

The CIA's triad of privacy, integrity, and availability is the foundation of information security. interchangeably.) However, this CIA triad emphasizes the relationship between accessibility and privacy, as well as security and privacy [16, 17]. Debate continues as to whether it is sufficient to meet rapidly changing technology and business requirements, with recommendations to consider scaling up. Sometimes other principles have been proposed, such as “responsibility”; It has been pointed out that issues such as non-denial do not correspond to the three basic concepts. It appears that the triad was mentioned in a 1977 NIST publication [18-22].

In 1992 and revised in 2002, the OECD Guidelines for the Security of Information Systems and Networks proposed nine generally accepted principles: awareness, responsibility, responsiveness, ethics, democracy, risk assessment, security design and implementation, security management and re-evaluation [23]. Based on these, in 2004 NIST's Engineering Principles for Information Technology Security He proposed 33 principles. It is derived from each of these guidelines and practices.

In 1998, Donn Parker proposed an alternative model to the classic CIA triad, which he called

the six atomic elements of information. The elements are privacy, ownership, integrity, authenticity, accessibility, and utility. The merits of the Parkerian Hex are a matter of debate among security experts.

II. Technical requirements for ensuring information security

IT security standards or cybersecurity standards are generally methods described in published materials that attempt to protect a user's or organization's cyber environment. This environment includes the users themselves, networks, devices, all software, processes, data in storage or transit, applications, services and systems that can be directly or indirectly connected to networks [24].

The main objective is to reduce risks, including preventing or mitigating cyber-attacks. These published materials include tools, policies, security concepts, security safeguards, guidelines, risk management approaches, measures, training, best practices, security and technologies.

Cybersecurity standards have existed for several decades, as users and providers have collaborated in many domestic and international forums to implement the necessary capabilities, policies, and practices - generally originating in the 1990s with work at the Stanford Consortium for Research on Information Security and Policy.

A 2016 U.S. security framework adoption study reported that 70% of surveyed organizations recognized the NIST Cybersecurity Framework as the most popular best practice for information technology (IT) computer security, but many considered it a significant investment. states that it requires Cross-border, cyber-exfiltration Law enforcement operations against international criminal activities on the dark web raise complex jurisdictional questions that remain to some extent unanswered. Tensions between domestic law enforcement efforts to conduct cross-border cyber-exfiltration operations and international jurisdiction are likely to continue. provides improved cyber security norms. The following subsections detail international standards related to cybersecurity.

III. Principles of building information protection systems

In our modern era, information protection systems are used on web pages. Programs are used to build these systems. Now, for the information protection we will look at, it is required to use visual studio and mysql programs.

Visual Studio IDE (integrated development environment) is software for developers to write and edit their code. Its user interface is used for software development to edit, debug, and build code. Visual Studio includes a code editor that supports IntelliSense (a code completion component), as well as code refactoring. The integrated debugger works as both a source-level debugger and a machine-level debugger. Other built-in tools include a code profiler, a designer for building GUI applications, a web designer, a class designer, and a database schema designer [25].

MySQL is a relational database management system (RDBMS) developed by Oracle and based on structured query language (SQL).

A database is a structured collection of data. This can be anything from a simple shopping list to an image gallery or a place to store large amounts of data on a corporate network [26]. Specifically, a relational database is a digital store that collects data and organizes it according to a relational model. In this model, tables consist of rows and columns, and the relationships between data elements all follow a strict logical structure. An RDBMS is simply a set of software tools used to implement, manage, and query such a database [27-29].

A new database is created in MySQL and this database will contain username and userpassword (Figure 1, 2).

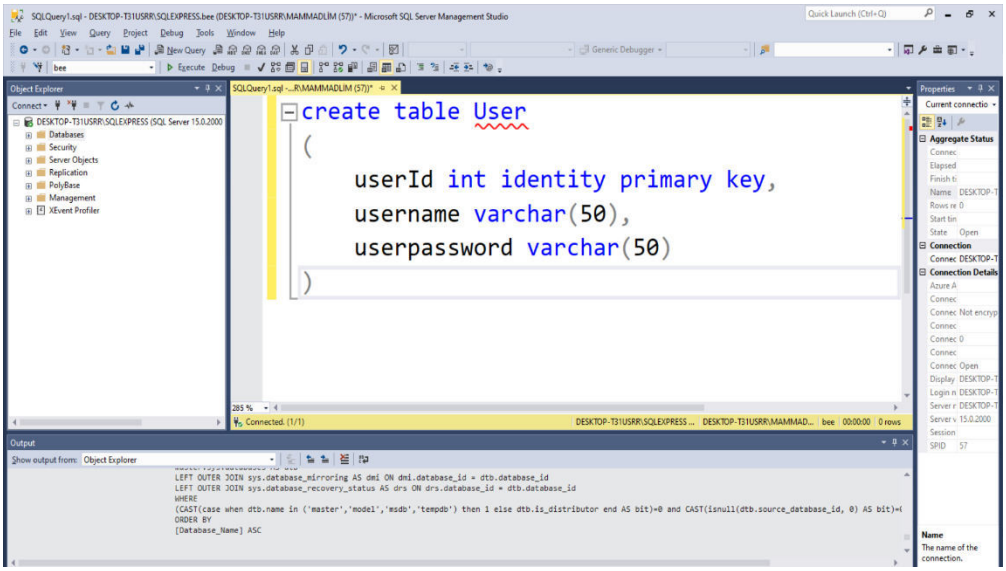


Figure 1: *User data*

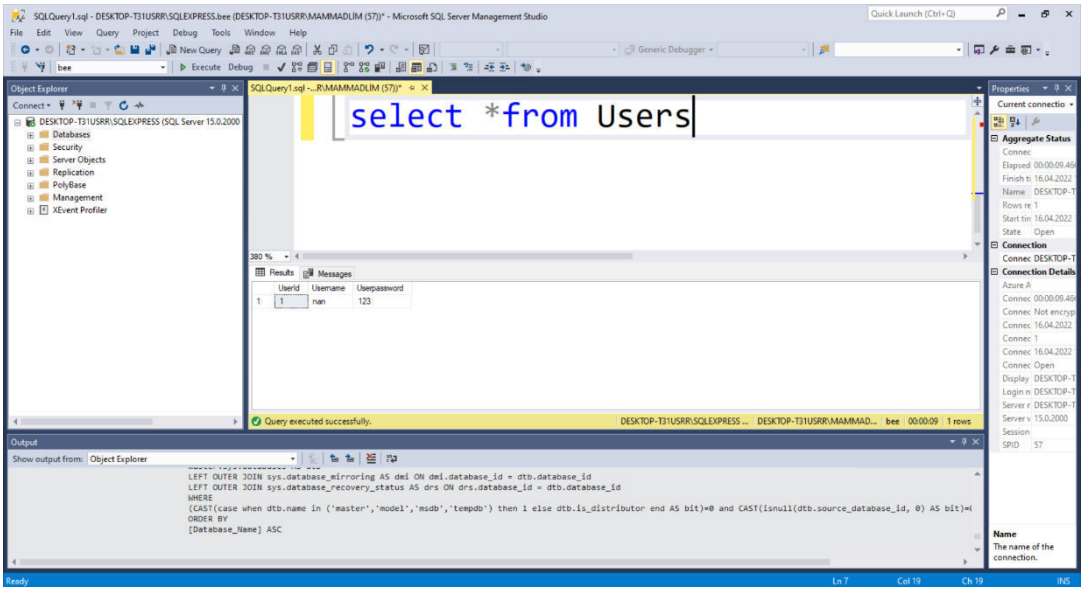


Figure 2: *User table*

At the next stage, we connect with visual studio and create a login in visual studio. With the connection we write in Visual Studio, we can get the information directly from MySQL.

IV. Ensuring security in information exchange

We provide security on web pages through the Visual Studio program and create security walls with the help of special codes. We receive a request from the database we have created under certain conditions, and this request is evaluated by the Visual Studio side. If the request is correct, then the user using our web page is redirected to the desired address, otherwise, if the user's request - user data is incorrect, then he will not be able to access the Web page (Figure 3).

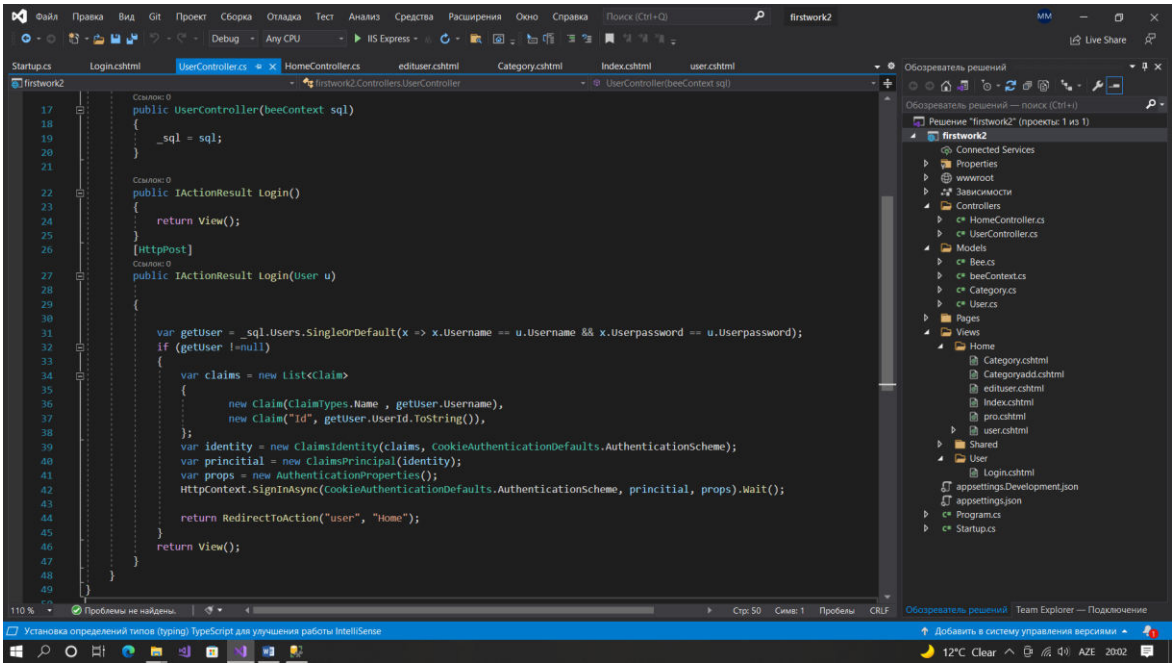


Figure 3: *Security codes*

V. Results and discussion

On our login page (Figure 2), enter your username and password, and enter your username and password for testing. Clicking on the sign in button here will redirect us to the user's personal page.

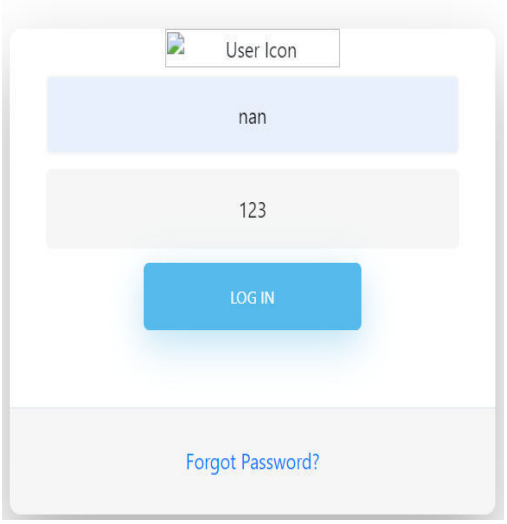


Figure 4: *Login*

The next stage is related to whether the user sent the information correctly or not. For this, we should look at the getuser code and Figure 5 was shown. The user will be redirected to the correct address as the information is already in the system.

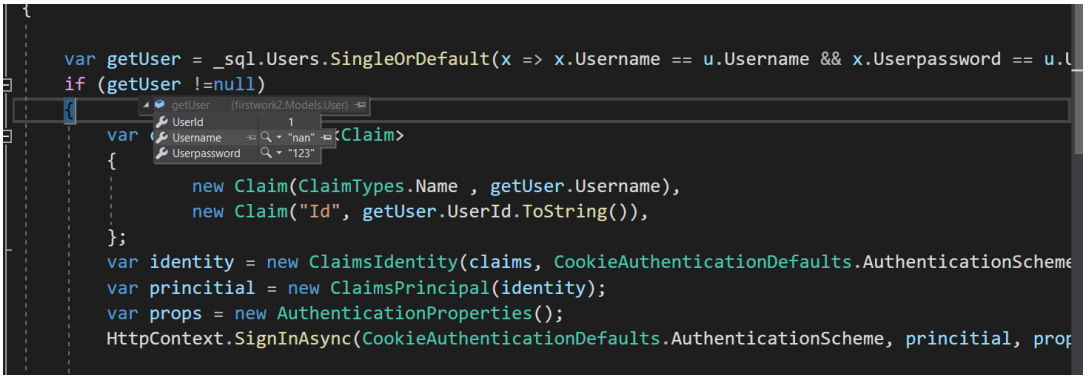


Figure 5: *Security check*

VI. Conclusion

The following results were obtained from the conducted research: Security control systems are used in information exchange, and these systems are also used in web pages. Firewalls are created on web pages and attacks are prevented by the software we use. From these pages, there is a registration page for users, and after registration, they can log into their personal accounts with the unique username and password provided to them.

References

[1] Kerstin., Fink (2004). Knowledge Potential Measurement and Uncertainty. Deutscher Universitätsverlag. ISBN 978-3-322-81240-7. OCLC 851734708.

[2] Keyser, Tobias (2018-04-19), "Security policy", The Information Governance Toolkit, CRC Press, pp. 57–62, doi:10.1201/9781315385488-13, ISBN 978-1-315-38548-8, retrieved 2021-05-28

[3] Ibrahimov B.G. Research and analysis of the efficiency of multiservice communication networks using the NGN architectural concept / B.G. Ibragimov, S.R. Ismaylova // T- Comm, Telecommunications and transport, - Moscow: - 2014. Vol. 8, No. 8, - pp. 47 - 50.

[4] Ibrahimov B.G., Ismaylova S.R. On one approach to assessing the quality of functioning of a signaling network link // All-Russian Scientific and Technical Conference "Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems", section - "Theory of Teletraffic", - Moscow: RUDN, - 2012, - pp. 38 - 40.

[5] Ibrahimov B.G., Mehdiyeva A.M., Bakhtiyarov I.N. Study of throughput indicators of corporate multiservice networks // Bulletin of Computer and Information Technologies, No. 5, Moscow, 2020. pp. 38 - 44.

[6] Ibrahimov B.G., Mehdiyeva A.M., Bakhtiyarov I.N. Mathematical model for assessing the level of noise immunity of the paths of systems for transmitting, processing and receiving packet messages // Proceedings of the XII International Scientific-Practical Conference. New Informational and Computer Technologies in Education and Science - IES-2020, – Ukraine, Vinnytsia: 2020, p .77–79.

[7] Tanenbaum E., Computer networks, Peter, 2003, pp. 361-370.

[8] Andreev R.N. Theory of electrical communication / R.N. Andreev, R.P. Krasnov, M. Hotline Telekom, 2014. 230 p.

[9] Sheluhin O.I. Modeling of information systems. O.I. Sheluhin. - Moscow: Hotline - Telekom, 2018. 516 p.

[10] Andreev R.N. Theory of electrical communication. Textbook for students / R.N. Andreev, R.P. Krasnov, M.Yu. Chepelev Hotline Telekom, 2014. 230 p.

[11] Michael P.F. Fundamentals of Communications Systems. Communications Engineering. New York: McGraw-Hill Companies, 2007. 436 p.

[12] Andreev R.N. Theory of electrical communication. Textbook for students / R.N. Andreev, R.P. Krasnov, M.Yu. Chepelev Hotline Telekom, 2014. 230 p.

[13] Bitner.V.I. Networks of the new generation-NGN. / V.I. Bitner, Ts.Ts. Mikhailova - Moscow: Hotline-Telecom. 2011. 228 p..

[14] Vasiliev K.K. Mathematical modeling of information communication systems. Moscow: Hotline Telekom. 2018. 236 p.

[15] Mehdiyeva A.M., Zeynalova, R.R., Safarova, A.A., Takhumova O.V., Nikolaevc P.P., Mozgovoy A.I. Development of an adaptive control system for the quality parameter in the lack of information. Proceedings of SPIE - The International Society for Optical Engineering, 12637, 1263707. doi: 10.1117/12.2681371, 2023, Fergana, Uzbekistan.

[16] Mehdiyeva A.M., Bakhtiyarov I.N., Bakhshaliyeva S.V. Increasing the Immunity of Information Transmission and Fault Tolerance of the Path. Lecture Notes on Data Engineering and Communications Technologies. Volume 166. Mobile Computing and Sustainable Informatics. Proceedings of ICMCSI 2023, 11-12 January 2023. Tribhuvan University, Nepal. pp. 775 784. http://icmcsi.com/2023.

[17] ISO/IEC 27000:2009 (E). (2009). Information technology – Security techniques – Information security management systems – Overview and vocabulary. ISO/IEC.

Committee on National Security Systems: National Information Assurance (IA) Glossary, CNSS Instruction No. 4009, 26 April 2010.

[18] Pipkin, D. (2000). Information security: Protecting the global enterprise. New York: Hewlett-Packard Company.

[19] McDermott, E., & Geer, D. (2001). Information security is information risk management. In Proceedings of the 2001 Workshop on New Security Paradigms NSPW '01, (pp. 97 – 104). ACM. doi:10.1145/508171.508187

[20] Anderson, J. M. (2003). "Why we need a new definition of information security". Computers & Security. 22 (4): 308–313. doi:10.1016/S0167-4048(03)00407-3.

[21] Michael P.F. Fundamentals of Communications Systems. Communications Engineering. New York: McGraw-Hill Companies, 2007. 436 p.

[22] Andreev R.N. Theory of electrical communication. Textbook for students / R.N. Andreev, R.P. Krasnov, M.Yu. Chepelev Hotline Telekom, 2014. 230 p.

[23] Bitner.V.I. Networks of the new generation-NGN. / V.I. Bitner, Ts.Ts. Mikhailova - Moscow: Hotline-Telecom. 2011. 228 p.

[24] Vasiliev K.K. Mathematical modeling of information communication systems. Moscow: Hotline Telekom. 2018. 236 p.

[25] Velichko V.V. Models and methods of increasing the durability of modern communication systems. Moscow: Hotline–Telekom 2016. 270 p.

[26] Velichko V.V. Models and methods of increasing the durability of modern communication systems. /V.V. Velichko, G.V. Popkov, V.K. Popkov. Moscow: Hotline.Telecom 2016. 270 p.

[27] Sheluhin O.I. Modeling of information systems. Teaching manual for universities. Moscow: Hotline – Telekom. 2018. 516 p.

[28] Sheluhin O.I. Modeling of information systems. Study guide for students. Moscow: Hotline Telecom. 2018. 516 p.

ELECTRICAL DISCHARGE IN DIELECTRIC STRUCTURES

Hikmet Aliyev[1], Hikmet Fattayev[1], Naib Hajiyev[1]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan
hikmetaliyev@aztu.edu.az, hikmet.fettayev@aztu.edu.az, naib.haciyev@aztu.edu.az

Abstract

Based on investigations on partial discharges under different physical conditions and when using different dielectric barriers, it was found that the discharge in the gas gap confined by dielectrics has a pulse character under all conditions. The discharge current is shown to consist of discrete pulses whose shape and frequency depend on the dielectric parameters, the type of gas in the gap and the pressure. One of the main reasons for the deterioration of the electrophysical properties of dielectrics is the development of electric discharges in dielectric structures consisting of dielectrics with significantly different properties. Therefore, the study of the influence of the properties of dielectrics and the electrophysical parameters of these dielectrics on the mechanism of development of the discharge under conditions of electric discharge is an urgent issue [1-3].

Keywords: Dielectric, microstructure, electrophysical parameters, electron-optical converter, discharge.

I. Introduction

Knowledge of microstructures and mechanisms of discharge formation, as well as the nature of physical and chemical processes occurring both in the volume of the gas medium and in the place of contact of microdischarges with the dielectric surface, is important for solving problems related to the creation of solid dielectrics operating under partial discharge conditions.

To reveal the microstructure, a spatial (optical) view of the discharge in the space between dielectrics under different electrophysical conditions was obtained. A special experimental apparatus was used, which allows to simultaneously record the electrical and optical picture of the discharge with the help of an oscillograph and an electron-optical converter (EOC) and a multi-frame optical time loop.

At present, it is not possible to fully understand the influence of the electrophysical properties of dielectrics on the microstructure and the mechanism of development of partial discharges.

The experimental apparatus used by us to study the electric discharge in the gas gap confined by dielectrics allowed simultaneous recording of electrical (current and voltage pulses, volt-coulomb properties) and optical scenes. The studied discharge was generated in an experimental core with disc-shaped electrodes made of stainless steel. Depending on the purpose of the research, the surface of the electrodes of the core is covered with different polymer dielectrics. The thickness of the polymer layers and their composites varies from 20 to 200 μm. The polymers are filled with various inorganic materials: semiconductors, magneto-piezoelectrics and semiconductor oxides. The electrodes are equipped with fixing means that create a reliable contact between the metal electrode and the dielectric. With the help of a micrometric screw it was possible to change the distance

between the electrodes from 0 to 15 mm with an accuracy of 10 μm. A high voltage with a frequency of 50 Hs was applied to the core under investigation.

The optical picture of the discharge was recorded by means of electron-optical time-lapse photography with multi-frame electron-optical converters (EOTC). The use of electron-optical converters, which provide multiple amplification of the image, is determined by the poor illumination and short lifetime of individual microdischarge channels. In our experiments, the maximum optical resolution reached 20 double lines/mm per frame across the entire screen of the EOCH converter.

The obtained optical images (EOC) (figure 1) show that the geometrical dimensions and microstructure of the channel of individual microcavities are highly dependent on the electrophysical parameters of the dielectrics bounding the gas gap. Naturally, it can be assumed that the value of the local residual voltage after the occurrence of an individual microdischarge in the air gap depends significantly on the dielectric properties.

Table 1 shows the average value of the diameter of the microdischarge channels for some dielectric anodes (positive electrode coated with dielectric as dielectric anode) and cathodes (negative electrode coated with dielectric). Other conditions being equal (constancy of the voltage amplitude at the electrodes, the thickness of the air gap), if polymer layers (PE, PP) are used as a dielectric, the diameter of the microdischarges near the anode is always the same as that of dielectrics with high dielectric permittivity, i.e. composites obtained on the basis of dielectrics dispersed with barium titanate ($BaTiO_3$) or non-organic dielectric particles of the LZT-19 type are smaller than those used.

PVDF + 20% vol. $BaTiO_3$ or PP + 10% vol. QST - 19 as a dielectric anode, i.e. the number of microdischarges in each series or in each hole of the gas gap (figure 1, b and c) is less than in the case with polymer layers (PP, PE) as anodes, i.e. dielectrics with low electrical conductivity and low dielectric permittivity (figure 1, a).

The amplitude of the applied voltage affects the number of micro-discharges in each series and therefore the amplitude of the voltage pulse generated during the perforation of the air gap.

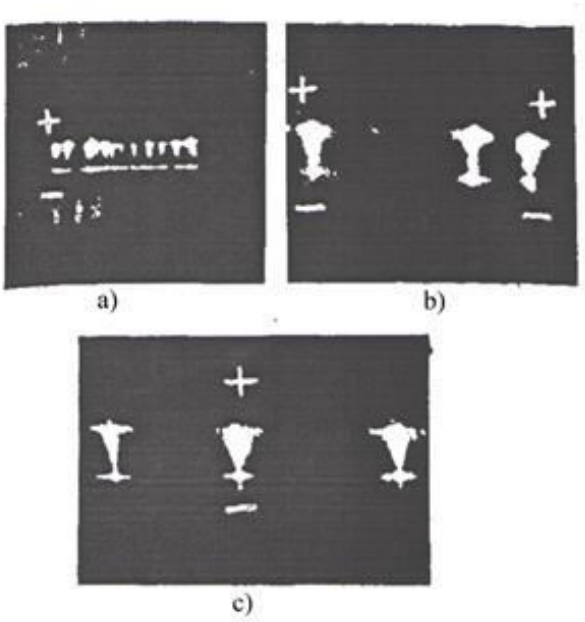


Figure 1: *Optical view of the discharge.*
a) PP - air gap - PP; ε = 2,3; b) PP + 20% vol. LZT-19 - air gap - PP + 20% vol. QST-19; ε = 10;
c) PVDF + 20% vol. LZT-19 - air gap - PVDF + 20% vol. LZT-19; ε = 18

Dielectric anode and cathode thickness 200 μm; thickness of the air gap between two-phase composites d = 4,5 mm.

Table 1: *Dependence of the diameter of the channel of individual microdischarges near the anode on the type of dielectric limiting the air gap*

Thickness of the air gap d, mm	Dielectric anode and cathode types			
	Diameter of micro-discharges D, mm			
	PVDF 20 % vol. $BaTiO_3$	PE	PP	PP+20 % vol. LZT - 19
1	1,15	07,	0,6	0,86
2	1,3	0,92	0,84	1,1
3	2,2	1,2	1,00	1,5
4	2,6	1,43	1,3	2,0
5	3,0	1,53	1,45	2,2
6	3,64	1,8	1,8	2,36

With a change in the thickness of the air gap between the dielectrics, the illumination brightness and geometrical dimensions of the microdischarge channels change. For example, with an increase in the thickness of the air gap under the same conditions, the diameter of the microdischarge channels near the dielectric anode increases significantly (table 1). In addition, under the conditions of our experiments, with an increase in the air gap, the number of microdischarges in the series, that is, with each puncture of the discharge gap, also decreases.

The study of the optical picture of the discharge showed that the properties and thicknesses of the dielectric anodes and cathodes have a significant influence on the optical picture of the microdischarge channel. The brightness, geometrical dimensions and shapes of the discharge channels depend significantly on the dielectric permittivity ε_g of the dielectric anode. The dielectric permittivity (ε_g) of the polymers was modified by dispersing them with barium titanate $BaTiO_3$ or LZT-19 particles.

The formation mechanism of different microdischarge channels in the polymer-air gap-polymer dielectric structure was considered, and it was found that the microdischarge channels expand near the dielectric anode when the positive electrode is coated with dielectric. It is known that the following can cause the microdischarge channels to widen:

1. Expansion due to thermal diffusion of electrons in the cross-sectional direction of the microdischarge channels.

2. Expansion due to electrostatic repulsion of the electron cloud.

3. Expansion due to the field of charges sitting on dielectric surfaces.

After that, by comparing the experimentally found values of D_{mb} - for EOCH grams with the calculated values of the microdischarge diameters, it was concluded that the expansion of microdischarges near the dielectric anode cannot be determined by the thermal diffusion of electrons and repulsion of the electron cloud. The subsequent expansion of the microdischarge channels is associated with the distortion of the electric field of charges sitting on dielectric surfaces in the air gap.

Indeed, with the passage of the first current, electrons settle on the surface of the dielectric anode and form a point of charge of any density there. The electric field of charges is characterised by both normal accumulation, directed against the external field, and tangential accumulation, determined by the unequal density of charges located at the point along the dielectric surface. It is believed that the presence of tangential accumulation of the field helps to attract charges from the centre of the microdischarge channel to the edges, i.e. leads to a noticeable bending of the trajectory of electrons near the dielectric anode.

Although the effect of the expansion of microdischarge channels near the dielectric anode has been extensively described, the reasons for the appearance of dark patches in the overall structure of microdischarge channels near the dielectric cathode have not been sufficiently investigated.

We found that when the air gap and dielectric thicknesses are unchanged, the structure of the microdischarge channel mainly depends on the dielectric permittivity of the walls. The obtained optical images show that the dark areas near the cathode decrease significantly with increasing ε_g. From the above, it can be concluded that the decrease of the dark field near the cathode with increasing εg can be explained as follows. In the polymer - air gap - polymer system, the voltage takes such a value that ionisation processes start in the air gap, after which the starting current is initiated from the place where the field intensity is greatest - from the dielectric surface of the cathode. As a result of the development of the initial flood, charges accumulate on the dielectric anode and cathode surfaces bounding the air gap at any $\Delta S^{/}$ - surface of the dielectric. In the air gap where the flow occurs, the electric field (ΔE) is the superposition of two fields: external (E_p) and the field of charges sitting on the surface of the dielectric (E_{ot}), i.e.,

$$\Delta E = E_p + |-E_{ot}|$$

We can estimate the capacitance of a dielectric barrier with area $\Delta S^{/}$, dielectric permittivity ε_g and thickness d_g as follows:

$$\Delta C^{/} = \frac{\varepsilon_0 \varepsilon_g \Delta S^{/}}{d_g} \quad (1)$$

The $\Delta q^{/}$ - load on the considered $\Delta S^{/}$ area creates a voltage increase $\Delta U^{/}$ equal to $\Delta q^{/}/\Delta C^{/}$. At this time, the voltage in the air gap will drop very low. As can be seen, the degree of voltage drop in the air gap after the initial onset of flooding when d_g = const is mainly determined by the dielectric permittivity ε_g. During the formation of the microdischarge, the formation of a flood of electrons along the channel continues until the field of settled charges reduces the value of the external field, so that after this value the development of ionisation processes is not possible [4-6]. Therefore, at a certain density of electric charges transported by means of an initial current to the surface of the dielectric barriers, the voltage drop slows down with an increase in εg in the air gap, and hence the number of currents generated from the same point increases on the surface of the dielectric cathode, which leads to the complete disappearance of individual micro-discharges. The increase in the number of electron flows with increasing ε_g is probably the main reason for the large retention of the dielectric anode surface (figure 1 b,c).

Let us look at possible reasons for the disappearance of the dark field in the microdischarge channel near the dielectric cathode with increasing dielectric permittivity. This topic is also important for determining the criteria and understanding the mechanism of electrical breakdown of dielectrics under the influence of discharge. Based on the experimental results obtained, several hypotheses can be proposed.

1. If the air gap is bounded by different types of dielectrics, the degree of decrease of the main (external) area near the dielectric electrodes at a constant value of the settled charges differs. Near the dielectric-covered electrode, which itself has a large dielectric permittivity (when d_g = const), the decrease in the base area during each current is smaller. We consider that after a discharge, the surface of the dielectric walls is not equipotential and the field is differentiated in the air gap near the electrodes due to the variation of ε_g, γ_s and d_g in the dielectric. Therefore, the electrons initiating the overflow processes can only obtain sufficient energy to excite and ionise the atoms of the gas after travelling a certain distance from the dielectric cathode surface in the air gap. This minimum distance can be distinguished as a dark area in the optical appearance of the microdischarge channels (figure 1, a).

2. The number of electron streams generated along the channel of an individual microdischarge, which ultimately leads to the formation of a microdischarge, can greatly influence the dimensions

of the dark field. The number of electron flows along the microdischarge channel can be determined using simultaneous plotting of electrical and optical images of the evolution of the discharge in the air between the dielectric walls as follows:

a) First we determine the value of the Q - charge in the half-period (T/2) of the applied voltage according to the oscillogram of the discharge and determine the number of micro-discharges according to the optical picture n in the period T/2. Then, the value of the charge Δq carried by the individual microdischarges is determined as follows:

$$\Delta q = \frac{Q}{n} \tag{2}$$

Knowing the value of the Δq charge, the number of N_s currents along the microdischarge channel is determined:

$$N_s = \frac{\Delta q}{2\Pi_e e_e} \tag{3}$$

where e_e - is the charge of an electron,

Π_e - is the number of electrons in a separate current and is determined by the following expression:

$$\Pi_e = n_0 \cdot exp\left(\int_0^d d \cdot dx\right) = n_0 e d_{or} \cdot d \tag{4}$$

Where n_0 - is the effective electron number,

α - is the Townsend coefficient and

d - is the thickness of the air gap in the polymer-air gap-polymer system.

In our experimental conditions, the value of α or is determined from the dependence graph $\alpha/p = f(E/p)$. As mentioned above, with the development of microdischarge, the charges sitting on the dielectric anode and cathode reduce the value of the electric field in the air gap during the formation time, and therefore the value of the pulse ionisation coefficient becomes an unstable quantity. Therefore, for a given value of α, E_{or} in the dependence $\alpha/p = f(E/p)$ should be assumed as follows:

$$E_{or} = \frac{U_d + U_b}{2d} \tag{5}$$

where U_d - is the breakdown voltage of the gap in the polymer-air gap-polymer system,

U_b - is the breakdown voltage of the individual microdischarge determined from the following expression:

$$\Delta q = \frac{\Delta s}{s}\left(C_g + C_b\right)\left(U_d + U_{sön}\right) \tag{6}$$

Here C_g and C_b are the dielectric and air gap capacitances, respectively.

b) The number of electron flows during the formation of local discharges is expressed as follows:

$$N_s = \frac{t_{mb}}{V_e^{-1} d} \tag{7}$$

where V_e is the electron drift velocity. The value of V_e is determined from the experimental graph $V_e = f(E_{or}/P)$ and t_{mb} is determined from the discharge voltage pulse oscillogram.

The intensity of the electric field is determined by the electric strength of the gas gap between the dielectric barriers, which is a constant quantity when the thickness of the gap d remains constant

and does not depend on the amplitude of the voltage applied to the polymer gas gap-polymer dielectric structure. The amplitude of the applied voltage mainly determines the number of discharges in the sinusoidal voltage period (figure 2).

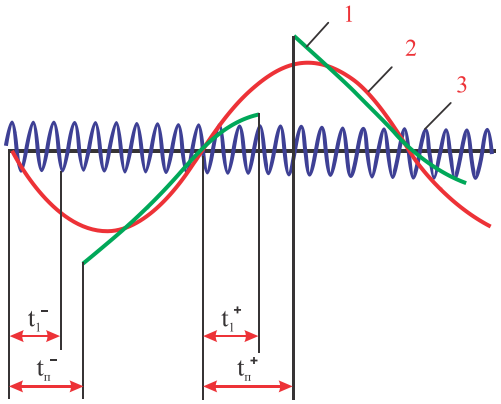


Figure 2: *Oscillograms of applied voltage and discharge voltage pulses:*
1 - discharge voltage pulses in the sinusoidal voltage period;
2 - variation of the applied voltage;
3 - calibration voltage (T = 1000 μs)

II. Conclusion

Thus, the application of EOC and simultaneously obtained electrical (figure 2) and optical (figure 1,a) images are used to reveal the microstructure of the electrical discharge in the dielectric - gas gap - dielectric system, their causes. The electrical corrosion of polymers and the formation of new electret, piezo and pyroelectric properties in them allow their proper utilisation.

References

[1] Aliyev H.S., Guliyev M.M., Ismayilova R.S. Relaxation Phenomena in Polyvinylchloride/graphite Composites. 18th IFAC Conference on Technology, Culture and International Stability TECIS 2018. IFAC Papers Online 51-30 (2018), 825-827 p.

[2] Aliyev H.S., Quliyev M.M. Effect of Barrier Discharge on Stability of Electret State of high – Density Polyethylene // Surface Engineering and Applied Electrochemistry, 2016, 52 (5), p. 494-498.

[3] Choudhury M., Mohanty S., Nayak K. Preparation and characterization of electrically and thermally conductive polymeric nanocomposites, Journal of Minerals and Materials, Characterization and Engineering. 11(7) (2012) p. 744-756.

[4] Kerimov, E. A. Study of the conductivity of carbon nanotubes deposited on an iridium-silicon silicide substrate. Russian Microelec-tronics, 2024, Pleiades Publishing, Ltd., 2024.

Vol. 53, No. 6, pp. 576-581.

[5] Jakubinek M.B., White A.M., M.M and Winey K.I. Temperature dependence of thermal conductivity enhancement in single-walled carbon nanotube/polystyrene composites. Applied Physics Lettres. 2010, v. 96, 083105, 3 p.

[6] Karami P. Polymer/nanodiamond composites - a comprehensive review from synthesis and fabrication to properties and applications // Adv. Colloid Interface Sci. Elsevier B.V., 2019, Vol. 269. p. 122-151.

METHOD FOR INCREASING THE ACCURACY OF INFORMATION EXCHANGE IN COMMUNICATION NETWORKS

Gunay Khasmammadova[1], Sevinj Bakhshaliyeva[2], Javanshir Zeynalov[3]

[1]Azerbaijan Technical University, Baku, Azerbaijan
[2]Azerbaijan State Oil and Industry University, Baku, Azerbaijan
[3]Nakhchivan State University, Nakhchivan, Azerbaijan
gunay.xasmemmedova@aztu.edu.az, sevinj.quliyeba@asoiu.edu.az, cavansirzeynaov@ndu.edu.az

Abstract

The presented article is devoted to the issue reducing losses in the information exchange channel. Satisfactory results have been obtained from the conducted research: calculations show that the implementation of the evolution of wireless communication network control center increases the performance of the network. It is based on G technology and its generations. Research shows that the LTE (4G) architecture is the way to improve mobile devices and data based on GSM/EDGE and UMTS/HSPA standards in the field of telecommunications. Ensuring the accuracy of information in wireless networks was thoroughly investigated, and for this, corrective filtering was proposed for receiving data in the processing process, and a positive result was obtained without filtering. Modeling of LTE was carried out in the OPNET modeling software package. During the research, simulation modeling was carried out in the Matlab environment, and satisfactory results were obtained.

Keywords: Information exchange, communication network, exchange channel, wireless communication, LTE technology.

I. Introduction

When considering the term wireless, it is clear that this term can be given different definitions. That is, the concept of wireless communication (or simply wireless when the context allows) is the transfer of information between two or more points that do not use an electrical conductor as a means of transmission [1, 2].

The most common and widely used wireless technologies use radio waves. The distances covered by radio waves can be short or long, for example a few meters for Bluetooth or up to millions of kilometers for deep space radio communication. It covers a wide variety of fixed, mobile and portable applications, including two-way radios, mobile phones, personal digital assistants (PDAs) and wireless networking [3, 4, 5]. Other examples of applications of radio wireless technology include GPS devices, garage door openers, wireless computer mice, keyboards and headsets, headphones, radio receivers, radio transmitters as well as satellite television, broadcast television and cordless telephones, etc. includes Somewhat less common methods of achieving wireless communication include the use of light, magnetic or electric fields, or other electromagnetic wireless technologies such as the use of sound.

The term wireless has been used twice in the history of communication, with slightly different meanings. It was first used from 1890 for early radio transmitter and receiver technology, such as wireless telegraphy, until around 1920 when the new word radio replaced it [6]. In countries around the world, non-portable radios in the UK continued to be called wireless sets into the 1960s. The term was revived in the 1980s and 1990s, mainly to distinguish digital devices that communicate wirelessly from those that require wires or cables, such as the examples listed in the previous sections. This became its primary use in the 2000s with the advent of technologies such as mobile broadband, Wi-Fi, and Bluetooth, allowing it to become more widespread [6, 7].

Wireless operations allow services such as cellular and interplanetary communications, i.e. global communications, that are impossible to implement with the use of wires [7, 8, 9]. The term is generally used in the telecommunications industry to refer to telecommunications systems (such as radio transmitters and receivers, remote controls, etc.) that use some form of energy (such as radio waves and acoustic energy) to transmit data without using wires. Information is transmitted in this way over both short and long distances.

The first wireless telephone conversation took place in 1880 when Alexander Graham Bell and Charles Sumner Tainter invented the photophone, which sent sound over a beam of light. Sunlight and a clear line of sight between the transmitter and receiver were required for that process to work. These factors greatly reduce the viability of the photophone in any practical experiment. It took several decades before the principles of the photophone found their first practical application in military communications and later in fiber-optic communications [9, 10, 11].

A number of wireless electrical signaling schemes, including sending electric currents through water and earth using the laws of electrostatics and electromagnetic induction, were explored for the telegraph in the late 19th century before practical radio systems became available.

These include Thomas Edison's patented induction system, William Preece's induction telegraph system for sending messages between water bodies, and several operational and proposed telegraphs, etc. includes In 1894, Guglielmo Marconi began to develop a wireless telegraph system using radio waves, which had been known since their existence in 1888 by Heinrich Hertz.

Soon Marconi developed a system that transmitted signals over a great distance that anyone could guess [12, 13, 14]. Marconi and Carl Ferdinand Braun were awarded the Nobel Prize in Physics in 1909 for their contributions to this form of wireless telegraphy. Millimeter wave communication was first studied by J. Chandra Bose in 1894-1896. He achieved a very high frequency of up to 60 GHz in his experiments. He also proposed the use of semiconductor compounds to detect radio waves when he patented the radio crystal detector in 1901 [1, 15, 16].

II. Connection of LTE technology with previous generation technologies and distinctive features

Nowadays, communication technologies are extremely important. We are surrounded by many telecommunication technologies, especially mobile phones, internet, satellite, wireless communication. 4G service provides faster and better transmission of data, internet and call services. For example, through this service, subscribers can open web pages in the blink of an eye, instantly upload photos to social networks, and watch HD movies online without interruption.

Data transfer speed in mobile networks reaches 326.4 Mbit/s through LTE technology, which belongs to the fourth generation technologies [15, 16, 17]. This allows providing mobile television, mobile finance, education, healthcare and other social services easily and with high quality. This technology also creates great opportunities for users operating in the corporate sector. So, they can carry out tasks requiring the exchange large volumes of files anywhere in the world through their mobile devices and participate in video conferences.

On average, a 5-minute song is downloaded in 100 seconds from mobile networks, 10 seconds with HSPA, and 0.38 seconds with 4G technology. In addition, a 1-hour movie in HD quality can be downloaded to our computer in 3 days from mobile networks, 8 hours with HSPA, and 17 minutes with 4G. These facts clearly demonstrate with visual evidence how much more powerful technology 4G is than its predecessors. In general, the connection speed in this technology is 100 mps in mobile phones and 1 Gbps in Wi-Fi networks [3, 17, 18]. At the same time, it is the same size as Wimax bandwidth. This is 10 times faster than 3G technology.

In general, more than half 4G connections worldwide are in North America (thanks to large-scale LTE networks established in the USA), 39% in developed Asia-Pacific countries (the largest networks are established in Japan and South Korea). If we look at the world trends, we will see that despite the active development of the fourth generation networks, international organizations, equipment manufacturers, operators and research centers are already working on the creation of the next telecommunication networks.

However, according to the research of "J'son & Partners Consulting" company, the golden age of using LTE networks was observed in 2012. Since that year, the 4G network has been available to local subscribers. Azercell Telekom was the first company to offer 4G services in Azerbaijan. Thus, in 2012, the company introduced the LTE service, which is the 4th generation network. With this, "Azercell" became the first company to introduce 4G technology for commercial use to customers not only in Azerbaijan, but also in the entire region [5, 8, 9].

Azercell Telekom was the first company to offer 4G services in Azerbaijan. Thus, in 2012, the company introduced the LTE service, which is the 4th generation network. With this, "Azercell" became the first company to introduce 4G technology for commercial use to customers not only in Azerbaijan, but also in the entire region. According to the results of the "GlobalWirelessSolutions" company, which conducted a survey of the mobile network quality at the invitation of the Ministry of Communications and High Technologies, "Azercell" network showed the highest indicators among the country's mobile operators according to many parameters. Also, according to the results of international systems such as "Opensignal" and "Testmy.net", which measure the network quality of radio signals, today the best quality and reliable mobile network and the fastest Mobile Internet signal in Azerbaijan are provided by "Azercell".

Ericsson company equipment is used to provide the network. "Azercell" is intensively developing the 4G network, taking into account the demand for fast mobile Internet of its subscribers [6, 9, 10]. During the first 5 months of 2016 alone, the company installed 177 new LTE stations in Baku and the Absheron Peninsula, and by the end of June, the number of stations increased to 220. Currently, the number of 4G stations in the mentioned area has reached 480.

With this, the level of 4G coverage of Baku and Absheron, which has the highest demand for high-speed mobile internet, reached 79%, and subscribers were able to use high-speed internet up to 100 Mb.s in the 4G network.

Good results on 4G coverage have already been achieved in Baku and Absheron Peninsula. The next target is other big cities of Azerbaijan. Already in 2017, the company aims to achieve 4G coverage in the regions [11, 14]. During this year alone, the investment made by "Azercell" in the 4G network amounted to 8 million manats.

III. IP Multimedia Subsystem VoLTE

The IP Multimedia Subsystem or IP Multimedia Core Network Subsystem (IMS) is a standardized architectural framework for delivering IP multimedia services. Historically, mobile phones have provided voice calling services over a circuit-switched network rather than strictly over an IP packet-switched network.

Alternative methods of delivering voice (VoIP) or other multimedia services have become available on smartphones, but they have not been standardized across the industry. IMS is an architectural framework that enables such standardization. IMS was originally designed by the wireless standards body the 3rd Generation Partnership Project (3GPP) as part of a vision for mobile networks evolving beyond GSM.

Its original formulation (3GPP Rel-5) was an approach for delivering Internet services over GPRS. This vision was later updated by 3GPP, 3GPP2 and ETSI TISPAN to require support for networks other than GPRS such as Wireless LAN, CDMA2000 and fixed lines [1, 2].

IMS uses IETF protocols whenever possible, such as Session Initiation Protocol (SIP). According to 3GPP, IMS is not designed to standardize applications, but to facilitate access to multimedia and voice applications from wireless and wired terminals, i.e. to create a form of fixed mobile convergence (FMC). This is done by having a horizontal management layer that isolates the access network from the service layer.

From a logical architecture perspective, services do not need to have their own management functions because the management layer is a common horizontal layer. However, this does not necessarily mean more reduced cost and complexity in implementation. Alternative and overlapping technologies for accessing and providing services over wired and wireless networks include Public Access Network, soft switches, and "bare-open" SIP connections.

IMS is widely adopted as it becomes increasingly easy to access content and contacts using mechanisms beyond the control of traditional wireless/fixed carriers.

Examples of IMS-based global standards include Voice over LTE (VoLTE), Wi-Fi Calling (VoWIFI), Video over LTE (ViLTE), SMS/MMS over WiFi and LTE, USSD over LTE, and MMTel, which is the basis for Rich. Communication Services (RCS), also known as Joyn or Advanced Messaging, RCS is a carrier application. RCS also performs Presence/EAB (enhanced address book) functionality.

IV. LTE network model in the OPNET modeling software package

OPNET Network simulator is a tool for simulating the behavior and performance of any type of network. The main difference between Opnet Network Simulator and other simulators is its power and versatility.

IT Guru provides pre-built models of protocols and devices. It allows you to create and simulate different network topologies. General Purpose Modeling (GPM) languages (eg UML, Southbeach Notation, IDEF) or Domain Specific Modeling (DSM) languages (eg SysML). Visual modeling in computer science had no standards before the 90s.

These include proprietary standards such as languages related to industry open standards (eg UML, SysML, Modelica), VisSim, MATLAB and Simulink, OPNET, NetSim, NI Multisim, and Reactive Blocks. Both VisSim and Reactive Blocks provide a downloadable viewer that allows anyone to simulate their models openly and interactively. Community publishing of Reactive Blocks also allows for full editing and building of models as long as the work is published under it. Eclipse Public License.

Visual modeling languages are an active area research that continues to develop, as evidenced by the growing interest in DSM languages. In the OPNET modeling software package, the LTE network model is three-tiered.

Three-tier OPNET hierarchy [2, 8, 11, 15]:

- Three domains: network, node and process
- The Node model defines an object in the network domain
- A process model specifies an object in the node domain.

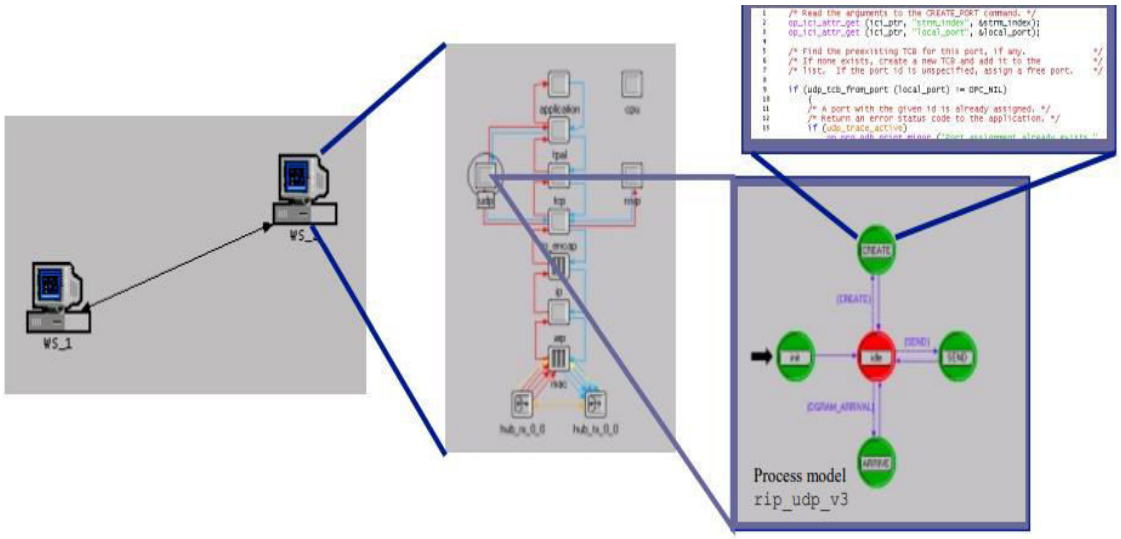


Figure 1: *OPNET model hierarchy*

V. Corrective filtering to increase accuracy

Filtering generally means separating the useful signal from other environmental noises that affect it. Since the measured signal is a tuned signal, solving this problem involves certain problems. The device that implements the filtering algorithm is called a "filter". Any node or processing operator operating during measurement information exchange in measurement systems should have any filtering property.

Selective filtering means the method of processing the useful signal and the error signal (noise) together with the aim of estimating the informative (current value) parameter of the measurement signal with the highest possible accuracy. Selective filtering is an indication of the level of accuracy that measurement systems can provide. Accuracy is the most objective efficiency indicator of any measuring tool [1, 16, 18]. Measurement information correction methods and tools are divided into 2 main classes, depending on the nature of the measurement signals processed in the measurement-processing exchange and the nature of the information exchange processes performed, that is, whether they are analog or discrete:

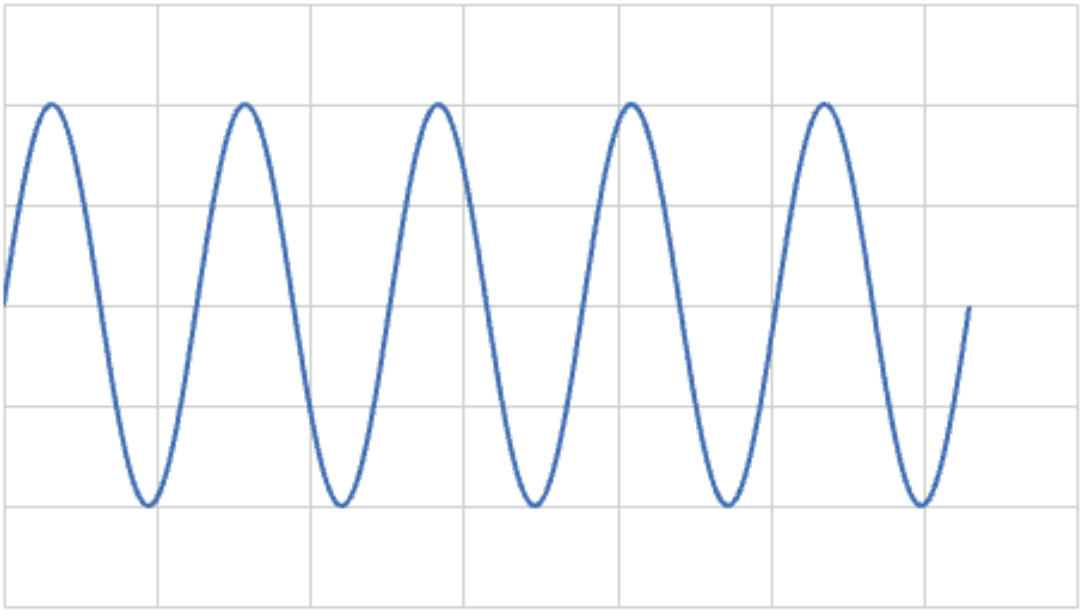
1) structure; 2) algorithmic.

Structural methods and tools serve to increase the accuracy and stability of measuring tools by including additional elements and blocks in the structure of the created functional node. The algorithmic method is implemented with software tools, their implementation is oriented towards classical and modern filtering algorithms. The most widely used of these modern filtering algorithms is the digital filtering algorithm.

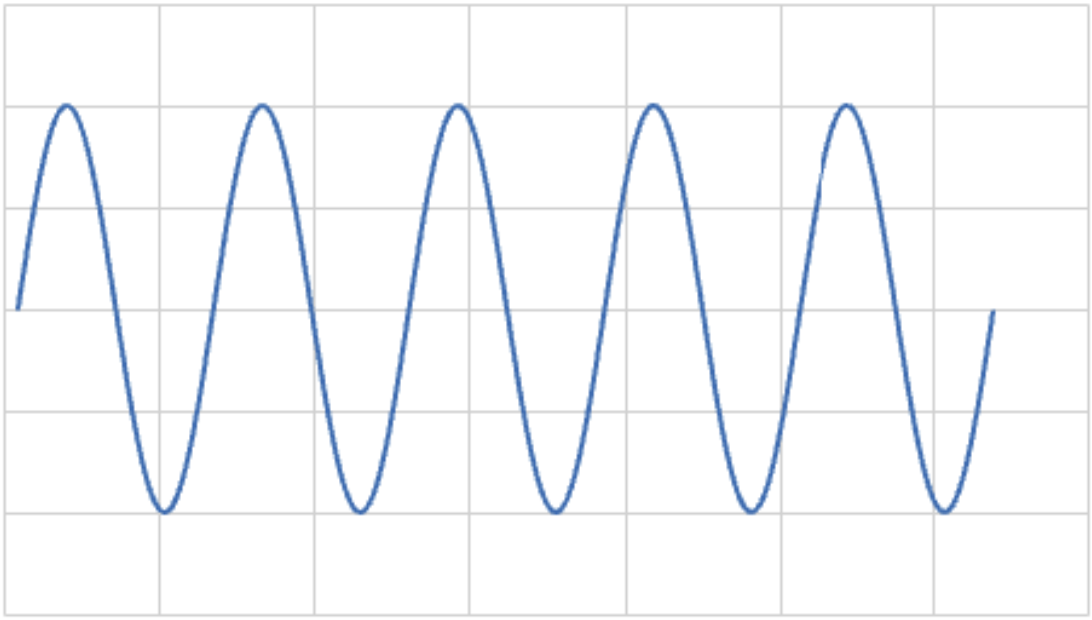
VI. Results and discussion

Many filters are used nowadays. To increase the measurement accuracy, the most widely used types of correction filters can be listed - Butterdort, Bessel, Chebyshev, Raised-Cosine, Cauer. There are a number of parameters that distinguish these filters from each other: system function, frequency transfer coefficient, dispersion and damping coefficient. The structural schemes and parameters of each can be determined using the MATLAB software package [7, 10]. Based on the used input signals, it is possible to determine the attenuation coefficient of the filters and compare the values

found to choose the most suitable filter. After completing the process, the most appropriate filter was selected based on the extinction coefficient, Chebyshev filter. Figure 2 shows the model of the Chebyshev filter created in the Matlab program.



a)



b)

Figure 2: *a) filter input, b) output signals*

VII. Conclusion

1. The presented article is devoted to the issue reducing losses in the information exchange channel. Satisfactory results have been obtained from the conducted research: calculations show that the implementation of the evolution wireless communication network control center increases the performance of the network.

2. It is based on G technology and its generations. Research shows that the LTE architecture is the way to improve mobile devices and data based on GSM/EDGE and UMTS/HSPA standards in the field of telecommunications.

3. Ensuring the accuracy of information in wireless networks was thoroughly investigated, and for this, corrective filtering was proposed for receiving data in the processing process, and a positive result was obtained without filtering. Modeling of LTE was carried out in the OPNET modeling software package.

References

[1] Shakhnovich I. Modern technologies of wireless communication.Technosphere.M.2006.287 p.

[2] Ibrahimov B.G. Research and analysis of the efficiency of multiservice communication networks using the NGN architectural concept / B.G. Ibragimov, S.R. Ismaylova // T- Comm, Telecommunications and transport, - Moscow: - 2014. Vol. 8, No. 8, - pp. 47 - 50.

[3] Ibrahimov B.G., Ismaylova S.R. On one approach to assessing the quality of functioning of a signaling network link // All-Russian Scientific and Technical Conference "Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems", section - "Theory of Teletraffic", - Moscow: RUDN, - 2012, - pp. 38 - 40.

[4] Ibrahimov B.G. Research of the efficiency of hybrid communication networks using signaling protocols / B.G. Ibrahimov, S.R. Ismaylova, F.I. Huseynov // Bulletin of Computer and Information Technologies. No.11, 2013. pp. 50 - 54.

[5] Ibrahimov B.G., Mehdiyeva A.M., Bakhtiyarov I.N. Study of throughput indicators of corporate multiservice networks // Bulletin of Computer and Information Technologies, No. 5, 2020. pp. 38 - 44.

[6] Ibrahimov B.G. Efficiency of the system and signaling in multiservice telecommunication networks / B.G. Ibrahimov, Sh.M. Mamedov. Baku: "Elm". 2015. 220 p.

[7] Ibrahimov B.G., Mehdiyeva A.M., Bakhtiyarov I.N. Mathematical model for assessing the level noise immunity of the paths systems for transmitting, processing and receiving packet messages // Proceedings of the XII International Scientific-Practical Conference. New Informational and Computer Technologies in Education and Science - IES-2020. Ukraine, Vinnytsia: 2020, p.77–79.

[8] Tanenbaum E., Computer networks, Peter, 2003, pp. 361-370.

[9] Falconer, D. and Ariyavisitakul, S. L., Frequency Domain Equalization for 2.11 GHz Fixed Broadband Wireless systems, Tutorial, presented during Session #11 of IEEE 802.16 in Ottawa, Canada, Jan. 22, 2001.

[10] Sheluhin O.I. Modeling of information systems. Moscow: Hotline - Telekom, 2018. 516 P.

[11] Andreev R.N. Theory of electrical communication. Textbook for students / R.N. Andreev, R.P. Krasnov, M.Yu. Chepelev Hotline Telekom, 2014. 230 p.

[12] Michael P.F. Fundamentals of Communications Systems. Communications Engineering. New York: McGraw-Hill Companies, 2007. 436 p.

[13] Bitner.V.I. Networks of the new generation-NGN. / V.I. Bitner, Ts.Ts. Mikhailova - Moscow: Hotline-Telecom. 2011. 228 p.

[14] Vasiliev K.K. Mathematical modeling of information communication systems. Moscow: Hotline Telekom. 2018. 236 p.

[15] Velichko V.V. Models and methods of increasing the durability of modern communication systems. /V.V. Velichko, G.V. Popkov, V.K. Popkov. Moscow: Hotline-Telecom.2016.270p.

[16] Mehdiyeva A.M., Zeynalova, R.R., Safarova, A.A., Takhumova O.V., Nikolaevc P.P., Mozgovoy A.I. Development of an adaptive control system for the quality parameter in the lack of information. Proceedings of SPIE - The International Society for Optical Engineering, 12637, 1263707. doi: 10.1117/12.2681371, 2023, Fergana, Uzbekistan.

[17] Mehdiyeva A.M., Bakhtiyarov I.N., Bakhshaliyeva S.V. Increasing the Immunity of Information Transmission and Fault Tolerance of the Path. Lecture Notes on Data Engineering and Communications Technologies. Volume 166. Mobile Computing and Sustainable Informatics. Proceedings of ICMCSI 2023, 11-12 January 2023. Tribhuvan University, Nepal. pp. 775 784. http://icmcsi.com/2023.

[18] Mehdiyeva A.M., Sardarova I.Z., Mahmudova Z.A. Development of an Information Accuracy Control System. Lecture Notes on Data Engineering and Communications Technologies. Volume 166. Mobile Computing and Sustainable Informatics. Proceedings of ICMCSI 2023, 11-12 January 2023. Tribhuvan University, Nepal. pp. 173 179. http://icmcsi.com/2023.

A NEW APPROACH TO NUMERICAL CALCULATION OF NON-STATIONARY PROCESSES IN COMPLEX MAIN GAS PIPELINES

Cherkez Yusubov[1], Habib Abbasov[1], Ankur Pan Saikia[2]

•

[1]Azerbaijan Technical University, Azerbaijan, Baku
[2]Assam down town University, Assam, India
cerkez.yusubov@aztu.edu.az, hebib.abbasov@baau.edu.az, ankur.saikia@adtu.in

Abstract

The indicators of the main gas pipeline complexes are analyzed and studied, and a new approach to the numerical calculation of non-static processes in the system is proposed. Issues of unsteady movement liquid and gas in pipes are significant importance for both the design and operation pipelines. Based on the new approach, a method for calculating the characteristics of non-static processes in main gas pipeline complexes has been constructed. Based on the calculation methods, some important analytical expressions for assessing the performance of main gas pipeline complexes have been obtained.

Keywords: Complex magistral gazoprovodov, gas industry, design, non-static processes, operation pipelines.

I. Introduction

Currently, the gas industry is one of the leading branches of Azerbaijan's economy. Gas supply data show that trunk gas pipelines operate in a non-stationary pumping mode for a significant part of the time [1].

In this regard, in conditions of non-stationarity during operation of trunk gas pipelines, it is often necessary to deal with various malfunctions of their technical condition, leading to a violation of the operating mode, contributing to the creation of pre-emergency and emergency situations in the system and, thereby, disrupting the rhythm of the pipeline operation [2, 3].

Considering that in many regions of Azerbaijan, gas is the main type of fuel and chemical raw material, guaranteed uninterrupted supply of gas during that time has important importance.

Accidents occurring on main gas pipelines lead to significant losses caused by gas leakage and stoppage of pumping, which in turn leads to stoppage processes, disruption of gas supplies to consumers, contamination of the surrounding area and the threat of death of plants and animals.

Therefore, it is of great practical importance to study the gas-dynamic condition of main gas pipelines under various conditions of violation of their working modes, with the aim of reducing gas loss and choosing a rational control method under technological and emergency modes.

Thus, the creation of methods of calculation of non-stationary modes of operation of main gas pipelines will allow more effective implementation of their design and operation [4, 5].

However, the solution to the problem developing sufficiently accurate mathematical relationships and algorithms describing non-stationary processes in gas transport systems, taking

into account real factors such as the influence of inertia and friction of the pumped gas, the distribution of parameters of the outlet sections, the influence of the characteristics of compressor stations, shut-off valves is still far from being implemented [6, 7].

In most of the studies, they were mainly carried out only for a particular case - the solution of the equations describing the non-stationary movement of gas along one section of the gas transport network, without considering the entire gas pipeline network, including compressor stations, taps, shut-off valves, gas consumers, in a single gas-dynamic mode, which Essentially narrows the circle of practical tasks [8, 9].

In this article, a simple universal numerical method is proposed, which allows to calculate non-stationary processes in complex main gas pipeline systems with a sufficiently high accuracy, taking into account the influence of inertial forces, friction, as well as the real characteristics of compressor stations, shut-off valves, taps when they are turned on and off, and arbitrary moments of time.

The proposed method is based on the use of a new approach in determining the values of the originals of the transfer functions by using a discrete analogue of the integral convolution equation, which allows you to move to the domain of discrete images of the desired functions.

In the domain of originals without finding the roots of the characteristic equation, while replacing the continuous integration operations by summation by the trapezoid formula, which significantly increases the accuracy of calculations, simplifies mathematical tabs and significantly expands the range of practical problems to be solved [1, 3, 6].

Examples of reasons causing non-stationary operating modes of gas pipeline systems may be:

- irregularity of gas reception from the field and gas supply from gas processing plants;
- stopping and starting compressor units, which may be planned or emergency due to failure of linear equipment, power outage, or activation of the automatic protection system;
- putting a gas pipeline into operation after construction, eliminating emergency situations or long-term shutdowns;
- turning on and off track discharges and gas injections along the pipeline route;
- changing the state of underground gas storage facilities:
- violation of the gas pipeline tightness, occurrence of leaks;
- stopping, switching from one mode of the gas pipeline system to another, which can be planned or emergency due to failure of equipment or a linear section.

In addition, the magnitude of the non-stationarity of the pumping process is influenced by a number of other factors: the process of filling and emptying the gas pipeline during pipe testing, the blowdown process, pipe ruptures and contamination with condensate, hydrate formation, corrosion processes, and finally, changes in the mode introduced by service personnel in the process of managing the long-distance gas transportation system.

In many cases, the occurrence of these reasons has a negative impact on the operation of the gas pipeline and creates the risk of emergency situations.

Nonstationary processes occurring in the main gas pipeline are described by partial differential equations of the parabolic type:

$$\begin{aligned} -\frac{\partial p}{\partial x} &= k_1 \omega \\ -\frac{\partial \omega}{\partial x} &= k_2 \frac{\partial p}{\partial t}, \end{aligned} \quad , 0 \le x \le l \qquad (1)$$

where $p = p(\chi, t)$, $\omega = \omega(\chi, t)$ – respectively, excess values pressure and speed of gas movement; $k_1 = 2a\rho$, $k_2 = 1/\rho c^2$, $\rho -$ gas density; c - speed of sound in gas; $2\alpha = \lambda \omega_{cp} / 2D$ – coefficient linearized according to I.A.

Charny; λ - coefficient of hydraulic resistance; D – internal diameter of the pipe; ω_{cp} - average gas flow velocity over the cross section in steady state [1]:

$$\omega_{\text{ъп}} = \frac{2}{3}\left(\frac{\omega_1^2 + \omega_0\omega_1 - 2\omega_0^2}{\omega_1 - \omega_0}\right) \tag{2}$$

where ω_0, ω_1 - steady-state values of average speed.

For this problem, the initial conditions are taken to be zero:

$$p(x,t)_{t=0} = 0, \quad \omega(x,t)_{t=0} = 0 \tag{3}$$

Boundary conditions may have different forms depending on the operating modes of the gas pipeline.

In the case under consideration, let us assume that the boundary conditions are as follows:

$$\omega(x,t)_{x=0} = \omega_H(t), \quad \mathrm{p}(x,t)_{x=1} = 0 \tag{4}$$

where $\omega_H(t)$ - arbitrary law of change in the speed of gas movement at the beginning of the pipe.

So, the gas pipeline length $l = 100$ km, the internal diameter of the gas pipeline $(D) = 1$ m, the coefficient hydraulic resistance $(\lambda_0) = 0{,}01$, the average gas pumping speed $(\omega_{cp}) = 10$ m/s operates in a steady state.

The case was considered when, at the beginning of the pipe at the moment of time $t \geq 0$, a pressure jump $\mathrm{P}(0,t) = \mathrm{P}_H(t) = 10^{-1} M\Pi a,$ occurs, and at the end of the pipeline a constant flow rate is maintained, i.e. $G(l,t) = G_K(t) = 0$ (Fig. 1) - since in this example $\mathrm{P} = \mathrm{P}(x,t)$, $G = G(x,t)$ - implies disturbances of pressure and flow rate above their stationary values.

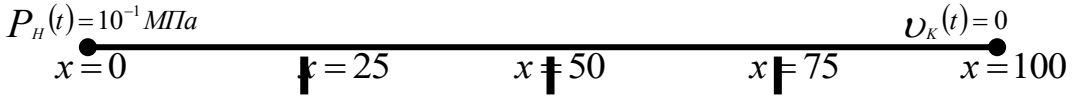


Figure 1: *Structure and description of the pipeline maintaining a constant flow rate*

The speed of sound in gas is 320 m/s.

The value of the reduced coefficient of linear friction is determined by the formula:

$$2\grave{a} = \frac{\lambda_0 \omega_{\tilde{n}\eth}}{2D} = \frac{0{,}01 \cdot 10}{2 \cdot 1} = 0{,}05 \ \ 1/c$$

According to the conditions of the problem, the initial conditions are zero.

Calculations were carried out on a computer at $\lambda = 10$.

As experimental studies show, when compressor units are turned on or off, or valves at the beginning of a gas pipeline are opened or closed, the pressure and speed of gas movement at any point in the pipe change according to an arbitrary law.

In connection with the widespread introduction of computer technology into the practice of engineering calculations, the use of numerical methods for calculating non-stationary processes in main gas pipelines is currently becoming especially effective.

In this case, a numerical method is given for calculating non-stationary modes in main gas pipelines taking into account the friction forces and inertia of gas, with an arbitrary law of change in the speed of gas movement at the beginning and pressure at the end of the pipeline (Figure 1).

The essence of the proposed method is based on the use of a discrete analogue of the integral convolution equation.

The advantage of the proposed approach is that it allows, without going into the area of discrete images, to make the transition from Laplace images of the sought functions to the area originals, without finding the roots of the characteristic equation, to replace the operations continuous integration with summation using the trapezoid method, which significantly simplifies mathematical calculations and increases the accuracy of calculations.

If non-stationary processes occur under the influence of a change in the gas velocity at the beginning of the gas pipeline, then as a result of this the pumping speed and gas pressure at any point of the gas pipeline also change.

This type of boundary conditions connects the gas velocity at the beginning of the gas pipeline in the time domain with the pressure at the end, which is most often used in the case of coordinating the productivity of gas pipeline sections.

The solution to the problem under consideration allows one to determine, for a given change in the velocity at the beginning and pressure at the end of the gas pipeline, the pressure and gas velocity in any section of the pipe.

From the point of view of operational management, it is of interest to obtain a dependence by which it would be possible to quantitatively assess the effect of a change in productivity at the beginning of a linear section on the pressure at its end.

In this case, the relationship between continuous time t and discrete time is as follows:

$$t = \frac{nT}{\lambda}$$

where $T = 2\tau$, $\tau = \frac{l}{c}$ - is the time of wave propagation to one end of the gas pipeline.

The results of calculations for this case, regarding the change in pressure in sections $x = 25км$, $x = 50км$, $x = 75км$ are shown in Figure 2 (curves $1 \div 3$), and regarding the change in flow rate – in figure 3 (curves $1' \div 3'$).

As can be seen from figure 2-3, when moving from one section to another during non-stationary processes, a delay and different rate of increase of the non-stationary process along the length of the main gas pipeline are clearly visible.

For example, as can be seen from Figure 2 (4.4) at t=6h the level of pressure increase in sections $x = 25км$, $x = 50км$, $x = 75км$ respectively is when, $\mathrm{P}(25,t) = 8{,}7МПа$, $\mathrm{P}(50,t) = 7{,}48МПа$, $\mathrm{P}(75,t) = 6{,}25МПа$.

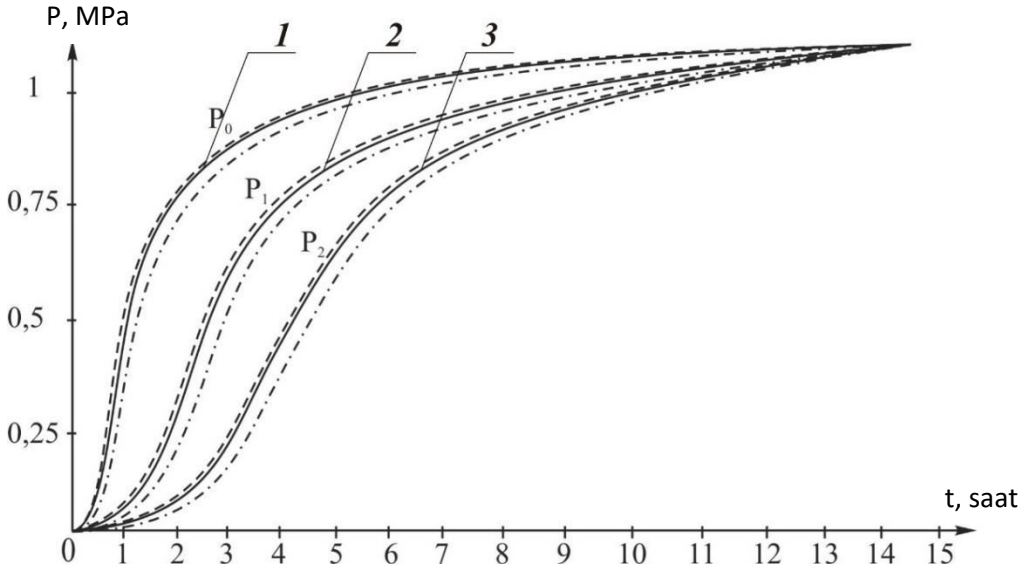


Figure 2: *Graphic dependence of pressure change in sections on time*

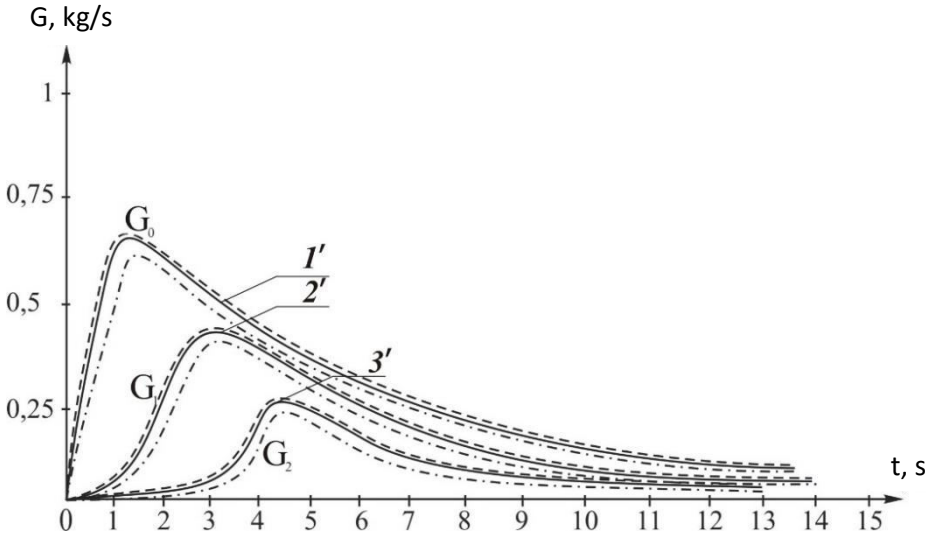


Figure 3: *Graphic dependence of non-stationary processes on time in a main gas pipeline*

In the case under consideration, when moving from one section to another, the pressure has a different rate of decrease, and the flow rate has a different rate of increase during non-stationary processes [8, 9].

From the analysis of the results shown in Figure 3, it follows that there is a different effect of delay of the non-stationary process relative to the pressure and flow rate from one section to another.

In addition, from the analysis of the obtained results (Figure 2-3) it follows that the nature of the change in pressure and flow rate during non-stationary processes occurring in main gas pipelines depends on the choice of the type of boundary conditions.

For comparison, Figure 2-3 also shows the results of this example obtained using the state variable method [4, 6] in the form of dotted lines. As can be seen from Figure 2-3, the calculation results obtained using both methods are completely identical. However, the method developed in this article significantly simplifies the mathematical calculations.

Since, in this case, to solve the problem, according to the method of state variables [4, 6]; the main gas pipeline is divided into four sections and for each section, in order to ensure the specified

calculation accuracy [4, 6], six approximated ordinary differential equations of the first order are written.

Further, with the joint solutions of the indicated equations, both for pressure and flow rate, non-stationary processes occurring in the main gas pipeline are determined, which makes the mathematical calculations extremely complicated.

Thus, a general disadvantage of the state variable method is the great difficulty of determining the number of approximated ordinary differential equations of the first order to ensure a given accuracy of calculation of a non-stationary process depending on the length of the linear sections of the pipeline [4, 6, 8].

In Figure 2-3, in the form of dotted lines, the results of this example are also shown, obtained according to the numerical method based on the theory of pulse systems and the mathematical apparatus of the discrete Laplace transform [5, 7] (with this approach, the operation of continuous integration is replaced by summation according to the rectangle formula).

In this case, the error of the numerical method [5, 7] in comparison with the methodology developed in this work, for this case with respect to pressure is 6%, and with respect to flow rate – 8%. The proposed numerical method can be widely used in the design and operation of complex main gas pipelines.

Calculations according to the developed algorithms are reduced to simple recurrence relations that allow easy implementation on a computer.

Calculations based on the developed algorithms are reduced to simple recurrence relations that can be easily implemented on a computer.

The proposed method can be widely used by research and design organizations in the design and operation of main gas pipelines. The developed algorithms make it possible to solve a wide range of problems on non-stationary processes in complex gas transportation systems and allow:

- determine dangerous increases and decreases in pressure at the discharge and suction compressor stations; determine dangerous excess pressure at the valve, as well as at characteristic points of linear sections and branch sections, when switching on and off gas pumping units of compressor stations, valves, associated gas consumers, in order to prevent emergency phenomena;
- develop a set of measures for operational dispatch control of non-stationary operating modes complex systems main gas pipelines, such as determining the sequence of switching compressor units of compressor stations; determining the sequence of switching on branches when distributing gas supply volumes between associated consumers.

The practical use of the developed methods allows to increase the efficiency of dispatch services and the reliability operation main gas pipelines, leads to an increase in the throughput capacity of the gas pipeline, and a decrease in gas losses, as well as the costs electricity consumed by compressor stations.

II. Conclusions

1. As a result, the study proposed a new approach to creating methods for calculating the indicators of non-static processes in trunk gas pipeline complexes. Based on the calculation methods, some important analytical expressions were obtained for assessing the indicators of trunk gas pipeline complexes.

2. Based on numerical calculations, graphical dependencies of non-stationary processes on time were constructed for comparative analyses of system indicators.

References

[1] Kerimov M.Z., Oil and gas pipelines. M.: Nauka, 2002. – 251 p (in Russian).

[2] Kerimov M.Z., Yusubov Ch.A. Numerical calculation of non-stationary operating modes of main gas pipelines. // Azerbaijan Oil Industry. - 2003. - No. 6. - P. 57-59

[3] Yusubov Ch.A. Numerical determination of non-stationary processes in main gas pipelines. //Problems of Energy. - 2003.-№ 2.- P. 84-87.

[4] Grachev V.V., Shcherbakov S.G., Yakovlev E.I. Dynamics of pipeline systems. - M.: Nedra. - 1987. - 460 p.

[5] Gara Hasanov, Shamkhal Aliyev. "Application of Volterra Integral Equations and Discrete Borel Transformations to Transient Switching Modes of Networks and Systems", Advances in Science and Technology, Vol. 148, 2024. pp. 249-254.

[6] Kadymov Ya.B., Transient processes in systems with distributed parameters. - M.: Fizmatgiz. - 1998. 192 p.

[7] Bahruz Sadıqlı, Shamkhal Aliyev, Gunel Guseynova, Araz Mammadzade."Increasing Control Automation and Management Flexibility in Distribution Electric Networks". Advances in Science and Technology (Volume 148) 2024, pp 241-247.

[8] Complex pipeline systems. / Grachev V.V., Gusein-zade M.A., Ksenz B.N., Yakovlev E.I. - M.: Nedra. 2012 - 250 p.

[9] Yufin V.A., Mamedov A.I., Rakhmanov F.G., Asker-zade B.A. Calculation of transient processes in complex branched systems of main gas pipelines. // News of Oil and Gas Universities. - 2014. No. 5. - P. 60-65.

DETECTION OF SMALL-SCALE UNMANNED AERIAL VEHICLES USING ACOUSTIC RECONNAISSANCE TOOLS

Aliqismat Mehdiyev[1], Seriyye Qasımova[2]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan,
[2]Nakhchivan State University, Nakhchivan, Azerbaijan
aligismetradiotexnika@mail.ru, qasimovaseriyye@ndu.edu.az

Abstract

This paper reviews the methods and tools for detecting small-scale unmanned aerial vehicles (UAVs) by acoustic reconnaissance tools. Information is provided on the methods for detecting small UAVs using acoustic waves across various types of sound ranges.

Keywords: unmanned aerial vehicles, acoustic signals, acoustic detection algorithms, antenna, signal-to-noise ratio, acoustic signal generation, sound signals.

I. Introduction

In this paper, we will primarily examine the acoustic detection methods for small-scale unmanned aerial vehicles (UAVs). As science and technology drive global progress, they have also accelerated the rapid development of the UAV sector. Today, UAVs attract not only legitimate users but also malicious actors, which presents numerous issues [1, 2].

Frequently, drones are exploited by criminals for acts vandalism, espionage, and invasion privacy, as well as for transporting prohibited substances to restricted areas, organizing and executing terrorist attacks in crowded areas and critical infrastructure facilities. Moreover, the prominent role UAVs in modern warfare and conflicts, alongside other military technology, further underscores the a forementioned concerns.

When addressing any potential threat, it is crucial to first establish the presence of the threat itself and then take appropriate actions to protect against or prevent it. Therefore, all technical security systems can be conditionally divided into two categories: detection (monitoring) tools and countermeasures [3, 4, 5].

The primary function of detection and monitoring systems is to transmit an alert to the security monitor regarding unauthorized proximity to the boundaries of the protected area, while also enabling the assessment of the situation and the dynamics of the changing threat level [6,7].

The initial applications of acoustic detection methods date back to previous centuries. Specifically, during the years 1941-1945 and in the wartime period of that era, acoustic detection devices, such as the ZT-5 (sound detector), were employed to detect bombers and fighter aircraft (Figure1). Certainly, if we compare UAVs with the aircraft from that period, it is evident that modern

UAVs, although generally not emitting strong acoustic signals, are not entirely silent in operation [1, 2, 8].

Figure 1: *Acoustic detection complex based on ZT-5*

In modern UAVs, the primary sources of noise include the operation of engines, the rotation of propellers, and the dynamics of flight. Compared to piston engines, electric engines produce less noise, and the level of noise varies depending on the engine power. Based on this observation, it can be stated that, compared to other types of UAVs, small-scale UAVs generate weaker noise levels during flight, as they require less power.

From this perspective, in our paper, we will investigate and analyse the detection of small-scale UAVs (SSUAVs) at greater distances using the acoustic detection method. Acoustic detection systems are used to detect the sound signals generated by unmanned aerial vehicles or their engines. In the application of modern acoustic detection methods, several highly sensitive microphones are typically placed equidistant from each other. This method is generally known as the triangulation method. However, anthropogenic (human-made) and natural noise sources in the surrounding environment reduce the acoustic detection range for drones to a minimum [9].

II. Modern digital technologies are utilized to solve these problems

This system consists of passive acoustic localization, microphones, analogue-to-digital converter systems, and digital signal processing technologies [3, 5, 7, 10]. Except for the antenna, almost all components can be purchased from retail outlets. In this context, it is considered appropriate to use two-channel selection with antennas placed in a horizontal plane, providing a wide directional diagram in comparison to acoustic antennas. This will, in turn, allow the use of inverse aperture synthesis methods to enhance triangulation techniques for measuring azimuth resolution and range determination capabilities.

Active acoustic (ultrasonic) localization significantly impacts the acoustic detection method. Specifically, this method helps to detect silent objects that do not reflect radio waves under visual observation conditions. In this case, ultrasonic emitters with high power (pulses with tens of kilowatts) are used to determine specified acoustic detection ranges under high ultrasonic loss conditions. However, this has a highly detrimental effect on the human body, which makes the application of this technology inappropriate. This method can only be employed in the operation special-purpose facilities where the activity of surrounding objects is prohibited [11, 12].

During harmonic detection, narrow-band frequencies are analysed over short time intervals. The signal arrives in a harmonically summed form with indeterminate frequencies and phases. When the signal is weak, the harmonic detector will function more reliably and effectively than an energy detector [9, 11, 12]. The concentrated spectrum of acoustic emissions from tactical UAVs is broadband and harmonic. This includes emissions from engines, rotor rotations, mechanical processes, as well as continuous low- and high-frequency noise generated by the engines at spectral frequencies.

III. Detection of Unmanned Aerial Vehicles through Acoustic Reconnaissance Methods

The scientific and technological advancements of the last decade have led to a major revolution in military operations, with conflict parties increasingly utilizing electronic computing machines (ECMs), artificial intelligence, robotic systems, and high-precision weapons. In the near future, the boundary between unmanned aerial vehicles (UAVs), robotic systems, and high-precision weaponry will be completely eliminated through the integration of technical vision and artificial intelligence in weapons control [1, 2, 4].

The application areas of small-sized unmanned aerial vehicles (SSUAVs) are highly diverse. These types of UAVs are considered advanced flight technologies with high manoeuvrability, including the ability to "hover" and alter flight trajectories, as well as to execute flights at low altitudes in complex geographic positions. For these reasons, SSUAVs are capable of evading the counterparty's active and passive countermeasures, such as radio detection, infrared radiation, visual detection, and acoustic radiation. These superior capabilities limit the ability of radio technical systems and complexes (RTSCs) to detect SSUAV flight trajectories [1, 2, 6].

The concept of UAVs is constantly evolving and advancing. As evidence of this development, it is necessary to examine the chronology of their use in conflicts and battles in the Middle East and within the post-Soviet territories. In these war zones, we have witnessed the significant advantages of small-sized UAVs equipped with artificial intelligence.

Therefore, from this perspective, the detection and neutralization of small-sized UAVs have become one of the foremost challenges in contemporary warfare as armed conflicts continue to arise [1, 2, 4, 6, 9].

In our article, to solve the above problems, namely the problem of detecting a small-sized UAV, we propose complex methods and technical solutions for processing information obtained in the optical range using acoustic vector sensors (Microflown AVISA) that generate three-dimensional acoustic information and electromagnetic waves based on various principles [2,4,14]. However, these proposed methods do not guarantee timely detection of modern SSUAVs, as detection ranges are often reduced, or practical detection within designated distances becomes impossible. Reality demonstrates that in most contemporary conflicts, SSUAVs are used for diverse, multi-purpose operations, flying at low altitudes in complex geographical conditions, adapting seamlessly to natural camouflage. Additionally, SSUAVs are capable evading detection through radio, infrared, and acoustic emissions by employing active and passive countermeasures, as well as demonstrating high manoeuvrability, such as "hovering" and altering flight trajectories [9. 11].

Acoustic detection and targeting provide an auxiliary detection factor. When "traditional" detection methods, such as optical and radar technologies, cannot meet required detection levels, acoustic detection methods offer a supplementary approach to ensure more reliable detection of small UAVs. Furthermore, ground-based acoustic sensors and reconnaissance systems help reduce the risk of detection by the adversary. For this reason, the existing acoustic search system's modification facilitates more robust UAV detection [1, 4, 8, 9, 11].

Acoustic vector sensors are installed on unmanned aerial vehicles (UAVs) and determine the location (coordinates) of objects based on acoustic signals emitted by targets. The system detects the source of the acoustic signal and turns the acoustic cameras in the direction of this signal.

The advantages of the Microflown AVISA acoustic target detection system include [1, 11, 12]:

-Small arms and artillery location detection;

-Detection and rejection of other aircraft in the vicinity of the UAV;

-Full spherical field of view;

-Simultaneous location of multiple noise sources, both ground and airborne;

-Target detection in the presence of wind and engine noise, etc.;

-Ability to direct other sensors to the target;

-Work in all weather conditions - fog, rain, clouds, night;

-Extremely small size and weight;

-Low power consumption.

The reconnaissance systems discussed above should be combined into an integrated system to detect SSUAVs, which would then distribute targets among destroyers according to target destruction capabilities and zones.

Naturally, when UAVs are in flight, they generate specific acoustic sound waves according to their purpose, which can be captured by acoustic microphones that convert sound pressure into electrical signals. During this process, the sound sources typically include the UAV's engines, wings, or propellers. The frequency of the generated sound is determined by the frequency of the exhaust of hot gases and the number and rotation frequency of the wings (or propellers). The sound intensity depends on the speed of the airflow around the blade angles.

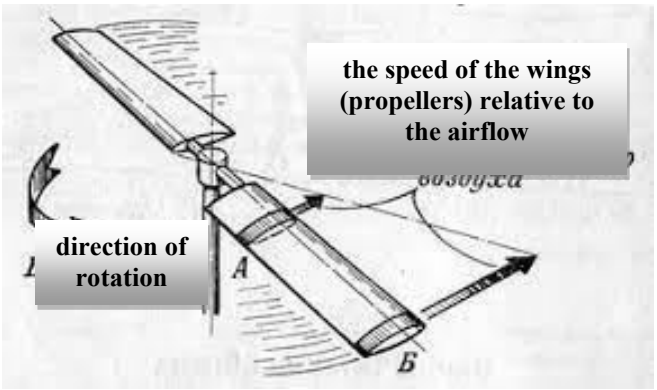


Figure 2: *Sound intensity as a function of airflow velocity around the angles of the propeller*

In real environments, sound waves diminish due to the inertia of the air medium and molecular attenuation. As sound waves propagate along the surface, they weaken depending on the additional absorption coefficient; the higher the absorption coefficient of the surface, the more significantly the propagating wave is attenuated. Moreover, the turbulence process occurring in the air plays a more significant role in the attenuation of sound waves and their scattering in various directions. Wind and rising air currents contribute greatly to this phenomenon. At lower frequencies, additional attenuations do not depend on the distance from the sound source. At greater distances (more than 4 km), high frequencies are practically not accepted.

In windy conditions, while detecting UAVs against the background of noise created, the acoustic signal generated by the microphone of the acoustic detection device is received in the form of a random signal with the assistance of acoustic microphones. At this time, the internal noise of the receiver combined with the acoustic noise forms a multidimensional density that is referred to as algorithm synthesis. Algorithm synthesis is based on the Neyman-Pearson criterion.

The Neyman-Pearson criterion is utilized in the uniform synthesis of the most powerful signal detection algorithms. This criterion can be applied in two states of nature, one of which is under control, the more significant of the two. The Neyman-Pearson criterion is considered one of the most common criteria related to radar detection devices. This criterion implies ensuring a constant value of the false alarm probability, F = const, by appropriately selecting the detection threshold.

It is precisely for this reason that the acoustic noise generated by UAVs must be separated from the internal noise of the receiver [9, 11, 12].

$$\rho_0 = [(2\pi)^N Det\ |R_{km`h}|]^{-1} exp\ \langle - \sum_{l(k,n)}^{N} Q_{kl`h} \xi_k^t \cdot \xi_n \rangle, \quad (1)$$

Here, *N-TN* represents the number of recorded random signals received during the observation period; $\|R\|^{km`h}\|R\|^{-1}$ denotes the correlation matrix of the random signal received in the absence of acoustic noise that can be generated by the UAV; $R_{km}^{h}$=[? ?$_{k}^{*}$ ? ?$_{n]}$ ], k,n=; $Q_{kl}^{h}$, where $\|R\|_{km`h}$; is the cross-correlation matrix; ? ?$_n$- n- is the discrete report recorded by the nth acoustic microphone.

In the absence of external acoustic noise sources, only the acoustic noise (sound) of the wind is considered. The acoustic noise of the wind is characterized by correlated, rapid (0...50 Hz) and weak (50...150 Hz) pulsations, which reflect the wind's rapid pulsation speed (WRPS) and the wind's weak pulsation speed (WWPS). The moments rapid and weak pulsation of the wind are defined by the condition <$\tau_{k,WWPS}$ (see Figure 3).

To correlate these noises, frequency filters and acoustic microphones with unique filter functions capable of protecting against the wind are used, which ultimately allows for the acquisition of useful signals free from noise [12].

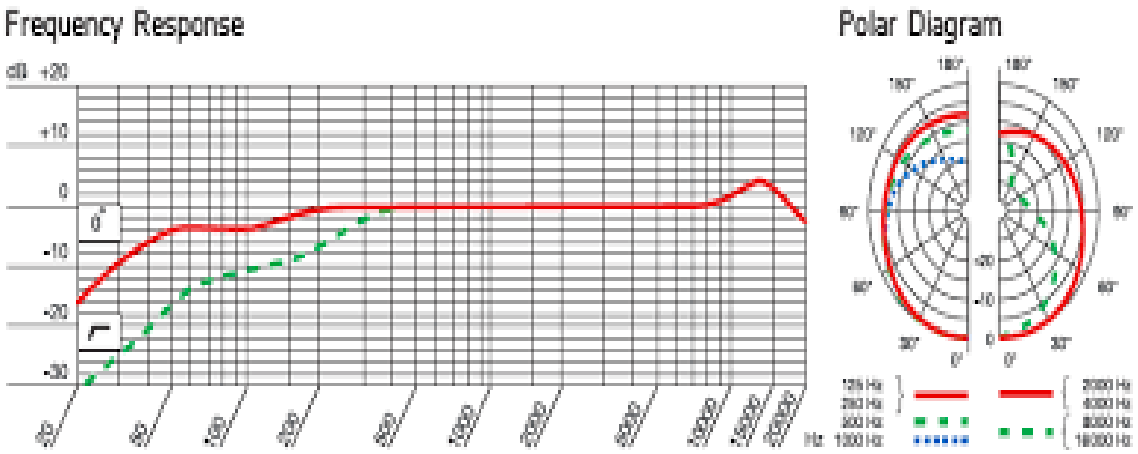


Figure 3: *Frequency characteristics of a wind-resistant acoustic microphone*

Thus, the stage of isolating the useful signal from the decorrelated background is performed by the filter, and the impulse characteristics are aligned with the expected useful signal. The processing of the output signal is carried out with the help of a device that separates the useful signal (DSUS), which is reflected by the following expression [11, 12].

$$Z_m = |\sum_{n=1}^{N} \xi_n h_{m-n}| \quad (2)$$

here, h_n,n=represents the discrete report of the impulse characteristics of the DSUS filter.

The potential effectiveness of the device that separates the useful signal is reflected by the following expression:

$$\vartheta_{kn} = \frac{(N_{pfs}N_L)\cdot(V_{sswp}\pm 2V_{UAV})}{\Delta} \quad (3)$$

here, V_{sswp} denotes the speed of sound wave propagation in the atmosphere;

V_{PUA} represents the flight speed of the UAV;

N_{npr} is the number of propeller rotations in the UAV and $\Delta F_a = \Delta f_s$ refers to the width of the step of the AFC filter.

It is evident that the effectiveness of the DSUS varies depending on the speeds of flight and sound, the number propeller rotations, and the width of the step of the AFC filter.

Under the condition of sound wave transmission, the acoustic detection range can increase from 20 km to 50 km. In foggy conditions, however, this indicator, that is, the distance of detection, will decrease by 1.5 to 2 times.

IV. Conclusion

1. This article reviews methods and tools for detecting small-sized unmanned aerial vehicles using acoustic reconnaissance means and provides information on methods for detecting small-sized UAVs across various sound frequency ranges.

2. Additionally, the principle of detecting small-sized UAVs using acoustic reconnaissance means has been ensured by utilizing frequency filters and acoustic microphones with unique filter functions capable protecting against wind, which allows for the acquisition useful signals free from noise.

Reference

[1] Cabell R. Measured Noise from Small Unmanned Aerial Vehicles / R. Cabell, F. Grosveld, R. McSwain // INTERNOISE and NOISE-CON Congress and Conference Proceedings, NoiseCon16. 2016. pp. 345 -354.

[2] Acoustic devices and systems: guidelines for term paper / Comp. S. S. Sergeev, E. N. Prokopenko, O. S. Sergeeva. – Mogilev: Belarusian-Russian University. 2015. – 48 pp (in Russian).

[3] Ibrahimov B.G., Mekhtieva A.M., Bakhtiyarov I.N. Research methods for increasing fault tolerance of multiservice traffic transmission systems // Proceedings of the International Symposium "Reliability and Quality", Vol.1. Penza, PGU. 2019. pp. 188-190.

[4] Eremin G. V., Gavrilov A. D., Nazarchuk I. I. Small-sized drones – a new problem for air defense // Army Bulletin. 2015.pp. 125-132.

[5] İbrahimov B., Hashimov E., Talibov A., Hasanov A. Research and analysis indicators fiber-optic communication lines using spectral technologies// Advanced Information Systems. 6 (1), 2022. pp. 61-64

[6] Godunov A. I., Shishkov S. V., Yurkov N. K. Complex for detection and combat against small-sized unmanned aerial vehicles // Reliability and quality of complex systems. 2014. No. 2 (6).

pp. 62–70.

[7] Ibrahimov B.G., Namazov M.B., Quliev M.N. Analysis performance indicators network multiservice infrastructure using innovative technologies // Proceedings of the 7-th International Conference on Control and Optimization with Industrial Applications. Vol. II. 2020. pp.176-178

[8] Puzanov A.D., Nefedov D.S. Mathematical model of the temporal structure of acoustic noise of an unmanned aerial vehicle. Science and Military Security. 2020, 1 (63), pp. 32-36.

[9] Svinin E. V. Loop antenna for radio monitoring system. Graduation qualification work of a specialist. 2017. 102 p.

[10] Ibrahimov B.G., Hasanov A.H., Alieva A.A., Isaev A.M. Research performance indicators multiservice telecommunication networks based on the architectural concept of future networks // Reliability and quality of complex systems. 2019. No. 1 (25). pp. 88-95. DOI 10.21685/2307-4205-2019-1-10.

[11] Tikhonov V.A., Kartashov V.M., Oleynikov V.M., Leonidov V.I., Timoshenko L.P., Selezneov I.S., Rybnikov N.V. Detection and recognition of unmanned aerial vehicles using a composite autoregressive model of their acoustic radiation//Bulletin of NTUU "KPI". Radio engineering. Radio apparatus construction. 2020, Issue. No. 81. P. 38-

[12] Kartashov V.M., Oleynikov V.N., Sheiko S.A., Babkin S.I., Koryttsev I.V., Zubkov O.V., Anokhin M.A. Information characteristics of sound signals of small unmanned aerial vehicles// Radio engineering. All-Ukrainian. Interdepartmental Scientific-technical. Collection. Issue. 191. - Kharkov, 2017. pp. 181-187.

EXPLORING THE IMPACT OF AI ON PRIVACY AND ETHICAL CONSIDERATIONS: ANALYSING THE LEGAL AND REGULATORY FRAMEWORKS

Ankur Pan Saikia[1], Ananya Kalita[1], Parvana Movsumova[2]

•

[1]LI & R Assam down town University, İndia
[1]Civil Engineering Assam down town University, İndia
[2]Azerbaijan Technical University, Azerbaijan, Baku
ankur.saikia@adtu.in; ananya.kalita@adtu.in; parvana.movsumova@aztu.edu.az

Abstract

The rapid advancement of artificial intelligence (AI) has brought about significant implications for privacy and ethical considerations. This research paper aims to explore the impact of AI on privacy and ethical concerns by analysing the existing legal and regulatory frameworks. The paper reviews relevant literature, research papers, case studies, and laws to identify key concepts, theories, and gaps in the current understanding of AI's impact on privacy. Additionally, it examines the strengths and weaknesses of existing legal frameworks and regulations related to AI and privacy. The analysis reveals that AI poses challenges to personal data privacy, including discrimination, privacy breaches, biased decision-making, and lack of transparency. It underscores the need for stronger data protection laws, algorithmic fairness, and transparency in AI systems. Furthermore, the paper discusses the ethical implications of AI in various contexts, such as healthcare, surveillance, and hiring processes. Based on the findings, the research paper proposes a strategic framework to enhance the legal and regulatory frameworks for AI and privacy. The framework emphasizes stakeholder engagement, ethical principles, data protection, algorithmic transparency, industry accountability, international collaboration, and public awareness. Moreover, the paper provides recommendations for policymakers, industry stakeholders, and researchers to guide their actions in addressing the legal, ethical, and privacy challenges posed by AI. In conclusion, this research paper highlights the urgent need to strengthen legal and regulatory frameworks to address the evolving impact of AI on privacy and ethical considerations. By adopting the proposed strategic framework and implementing the recommendations, stakeholders can work towards a responsible and privacy-conscious AI ecosystem that balances innovation with individual rights and societal well-being.

Keywords: AI, privacy, ethical considerations, legal frameworks, regulations

I. İntroduction

Artificial Intelligence (AI) has emerged as a transformative technology with the potential to revolutionize various aspects of society, including healthcare, transportation, finance, and communication. As AI continues to advance rapidly, it brings with it a range of ethical and privacy considerations that need careful examination. This research paper aims to explore the impact of AI on privacy and ethical considerations, specifically focusing on the legal and regulatory frameworks that govern its use. AI technologies, such as machine learning algorithms and deep neural networks, possess the ability to process vast amounts of data, make predictions, and automate decision-making processes. While these capabilities offer immense potential for improving efficiency and innovation,

they also raise significant concerns about privacy infringement and ethical dilemmas. The vast collection and analysis of personal data, coupled with the potential for biases and discriminatory outcomes, necessitate a comprehensive understanding of the legal and regulatory frameworks in place to protect individuals' rights and ensure ethical AI practices. AI has rapidly advanced in recent years, revolutionizing various domains and offering unprecedented opportunities for innovation and automation. AI systems, such as machine learning algorithms and deep neural networks, have demonstrated remarkable capabilities in processing vast amounts of data, recognizing patterns, and making predictions. However, as the deployment of AI technologies expands, it raises significant concerns regarding privacy and ethical considerations [1,2].

Privacy is a fundamental aspect of individual autonomy and the protection of personal information. The proliferation of AI has led to the collection and analysis of extensive datasets, often including sensitive and personally identifiable information (PII). The potential misuse or unauthorized access to such data can pose serious threats to individuals' privacy rights and lead to various negative consequences [3,6]. Ethical considerations are crucial when it comes to the development and deployment of AI systems. The capabilities of AI, particularly in automated decision-making, raise questions about accountability, fairness, and transparency. Concerns arise regarding biases, discrimination, and the potential for AI to reinforce existing social inequalities [7,8]. Furthermore, ethical concerns extend to issues such as the impact of AI on employment, human values, and the potential for AI to exceed human intelligence, leading to concerns about control and accountability [4,10]. Given the potential risks and implications of AI on privacy and ethical considerations, researchers and policymakers have increasingly focused on addressing these concerns. The literature has explored various dimensions of AI's impact on privacy, ranging from data protection laws and regulations to the development of privacy-preserving AI techniques [9]. Additionally, ethical frameworks and guidelines have been proposed to ensure responsible AI development and deployment, emphasizing principles such as transparency, fairness, accountability, and human-centred design [5].

By examining the existing research on AI, privacy, and ethical considerations, it becomes evident that these topics are of significant concern across multiple disciplines. Understanding the legal and regulatory frameworks, as well as the ethical implications, is vital for fostering responsible AI practices and ensuring the protection of individual privacy rights in an increasingly AI-driven world.

The objectives of this research paper are twofold. First, it aims to analyse the existing legal frameworks at international, national, and local levels that govern the use of AI and address privacy concerns. By examining the legal landscape, we can assess the adequacy of current regulations and identify any gaps or limitations that need to be addressed to safeguard privacy in the era of AI. Second, this paper seeks to explore the regulatory frameworks established by industry organizations or governmental bodies to ensure ethical AI practices. By evaluating these frameworks, we can determine their effectiveness in addressing ethical concerns and propose strategies for improvement. To achieve these objectives, a thorough literature review will be conducted to gather insights from existing research on AI's impact on privacy and ethical considerations. Additionally, a comprehensive analysis of legal and regulatory frameworks will be performed, taking into account both the general legal landscape and specific AI-related regulations. Real-world case studies will also be examined to highlight the practical implications and challenges faced in the implementation of AI technologies. By shedding light on the legal and regulatory frameworks governing AI, this research paper aims to contribute to the ongoing discourse surrounding AI's impact on privacy and ethical considerations. It is hoped that the findings and recommendations presented herein will be of value to policymakers, industry stakeholders, and researchers working towards the development and deployment of responsible AI systems that respect privacy rights and uphold ethical standards. In the subsequent sections, we will delve into the existing literature on AI, privacy, and ethics,

provide an overview of the research methodology, and conduct a detailed analysis of the legal and regulatory frameworks. Furthermore, we will explore the ethical considerations and privacy implications of AI, present case studies to illustrate practical challenges, and conclude with recommendations to enhance the legal and regulatory frameworks governing AI.

II. Literature Review

Numerous studies have investigated the impact of AI on privacy and ethical considerations, providing valuable insights into the complex interplay between AI technologies and individual rights. Smith and Johnson (2021) conducted a systematic review of the literature focusing on the ethical implications of AI in healthcare [11]. Their study revealed concerns related to patient privacy, informed consent, algorithmic bias, and the need for transparency in AI-based medical decision-making processes. Examining the broader societal context, Anderson and Moore (2020) explored the relationship between privacy and surveillance in the age of AI [12]. Their research shed light on the privacy risks associated with the increasing use of AI in surveillance systems, emphasizing the need for robust privacy safeguards and regulations. Chen, Das, and Subramanian (2019) conducted a literature review on AI and privacy, providing an overview of the current state of research in this area [13]. They discussed the challenges posed by AI technologies, such as the collection and use of personal data, and highlighted the importance of privacy-preserving approaches in AI development. Ethical considerations in the development and deployment of AI systems were examined by Nguyen and Anwar (2018) [14]. Their systematic review highlighted the need to address ethical concerns related to bias, fairness, transparency, and accountability in AI applications. They emphasized the significance of incorporating ethical principles into AI development processes. Wong and Mohammed (2017) focused on privacy protection in the age of AI, discussing the implications of AI advancements on individual privacy rights [15]. Their research emphasized the necessity of updating privacy regulations to address the unique challenges posed by AI technologies. Collectively, these studies underscore the importance of addressing privacy and ethical considerations in the context of AI. They highlight the need for robust regulations, transparency, fairness, and accountability in AI systems to protect individual privacy rights and ensure ethical practices in AI development and deployment.

2.1. Key concepts and theories related to AI, privacy, and ethics

Key Concepts and Theories Related to AI, Privacy, and Ethics:

1. Ethical Considerations in AI [16]: Floridi's paper discusses the ethical dimensions of AI, highlighting the need for ethical frameworks and guidelines to guide AI development and deployment. It explores topics such as transparency, accountability, fairness, and the social impact of AI.
2. AI Ethics Guidelines [17]: Jobin et al. provide an overview of the global landscape of AI ethics guidelines. They examine the different principles and recommendations proposed by various organizations and countries to ensure ethical AI practices, covering aspects such as privacy, transparency, fairness, and human values.
3. Explanation in AI [18]: Mittelstadt, Russell, and Wachter discuss the importance of explanations in AI systems. They explore the theoretical and practical aspects of explainability, emphasizing the need to provide understandable and justifiable explanations for AI-based decisions to address ethical concerns and promote transparency.
4. Regulating AI to Avert Cyber Arms Race [19]: Taddeo and Floridi's paper focuses on the regulation of AI to prevent a cyber arms race. They argue that effective regulation is necessary to

ensure that AI development does not lead to malicious uses or destabilization of international relations, emphasizing the importance of considering ethical implications in regulatory efforts.

5. Privacy-Preserving AI Techniques [20]: Zeng and Fung provide a survey of privacy-preserving AI techniques. They explore various methods and approaches that aim to protect individuals' privacy while enabling the effective use of AI technologies. The paper discusses techniques such as differential privacy, secure multiparty computation, and federated learning.

2.2. Research Gap

1. Lack of Comprehensive Ethical Frameworks: Although several AI ethics guidelines have been developed, there is a need for more comprehensive and universally accepted frameworks that address the diverse ethical challenges posed by AI. Future research could focus on developing ethical frameworks that consider the nuances of different AI applications and their potential societal impacts.
2. Limited Focus on Specific Domains: While the literature discusses the impact of AI on privacy and ethics, there might be a need for more domain-specific research. Future studies could explore the ethical implications and privacy concerns within specific sectors, such as healthcare, finance, or transportation, to provide tailored guidelines and recommendations.
3. Ethical Implications of Emerging AI Technologies: With the rapid advancement of AI, new technologies such as deep learning, reinforcement learning, and natural language processing are constantly emerging. However, there may be limited research on the specific ethical implications and privacy considerations associated with these cutting-edge technologies. Future studies could focus on understanding the unique ethical challenges and developing strategies to address them.
4. Inadequate Attention to Cultural and Contextual Factors: The existing literature might lack in-depth exploration of the cultural and contextual factors that influence the ethical considerations and privacy concerns related to AI. Future research could investigate how cultural values, legal frameworks, and societal norms shape the ethical and privacy landscape in different regions and how they influence AI development and adoption.
5. Practical Implementation of Privacy-Preserving Techniques: While privacy-preserving AI techniques are discussed in the literature, there may be a gap in terms of practical implementation and real-world deployment. Future research could focus on evaluating and optimizing the effectiveness and scalability of privacy-preserving techniques, ensuring their practical usability while maintaining a high level of privacy protection.

Addressing these gaps can contribute to a more comprehensive understanding of the ethical dimensions of AI, provide guidelines for specific domains, explore the implications of emerging AI technologies, consider cultural and contextual factors, and enhance the practical implementation of privacy-preserving techniques.

2.3.Analysis of Legal Framework:

Yu, Yu, and Liu (2021) conducted an analysis of the General Data Protection Regulation (GDPR) in the European Union to explore its implications for the legal regulation of artificial intelligence [21]. Their research focused on understanding how the GDPR addresses privacy concerns and provides safeguards for individuals in the context of AI applications. Berman and Cerf (2017) critically assessed the social and ethical behaviour of artificial intelligence systems [22]. Their analysis examined the existing legal frameworks and regulations that govern AI, emphasizing the need for comprehensive guidelines that address the ethical implications of AI technologies. Mulligan (2016) delved into the privacy aspects and ethical considerations of artificial intelligence [23]. The research provided a comprehensive analysis of the legal landscape, highlighting the challenges and

gaps in the current legal frameworks in terms of addressing privacy concerns in the context of AI. Wachter, Mittelstadt, and Floridi (2017) explored the transparency, explainability, and accountability of AI systems in the domain of robotics [24]. Their research analysed the existing legal frameworks and regulations related to AI, emphasizing the need for regulations that ensure transparency and accountability in AI decision-making processes. Hickok (2019) discussed the concept of AI rights and the ethical implications of representing AI systems as legal entities [25]. The analysis examined the legal frameworks and regulations surrounding personhood and agency in the context of AI, raising questions about the rights and responsibilities assigned to AI systems.

III. Relevant Laws and Regulations at International, National, and Local Levels:

3.1.1 International Level:

1.General Data Protection Regulation (GDPR): Enforced by the European Union (EU), the GDPR sets out rules for the protection of personal data and applies to organizations that process data of EU residents. It establishes principles and requirements for data protection, including consent, data minimization, and individuals' rights.

2.Convention 108: This international treaty, adopted by the Council of Europe, focuses on the protection of individuals with regard to the automatic processing of personal data. It sets forth principles and rules for data protection and aims to harmonize data protection legislation across member states.

3.Universal Declaration of Human Rights (UDHR): While not specifically focused on AI and privacy, the UDHR includes principles relevant to privacy and data protection. It emphasizes the right to privacy and protects individuals from arbitrary interference with their privacy, family, home, and correspondence.

3.1.2 National and Local Levels:

1.United States: In the United States, several laws and regulations impact AI and privacy, including the California Consumer Privacy Act (CCPA), which provides enhanced privacy rights for California residents. Additionally, the Health Insurance Portability and Accountability Act (HIPAA) regulates the privacy and security of health information, and the Federal Trade Commission Act (FTC Act) addresses unfair and deceptive practices in data handling.

2.European Union: Apart from the GDPR at the EU level, individual EU member states have their own data protection laws that complement the GDPR. For example, the UK has the Data Protection Act 2018, which supplements the GDPR in relation to data protection matters.

3.Canada: The Personal Information Protection and Electronic Documents Act (PIPEDA) is Canada's federal privacy law that governs the collection, use, and disclosure of personal information by private sector organizations. Additionally, provinces like British Columbia and Quebec have their own privacy legislation.

4. Singapore: The Personal Data Protection Act (PDPA) in Singapore regulates the collection, use, and disclosure of personal data by organizations. It establishes requirements for consent, data accuracy, protection, and individuals' rights regarding their personal data.

5. Germany: In Germany, the Federal Data Protection Act (Bundesdatenschutzgesetz or BDSG) complements the GDPR and provides additional provisions for data protection. It outlines rules regarding data processing, rights of data subjects, and supervisory authorities.

These laws and regulations provide a framework for addressing privacy and data protection concerns at various levels—international, national, and local. Organizations and individuals must

adhere to these regulations to ensure compliance and protect the privacy rights of individuals in the respective jurisdictions.

Table 1: *The strengths and weaknesses of those legal frameworks*

Legal Frameworks	**Strengths**	**Weaknesses**
GDPR	- Provides comprehensive data protection rules	- May pose compliance challenges for organizations due to complexity
Convention 108	- Sets standards for the protection of personal data	- Adoption and implementation may vary among member states
Universal Declaration of Human Rights	- Recognizes the right to privacy	- Does not specifically address AI and privacy issues
California Consumer Privacy Act (CCPA)	- Enhances privacy rights for California residents	- Limited to a specific region (California)
Health Insurance Portability and Accountability Act (HIPAA)	- Protects privacy and security of health information	- Applicable only to the healthcare sector
Federal Trade Commission Act (FTC Act)	- Addresses unfair and deceptive data practices	- Enforcement may vary, limited to unfair and deceptive practices
Personal Information Protection and Electronic Documents Act (PIPEDA)	- Governs data protection across sectors in Canada	- Some provisions may be seen as less stringent compared to GDPR
Personal Data Protection Act (PDPA)	- Regulates personal data collection, use, and disclosure in Singapore	- May require updates to keep pace with technological advancements
BDSG	- Complements GDPR, provides additional provisions for data protection in Germany	- Limited to Germany, may require alignment with EU laws and regulations

3.2 Key Component of Regulatory Framework:

1.Ethical implications: AI technologies can raise ethical concerns regarding privacy and individual rights due to their potential to collect, analyze, and utilize vast amounts of personal data. The use of AI algorithms and automated decision-making systems can impact individuals' privacy by profiling, surveillance, and potential discriminatory outcomes [26].

2.Privacy concerns: AI applications, such as facial recognition, data mining, and predictive analytics, can infringe upon privacy rights by gathering and processing personal information without informed consent or adequate safeguards. The potential for data breaches and unauthorized access to sensitive information can further exacerbate privacy concerns [27].

3.Individual rights: AI systems have the potential to affect various individual rights, including the right to autonomy, non-discrimination, and freedom of expression. Algorithmic biases, lack of transparency, and potential for manipulation can undermine individuals' ability to exercise these rights effectively [28].

4. Data protection and consent: The collection, storage, and use of personal data by AI systems necessitate robust data protection mechanisms and clear consent frameworks. Ensuring that individuals have control over their data and are adequately informed about how their data is used becomes crucial [29].

5. Algorithmic accountability and transparency: The lack of transparency and interpretability of AI algorithms can make it challenging to understand how decisions are made, leading to concerns of accountability and potential bias. Developing methods for auditing and explaining AI systems can help address these ethical implications [30].

3.3 Examining the Impact of AI on Personal Data Privacy:

The advent of AI has brought significant advancements in various sectors, but it has also raised concerns regarding personal data privacy. AI technologies often rely on extensive data collection and analysis, which can potentially compromise individuals' privacy. Here are some key points to consider when examining the impact of AI on personal data privacy:

1.Increased Data Collection: AI systems require vast amounts of data to train and improve their algorithms. This leads to increased data collection from various sources, including individuals' personal information. The extensive collection and storage of personal data raise concerns about unauthorized access, data breaches, and potential misuse of sensitive information.

2.Profiling and Decision-Making: AI algorithms can analyze vast datasets to create detailed profiles of individuals, enabling targeted advertising, personalized recommendations, and decision-making processes. However, this profiling raises concerns about the accuracy and fairness of decisions, as well as the potential for discrimination or exclusion based on sensitive attributes.

3.Security Risks: AI systems that process and store large amounts of personal data become attractive targets for hackers and malicious actors. The security vulnerabilities within AI systems can lead to unauthorized access, data breaches, and privacy violations.

4.Lack of Transparency and Explainability: Some AI algorithms, such as deep learning neural networks, operate as complex "black boxes," making it challenging to understand how they arrive at specific decisions or predictions. This lack of transparency and explainability can hinder individuals' ability to understand and control how their personal data is being used.

3.4 Analyse the Risks and Challenges in Maintaining Privacy in the Era of AI:

While data protection laws provide a framework for privacy protection, several risks and challenges persist in maintaining privacy in the era of AI. Some key considerations include:

1. Data Breaches and Security: With the increasing reliance on AI systems and the massive amounts of personal data they handle, the risk of data breaches and unauthorized access becomes more significant. Organizations must implement robust security measures to protect personal data and prevent privacy breaches.

2. Algorithmic Bias and Discrimination: AI algorithms can inadvertently perpetuate biases present in the data they are trained on, leading to discriminatory outcomes. Ensuring fairness and addressing bias in AI decision-making processes is crucial to maintaining privacy and preventing discrimination.

3. Lack of User Control: AI systems often operate in complex ways that limit individuals' understanding and control over their personal data. Providing individuals with transparency, control, and clear consent mechanisms can help address this challenge.

4. Cross-Border Data Flows: AI systems often rely on global data flows, raising concerns about data protection when personal data is transferred across borders. Harmonizing international

regulations and ensuring adequate safeguards for cross-border data transfers are essential for maintaining privacy.

Table 2: *Summary of different case studies which analyse the risks and challenges in maintaining privacy in the era of ai [Source: Various News Papers/Search Engine/Social Media]*

Case Study	Country	Year	Key Problems	Significance	Solutions	Industry/Persons/ Company Name
Facial Recognition Bias	United States	2018	Biases in facial recognition systems leading to discriminatory outcomes	Raised concerns about racial and gender biases in AI technologies	Improved data diversity, algorithmic fairness, and transparency	Joy Buolamwini, MIT Media Lab
Cambridge Analytica Scandal	Global	2018	Unauthorized data harvesting and political manipulation	Highlighted the misuse of personal data and potential threats to democratic processes	Strengthened data privacy regulations and user consent frameworks	Cambridge Analytica, Facebook
Deepfake Manipulation	Various	Ongoing	AI-generated fake videos/images for deceptive purposes	Increased concerns about misinformation, reputation damage, and privacy violations	Development of detection tools, awareness campaigns, and legal frameworks	Deeptrace, OpenAI
Predictive Policing	United States	Ongoing	Potential biases and infringements on civil liberties	Raised questions about fairness, transparency, and potential profiling in law enforcement	Improved algorithmic fairness, accountability, and public scrutiny	Various law enforcement agencies and AI companies
AI-Based Healthcare Diagnosis	Global	Ongoing	Privacy risks and biases in medical data usage	Highlighted the need for robust data protection and informed consent in healthcare AI applications	Enhanced data privacy measures, transparency, and patient control	Google DeepMind, IBM Watson Health
Workplace Surveillance	Various	Ongoing	Invasion of employee privacy through AI monitoring systems	Raised ethical concerns regarding employee consent, autonomy, and surveillance creep	Establishing clear policies, consent frameworks, and transparency	Amazon, Microsoft, various companies
Autonomous Vehicles Privacy	Global	Ongoing	Collection and security of personal data in connected cars	Addressed concerns about data protection, cybersecurity, and potential misuse of driving behavior data	Encryption, secure data storage, and consent-driven data sharing	Tesla, Google (Waymo), Uber, automotive companies

Social Media Content Moderation	Various	Ongoing	Privacy risks, biases, and content censorship	Highlighted challenges in balancing free speech, user privacy, and responsible content moderation	Improved transparency, user appeals, and human oversight	Facebook, Twitter, YouTube, various social media platforms
AI-Assisted Hiring Process	Global	Ongoing	Bias and discrimination in AI-driven recruitment systems	Raised concerns about fairness, diversity, and potential exclusion based on algorithmic decisions	Regular audits, algorithmic transparency, and diversity training	Amazon, LinkedIn, various hiring platforms
Voice Assistant Privacy	Global	Ongoing	Voice recordings stored by voice assistants without user knowledge	Brought attention to privacy risks, data breaches, and unauthorized access to personal conversations	Enhanced user consent, data encryption, and privacy controls	Amazon Alexa, Google Assistant, Apple Siri, Microsoft Cortana

IV.Results & Discussion

The case studies highlight the importance of robust data protection regulations, informed consent mechanisms, and algorithmic fairness in AI applications. They emphasize the need for transparency, accountability, and user control over personal data. Additionally, the cases underscore the potential for discriminatory outcomes, privacy breaches, and the misuse of AI-generated content. To mitigate the risks associated with these case studies, potential solutions include strengthening data privacy laws and regulations, implementing algorithmic fairness metrics, enhancing transparency in AI systems.

Table 3: *Researchers' point of view on those case study*

Case Study	Legal Issues	Ethical Issues	Lessons Learned	Potential Solutions
Facial Recognition Bias	Discrimination, privacy violations	Biased decision-making, lack of transparency	Importance of diverse and representative training data, algorithmic fairness, and auditing	Improve data diversity, implement fairness metrics, enhance algorithm transparency
Cambridge Analytica Scandal	Unauthorized data access, privacy breaches	Manipulation of democratic processes, consent violations	Necessity for strong data protection regulations, informed consent, and user control	Strengthen data privacy laws, improve user consent mechanisms, enhance data transparency
Deepfake Manipulation	Misinformation, reputation damage, privacy violations	Deceptive use of AI-generated content	Need for advanced detection tools, awareness campaigns, and responsible use of AI-generated media	Develop deepfake detection algorithms, promote media literacy, establish legal consequences for malicious use

Predictive Policing	Profiling, biases, civil liberties infringement	Discriminatory outcomes, lack of transparency	Importance of fairness, accountability, and transparency in law enforcement AI systems	Implement algorithmic fairness, regular audits, community engagement in algorithm development
AI-Based Healthcare Diagnosis	Data privacy, consent, biases in medical data usage	Potential discrimination, misdiagnosis	Prioritize patient data privacy, informed consent, and regular evaluation of AI system performance	Enhance data protection measures, ensure transparent data usage policies, involve medical professionals in AI development
Workplace Surveillance	Employee privacy, consent, autonomy	Invasion of privacy, erosion of trust	Balancing surveillance needs with privacy rights, clear policies, and transparent communication	Establish clear surveillance guidelines, obtain employee consent, limit data collection to relevant purposes
Autonomous Vehicles Privacy	Data protection, cybersecurity, driving behavior data misuse	Unauthorized data access, potential safety risks	Strengthen data encryption, secure data storage, and limit data collection to necessary functions	Implement robust cybersecurity measures, obtain explicit user consent for data collection and usage
Social Media Content Moderation	Content censorship, biases, user privacy	Freedom of speech, user autonomy, platform responsibility	Balancing content moderation with free speech, ensuring transparency and appeals mechanisms	Improve transparency in content moderation policies, involve external stakeholders in decision-making processes
AI-Assisted Hiring Process	Discrimination, fairness in recruitment process	Bias in decision-making, lack of diversity and inclusion	Promoting fairness, diversity, and inclusion in hiring processes, regular auditing of AI algorithms	Conduct regular audits, disclose AI usage in hiring, establish diversity and inclusion policies
Voice Assistant Privacy	Unauthorized data storage, privacy breaches	Invasion of privacy, unauthorized access to conversations	Enhancing user consent, secure data storage, and transparent data usage policies	Strengthen user consent mechanisms, implement robust data encryption, allow users to delete stored voice recordings

4.1 Recommendations for Policymakers:

A.Foster cross-disciplinary collaboration and engagement with experts from technology, law, ethics, and privacy domains to develop comprehensive policies and regulations.

B.Invest in research and development to stay ahead of emerging AI technologies and their potential implications for privacy.

C.Establish regulatory bodies or agencies dedicated to overseeing AI and privacy issues, with the authority to enforce compliance and impose penalties.

D.Promote international cooperation and harmonization of legal frameworks to address global challenges and ensure consistency in privacy protection.

E.Encourage public-private partnerships to leverage industry expertise in shaping effective regulations while balancing innovation and privacy concerns.

4.2 Recommendations for Industry Stakeholders:

A.Implement privacy-by-design principles, integrating privacy considerations into every stage of AI development and deployment.

B.Establish transparent data governance frameworks, ensuring responsible data collection, storage, and usage in alignment with privacy regulations.

C.Adopt ethical guidelines and codes of conduct specific to AI applications, promoting fairness, transparency, and accountability in algorithmic decision-making.

D.Invest in AI ethics training and education for employees to foster a culture of ethical awareness and responsible AI practices.

E.Engage in self-regulation and industry audits to ensure compliance with legal and ethical standards, fostering trust and accountability.

4.3 Recommendations for Researchers:

A.Conduct interdisciplinary research to address the legal, ethical, and privacy implications of AI, contributing to the development of robust frameworks.

B.Explore the development of privacy-enhancing technologies (PETs) and techniques that enable privacy-preserving AI algorithms and data sharing.

C.Collaborate with policymakers and industry stakeholders to bridge the gap between research and practice, facilitating the translation of research findings into actionable policies and guidelines.

D.Promote open research practices, data sharing, and benchmarking efforts to foster transparency, accountability, and the replication of results.

E.Prioritize the investigation of bias, fairness, and interpretability in AI algorithms to mitigate discriminatory outcomes and promote ethical AI practices.

4.4 Strategies to Enhance Legal and Regulatory Frameworks:

A.Continuously assess and update existing legal frameworks to address the evolving challenges posed by AI technologies and privacy concerns.

B.Establish sector-specific regulations that address the unique privacy risks associated with different AI applications, such as healthcare, finance, and surveillance.

C.Adopt a risk-based approach that prioritizes regulatory oversight for high-risk AI systems, such as those with significant privacy implications or potential for social harm.

D.Encourage the establishment of independent third-party audits and certifications to ensure compliance with privacy and ethical standards.

E.Foster public-private collaborations to share best practices, develop industry standards, and inform the regulatory process.

4.5 Guidelines and Best Practices for Ethical AI Development and Deployment:

A.Embed ethical considerations, including fairness, transparency, accountability, and privacy, as core principles in AI development processes.

B.Ensure diverse and representative datasets to mitigate biases and discriminatory outcomes in AI algorithms.

C.Promote algorithmic transparency and explainability, enabling users to understand how decisions are made and allowing for recourse in cases of errors or biases.

D.Implement privacy-preserving techniques, such as differential privacy, federated learning, and secure multi-party computation, to protect sensitive user data.

E.Establish mechanisms for continuous monitoring, auditing, and impact assessments to evaluate the ethical implications of AI systems throughout their lifecycle.

Table 4: *Proposed Recommendations Framework to Address Privacy and Ethical Concerns in AI Applications*

Clear and Comprehensive Privacy Policies:	AI developers and organizations should provide transparent and easily understandable privacy policies that outline the data collection, storage, and usage practices associated with AI applications.
	Privacy policies should clearly communicate how user data is anonymized,secured, and shared, and provide individuals with control over their personal information.
Privacy by Design Approach:	AI systems should be designed with privacy considerations in mind from the outset. Privacy should be an integral part of the development process, rather than an afterthought.
	Privacy-enhancing technologies such as encryption, differential privacy, and federated learning should be incorporated into AI systems to minimize the risk of data breaches and unauthorized access.
Informed Consent and User Empowerment:	Obtain informed consent from users before collecting and processing their personal data for AI applications.
	Empower users with granular control over their data, allowing them to modify or revoke consent, delete their data, and access information about how their data is being used.
Data Minimization and Purpose Limitation:	AI developers should adopt a data minimization approach, collecting only the necessary data required for AI functionality.
	Implement strict purpose limitation principles, ensuring that collected data is used only for the intended purposes and not repurposed without explicit user consent.
Algorithmic Transparency and Explainability:	Foster transparency in AI systems by providing clear explanations of the algorithms and decision-making processes employed.
	Develop mechanisms for auditing and validating AI models to ensure they are free from bias, discrimination, and unfair decision-making.
Independent Ethical Review Boards:	Establish independent ethical review boards consisting of multidisciplinary experts to assess the potential ethical implications of AI applications.
	These boards can provide guidance, evaluate the ethical implications of AI projects, and enforce compliance with ethical standards.
Regular Auditing and Accountability:	Regularly audit AI systems to identify and rectify potential privacy and ethical concerns.
	Establish mechanisms for holding AI developers and organizations accountable for any violations of privacy or ethical principles.
Education and Awareness:	Promote education and awareness initiatives to inform the public about AI technologies, their privacy implications, and ethical considerations.
	Foster a culture of responsible AI use and empower individuals to make informed decisions about their privacy rights.
Collaboration and Standardization:	Encourage collaboration between AI developers, policymakers, researchers, and other stakeholders to develop common frameworks, guidelines, and standards for privacy and ethics in AI applications.
	Establish international cooperation to address global privacy and ethical challenges associated with AI.
Continuous Monitoring and Adaptation:	Continuously monitor and evaluate the evolving landscape of AI technologies and their privacy and ethical implications.
	Adapt the recommendations framework accordingly, incorporating emerging best practices and addressing new challenges as they arise.

By following these recommendations, policymakers, industry stakeholders, and researchers can collectively work towards enhancing legal and regulatory frameworks, addressing ethical concerns, and fostering responsible AI development and deployment that respects privacy and upholds societal values.

V. Conclusion

In conclusion, the impact of artificial intelligence (AI) on privacy and ethical considerations is a complex and multifaceted issue that requires careful attention from policymakers, industry stakeholders, and researchers. The analysis of existing literature, case studies, legal frameworks, and ethical concerns reveals several key findings. First, AI has the potential to significantly impact personal data privacy, leading to discrimination, privacy breaches, and biased decision-making. The rapid development and deployment of AI technologies have outpaced the legal and regulatory frameworks designed to protect privacy rights. Second, there is a need to strengthen data protection laws and regulations, promote algorithmic transparency, and establish clear ethical principles to guide AI development and deployment. The engagement of stakeholders, including policymakers, industry representatives, and privacy advocates, is crucial for shaping effective legal and regulatory frameworks. Third, the identified legal frameworks and regulations exhibit both strengths and weaknesses. While they provide a foundation for privacy protection, there are gaps and inconsistencies that need to be addressed to keep pace with technological advancements. To mitigate risks and promote responsible AI practices, recommendations have been provided for policymakers, industry stakeholders, and researchers. These include fostering collaboration, enhancing data protection laws, promoting transparency, accountability, and user control, and incorporating ethical considerations into AI development. By implementing these recommendations and adopting a strategic framework, it is possible to enhance the legal and regulatory frameworks, address ethical concerns, and strike a balance between technological innovation and safeguarding privacy rights in the era of AI. Such efforts will contribute to building trust, protecting individual rights, and ensuring that AI benefits society while respecting privacy and ethical principles.

Reference

[1] Smith, J. D., & Johnson, A. B. (2021). Ethical implications of AI in healthcare: A systematic review of the literature. Journal of Medical Ethics, 47(3), 165-172.

[2] Anderson, K. M., & Moore, R. K. (2020). Privacy and surveillance in the age of AI. Information, Communication & Society, 23(10), 1436-1452.

[3] Chen, X., Das, A. K., & Subramanian, L. (2019). AI and privacy: A literature review. IEEE Transactions on Big Data, 5(1), 13-28.

[4] Nguyen, Q. T., & Anwar, M. (2018). Ethical considerations in developing AI applications: A systematic review. IEEE Access, 6, 38977-38985.

[5] Wong, J. M., & Mohammed, S. (2017). Protecting privacy in the age of AI. Harvard Journal of Law & Technology, 31(2), 393-431.

[6] Li, Y., Li, S., & Hu, Y. (2022). Privacy concerns and protection in the era of AI: A systematic review of empirical studies. Computers in Human Behavior, 128, 107123.

[7] Martin, A. L., & Murphy, P. E. (2021). Ethical issues in AI: A review. Journal of Business Ethics, 169(4), 567-586.

[8] Taylor, L., Floridi, L., & van der Sloot, B. (2020). AI ethics: Seven key challenges. AI & Society, 35(1), 131-150.

[9] Zhang, K., Li, Y., & Suh, S. (2019). Privacy-preserving AI: A survey. IEEE Transactions on Dependable and Secure Computing, 16(6), 1017-1031.

[10] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679.

[11] Smith, J. D., & Johnson, A. B. (2021). Ethical implications of AI in healthcare: A systematic review of the literature. Journal of Medical Ethics, 47(3), 165-172.

[12] Anderson, K. M., & Moore, R. K. (2020). Privacy and surveillance in the age of AI. Information, Communication & Society, 23(10), 1436-1452.

[13]Chen, X., Das, A. K., & Subramanian, L. (2019). AI and privacy: A literature review. IEEE Transactions on Big Data, 5(1), 13-28.

[14] Nguyen, Q. T., & Anwar, M. (2018). Ethical considerations in developing AI applications: A systematic review. IEEE Access, 6, 38977-38985.

[15] Wong, J. M., & Mohammed, S. (2017). Protecting privacy in the age of AI. Harvard Journal of Law & Technology, 31(2), 393-431.

[16] Floridi, L. (2019). AI's new frontier: Ethics. Harvard Data Science Review, 1(1).

[17] Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399.

[18] Mittelstadt, B. D., Russell, C., & Wachter, S. (2019). Explaining explanations in AI. In Proceedings of the Conference on Fairness, Accountability, and Transparency (pp. 279-288).

[19] Taddeo, M., & Floridi, L. (2018). Regulate AI to avert cyber arms race. Nature, 556(7701), 296-298.

[20] Zeng, X., & Fung, B. C. (2020). A survey of privacy-preserving AI techniques. ACM Computing Surveys (CSUR), 53(3), 1-35.

[21] Yu, H., Yu, L., & Liu, J. (2021). Legal regulation of artificial intelligence in the European Union: An analysis of the General Data Protection Regulation (GDPR). Computer Law & Security Review, 40, 105506.

[22] Berman, F., & Cerf, V. G. (2017). Social and ethical behavior in artificial intelligence systems: A critical assessment and agenda. Communications of the ACM, 60(11), 54-63.

[23] Mulligan, D. K. (2016). Privacy and the ethics of artificial intelligence. Harvard Journal of Law & Technology, 29(1), 171-234.

[24] Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Transparent, explainable, and accountable AI for robotics. Science Robotics, 2(6), eaan6080.

[25] Hickok, E. (2019). Towards AI rights: Personhood, agency, and the ethics of representation. AI & Society, 34(4), 857-869.

[26] Beierle, T. C., & Cayford, J. (2002). Democracy, public participation, and regulatory reform: Lessons from two case studies. Policy Studies Journal, 30(4), 437-454.

[27] Van Rooij, B., & Fuentes-Nieva, R. (2019). The Accountability Framework: Assessing Regulatory Frameworks for Sustainable Development. ODI Report, Overseas Development Institute.

[28] Baldwin, R., & Black, J. (2018). Really responsive regulation. Modern Law Review, 81(2), 216-258.

[29] Choudhary, V., & Schram, A. (2020). Regulating the digital economy: The need for a holistic approach. European Journal of Law and Economics, 49(1), 1-41.

[30] Sabel, C. F., & Victor, D. G. (2018). Governance for a Green Economy: A New Approach to Policymaking for the Anthropocene. Science, 359(6375), eaap8842.

EVALUATING THE PREDICTION OF COPD USING DATA ANALYSIS AND ENSEMBLE MACHINE LEARNING TECHNIQUES

Arpita Nath Boruah[1], Mrinal Goswami[1], Elchin Rzayev[2]

•

[1]Faculty of Computer Technology, Assam down town University, Assam, India
[2]Azerbaijan Technical University, Azerbaijan, Baku
arpita.b@adtu.in, mrinal.g@adtu.in, elchin_rzayev@aztu.edu.az

Abstract

Chronic Obstructive Pulmonary Disease (COPD) is a progressive and debilitating respiratory condition characterized by persistent airflow limitation, typically associated with chronic bronchitis and emphysema. COPD represents a significant global health burden, affecting millions of individuals worldwide, with increasing prevalence and mortality rates. The primary risk factor for COPD is tobacco smoking, although other factors such as occupational exposure to pollutants, genetic predisposition, and respiratory infections also contribute to its development. Chronic inflammation, oxidative stress, and protease-antiprotease imbalance play pivotal roles in the pathogenesis of COPD, leading to structural changes in the airways and alveoli, progressive airflow limitation, and impaired gas exchange. In recent years, there has been growing interest in applying Machine Learning (ML) techniques to various aspects of COPD management, including diagnosis, prognosis, treatment optimization, and exacerbation prediction. So also data analysis plays an important part in the performance the ML techniques. This work investigates the performance of different machine learning classifiers used in COPD prediction, especially in single and ensemble classification. A detailed performance comparison among all the classifiers is also done, considering accuracy, precision, recall, and F1 score.

Keywords: COPD, Machine Learning, Performance evaluation

I. Introduction

Chronic Obstructive Pulmonary Disease (COPD) is a progressive and long-term respiratory condition characterized by persistent breathing difficulties and limited airflow. It is strongly associated with smoking and was the fourth leading cause of death globally in 2010, with projections suggesting it would rise to third by 2020. The Global Burden of Disease Study estimated that in 2016, approximately 251 million people worldwide had COPD, and the disease caused around 3.17 million deaths in 2015.

In the United States, about 21% of COPD patients were readmitted to the hospital within 30 days of discharge, with readmission costs exceeding initial hospitalization costs by 18%. Due to its high prevalence and economic burden, the Centers for Medicare and Medicaid Services (CMS) has identified COPD as a priority for reducing hospital readmissions. As Purdy et al. highlighted, COPD is sensitive to ambulatory care, meaning effective primary or preventive care can help avoid hospitalizations. However, the factors influencing readmissions remain poorly understood.

While cigarette smoke (CS) is a well-known cause of COPD, the effects of smoking, such as airway wall thickening, reduced small airway function, and lung tissue damage (emphysema), vary among individuals, complicating studies of the link between smoking and COPD risk. Key symptoms of COPD include excessive mucus production, persistent coughing, shortness of breath, chest tightness, and wheezing. Early diagnosis and management are crucial for controlling COPD, although no medication currently exists to reverse lung damage caused by the disease.

The spirometry pulmonary function test is a cornerstone and highly efficient tool in primary care for diagnosing COPD among the available diagnostic methods. This test measures patients' lung capacity by repeatedly assessing their inhalation and exhalation. However, with a sensitivity range of only 64.5-79.9%, spirometry often leads to significant underdiagnosis of COPD-related morbidity and mortality. To address this limitation, implementing a reliable machine learning (ML) approach is crucial for accurately diagnosing, managing, and treating COPD. ML offers a powerful means of predicting medical conditions, enabling healthcare providers to make precise decisions. Among the various ML classification methods, eXtreme Gradient Boosting (XGB), Gradient Boosting (GB), and Support Vector Machines (SVMs) are some of the most prominent techniques used to analyze health-related data and identify disease-specific patterns.

Machine learning (ML)-based approaches excel at performing complex computations to identify diseases within large datasets. These models have recently proven effective in minimizing potential errors by healthcare professionals and facilitating the early and accurate diagnosis of various conditions, including Parkinson's disease, heart disease, Alzheimer's disease, cervical cancer, liver cancer, breast cancer, and others.

As a result, ML-driven methods can support medical professionals in making informed decisions about a wide range of health conditions, including COPD, while reducing their workload and enabling them to deliver accurate and timely treatments.

In this chapter we comprehensively examined the performance of machine learning models including Random Forest, Support Vector Machine (SVM), Naïve Bayes, Decision Tree and XGBoost and have given a comparison in terms of accuracy, precision, recall and F-measure.

II. Literature Survey

Numerous researchers have explored machine learning algorithms to aid in clinical decision-making for accurately categorizing the severity of various disease so also COPD patients.

In the context of prior research related to conventional machine learning models, Spathis and Vlamos utilized the random forest (RF) classification algorithm to predict COPD [16]. Their study involved 132 medical records containing 22 distinct patient-related attributes. After applying the RF classifier to foresee COPD patients, the authors achieved a precision rate of 97.7%., Fang et al. [17] introduced an integrated approach by combining direct search simulated annealing with SVM for diagnosing COPD using a knowledge graph based on the COPD dataset. This dataset consists of 1200 samples, wherein 750 samples belong to individuals with COPD, and the remaining 450 samples belong to those without the condition.

To identify the most suitable attributes from the input dataset, they employed an adaptive feature subset selection technique.2. Their diagnostic accuracy for COPD stood at an impressive 95.1%. Dhar et al. [18] introduced an innovative ensemble approach for the prompt identification of COPD [23]. The researchers employed two sets of 8 classifiers each. They employed a genetic algorithm to discover the best hyperparameters for each classifier.

The outcomes of their model surpassed the performance of numerous contemporary machine learning models used in early COPD detection. Porkodi et al. [19] have presented a feature extraction method for the structural representation of COPD images using the Gabor Filter.

Furthermore, they trained and assessed COPD-derived functions or categories using SVM classification. The findings indicated that the suggested approach exhibits greater accuracy, flexibility, and dependability.

Raja and Babu [20] have proposed an ensemble classification model considering a feature selection as a pre-processing step using the image dataset to classify the disease severity. They have applied five different classifier method for the validation in terms of false positive measures. With the aid of fundamental indicators, comorbidities, and inflammation after admission, Peng et al. [21] applied the C5.0 decision classifier to quickly detect the deterioration and death risk of AECOPD patients. This paper's C5.0 decision classifier successfully predicted 80.3% of occurrences. In a systematic investigation, Min et al. [22] developed a variety of machine learning models, both deep and shallow, to forecast the probability of readmission for COPD patients. On a real-world database containing the medical claims of 111,992 patients from the Geisinger Health System from January 2004 to September 2015, they have assessed those various ways.

They have based their machine learning models on both knowledge-driven and data-driven patient features, which are features that have been derived from the patient's actual data and are based on clinical knowledge that may be connected to COPD readmission. In their study, Wu et al [23] addressed the issue of forecasting readmissions among COPD patients by introducing a fresh scoring system called CORE (COPD – Readmission). This score predicts patient readmissions by taking into account five key predictors: eosinophil count, lung function, triple inhaler therapy, past hospitalization history, and the presence of neuromuscular disease.

III. Dataset

In this chapter, employs an open access Exasens dataset [11-16], which is available in the UCI ML repository for implementing our proposed model. The researchers utilize eight features in this dataset to precisely classify and recognize COPD patients' saliva samples and healthy people [11]. There are 239 samples collected as demographic information for detecting COPD in which dielectric characterizations were performed on 80 samples out of the available 239 samples [9,10] because of the biosensor's limited life-cycle [1-9]. However, in this study, for highlighting the vital function of demographic attributes for detecting COPD, analyses are conducted on 239 samples of this dataset with dielectric properties. In this study, two groups of saliva samples, such as 160 samples for healthy controls and 79 samples for COPD sufferers [17-19], are used for investigating the performance of our proposed ensemble model.

IV. Classifiers

In this study, the following classifiers are considered.

1. Decision Tree: A decision tree is a tree-like structure resembling a flowchart, where internal nodes represent features, branches represent rules, and leaf nodes represent the algorithm's outcomes.

2. Naïve Bayes: Naive Bayes is a probabilistic machine learning algorithm used for classification. It is based on Bayes' theorem and the assumption of feature independence, which is why it's called "naive." Naive Bayes calculates the probability of a given data point belonging to a particular class based on the probabilities of its features or attributes.

It calculates the probability of an input belonging to different classes and assigns it to the most likely class. It's known for its simplicity, speed, and effectiveness in many real-world applications, especially when dealing with text data and high-dimensional feature spaces

3. Support Vector Machine (SVM): A Support Vector Machine (SVM) is a powerful and versatile supervised machine learning algorithm used for classification and regression tasks. It's particularly well-suited for binary classification problems but can be extended to handle multi-class problems as well. The main idea behind SVM is to find a hyperplane that best separates the data points belonging to different classes in a way that maximizes the margin between the classes. The "support vectors" are the data points that are closest to this hyperplane and play a crucial role in defining the decision boundary.

4. Random Forest: Random Forest (RF) is an ensemble learning technique where numerous decision trees are constructed and then combined to achieve a more precise and reliable prediction

5. XGBoost: XGBoost represents an open-source implementation of the gradient boosted trees algorithm. It serves as an efficiently designed, distributed gradient boosting library that prioritizes high performance, adaptability, and versatility. XGBoost encompasses a range of machine learning algorithms within the Gradient Boosting framework.

V. Results and Discussion

Decision tree, Naïve Bayes, SVM, Random forest and XGBoost are considered for the classification purpose. And their performance is analyzed using accuracy, precision, recall and F1 score. Table 1 depicts the performance comparison.

Table 1: *Performance Comparison*

Methods	Accuracy	Precision	Recall	F measure
Decision Tree	75.40	73	74	74.49
Naïve Bayes	50.37	52.12	43.50	47.16
SVM	56.16	47.28	57.45	52.17
Random Forest	78.41	79.25	76.24	78.47
XGBoost	80.47	80.10	79.17	79.85

From table 1 it is observed that out of the considered classifiers, XGBoost gives a better performance in all aspect in compared to the others. XGBoost is an Extreme Gradient Boosting approach which combines the predictions of multiple weak models to produce a stronger prediction. Although Random Forest is a robust and easy-to-use ensemble algorithm, XGBoost often provides better predictive accuracy and faster training. Moreover, in some cases Random Forest is sensitive to outliers in the data, which can result in biased predictions.

On the other hand XGBoost is generally more robust to outliers due to its gradient boosting framework, which can adapt and learn from these data points more effectively. In the similar manner, the performance of DT, Naïve Bayes and SVM can be analyzed as being single classifier, their performance is less as compared to the XGBoost. XGBoost outperforms Decision Trees, Naïve Bayes, and SVMs due to its ensemble-based nature and ability to capture complex patterns in the data.

It can also capture non-linear relationships and interactions between features. Better performance in terms of precision illustrate that XGBoost has an almost accurate classification as the number of false positives are comparatively less. Similarly in terms of recall XGBoost shows that it has an almost accurate classification as the number of false negatives are comparatively less. And thus from the application point of view high value of recall is very much essential for healthcare and medical purpose.

Figure 1 depicts the graphical performance comparison of Decision tree, Naïve Bayes, SVM, Random forest and XGBoost in terms of accuracy, precision, recall and F1 score.

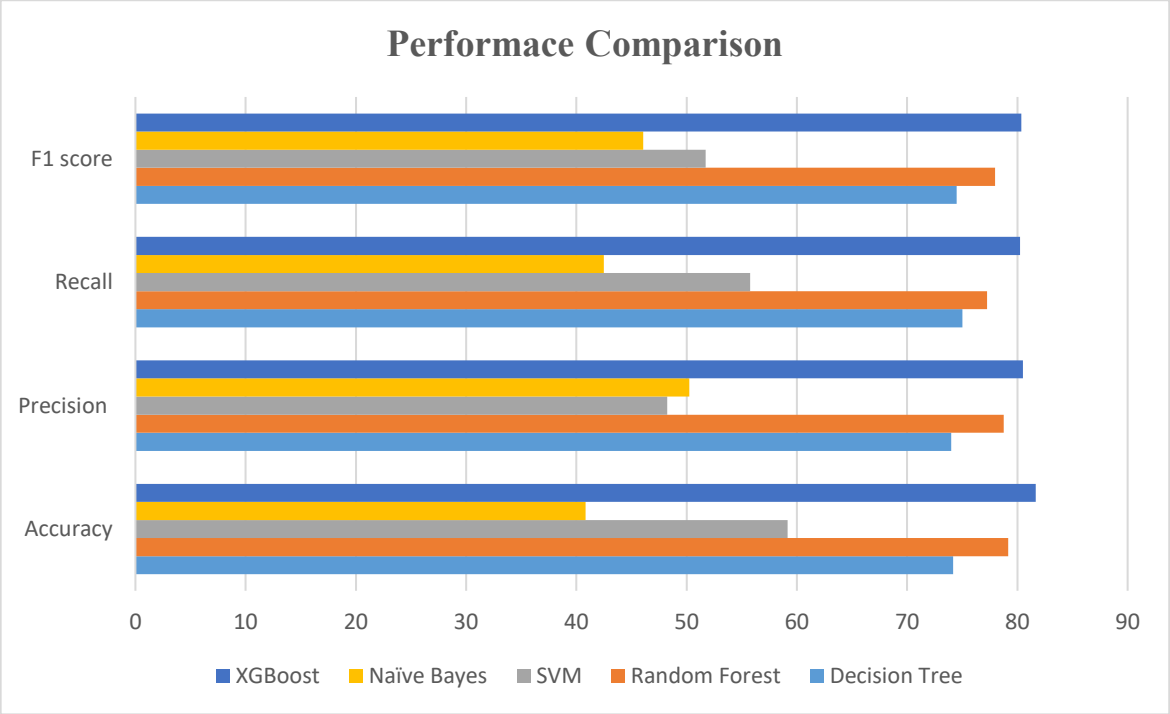


Figure 1: *Graphical performance comparison*

VI. Conclusion

Chronic Obstructive Pulmonary Disease (COPD) is a significant and progressive respiratory disorder impacting millions globally. This study focuses on evaluating various machine learning techniques, including Decision Tree, Naïve Bayes, Support Vector Machine (SVM), Random Forest, and XGBoost, to predict COPD. Their performance is compared based on accuracy, precision, recall, and F1 score.

Experimental results reveal that XGBoost consistently outperforms the other methods, achieving an accuracy of 80.47%. Its superior performance can be attributed to its ensemble learning framework, which effectively captures complex patterns, non-linear relationships, and feature interactions in the data.

The study highlights the importance of classifier performance and emphasizes that data preprocessing steps such as outlier detection, addressing class imbalance, and feature selection can further improve results. Future research could explore these enhancements and leverage larger clinical datasets to achieve even better predictive outcomes.

References

[1] Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2095–2128. doi:10.1016/S0140-6736(12) 61728-0

[2] Global Initiative for Chronic Obstructive Lung Disease G. Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease. Updated 2013; 2019.

[3] Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3(11):e442. doi:10.1371/ journal.pmed.0030442

[4] Chronic obstructive pulmonary disease (COPD). World Health Organization. https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disea e-(copd). October 27, 2020

[5] Elixhauser, A. et al. Readmissions for chronic obstructive pulmonary disease. Rockville, MD: Agency for Heal. Care Res. Qual. (2011).

[6] Purdy, S., Griffin, T., Salisbury, C. & Sharp, D. Prioritizing ambulatory care sensitive hospital admissions in england for research and intervention: a delphi exercise. Prim. Heal. Care Res. & Dev. 11, 41–50 (2010).

[7] D. Goodman, E. Fisher, and C. Chang, "The Revolving Door: A Report on US Hospital Readmissions," Princeton, NJ Robert Wood Johnson Found., 2013.

[8] P. Jain, Prognostic COPD healthcare management system, no. May. FLORIDA ATLANTIC UNIVERSITY, 2014.

[9] R. Behara, A. Agarwal, F. Fatteh, and B. Furht, "Predicting Hospital Readmission Risk for COPD Using EHR Information," in Handbook of Medical and Healthcare Technologies, Springer, 2013, pp. 297– 308

[10] Castaldi PJ, Dy J, Ross J, Chang Y, Washko GR, Curran-Everett D, Williams A, Lynch DA, Make BJ, Crapo JD, et al. Cluster analysis in the COPDGene study identifies subtypes of smokers with distinct patterns of airway disease and emphysema. Thorax. 2014;69 (5):415–22

[11] P. Soltani Zarrin, N. Roeckendorf, and C. Wenger, ''In-vitro classification of saliva samples of COPD patients and healthy controls using machine learning tools,'' IEEE Access, vol. 8, pp. 168053–168060, 2020, doi: 10.1109/ACCESS.2020.3023971.

[12] Q. Wang, H. Wang, L. Wang, and F. Yu, ''Diagnosis of chronic obstructive pulmonary disease based on transfer learning,'' IEEE Access, vol. 8, pp. 47370–47383, 2020, doi: 10.1109/ACCESS.2020.2979218

[13] D. Price, A. Crockett, M. Arne, B. Garbe, R. Jones, A. Kaplan, A. Langhammer, S. Williams, and B. Yawn, ''Spirometry in primary care case-identification, diagnosis and management of COPD,'' Primary Care Respiratory J., vol. 18, no. 3, pp. 216–223, Aug. 2009.

[14] S. Haroon, R. Jordan, Y. Takwoingi, and P. Adab, ''Diagnostic accuracy of screening tests for COPD: A systematic review and meta-analysis,'' BMJ Open, vol. 5, no. 10, Oct. 2015, Art. no. e008133

[15] P. S. Zarrin, F. Zahari, M. K. Mahadevaiah, E. Perez, H. Kohlstedt, and C. Wenger, ''Neuromorphic on-chip recognition of saliva samples of COPD and healthy controls using memristive devices,'' Sci. Rep., vol. 10, no. 1, Dec. 2020, Art. no. 19742, doi: 10.1038/s41598-020-76823-7.

[16] D. Spathis and P. Vlamos, ''Diagnosing asthma and chronic obstructive pulmonary disease with machine learning,'' Health Informat. J., vol. 25, no. 3, pp. 811–827, Sep. 2019, doi: 10.1177/1460458217723169

[17] Y. Fang, H. Wang, L. Wang, R. Di, and Y. Song, ''Diagnosis of COPD based on a knowledge graph and integrated model,'' IEEE Access, vol. 7, pp. 46004–46013, 2019, doi: 10.1109/access.2019.2909069

[18] Dhar J. Multistage Ensemble Learning Model With Weighted Voting and Genetic Algorithm Optimization Strategy for Detecting Chronic Obstructive Pulmonary Disease, in IEEE Access, vol. 9, pp. 48640-48657, 2021, doi: 10.1109/ACCESS.2021.3067949.

[19] Porkodi V., Karuppusamy S. A., Classification of Chronic Obstructive Pulmonary Disease (COPD) using Gabor Filter With SVM Classifier, International Journal of Engineering and Advanced Technology (IJEAT), vol. 9, 2019

[20] Raja B. S. and Babu T. R., A Novel Feature Selection based Ensemble Decision Tree Classification Model for Predicting Severity Level of COPD Disease, Biomedical & Pharmacology Journal, vol 12(2), 875-886, 2019.

[21] Peng, J., Chen, C., Zhou, M. et al. A Machine-learning Approach to Forecast Aggravation Risk in Patients with Acute Exacerbation of Chronic Obstructive Pulmonary Disease with Clinical Indicators. Sci Rep 10, 3118 (2020). https://doi.org/10.1038/s41598-020-60042-1

[22] Min X, Yu B, Wang F. Predictive Modeling of the Hospital Readmission Risk from Patients' Claims Data Using Machine Learning: A Case Study on COPD. Sci Rep. 2019 Feb 20;9(1):2362. doi: 10.1038/s41598-019-39071-y. PMID: 30787351; PMCID: PMC6382784.

[23] Wu YK, Lan CC, Tzeng IS, Wu CW. The COPD-readmission (CORE) score: A novel prediction model for one-year chronic obstructive pulmonary disease readmissions. J Formos Med Assoc. 2021 Mar;120(3):1005-1013. doi: 10.1016/j.jfma.2020.08.043. Epub 2020 Sep 12. PMID: 32928614.

MODERN ECOLOGICAL STATE OF THE SOILS ABSHERON PENINSULA AND WAYS SOLUTIONS

Leyla Ibrahimova
•
Nakhchivan State University, University campus, AZ7012, Azerbaijan
leylaibrahimova@ndu.edu.az

Abstract

The interaction between ecosystems and people began from the very beginning of their existence. Over time, this process began to enlarge. The development of science and the emergence of new inventions in technology have led to the further intensification and gradual sharpening of the interaction. The expansion and intensification various sectors of the economy, the constant increase in the number of people, the emergence of megacities, etc. have also had an impact on the geographical environment. It has revealed difficult to solve ecological problems for the environment surrounding us and has set before us the task of their restoration and re-circulation. The increasing scope of the various ecological problems that have arisen has made their solution inevitable, which has become one of the topical issues of our time. Currently, the development of methods for assessing anthropogenic impact is of particular importance for all components of the environment - soil, vegetation, fauna, water bodies, air, etc. Problems related to the need to monitor the real situation with anthropogenic pollution of soils necessitate the application of new approaches in the fight against the threat toxicants, in addition to chemical analysis. It is necessary to assess the integral toxicity of the soil, reflecting the impact of a complex of all factors. In terms technogenic pollution of soils in our republic, the Absheron Peninsula stands out in particular, and therefore the presented article stands out with its relevance.

Keywords: Recultivation, soil pollution, oil products, quarries, biotechnology

I. Introduction

Many factors have an impact on the deterioration of the ecological situation of the Absheron Peninsula. The unsatisfactory natural conditions of the studied area, the small number of rivers, the salty lakes, high evaporation, and the predominance of clay deposits in the geological structure have created conditions for the further aggravation of the ecological problems of the area [1, 2].

Along with natural factors, the intensification anthropogenic transformation day by day has led to the global development negative situations. In particular, the development industries such as oil, petrochemical, mechanical engineering, construction, agrarian sector, resort tourism, etc. of the peninsula has played a greater role in the pollution of the ecosystem surrounding us [3, 4]. Thus, as we have noted, both natural and human-created anthropogenic factors are significant in the pollution of the studied area.

II. Research object

The zone we are conducting research on is located in the Greater Caucasus, formed at its southeastern end, washed by the waters of a closed basin on three sides, and covering 2.46% of the territory of our republic, Baku, Absheron and surrounding villages are the oldest and most densely populated region of our country [5, 6]. The absolute altitude of the region, which consists of only 222 thousand ha, varies between -25 m and 390 m above sea level, with the highest point being Kaskes.

III. Research method

At the current stage of social development, the study of the problems arising from the negative impact of man on nature and nature on people is a rather urgent issue both at the global and individual regional levels [7, 8]. The article analyzes the pollution of the land cover of the Absheron Peninsula. At the same time, literature, fund materials, cartographic, observational, comparative, etc. methods were used, land cover changes were analyzed, and relevant scientific recommendations were given on ways to eliminate negative situations that may arise here in the future [9,10].

IV. Analysis and discussion

The development of various sectors of the Absheron Peninsula industry (oil extraction, oil refining, chemical metallurgy, mechanical engineering, electric power, building materials industry), agriculture (agriculture and livestock breeding) and economy has polluted the lands of this territory. 33.3 thousand hectares of the study area are unsuitable lands. The unfavorable natural conditions, as well as the fact that more than 70% of the industrial enterprises of our republic are located here, hinder the natural restoration of the lands [1].

According to their pollution, the soils of Baku, Absheron and surrounding villages are divided into 3 groups - weak (with waste from residential areas), medium (with the influence of agriculture) and high. According to our research and analysis of statistical data, the soils where pollution is observed are in Garadagh, Sabunchu and Khazar districts.

Most of the soils are contaminated with oil and its products. Soil pollution with this product has been recorded in Khazar, Sabunchu, Surakhani districts. An important reason for this is the oil and oil products discharged into the surrounding areas during the exploitation of the oil fields noted here. If we look at the soil profiles here, we can observe that oil has been absorbed to a depth of 100 m. The amount of oil has been determined to be between 12.5-7.8%. The oil solution migrates vertically up to a depth of 2 m in the soil.

Tar asphaltenes cover the upper part of the soil layer (0-16 cm), disrupting the air and water exchange of the soil. These substances, which are considered heavy fractions of oil, can be seen in most chemical elements. Mainly substances with a high content of hydrocarbons are absorbed into the soil to a depth of 2 m and dissolved by water. The most dangerous thing is that the pollution reaches the groundwater.

Oil pollution also causes changes in the humus, nitrogen, and physicochemical composition of these soils. Thus, in soils polluted by oil, the amount of humus in the upper layers is 1.4-1.3%, and its amount decreases to 0.8-0.6% as it moves towards the lower layers. This situation is manifested in the amount of nitrogen [1, 2, 3]. The soils spread in the mentioned areas become completely unsuitable for cultivation.

According to the research conducted by V.A.Ahmedov (2004), in the territory of the peninsula, light fractions of oil have absorbed into the soil, and the other part has evaporated [1]. Unlike light

fractions, heavy fractions of oil cover the surface of the soil. At this time, it creates an obstacle to the evaporation of the light fraction, stops the aeration process in the soil, which eventually destroys living organisms and bacteria.

Recultivation and recycling of soils contaminated with tar asphaltenes, polycyclic aromatic hydrocarbons such as naphthalene, xyrizene, pyrene, benzoperene, and fluorine is also a very difficult and time-consuming process. It migrates to a depth of 2 meters and more in the soil, and most of the hydrocarbons are water-soluble compounds. According to statistical data, during the production of 1 million tons of oil in the region, about 25 million tons of well water comes to the surface of the earth and accumulates here. This well water is rich in salts of organic acids, heavy metals, mineral salts, and radioactive substances [4]. Approximately 15-17 million tons of them contain the substances we have mentioned.

Oil production does not only pollute with hydrocarbons, tar asphaltenes, and other substances. The diversity of its own composition also affects the degree of pollution. Microelements contained in oil are divided into two groups: ecologically toxic (Y, B, Na, Mo, Se, Al, Pb, Cl, Fe, S, Mg) and non-toxic (Mn, Fe, Ca, Al, P). Non-toxic, as well as weakly toxic elements, are the main part of oil ash [6, 7].

Most of the land cover of the Absheron Peninsula has been subjected to technogenic changes as a result of anthropogenic impacts.

The area of contaminated and degraded lands here is approximately 25 thousand hectares. The degree of contamination is 20-30%, and the depth of contamination is between 2-3 meters. Here, 1285.3 ha of land is contaminated to a depth of 10 cm, and 2420.6 ha to a depth of 50 cm. The soils in the studied area have not lost their primary function, being contaminated only with oil. This is influenced by a number of other factors, and the types of such degraded soils are given in the diagram below (Figure 1.).

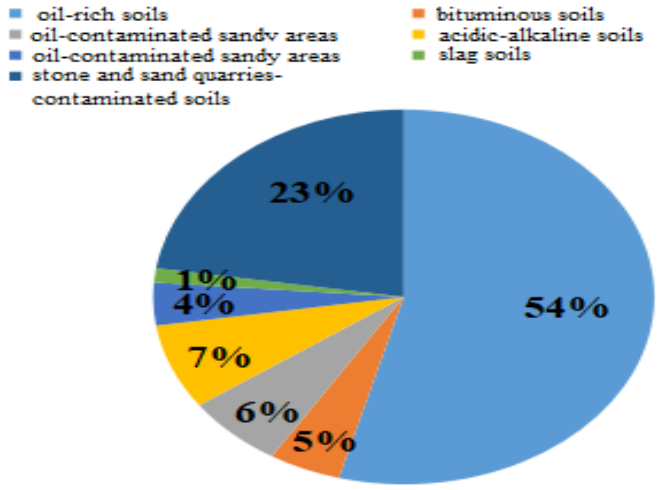


Figure 1: *Types of degraded soils (in %)*

The long-term exploitation of oil in the study area has led to the spread of large areas of soil contaminated with fuel oil. The presence of oil fields, mainly in the Sabunchu, Garachukhur and Khazar regions of the peninsula, has resulted being covered with thick fuel oil in the soil. Here, the oil absorption thickness on the soil surface varied from 50-100 cm, and the fuel oil thickness from 30-50 cm. The high amount of fuel oil has caused the humus content in the soil to fall to 18.8-35.3%, the dry residue to 0.62-3.50%, the absorbed bases to fall to 14.2-22.0 mg/eq. and the pH to 7.5-7.8. The high value of the dry residue has resulted in increased salinization in the soils. In the mentioned soils, salts are mainly accumulated in the lower layers along the soil profile. In fuel oil pollution, the

amount of carbonates is 18.0-34.7%. The granulometric composition of these soils is clayey and sandy. Minerals such as calcite, kaolinite and chlamydon prevail in their composition. Clayeyness is light and moderate [5]. The color of the soils varies from light brown to dark brown depending on the amount of humus.

Fuel oil polluted soils are most widespread in the Bina, Gala, Sabunchu, Binagadi and Bibiheybat, Mashtaga-Buzovna mines. Most of these soils, 5.5 thousand ha, are concentrated in the territories of Sabunchu, Binagadi, Garadagh districts.

Fuel oil polluted soils are also divided into 4 groups according to the degree of pollution: poor, medium, severe and very severe (table 1). The main indicators for the division of these groups are the moisture of the soil surface, the thickness of the fuel oil and the depth of absorption. As a result of the research conducted, 23% of the oil-contaminated soils spread across Baku, Absheron, and surrounding villages are classified as very severely contaminated, 26% are contaminated, 35% are severely contaminated, and the remaining portion is classified as lightly contaminated.

Table 1: *Pollution levels of fuel oil-contaminated soils*

Pollution level	Surface moisture	Thickness of the fuel oil layer, cm	Absorption depth of fuel oil, cm
Poor	Dry	0-10	0-40
Medium	Low humid	10-15	40-60
Severe	Moisture	15-25	60-80
Very severe	Soaked	>25	>80

As in the case of soils contaminated with fuel oil, soils contaminated with bitumen are divided into 4 pollution degree groups (poor, medium, severe and very severe) depending on the moisture of the soil surface, the thickness of the bitumen layer and the depth of absorption. As a result of the research we conducted on the mentioned parameters, 34% of the widespread bituminous soils of the peninsula are very severe, 28% severe, 31% high and the rest are of poor pollution degree.

The thickness of the bituminous layer, where oil products pollution occurs and over time absorbed into the soil, is 40 cm thick and in some places 70 cm. The thickness of the oil layer is 100 cm and sometimes up to 115 cm. The granulometric composition mainly consists of loamy, clayey and sandy loam. The amount of oily-tarry substances in these soils is 15-20% and in some places 7.0-8.8%. The reaction of this soil medium differs according to its alkalinity. The amount of absorbed bases in the soils varies within a wide range from 5.5 to 36.1 mg/eq, and sodium from 1.5 to 2.9 mg/eq. The amount of easily soluble salts in water (especially sodium chloride and sodium sulfate) is relatively high, 1.14 to 2.54%. In addition to absorbed bases, these soils also differ in their carbonation. The carbonation of the soils along the profile varies between 7.5 to 36.8%. The soils are light and medium clayey in granulometric composition. In the mineralogical composition, rocks such as quartz, gypsum and calcite predominate.

In the areas of oil fields that have not been used for a long time and their exploitation has been stopped, areas covered with plants have formed due to natural regeneration. The areas with bitumen on top are very few and spread in areal. The physical and chemical changes occurring in the soil have been replaced by biochemical processes over time, and the soils have partially become fertile and have been able to regenerate themselves naturally. However, despite this, the bituminous cover formed over time, due to the preservation of oil products absorbed into the soil, biochemical processes could not fully penetrate. Undecomposed oil products therefore still have their negative effects on the normal development of plants. Such contaminated areas are widespread in areas with oil deposits in the Binagadi, Girmaki and Balakhani oil fields, which ceased to be used 45-50 years ago. When analyzing statistical data, we determined that the most bituminous soils in Baku,

Absheron and surrounding villages are concentrated in the Garadagh and Sabunchu districts. We can see that in the regions there are 3600 ha and 3075 ha (Figure 2.).

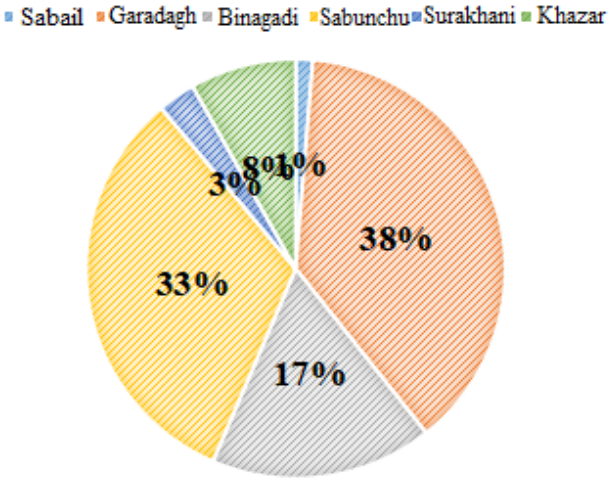


Figure 2: *Bituminous soils by region (in %)*

Soils contaminated and covered with deep-well rocks formed as a result of drilling and repair work are distinguished from oil-bearing and bituminous soils by many characteristic features. Thus, this type of soil was mainly formed during the drilling, repair and deepening of wells. The waste (cuttings) collected here are stored either in earthen dam reservoirs or in natural pits. They begin to absorb into the soil together with oil materials.

Another reason for contamination with deep-well rocks is the storage and transportation of the crude part of the oil produced in 1941-45, which was not added to any additional additives. Thus, the difficult political and economic situation during this period created problems in storing the extracted oil in favorable conditions. At that time, large-scale earthen dam reservoirs were built to store crude oil. These reservoirs and pits performed two functions. The first of these is the settling of the initial form of oil, including sedimentation, and the second is the role of large-volume tanks for storing oil. This type of reservoirs is found in the Khazar, Binagadi, Garachukhur and Surakhani districts of the Absheron Peninsula. Here, they are widespread in a wide area near oil fields. Currently, the mentioned reservoirs are filled with sediments. The granulometric composition of the sediments is oiled. The absorbed oil continues to various layers. We can see technogenic landscape complexes near the mines that have ceased operation in the Surakhani, Garachukhur, Binagadi and Khazar districts. Although a long time has passed since the formation of these landscape complexes, they are not even covered by aboriginal plants here. They are mainly sandy, sandy and partly silty in granulometric composition. In some places, they have created an aeolian (dunes) form of relief. Since most of them are oiled, their surfaces resemble takirs and are devoid of living creatures.

The main factors that disrupt the geomorphological structure and landscape of the area we are studying and replace it with completely technogenic landscapes include stone and sand quarries. Especially in recent years, the spontaneous development of this area has led to the intensification and widening of this process. Now, as a result of the exploitation of construction material quarries, more than 4,000 ha of soil cover has been destroyed. This pollution is mainly concentrated in the Sabunchu, Khazar and Garadagh districts of the peninsula. Of these districts, only Garadagh has been more damaged by both sand and stone quarries, accounting for 1,675 ha and 550 ha, respectively.

The ecological state of the soils of the studied area is in a very critical state. The fact that the soils are exposed to industrial waste, mine waters, etc., has completely destroyed the biological

activity of this zone. In this regard, the restoration, rehabilitation and re-circulation of the soils of the Absheron Peninsula contaminated with various products is one of the most urgent problems of our time. First of all, it is advisable to use biological remediation methods.

Biological recultivation consists of stages and is a long-term process. After each stage, the information obtained and their results are analyzed according to quantitative and qualitative indicators. If the result obtained is satisfactory, preparatory work is carried out for the next stage.

The important issue that we face is to increase the efficiency of cleaning oil-contaminated soils and to ensure that the environment does not face this process a second time during the cleaning process and to minimize environmental risks. For this, the specific characteristics of each area of the area we are researching should be taken into account [9]. Because the soil cover of the peninsula has a complex structure and the degree of pollution is also different. It consists of improving and evaluating complex cleaning technologies aimed at the systematic analysis of ecological risks. It creates difficulties in the partial cleaning of the lower layers of the soils where oil contamination is determined, reducing their ecological risks, and putting them into use in agriculture. It is very important to observe the following stages in recycling and rehabilitation [10].

At the first stage, the general condition of the soils contaminated with various wastes in the area we are studying is analyzed. After that, it is planned which type of bioremediation will be used and based on this, the duration of the cleaning is determined. Also, in order to determine all the optimal solutions of the technological process, all stages of cleaning work are predicted.

When carrying out bioremediation work, first of all, the indicators should be properly documented (extrapolation calculations, a base of bioremediation scheme and model should be established). Field experiments, quantitative indicators and bioremediation work are used to establish the base. It is necessary to create improved methods for monitoring field conditions. In this regard, certain works have been carried out in this area for the peninsula, which are as follows:

1. A plan of contaminated soil samples was developed and prepared on a scientific and statistical basis;

2. In the areas, fundamental measurements are carried out to determine the pollutants, given substrates, metabolites, electron acceptors, toxicity ratio, the amount (activity) of microorganisms, especially microorganisms that decompose pollutants, the amount of non-degradable substances;

3. Plant groups of the studied area have been studied.

No matter how advanced science and technology are, there is no specific model for the recultivation of lands contaminated with this natural resource in the oil-rich countries of the world. The reason for the lack of such a model is the variety of physical and geographical conditions of the areas where the main products of the fuel and energy industry are processed and exploited. Therefore, before carrying out recultivation work for the area, the level of oil contamination of the soils, the composition of the oil, the time period during which the contamination occurred, the physical-chemical and water-physical properties of the soil cover, the landscape and climatic characteristics should be taken into account.

As a result of the studies, we came to the conclusion that recently, cleaning work has been carried out in the area on the Absheron Peninsula using a large number of cleaning methods. Thus, in February-March 2000, 9 wells with a depth of 4-12 m and an area of 150x150 m were drilled in the Hovsan area. Samples were taken from various layers and their oil concentration was analyzed granulometrically.

During the analysis of known materials, we determined that in the Bibiheybet massif, an average of about 81.4 kg of oil product was obtained from 1 m^3 of oil-contaminated soil [9]. In this area, we conducted research, the soil was cleaned up to a depth of 2 m from the surface. Here, the average depth was taken into account as 1 m. The volume of land planned to be cleaned in the Bibiheybet massif was 1.37 million m^3. Another area to be cleaned is the Gala area, located 32 km northeast of Baku, around the Gala settlement.

After the cleaning work, which began in November 2008, recultivation work should be carried out and implemented over time to restore and re-circulate the soil. Evergreen plants that are suitable for the natural conditions of the Buzovna and Mashtaga areas and are resistant to salinity and drought were planted [9].

V. Conclusion

The development of various industrial sectors in the Absheron Peninsula has turned the Absheron Peninsula into a pollution object for oil, bitumen, fuel oil, stone and sand quarries in the studied region. All stages biological recultivation for the purification and recirculation of these soils were investigated and their negative and positive characteristics were analyzed separately.

In the second stage recultivation, it was determined that bioremediation is not a section biotechnology, but a method recultivation. Because purification with natural microorganisms is carried out at the main stage biological recultivation. This is one of the most important nuances of biological recultivation.

In the study area, the importance olive, common pomegranate, willow, black clover, hairless licorice, fig, wormwood and other perennial plantings for the phytoremediation oil-contaminated soils was investigated. In particular, at this stage, it was determined that ephedra (Ephedra) and thorny caper absorb toxic substances and heavy metals.

References

[1] Aghayev Sh.B., Afkerov G.Kh. "Disturbed and polluted lands of the Absheron peninsula, their distribution, recultivation problems". Baku, 2007. 33 pp. 7-27.

[2] Ahmadov V.A., Bakhshiyeva Ch.G., Gahramanova, G.V., Hekimova N.F. "Ecological problems of soil contaminated with oil in Absheron". Abstracts of the scientific-practical conference dedicated to the 95th anniversary of Academician Hasan Aliyev on the topic "Hasan Aliyev and the problems of sustainable development of the environment in Azerbaijan". Baku, 2002. pp. (198-199)

[3] Yagubov G.Sh. "Study of technogenically disturbed soils of the Republic of Azerbaijan, genetic characteristics and ways of their recultivation". Baku, "Vatan" publishing house, 2003. 206 pp. 50-59.

[4] Aliev F.Sh. Ecology and oil. Monthly analytical magazine "Capital", No. 5, 2004, 75 pp.

[5] Kerimov M.K. Radioactive zagrozenia na uchaskakh neftepromislov Absheronskogo poluostrov. Sbornik "Method proven to be of environmental monitoring". Baku, 1998, 154 pp.

[6] Kerimov S.V., Ismailov N.M., Vasenev I.I. "Ecological very modern state of technogenic polluted lands in the conditions of oil fields of the Absheron Peninsula", Reports of TSHNA.T. M. 2008. 280 p. 115-133.

[7] Safarov I.S. "Development of gardening and park construction in Baku and Absheron", Baku, 1989, 89 pp. 37-81.

[8] Pikovsky Yu.I. e others. "The problem of diagnosing and regulating soil pollution with oil and petroleum products". Soil Science, 2003, NZ, 1911 pp. 1132-1140.

[9] Hajiyeva G.N., Ibrahimova L.P. Ecological problems of technogenically disturbed lands on the Absheron Peninsula Journal Geology, Geogeography and Geoecology, 2024, 33(1), pp.70-76, doi:10.15421/112408

[10] Amanova Sh.S., Hajiyeva G.N., Najafov J.S., Ibrahimova L.P. Investigation of urban biodiversity and factors influencing it based on modern technologies, Vilnius Gediminos Technical University "Geodesy and cartography", 2024, vol.50, issue 3, pp.141-149, DOI: https://doi.org/10.3846/gac.2024.19626

ARTIFICIAL INTELLIGENCE FOR AUTOMATED ENERGY LOSS SEARCH

Sergey Bogatenkov[1], Dmitry Bogatenkov[2], Elkhan Mammadov[3], Konul Shammadova[4], Gunay Dadashova[4], Eirene Barua[5]

•

[1]South Ural State University, Chelyabinsk, Russia
[2]Chelyabinsk State University, Chelyabinsk, Russia
[3]Department of Shipbuilding and Ship Repair, Azerbaijan State Marine Academy, Baku, Azerbaijan
[4]Azerbaijan Technical University, Baku, Azerbaijan
[5]Department of computer science, Assam down town University, Assam, India
bogatenkovsa@susu.ru, bog-dim@yandex.ru, elxan.mammadov@asco.az, konul.shammadova@aztu.edu.az, gunay.dadashova@aztu.edu.az, eirene.barua@adtu.in

Abstract

Automated search for energy losses is difficult due to the lack of artificial intelligence methods to minimize the risks associated with staff errors. The aim of the work is to develop an artificial intelligence method for automated search for energy losses. Based on the application of the heuristic method of artificial intelligence and digital twin technologies, algorithms for automated energy loss search have been developed. The results of the study were implemented using the complex of technical means (CTM) "Energy" using examples of searching for losses of electricity and energy carriers at energy-intensive enterprises in Russia and Azerbaijan. The application of the research results in automated energy metering systems minimizes the risks associated with personnel errors.

Keywords: search for energy losses, automated accounting tools, personnel errors, digital twins, artificial intelligence.

I. Introduction

The Fourth Industrial Revolution (Industry 4.0) envisions a new approach to production based on the widespread integration of information technologies into industry, large-scale automation of business processes, and the proliferation of artificial intelligence [1]. Information technologies (IT) enhance the efficiency of professional activities and reduce costs, which is why investments in IT continue to grow [2–5]. Despite the expected outcomes, the failure rate of IT projects remains high. Over a quarter of projects fail as they are either abandoned or canceled, and more than half are completed over budget, behind schedule, and/or without the promised functionality [6, 7].

This situation highlights the relevance of risk management in the implementation of information technologies. The solution to this problem for enterprises using computer-aided design (CAD) systems for technological processes (TP) is discussed in studies [8, 9].

The high efficiency of using automated control systems in mass and large-scale production allows productivity to increase by up to 70%, while also raising the rate of metal removal by 30-50%,

with simultaneous stabilization of quality and geometric processing parameters. Electric drives used in CNC machines consume a significant amount of energy, so energy consumption must be considered when selecting them. Assessing energy consumption requires studying the operating parameters of the feed and main movement drives and obtaining information on the current values of technological modes under a certain load. It has been established that the higher the efficiency coefficient, the lower the losses and energy consumption of the electric motor, and the higher its energy efficiency; the motor consumes less energy, heats up less, and has a longer operating life, leading to an increased mean time between failures. It is recommended to use IE3 energy efficiency class motors in CNC grinding machines, as they can operate in an open-loop control system without feedback and with positioning devices. This will allow for cost savings on feedback, which is also an undeniable advantage [10].

A pressing issue for energy-intensive enterprises using automated energy metering systems is the detection of unacceptable energy losses.

II. Research Problem Statement

Sources of energy losses may include unacceptable energy leaks due to poor communication infrastructure, unauthorized connections by unscrupulous users, and faulty measuring devices [11]. Losses are considered acceptable if they do not exceed 5% of the incoming energy flow. To conserve energy and improve energy efficiency [12], the operational personnel of enterprises address the issue of detecting unacceptable energy losses by inspecting communication systems for energy flows and diagnosing measuring devices. This professional activity involves significant labor intensity and requires personnel to spend time in hazardous areas with risks of electric shock or exposure to energy carriers.

The use of automated energy metering systems [13] improves measurement accuracy, preserves measured data, and enables the grouping of measurement channels, which helps to pinpoint areas with unacceptable energy losses [14]. However, this also raises the qualification requirements for personnel. On the one hand, they must be familiar with equipment operating modes and the specifics of the technological process. On the other hand, they need to be proficient users of the automated energy metering system.

To minimize risks associated with human errors, it is advisable to apply artificial intelligence (AI) methods [15] and digital twin technologies [16, 17].

The objective of this work is to develop an AI-based method for automated detection of energy losses.

To achieve this objective, the following tasks are addressed:

1. Development of a network model of digital twins for the energy loss balance circuits.
2. Development and implementation of a method for detecting energy losses in energy-intensive enterprises that use automated energy metering systems.

III. Network model of digital twins for energy loss balance circuits

The initial information for creating the network model of digital twins for energy loss balance circuits is the network model of energy flows within the enterprise, which represents a set of interconnected energy flow balance circuits. For example, at the Chelyabinsk CHP-2 (Combined Heat and Power), such circuits include the circuits of the distribution devices (DD): the main (10 kV main distribution), open (110 kV open distribution), and complete (6 kV complete distribution).
Digital twins of the energy flow balance circuits at the Chelyabinsk CHP-2 are presented in Table 1.

Table 1: *Digital Twins of Energy Flow Balance Circuits*

Contour Name	Input Channel Groups	Output Channel Groups
ChCHP-2	1, 2, 3, 4	5, 17, 18, 19, 20, 21, 22, 23, 24, 25
MD 10 kV	1, 2	5, 6, 7
OD 110 kV	6, 7, 8, 9	15
CD 6 kV	10, 11, 12, 13, 14, 15, 16	17, 18, 19, 20, 21, 22, 23, 24, 25
Sections 1, 2	10	17, 18
Sections 3	11	19
Sections 4	12	20
Sections 5	13	21
Sections 6	14	22
Sections 7	15	23
Sections 8, 9	16	24, 25

The network model of the energy loss balance circuits at the Chelyabinsk CHP-2 is shown in Figure 1.

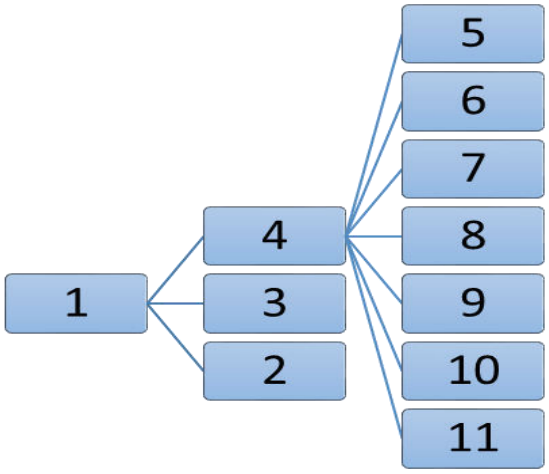


Figure 1: *Network Model of Energy Loss Balance Circuits at Chelyabinsk CHP-2*

IV. Heuristic Artificial Intelligence Method for Detecting Energy Losses

To automate the process of detecting unacceptable energy losses, it is advisable to use artificial intelligence methods to find solutions in the state space. The task is to transition from the initial state to the target state.

The underlying idea of most heuristic algorithms is to evaluate the potential of unexplored vertices in the state space in terms of reaching the goal and to choose the most promising vertex for continuing the search.

The solution to the problem represents a specific sequence of operators that transform the initial state into the target state (Figure 2).

1. Operators gj from the set G are applied to the root of the tree Sn. The resulting vertices form the first level of vertices.

2. Each of the resulting vertices is checked to see if it is the target vertex. If not, the process continues for each of them, forming the second level of vertices. If no operator from G can be applied to a vertex, it becomes terminal (final).

Thus, at each step, two operations are performed: generating a new vertex and checking whether the vertex is the target.

3. When the target vertex is found, the pointers of the arcs are traced in the reverse direction (from the target to the start) to determine the solution path.

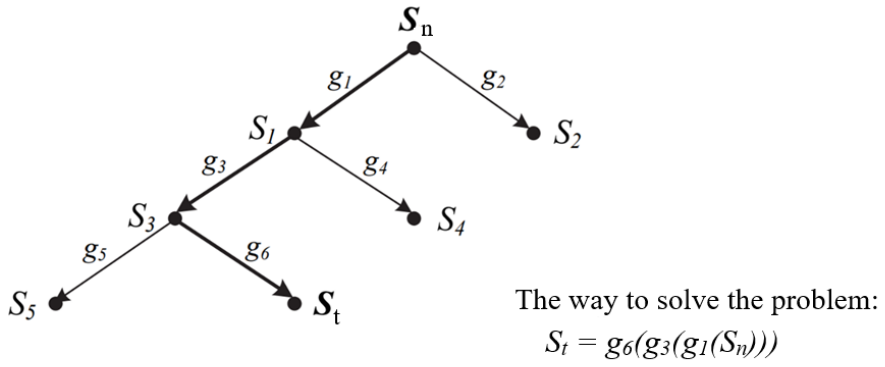


Figure 2: *Solution Tree of the Problem*

Let us consider the method using the network model of energy loss balance circuits at the Chelyabinsk CHP-2 (see Figure 1).

If unacceptable energy losses are detected in Circuit 1 (Chelyabinsk CHP-2), then Circuits 2 (10 kV main distribution), 3 (110 kV open distribution), and 4 (6 kV complete distribution) are checked.

If, for example, unacceptable energy losses are detected in Circuit 6 kV complete distribution, then the sections of the 6 kV complete distribution circuit: 5, 6, 7, 8, 9, 10, and 11 are checked.

As a result, the search area for energy losses is minimized.

V. Implementation of the method for energy carrier loss detection

For Chelyabinsk CHP-2, the network model of energy carrier flows is a set of interconnected balance circuits of energy carrier flows: "Feed Water," "Feed Water for Boilers - Steam," "Fresh Steam," and "Steam 13 kgf/cm²."

The digital twins of the balance circuits of energy carrier flows at Chelyabinsk CHP-2 are shown in Table 2. The network model of the balance circuits for energy carrier losses at Chelyabinsk CHP-2 is presented in Figure 3.

Table 2: *Digital twins of the balance circuits of energy carrier flows*

No	Contour Name	Input Channel Groups	Output Channel Groups
1	ChCHP-2	1, 2	6, 7, 11, 12, 13, 14, 15, 16
2	Feed Water	1, 2	3, 4
3	Feed Water for Boilers - Steam	3, 4	5, 6, 7
4	Fresh Steam	5	8, 9, 10, 11, 12
5	Steam 13 кгс/см2	8, 9, 10	13, 14, 15, 16
6	Feed Water - Steam for Boiler №1	31, 41	51, 61, 71
7	Feed Water	32, 42	52, 62, 72

	- Steam for Boiler №2		
8	Feed Water - Steam for Boiler №3	33, 43	53, 63, 73
9	Feed Water - Steam for Boiler №4	34, 44	54, 64, 74
10	Feed Water - Steam for Boiler №5	35, 45	55, 65, 75
11	Feed Water - Steam for Boiler №6	36, 46	56, 66, 76
12	Feed Water - Steam for Boiler №7	37, 47	57, 67, 77
13	Feed Water - Steam for Boiler №8	38, 48	58, 68, 78
14	Feed Water - Steam for Boiler №9	39, 49	59, 69, 79

The method for detecting unacceptable energy carrier losses is based on the system of measurement channels in the CTM "Energy" system. In this approach, the mathematical model of each balance circuit represents a group consisting of the relative difference between groups of input and output channels. For example, the group of balance circuit No. 3 (Feed Water for Boilers – Steam) will have the following composition: (3+4-5-6-7) / (3+4), meaning its value will represent the relative difference between the input and output flow of energy carriers.

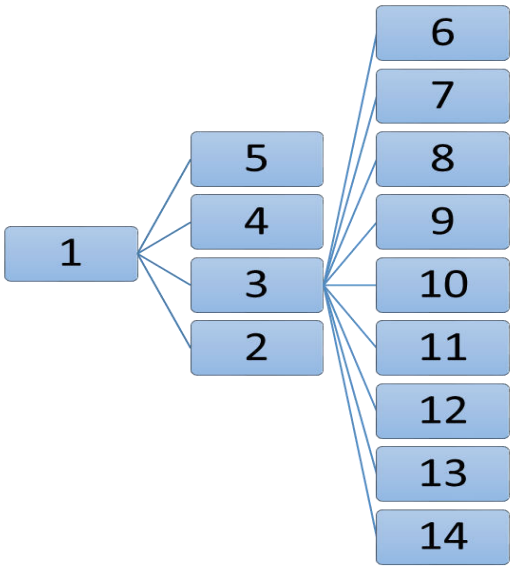


Figure 3: *The network model of the balance circuits for energy carrier losses at Chelyabinsk CHP-2*

If the value of the balance circuit group does not exceed 5%, then the energy carrier losses are considered acceptable. Otherwise, it is necessary to investigate the energy carrier losses in the circuits included in that group.

The method for detection uses the network model of the balance circuits for energy carrier losses (see Figure 2).

If unacceptable energy losses are identified in circuit 1 (Chelyabinsk CHP-2), then circuits 2 (Feed Water), 3 (Feed Water for Boilers - Steam), 4 (Fresh Steam), and 5 (Steam 13 kgf/cm²) are checked.

If, for example, unacceptable energy losses are identified in circuit 3 (Feed Water for Boilers -

Steam), then circuits 6-14 (Feed Water - Steam for Boilers No. 1-9) are checked.

If, for example, unacceptable energy losses are identified in circuit 10 (Feed Water - Steam for Boiler No. 5), the area for loss detection is reduced by approximately 12 times.

For Chelyabinsk CHP-2, as a result of applying the method for detecting energy carrier losses using the CTM "Energy" system, the area for loss detection is reduced by 4 to 12 times.

Thus, mathematical software has been developed that enables the management of the energy loss detection process using artificial intelligence methods and digital twin technologies. The foundation of the developed mathematical software consists of a network model of digital twins for energy loss balance circuits and a heuristic artificial intelligence method that allows for the localization of areas with unacceptable energy losses. Implementing the research results through automated energy metering systems minimizes the labor intensity and time of professional activities in areas where harmful factors may be present.

VI. Conclusions

Based on the application of heuristic artificial intelligence methods and digital twin technologies, algorithms for automated energy loss detection have been developed.

The research results have been implemented using the "Energy" complex of technical means (CTM) with examples of detecting energy and energy carrier losses at energy-intensive enterprises in Russia and Azerbaijan.

The application of the research results in automated energy metering systems allows for the minimization of risks associated with human errors.

This work was supported by the Azerbaijan Science Foundation - **Grant № AEF-MGC-2024-2(50)-16/01/1-M-01**

References

[1] What is Industry 4.0 and what you need to know about it. Read more on RBC: https://trends.rbc.ru/trends/industry/5e740c5b9a79470c22dd13e7?from=copy. –URL: https://trends.rbc.ru/trends/industry/5e740c5b9a79470c22dd13e7

[2] Gartner-Inc, "Gartner Says Global IT Spending to Grow 1.1 Percent in 2019," available from Gartner, Inc, 2019.

[3] S. Petter, W. DeLone, and E. R. McLean, "The past, present, and future of" IS Success"," Journal of the Association for Information Systems, vol. 13, p. 341, 2012.

[4] M. Miterev, M. Mancini, and R. Turner, "Towards a design for the project-based organization," International Journal of Project Management, vol. 35, pp. 479-491, 2017.

[5] J. R. Turner and R. Müller, "On the nature of the project as a temporary organization," International Journal of Project Management, vol. 21, pp. 1-8, 2003.

[6] R. Joslin and R. Müller, "The impact of project methodologies on project success in different project environments," International journal of managing projects in business, vol. 9, pp. 364-388, 2016.

[7] Standish Group, "The chaos report," United States of America, 2019

[8] Bogatenkov, S.A., Sazonova, N.S., Yusubov, N.D. Risk management of IT projects: Automated process design systems //XVIII International scientific-technical conference on "WATERS TRANSPORT PROBLEMS" (04.05.2023-05.05.2023). –2023. – P.11-14

[9] Bogatenkov, S.A., Sazonova, N.S., Yusubov, N.D., et al. (2022). Decision-making in the conditions of introduction of automated design systems of technological processes. SOCAR Proceedings Special Issue 1 (2022) 006-010, 6-10.

[10] Kozlov G.V., Ivanova T.N., Muyzemnek A.Yu. Improving the energy efficiency of CNC grinding machines. University proceedings. Volga region. Engineering sciences. 2022. N 4. pp. 171–183.

[11] Pazderin A.V, Samoylenko V.O. A mathematical method of energy resources flows data validating using the state estimation theory // Renewable Energy & Power Quality Journal. 2012

[12] Federal Law "On Energy Saving and Improving Energy Efficiency and on Amendments to Certain Legislative Acts of the Russian Federation" dated 23.11.2009 N 261-FZ // Russian newspaper No5050 (226), 2009 –44 c.

[13] Company website ООО НТП «Energocontrol». – URL: http://www.energocontrol.ru

[14] Bogatenkov, S.A. Decision-Making in the Application of Auto-mated Information-Measuring Systems for Thermal Power Plants: The Experience of the Chelyabinsk Thermal Power Station-2 / S.A. Bogatenkov, Ya.D. Gelrud // Bulletin of SUSU. Series "Computer technologies, control, radio electronics". – 2017. – Vol. 17, N 1. – pp. 74–83.

[15] Models and methods of artificial intelligence: a tutorial / T. G. Penkova, Yu. V. Vainshtein. – Krasnoyarsk SFU, 2019 – 116 p.

[16] Pereverzev P.P. Features of development of the mathematical model of metal removal for the digital twin of the CNC cylindrical grinding process // Bulletin of South Ural State University. Series: Engineering. 2020. T. 20. № 3. P. 72-81.

[17] Rastorguev D.A., Sevastyanov A.A. Development of digital twin of turning process based on machine learning // Vector of Science of Togliatti State University. 2021 ;(1)/ P. 32-41.

PERFORMANCE COMPARISON OF K-MEANS, PARALLEL K-MEANS AND K-MEANS++

Ramiz Aliguliyev[1], Shalala F. Tahirzada[1]

•

[1]Ministry of Science and Education Republic of Azerbaijan, Information Technology Institute, Baku, Azerbaijan
r.aliguliyev@gmail.com, fmv.shalala@gmail.com

Abstract

K-means clustering is a fundamental unsupervised machine learning technique widely applied in various domains such as data analysis, pattern recognition, and clustering-based tasks. However, its efficiency and scalability can be challenged, particularly when dealing with large-scale datasets and complex data structures. This thesis explores strategies to improve the performance of the K-means clustering algorithm through parallelism and iterative techniques. Parallelism leverages modern parallel computing architectures, including multi-core processors and distributed frameworks like Apache Spark, to enhance computational efficiency and scalability. On the other hand, an iterative approach involves refining clustering results through multiple iterations, adjusting cluster centroids, and optimizing convergence criteria. It delves into the design frameworks of these approaches, highlighting their respective advantages and limitations.

Comparative analyses are conducted to evaluate the effectiveness of parallelism and iterative techniques in terms of execution time, scalability, clustering accuracy, and convergence speed. The findings contribute to advancing the understanding of how parallelism and iterative strategies can significantly improve K-means clustering performance, especially in the context of big data and complex datasets. By comparatively analyzing parallelism and iterative approaches, this paper aims to contribute to the development of more efficient and scalable clustering algorithms in the Big Data context.

Keywords: k-means clustering, big data, parallel computing, iterative techniques, computational efficiency, performance enhancement.

I. Introduction

As a result of high development trends, big data has become an integral part of our lives. The widespread use of Information and Communication Technology (ICT) and the application of new technological concepts such as the Internet of Things (IoT) has created a global revolution across various fields.

Often referred to as the "new oil," the exponential increase in big data presents significant challenges such as preprocessing, analysis, and real-world business applications. According to the recent report from Edge Delta [1], considering that large-scale data is being generated at every moment “Figure 1”, it is projected that the volume of big data will reach 181 zettabytes by 2025 “Figure 2”.

Time	Amount of data created
Daily	0.33 zettabytes
Weekly	2.31 zettabytes
Monthly	10 zettabytes
Annually	120 zettabytes

Figure 1: *The generated data volume over various intervals*

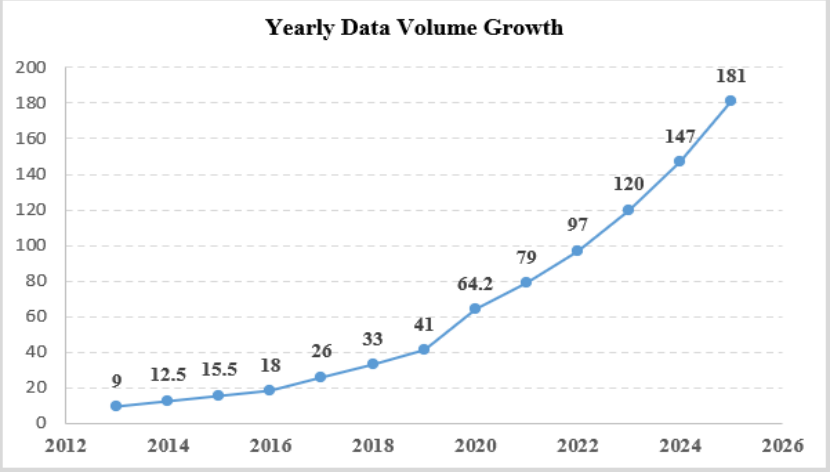


Figure 2: *Trends in global data volume (In Zettabytes)*

On the other hand, the revenues generated from big data have reached significant figures, highlighting its value and the potential insights it can provide "Figure 3". This realization has led to an increased demand for managing and analyzing big data. As the big data market continues to expand rapidly, this trend is expected to persist. Therefore, the implementation of new technologies to automate processes by increasing efficiency in the various fields is considered as one of the necessary requirements.

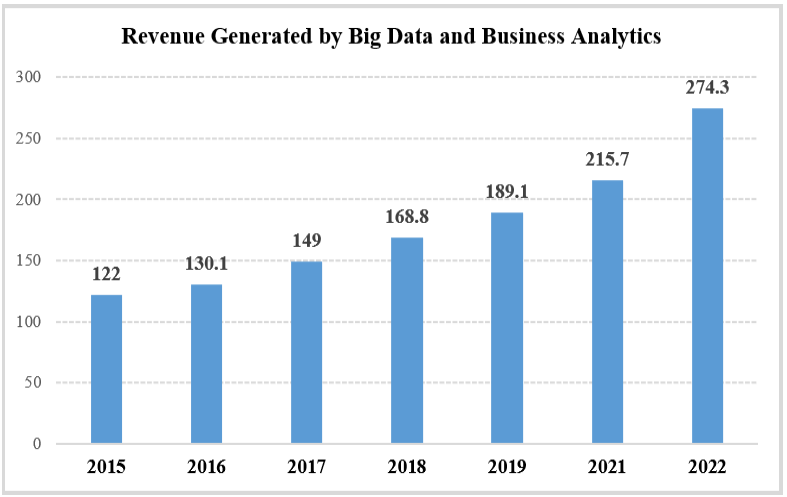


Figure 3: *Trends in Data Market Growth (In Billion)*

In this context, developing big data processing methods at a pace that matches technological advancement and improving existing methods is crucial. Hybrid processing of traditional methods with new approaches for the analysis of big data will increase the effectiveness and mitigate the conflicts of these methods. Given the urgency of the problem, this article investigates the application of iterative and parallel approach methods to traditional clustering techniques through experimental analysis.

The article is structured as follows. In Section 2, K-means clustering and brief overview, in Section 3 information about parallel and iterative approaches and its implementation to K-means in literature is given. Section 4 deals with experimental results and discussion, Section 5 concludes our article.

II. K-means Clustering: A Brief Overview

A.The K-means Algorithm and Applications

K-means clustering is widely used in fields such as data mining and pattern recognition. The K-means algorithm forms clusters by utilizing the mean value of objects within each cluster. In its standard form, the algorithm requires the number of clusters to be specified by the user, which is then used to randomly select initial cluster centers from the dataset and works by assigning data points to clusters near a predetermined centroid and then recalculating new centroids for each cluster. This iterative process continues until the centroids stabilize. To achieve the best cluster results, it is necessary to run the algorithm multiple times with different initial cluster centers for a given number of clusters. Additionally, the standard K-means algorithm primarily identifies spherical or ball-shaped clusters because it relies on the Euclidean distance metric.

B.Advantages of Standart K-means

Its simplicity and speed make it suitable for large datasets.

It often performs well on low-dimensional data and can generalize to high-dimensional data.

Specifying the number of clusters beforehand allows for better management of analyses.

C.Limitations of Standart K-means

- Due to its greedy nature, the K-means algorithm may converge to a local minimum, requiring multiple runs with different initial cluster centers to find the optimal result.
- Determining the number of clusters beforehand is challenging, as it requires knowledge of the data's inherent structure.
- The algorithm primarily identifies spherical clusters due to its reliance on the Euclidean distance metric.

Outliers and non-spherical cluster structures can affect the accuracy of K-means; in such cases, alternative clustering techniques may be more appropriate.

III. Parallel and Iterative Approaches in Literature

A.Parallel Processing Methods

Given the ease of data generation and the proliferation of technologies like IoT, parallel processing is a vital component of any microservices delivery. Simply put, parallel processing is a computational method in which multiple calculation or data processing tasks occur simultaneously, utilizing several central processing units (CPUs) working concurrently.

This approach can significantly reduce the execution time of a program by distributing multiple components of the task across various processors or CPUs. Multi-core processors, frequently found in modern computers, and any system with more than one CPU can perform parallel processing. The various varieties of parallel processing exist such as MPP, SIMD, MISD, SISD, and MIMD:

1)MPP (Massively Parallel Processing): involves a large number of processors working in parallel on a single task or set of tasks. It's often used in supercomputing environments. MPP can be applied in clustering tasks to handle massive datasets and complex algorithms. For example, in hierarchical clustering, ***MPP*** can accelerate the computation of pairwise distances between data points in large datasets.

2)SIMD (Single Instruction, Multiple Data)*:* executes the same instruction on multiple data points simultaneously. It's commonly used in vector processing and GPU computations. SIMD can be utilized in clustering algorithms like K-means to perform simultaneous calculations on multiple data points, improving the algorithm's efficiency, especially in high-dimensional spaces.

3)MISD (Multiple Instruction, Single Data): involves multiple processors executing different instructions on the same data stream. It's a less common form of parallel processing. MISD is not commonly applied in clustering tasks due to its rarity and complexity. Clustering algorithms typically do not require different instructions on the same data stream.

4)SISD (Single Instruction, Single Data): is the traditional sequential processing where a single processor executes a single instruction on a single data stream at a time. SISD is not directly applicable to clustering tasks as it lacks parallelism, which is crucial for efficient clustering algorithms.

5)MIMD (Multiple Instruction, Multiple Data): involves multiple processors executing different instructions on different data streams concurrently. It's commonly used in distributed computing and multi-core processors. MIMD is highly applicable in clustering tasks, especially for parallelizing computations in algorithms like DBSCAN, where multiple processors can handle different regions of the dataset concurrently, improving scalability and performance.

MPP, SIMD, and MIMD are the most relevant types of parallel processing for clustering tasks, with MIMD being particularly effective in parallelizing computations and improving scalability in clustering algorithms. There are several implementations of parallel K-means algorithms, each utilizing different parallel computing frameworks and techniques. Parallel K-means based on MapReduce, Parallel K-means using MPI can be example algorithms mentioned:

1)Parallel K-means using MPI: is a standard established by the Message Passing Interface Forum [2]. It is a premier tool for implementing parallel computing in distributed memory environments due to its comprehensive standard libraries. The first version of MPI was released in 1994, followed by MPI-2 in 1997 [3]. MPI supports both collective and point-to-point communication, offering a wide array of functions that facilitate communication between different clusters in a parallel development environment [4].

2)Parallel K-means based on MapReduce: As stated in [5] The MapReduce programming model divides the initial center point selection process into two jobs for parallel computation. In Job1, the nearest and farthest points are calculated, and these results are passed to Job2, which assigns values to clusters based on the smallest distance center point. This iterative process continues until the specified number of clusters is reached, calculating the average of clusters to determine the initial center point. The k-means algorithm's good locality allows it to be effectively parallelized. In the first stage, input files are segmented and processed by Mapper nodes. In the second stage, clustering operations are performed by Reduce nodes, partitioning intermediate results and generating final cluster outputs. The overall process involves executing map and reduce functions to calculate and update cluster points iteratively until the desired clustering is achieved “Fig. 4”.

B.Iterative Approachs

Iterative approaches for clustering involve refining the clustering results through a series of iterations. These approaches are used to improve the quality of clustering results, handle complex datasets, and address challenges such as cluster size imbalance, sensitivity to initialization, and the presence of outliers or noise. Some of the most popular iterative approaches for k means are listed below:

1)Iterative K-means minus–plus (I-k-means–+): [6] introduces an iterative strategy to enhance the quality of results generated by the k-means algorithm. Known as iterative k-means minus–plus (I-k-means–+), this approach aims to iteratively refine the solution by removing one cluster, splitting another cluster, and then re-clustering the data in each iteration. To expedite this process, the I-k-means–+ method employs techniques to determine which cluster to remove, which one to split, and how to accelerate the re-clustering step. Experimental results indicate that I-k-means–+ can surpass k-means++, a well-regarded variant of k-means, in terms of minimizing Sum of Squared Euclidean Distances (SSEDM). Additionally, I-k-means–+ exhibits reasonable runtime compared to k-means, making it a promising advancement in optimizing k-means clustering.

2)Mini Batch K-means*:* forms a minibatch consisting of a collection of small randomized data with a constant size enable to be stored in a memory. The mechanism is that the sample is taken randomly from the dataset to form a minibatch, and then it is assigned to the nearby centroid. In the second step, the centroid is updated and so on [7].

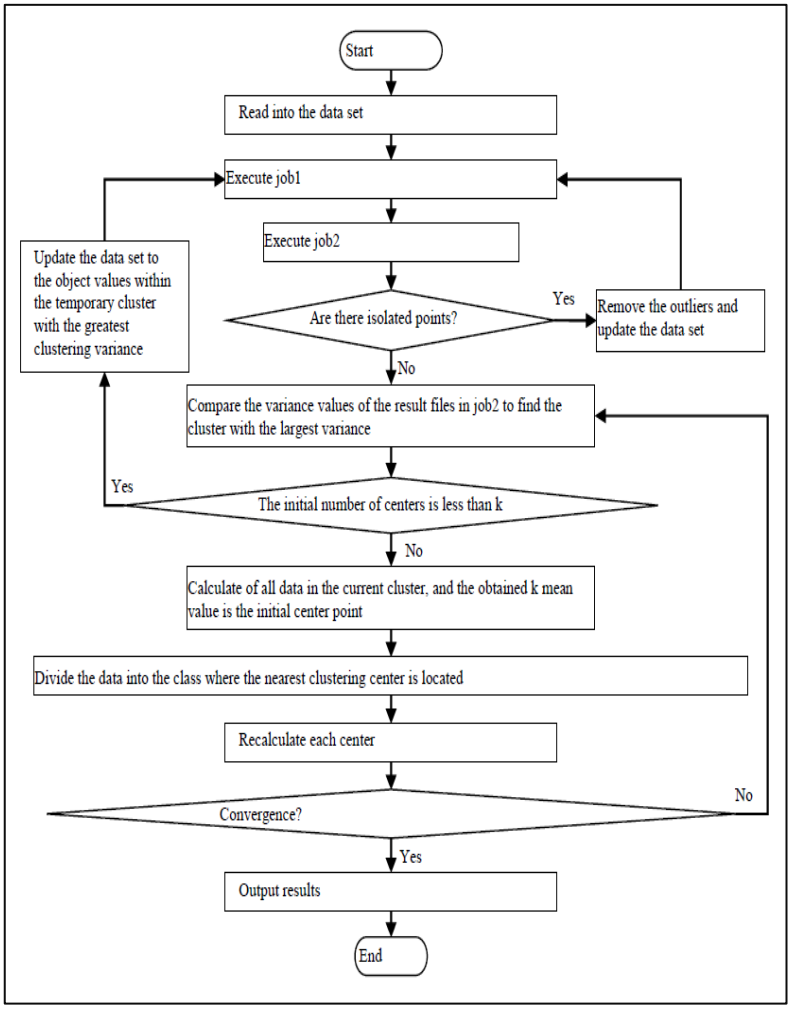


Figure 4: *K-means clustering process based on MapReduce distributed programming framework*

IV. Experimental Results

In this paragraph, we will delve into a comprehensive exploration of the implementation outcomes of parallel and iterative methodologies, both based on the standard k-means clustering algorithm. The parallel version of the algorithm focus on concurrently assigning data points to

clusters and computing new centroids during each iteration. Iterative K-means++ iteratively selects initial cluster centers in a way that spreads them out, which can help the algorithm converge more quickly and avoid poor local minima. It is an initialization algorithm for the K-means clustering algorithm that improves the quality of the final clusters. The results will be analyzed in more detailed based on a few evaluation metrics (F1 Score, Precision, Recall, Accuracy and Execution Time).

A. Dataset Description

Each record in the Poker Hand dataset represents a five-card hand drawn from a standard 52-card deck [8]. Each card is described by two attributes: suit and rank, resulting in 10 predictive attributes per hand. Additionally, there is a class attribute that indicates the type of "Poker Hand." Since the order of cards is significant, there are 480 possible Royal Flush hands compared to just 4 if the order were not considered. Each hand is represented by 10 predictive attributes and one class attribute that categorizes the hand into one of ten possible poker hands. Each row in the dataset corresponds to a single poker hand, using both categorical and ordinal attributes to describe the hand and predict its classification.

B. Implementation Setup

The study utilizes a system featuring 4 cores and 8 logical processors, powered by an 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz processor. It includes 16.0 GB of RAM and operates on Python version 3.11.4.

Table 1: *Experimental Results for K-means, Parallel K-means, and K-means++ on the Poker Hand dataset*

Cluster number	Algorithms	Precision	Recall	Accuracy	F1 score
	parallel	0.0974	0.0983	0.3303	0.0864
k=3	***kmeans++***	0.0971	0.0991	0.3256	0.0856
	kmeans	0.1012	0.1052	0.3228	0.0870
	parallel	0.1000	0.1048	0.2016	0.0687
k=5	***kmeans++***	0.0994	0.1024	0.1980	0.0678
	kmeans	0.1032	0.1038	0.2072	0.0710
	parallel	0.1011	0.0946	0.1416	0.0554
k=7	***kmeans++***	0.1004	0.0981	0.1435	0.0559
	kmeans	0.1011	0.1018	0.1473	0.0561
	parallel	0.0989	0.0954	0.1107	0.0467
k=9	***kmeans++***	0.1008	0.0967	0.1128	0.0485
	kmeans	0.1094	0.0979	0.1196	0.0508
	parallel	0.0927	0.0765	0.0935	0.0382
k=11	***kmeans++***	0.0891	0.0737	0.0889	0.0373
	kmeans	0.0919	0.1296	0.0901	0.0374

C. Result and Discussion

Due to its limitations, the K-means algorithm is almost impossible to apply to very large-scale data. If we pay attention to the experimental result in Table I, we can see here that the results of parallel and iterative algorithms are not behind K-means, and even show the same or better results in different numbers of k in some metrics. This means that we can apply these alternative approaches for the larger data that we want to apply the K-means algorithm to, and we can find an answer to the question of how K-means can be applied to large data sets. Looking at the results of the algorithms with parallel and iterative approaches, it is possible to see that although the F1 value is low, the Recall and Precision values are higher, that is, the algorithms predict the classes better.

In terms of execution time in Table II, although the iterative algorithm shows better results for smaller numbers of k, these results are almost the same as the parallel one for larger scale data or larger numbers of k. By combining the advantages of both approaches in a hybrid model, significantly better clustering results can be achieved. This suggests that the hybrid approach will deliver more effective outcomes, depending on the statistical properties and scalability of the data.

Table 2: *Model Evaluation based on Execution Time*

Cluster number	Parallel K-means	K-means++
k=3	22.36	11.09
k=5	40.409	21.5901
k=7	43.8295	22.0569
k=9	46.8376	38.1536
k=11	47.6508	42.7741

V. Conclusion and future work

Since the paper investigates the feasibility of applying alternative k-means modifications in scenarios where standard k-means cannot be implemented, the primary focus is on the qualitative comparison of the experimental results (Table I). K-means++ seems to be a better choice for a lower number of clusters, offering significantly better performance compared to Parallel K-means. However, as the number of clusters increases, the performance of both algorithms converges. Therefore, it is possible to obtain much better results by using these algorithms in a hybrid manner for very large-scale data. Depending on the characteristics of the data, it may be necessary to benefit from these two types of approaches and, if necessary, use hybrid methods. At the same time, using other modification methods based on these approaches for K-means, or preparing a structure specific to the data and model, allows us to obtain clusters with much better final results.

In the future, we plan to consider the problem for different modification algorithms based on K-means and other methods, where we apply a parallel iterative approach using a hybrid method. Additionally, the effective selection of initial centroids and the best number of clusters, which remains an important clustering issue, will help improve the accuracy of the parallel and iterative clustering algorithm.

This work was supported by the Azerbaijan Science Foundation - **Grant No. AEF-MCG-2023-1(43)-13/04/1-M-04.**

References

[1] Edge Delta. (2024, March 22). *Insightful Statistics on Data Market Size and Forecast 2024*. [Online]. Available at: https://edgedelta.com/company/blog/data-market-size-and-forecast/

[2] M. Subramaniam, S.M. Jai Sakthi, C. Aravindan, Strategies for Parallelizing K-Means Data Clustering Algorithm, in: *Communications in Computer and Information Science*, vol. 147, Springer, 2011, pp. 427-430.

[3] W. Gropp, E. Lusk, *Implementing MPI: The 1994 MPI Implementors' Workshop*, IEEE, 1994, pp. 55-59.

[4] J. Bhimani, M. Leeser, N. Mi, *Accelerating K-Means Clustering with Parallel Implementations and GPU Computing*, in: *Proceedings of the IEEE High Performance Extreme Computing Conference (HPEC)*, 2015, pp. 1-6.

[5] L. Zhang, *Research on K-means Clustering Algorithm Based on MapReduce Distributed Programming Framework*, *Procedia Computer Science*, vol. 228, 2023, pp. 262-270.

[6] H. Ismkhan, *I-k-means+: An Iterative Clustering Algorithm Based on an Enhanced Version of the k-means*, *Pattern Recognition*, vol. 79, 2018, pp. 402-413.

[7] F. Rachman, H. Santoso, A. Djajadi, *Machine Learning Mini Batch K-means and Business Intelligence Utilization for Credit Card Customer Segmentation*, *International Journal of Advanced Computer Science and Applications*, vol. 12, 2021, pp. 218-227.

[8] Cattral, R., & Oppacher, F. (2002). *Poker Hand* [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C5KW38.

EXPLORING BIG DATA CLUSTERING: APPROACHES, ALGORITHMS, AND PLATFORMS

Ramiz Aliguliyev[1], Tural Badalov[1]

•

[1]Institute of Information Technology, Baku, Azerbaijan
r.aliguliyev@gmail.com, bedelov.tural@gmail.com

Abstract

Clustering as the problem of discovering natural grouping in data has gotten a lot of attention due to its wide range of applications in health care, customer segmentation, image processing & transformation, market and recommendation systems, social network analysis, etc. It is an unsupervised learning task used to discover similar objects in a large dataset without relying on any prior information and gathering them into the same group. With the rapid growth of big data as result of data sets acquired by mobile devises, cameras, various sensors and other sources has necessitated research into extracting valuable information from enormous data sets. In this paper, we looked at different big data clustering approaches in the context of general clustering methods. In addition, we discussed several similarity measures as well as key clustering challenges such as cluster tendency assessment and cluster validity.

Keywords: big data, clustering, vertical scaling platforms, GPU, FPGA, MapReduce, Apache Spark.

I. Introduction

The amount of data generated by many sources is enormous, and this process is speeding up due to the high level of technology employed for various purposes. It is needed to deal with this huge data to extract useful information that benefits both businesses and individuals. So as a data mining tool cluster analysis or clustering is used to divide the data objects into subgroups or clusters so that objects within the same cluster resemble one another but differ greatly from those in other clusters. The fundamental advantage of clustering over classification is that class label for an object is not known in advance, allowing it to be used in a variety of fields such as, biology, business intelligence, image pattern recognition, Web search and etc. In Biology, It is typical to use cluster analysis to evaluate the gene expression data. For example, capturing and analyzing thousands of data points on gene expression level of cancer cells can be used to predict disease prognosis.

Clustering can be used in business intelligence to organize a large number of consumers into groups with strong common qualities. This makes it easier to come up with corporate strategies to improve customer relationship management. In Web search, clustering is used for organizing the search results according to the keyword into groups presenting them in a clear and accessible manner. Furthermore, approaches for grouping texts into subjects have been developed.

The rest of the paper is divided into three parts. Section 2 presents related work that provides a literature review of large data clustering approaches. Section 3 discusses the clustering of big data. The conclusion is made in Section 4.

II. Related Work

In [1] author gives a comparison of some existing algorithms with different attribute reduction techniques and address the challenges in IoT Big Data Clustering.

Researchers in [2] provides an overview of different clustering algorithms and make a taxonomy of different clustering techniques and analyze their suitability for clustering Big Data.

Authors in [3] evaluate clustering methods for big data. They enlist the criterion according to the properties of big data like volume, variety and velocity. By using results of experiments on real datasets it has been selected some algorithms over others according to which extent the listed criterion is satisfied.

In [4] Authors split big data clustering into two categories, single-machine and multi-machine clustering techniques. They give the characteristics of various clustering algorithms in each category and point out the some advantages of GPU based MapReduce technique over another algorithms.

Annad Nayyar's paper [5] represents a comprehensive analysis of different clustering algorithms classifying into the following groups: Partitioning-based, Hierarchical, Density-base, Grid-based and Model –based methods. He concludes his work in the summary table mentioning the competence of the algorithms for large data clustering.

III. Big data clustering algorithms and methods

Clustering big data is a hard task to do in data mining and most of the existing algorithms cannot deal with that. So, It requires modification of some appropriate algorithms or a totally new approach to cope with the problem. Generally, it is considered single- machine based techniques and multiple-machine based techniques.

A) Single-Machine Based Techniques

Single machine techniques use the resources of a single machine, implement on comparatively small data set, and lead to a conclusion for the entire data set. The idea of reducing the size of data distinguishes Sampling based techniques form dimension-reduction techniques in this category. There are different sampling methods involving probability and non-probability sampling and there have been conducted various studies on data sampling in the context of big data [21], [22], [23].

The main advantages and disadvantages of single-machine-based clustering techniques are demonstrated in table1.

a) *Partitioning –Based Methods:* Mini-batch k-means (MBKM) [6] is an analog of K means for large data that achieves local optimum with minimal computation cost by using a randomly picked section of the whole data in each iteration.

ClusteringLargeApplications(CLARA) [7] is a variation of PAM, that was created to address PAM's drawbacks, as it employs a random subset of data rather than a complete dataset. The main problem with Clara is that it may produce incorrect clustering if one or more sampled medoids are far away from real medoids. To resolve this issue, it was introduced Clustering Large Applications based on Randomized Search(CLARANS) [8]. Instead of selecting a random subset, this algorithm picks sample of neighbors dynamically which is then specified as a parameter, to be examined for the best medoid in each iteration.

A single pass fuzzy-c-means algorithm was presented in [24] The algorithm doesn't require fitting the whole dataset in the memory. Here original dataset is divided into n equal parts and in each step only one part is loaded into the main memory and then by using Fuzzy C means clustering algorithm data is partitioned into c clusters. Then data in memory is compressed into c weighted points, and clustering is then continued with a newly loaded chunk. When the entire dataset has been processed, the procedure is finished.

b) *Hierarchical-Based Methods:* BIRCH [9] is a two-phase clustering, in the first phase it scans the entire data to build CF –tree, and then by using any known algorithm it clusters leaf nodes of CF-Tree. So It offers flexibility to be used in conjunction with other clustering methods.

CF (Clustering Feature) as a hierarchical data structure for the algorithm is a statistical summary of the data points consisting of three components : $CF = (N, \overrightarrow{LS}, SS)$. N –number of data points in the cluster, $\overrightarrow{LS}$ - is a linear sum of N data points, SS- is a square sum of N data points. This method has a linear scalability, with just one scan, it can identify a solid clustering. However, the algorithm uses radius or diameter as a hyper-parameter, that is why the clusters frequently have a spherical shape. All the mentioned algorithms above employ a single point as a cluster representative, so the clusters tend to be spherical. This problem is addressed in [10]. CURE algorithm uses multiple representative points to describe the cluster. First, it chooses data samples at random that will fit in main memory. The selected data is then divided into clusters by employing any clustering technique, most often a hierarchical approach. A few typical points for each cluster serve as an explanation. So that enables CURE to cluster of arbitrary shape.

c) Density-Based Methods: Algorithms in this category are generally not seen to be very effective in clustering large amounts of data. However, DENCLUE [11] is good at handling arbitrary shaped clusters in high dimensional dataset. This algorithm uses statistical density function to determine the influence of each data point. The total sum of influence functions applied to all data points constitute what is referred to as the data space's overall density. Density attractors, which are the local maxima of the overall density function, then identify clusters.

It has different variations. In order to improve time complexity, it was introduced DENCLUE-IM [12]. According to the experiments it has been observed that execution time of DENCLUE-IM is reduced multiple times compared to DENCLUE and its other variants.

d) Grid-Based Methods: Grid-Based Methods represent a distinctive approach to clustering, particularly well-suited for high-dimensional data and scenarios where data distribution exhibits complex shapes or patterns. Instead of directly working with individual data points, these methods partition the data space using a grid structure, and clustering is performed on these grid cells. One notable algorithm in this category is OptiGrid [13], which excels at clustering high-dimensional data with arbitrary shapes. Grid-Based Methods divide the feature space into a grid of cells. The size and granularity of the grid are crucial parameters that influence the clustering results. The grid cells serve as the basic units for clustering. Each grid cell often contains statistical information summarizing the data points within it. This information can include the mean, variance, or other relevant statistics. By summarizing the data in this way, grid-based methods reduce the dimensionality of the problem, making it more manageable for clustering algorithms. Grid-based clustering is not directly dependent on individual data points but rather on the characteristics of the grid cells. This can be advantageous in high-dimensional data scenarios where traditional clustering algorithms may struggle due to the curse of dimensionality. High-dimensional data often poses challenges for clustering algorithms, as the distance metrics become less effective in high-dimensional spaces. Grid-Based Methods can mitigate these challenges by focusing on the relationships between grid cells rather than individual data points. OptiGrid, in particular, is known for its ability to identify clusters of arbitrary shapes in high-dimensional data. This is a valuable feature, as many real-world datasets do not conform to simple geometric shapes. This algorithm can be highly scalable, making it suitable for large-scale and big data clustering tasks. The grid structure allows for efficient parallelization and distributed processing.

Table 1: *Main Advantages and Disadvantages of Single Machine Based Clustering Techniques*

Different Approaches to Big Data Clustering	**Advantages**	**Disadvantages**
Partitioning based methods	Scalability Flexibility in choosing the number of partitions Allowance for parallel processing	Sensitivity to Initial Partitioning Difficulty in handling high-dimensional data
Hierarchical-Based Methods	No prior knowledge of cluster number Interpretable results	Computational complexity Lack of support for distributed computing Difficulty in handling high-dimensional data
Density-Based Methods	Robustness to noise Flexibility in cluster shape and size	Sensitivity to parameter settings Lack of support for distributed processing computationally expensive, especially when dealing with large-scale datasets
Grid -based Methods	Scalability Parallel processing capabilities	Challenge in Grid cell size determination Grid shape constraints (assume that data distribution is uniform and etc.) Sensitivity to input order
Dimension Reduction Techniques	Making clustering algorithms more efficient and scalable for big data Improved Interpretability Handling Curse of Dimensionality	Information Loss Subjectivity in Feature Selection Increased Complexity- adds an extra step to the clustering pipeline
Vertical Scaling Platforms	Performance- fast processing Scalability- handle larger data volumes without the need for distributed systems Simplicity- more user-friendly compared to distributed systems	Limited scalability-physical limit to how much a single machine can be scaled up Cost- they require high-end hardware to support the increased resource requirements. Lack of fault tolerance

B) Dimension Reduction Techniques – One of the major problems with clustering, especially in image processing and text documentation is the high dimensionality of the dataset, commonly known as the "curse of dimensionality". In literature there have been suggested several dimension reduction techniques, basically they are divided into two types: feature selection and feature transformation.

The primary goal of feature selection is to identify and retain a subset of the most relevant and informative features from the original dataset while discarding the less important ones. This process

aims to reduce the dimensionality of the data while preserving its essential characteristics, making it more suitable for clustering algorithms. It can be performed through manual selection by domain experts or through automated algorithms that evaluate the importance or relevance of each feature. For example: fast correlation based filter (FCBF) [14], fast clustering-based feature selection algorithm (FAST) [15], Markov Blanket Filter (MBF) [16]. FCBF measures feature relevance using correlation and conditional mutual information. It identifies and selects features that are highly correlated with the class labels. FAST employs a clustering technique to group similar features and selects representatives from each cluster. This reduces redundancy and retains essential information. MBF identifies features that are conditionally independent of the class label when other features are known. It aims to find a minimal set of features that predict the class label effectively. Even though both methods are utilized to lessen the number of attributes in a dataset, future selection extracts attributes without affecting them, but feature transformation generates new combinations of the original features to produce a reduced-dimensional representation of the data. These combinations are linear or nonlinear transformations of the original attributes. Principal component analysis (PCA) [17] Random projection (RP) [18] Linear Discriminant Analysis (LDA) [19] Canonical Correlation Analysis (CCA) [20] are some popular feature transformation algorithms. PCA is a widely-used linear feature transformation technique that projects the data onto a set of orthogonal axes (principal components) that capture the maximum variance. It reduces dimensionality while preserving as much variance as possible. RP is a dimensionality reduction technique that uses random projections to map high-dimensional data to lower-dimensional space while preserving pairwise distances to some extent. LDA is a supervised feature transformation technique that finds linear combinations of features that maximize the separation between classes, making it particularly useful for classification tasks. CCA is used for multivariate data analysis. It finds linear combinations of features from different datasets that maximize their correlation, which can be beneficial when dealing with data from multiple sources or domains.

C) Clustering Using Vertical Scaling Platforms- By adding more power, such as upgrading hardware on the existing system, we may increase the performance of the algorithms performed on a single machine. The most widely used vertical scale-up paradigms include multicore CPUs, Graphical Processing Units (GPU), and Field Programmable Gate Arrays (FPGA) [25].

Multicore CPUs- The goal of conventional CPU optimizations was to accelerate the serial execution of a single thread. A processor with several cores is referred to as a multicore CPU and parallel computation is accomplished using the multi-threading paradigm [26]. The task is divided into threads, each of which is executed simultaneously across many CPU cores. The primary disadvantage of CPUs is that system memory limits the amount of data that they can process. Some parallel clustering techniques based on multicore processor have been presented. For example, In [27] It is suggested modification of the FDBSCAN algorithm for multicore platforms. Authors named their algorithm M-FDBSCAN. Here the dataset is distributed equally among the cores, then FDBSCAN is applied to each subset. To obtain the end result, intermediate sub dataset pairs are merged according to ε neighborhood of the splitting line.

Graphical Processing Units- Originally GPU was designed to perform calculations at three-dimensional (3 D) graphics. The input primitives for the GPU are the vertices of the triangles that make up the three-dimensional representation of the data. Each vertex needs to be converted into pixels and shaded, then mapped onto the screen. The final image is created by combining those pixels. The same program is used by GPU to process several components (vertices) simultaneously. According to the SPMD (single-program multiple-data) programming paradigm, components are independent of one another and are unable to communicate [28]. That is why there is limited software integrated with GPUs. With the release of the CUDA framework, Nvidia opened up GPU programming to any programmers without requiring them to understand the specifics of the hardware. With this framework, authors in [29] represented implementation of Markov Clustering

algorithm (MCL) on GPU. Two operations, expansion and inflation, which are sparse matrix-matrix multiplication and Markov matrix normalization, respectively, define the MCL's time complexity. The performance of the method MCL was enhanced by computing parallel tasks for both expansion and inflation procedures in this study. In [30] parallel implementation of K-means algorithm on GPU via CUDA platform is discussed. The suggested variant is compared with optimized sequential k-Means. Experiments on synthetic data showed that parallel K-Means with CUDA outperform by 4 to 43 times for large dataset.

Field Programmable Gate Arrays (FPGA)- FPGA is an integrated circuit device consisting of logic blocks, programmable routing and I/O blocks to make connections with external devices. Logic blocks perform basic computations and are used as a storage element. The programmable routing consists of wires and programmable switches that provides connection among logic blocks and I/O blocks [31]. On the programming side, hardware description language (HDL) is used to generate PFGA designs . FPGAs are usually compared with Application Specific Integrated Circuit (ASIC), where the latter is power-efficient by 12 times on average, faster by approximately 3.2 times, and requires less area by roughly 23 to 55 times [32].In spite of having these disadvantages, FPGAs are very suitable for a specific set of applications because of their flexible and adaptable nature. On FPGAs, parallelism is made possible by the programmable routing circuit design and architecture, while many applications rely on the parallel execution of the same tasks. There are several practical uses for FPGAs, including in medical electronics, video and image processing, search engine algorithms and other areas. In regard with clustering FPGA-based implementation of DBSCAN is considered in [33]. The authors of the study assert that since extended neighbourhood of point queries are independent of data and take up the majority of execution time, they can be carried out in parallel. The suggested technique is on average 32 times and 202 times quicker than established software algorithms, according to tests on two-dimensional real and synthetic datasets. Hardware implementation of KMeans on FPGA is represented in [34]. This approach considers a large number of clusters (up to 256) and is intended for color images. In order to accelerate the algorithm the author used simplified version of filtering and FEKM algorithms. The former technique is employed to attain high performance by restricting comparisons for the closet centers to 24. FEKM technique reduces the number of data points scanned during each iteration, allowing it to analyze large datasets effectively. Experiments on images of 512*512 and 640*480 display that the performance of the suggested implementation of K-means with one FPGA is more than 30 frames per second.

Clustering using vertical scaling platforms involves enhancing the computational power of a single machine to improve the performance of clustering algorithms. The choice of platform (e.g., multicore CPU, GPU, FPGA) depends on the specific clustering algorithm, dataset size, and available resources. It's essential to carefully evaluate the benefits and costs of vertical scaling to determine if it's the right approach for a particular clustering task. It's important to note that not all clustering algorithms can be easily parallelized or accelerated using these platforms. Some algorithms are inherently sequential and may not benefit significantly from vertical scaling platforms.

D) Multiple-Machine Based Techniques

The volume of big data is now measured in petabytes, and it is constantly growing. Even with the use of any dimension reduction techniques, processing this much data on a single machine would be challenging or impossible. The majority of conventional clustering algorithms, on the other hand, struggle because of their reliance on input parameters, data order, and computation costs. Due to these constraints, researchers have led to design different algorithms to perform in a parallel and distributed environment using MapReduce or Spark frameworks. Table 2 presents a comprehensive summary of clustering algorithms, highlighting their applicability based on data characteristics and providing an analysis of their scalability factor.

Table 2: *A brief overview of clustering methods*

Clustering Method	Data Characteristics	Scalability	Examples
Partitioning-Based	Low-dimensional, well-separated clusters	Medium scalability for small to moderate data	K-Means, Mini-batch K-Means, CLARA
Hierarchical-Based	Data without prior knowledge of clusters, interpretable structure	Limited scalability (high computational cost)	BIRCH, CURE
Density-Based	Arbitrary-shaped clusters, noisy data	Low scalability (parameter-sensitive)	DBSCAN, DENCLUE
Grid-Based	High-dimensional data, complex distributions	High scalability (supports parallel processing)	OptiGrid
Multi-Machine	Very large datasets, distributed computing needed	Very high scalability (e.g., petabyte-scale data)	MapReduce, Apache Spark
GPU/FPGA-Based	Real-time processing, iterative tasks	Medium scalability (requires specialized hardware)	CUDA K-Means, FPGA DBSCAN

MapReduce- MapReduce is highly scalable and fault-tolerant programming model that is used in Hadoop to process massive data in parallel fashion. There are two primitive functions, Map and Reduce functions and programmer defines their work on these two functions without taking care of parallel execution across nodes. Another advantage of MapReduce is its flexibility that offers to process structured or unstructured data. In spite of having many advantages, MapReduce inherits some disadvanges by its nature. Firstly, it does not support any high level language like SQL in DBMS, which would optimize coding, and programmers should write their operations with Map and Reduce only. So it is hard to implement most of complex algorithms in this framework. In addition to that , it takes much time for frequent I/O operations, since the intermediate results between queries are stored in local disks. That causes low efficiency.

In [35] it is represented Multiplex KMeans algorithm with MapReduce. The primary concept of the approach is that in order to increase clustering quality, they run multiple KMeans concurrently using various centroid groups and choose the best one among them. The proposed algorithm firstly runs KMeans processes, then it evaluates quality of clustering result by Total Within-Cluster Variation value (TWCV) for each centroid groups, at the end of each iteration. Those which have high TWCV values are pruned in the next step. In the third step, The Permute job is deployed to determine similar centroids and to provide the same location in each group for similar centroids by rearranging them. New centroid groups are formed in the last step. Experiments on real-world datasets show that Mux-KMeans outperform naive KMeans in terms of clustering quality.

Qing Liao et al. in [36] suggested another parallel implementation of Kmeans using MapReduce. They achieved an improvement in comparison to traditional parallel KMeans by adjusting distance measure and initial centroids. In this study, Euclidean distance is chosen as the default distance metric, while Manhattan distance is utilized conditionally. The Initial centroids are picked from the high density region that are the furthest apart.

The work proposed in [37] is a parallel implementation of well-known DBSCAN algorithm using MapReduce. In this work, genetic algorithm (GA) is utilized for adjusting hyper-parameters (minPts and Eps). In order to overcome the time consuming problem of GA-DBSCAN algorithm , it is deployed on MapReduce framework.

Spark- Spark is an open-source distributed computing engine for analyzing large data. It is designed as an alternative to Hadoop to overcome the I/O constrains and enhance performance. When compared to Hadoop MapReduce, they are both fault-tolerant, well scalable and very effective for processing massive data across clustered computers, but Apache Spark outperforms in terms of real-time and iterative processing. Spark employs the concept of RDD (Resilient Distributed Dataset), which makes it perform in-memory computation. Many studies have shown that Spark is up to 100 times quicker than Hadoop MapReduce when the data can fit in memory and up to 10 times faster for batch processing.

Apache Spark provides four higher-level libraries: Spark SQL and Data Frames, Spark Streaming, Spark's Machine Learning Library (MLlib), and GraphX. MLlib offers more than 55 common algorithms for distributed data modeling, such as classification, regression, clustering, feature transformations and etc. [38]. Few clustering algorithms that have been implemented using Spark are included in MLlib as an open-source package. Several studies have been undertaken in order to solve specific shortcomings of the clustering algorithms using Spark or to give completely new techniques. Majority of the recent work on clustering with Spark is related to two prominent algorithms: KMeans and DBSCAN.

One of the first implementation of KMeans based on Spark is suggested in [39]. The proposed algorithm works in two phases: In the first phase, the Basic KMeans algorithm runs for the large number of k. Instead of random selection of initial centroids the algorithm utilize probability sampling, which is aimed to cut off number of iterations to converge. In the second phase, resulting centroids are merged according to certain criteria. Experiments on synthetic large scale datasets display that the proposed method solves the problem of over-resolution without degrading clustering performance.

S_DBSCAN [40] is a performance-oriented extension of DBSCAN that makes use of SPARK. The proposed algorithm consists of three steps: In the first step, data partition is performed by using a random sampling method according to the number of worker nodes. The local DBSCAN algorithm then runs in parallel to build intermediate clusters and each resulting cluster is saved to HDFS as a new RDD. In the last stage, the algorithm combines the partial clusters to generate global clustering results based on the proximity of centroids. The experiments demonstrate that the suggested method is as accurate as regular DBSCAN but more efficient when dealing with large amounts of data.

There are various other articles related to the Spark-based implementation of KMeans, such as [41],[42] and the Spark-based implementation of DBSCAN, such as [43], [44].

IV. Conclusion

Which algorithm and which platform to choose to cluster big data? There is no straight answer to the question and it usually depends on application. Time complexity and scalability are generally traded off. In other words, if it is required to get result in real-time processing, then it is most suitable to run the modification of the algorithms on GPU or FPGA platforms. If it is needed to deal with

huge amount of data like hundreds of gigabytes or even petabytes , then it is compulsory to choose multiple machine techniques, like MapReduce or Apache Spark. It is worth noting that Spark would be the best option of these two for implementing suitable clustering algorithms, because of its fast and easy big data processing characteristics. In the future, a hybrid method to clustering massive data is more suited, taking advantage of Spark's huge scalability and GPU's processing speed.

This work was supported by the Azerbaijan Science Foundation - Grant No. AEF-MCG-2023-1(43)-13/04/1-M-04.

References

[1] Nweso Emmanuel Nwogbaga, "A Review of Big Data Clustering Methods and Research Issues", International Journal of Science and Research (IJSR) vol. 9 no. 5, pp. 253-264, 2020.

[2] K. Djouzi and K. Beghdad-Bey, "A Review of Clustering Algorithms for Big Data," 2019 International Conference on Networking and Advanced Systems (ICNAS), pp. 1-6, 2019.

[3] A. Fahad et al., "A Survey of Clustering Algorithms for Big Data: Taxonomy and Empirical Analysis," in IEEE Transactions on Emerging Topics in Computing, vol. 2, no.3, pp. 267-279, 2014.

[4] Shirkhorshidi, A.S., Aghabozorgi, S., Wah, T.Y., Herawan, T. Big Data Clustering: A Review. In: , et al. Computational Science and Its Applications – ICCSA 2014. ICCSA 2014. Lecture Notes in Computer Science, vol 8583. Springer, (2014).

[5] Anand Nayyar, Vikram Puri, "COMPREHENSIVE ANALYSIS & PERFORMANCE COMPARISON OF CLUSTERING ALGORITHMS FOR BIG DATA", Review of Computer Engineering Research, vol 4, no. 2, 2017

[6] D. Sculley, "Web-scale k-means clustering," in Proceedings of the 19th international conference on World wide web. ACM, pp. 1177–1178, 2010.

[7] P. J. Rousseeuw and L. Kaufman, Finding Groups in Data. Wiley Online Library, 1990.

[8] R.T. Ng and J. Han, "Clarans: A method for clustering objects for spatial data mining," IEEE transactions on knowledge and data engineering, vol. 14, no. 5, pp. 1003–1016, 2002.

[9] T. Zhang, R. Ramakrishnan, and M. Livny, "Birch: an efficient data clustering method for very large databases," in ACM Sigmod Record, vol. 25, no. 2. ACM, pp. 103–114, 1996.

[10] S. Guha, R. Rastogi, and K. Shim, "Cure: an efficient clustering algorithm for large databases," Information Systems, vol. 26, no. 1, pp. 35–58, 2001.

[11] A. Hinneburg, D. A. Keim et al., "An efficient approach to clustering in large multimedia databases with noise," in KDD, vol. 98, pp. 58–65, 1998.

[12] H. Rehioui, A. Idrissi, M. Abourezq, F. Zegrari "DENCLUE-IM: A New Approach for Big Data Clustering", Procedia Computer Science, vol. 83, pp. 560 – 567, 2016.

[13] A. Hinneburg and D. A. Keim, ''Optimal grid-clustering: Towards breaking the curse of dimensionality in high-dimensional clustering,'' in Proc. 25th Int. Conf. Very Large Data Bases (VLDB), pp. 506–517, 1999.

[14] Yu. Lei, Huan Liu, Efficient feature selection via analysis of relevance and redundancy, J. Mach. Learn. Res. 5, 1205–1224, (Oct. 2004)

[15] Qinbao Song, Jingjie Ni, Guangtao Wang, A fast clustering-based feature subset selection algorithm for high-dimensional data, IEEE Trans. Knowl. Data Eng. 25 (1), 1–14, (2013).

[16] Y. Wang, J. Wang, H. Liao, and H. Chen, "An efficient semi-supervised representatives feature selection algorithm based on information theory," Pattern Recognition, vol. 61, pp. 511–523, 2017.

[17] X. Kong, C. Hu, and Z. Duan, "Generalized principal component analysis," in Principal Component Analysis Networks and Algorithms. Springer,pp. 185–233, 2017.

[18] S. Tasoulis, L. Cheng, N. Valim ¨ aki, N. J. Croucher, S. R. Harris, ¨ W. P. Hanage, T. Roos, and J. Corander, "Random projection based clustering for population genomics," in Big Data (Big Data), 2014 IEEE International Conference on. IEEE, pp. 675–682,2014.

[19] D. Chu, L.-Z. Liao, M. K.-P. Ng, and X. Wang, "Incremental linear discriminant analysis: a fast algorithm and comparisons," IEEE transactions on neural networks and learning systems, vol. 26, no. 11, pp. 2716–2735, 2015.

[20] Sakar, C.O., Kursun, O. & Gurgen, F. "Ensemble canonical correlation analysis". Applied Intelligence, vol. 40, pp. 291–304 (2014).

[21] Ying Yan, Liang Jeff Chen, and Zheng Zhang. Error-bounded sampling for analytics on big sparse data. PVLDB, 7(13):1508–1519, 2014

[22] Z. Liu and A. Zhang, "Sampling for Big Data Profiling: A Survey," in IEEE Access, vol. 8, pp. 72713-72726, 2020.

[23] Julian Ramos Rojas, Mary Beth Kery, Stephanie Rosenthal, and Anind K. Dey. Sampling techniques to improve big data exploration. In 7th IEEE Symposium on Large Data Analysis and Visualization, LDAV 2017, Phoenix, AZ, USA, October 2, 2017, pages 26–35, 2017

[24] P. Hore, L. O. Hall, and D. B. Goldgof, "Single pass fuzzy c means," in 2007 IEEE International Fuzzy Systems Conference. IEEE, pp.1–7, 2007.

[25] Singh, Dilpreet, and Chandan K Reddy. "A Survey on Platforms for Big Data Analytics." Journal of Big Data 2, no. 1 (2014).

[26] Sodan, Angela & Machina, Jacob & Deshmeh, Arash & Macnaughton, Kevin & Esbaugh, Bryan." Parallelism via Multithreaded and Multicore CPUs". Computer. 43. 24 – 32, (2010).

[27] Erdem, Atakan & Gündem, Taflan. "M-FDBSCAN: A multicore density-based uncertain data clustering algorithm". TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES. 22. 143-154, (2014).

[28] Owens, J.D., M. Houston, D. Luebke, S. Green, J.E. Stone, and J.C. Phillips. "GPU Computing." Proceedings of the IEEE 96, no. 5 , pg.879–899, (2008).

[29] A. Bustamam, K. Burrage and N. A. Hamilton, "Fast Parallel Markov Clustering in Bioinformatics Using Massively Parallel Computing on GPU with CUDA and ELLPACK-R Sparse Format," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 9, no. 3, pp. 679-692, May-June 2012.

[30] Zechner, Mario & Granitzer, Michael. (2009). "Accelerating K-Means on the Graphics Processor via CUDA". Proceedings of the 1st International Conference on Intensive Applications and Services, INTENSIVE, 7-15, 2009.

[31] Ian Kuon, Russell Tessier and Jonathan Rose, "FPGA Architecture: Survey and Challenges", Foundations and Trends® in Electronic Design Automation: Vol. 2: No. 2, pp 135-253, (2008).

[32] I. Kuon and J. Rose, "Measuring the gap between FPGAs and ASICs," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 26, no. 2, pp. 203–215, (2007).

[33] Scicluna, Neil & Bouganis, Christos.A Multidimensional FPGA-Based Parallel DBSCAN Architecture. ACM Transactions on Reconfigurable Technology and Systems. 9. 1-15, (2015).

[34] Maruyama, Tsutomu." Real-time K-Means Clustering for Color Images on Reconfigurable Hardware". Proceedings - International Conference on Pattern Recognition. 2. 816-819, (2006)

[35] Chen Li, Yanfeng Zhang, Minghai Jiao, and Ge Yu. 2014." Mux-Kmeans: multiplex kmeans for clustering large-scale data set". In Proceedings of the 5th ACM workshop on Scientific cloud computing (ScienceCloud '14). Association for Computing Machinery, New York, NY, USA, 25–32. 2014

[36] Liao, Qing & Yang, Fan & Zhao, Jingming." An improved parallel K-means clustering algorithm with MapReduce", 15th IEEE International Conference on Communication Technology (ICCT), 764-768, (2013)

[37] Hu, Xiaojuan & Liu, Lei & Qiu, Ningjia & Yang, Di & Li, Meng. A MapReduce-based improvement algorithm for DBSCAN. Journal of Algorithms & Computational Technology. 12.(2017).

[38] Assefi, Mehdi & Behravesh, Ehsun & Liu, Guangchi & P. Tafti, Ahmad. "Big Data Machine Learning using Apache Spark MLlib". IEEE Big Data , 2017.

[39] Sinha, Ankita & Jana, Prasanta." A novel K-means based clustering algorithm for big data". International Conference on Advances in Computing, Communications and Informatics (ICACCI), 1875-1879.(2016)

[40] G. Luo, X. Luo, T. F. Gooch, L. Tian and K. Qin, "A Parallel DBSCAN Algorithm Based on Spark," 2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom), pp. 548-553, 2016.

[41] Wang B, Yin J, Hua Q, Wu Z, Cao J (2016) Parallelizing k-means-based clustering on spark. In: International conference on advanced cloud and Big Data (CBD). IEEE, Chengdu, pp 31–36.

[42] Zayani A, Ben N'Cir CE, Essoussi. N Parallel clustering method for non-disjoint partitioning of largescale data based on spark framework. In: IEEE international conference on big data (Big Data). IEEE, Washington, DC, pp 1064–1069. (2016)

[43] D. Han, A. Agrawal, W. -k. Liao and A. Choudhary, "Parallel DBSCAN Algorithm Using a Data Partitioning Strategy with Spark Implementation," 2018 IEEE International Conference on Big Data (Big Data), pp. 305-312, 2018.

[44] Gong, Y., Sinnott, R.O., Rimba, P. RT-DBSCAN: Real-Time Parallel Clustering of Spatio-Temporal Data Using Spark-Streaming. In: , et al. Computational Science – ICCS 2018. ICCS 2018. Lecture Notes in Computer Science(), vol 10860. Springer, Cham.(2018).

ARCHITECTURAL-TECHNOLOGICAL PRINCIPLES OF A PATIENT-CENTERED DIGITAL TWIN AND ITS VISUALIZATION ALGORITHM IN CLINICAL PRACTICE

Masuma Mammadova[1], Zarifa Jabrayilova[1], Aytan Ahmadova[1]

[1]Ministry of Science and Education Republic of Azerbaijan, Information Technology Institute, Baku, Azerbaijan
mmg51@mail.ru , djabrailova_z@mail.ru , ayten.adia1996@gmail.com

Abstract

This article presents the essence and architectural-technological principles of digital twin technology, and presents an algorithm for creating a digital twin based on the visualization of medical data through Power BI. The digital twin supports the early prediction of diseases and the analysis of personalized medical indicators by simulating the human body and enabling real-time monitoring of the condition. The article proposes a generalized architecture for building a digital twin of the human body at the level of organs and relevant diseases, the stages of its formation, and develops a schematic description of a virtual object. The Power BI platform is chosen to visually present the virtual object update simultaneously with any changes occurring in the physical object, which is the main architectural component of the medical digital twin. As a real source of information, it is referred to the available national database where doctors periodically collect information about patients in traditional clinical practice. Dashboards are developed on the Power BI platform to form the trajectory of change of medical indicators based on their values given in a time series. The digital twin created based on this approach enables the real-time monitoring of dynamic changes in test results, which supports the acceleration of physician decision-making related to health condition management, the improvement of service quality, and the provision of more prompt and personalized patient care.

Keywords: digital technology, e-health data, digital twin, visualization, Power BI, digital twin dashboard.

I. Introduction

Nowadays, digital technology is applied in a complex and targeted manner in almost all areas of healthcare. Today it is possible to process heterogeneous health data in real time, whereas, previously, the generation, collection and joint processing of this data was time consuming and performed offline.

The acceleration of the trends in the use of digital technology expands its boundaries and possibilities, and accordingly, conditions the emergence of the phenomenon of "digital transformation". Digital transformation ensures the data transformation into information and knowledge and its use for effective decision-making based on the new knowledge acquired [1]. The digital twin (DT), one of the basic technologies of digital transformation, acts as a digital copy of the

physical object or process they represent and enables their real-time monitoring and situation assessment, regardless of location [2].

The present article proposes an algorithm for creation of a DT based on the Power BI platform to visually present the essence, architecture of a medical DT and update of a virtual object.

II. Related work

In the 1950s, NASA (National Aeronautics and Space Administration), GE (General Electric) and other industrial manufacturers created abstract digital models of equipment to account for equipment productivity and life cycle through simulation modeling [3].

In 2002, DT was presented as a "virtual, digital equivalent of a physical object" [4], and proposed as a conceptual model based on Product Lifecycle Management (PLM). It was technologically extremely difficult and expensive to bring to the use. Ten years later, the rapid development of "big data", "Internet of Things" (IoT) and "Internet of Medical Things" (IoMT) technologies enabled the implementation of this idea. In 2010, Jon Vickers from NASA introduced the term "digital twin" to science. In 2014, the concept of "virtual digital equivalent of a physical product" was proposed by M. Grieves [5].

Combining sensor data, computer modeling and artificial intelligence algorithms, medical DT acts as a digital (virtual) mirror of the real object [6].

DT in healthcare is designed to provide more effective and flexible medical interventions, assisting physicians and medical staff to understand the patient's health status [7]. One of the main conditions for the application of DT in medicine is its real physical object. A medical DT is a dynamic digital model that contains all the information about a physical object or medical system.

Establishing a complete DT of healthcare is a long-term and multi-level process. The first, that is, the health level, includes medical activity, pharmacy and medical education. The second level is the subject level, which includes departments such as ambulatory care, emergency care, dentistry, drug production and drug sales. The third and fourth levels, respectively, consist of the creation of personalized (patient-centered) DT on individual human organs and diseases related to these organs [8].

The benefits of DT can be applied in disease treatment, health management and many areas of personalized medicine in the healthcare sector. DT can provide different treatment options for different patients with the same disease and accelerate the development of personalized medicine. The essence of patient-centered medicine is the individualization of drug therapy according to the personal data and genotype of a specific patient. These opportunities created by DT are very important in terms of providing medical services to patients, which is the main component of healthcare. By transferring the patient's physical characteristics and health data to the digital environment, DT offers innovative solutions in determining the accurate diagnosis and treatment process, as well as in the formulation of personalized medicine. Digital twins of the human body are built to model its organs, cells, individual genetic composition, physiological characteristics and habits (lifestyle) of each person, and enable the creation of individual drugs and treatment plans [9]. These copies (the digital patient), which are the internal subsystems of the human body as a whole, can improve healthcare and patient care by personalizing diagnostics.

Medical digital twins can be used to solve the following problems for the formulation of personalized medicine:

- making an early diagnosis at the initial stage of the disease;
- monitoring of further development trajectories of the disease;
- optimizing the time of medical aid;
- improving personalized medicine;

• clarifying the mechanisms of drug action on the patient, etc.

Along with listed solutions, DT allows to test and visualize relevant results and predict dynamic variables [10].

[11] defines the main features of DT as follows:

1. Connectivity- A connection is required between the DT and its real-life equivalent that ensures a continuous and reliable data flow.

2. Homogeneity - the data from different sources should be homogenized.

3. Reprogrammable- as the volume of regularly collected data increases, DT also develops and provides smarter decision-making through artificial intelligence.

4. Modularity- DT can allow the manufacturer to learn which specific components of the device are underperforming.

A connection between the DT and its real-life equivalent that ensures a continuous and reliable data flow is a prerequisite. Thus, DT is updated and changes as its physical equivalent changes. It uses IoT to understand the physical object's current state, location, function, composition, etc. The generated unstructured data is then transmitted to the virtual twin in real time, resulting in a dynamic representation of the situation. DT can learn from this data through machine learning, from human experts with deep knowledge in the subject field, as well as from identical twins [12].

Thus, the general architecture of the medical DT consists of physical and virtual objects, technologies that ensure the connection between them, a unified health database, a centralized management system and visualization systems (Figure 1) [13].

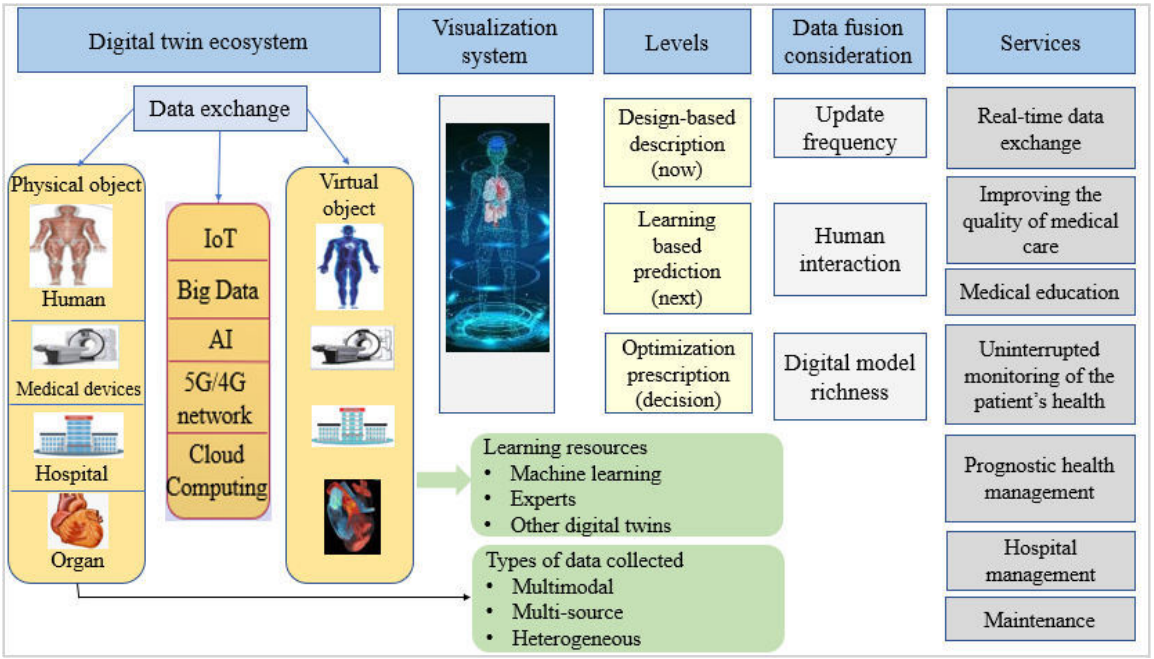


Figure 1: *Generalized architecture of medical digital twins [13]*

The **physical objects** of a medical digital twin refer to patient, hospital, doctor, medical devices, organs, etc.

Table 1 presents the information required for building the DT of the physical objects of the health care system.

Table 1: *Information required for building the DT*

№	Physical objects of healthcare system	Information required for building the DT
1	Patient	Genetic data, laboratory results, medical images, biomedical signals, personal data, social determinants
2	Hospital	Medical resources (equipment, medicinal preparations, etc.), personnel resources, intra-hospital operational data, building layout
3	Health data collectors	Sensor data, quality indicators, environmental data

Along with all examination data of the relevant physical object, data regularly collected through 4G/5G, IoT, IoMT are used to form a database of personalized medical DT [14,15]. The data collected from physical objects are large-scaled and diverse in nature, but also obtained from different sources, and in this regard, they can be classified as follows [16]:

- IoT devices: Blood pressure, heart rate and other biometric parameters;
- Electronic Health Records (EHR): Laboratory results, treatment plans and health regords;
- Artificial intelligence and analytical models: Data for patient health predictions and early disease detection, etc.

This data is stored in databases such as cloud systems or SQL servers.

A **virtual object** is software consisting of a set of applications that explain the behavior of a medical object on a computer. Applications are used as a real-time control system of the object's activity, and this system operates throughout the entire life cycle of the object.

Visualization system displays data management models, information from various sources and results in a visual form in different fields of medicine. This enables effective monitoring, control and management of processes occurring in a medical physical facility.

III. Problem statement

I. Establishment of the DG at the level of the human body organs and relevant diseases;
II. Creating a digital twin through health data visualization.

IV. Problem solution

I. The study proposes the implementation of the following stages of the DT at the level of the human body organs and relevant diseases:

1. first stage is characterized by the collection of necessary medical data on each relevant physiological level. This stage builds a general (standard) digital model of the physical medical object based on available medical data;

2. second stage performs collection, integration and analysis of data about a specific individual, and creates an individual model of a physical object through modeling;

3. third stage analyzes the differences between the general and individual model built:

a) deviations in the considered parameters according to all physiological levels are determined;

b) an intellectual analysis of the current situation is performed and an appropriate decision is made;

c) obtained positive decisions are used for examination, diagnosis, prediction of diseases and prescription of the therapy of the patient.

Figure 2 presents a schematic illustration of the proposed functionalization of the DT based on medical data regularly collected from patients.

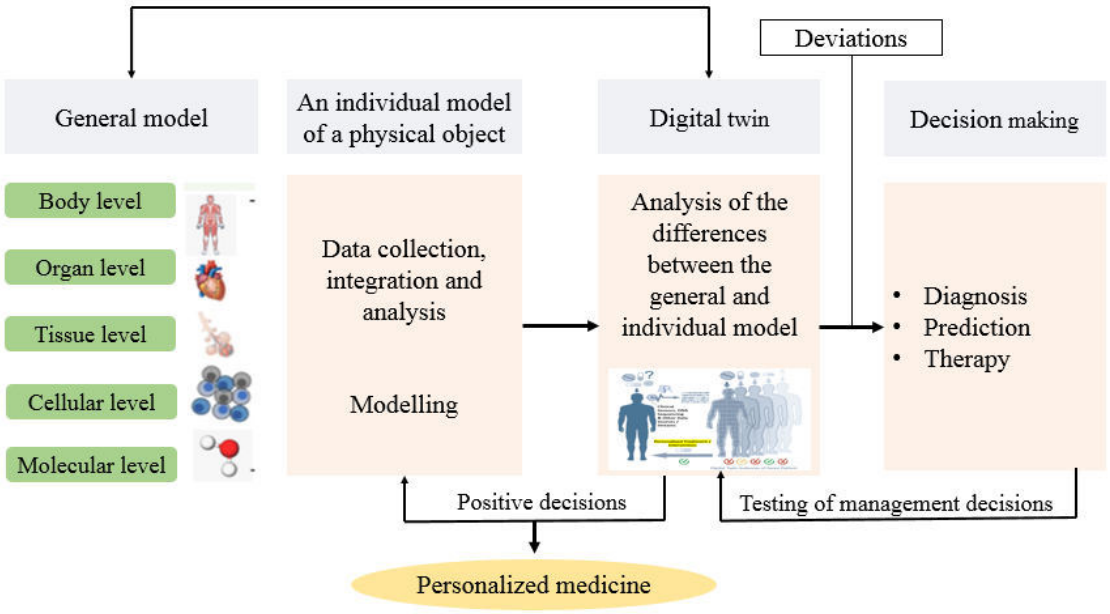


Figure 2: *Formation of DT based on medical data collected from patients*

II. Visualization of a digital twin is a visual representation of a virtual object based on continuously collected data to update it simultaneously with the changes occurring in the physical object [17]. This article uses Power BI software, a cloud-based data analytics platform developed by Microsoft, to visualize real health data. Power BI collects and cleans data from external sources, creating data models, analyzing and visualizing data. As the visualization tool, Power BI has a more user-friendly interface than others. Power BI enables intuitive visualization and continuous updating of digital twin data by linking the built-in dashboard with data collected from a real physical object. Visualization of medical digital twin data through Power BI is a viable approach to optimize healthcare services and improve decision-making based on real-time data [18].

The algorithm for creating a digital twin with the visualization of health data based on Power BI is performed in the following stages:

Stage 1: Selection of the base. In this regard, data can be connected to Power BI from sources such as Azure Data Lake, SQL Database, APIs or Excel.

The present study refers to existing national database, which is periodically collected by doctors in traditional clinical practice regarding patients, as a real source of information. The real database is a database of cirrhosis patients created in Excel by doctors of the Surgical Diseases Clinic of Azerbaijan Medical University (Figure 3).

	Name	Gender	Date	Diagnosis	Age	Interval	Birth date	BWI	MELD	HB	Leuk
	Patient 14		11/2/2022	Cirrhosis			1960		9	14.5	5.54
	Patient 14		6/23/2023	Cirrhosis			1960		8	13.1	5.27
	Patient 14		11/15/2023	Cirrhosis			1960		10	13	5.02
	Patient 15	Female	10/12/2019	HBV, Cirrhosis	52	50-60	1972			13.4	8.44
	Patient 15		11/4/2022	HBV, Cirrhosis						13.3	7.44
	Patient 15		4/5/2023	HBV, Cirrhosis					9	13.5	7.53
	Patient 15		4/27/2023	HBV, Cirrhosis					12	13.4	6.64
	Patient 15		7/24/2023	HBV, Cirrhosis					12	12.9	6.72
	Patient 15		11/1/2023	HBV, Cirrhosis					8	12	5.96
	Patient 16	Male	5/10/2023	Alk, Cirrhosis	53	50-60	1971		12	11.4	4.25
	Patient 16		6/12/2023	Alk, Cirrhosis			1971		12	11	3.65
	Patient 16		10/27/2023	Alk, Cirrhosis			1971		12	11	3.47
	Patient 17	Female	4/30/2021	Cirrhosis	67	60-70	1957			13.7	5.72
	Patient 17		6/29/2021	Cirrhosis			1957		10	11.2	5.6
	Patient 17		10/4/2021	Cirrhosis			1957		10	11.6	4.51
	Patient 17		7/13/2022	Cirrhosis			1957		12	11.8	4.66
	Patient 17		6/19/2023	Cirrhosis			1957		15	11	4.48
	Patient 17		8/2/2023	Cirrhosis			1957			11.3	3.93
	Patient 18	Male	7/31/2023	Alk, Cirrhosis	59	50-60	1965			13	3.74
	Patient 18		8/15/2023	Alk, Cirrhosis						11	2.95
	Patient 18		9/15/2023	Alk, Cirrhosis					15	13	3.02

Figure 3: *Fragment from the database created by the doctors of the Surgical Diseases Clinic of Azerbaijan Medical University on cirrhosis diseases*

Stage 2: Uploading the data collected in the Excel database to the visualization program.

Stage 3: Data transformation is performed. Data transformation in Power BI is implemented through Power Query. At this stage, the following main tasks are executed:

- If empty cells are found in the medical records, they are either deleted or filled in accordingly;
- Detection of outliers, i.e., examination of abnormal values;
- Generated unstructured data is structured and so on.

Structuring the data from a physical object provides the following opportunities:

- Analyzing patient data more accurately;
- Detecting risky situations in advance;
- Making decisions based on real-time (live) indicators.

By the implementation of the listed stages of transformation, it is possible to create visual images for each medical parameter characterizing the patient's condition.

Stage 4: Building a digital twin dashboard. Data visualization enables healthcare professionals to monitor patient conditions in real time and make flexible decisions. Dashboards in Power BI provide detailed analysis and targeted monitoring.

Fig. 4 shows the control panel of the digital twin based on dynamic examination data of cirrhosis patients: a) by the cirrhosis types (HCV cirrhosis, HBV cirrhosis, ALK cirrhosis, cirrhosis); b) by the demographic distribution of patients (age, gender).

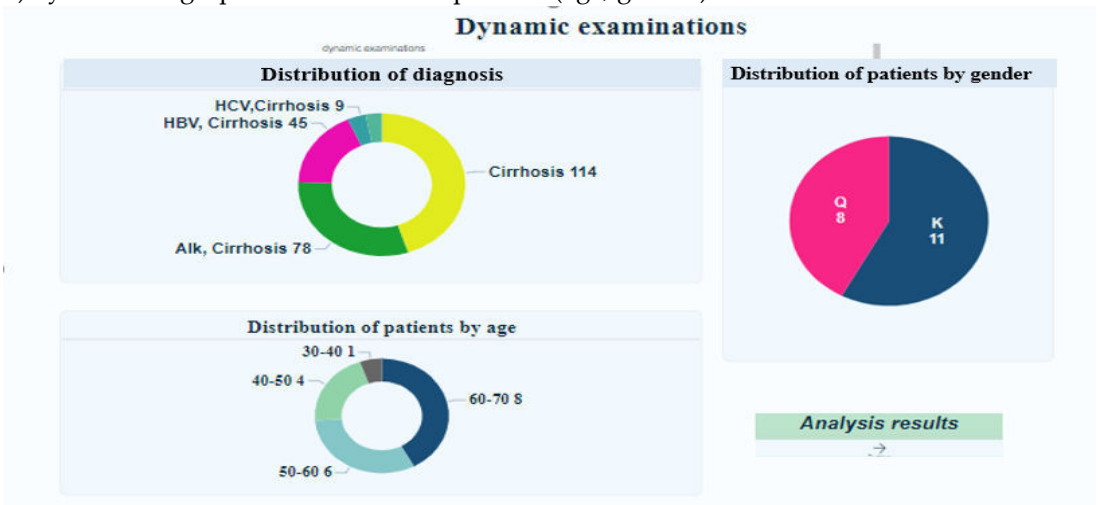


Figure 4: *Digital twin control panel: a) by the cirrhosis types (HCV cirrhosis, HBV cirrhosis, ALK* cirrhosis, cirrhosis); b) by the demographic distribution of patients (age, gender)

As shown in the figure, 11 patients are male, 8 are female, 1 patient is aged 30-40, 4 patients aged 40-50, 6 patients aged 50-60 and 8 patients aged 60-70.

When we click on the "Analysis results" button, the window illustrated in Figure 5 opens.

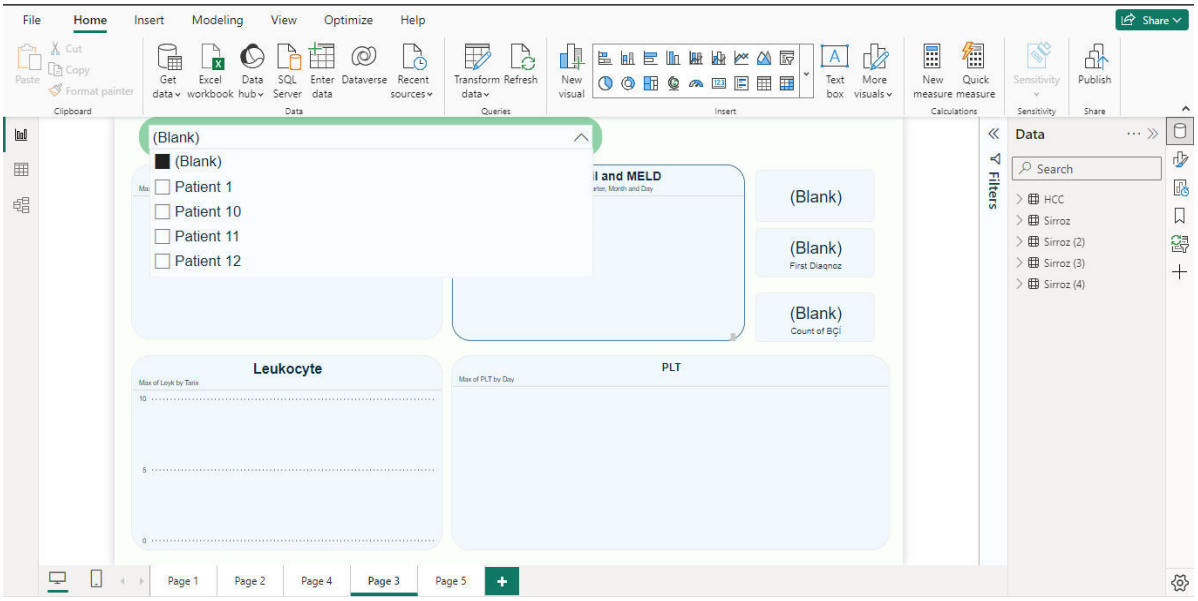


Figure 5: *"Analysis results" login window*

To view individual analysis results of any patient in the login window, select the patient's name from the list. In the window depicted in Figure 6, it is possible to see the patient's year of birth, type of disease, and how laboratory analysis results change over time through visual tools, such as Gauge chart, line chart, and slicer for a specific selected patient. By approaching any point in those graphs, it is possible to see the date and result of the analysis.

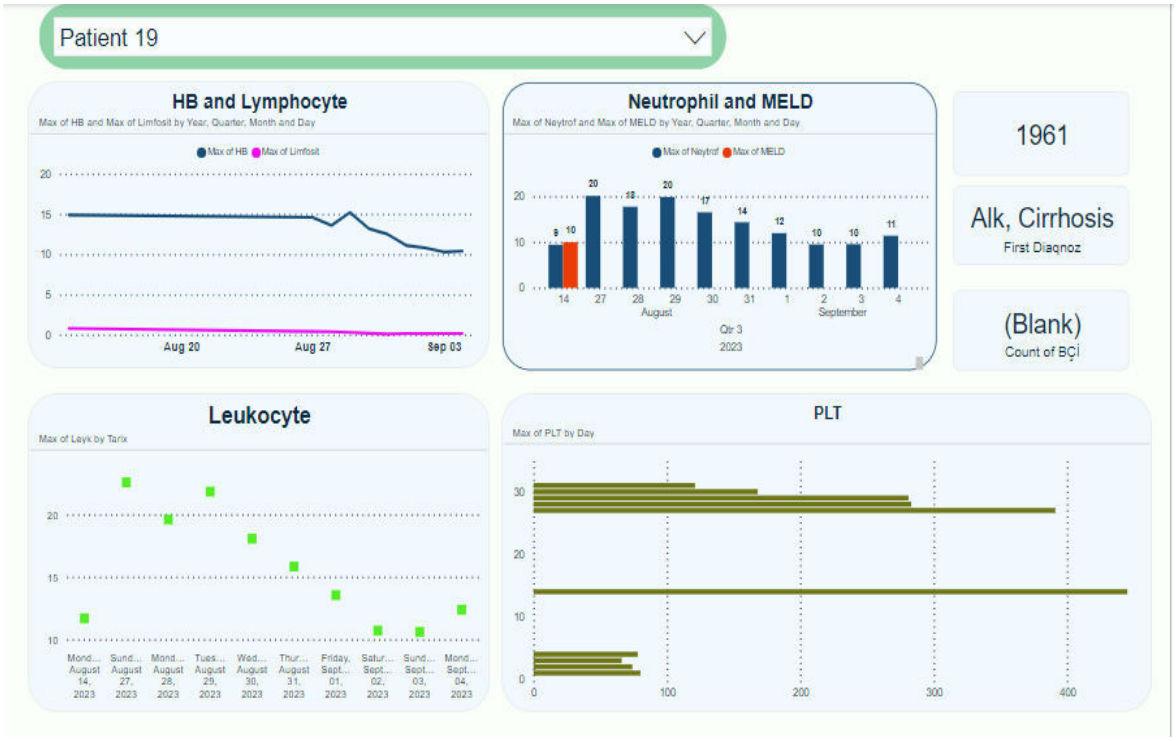


Figure 6: *Visual image of a digital twin of a cirrhotic patient*

Real-time visualization of data from a physical object through dashboards created in Power BI provides the following opportunities:

- The dashboards represents the dynamics of time series formed on each medical parameter of the patient;
- Early prediction of diseases and optimization of treatment processes are ensured;
- Decision-making process accelerates for physicians by creating a visual representation of personalized indicators related to each patient.

Such a "digital twin" contains information about the patient, shows trends in the development of the disease, and shapes the patient's health trajectory by updating the data after each medical examination.

Based on our proposed algorithm, it is possible to create a digital twin by visualizing examination data related to other diseases in a similar manner.

Thus, the visualization of personal health data through Power BI allows for continuous monitoring of patients' conditions and timely medical interventions through dashboards created in real time.

V. Summary

The use of medical digital twins has great potential for improving individual treatment programs in the healthcare sector. This article presented the architecture of creating a DT - the transformation of a physical object into a virtual object based on data collected from patients. It proposed the architectural and technological principles of DT at the level of the human body organs and relevant diseases and the stages of virtual object formation and a schematic description. When creating a DT, data collected regularly under the supervision of doctors in clinical practice were referred to as real-time data, and an algorithm for building a visualization system on the Power BI platform was given. The Power BI visualization tool visually depicted the trajectory of change of that parameter according to parameters recorded at different points in time, creating the conditions for viewing the patient's vital medical indicators in context with time and space components, which supported doctors in assessing the criticality of the current situation. In further studies, it is planned to conduct research on the prediction of the patient's health based on clinical data through the integration of Power BI with artificial intelligence models.

References

[1].J.Reis, M. Amorim Melão, P. Matos, Digital transformation: a literature review and guidelines for future research, World Conference on Information Systems and Technologies, (2018) 411–421.

[2]. K. Zhang, H. Zhou, T. D. Baptista-Hon, etc. Concepts and applications of digital twins in healthcare and medicine, Patterns, 5, (2024), https://doi.org/10.1016/j.patter.2024.101028.

[3]. K. Y. H. Lim, P. Zheng, C. H. A. Chen, State-of-the-art survey of digital twin: Techniques, engineering product lifecycle management and business innovation perspectives, Journal of Intelligent Manufacturing, 31, (2020) 1313–1337. DOI: 10.1007/s10845-019-01512-w.

[4]. M. Grieves, J. Vickers, Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems, Transdisciplinary Perspectives on Complex Systems, (2017) 85–113.

[5].M. Grieves, Digital twin: Manufacturing excellence through virtual factory replication, 2014, pp. 8 .

[6]. L. El-Warrak , C.M. Farias. The State of the Art of Digital Twins in Health—A Quick Review of the Literature, Computers, 13 (2024) 9 https://doi.org/10.3390/computers13090228.

[7]. G. Coorey, G.A. Figtree, D.F. Fletcher, et al. The health digital twin to tackle cardiovascular

disease—a review of an emerging interdisciplinary field, Digit. Med. 126 (2022) 5, https://doi.org/10.1038/s41746-022-00640-7.

[8]. A. Ahmadova, Applications of digital twins in medicine and the ontological model of medical digital twins, Problems of Information Society, 15, (2024) 98-105, http://doi.org/10.25045/jpis.v15.i1.10.

[9].T. Sun, X. He, X, Z. Li, Digital twin in healthcare: Recent updates and challenges, Digit Health, 9 (2023) 1–13, doi:10.1177/20552076221149651.

[10].D. Silva, M. Azevedo, A. L. Soares, A Vision for a Platform-based Digital-Twin Ecosystem, IFAC-PapersOnLine, 54, (2021) 761-766, https://doi.org/10.1016/j.ifacol.2021.08.088 .

[11].B. Barricelli, E. Casiraghi, D. Fogli, A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications, IEEE Access, 7 (2019) 167653-167671, doi: 10.1109/ACCESS.2019.2953499 .

[12].R. Liyanage, N. Tripathi, Päivärinta, Digital Twin Ecosystems: Potential Stakeholders and Their Requirements. In: Carroll, N., Nguyen-Duc, A., Wang, X., Stray, V. (eds) Software Business. ICSOB 2022. Lecture Notes in Business Information Processing, 463. (2022) https://doi.org/10.1007/978-3-031-20706-8_2.

[13].M. Mammadova, A. Ahmadova, Development of digital twin ecosystem and ontology in medicine. Technology transfer: fundamental principles and innovative technical solutions (2023) 21–23. doi: https://doi.org/10.21303/2585-6847.2023.003203

[14].M. Mammadova, Z. Jabrayilova, Synthesis of decision making in a distributed intelligent personnel health management system on offshore oil platform. EUREKA: Physics and Engineering, 4 (2022) 179–192. https://doi.org/10.21303/2461-4262.2022.002520.

[15].M. Mammadova, Z. Jabrayilova, Decision Making in a Distributed Intelligent Personnel Health Management System on Offshore Oil Platform. Studies in Fuzziness and Soft Computing. 423 (2023) 145–153.

[16].O.S. Kobyakova, V.I. Starodubov, N.G. Kurakova, L.A. Tsvetkova, Digital Twins in Healthcare: Assessment of Technological and Practical Prospects, Annals of the Russian Academy of Medical Sciences, 76 (2021) 476–487.

[17].Y. Fan, J. Yang, J. Chen, X Wang, B. Zhou, A digital-twin visualized architecture for Flexible Manufacturing System, Journal of Manufacturing Systems, 60 (2021) 176-201, https://doi.org/10.1016/j.jmsy.2021.05.010Get rights and content.

[18].S. M. Zulkafli, M. M. Ariffin, A. Zakariya, Data Analytics and Visualization of Remote Healthcare Monitoring System, 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA, 2022, 1-6, doi: 10.1109/ICCUBEA54992.2022.10010938.

OPTIMIZED CNN-BASED APPROACH FOR ALZHEIMER'S DISEASE BY TACKLING CLASS IMBALANCE IN MRI CLASSIFICATION

Soraisam Gobinkumar Singh[1], Dulumani Das[1], Utpal Barman[1], Hasan Huseynov[2]

•

[1]Faculty of Computer Technology, Assam down town University, Panikhaiti, Guwahati, 781026, Assam, India.

[2]Azerbaijan Technical University, Baku, Azerbaijan

gobinkumarsoraisam@gmail.com, dulumoni.d@adtu.in , dean.foct@adtu.in, tk_xt@aztu.edu.az

Abstract

Accurate and early diagnosis of Alzheimer's Disease (AD) is crucial for effective intervention and treatment. This study presents a Convolutional Neural Network (CNN)-based approach for the classification of brain MRI images into four categories: Mild Demented, Moderate Demented, Non-Demented, and Very Mild Demented. To address the challenges of class imbalance inherent in the dataset, we employed class weighting and focal loss during training. Class weighting ensured that underrepresented classes received adequate attention, while focal loss emphasized harder-to-classify examples, resulting in improved model performance on minority classes. The model achieved remarkable results, with an accuracy of 97.66%, precision of 97.66%, recall of 97.66%, F1-score of 97.66%, specificity of 98.98%, and Cohen's Kappa of 96.14%, indicating a robust performance across all metrics. A comparative analysis with state-of-the-art methods demonstrated that our approach outperformed many existing models, including Siamese CNNs, 3D DenseNet ensembles, and other transfer-learning-based techniques. The ROC-AUC analysis further highlighted the model's ability to distinguish between classes with near-perfect curves for all categories. These results underscore the effectiveness of combining CNN architectures with class imbalance-handling strategies for medical image classification. The proposed method holds promise for improving diagnostic accuracy and early detection in AD, thereby supporting clinical decision-making.

Keywords: AD; CNN; ROC-AUC.

I. Introduction

Neurological disorders (NLD) impact the central nervous system, including the brain, spinal cord, and nerves (both cranial and peripheral). Even slight disruptions in the functioning of these critical systems can lead to severe physiological conditions. Alzheimer's disease (AD) is one prominent example of an NLD, currently affecting 55 million people worldwide, as highlighted in the latest World Alzheimer Report [1], [2]. This figure is projected to rise to 139 million by 2050, according to the World Health Organization (WHO). Additionally, dementia, which includes AD, incurs an annual global cost of $1.3 trillion as of 2019—a figure expected to exceed $2.8 trillion by 2030 due to the aging population. Every three seconds, someone develops dementia, underscoring its significant global impact[1]. As one of the leading causes of death worldwide, AD is an incurable, progressive, and life-altering neurodegenerative condition. Within the brain cells, protein

structures—referred to as plaques and tangles—gradually degrade when impacted by AD. This protein damage results in a substantial decline in cognitive abilities, ultimately causing significant impairments in both personal and social aspects of life[2], [3] .

Alzheimer's disease significantly impacts individuals, leading to memory impairment, behavioral disturbances, and various physical challenges, including difficulties with vision and mobility. One of the primary obstacles to early detection is the general lack of public awareness about the disease[2]. This often results in cognitive decline and associated behaviors being misinterpreted as typical aspects of aging or symptoms of other psychiatric conditions. Additionally, issues such as geographical isolation, a shortage of trained caregivers, and limited access to specialists and advanced diagnostic tools exacerbate the challenges faced by patients [1]. These barriers can severely affect their ability to maintain independence in daily and social activities. Therefore, early detection of AD is crucial to alleviate the burden on patients and their families.

Alzheimer's disease is primarily diagnosed through observation of patient symptoms, a process that can often take years to confirm [4]. However, advances in diagnostic research have identified several biomarkers, such as Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), Computed Tomography (CT), and blood tests that support early detection. When integrated with Artificial Intelligence (AI), these biomarkers enable healthcare professionals to achieve more precise diagnoses and improve patient care. Machine learning (ML) classifiers have been widely adopted across healthcare, demonstrating significant effectiveness in AD classification [2]. In recent years, Deep learning (DL) techniques have become increasingly prominent in the healthcare field due to their ability to develop accurate end-to-end models using complex datasets [5]. This surge in DL applications has revolutionized the identification of neurological disorders, including AD, by enhancing diagnostic accuracy. Coupling DL with neuroimaging has provided critical insights into brain activity and associated disorders [6]. Various computer-aided diagnosis (CAD) systems have been proposed for predicting AD using neuroimaging techniques like functional MRI (fMRI), structural MRI (sMRI), and PET. Structural MRI, in particular, provides crucial details such as brain white matter (WM), gray matter (GM), cortical thickness, and volumetric measurements. These metrics are essential for assessing the neurodegenerative processes that contribute to AD. By enhancing high-resolution imaging data and the robust feature extraction capabilities of DL, clinicians can make informed decisions about complex AD cases.

To classify the imbalanced classes MildDemented, ModerateDemented, NonDemented, and VeryMildDemented, a comparison was made between a standard Convolutional Neural Network (CNN) model and a ResNet-based transfer learning approach using an MRI dataset sourced from Kaggle [7]. While CNNs are capable of extracting spatial features, ResNet's residual learning framework allows deeper network architectures, addressing vanishing gradient issues and making it well-suited for handling complex neuroimaging datasets[8]. To address the class imbalance inherent in the dataset, a combination of class weighting and focal loss was utilized. Class weights were assigned inversely proportional to the class frequencies, ensuring better representation of minority classes. Focal loss further emphasized hard-to-classify samples, reducing the dominance of well-classified instances during training[9]. This dual approach enhanced the model's ability to handle imbalanced data. Moreover, the Alzheimer’s Disease Neuroimaging Initiative (ADNI) and Open Access Series of Imaging Studies (OASIS) datasets, recognized for their comprehensive neuroimaging data, provided the foundational context for this work, emphasizing the importance of MRI-based analysis in Alzheimer's disease classification [10], [11]. Results demonstrated that ResNet outperformed the standard CNN model, enhancing transfer learning to extract high-level features and achieve better generalization across imbalanced classes.

The objective of this study is to explore and evaluate the use of a Kaggle-sourced dataset for Alzheimer's disease classification, conducting a comparative analysis of different modeling approaches alongside experiments to validate the effectiveness of the implemented techniques. The

following are the key contributions of this work:

1. A comparative analysis of CNN architectures, evaluating their effectiveness in classifying imbalanced classes of Alzheimer's disease stages using the Kaggle dataset.
2. Implementation of a combined approach utilizing class balancing and focal loss to address class imbalance, enhancing the model's ability to accurately classify minority classes.

Comprehensive evaluation of the proposed methods on the Kaggle MRI dataset, providing insights into their practicality and performance in real-world Alzheimer's disease classification scenarios.

II. Related Works

Recent advancements in neuroimaging and machine learning have enabled significant progress in Alzheimer's disease diagnosis and classification. Numerous studies have explored the use of deep learning models, such as CNNs and transfer learning-based architectures, to analyze neuroimaging data and address challenges like class imbalance and feature extraction.

[4] proposed a Siamese Convolutional Neural Network (SCNN) using a triplet-loss function to generate k-dimensional embeddings of MRI images for 4-way Alzheimer's Disease classification. Both pre-trained and non-pretrained CNNs were utilized for embedding generation. The model achieved accuracies of 91.83% on the ADNI dataset and 93.85% on the OASIS dataset, outperforming comparable methods in the literature.

[6] proposed a 3D DenseNet ensemble achieved 83.33% accuracy in 4-way classification using the ADNI dataset, distinguishing AD, healthy controls, EMCI, and LMCI. Dense connections and a probability-based fusion method enhanced feature extraction and improved performance over state-of-the-art models.

[12] utilized VGG16, Xception, and a customized CNN model with transfer learning to classify four stages of Alzheimer's Disease using 2D MRI images. The customized CNN achieved superior performance, with 94.77% accuracy and an F1-score of 0.9481. This approach demonstrated improved efficiency and reduced complexity compared to 3D MRI-based CNN models and conventional SVM techniques.

[13] introduced a swarm multi-verse optimizer with a deep neuro-fuzzy network (CSMVO + DNFN) for Alzheimer's Disease classification using MRI. Preprocessing involved a median filter, followed by segmentation with a channel-wise feature pyramid network module (CFPNet-M). Extracted features included Haralick, CNN, and texture attributes. The model achieved 89.9% accuracy, 89.6% sensitivity, and 87.0% specificity, demonstrating efficiency in classifying AD stages.

[14] investigated automated pre-detection of Alzheimer's Disease symptoms using the ADNI dataset. An initial experiment employed SVM for AD detection, achieving 84.4% accuracy, 95.3% sensitivity, and 71.4% specificity. Due to suboptimal results, a CNN-based approach was explored, incorporating various image segmentation methods. The best segmentation method achieved 96% accuracy, 96% sensitivity, and 98% specificity, highlighting the potential of deep learning for early AD diagnosis.

[15] purposed ensemble transfer-learning techniques for early Alzheimer's Disease diagnosis using structural brain MRI from the ADNI dataset. They compared an ensemble of five pretrained architectures, a 3D CNN trained from scratch, and a fusion of conventional SVM-based classifiers. The transfer-learning ensemble achieved 90.2% AUC for AD vs. CN, 83.2% for MCIc vs. CN, and 70.6% for MCIc vs. MCInc, performing slightly lower than the SVM-based fusion. The 3D CNN underperformed due to limited training data, highlighting transfer learning's potential in generic image pretraining for neuroimaging tasks.

III. Materials and methods

In this study, a Convolutional Neural Network (CNN) model was employed for classifying Alzheimer's disease into four classes (NonDemented, VeryMildDemented, MildDemented, and ModerateDemented) using an imbalanced MRI dataset sourced from Kaggle shown in figure 1.

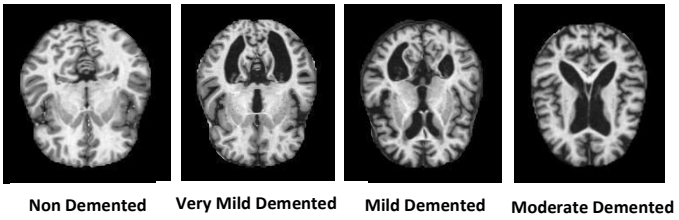


Figure 1: *Show sample images from each of the four imbalance classes*

To address the imbalance issue, a combination of class weighting and focal loss was integrated into the model training process. The dataset was preprocessed by resizing all images to a uniform size of 128×150×3 pixels from the original size 176x208x3, ensuring compatibility with the CNN model. Training and validation datasets contained 5,119 for training and 1,281 for validation images, respectively, distributed across the four classes shown in Figure 2.

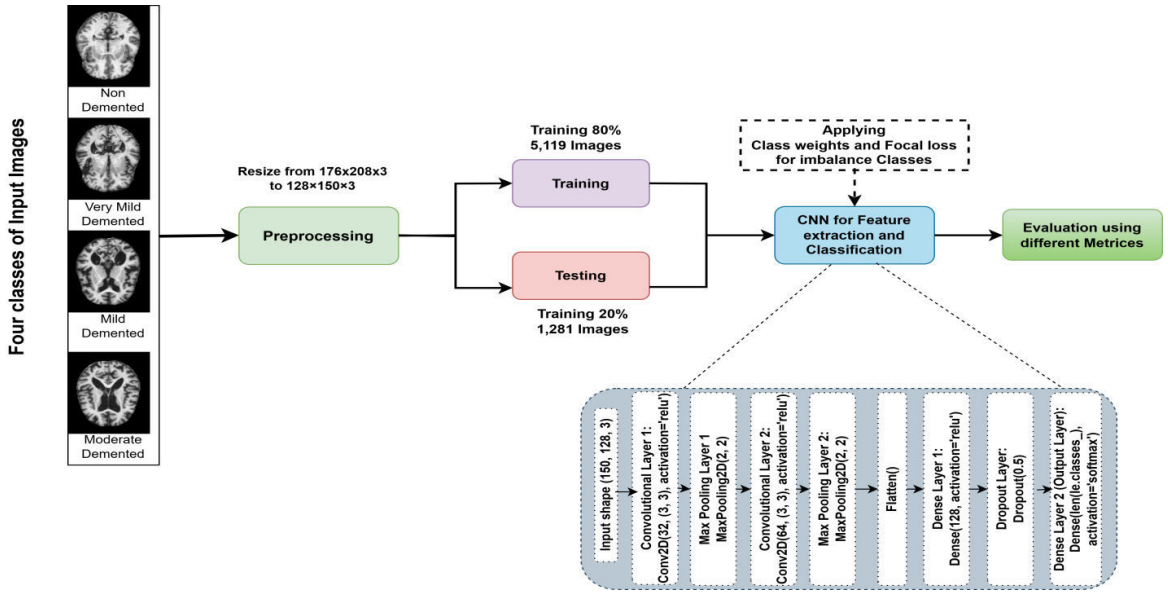


Figure 2: *The purpose model using CNN with Class weights and Focal Loss*

The class distribution of training and validation datasets for four classes: VeryMildDemented, NonDemented, ModerateDemented, and MildDemented. The training dataset is significantly imbalanced, with the majority class being NonDemented (2560 images) and the minority class being ModerateDemented (51 images). Similarly, the validation dataset shows a similar imbalance, with NonDemented having the highest count (640 images) and ModerateDemented the lowest (13 images) as depicted in Figure 2. This imbalance highlights the need for techniques like class weighting and focal loss to improve model performance.

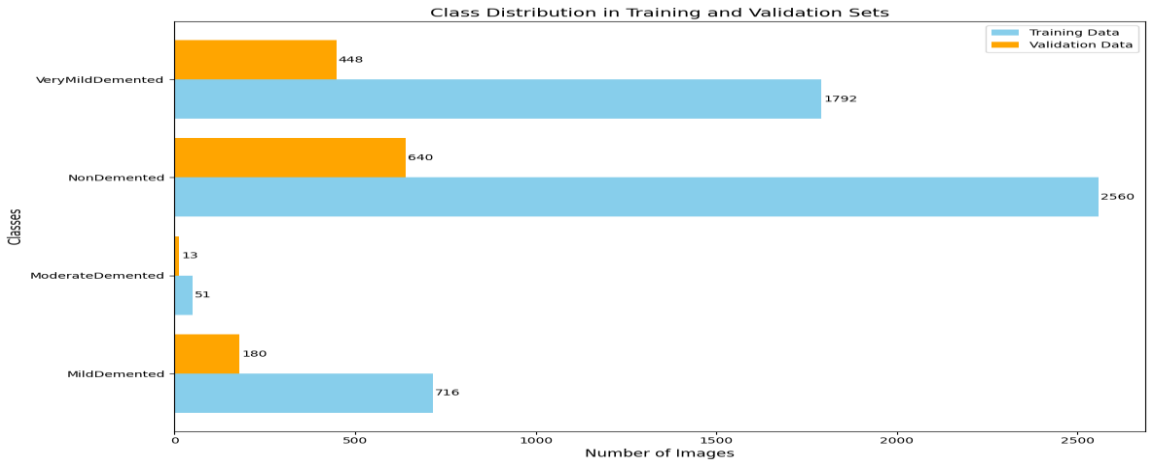


Figure 3: *Show the different class distribution of training and validation datasets*

The class weights computed to address the class imbalance in the dataset. The weight for Moderate Demented is the highest at 25.09, as it is the most underrepresented class. In contrast, Non-Demented has the lowest weight (0.50), reflecting its dominance in the dataset. These weights ensure that the model pays proportionally more attention to minority classes during training Figure 3. Additionally, a focal loss function is employed to further handle the class imbalance. Focal loss dynamically scales the standard cross-entropy loss by focusing more on hard-to-classify examples. It does so by down-weighting the loss for well-classified samples (where the predicted probability is high) and up-weighting the loss for misclassified ones. This is controlled by parameters alpha (0.25) and gamma (2.0). Combining class weights and focal loss enhances the model's ability to learn from imbalanced data effectively.

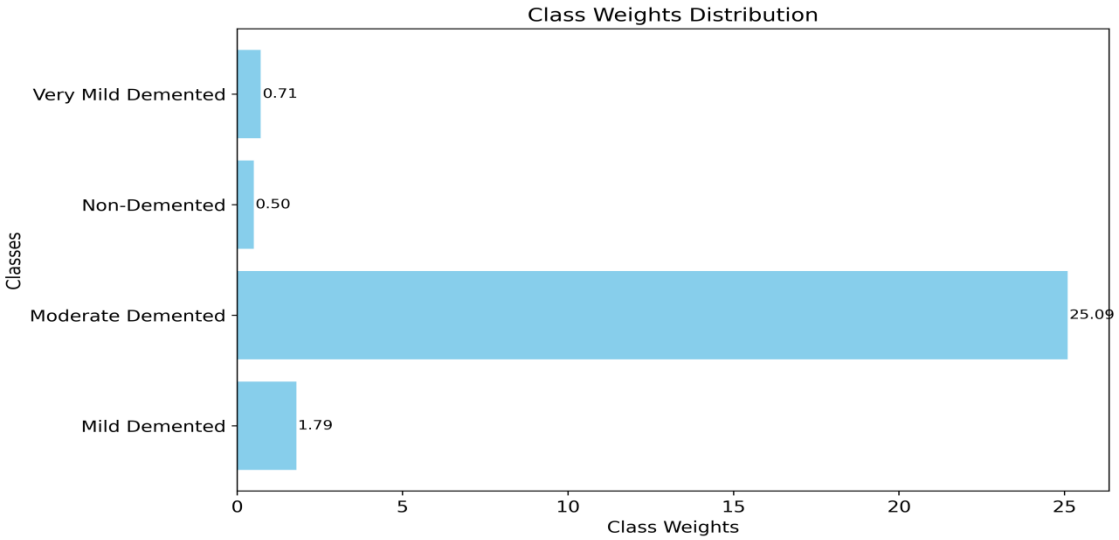


Figure 4: *Show the weighted class value distributions for the imbalance class*

The spatial dimensions of the input image (150×128×3) evolve through each layer of the CNN model up to the flatten layer. The first convolution reduces the dimensions to 148×126×32, followed by max pooling, which downsamples it to 74×63×32. The second convolution further reduces it to 72×61×64, and another pooling layer brings it to 36×30×64. Finally, the Flatten layer reshapes this 3D tensor into a 1D vector of size 69,120, which serves as the feature input for fully connected layers. After the feature input first dense layer is a fully connected layer with 128 neurons and ReLU activation, enabling the model to learn complex features. A dropout layer with a 50% rate is applied

next to prevent overfitting by randomly dropping connections during training. Finally, the output layer is a fully connected layer with neurons equal to the number of classes (4 in this case) and softmax activation to produce class probabilities for multi-class classification.

IV. Results

The Results section presents the evaluation of the proposed model's performance, highlighting its ability to address the challenges posed by the imbalanced medical image dataset. Key metrics such as accuracy, precision, recall, F1 score, specificity, and Cohen's kappa are analyzed to demonstrate the effectiveness of the CNN based approach with class balancing and focal loss model. These metrics rely on the confusion matrix, which summarizes the performance of a classification model using the following components: True Positives (TP), True Negatives (TN), False Positives (FP), False Negatives (FN).

1.Performance Analysis

Accuracy: Accuracy measures the proportion of correctly classified instances out of the total instances shown in equation(i).

$$\text{Accuracy (ACC)} = (Tp+Tn)/ (Fp + Fn + Tp + Tn) \quad (1)$$

Precision (Positive Predictive Value): Precision indicates the proportion of true positive predictions out of all positive predictions shown in equation(ii).

$$\text{Precision} = Tp/(Tp+Fp) \quad (2)$$

Recall (Sensitivity or True Positive Rate): Recall shows the proportion of actual positives correctly identified shown in equation(iii).

$$\text{Sensitivity (Recall)} = Tp/(Tp+Fn) \quad (3)$$

F1-score: F1-score combines precision and recall into a single metric, emphasizing their harmonic mean shown in equation(iv).

$$F\ 1 - \text{Score} = 2 * Tp/2*(Tp + Fp + Fn) \quad (4)$$

Specificity (True Negative Rate): Specificity measures the proportion of actual negatives correctly identified in equation(v).

$$\text{Specificity} = Tn/ (Tn+Fp) \quad (5)$$

Cohen's Kappa: It relies on observed and expected agreements, accounting for randomness in predictions.

$$\text{Kappa}=p_o - p_e / 1 - p_e \quad (6)$$

Where:

p_o = observed agreement (accuracy)

p_e = expected agreement (calculated based on random chance)

Indicates the level of agreement between predictions and ground truth, accounting for chance agreement.

2. CNN with class balancing and focal loss

The model employs a CNN architecture integrated with class balancing techniques and focal loss, ensuring robust learning even in the presence of class imbalances. The plots demonstrate the training and validation performance of the model in Figure 4. The loss curve indicates a consistent decrease in training and validation loss, confirming the model's ability to learn effectively. The accuracy curve highlights a steady improvement in accuracy for both training and validation datasets, demonstrating good generalization without overfitting. The model was compiled using the focal loss function with parameters α=0.25 and γ=2.0, specifically designed to address class imbalance by focusing more on hard-to-classify samples.

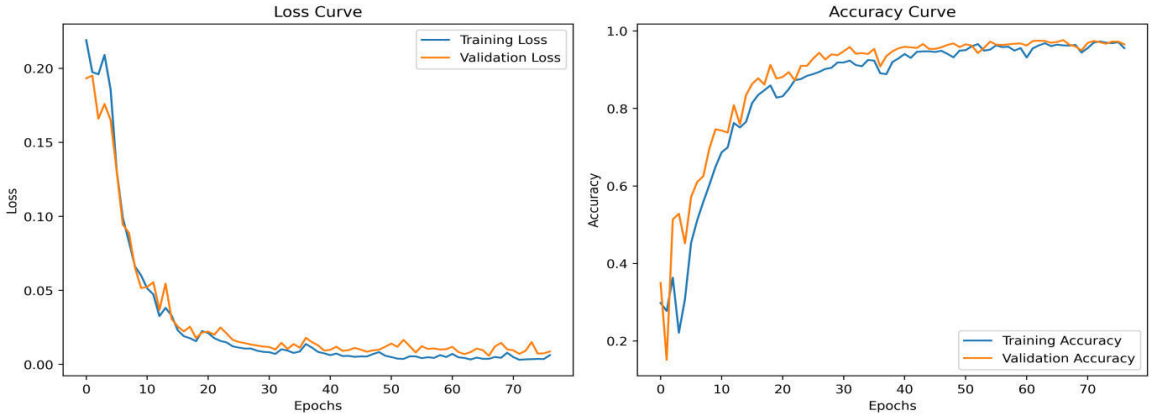


Figure 5: *Show the Performance Analysis of Loss and Accuracy Over Epochs using CNN*

An Adam optimizer with a learning rate of 0.001 was used, enabling adaptive learning. Training was performed with a batch size of 32 and incorporated early stopping with a patience of 10 epochs, ensuring the model did not overtrain while restoring the best weights. Additionally, class weights were applied to further handle the class imbalance effectively.

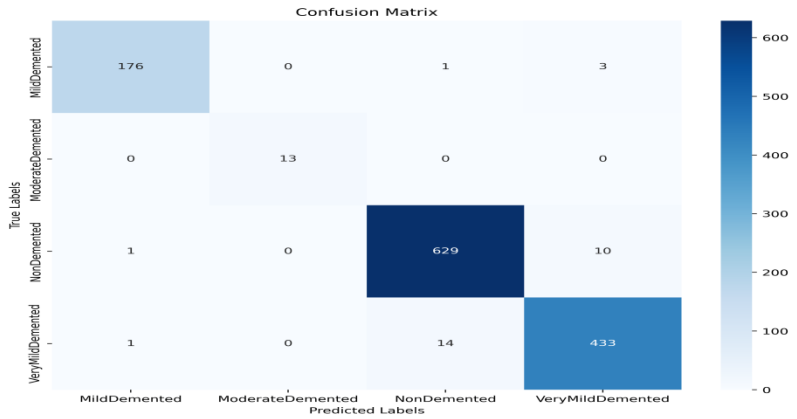


Figure 6: *Confusion matrix of the four classes' datasets*

The performance metrics of the trained CNN model, evaluated on the test dataset, are shown in Figure 6. The model achieved exceptional accuracy (97.66%) in classifying the data, demonstrating its robustness and reliability. Precision, recall, and F1-score are all equally high at 97.66%, indicating a balanced performance in identifying both positive and negative cases accurately. Specificity, at 98.98%, highlights the model's ability to correctly identify true negatives, which is crucial for avoiding false positives in medical diagnoses. The Cohen's kappa score of 0.9614 further confirms strong agreement between the predicted and true labels, accounting for chance agreement. These results showcase the effectiveness of using focal loss and class balancing strategies in enhancing the model's performance on imbalanced datasets.

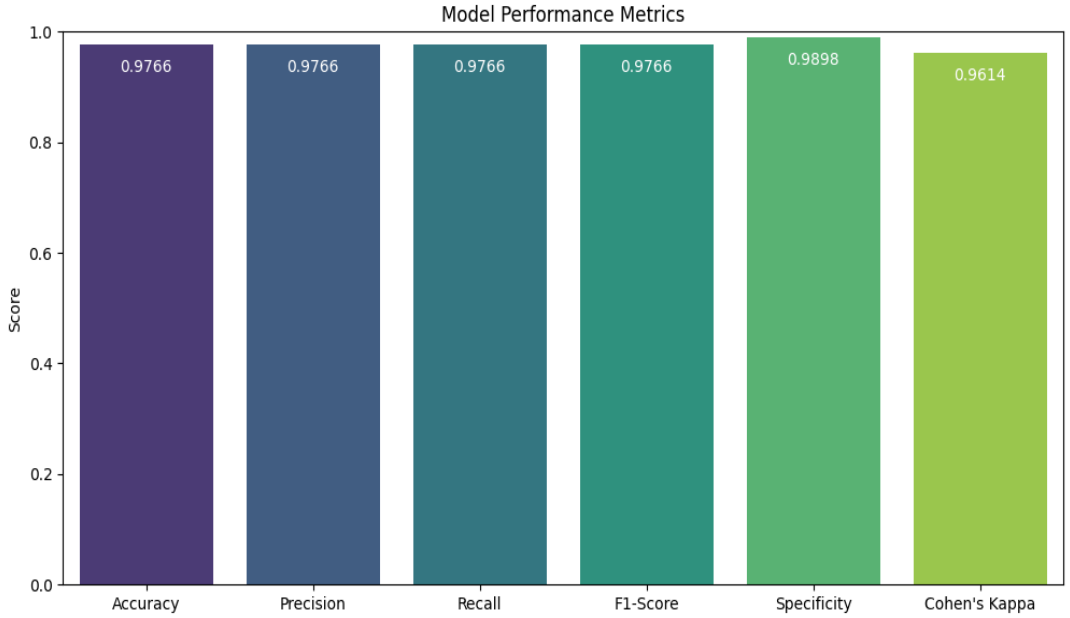


Figure 7: *The performance of CNN model with utilizing class balancing and focal loss*.

The ROC AUC curve shows the model's exceptional performance in distinguishing between the four classes: "Mild Demented," "Moderate Demented," "Non-Demented," and "Very Mild Demented." Each curve represents the one-vs-all ROC for a class, with AUC scores above 0.99 for all classes, indicating near-perfect classification as shown in Figure 7.

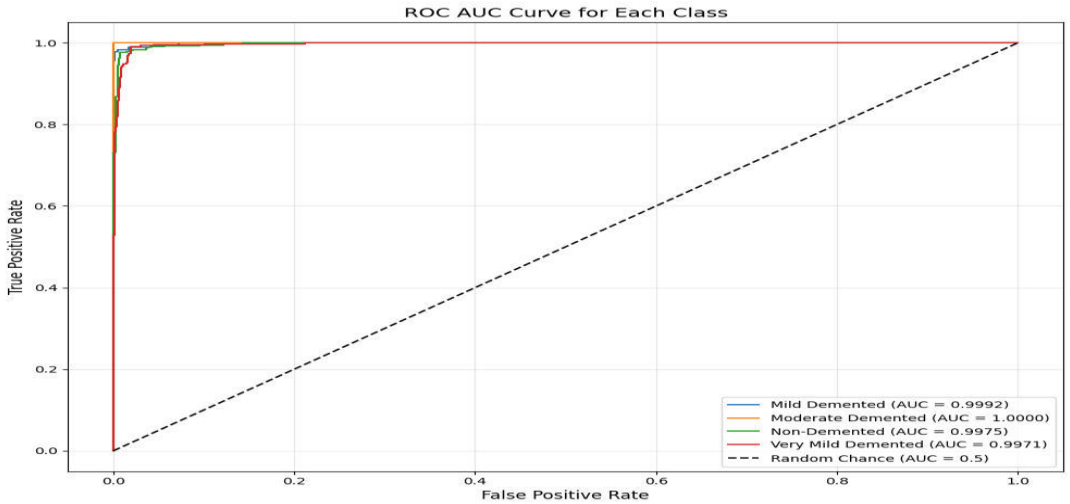


Figure 8: *The ROC AUC curve of the four AD classes*

The "Moderate Demented" class achieves a perfect AUC of 1.0, while the other classes—though slightly lower—still display excellent discrimination. The curves are close to the top-left corner, reflecting a high true positive rate (TPR) and low false positive rate (FPR) for each class. This validates the model's robustness, even with class imbalance, confirming its suitability for this multi-class classification task.

V. Discussions

In this study, a Convolutional Neural Network (CNN) was employed to classify brain MRI images into four classes: "Mild Demented," "Moderate Demented," "Non-Demented," and "Very

Mild Demented." Addressing the challenge of class imbalance, we integrated class weighting and focal loss into the training process. Class weighting ensured that underrepresented classes received higher penalties during misclassification, thereby guiding the model to pay balanced attention to all classes. Meanwhile, focal loss effectively reduced the impact of easy-to-classify samples and focused on harder examples, further enhancing the model's performance on minority classes.

Table 1: *Comparing of the combining CNN with class weighting and focal loss with other approaches*

Study	Model/Method	Dataset	Accuracy	Key Features
[4]	Siamese Convolutional Neural Network (SCNN) with triplet-loss	ADNI, OASIS	91.83% (ADNI), 93.85% (OASIS)	Generated k-dimensional embeddings with pre-trained and non-pretrained CNNs for robust classification.
[6]	3D DenseNet Ensemble	ADNI	83.33%	Used dense connections and probability-based fusion for better feature extraction.
[12]	Customized CNN with Transfer Learning	OASIS	94.77%	Outperformed traditional 3D CNN and SVM models; focused on efficiency with 2D MRI.
[13]	CSMVO + DNFN (Swarm Multi-Verse Optimizer with Deep Neuro-Fuzzy Network)	Unspecified	89.90%	Integrated segmentation and advanced feature extraction (Haralick, CNN, texture features).
[14]	CNN with optimized segmentation methods	ADNI	96.00%	Improved pre-detection of AD symptoms with segmentation techniques.
[15]	Ensemble transfer-learning methods	ADNI	90.20% (AUC)	Used pretrained architectures and conventional SVM-based classifiers for early diagnosis.
Our Model	**CNN with weighted class and focal loss**	**OASIS (kaggle)**	**97.66%**	**Balanced class performance using weighted loss and focal loss for imbalanced datasets.**

From the comparison, it is evident that the proposed CNN model achieves a superior accuracy of 97.00%, outperforming many state-of-the-art methods such as the Siamese CNN in [4] and the customized CNN in [12]. This improvement can be attributed to the integration of class weighting and focal loss, which effectively addressed the class imbalance challenge. Additionally, the proposed method demonstrated consistent performance across all classes, as seen in the ROC AUC curves,

further validating its robustness. While segmentation techniques in [14] achieved slightly higher accuracy, the complexity of preprocessing makes our method more efficient and easier to implement in real-world scenarios.

VI. Conclusions

In this study, we proposed a CNN model integrated with class weighting and focal loss to address the challenges of class imbalance in the classification of brain MRI images into four stages of Alzheimer's Disease: Mild Demented, Moderate Demented, Non-Demented, and Very Mild Demented. The performance metrics of the model, as depicted in the bar chart, demonstrate exceptional results across all key measures: accuracy (97.66%), precision (97.66%), recall (97.66%), F1-score (97.66%), specificity (98.98%), and Cohen's Kappa (96.14%). These high values indicate the model's robustness in correctly identifying all classes, even in the presence of imbalanced data.

By enhancing class weighting, the model ensured that underrepresented classes were prioritized during training, minimizing the risk of bias towards majority classes. The incorporation of focal loss further enhanced the model's ability to focus on harder-to-classify samples, improving performance on minority classes. Compared to state-of-the-art methods, the proposed approach achieves competitive or superior performance while maintaining simplicity and computational efficiency. These results underscore the potential of combining CNN architectures with tailored loss functions for effective medical image classification and early diagnosis of Alzheimer's Disease.

References

[1] "World Alzheimer Report 2024: Global changes in attitudes to dementia," Sep. 2024, Accessed: Nov. 28, 2024. [Online]. Available: https://www.alzint.org/resource/world-alzheimer-report-2024/

[2] S. G. Singh, D. Das, U. Barman, and M. J. Saikia, "Early Alzheimer's Disease Detection: A Review of Machine Learning Techniques for Forecasting Transition from Mild Cognitive Impairment," *Diagnostics*, vol. 14, no. 16, p. 1759, 2024.

[3] L. Rizzi, I. Rosset, and M. Roriz-Cruz, "Global Epidemiology of Dementia: Alzheimer's and Vascular Types," *BioMed Res. Int.*, vol. 2014, pp. 1–8, 2014, doi: 10.1155/2014/908915.

[4] F. Hajamohideen *et al.*, "Four-way classification of Alzheimer's disease using deep Siamese convolutional neural network with triplet-loss function," *Brain Inform.*, vol. 10, no. 1, p. 5, Dec. 2023, doi: 10.1186/s40708-023-00184-w.

[5] M. Mahmud, M. S. Kaiser, A. Hussain, and S. Vassanelli, "Applications of deep learning and reinforcement learning to biological data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2063–2079, 2018.

[6] J. Ruiz, M. Mahmud, M. Modasshir, M. Shamim Kaiser, and F. T. Alzheimer's Disease Neuroimaging In, "3D DenseNet Ensemble in 4-Way Classification of Alzheimer's Disease," in *Brain Informatics*, vol. 12241, M. Mahmud, S. Vassanelli, M. S. Kaiser, and N. Zhong, Eds., in Lecture Notes in Computer Science, vol. 12241. , Cham: Springer International Publishing, 2020, pp. 85–96. doi: 10.1007/978-3-030-59277-6_8.

[7] "Alzheimer MRI 4 classes dataset." Accessed: Nov. 28, 2024. [Online]. Available: https://www.kaggle.com/datasets/marcopinamonti/alzheimer-mri-4-classes-dataset

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778. Accessed: Nov. 28, 2024. [Online]. Available: http://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html

[9] T. Lin, "Focal Loss for Dense Object Detection," *ArXiv Prepr. ArXiv170802002*, 2017, Accessed: Nov. 28, 2024. [Online]. Available: http://openaccess.thecvf.com/content_iccv_2017/html/Lin_Focal_Loss_for_ICCV_2017_paper.html

[10] C. R. Jack *et al.*, "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *J. Magn. Reson. Imaging*, vol. 27, no. 4, pp. 685–691, Apr. 2008, doi: 10.1002/jmri.21049.

[11] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *J. Cogn. Neurosci.*, vol. 19, no. 9, pp. 1498–1507, 2007.

[12] R. A. Hridhee, B. Bhowmik, and Q. D. Hossain, "Alzheimer's Disease Classification From 2D MRI Brain Scans Using Convolutional Neural Networks," in *2023 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, IEEE, 2023, pp. 1–6. Accessed: Nov. 28, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10101539/

[13] A. Allada, R. Bhavani, K. Chaduvula, and R. Priya, "Alzheimer's disease classification using competitive swarm MULTI-VERSE optimizer-based deep NEURO-FUZZY network," *Concurr. Comput. Pract. Exp.*, vol. 35, no. 21, p. e7696, Sep. 2023, doi: 10.1002/cpe.7696.

[14] K. Gunawardena, R. N. Rajapakse, and N. D. Kodikara, "Applying convolutional neural networks for pre-detection of alzheimer's disease from structural MRI data," in *2017 24th international conference on mechatronics and machine vision in practice (M2VIP)*, IEEE, 2017, pp. 1–7. Accessed: Nov. 28, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8211486/

[15] L. Nanni *et al.*, "Comparison of transfer learning and conventional machine learning applied to structural brain MRI for the early diagnosis and prognosis of Alzheimer's disease," *Front. Neurol.*, vol. 11, p. 576194, 2020.

EFFECTIVE USE OF ARTIFICIAL INTELLIGENCE METHODS FOR THE IMPLEMENTATION OF PREREQUISITES IN CURRICULUMS

Zafar Jafarov[1], Vahid Garuslu[2], Atif Namazov[1]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan
[2]Queen's University Belfast, Belfast, UK
zafar.cafarov@aztu.edu.az, v.garousi@qub.ac.uk, atif.namazov@aztu.edu.az

Abstract

Prerequisites play a critical role in the construction of a curriculum for courses where student success heavily relies on previously acquired knowledge or skills. The application of artificial intelligence (AI) methods for the implementation of prerequisites in curriculums can significantly enhance the efficiency and effectiveness of educational planning. The article argues that association rules can be effectively applied in the implementation of prerequisites in the curriculum.

Keywords: curriculums, program schema, prerequisites, artificial intelligence, association rules.

I. Introduction

The curriculum in academic education refers to the structured set of courses, learning experiences, and assessments designed to provide students with the knowledge, skills, and competencies required to achieve educational objectives. It is a comprehensive plan that outlines what students are expected to learn, how they will learn it, and how their learning will be assessed. Curriculum is usually plotted against Bachelor of Science (B.Sc.) program.

A Bachelor of Science (B.Sc.) program typically includes a mix of core courses, major-specific courses, electives, labs, and possibly a capstone project or thesis. A program schema for a Bachelor of Science (B.Sc.) typically includes the structure and organization of courses that a student needs to complete to earn the degree.

Prerequisites in a Bachelor of Science (B.Sc.) program schema are essential for several reasons, as they establish a foundation of knowledge and skills necessary for students to succeed in their academic and professional endeavors. By leveraging AI methods, educational institutions can create more effective and personalized learning experiences, ensuring that students are well-prepared for advanced coursework and ultimately improving their academic success. Association rules can be very useful for determining the implementation of prerequisites in educational curriculums.

II. Program schema and prerequisites

A Bachelor of Science program schema typically consists of various components designed to provide a comprehensive education in the sciences. Here's a general outline of what a B.Sc. program

schema might look like [1]:

1. Core Courses (CC)

These are mandatory courses that provide foundational knowledge in the chosen field of study.

2. Major-Specific Courses (MSC)

These courses provide specialized knowledge in the chosen field of study. Here are examples for different majors

Courses specific to the student's major that provide in-depth knowledge and specialized skills.

3. Electives Courses (EC)

Students can choose electives based on their interests, which can be within or outside their major. Electives allow students to explore other disciplines or deepen their knowledge in their field of study.

Table 1 shows a fragment of the program scheme for the Bachelor of Science in Information Technology.

Table 1

Fourth Semester				
S.No.	Course type	Course Title	Prerequisite	Credit Hours
1	CC	Probability & Statistics	None	3
2	MSC	Database Systems	IT Fundamentals	7
3	MSC	Computer Networks	Computer architecture	8
4	CC	Discrete mathematics	None	3
5	CC	Business and academic communication in English - IV	Business and academic communication in English - III	3
6	EC	IT Elective I	None	6

Prerequisites play a critical role in the construction of a curriculum for courses where student success heavily relies on previously acquired knowledge or skills. Here are some key reasons why prerequisites are important:

- Ensuring Readiness: Prerequisites help ensure that students have the necessary background knowledge and skills to grasp the new material. This readiness is crucial for courses that build on complex concepts introduced in earlier coursework.
- Maintaining Academic Standards: By requiring students to complete certain courses before advancing, institutions can maintain a high standard of education and ensure that students are adequately prepared for the challenges of more advanced material.
- Improving Success Rates: Students who meet the prerequisite requirements are more likely to succeed in subsequent courses. They are better equipped to understand and engage with the material, which can lead to higher grades and retention rates.
- Streamlining Learning: Prerequisites help create a logical progression through a curriculum, allowing for a more structured and coherent learning experience. This structure helps students build on their knowledge systematically.
- Preventing Overwhelm: Courses without prerequisites might result in students feeling overwhelmed by material for which they are not prepared, potentially leading to frustration, disengagement, and higher dropout rates.
- Efficient Use of Resources: Prerequisites help instructors plan their courses more effectively, as they can assume a certain level of knowledge among their students. This allows for more efficient use of class time and resources.

Prerequisites are essential in curriculum design to ensure that students are adequately

prepared, which in turn enhances their chances of success and maintains the integrity and quality of the educational program.

Here are the key points explaining the importance of prerequisites:

1. Foundation of Knowledge

- Building Blocks: Prerequisites ensure that students have the basic understanding and foundational knowledge required for advanced courses. Without this foundation, students may struggle with complex concepts and techniques.
- Sequential Learning: Many subjects, especially in science and mathematics, build upon previously acquired knowledge. Prerequisites ensure that students progress through their education in a logical and effective sequence.

2. Academic Preparedness

- Readiness for Advanced Material: Prerequisites verify that students are academically prepared for the challenges of more advanced courses, reducing the likelihood of failure or the need for remedial instruction.
- Enhanced Learning Experience: When students enter a course with the necessary background, the entire class can engage more deeply with the material, leading to a richer and more productive learning experience.

3. Efficient Use of Resources

- Optimal Use of Faculty Time: Ensuring students meet prerequisites allows instructors to focus on teaching advanced material rather than re-teaching basic concepts.
- Classroom Dynamics: Classes can proceed at the intended pace without having to accommodate students who are not adequately prepared, leading to a more efficient and effective use of instructional time.

4. Student Success and Retention

- Higher Success Rates: Students who meet prerequisites are more likely to succeed in their courses, which can improve overall retention and graduation rates.
- Confidence and Motivation: When students feel prepared and capable of handling course material, their confidence and motivation increase, positively impacting their academic performance and engagement.

5. Professional and Academic Standards

- Maintaining Standards: Prerequisites help maintain high academic standards within the program, ensuring that graduates possess the knowledge and skills expected by employers and graduate schools.
- Accreditation Requirements: Many academic programs have accreditation standards that require adherence to specific prerequisite structures to ensure the quality and rigor of the education provided.

6. Interdisciplinary Integration

- Cohesive Knowledge Base: For programs that integrate multiple disciplines, prerequisites ensure that students have the necessary background in all relevant areas, facilitating better interdisciplinary understanding and collaboration.
- Preparation for Capstone Projects: Prerequisites ensure students are adequately prepared for capstone projects or theses, which often require integrating knowledge from various courses and disciplines.

7. Time Management and Planning

- Efficient Course Planning: Understanding prerequisites helps students plan their course schedules effectively, ensuring they meet all necessary requirements in a timely manner without delaying their graduation.
- Avoiding Overload: Properly structured prerequisites help prevent students from enrolling in courses they are not ready for, which can lead to academic overload and burnout.

8. Career Readiness

- Skill Competency: Prerequisites ensure that students develop essential skills progressively, making them more competent and competitive in their chosen careers.
- Preparation for Professional Exams: For fields requiring certification or licensure, prerequisites prepare students for professional exams by ensuring they have covered all necessary material.

Prerequisites in a B.Sc. program schema are crucial for ensuring that students are adequately prepared, facilitating effective learning, maintaining academic standards, and ultimately contributing to student success and readiness for professional and academic challenges.

III. The application of AI methods for the implementation of prerequisites in curriculums

The application of AI methods for the implementation of prerequisites in curriculums can significantly enhance the efficiency and effectiveness of educational planning. Here are several ways AI can be applied [2]:

- Data Analysis and Predictive Analytics:
- Student Performance Prediction: AI algorithms can analyze historical student performance data to predict which prerequisites are most strongly correlated with success in advanced courses. This can help in identifying essential prerequisite courses.
- Early Warning Systems: AI can identify students who are likely to struggle in a course based on their performance in prerequisite courses, allowing for early interventions.
- Personalized Learning Paths:
- Adaptive Learning Systems: AI-driven platforms can create personalized learning paths based on individual student performance, strengths, and weaknesses. These systems can dynamically adjust prerequisites for each student, ensuring they receive the preparation they need.
- Recommendation Engines: Similar to how e-commerce sites recommend products, AI can recommend courses that students should take next based on their academic history and performance in prerequisite courses.
- Curriculum Development and Optimization:
- Curriculum Mapping: AI can assist in mapping out the entire curriculum, identifying gaps, redundancies, and optimal prerequisite structures. This ensures a more streamlined and efficient curriculum.
- Scenario Analysis: AI can simulate different curriculum structures and prerequisite requirements to determine the most effective configurations for student success.
- Natural Language Processing (NLP) and Text Mining:
- Prerequisite Identification: NLP techniques can analyze course descriptions, syllabi, and other educational materials to automatically identify and recommend prerequisites for new or existing courses.
- Curriculum Alignment: Text mining can help ensure that the content of prerequisite courses aligns well with the requirements of subsequent courses.
- Automated Advising Systems:
- Virtual Advisors: AI-powered chatbots and virtual advisors can provide students with guidance on course selection, ensuring they understand and fulfill prerequisite requirements.
- Degree Planning Tools: AI-driven tools can help students plan their entire academic journey, taking into account prerequisites, course availability, and personal preferences.
- Assessment and Feedback:
- Intelligent Tutoring Systems: These systems can assess students' knowledge and skills in

real-time, providing immediate feedback and identifying areas where prerequisite knowledge may be lacking.

Formative Assessments: AI can design and administer formative assessments to gauge students' readiness for advanced courses, ensuring they have mastered prerequisite content.

IV. Association rules for the Implementation of Prerequisites in curriculums

Association rules, a concept widely used in data mining, can be effectively applied to the implementation of prerequisites in curriculums. These rules can help in identifying patterns and relationships between different courses, thereby aiding in the design of prerequisite structures. Here's how association rules can be utilized:

- Identifying Strong Course Relationships:

- Mining Historical Data: By analyzing historical enrollment and performance data, association rules can help identify strong relationships between courses. For example, if a significant number of students who succeed in Course A also succeed in Course B, it might suggest that Course A could be a good prerequisite for Course B.

- Support and Confidence Metrics: Use metrics such as support (the frequency with which items appear together) and confidence (the likelihood of item B appearing in transactions containing item A) to determine the strength of associations between courses.

- Developing Prerequisite Structures:

- Frequent Itemsets: Identify sets of courses that frequently appear together in successful student records. These sets can suggest natural groupings of prerequisites and subsequent courses.

- Rule Generation: Generate rules that can guide the creation of prerequisite structures. For example, a rule like {Course A, Course B}→{Course C} might indicate that students who complete Courses A and B are well-prepared for Course C.

- Enhancing Student Advising:

- Personalized Recommendations: Use association rules to provide personalized course recommendations to students. If a student has successfully completed certain courses, the system can recommend the next courses based on association rules mined from data of similar students.

- Path Analysis: Analyze the paths of successful students to recommend optimal course sequences for current students, ensuring they meet prerequisite requirements efficiently.

- Improving Course Success Rates:

- Predictive Analytics: Use association rules to predict potential challenges students might face in certain courses based on their performance in prerequisite courses. Interventions can then be designed to support at-risk students.

Early Detection: Detect early signs of struggle in prerequisite courses and provide additional support to ensure students are adequately prepared for advanced courses.

V. Example Workflow for Implementing Association Rules

Data Collection: Gather historical data on student enrollments, grades, and course sequences.

Data Preprocessing: Clean and preprocess the data to ensure it is suitable for mining. This might involve removing incomplete records and normalizing grades.

Frequent Itemset Mining: Use algorithms like Apriori or FP-Growth to find frequent itemsets in the data.

Rule Generation: Generate association rules from the frequent itemsets, focusing on rules with high support and confidence.

Rule Evaluation: Evaluate the generated rules to ensure they make logical and educational

sense. Discard any rules that do not provide valuable insights.

Implementation: Implement the validated rules into the curriculum design process, updating prerequisite structures and advising systems accordingly.

Monitoring and Adjustment: Continuously monitor student performance and feedback, adjusting the rules and prerequisites as needed to ensure optimal outcomes.

By using association rules, educational institutions can create data-driven prerequisite structures that enhance student preparedness and success.

VI. Association Rule Learning

Association rule learning is a type of unsupervised learning technique that checks for the dependency of one data item on another data item and maps accordingly so that it can be more profitable. It tries to find some interesting relations or associations among the variables of dataset. It is based on different rules to discover the interesting relations between variables in the database. Association rule learning works on the concept of If and Else Statement, such as if A then B. Here the If element is called antecedent, and then statement is called as Consequent. These types of relationships where we can find out some association or relation between two items is known as single cardinality. It is all about creating rules, and if the number of items increases, then cardinality also increases accordingly. So, to measure the associations between thousands of data items, there are several metrics. These metrics are given below: *Support; Confidence; Lift*. Let's understand each of them [3].

Support is the frequency of A or how frequently an item appears in the dataset. It is defined as the fraction of the transaction T that contains the itemset X. If there are X datasets, then for transactions T, it can be written as:

$$Supp(T) = \frac{Freq(X)}{T} \tag{1}$$

Confidence indicates how often the rule has been found to be true. Or how often the items X and Y occur together in the dataset when the occurrence of X is already given. It is the ratio of the transaction that contains X and Y to the number of records that contain X.

$$Confidence(T) = \frac{Freq(X,Y)}{Freq(X)} \tag{2}$$

Lift, is the strength of any rule, which can be defined as below formula

$$Lift(T) = \frac{Supp(X,Y)}{Supp(X) \times Supp(Y)} \tag{3}$$

It is the ratio of the observed support measure and expected support if X and Y are independent of each other. It has three possible values:

- If Lift= 1: The probability of occurrence of antecedent and consequent is independent of each other.
- Lift>1: It determines the degree to which the two itemsets are dependent to each other.

Lift<1: It tells us that one item is a substitute for other items, which means one item has a negative effect on another.

VII. Problem Statement

When we study, we have a standard list of subjects that we learn. Each student has a different list of electives, depending on their needs and preferences. Students can choose from a variety of subjects. These electives can help enhance learning in several ways. If there is a pair of subjects X and Y that are often studied together [4]:

Both X and Y can be placed on the same syllabus so that students who study one subject are encouraged to study the other.

Although we know that certain subjects are often studied together, the question arises: how can we detect these associations?

Association rules can also be used in academic education. For example, they can determine the required sequence of subjects (prerequisites) included in the curriculum and help students successfully earn credits.

VIII. Apriori Algorithm

The algorithm was first proposed in 1994 by Rakesh Agrawal and Ramakrishnan Srikant. Apriori algorithm finds the most frequent itemsets or elements in a transaction database and identifies association rules [5] between the items just like the above-mentioned example.

The model: data

- $I = \{i_1, i_2, \ldots, i_m\}$: a set of items
- Transaction t: t a set of items, and $t \subseteq I$.
- Transaction dataset T: a set of transactions $T = \{t_1, t, \ldots, t_n\}$

To construct association rules between elements or items, the algorithm considers 3 important factors which are, support, confidence and lift.

The support of item X is defined as the ratio between the number of transactions containing the item X by the total number of transactions expressed in formula (1). Support indicates how popular an itemset is, as measured by the proportion of transactions in which an itemset appears. In Table 2 below, the support of {IT Fundamentals} is 4 out of 8, or 50%. Itemsets can also contain multiple items. For instance, the support of {IT Fundamentals, Computer Networks, Database Systems} is 2 out of 8, or 25%. If you find that test results outside of a certain ratio have a significant impact on your success, you may want to consider using that ratio as a support threshold. You can then define itemsets with support values above this threshold as significant itemsets.

Table 2

Transactions (successful exam of students)	Items (subjects)
Transaction 1	“IT Fundamentals”, “Computer Networks”, “Database Systems”, “IT Elective”
Transaction 2	“IT Fundamentals”, “Computer Networks” , “Database Systems”
Transaction 3	“IT Fundamentals”, “Computer Networks”
Transaction 4	“IT Fundamentals”, “Discrete mathematics”
Transaction 5	**“**Computer architecture”, “Computer Networks”, “Database Systems”, IT Elective”
Transaction 6	**“**Computer architecture”,, “Computer Networks”, “Database Systems”
Transaction 7	**“**Computer architecture”,, “Computer Networks”
Transaction 8	**“**Computer architecture”, “Discrete mathematics”

This is measured by the proportion of transactions with item X, in which item Y also appears. The confidence between two items X and Y, in a transaction is defined as the total number of transactions containing both items X and Y divided by the total number of transactions containing X (formula (2)). Confidence says how likely item Y is purchased when item X is purchased, expressed as {X→Y}. This is measured by the proportion of transactions with item X, in which item Y also appears. In Table 2, the confidence of {IT Fundamentals →Computer Networks} is 3 out of 4, or 75%.

One drawback of the confidence measure is that it might misrepresent the importance of an association. This is because it only accounts for how popular "IT Fundamentals"s are, but not "Computer Networks"s. If "Computer Networks"s are also very popular in general, there will be a higher chance that a transaction containing "IT Fundamentals"s will also contain "Computer Networks"s, thus inflating the confidence measure. To account for the base popularity of both constituent items, we use a third measure called lift.

Lift is the ratio between the confidence and support. Lift says how likely item Y is purchased when item X is purchased, while controlling for how popular item Y is.

$$Lift(X \rightarrow \mathrm{Y}) = \frac{Supp(X,Y)}{Supp(X) \times Supp(Y)}$$

In Table 2, the lift of {IT Fundamentals→Computer Networks} is 1 which implies no association between items. A lift value greater than 1 means that item Y is likely to be bought if item X is bought, while a value less than 1 means that item Y is unlikely to be bought if item X is bought. (*here, X represents "IT Fundamentals" and Y represents "Computer Networks"*).

IX. Summary

In summary, prerequisites are essential in curriculum design to ensure that students are adequately prepared, which in turn enhances their chances of success and maintains the integrity and quality of the educational program. By using association rules, educational institutions can create data-driven prerequisite structures that enhance student preparedness and success. By leveraging these AI method, educational institutions can create more effective and personalized learning experiences, ensuring that students are well-prepared for advanced coursework and ultimately improving their academic success.

References

[1] A.V. Kelly, The Curriculum: Theory and Practice" ISBN: 978-1446273631

[2] Bailey, J. AI in Education: The leap into a new era of machine intelligence carries risks and challenges, but also plenty of promise. Education Next, 23(4), 28-35.

[3] Information on https://www.javatpoint.com/association-rule-learning

[4] Information on https://www.kaggle.com/code/rockystats/apriori-algorithm-or-market-basket-analysis

[5] Information on https://www.geeksforgeeks.org/apriori-algorithm/

GAME THEORY-BASED OPTIMIZATION FOR SUSTAINABLE ENERGY TRADING SYSTEMS

Elbey Rustemzade[1], Nurali Yusifbayli[1]

•

[1]Azerbaijan Technical University, H. Cavid ave. 25, Baku, Azerbaijan AZ 1073
elbeyrustemzade@gmail.com, yusifbayli.n@gmail.com

Abstract

This article explores the transition from traditional energy networks to Microgrid networks, highlighting the implications of Energy 4.0 and the adoption of modern innovative approaches in energy systems. As energy demands evolve, decentralized energy trading systems are emerging as crucial mechanisms that empower consumers and enhance grid resilience. The paper presents a technical solution to the challenges posed by this transition, focusing on optimizing energy trade through Corporate Game Theory methods. By employing these strategic frameworks, the study aims to improve decision-making processes among market participants, ultimately leading to more efficient energy transactions. The results demonstrate significant advancements in optimizing energy trade, showcasing the potential for increased efficiency and sustainability in decentralized energy markets. This research contributes to a deeper understanding of how innovative strategies can facilitate the effective integration of Microgrid networks within the broader context of Energy 4.0.

Keywords: Energy4.0, Game theory, peer to peer energy trade,microgrid.

I. Introduction

Energy issues represent one of the most complex challenges of today. The increasing global demand for energy, the urgent need to combat climate change, and the integration of renewable energy sources necessitate innovative solutions to create more sustainable and efficient energy systems. In this context, game theory emerges as a powerful tool for optimizing strategic decision-making processes in energy trading and management [1].

Game theory is a mathematical framework that models situations where individuals or groups interact to achieve specific objectives. In the energy sector, these interactions can often be competitive or collaborative. By analyzing these dynamics, game theory helps balance the interests of various stakeholders - such as producers, consumers, and regulatory bodies - leading to better outcomes in energy distribution and consumption [2].

The rise of renewable energy sources has introduced new dynamics into energy trading. Innovative approaches like *peer-to-peer (P2P)* energy trading allow individuals to trade their energy directly with one another, challenging traditional market structures. In this process, game theory can provide crucial insights into pricing strategies and resource allocation. Furthermore, cooperative game theory offers a framework for different stakeholders to collaborate effectively toward common goals [3].

However, the complexity of energy issues extends beyond technical challenges; it encompasses social, economic, and environmental dimensions as well. For instance:

Increasing Energy Demand: Population growth and industrialization are driving a rapid rise in energy demand. This escalation strains existing energy infrastructures and heightens the need for new resources [4].

Climate Change: The reliance on fossil fuels significantly contributes to climate change, prompting countries to seek sustainable energy solutions. Yet, integrating renewable sources into existing systems is a multifaceted process fraught with conflicting stakeholder interests [4].

Regulatory Uncertainties: Changing regulations and policies in energy markets create uncertainty for investors, negatively impacting market dynamics and long-term planning [4].

Technological Advancements: The necessity for innovative technologies presents both opportunities and challenges. Game theory can guide strategic decision-making while influencing the adoption of new technologies.

Given these complexities, the role of game theory in addressing energy issues becomes increasingly evident. By developing strategies informed by game-theoretic principles, stakeholders can better navigate their interactions and conflicts, ultimately leading to more sustainable and efficient energy systems [4].

Several real-world applications illustrate how game theory is being used to address energy challenges:

Brooklyn Microgrid: This project enables local energy trading among participants using renewable energy sources. Game theory was used to model the interactions between prosumers (producers and consumers) to optimize trading strategies. By applying game theory, the project achieved efficient price-setting mechanisms, allowing participants to maximize their economic benefits while promoting local energy resilience [5].

California's Demand Response Program: This program utilizes game theory to incentivize residential consumers to reduce their energy usage during peak periods. The program models consumer behavior to determine optimal pricing strategies. The application of game theory led to increased participation rates and significant reductions in peak demand, contributing to grid stability [6].

UK Energy Market Auctions: The UK energy market employs auction mechanisms for capacity allocation, where game theory is used to analyze bids from various energy producers. The goal is to design auctions that maximize efficiency and fairness. By applying game theory principles, the auction design improved competition and ensured optimal resource allocation among diverse energy suppliers [7].

Sonnen Community Project in Germany: This initiative facilitates peer-to-peer energy trading among residents using a blockchain platform. Game theory is used to develop trading algorithms that optimize energy exchanges based on real-time consumption patterns. The project has demonstrated successful local energy trading, allowing consumers to benefit economically while enhancing the use of renewable energy sources [8].

Electric Vehicle (EV) Charging Optimization: Projects focusing on EV charging optimization utilize game theory to manage charging behaviors among users while minimizing grid impacts. Game theory models helped optimize charging schedules, balancing user convenience and grid reliability—crucial as EV adoption increases [9].

The aforementioned examples illustrate the practical applications of game theory in various energy-related projects, highlighting its importance in optimizing resource allocation, enhancing market efficiency, and facilitating the integration of renewable energy sources.

This paper will explore how game theory can be applied to solve various energy-related problems, emphasizing its potential in areas such as p2p trading, dynamic pricing models, and cooperative strategies among stakeholders. Through this examination, we aim to illustrate how strategic thinking rooted in game theory can help overcome the intricate challenges facing the energy sector today.

II. Challenges in Microgrid Protection Systems

The rise of micro grids has transformed energy trade, introducing both opportunities and significant challenges. One of the main issues is the integration of diverse renewable energy sources, such as solar and wind, which are inherently intermittent. This variability complicates forecasting and reliable trading agreements, making it essential for market participants to develop robust prediction models [10,11].

The integration of alternative energy resources especially wind and solar energy into electrical grids presents challenges due to their dynamic characters. Especially when the share of these renewable energy sources exceeds 5-10%, energy managers need to develop strategies. First of all, it is important to prevent local concentrations. Additionally, renewable energy plants must have the ability to *stabilize the grid*. Accurately forecasting electricity production plays a critical role in the *operational planning* of other plants. These forecasts help *optimize energy flows*. There is significant experimental data and theoretical models related to high shares of renewable energy. When the right measures are taken, a *reliable energy transition* can be achieved [11].

Regulatory frameworks also pose challenges, as differing rules across jurisdictions can create confusion and hinder local energy exchanges. A cohesive regulatory approach is necessary to facilitate seamless operations and ensure fair participation in energy markets. [12,13]

Grid interoperability is another critical issue. Effective communication between microgrids and the main grid is vital for enabling energy transactions, yet the lack of standardized communication protocols can lead to inefficiencies [13].

Financial challenges arise in establishing fair pricing mechanisms that reflect the value of renewable energy while ensuring consumer affordability. Innovative pricing models that consider market demand and generation variability are essential for successful trading [14].

Additionally, trust and transparency in energy trading are crucial. The implementation of blockchain technology and smart contracts can enhance transaction security and reliability, ensuring that all parties adhere to agreed-upon terms while streamlining the trading process [15].

In summary, while energy trading in microgrids offers exciting opportunities for localized energy management, addressing challenges related to renewable integration, regulatory consistency, interoperability, pricing, and transparency will be key to optimizing these systems.

III. Optimizing Energy Trade by Corporate Game Theory Method

Game theory is a mathematical framework that applies to economic theory, where the economy is viewed as a complex game involving producers, consumers, and market intermediaries. Introduced by John von Neumann in 1928, it initially focused on two-person zero-sum games, later expanded by von Neumann and Oskar Morgenstern in their 1944 work, "Theory of Games and Economic Behavior." The significance of game theory is underscored by 11 Nobel prizes awarded for research in various fields that utilize its principles. At its core, game theory analyzes decision-making processes among rational actors, examining how their strategies and choices are shaped by the interactions with others [16].

Modern electric power systems face numerous challenges due to deregulation and increased competition, complicating the decision-making process. Additionally, there has been a shift from a vertical to a horizontal control and operational structure, which has heightened the complexity of issues related to reliability, operation, control, and management. Traditional models struggle to address the interdependent decision-making processes within power systems, as they often treat participants as static entities [16].

Game theory emerges as a valuable tool for tackling these contemporary challenges. It serves as an analytical framework for understanding strategic interactions among rational decision-makers, where each player's actions are influenced by the actions of others. Game theory is applicable across various fields, including computer science, economics, biology, political science, and psychology .It provides insights into a wide range of human and computer interactions, establishing itself as a science of logical decision-making. Primarily, game theory applies to economic theory, conceptualizing the economic system as a complex game between producers and consumers facilitated by market intermediaries [17].

Literature reviews on electricity markets provide a thorough examination of the diverse applications of game theory across multiple dimensions, including energy management, trading, pricing strategies, bidding tactics, and demand-side management. Within this context, games can be classified into four distinct types, Fig. 1., [17,18,19,20]:

Non-corperative Game: In this category, players make decisions simultaneously, meaning that the payoff for each participant is contingent upon the choices made by all others involved. This interdependence necessitates careful consideration of rivals' potential actions. In some academic literature, it is referred to as a "Strategic Game."

Dynamic Game: This type allows players to make decisions at various points in time, enabling strategies to adapt and evolve based on earlier actions and outcomes. Such temporal dynamics provide rich opportunities for strategic maneuvering and long-term planning.

Cooperative Game: This classification emphasizes the formation of coalitions among groups of players who can negotiate binding agreements to enhance their collective outcomes. The ability to collaborate strategically can lead to more favorable results than isolated decision-making.

Evolutionary Game: This approach investigates how strategies develop and change over time, influenced by their success relative to others within a population. It often draws on concepts of natural selection, illustrating how certain strategies can become more prevalent as they prove more effective in achieving desired outcomes.

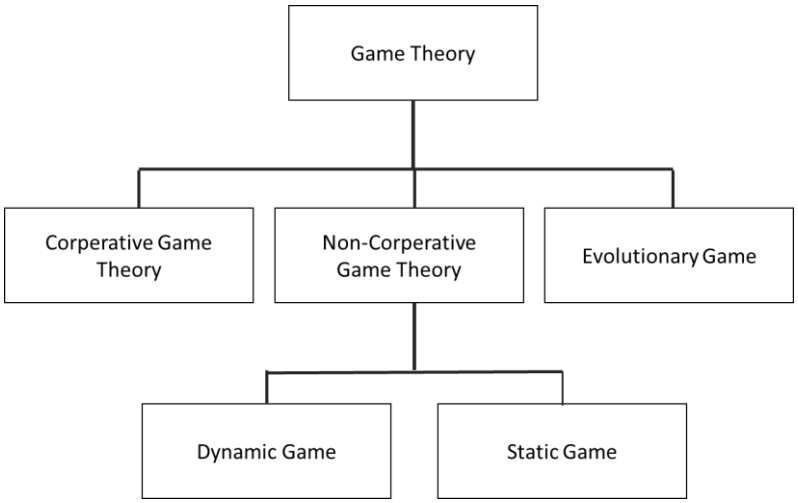


Figure 1: *Classification of Game Theory method. [20]*

Table 1 presents the examined decision-making challenges within electric power systems, utilizing various classifications of game theory methodologies [19,20].

Table 1: *Classifications of game theory methodologies*

Cooperative game theory	Analyzing how losses can be fairly distributed among participants
	Evaluating strategies employed by market agents to enhance system efficiency
	Developing frameworks that facilitate effective energy management
	Assessing the reliability of power systems under cooperative strategies
	Investigating the construction and optimization of distributed heating networks
	Addressing the complexities of demand management within microgrids
	Exploring how distributed resources can engage in market activities
	Analyzing cost allocation methods for transmission expansion
	Focusing on scheduling methodologies for renewable energy sources
Static game theory	Strategies to optimize consumer demand response
	Examining transaction dynamics in electricity markets
	Implementing energy management strategies in smart homes
	Addressing energy management challenges in hybrid systems
	Further insights into microgrid scheduling strategies
	Analyzing bidding strategies for VPPs in competitive markets.
	Exploring management techniques for distribution systems.
	Continued focus on renewable scheduling practices
	Investigating attacker-defender dynamics within power systems.
	Reiterating the importance of distributed resources in market participation.
Dynamic game theory	Strategies for integrating distributed resources into distribution systems
	Time-sensitive strategies for managing consumer demand
	Dynamic aspects of electricity market transactions
	Continued exploration of scheduling within microgrids over time
	Enhancing reliability through dynamic strategies
	Further analysis of attacker-defender interactions over time
	Power transactions of generating entities in an electricity market
Evolutionary game theory	Strategies for managing electric vehicle loads within decentralized frameworks
	Understanding adaptive transaction behaviors over time
	Adaptive approaches to demand-side strategies
	Evolutionary perspectives on microgrid scheduling practices
	Exploring innovative supply chain policies for renewable energy sources

This structure emphasizes the significance of each type of game theory while providing a clearer overview of their applications in energy systems. [17,18,19,20]:

IV. Working Principle of the Game Framework

Energy trading in Integrated Energy Systems (IES) involves two main levels of decision-making. At the upper level, Energy Resources (ERs) set pricing strategies for the demand side, while the supply side makes quantitative decisions about how much energy to provide. These decisions are important because they affect consumer willingness to pay and supplier readiness to deliver energy. All players—ERs, energy suppliers (ESs), and users—are assumed to be rational and focused on optimizing their outcomes in a competitive market.

At the lower level, suppliers decide how much energy to bid based on ER prices, and users adjust their consumption in response to these prices. This interaction is dynamic, with each side

influencing the other. The situation can be modeled using *Stackelberg evolutionary game theory:* where ERs act as leaders setting prices, and ESs and users follow by adjusting their strategies. This framework helps analyze how strategic interactions among participants contribute to market efficiency and stability in energy trading within IES.

Fig. 2 the game framework mentioned in the sources is based on a two-stage game model used to model interactions among Energy Retailers (ER), Energy Suppliers (ES), and users.

Stage 1: Vertical Stackelberg Game

Leader (ER): The ER develops energy purchasing strategies for the supply side and pricing strategies for the demand side based on market information. The goal is to maximize the ER's revenues.

Followers (ESs and Users): ESs determine their bids according to the ER's energy purchasing strategy, while users adjust their Integrated Demand Response (IDR) based on the energy selling prices set by the ER. ESs aim to maximize their revenues, while users seek to minimize their energy costs.

Stage 2: Horizontal Non-Cooperative Game

Participants (ESs): ESs compete with each other as independent and rational individuals to maximize their revenues.

Strategy (ESs): Each ES sets its energy selling price offer to the ER.

Revenues (ESs): Each ES aims to maximize the difference between the revenue obtained from selling energy to the ER and its operational costs. [21]

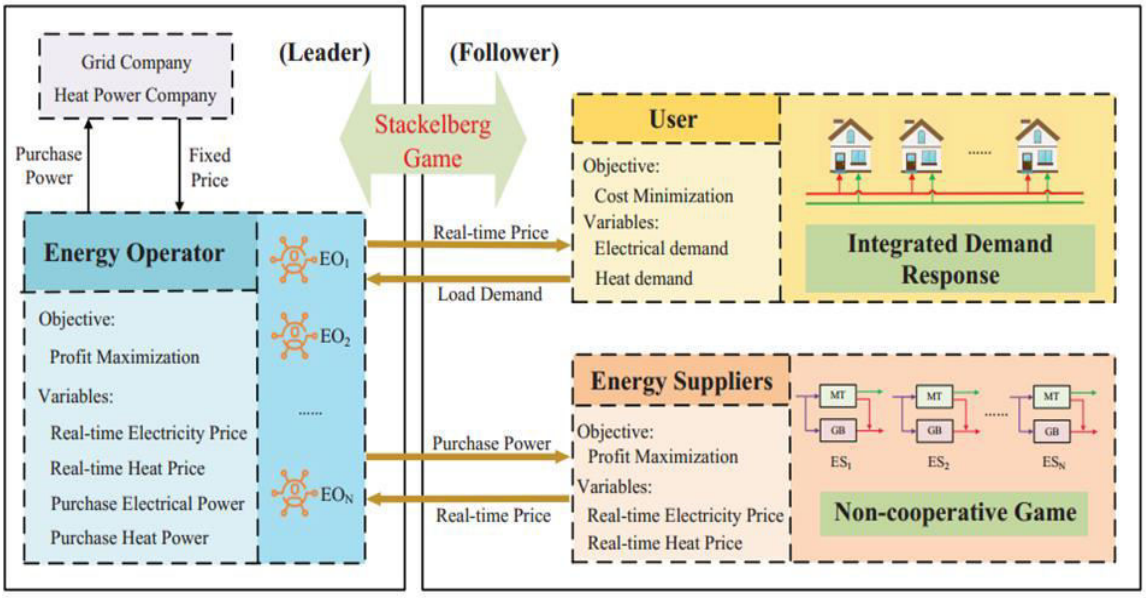


Figure 2: *IES game framework. [21]*

This model includes a vertically structured Stackelberg game where ER acts as the leader, while ES and users are followers, and a horizontally structured non-cooperative game among ESs [20]. The primary objective of the game is to ensure that the system operates in a balanced and stable manner, assists both supply and demand sides in achieving their goals, and optimizes the revenues of all parties involved [21].

V. Implementation of the Game Framework

The implementation of the game framework is carried out using a distributed approach that combines genetic algorithms and second-order programming. Since the decisions of the ER represent a multivariable and nonlinear optimization problem, genetic algorithms are employed. In contrast, since the decisions of the ESs are second-order optimization problems, a second-order programming approach is utilized [21].

Enables active participation of ESs in the energy market, increasing their revenues, allows users to optimize their energy consumption through IDR and reduce their costs, Ensures that the ER maintains a balanced and stable operation of the system, Achieves distributed autonomy and collaborative optimization of Integrated Energy Systems (IES) is the main benefits of the game Framework [21,22].

VI. Conculusion

This study proposes using game theory to improve how we integrate renewable energy sources and make shared energy storage systems more practical. As energy systems grow more complex and interconnected, finding ways to balance different stakeholders' interests and strategies becomes essential. Game theory provides a way to design collaborative solutions that maximize benefits for everyone involved, enhancing overall system efficiency.

By applying cooperative game theory concepts, we developed fair benefit-sharing methods. These methods aim to make shared storage systems more appealing by ensuring that all participants gain from the arrangement.

Overall, this study underscores the potential of game theory to optimize energy systems, facilitate renewable energy integration, and encourage shared storage adoption. Moving forward, research might focus on more advanced game-theory models, including dynamic pricing and smart grid technologies. Exploring the impact of these solutions in various cultural and socioeconomic contexts could help create a more inclusive and sustainable future for energy management.

References

[1] Recent Advances in Local Energy Trading in the Smart Grid Based on Game–Theoretic Approaches .Mar 16 2024.Matthias Pilz, and Luluwah Al-Fagih M.Pilz and L.Al-Fagih are with Kingston University London.Faculty of Science, Engineering and Computing.

[2] Influence decision models: From cooperative game theory to social network analysis Volume 39, February 2021, 100343 , Xavier Molinero a b, Fabián Riquelme c

[3] Recent Advances in Local Energy Trading in the Smart Grid Based on Game-Theoretic Approaches, 18 October 2017,Matthias Pilz; Luluwah Al-Fagih

[4] A review of microgrid protection for addressing challenges and solutions
Kunal Kumar ,Prince Kumar , Susmita Kar

[5] Designing microgrid energy markets A case study: The Brooklyn Microgrid Esther Mengelkamp a,⇑ , Johannes Gärttner a , Kerstin Rock b , Scott Kessler b , Lawrence Orsini b , Christof Weinhardt

[6] C. E. T. R. Su et al. (2023). "Game Theory-Based Demand Response in California: A Case Study." Energy Economics.

[7] G. G. M. I. L. W. de Vries et al. (2024). "Game Theory in UK Energy Auctions: Enhancing Market Efficiency." Journal of Energy Markets.

[8] Summary case study report: The sonnenCommunity ,February 2022 DOI: 10.5281/zenodo.6500618.Jake P BarnesJake P BarnesPaula HansenPaula HansenS. J. DarbyShow all

5 authorsMark A. AndorMark A. Andor

[9] M. K. Alabed et al. (2023). "Game-Theoretic Optimization of Electric Vehicle Charging in Smart Grids." IEEE Transactions on Power Systems.

[10] Smith, J., Brown, A., & Wilson, T. (2024). "Challenges in Local Energy Trading: Opportunities and Solutions." Journal of Energy Markets, 15(1), 1-20.

[11] SOME PROBLEMS OF ENERGY SECURITY IN THE CONTEXT OF WIDESPREAD USE OF RES N. A. Yusifbayli1 *, V. X. Nasibov2 2023.1Azerbaijan Technical University, Baku, Azerbaijan 2Azerbaijan Research and Design–Prospecting Institute of Energetics, Baku, Azerbaijan

[12] Johnson, R., & Lee, M. (2024). "Regulatory Barriers to Microgrid Energy Trading: A Comparative Analysis." International Journal of Energy Research, 48(2), 110-125.

[13] Chen, L., Kim, S., & Davis, E. (2024). "Interoperability Challenges in Microgrid Energy Trading Systems." Renewable Energy Systems Review, 29(3), 205-220.

[14] Patel, N., & Zhao, Q. (2024). "Innovative Pricing Models for Renewable Energy in Microgrid Markets." Energy Reports, 10(2), 75-90.

[15] Thompson, K., White, R., & Green, D. (2024). "Enhancing Trust in Microgrid Energy Transactions: The Role of Blockchain." Journal of Modern Power Systems and Clean Energy, 12(5), 400-415.

[16] Energy 4.0: AI-enabled digital transformation for sustainable power networks. Muhammad Khalid Volume 193, July 2024, 110253

[17] D. A. Hirth et al. (2022). "Transactive Energy Systems in Microgrids: A Game-Theoretic Approach." IEEE Transactions on Smart Grid.

[18] Game Theory Approaches for the Solution of Power System Problems:A Comprehensive Review .Saee d Abapou r1 • Morteza Nazari-Heris1 • Behnam Mohammadi-Ivatloo1 • Mehrdad Tarafdar Hagh1.19 November 2018

[19] Nurcan Yarar 1, Yeliz Yoldas 2, Serkan Bahceci 3 , Ahmet Onen 4 and Jaesung Jung 5,A Comprehensive Review Based on the Game Theory with Energy Management and Trading

[20] Game Theory Approaches for the Solution of Power System Problems: A Comprehensive Review. Saee d Abapou r1 • Morteza Nazari-Heris1 • Behnam Mohammadi-Ivatloo1 • Mehrdad Tarafdar Hagh1 2018

[21] Distributed collaborative operation strategies in multi-agent integrated energy system considering integrated demand response based on game theory Ke Li∗, Ning Ye, Shuzhen Li, Haiyang Wang, Chenghui Zhang .School of Control Science and Engineering, Shandong University, Jinan, 250061, Shandong Province, China

[22] Innovative Energy Systems: The Role of Energy 4.0 in Shaping Azerbaijan’s Sector .Yusifbayli N.A, Rustemzade E.Sh DOI - 10.61413/BYCI6990

SECURING THE FUTURE OF ENERGY TRADING: ENHANCING CYBERSECURITY WITH BLOCKCHAIN IN WEB 3.0

Elbey Rustemzade[1], Nurali Yusifbayli[1]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan
aelbeyrustemzade@gmail.com, byusifbayli.n@gmail.com

Abstract

Cybersecurity is becoming increasingly important in energy trading, especially as the sector becomes more interconnected and reliant on digital technologies. With the rise of distributed energy resources and peer-to-peer (P2P) trading, new cyber threats are emerging, making it essential to protect data and transactions to maintain trust and stability in energy systems. Web 3.0 represents a significant shift towards decentralization and user empowerment, primarily driven by blockchain technology. This innovative approach allows for secure transactions without intermediaries, enabling transparent exchanges between consumers and producers in the energy market.

The integration of Ethereum platforms is crucial for enhancing cybersecurity in energy trading. Ethereum's advanced smart contract capabilities facilitate automated and secure transactions, greatly reducing the risks of human error and fraud. Provides a robust environment for testing these smart contracts, allowing developers to identify vulnerabilities before deployment. For Azerbaijan, which is modernizing its energy infrastructure and diversifying its energy sources, adopting these technologies can significantly improve security and efficiency in its energy markets. By focusing on cybersecurity measures within Ethereum frameworks, stakeholders can foster a safer, more efficient energy market.

Keywords: Web3.0, energy trade, microgrid, blockchain, cybersecurity

I. Introduction

Energy trading has traditionally relied on centralized systems, which can expose it to significant challenges such as security breaches, inefficiencies, and a lack of transparency. These issues not only compromise the integrity of transactions but also undermine the trust that participants have in the market. As a result, many stakeholders are actively seeking more secure and efficient alternatives that can foster confidence and collaboration within this crucial sector [1].

Recent research highlights that adopting decentralized models, such as *peer-to-peer energy trading* powered by *blockchain technology*, offers promising solutions to these challenges. For instance, blockchain technology enhances security through its decentralized and immutable ledger, which can significantly reduce the risk of fraud and unauthorized access [1]. Additionally, smart contracts facilitate dynamic trading by automating transactions and ensuring compliance with predefined conditions, thus improving efficiency and reducing transaction costs [2].

Furthermore, studies show that blockchain can improve transparency in energy trading by providing a clear and tamper-proof record of all transactions, enabling participants to verify the

authenticity of trades [3]. This increased transparency not only builds trust among participants but also encourages greater participation in local energy markets.

Web 3.0 is seen as a major leap in internet technology, addressing challenges through its core principles. By using decentralized systems like blockchain, it aims to give users more control, make peer-to-peer interactions easier, and encourage innovation in areas like energy trading. This shift highlights its potential to reshape how we approach technology and collaboration.

Web 3.0 is the next evolution of the World Wide Web, with 5.3 billion users as of October 2023, representing 65.7% of the global population. Unlike Web 2.0, which is dominated by large tech companies, Web 3.0 aims to decentralize the internet, giving users full control over their data and digital assets through technologies like blockchain and artificial intelligence. Ethereum exemplifies innovation in Web3, showcasing smart contract technology and diverse token standards like ERC-20 and ERC-721. Its thriving DeFi sector, ongoing upgrades for scalability, and vibrant community solidify its role as a foundational element of Web3 protocols. With the transition to Ethereum 2.0, it is set to enhance interactions with digital applications and finance. Ethereum's evolution continues, shaping a future where decentralized systems transform our digital experiences. [4].

This new framework also allows content creators to earn directly from their work, eliminating the need for intermediaries. Right now, most of our online data and activities are controlled by large platforms in a centralized setup, which can compromise user privacy and limit control over personal data. Web 3.0, though, envisions a shift to a more decentralized internet where individuals have much more control. Using blockchain technology, smart contracts, and decentralized applications (DApps), Web 3.0 allows for secure, trust-free transactions directly between users. This setup is already making waves in areas like decentralized finance (DeFi), where people can handle financial transactions without banks or other intermediaries [5, 6].

Figure 1 illustrates the architecture proposed by Jianjun Zhu, Fan Li, and Jinyuan Chen, detailing the sequential flow of data processing within the system. This architecture is structured into multiple layers, including the infrastructure layer, which comprises blockchain technology and decentralized storage solutions, enabling secure peer-to-peer transactions. The network layer facilitates communication through peer-to-peer protocols and decentralized identity systems, while the protocol layer establishes standards for data transmission and governance, promoting interoperability among diverse blockchain networks. Furthermore, the cognitive layer integrates artificial intelligence and machine learning to enrich user experiences through data analysis and personalization. Finally, the use case layer showcases practical applications such as decentralized finance (DeFi), non-fungible tokens (NFTs), and decentralized applications (dApps), which exemplify the transformative potential of Web 3.0 technologies in creating a more secure, transparent, and user-centric internet landscape [5].

Traditional energy technologies often struggle to adapt to the complexities of modern microgrids, which require rapid peer-to-peer interactions and transparent data exchanges. In this context, blockchain technology emerges as a powerful solution, effectively addressing these demands. The shift towards decentralized energy systems underscores the necessity for innovative approaches that can meet the evolving needs of energy grids. With the growing adoption of Web 3.0 and blockchain, decentralized microgrids serve as a prime example of how these technologies can facilitate direct and real-time data management, enabling seamless peer-to-peer interactions among users. The challenges faced by outdated centralized technologies highlight their inability to provide the flexibility and speed required in today’s decentralized networks, making a compelling case for the integration of blockchain into energy management systems[7].

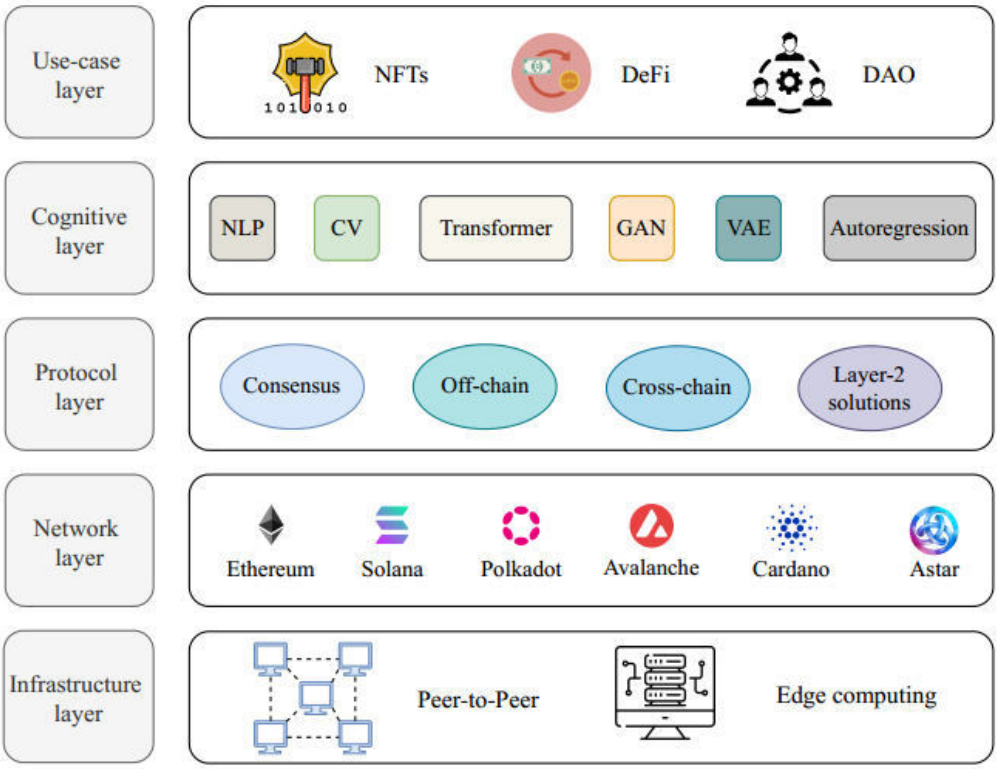


Figure 1: *Web 3.0 stack architecture[5]*

Ultimately, Web 3.0 holds the promise of a more open and secure internet, giving users back control over their online activities. However, this shift brings new challenges too, including scalability and regulatory hurdles, which researchers and developers are actively working to address. This article is written to contribute to the literature on these topics.

II. Energy trading challenges in Web 3.0 apllications

Blockchain technology shows promise for energy trading because it allows peer-to-peer transactions without a central authority, but it also brings major security challenges. Without centralized oversight, blockchain relies on robust security measures to protect against attacks like selfish mining and majority attacks, which could disrupt or manipulate the system [9,10]

Security issues also connect directly to transparency and fairness. For participants to trust the system and receive fair compensation, data must be tamper-proof. Weak security could let attackers skew transaction data or unfairly benefit some users [8]. As more users join, scalability also becomes crucial: slow transaction speeds and inconsistent validations can leave the system vulnerable to interruptions, further stressing security needs [7].

In short, blockchain's success in energy trading depends on overcoming these intertwined security, transparency, and scalability issues to create a reliable, trust-based network. Table 1 presents a categorization of challenges identified in the literature.

Table 1: *Security Challenges for Blockchain's success in energy trade systems*

Challenge	Description	Citations
Transparency Issues	Achieving full transparency in energy trading systems is challenging, which raises concerns about trust and accountability in transactions.	[7]
Fair Evaluation and Participation	Ensuring fair evaluation of prosumers' contributions is crucial for fostering participation in peer-to-peer (P2P) energy trading markets. Reliable mechanisms for assessing energy commitments are often lacking.	[8]
Security and Reliability	As systems become decentralized, securing transactions against malicious attacks while maintaining reliability in energy supply and demand balancing is a significant challenge.	[7], [8]
Scalability	The scalability of blockchain solutions to handle large volumes of transactions efficiently remains a concern, especially as more users join decentralized energy markets.	[7, 9,10]
Integration with Existing Infrastructure	Transitioning to decentralized models requires integrating new technologies with existing infrastructure, which can be complex and costly for stakeholders.	[9]

III. Enhancing Cyber Security in Energy Trading systems through Ethereum Smart Contracts

Ethereum, developed by Russian-Canadian programmer Vitalik Buterin in 2013 and launched in 2015, marks a transformative advancement in blockchain technology by introducing smart contracts, which automate and enforce agreements without intermediaries. Building on concepts proposed by Nick Szabo in 1994, Ethereum is the first platform to widely implement smart contracts, allowing users to facilitate a diverse array of transactions—ranging from digital currencies to real estate—within a secure, decentralized framework. This innovation has also led to the emergence of decentralized applications (DApps) and organizations (DAOs), which operate autonomously through smart contracts on the Ethereum blockchain. DAOs enable collective decision-making among token holders, thereby democratizing organizational structures and enhancing transparency and efficiency in various processes. As Ethereum continues to evolve, its capabilities are reshaping industries by providing automated solutions that streamline operations and reduce reliance on traditional contractual frameworks [11].

Ethereum represents a highly suitable platform for energy trade system applications, owing to its decentralized architecture, robust security features, and environmentally sustainable infrastructure.

Figure 2 illustration is the architecture of an energy trading solution utilizing the Ethereum blockchain and smart contracts. This solution includes traditional energy market participants such as consumers, producers, and retailers, as well as an Energy Authority that activates the smart contract.

The fundamental operating principles of the solution are as follows: [12]

System Registration: Producers and retailers register their Ethereum addresses, verified by the Energy Authority. This helps ensure that the identities of all parties participating in the system are authenticated.

Energy Production: Producers generate energy and transfer it to retailers via smart meters connected to the blockchain. When energy transfer occurs, the smart contract sends tokens to the producer at a predetermined rate and creates a "Transfer" event to record the transaction.

Energy Demand: Consumers request energy by creating a demand indicating the amount needed. This request initiates a blind auction within the smart contract.

Auction: Producers participate in the auction by submitting price bids for a specific energy supply. Bids are kept confidential from other participants and are protected through encryption.

Determining the Winner: Once the auction concludes, the smart contract identifies the best bid and selects the winning producer.

Agreement: The winning producer and consumer reveal their identities and exchange energy tokens and payment (ether) through the smart contract.

Conversion of Tokens to Energy: Consumers and producers can convert their tokens into energy by sending them to retailers. Retailers then deliver the corresponding amount of energy to the consumer via smart grid infrastructure [12].

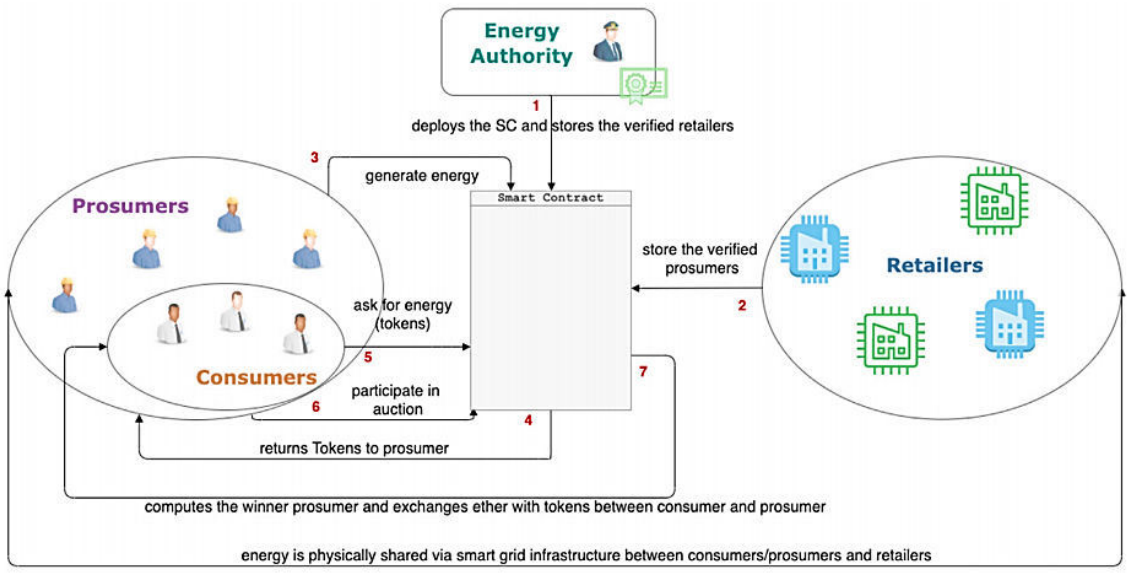


Figure 2: *Smart Contract Architecture [12]*

Figure 2 demonstrates how these steps interact to create a secure, transparent, and efficient energy trading system. The use of smart contracts helps automate the process while minimizing the need for human intervention, thereby fostering trust.

In this section, we outline key security properties that underpin a secure and reliable energy trading solution: [12]

1. Data Confidentiality: Protecting sensitive information is essential to prevent unauthorized access. Here, this means that real bid values in an auction are concealed from other participants until the auction ends, ensuring fair competition and protecting user data from potential exploitation [12].

2. Data Integrity: For data to remain complete, accurate, and unaltered, it must be safeguarded against tampering. In the context of energy trading, this means that once a bid price is entered, it remains immutable, preventing any unauthorized changes that could skew auction outcomes or fairness [12].

3. Privacy: User privacy is preserved by ensuring no unnecessary identifying information is disclosed. Both prosumers (producers/consumers) and consumers' identities should remain private during auctions, creating an equitable environment without bias or undue influence [12].

4. Authentication: It's crucial to verify the identities of participants in the system. Post-auction, participants are assured of each other's identities, reinforcing trust and accountability among all parties involved [12].

5. Accountability: To maintain transparency and reliability, every transaction within the system is recorded and accessible for verification. This ensures that all operations are carried out openly and with accountability among actors [12].

6. Reliability: A reliable system guarantees consistent performance of its functions over time. For energy trading, this means participants can count on system operations, such as energy requests, auctions, and agreements, being conducted reliably to ensure continuity and dependable service for all parties involved [12].

In other hand,to ensure security in Smart Grid (SG) systems, various protocols are employed. These protocols include Advanced Metering Infrastructure Security (AMI-SER), IEC 62351, NERC-CIP, and ISO/IEC standards. These standards aim to enhance security by addressing issues such as data integrity, confidentiality, and authentication in energy management and distribution. For instance, the Open Smart Grid Protocol (OSGP) secures data transmission through encryption techniques, while the integration of blockchain technology addresses existing security vulnerabilities, providing a more robust framework. Additionally, new standards and protocols are continuously being developed to improve the effectiveness of these systems [13].

IV. Case Study

To enhance the practical applicability of the topic, we want to present a discussion of a sample project also in this paper. In Australia, high electricity prices and favorable conditions for solar energy present an opportunity for change, making it a prime location for Power Ledger's innovative solutions. Power Ledger aims to address the needs of three key groups: energy consumers seeking cheaper and greener options, producers wanting better profits for excess power, and providers needing strategies to enhance their electric grids.Power Ledger offers significant benefits that could transform the energy market. Firstly, it enables peer-to-peer energy trading, allowing consumers to buy and sell excess renewable energy directly with their neighbors, which promotes the use of green energy and reduces reliance on centralized utilities. This decentralized model not only makes electricity more affordable but also empowers consumers to take control of their energy consumption. Additionally, This platform utilizes a dual blockchain system to facilitate its energy trading platform, employing two main protocols: the public Ethereum blockchain and a private consortium blockchain known as Ecochain. The Power Ledger Token (POWR) serves as the utility token that allows application hosts and participants to access the platform. Users must hold a sufficient amount of POWR to engage in transactions, which are secured through Ethereum Smart Bonds that facilitate the exchange of another token called Sparkz. Sparkz represents electricity credits and is pegged to local fiat currencies, allowing for stable pricing in various markets.

Sparkz is a trading token designed for seamless buying and selling of electricity measured in kWh. It is created for specific transactions, pegged to local currencies upon creation, and destroyed when redeemed for fiat or POWR tokens. POWR, on the other hand, is an ERC-20 utility token used by energy producers to operate within the Power Ledger ecosystem, facilitating smart contracts that manage trades and power distribution. Sparkz transactions occur using fiat currencies through various trading platforms that support closed-loop exchanges for energy and Sparkz. This dual-token system enhances market flexibility and ensures stable electricity pricing based on local currency values while separating it from the fluctuating value of POWR tokens. Overall, this framework which is describe in Figure 3 allows for efficient energy trading and supports the growth of decentralized energy markets. [14].

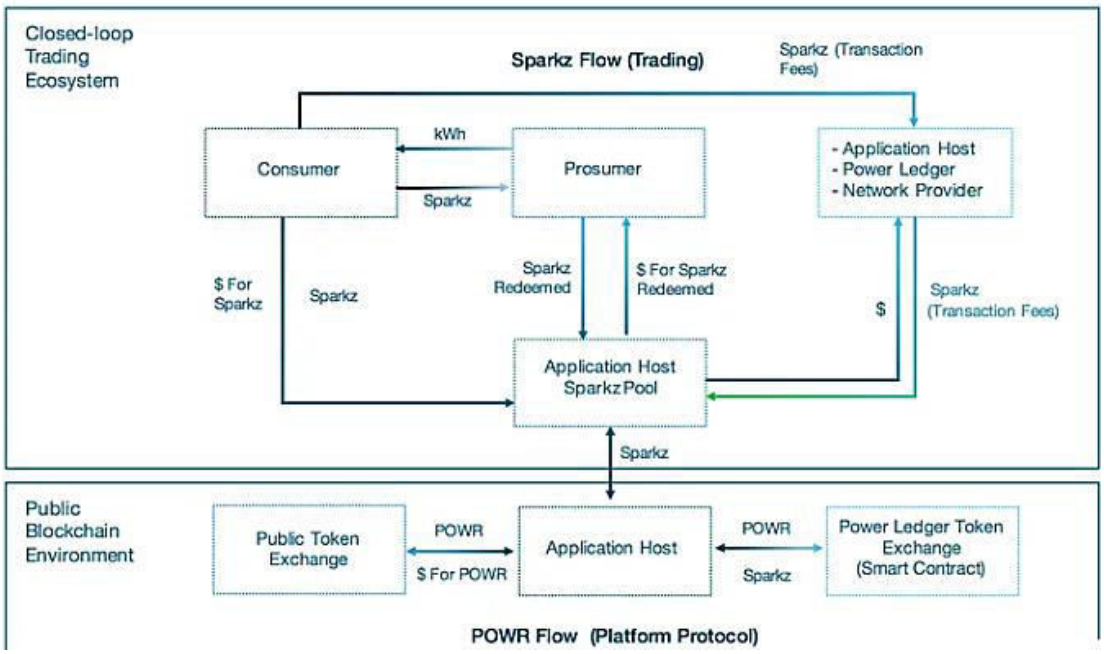


Figure 3: *Power Ledger Platform [14]*

In addition to these tokens, Power Ledger is at the forefront of revolutionizing energy markets through its innovative Peer-to-Peer (P2P) trading model, which allows individuals and businesses to generate, share, and trade energy directly without relying on centralized utilities. By leveraging blockchain technology, Power Ledger ensures secure, transparent transactions that enable prosumers to sell excess energy efficiently, promoting greater energy autonomy and increased use of renewable resources. The platform has successfully implemented various projects globally, demonstrating its effectiveness in enhancing community energy self-sufficiency and lowering costs. Additionally, Power Ledger's research focuses on the technological and economic implications of decentralized energy markets, exploring how Gen3 blockchains like Solana can optimize transaction management in P2P trading environments.In a result Integration with the Solana blockchain, enhancing its capabilities and scalability in the sustainability sector. This integration signifies a strategic shift from relying solely on its own blockchain to leveraging Solana's robust ecosystem, thereby accelerating innovation in decentralized energy markets. The combination of these protocols enables Power Ledger to create transparent, efficient, and secure energy trading solutions that promote renewable energy use and empower consumers in their energy choices. [14,15].

The platform's innovative approach has already demonstrated tangible savings for users, with reports indicating that energy consumers have saved up to $900 annually on their electricity bills.Power Ledger is partnering with global organizations to advance blockchain-based energy trading and renewable energy solutions. In Thailand, BCPG is launching a peer-to-peer energy trading project with renewable assets in Bangkok, offering electricity at competitive rates. In India, Tech Mahindra is implementing Power Ledger's platform to manage microgrids in rural areas, with potential applications in global markets. Australia's government has granted $8 million for a smart city energy project in Fremantle, integrating distributed energy systems and blockchain technology. Additionally, Power Ledger collaborates with Origin Energy for energy trading trials and with the Liechtenstein Institute for Strategic Development to promote renewable energy microgrids in Europe. [14,15,16].

Another successful project called "Choose your mix" enabled customers of French green energy retailer ekWateur to select and receive their preferred sources of energy – such as wind, solar or hydro – via Power Ledger Platforms. [14,16].

Overall, Power Ledger's integration of blockchain technology enhances transparency, security, and efficiency in energy management, positioning it as a leader in the transition towards sustainable energy solutions.

V. Conclusion

Web3.0 and Blockchain technology represents a transformative solution for enhancing the security of smart grid systems, particularly within the framework of Azerbaijan's energy sector modernization. As smart grids evolve into increasingly complex and decentralized networks, they become more susceptible to cyberattacks that can compromise the integrity of energy data and disrupt power distribution. The distinctive features of block chain—its distributed ledger system, transparency, and immutability—make it exceptionally well-suited to address these security concerns. By facilitating secure management of energy transactions and data, block chain can effectively mitigate risks associated with unauthorized access and data manipulation.

In Azerbaijan, where the energy sector is actively integrating renewable energy sources, the adoption of block chain technology can enable peer-to-peer energy trading and enhance consumer engagement. This decentralized approach not only empowers consumers by granting them greater control over their energy choices but also fosters improved efficiency in energy distribution. Moreover, block chain's capacity to provide a tamper-proof record of transactions builds trust among market participants. As Azerbaijan seeks to modernize its energy infrastructure, blockchain can play a pivotal role in strengthening system security through several key mechanisms:

Strengthening Cyber Resilience: The implementation of block chain fortifies Azerbaijan's energy infrastructure against cyber threats by providing a robust defense mechanism that is inherently resilient to attacks.

Facilitating Secure Renewable Energy Integration: As the country incorporates more renewable energy sources into its grid, block chain ensures secure transactions and data exchanges between decentralized energy producers and consumers, mitigating risks linked to increased complexity.

Regulatory Compliance and Audit Trails: The transparent nature of block chain simplifies auditing processes and compliance with energy regulations, enhancing trust among stakeholders while streamlining reporting related to cybersecurity incidents.

Attracting Investment through Security Assurance: A secure energy sector supported by block chain technology can attract foreign investments by demonstrating a commitment to safeguarding infrastructure from cyber threats, thereby fostering economic growth.

Ultimately, the integration of block chain technology in Azerbaijan's smart grid systems holds significant potential for improving security, efficiency, and sustainability in the energy sector. By enhancing data integrity, ensuring confidentiality, and providing real-time monitoring capabilities, block chain can effectively counteract the risks posed by cyberattacks. As Azerbaijan advances toward a modernized energy landscape, embracing block chain will not only bolster system security but also position the country as a forward-thinking leader in energy technology. This strategic integration will contribute to a more resilient and secure energy future for Azerbaijan, ensuring reliable service delivery while protecting critical infrastructure from evolving cyber threats.

We will start real project implementation in simulation platform such as Solidity and Ganash and research in the field of establishing and implementing such platforms in Azerbaijan. We will provide detailed information about these projects in our next articles in the near future. Additionally, as Azerbaijan continues to modernize its energy sector and integrate renewable sources, we

anticipate that our initiatives will align with ongoing projects, such as the recent agreements for large-scale solar developments, which aim to increase the share of renewables in the national energy mix. This alignment not only underscores the relevance of our work but also highlights the potential for collaboration with key stakeholders in Azerbaijan's energy landscape.

References:

[1] Khan et al. (2023). Blockchain Technology on Smart Grid, Energy Trading, and Big Data: Security Issues, Challenges, and Recommendations. Link

[2] Smith & Johnson (2024). Ethereum Based Smart Contract For Peer-To-Peer Energy Trading Using Blockchain Technology. Link

[3] Zhang et al. (2023). Decentralized Energy Trading Using Blockchain Technology: A Novel Approach to Democratizing the Energy Market. Link

[4] Ethereum Foundation on Web3's current landscape. Dustin Perez October 14, 2024 https://cask.fi/ethereum-the-foundation-of-web3-innovation/

[5] Brickendon Consulting on the fundamental principles of Web 3.0.July31,2023

[6] Mingzhe Liu et al., published on MDPI, on decentralized applications in Web https://doi.org/10.3390/blockchains1020008

[7] Al-Turjman, F., & Abujubbeh, M. (2019). Blockchain in Smart Grids: A Review on Different Use Cases. IEEE Access, 7, 40498-40509. doi:10.1109/ACCESS.2019.2904522

[8] Chibli, S., Alamri, A., & AlKassem, A. (2022). Fair Energy Trading in Blockchain-Inspired Smart Grid: Technological Barriers and Future Trends in the Age of Electric Vehicles. MDPI, 14(7), 150-165.

[9] Mingzhe, L., Zhao, Z., & Guo, S. (2021). Blockchain for Electric Vehicles Energy Trading: A Survey on Concepts and Solutions. IEEE Transactions on Industrial Informatics, 17(4), 2535-2545. doi:10.1109/TII.2020.3034745

[10] Vukolic, M. (2016). The Quest for Scalable Blockchain Fabric: Proof-of-Work vs. BFT Replication. IFIP International Conference on Open Problems in Network Security, 112-125. doi:10.1007/978-3-319-39028-4_9

[11] BLOCKCHAIN TECHNOLOGIES: COMPONENTS, APPLICATIONS AND PROBLEMS Imamverdiyev Y.N. ANAS Institute of Information Technologies, Baku, Azerbaijan DOI: 10.25045/jpis.v10.i2.02 Information society problems, 2019, No. 2, 18–32

[12] An Ethereum-based solution for energy trading in smart grids Francesco Buccafurri *, Gianluca Lax, Lorenzo Musarella, Antonia Russo 2023

[13] Blockchain Technology on Smart Grid,Energy Trading,and BigData:Security Issues , Challenges , and Recommendations 2022 Mohammad Kamrul Hasan ,1 Ali Alkhalifah,2 Shayla Islam ,3 Nissrein B. M. Babiker,4A.KM. Ahasan Habib,1 Azana Hafizah Mohd Aman,1 and Md. Arif Hossain1

[14] Power Ledger’s Disruptive Decentralized Energy Market, Nir Kabessa ·Dec 24, 2017

[15] P2P Trading-enabled Local Energy Market Supplemented with Blockchain Technology: An Australian Case Study, Liaqat Ali; M. Imran Azim; Jan Peters; Nabin B. Ojha; Vivek Bhandari; Anand Menon DOI: 10.1109/GreenTech56823.2023.10173850 19-21 April 2023

[16] The Future of Decentralised Energy: Powerledger’s Research in Blockchain and P2P Trading,October 22, 2024 , https://powerledger.io/media/the-future-of-decentralised-energy-powerledgers-research-in-blockchain-and-p2p-trading/

RISKS OF CASCADING FAILURES IN CRITICAL INFORMATION INFRASTRUCTURE

Arzu Babayeva[1], Yadigar Imamverdiyev[1]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan
arzu.babayeva@aztu.edu.az , yadigar.imamverdiyev@aztu.edu.az

Abstract

The article presents an analysis of risk research related to cascading failures in critical information infrastructure. An example of a systems approach implemented in the areas of critical information infrastructure (CII) activity is considered. A model for the spread of cascading failures between interdependent CII objects is proposed, using a weighted graph model. Based on this model, formulas for calculating reliability dependencies and assessing risks between CII objects are presented. Additionally, a risk assessment method accounting for cascading effects for CII objects has been developed and applied.

Keywords: Critical Information Infrastructure (CII); Critical Information Infrastructure Object; Critical Information Infrastructure Subject; Failure; Cascading failure; Risk; Risk assessments; The graph; Probability; Information security.

I. Introduction

Critical Information Infrastructure (CII) refers to a set of information systems, automated control systems, and information-communication networks that support activities in essential areas, the disruption of which can cause significant harm to the interests of citizens, society, and the state [1]. Worldwide, and especially in the Republic of Azerbaijan, in recent years, issues related to ensuring the security and stability of these systems have become particularly relevant, considering their high degree of interconnection.

One of the most important aspects of risk management in CII is the prevention and minimization of the consequences of potential failures. One of the threats that is of significant interest for research is cascading failures.

A cascading failure is a process in a system of interconnected parts in which the failure of one or several parts can trigger the failure of other parts, and so on. Cascading failures in CII can be especially dangerous, as they can lead to widespread consequences, affecting not only one object or component of the system but also causing failures in other parts of critical infrastructure. Cascading failures are extensively studied in power grids. Effective risk assessment of cascading failures includes the probability of failures, potential consequences, and the impact of one object's failure on others [2-11].

The goal of this work is to study the interconnected and interdependent risks of cascading failures between CII objects.

II. Analysis of the Research Area

The main components and assets of Critical Information Infrastructure (CII) are the objects and subjects of CII, as well as the critical processes that take place within them. According to the Law of the Republic of Azerbaijan "On Information, Informatization, and Information Protection," a CII object is an information system, an automated control system, and/or an information-communication network. The subjects of CII are government bodies (institutions), legal entities, and individual entrepreneurs. According to this law, the most important areas of CII activity in the Republic of Azerbaijan include government administration, defense, healthcare, the financial market, energy, transport, information technology, telecommunications, water supply, and ecology [1].

In the work [2], the subject of CII is considered as a system of interacting CII objects, as well as the means of their communication, owned by a specific CII subject. Based on this, it can be assumed that such systems are implemented in 10 areas of CII activity. As a result, interconnected cascading failures in CII can occur in the following directions: between objects within the same system, between different systems within the same activity area, and between systems from different CII activity areas. The third case occurs, for example, when a failure in the energy sector system can disable a system in the water supply sector.

The article [3] provides an example of a risk dependency graph for interdependent CIIs, where the consequences of one risk arising in one CII and affecting other dependent critical infrastructures are assessed. Operational risk assessment methods adopted in financial organizations are used as a method.

The values of interconnection and interdependence are determined by the concept of mutual influence. The authors [3], [4], [5] in their works have examined the types of mutual influences. The main types can be highlighted as follows: physical, geographical, and informational.

- The physical type defines the influence caused by the exchange of energy or matter between systems, which changes their state.
- The informational type refers to the influence arising from the exchange of data, signals, or information between systems, leading to a change in their state.
- The geographical type is determined by the spatial arrangement of systems, which spreads the consequences based on their proximity.

The article [5] discusses a method for assessing the risks of cascading failures (breakdowns) using dynamic criticality matrices. The method is based on the application of linguistic approximation, which allows for heuristic forecasting of accident developments.

The work [6] presents a model for the spread of cascading failures of elements in an electric power network across its cuts based on a dependency graph. Within this model, network reliability indicators were introduced. Methods for exact and approximate reliability calculation of the network were developed based on the Monte Carlo method, as well as a cumulative method for refining reliability boundaries.

III. Model of Cascade Failure Propagation Between Interdependent CII Objects

Probabilistic models are typically used to analyze the reliability of a network and assess risk in the event of cascading failures, which consider the failure dependencies between its elements [6].

To study the model of interdependent CII objects, we will use a weighted graph G=(V, E) (Figure 1).

Let the vertices of the graph V represent the CII objects (O1 …On), and the edges of the graph E represent the failure dependencies between these objects. Assume that the vertices of the graph

occur with probability p_v=1, while the edges occur with probability 0≤p_e≤1 within a specified time interval. Suppose that for each object i in the system, there is a weight w_i that can be used to weight the probability of failure of the object (or the vulnerability degree of the nodes).

Thus, one approach would be to use a calculation that takes into account the contribution of each object to the system's failure. Formula (1) allows for the calculation of the probability that at least one object i will be responsible for the system's failure, considering its weight.

$$P(A) = 1 - \prod_{i=1}^{N}(1 - w_i * p_i) \tag{1}$$

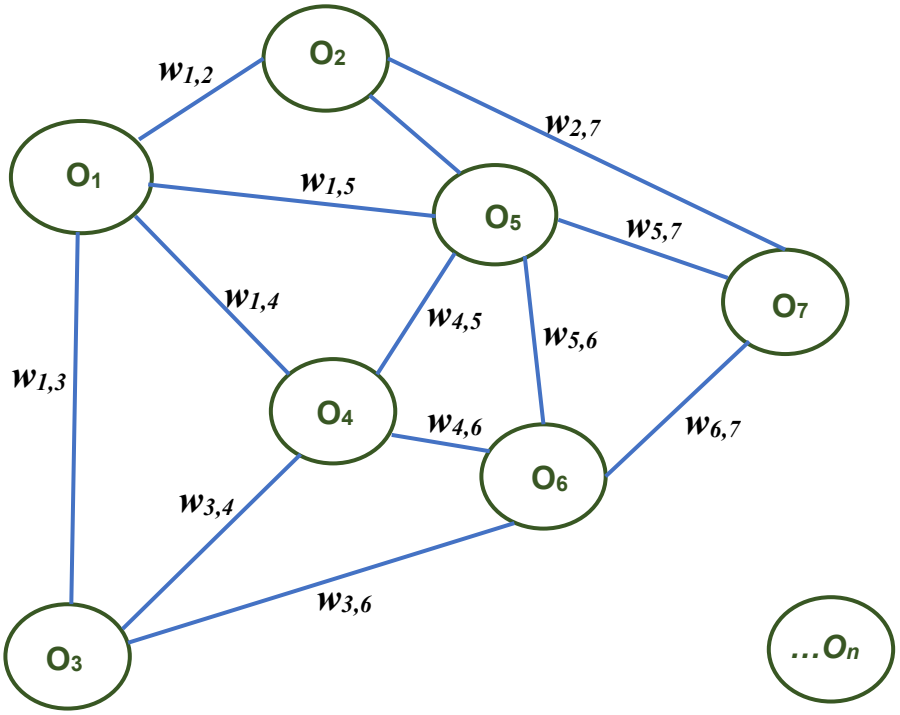


Figure 1: *Example of a weighted graph model of CII objects*

Now, if we want to take the weights into account when calculating the probability that a specific object i will be responsible for the system's failure, we can use the following formula (2):

$$P(B_i) = \frac{w_i * q_i}{\sum_1^N w_j * q_j} \tag{2}$$

And finally, to calculate the probability that the failure of a specific object i will lead to the failure of the entire system, considering its weight, the formula will be as follows: Formula (3):

$$Q_i = P(A) * P(B_i) = \frac{(1-\prod_1^N p_i)*w_i*q_i}{\sum_1^N w_j*q_j} \tag{3}$$

If the failure of one or more objects in the system can lead to a complete system breakdown, then this probability becomes an important element for assessing risk factors. The connection

between probabilistic failure models and risks is that these models not only allow for the estimation of failure probabilities but also enable the analysis of how the failure of each object may affect the overall risk of the system.

Formula (1) takes into account how the risks of individual objects (their failure probability and importance to the system) influence the overall risk of the system as a whole.

Formula (2) allows the calculation of the probability that the failure of a specific object i will be responsible for the total system failure. This helps to identify key risks within the system — objects with high values of w_i (object weight) and q_i (failure probability), which can significantly impact the system.

Formula (3) helps link the probability of failure of a specific object i to the overall system failure probability. By considering Q_i, we can assess how the failure of a specific object affects the risk to the entire system.

IV. Risk Assessment of Cascade Failures

Risk assessment in information security (IS) plays a crucial role in ensuring the protection of information infrastructure. Risk assessment can be performed using two main methods:

- Qualitative risk assessment;
- Quantitative risk assessment.

Qualitative risk assessment methods include: expert judgment method, rating assessment method, analogy method, and others. Quantitative methods include techniques such as scenario analysis, simulation modeling (Monte Carlo simulation method, historical simulation method), situation modeling based on game theory, tree-building methods, methods based on artificial intelligence systems, and so on [7].

According to the work [8], the qualitative and quantitative risk assessment of the information security (IS) of a CII object is determined as the product of C — the potential damage (consequences) caused to the i -th information resource of the designated security zone, and the probability P_j of the occurrence of the j-th threat and the probability P_k of exploiting the k -th vulnerability. The formula (4) is as follows:

$$R_i = P_j * P_k * C_i \tag{4}$$

A threat can arise from both natural and artificial phenomena and includes the probability that it exists or may occur. Vulnerability is a weak point or limitation that a threat can exploit. Consequences refer to the cost and loss coefficient for risk assessment [9].

For a more detailed assessment of the risks of cascading failures in a system of critical information infrastructure (CII) objects, we need to consider not only the direct failure risks for each object but also their interconnection. When one object fails, it can affect other objects, causing additional failures, leading to a cascading effect. This effect can be modeled using a risk matrix to systematize and quantitatively assess the impact of failures of different system objects on the operation of the entire CII.

So, to account for cascading effects, we need to adjust the overall risk for each object, taking into account the failure probability of not only the object itself but also the probabilities of failure of other objects that may be triggered by cascading failures. Formula (5):

$$R_i = P_i * w_i * C_i + \sum_{j \neq 1} P_j * w_j * C_j * E(i \to j) \tag{5}$$

Where:

- P_i- is the probability of failure of object i,
- w_i- is the weight of the object,
- C_i is the consequences of the failure of object i,
- $E(i \to j)$ is the cascade effect $(i \to j)$ coefficient, which shows the impact of the failure of object i on other objects j.

Based on formula (5), an example of risk assessment for cascade failures of CII objects can be presented.

Example:

Suppose we have three objects in one CII system with the following data (Table 1, 2):

Table 1: *Cascade Effect Matrix for All Pairs of CII Objects* $(i \to j)$

Object i/ Object j	O_1	O_2	O_3
O_1	*0*	*0.4*	*0.3*
O_2	*0.3*	*0*	*0*
O_3	*0.2*	*0*	*0*

Table 2: *Risk Assessment Matrix for Cascade Failures of CII Objects*

Objects	P_i	w_i	C_i
O_1	*0.1*	*2*	*8*
O_2	*0.05*	*1*	*7*
O_3	*0.2*	*1*	*6*

$$R_1 = 0.1*2*8+0.05*1*7*0.4+0.2*1*6*0.3 = 1.6+0.14+0.36 = 2.1$$

Output:

The risk of failure of object 1 (O_1), considering the cascade effects for the system, is *2.1*. This means that the failure of object 1 not only has its own risk (*1.6*) but can also impact the failure of other objects (*0.14* for Object 2 and *0.36* for Object 3), thereby increasing the overall risk for the entire system.

V. Conclusion

Cascade failure (breakdown) is considered the most complex scenario in terms of design, security assessment, and response to these failures. During the analysis of the topic under study, a systems approach was considered, which is implemented in various areas of CII activities.

A model of cascade failure propagation between interdependent CII objects is proposed, using a weighted graph model as an example. Based on this model, formulas for calculating the reliability of dependencies and risk assessment between critical information infrastructure objects were presented. Additionally, based on both qualitative and quantitative risk assessment methods for information security, a risk assessment method taking into account cascade effects for CII objects was developed and applied.

References

[1] The Law of the Republic of Azerbaijan "On Information, informatization, and information Protection". Official website e-qanun.az - unified electronic database of legal acts. 2022.

[2] Maksimova Elena A., and Sadovnikova Natalya P. "Intersubjective interaction as a source of destructive Influences on the subject of critical information infrastructure". Caspian journal: Control and High Technologies 2 (54), pp. 71-80. 2021.

[3] Imamverdiev Y. N. "Method of Protecting Distributed Networks Through the Formation About Information Exchange". Information Systems and Technologies 111.1, 2019.

[4] Panteleev V. A., Kirillov I. A., Berberova M. A., Klimenko S. V. " Method Cascade and intersystem accidents scenarios description." SCVRT2017, pp. 239-244. 2017.

[5] Brezhnev E.V. "Method of Cascading Failure (Accident) Risk Assessment Based on Application of Dinamic Criticality Matrix". Science and technology of the Air Forces of the Armed Forces of Ukraine No. 1 (18). pp.187-190. 2015.

[6] Migov D. A., Korotkov A. N. "Cuts Using for Modeling the Propagation of Cascading Failures in Electrical Power Grids." Problems of informatics (3). 2021.

[7] Petrova A.V. «Methods of Assessing the Level of Risk in the Enterprise». Economics and Efficiency of Production Organization, (21). pp. 97-104, 2014.

[8] Vulfin Aleksey Mikhaylovich. " Models and methods for comprehensive assessment of security risks of critical information infrastructure objects based on intelligent data analysis." System engineering and information technology 5.4 (13). pp.50-76.2023.

[9] Abdulova E. A., Kalashnikov A.O." On The Issue of Risk Management of Critical Information Infrastructure." Papers. pp.1275-1282. 2021.

[10] Komendantova Nadejda P. " Risk Governance and Vulnerability Factors of Critical Infrastructure." Russian Digital Libraries Journal V. 20. No 1. pp.88-108. 2017.

[11] Kotenko Igor, Saenko Igor, Doynikova Elena. "Risk Assessment in Computer Networks of Critical Infrastructures". Innovations in science 16-1. pp. 84-88. 2013.

INVESTIGATION OF SURFACE ROUGHNESS IN HYDROABRASIVE MACHINING DEPENDING ON CHANGES IN ABRASIVE GRAIN SIZE AND PRESSURE

Sylvio Simon[1], Nizami Yusubov[2], Samir Amirli[2]

[1]Brandenburg University of Technology, Cottbus-Senftenberg, Senftenberg, Germany
[2]Azerbaijan Technical University, Baku, Azerbaijan Republic
sylvio.simon@b-tu.de, nizami.yusubov@aztu.edu.az, amirlsam@b-tu.de

Abstract

The article investigates the surface roughness generated on HARDOX-500 chromium-nickel steel blanks in hydroabrasive machining as a function of changes in abrasive grain sizes. The study examines the intervals of roughness variation based on various technological, kinematic, structural, and processing environment factors, and identifies optimal roughness values. Experimental results indicate that as the granularity of abrasive grains increases, the surface roughness on the cut surface of the blanks also increases. However, as the pressure of the water-abrasive mixture and the consumption of abrasive grains rise, the height of the resulting surface roughness decreases.

Keywords: hydroabrasive machining, steel blanks, abrasive grains, surface roughness, regression coefficient, granularity, pressure.

I. Introduction

Chromium-nickel alloyed steels of the HARDOX-500 grade find extensive application in various fields of mechanical engineering, including the aerospace, shipbuilding, and other sectors. The alloying of steel with up to 1.5% chromium and nickel enhances its resistance to bending, wear, and friction, but also complicates its mechanical processing to some extent. Therefore, performing cutting operations on HARDOX-500 steel blanks using the hydroabrasive method can improve productivity and quality parameters in the production of machine parts. In this process, steel sheets and blanks are cut by a waterjet mixed with abrasive particles, applied to the surface of the processed blank at high pressure (3500 bar) on a specialized machine [1]. Depending on the thickness of the cut blank, the surface roughness and wave patterns produced in hydroabrasive cutting can vary across the height of the cut surface [11]. Studies have shown that the formation of geometric surface features is influenced by the physical and technological characteristics of the hydroabrasive cutting process, cutting parameters, as well as the physical-mechanical properties and composition of the material being processed. Thus, examining the roughness, waviness, and other geometric parameters of hydroabrasively cut surfaces in HARDOX-500 steel blanks is one of the relevant challenges in the mechanical engineering industry [2].

One of the primary objectives of studying surface roughness and waviness in the hydroabrasive cutting of chromium-nickel steel is to efficiently determine the sequence of preliminary mechanical processing technology for parts to be manufactured from these blanks. By analyzing the topography

of roughness and form errors on the cut surface, we can evaluate the resulting geometric features and irregularities to ensure proper control over the chip formation process when directing abrasive particles onto the blank surface with a waterjet across the cut thickness in HARDOX-500 steel [3-5].

Theoretical research has indicated that explaining the formation of quality parameters in manufacturing machine parts from hard-to-process chromium-nickel materials using hydroabrasive cutting is theoretically complex. Therefore, experimentally studying the output parameters and synthesizing these findings with theoretical results is crucial for determining the optimal technological parameters for hydroabrasive cutting. Experimental studies examine quality parameters such as surface roughness (R_a), (R_z), dimensional accuracy (Δ), microhardness of the processed surface (H_μ), processing time (t), processing efficiency, and other factors [2, 6-9].

The symbols and units of measurement for the parameters given in the article are provided in the table 1.

Table 1: *Symbols and Units of Measurement for Parameters*

R_a [μm]	arithmetic average roughness
Rz [μm]	mean roughness depth
Z [μm]	abrasive particle size
P [Mpa]	pressure of the water-abrasive jet
Q [g/l]	abrasive consumption
S_{long} [mm/min)	feed rate
H_μ [μm]	microhardness of the processed surface
t [sec]	processing time
Δ [μm]	dimensional accuracy

Several articles dedicated to the study of the problem posed by us have been published in periodicals.

In abrasive waterjet processing, surface roughness depends on cutting parameters, including the pressure of the water-abrasive jet, the feed rate of the mixture in the cutting zone, the variation speed of the longitudinal feed movement of the nozzle or workpiece, the thickness of the workpiece, and other factors. Accordingly, experimental research has been conducted on surfaces processed with a FLOW-Gut model CNC-controlled abrasive waterjet machine in the "Metal-Cutting Machines" Department at Brandenburg University, Germany, using various devices based on a methodology developed for these experimental studies.

The research has shown that the surface roughness of a workpiece processed by abrasive waterjet cutting varies significantly depending on the thickness of the workpiece. Given the sharp variations in cross-sectional shape and dimensions as the thickness of the workpiece changes, examining the resulting surface roughness enables an exploration of the technological capabilities of this operation [9].

Purpose of the study. The study investigates the surface roughness obtained in the cross-section during the abrasive waterjet cutting of HARDOX-500 steel workpieces with thicknesses of 5, 10, and 15 mm, focusing on the effects of variations in abrasive particle size, granularity, abrasive consumption, and the pressure of the water-abrasive jet.

II. Methodology

The investigation involves studying the surface roughness obtained during the abrasive waterjet cutting of workpieces with thicknesses of 5 mm, 10 mm and 15 mm, using a feed rate of S_{long}=26,7 mm/min, an abrasive particle size of 80÷200 μm, and an abrasive consumption of

Q=125g/l, while varying the pressure of the water-abrasive jet from 200 Mpa to 350 MPa. This study encompasses the establishment of various curves, corresponding mathematical equations, and regression coefficients using experimental and theoretical values in an Excel program, as well as discussing the obtained results.

One of the output parameters from the conducted experiments is the roughness, which is determined with high accuracy using the "JENOPTIK" device, designed to measure the surface roughness based on the granularity of the abrasive and the longitudinal feed rate [3].

III. Discussion of the results

The surface roughness formed during the cutting of HARDOX-500 steel workpieces varies based on numerous technological, kinematic, structural, and processing environment factors. Consequently, one of the most important technological factors affecting the roughness formed on the surface during abrasive waterjet cutting is the pressure of the water-abrasive mixture applied to the cutting zone. An increase in the pressure of the water-abrasive mixture enhances the cutting capabilities of the water-abrasive, which plays the role of the cutting tool in the steel cutting process, thereby intensifying the cutting of the workpiece. Our research has determined that the study of the surface roughness dimensions in hydroabrasive machining varies widely depending on the processing conditions and regime parameters of the process. Therefore, examining the regularities of roughness changes during the machining of the selected material is one of the important tasks for identifying the advantages of the process. In this case, the roughness formed on the processed surface, depending on the optimal cutting process, has been determined through experimental research to remain within the required limits. In the experiments, a workpiece thickness of 15 mm was taken, with a longitudinal feed rate of S_{long}=26,7 mm/min, an abrasive particle size of 80 μm, and an abrasive consumption of Q=125g/l, while the surface roughness obtained during abrasive waterjet cutting [3] is shown in Figure 1, and the dependence of the obtained experimental and theoretical values on the influence of the water-abrasive jet is presented in Tables 2 and 3.

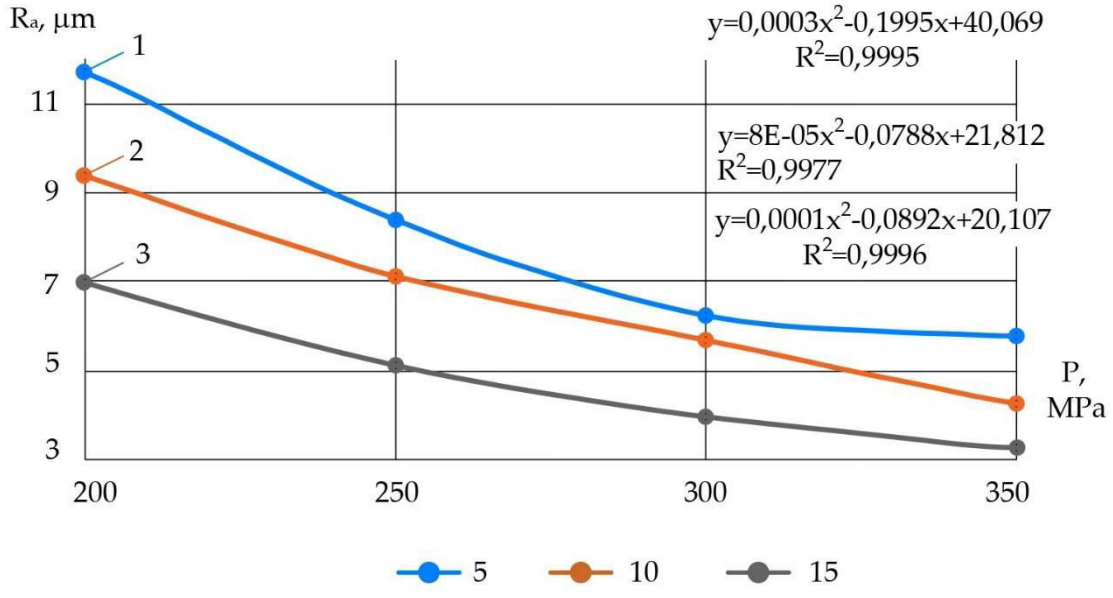


Figure 1: *Dependencies of surface roughness obtained on the cut surface during abrasive waterjet machining on the pressure of the water-abrasive jet*

Table 2: *Experimental values of surface roughness obtained on the cut surface during abrasive waterjet machining*

$00P$, [MPa]	200	250	300	350
R_{a1},[μm]	11,732	8,365	6,214	5,742
R_{a2},[μm]	9,375	7,115	5,675	4,234
R_{a3},[μm]	6,98	5,113	3,964	3,273

The theoretical values of surface roughness obtained on the cut surface during abrasive waterjet machining are presented in Table 3.

Table 3: *Theoretical values of surface roughness obtained on the cut surface during abrasive waterjet machining*

P_0, $[MPa]$	200	250	300	350
R_{a1} $[\mu k]$	12,169	8,944	7,219	6,994
$R_{a2}[\mu k]$	9,252	7,112	5,372	4,032
$R_{a3}[\mu k]$	6,267	4,057	2,347	1,137

The mathematical equations of the obtained graphical curves (Figure 1) are presented in equation (1).

$$\begin{aligned} R_{a1} &= 40{,}069\text{-}0.1995P_0 + 0{,}0003\ P_o^2 \\ R_{a2} &= 21{,}812\text{-}0{,}0788P_0 + 0{,}00008\ P_o^2 \\ R_{a3} &= 20.107\text{-}0.0892P_0 + 0{,}0001P_o^2 \end{aligned} \quad (1)$$

Figure 1 shows that the first curve corresponds to a thickness of 5 mm, the second curve to a thickness of 10 mm, and the third curve to a thickness of 15 mm, representing the surface roughness of the cut materials as the pressure of the water-abrasive jet varies from 200 MPa to 350 MPa. As seen from the graphs, in all cutting cases, the surface roughness of the cut material decreases as the pressure of the water-abrasive jet increases from 200 MPa to 350 MPa.

Experiments reveal that as the pressure of the water-abrasive jet increases, the forces generated upon impact of the abrasive particles on the processed surface also increase, resulting in a higher number of broken particles. As the incidence of abrasive particle breakage rises, the number of sharp edges in the newly fractured abrasive particles increases, which in turn reduces the thickness of the resulting chips. Consequently, the average height of the generated surface roughness decreases.

Research has shown that as the thickness of the workpiece increases, the surface roughness in the cutting zone decreases. This can be explained by the fact that at smaller workpiece thicknesses, such as h=5mm (Curve 1), fewer abrasive particles concentrate in the cutting layer during hydroabrasive cutting, resulting in a lower number of abrasives per unit surface area. Consequently, the thickness of the chip layer removed from the surface increases, which leads to greater roughness. Additionally, with fewer abrasive particles directed at the cut, the total impact forces decrease, leading to less breakage of abrasive particles and thus a reduced likelihood of new cutting edges forming, which also contributes to increased roughness. As the thickness of the workpiece increases (e.g., to 10 mm, 15 mm), the water-abrasive jet supplied to the cutting zone does not escape beyond the contact area, resulting in an increased number of cutting particles

forming chips. Consequently, the volume of chips removed by each abrasive particle decreases, which reduces the height of the roughness. It has been observed that the trend of decreasing roughness with increasing workpiece thickness holds at all pressures of the water-abrasive jet.

One of the critical factors in the cutting of metals with free abrasive particles, i.e., in

hydroabrasive cutting, is the geometric shape and size of the abrasive particles, with granularity defined according to existing standards [10, 12]. During the experiments, particles with sizes of 80 μm, 120 μm, 160 μm and 200 μm were used in the hydroabrasive process. The variation in roughness as a function of abrasive particle size is shown in Figure 2, and the experimental values for the effect of particle size on roughness are provided in Table 4. Particle size is denoted by Z, and its impact on roughness formation is studied as Z increases.

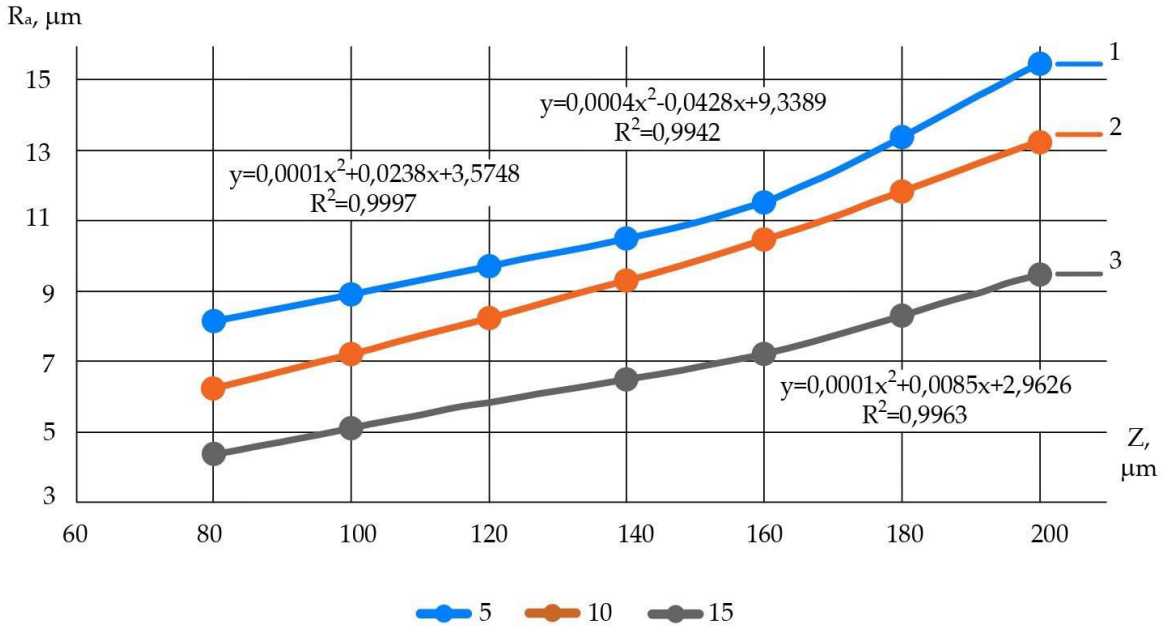


Figure 2: *Graph of the variation in roughness based on abrasive particle size and granularity*

Table 4: *Experimental values of the effect of abrasive particle size (Z μm) and granularity on roughness*

Z, [μm]	80	120	160	200
R_{a1}, [μm],h=5mm	8,16	9,743	11,564	15,485
R_{a2}, [μm],h=10mm	6,245	8,26	10,465	13,265
R_{a3}, [μm],h=15mm	4,354	5,852	7,216	9,475

The equations determined by solving the mathematical expressions of the curves based on the values in Table 4 are presented in (2).

Table 5 provides the theoretical values for the dependence of roughness on Z.

Table 5: *Theoretical values of the effect of abrasive grain size (Z) and granularity on roughness*

Z, [μm]	80	120	160	215
R_{a1}, [μm],h=5mm	8,4749	9,9629	12,7309	16,7789
R_{a2}, [μm],h=10mm	5,9149	4,2029	2,4909	0,7789
R_{a3}, [μm],h=15mm	4,2826	5,4226	6,8826	8,6626

Equations determined by solving the mathematical expressions of the curves derived from the values shown in Table 4.

$$R_{a1}=9,3389-0,0428Z+0,0004Z^2$$
$$R_{a2}=3,5748+0,0238Z+0,0001Z^2 \quad (2)$$
$$R_{a1}=2,9626+0,0085Z+0,0001Z^2$$

The graphs shown in the figure correspond to cutting samples with thicknesses of h=5 mm (Curve 1), h=10 mm (Curve 2), and h=15 mm (Curve 3). The dependency of the average surface roughness (R_a) on changes in abrasive particle changing from 80 μm to 200 μm has been analyzed. It has been found that, for all three thicknesses, the average roughness (R_a) increases as the abrasive grain size increases. This result is explained by the fact that as the grain size increases, the number of particles in a unit volume of the waterjet participating in the cutting process decreases sharply. As particle size increases, their geometrical dimensions increase, resulting in larger cutting edges forming the chip, which, in turn, increases the chip size in the cutting zone. Additionally, as abrasive particle size increases, their resistance to cutting forces also rises, reducing the number of particles subjected to breakage. Consequently, dulling of the abrasive particles due to edge wear becomes more prominent, making the chip formation process more challenging and leading to a higher average roughness.

Nevertheless, since the number of abrasive particles involved in cutting process decreases, the roughness height increases. Studies have shown that while the rate of increase in average roughness height is lower with smaller abrasive particles (eg., 80 μm), as particle size and workpiece thickness increase, the roughness range becomes significantly higher (see the values obtained with a particle size of 200 μm in Figures 1, 2, and 3).

Thus, for hydroabrasive cutting, it is essential to select abrasive particles of an optimal size that meets the required roughness limits. Considering that an average roughness height of 4,5 μm to 6,5 μm is typically required for machine parts made from chrome-nickel steel, it is recommended to use abrasives with particle which sizes between 80 μm and 120 μm for cutting HARDOX-500 steel parts with thicknesses of 15 mm to 20 mm. Cutting with these recommended abrasive sizes results in fine, thin chips, ensuring that the roughness of the processed surface meets the required conditions.

One of the factors influencing the formation of roughness on the cut surface in hydroabrasive processing is the mass of abrasive particles mixed with the waterjet. As the mass of abrasive particles mixed into the waterjet increases, the amount of abrasive involved in cutting also rises, thereby increasing the volume of chips removed from the contact zone per unit time in hydroabrasive processing. This is because, as the weight of the abrasive particles increases, the number of particles actively cutting at any given time also rises.

IV. Results

1. In hydroabrasive cutting, the change intervals of roughness are studied based on the size of the abrasive particles, their weight consumption, the feed rate of the cutting motion, the pressure of the water-abrasive mixture, and other factors, with optimal values determined.

2. Experimental studies have shown that as the pressure of the water-abrasive mixture and the consumption of abrasive particles increase, the height of roughness formed on the processed surface decreases.

3. Research indicates that as the granularity of the abrasive particles increases, the values of roughness formed on the cut surface of the workpiece in hydroabrasive cutting rise, which is why it is recommended to select abrasive particle which sizes between 80 μm and 125 μm for this process.

This work was supported by the Azerbaijan Science Foundation- **Grant № AEF-MGC-2024-2(50)-16/01/1-M-01**

References

[1] Amirov F. G., Simon S., Steffen W., Amirli S.F., Frana K. (2021). Determining the Accuracy of Water Pressure Processing using 3D Scanning. Herald of the Azerbaijan Engineering Academy, Vol.13, No.3, pp. 38–44.

[2] Simon S., Yusubov N.D., Amirli S.F. (2024). Formation of Geometric Parameters of the Surfaces of Cylindrical Parts during Waterjet Cutting. Advances in Science and Technology Submitted, Vol.148, pp. 59–64. https://doi.org/10.4028/p-l5wSEu

[3] Simon, S., Yusubov, N.D., Amirli, S.F. et al. Waterjet Cutting of HARDOX-500 Workpiece. Russ. Engin. Res. 44, 1572–1576 (2024). https://doi.org/10.3103/S1068798X24702927

[4] Simon S., Yusubov N.D., Amirli S.F., Amirov F.G. Planning of Full Factorial Experiments for the İnvestigation of Roughness in Hydroabrasive Waterjet Cutting of Hardox-500 Steel.// Pakistan Journal of Life and Social Sciences. (2024), 22(2):p. 21590-21595.

[5] Simon Sylvio, Yusubov N.D., Amirli S.F., Amirov F.G. Certain features of chip formation in the waterjet machining. BMSTU Journal of Mechanical Engineering, 2024, no. 11, pp. 53–61.

[6] Steffen W. (2014). Schneiden mit dem Wasserstrahl. Masterthesis, BTU Cottbus – Senftenberg, p. 38.

[7] Schulz, F. (2015). Fertigungstechnik. Springer, Vol.11, p. 408.

[8] Kolb, M. (2006). Wasserstrahlschneiden: Materialbearbeitung mit einem Hochdruckwasserstrahl, p.70.

[9] Tamarkin M. A., Tikhonov A. A. Tishchenko E. E. (2014). Metal removal in hydroabrasive machining. Russian Engineering Research, Vol. 34, No. 3, pp. 175–177.

[10] Metal-Cutting Tools. (1989). Textbook for university students in "Manufacturing Technology" and "Metal-Cutting Machines and Tools"/M. : Mashinostroenie, 328 pages.

[11] Natarajana Y., Murugesanb P.K., Mohanb M., Khanc S.A.L.A. (2020). Abrasive Water Jet Machining process: A state of art of review. Journal of Manufacturing Processes, Vol. 49, pp. 271-322.

[12] CL Trade Services. https://www.cltradeservices.de

EFFECT OF HIGH-SPEED SINTERING ON THE STRUCTURE AND PROPERTIES OF MOLYBDENUM POWDER STEELS

Subhan Namazov[1], Shahin Mashayev[1], Taleh Taghiyev[1]

•

[1]Azerbaijan Technical University, Huseyn Javid Ave.-25, Az-1073, Baku
subhan.namazov@aztu.edu.az, shahin.mashayev@aztu.edu.az, taleh.taqiyev@aztu.edu.az

Abstract

Due to the fact that the technology of sintering powder steels is carried out at different tem-peratures and conditions, the formation of their structure and properties is obtained some what dif-ferently. Proper preparation of the sintering technology allows to obtain high density and properties of smooth steel. In most cases, the reason for the reduction of the properties of powder steels is that the diffusion process, which occurs due to the low sintering temperature and sintering time, is weak or not at all. As we know, the density and many physical and mechanical properties of the product increase due to pore diffusion during sintering.

Keywords: effect, high-speed sintering, powder, steel, drilling, sintering, technology, molybdenium powder steel.

I. Introduction

The advantage of the process is the availability of the possibility of making details and products of very complex shapes by this method. The main purpose of the work is to investigate the possibility of obtaining high density and properties of ovate steels using the SPS method [1]. Many technologies are used to obtain high density molybdenum scrub steels. The most modern of these technologies is plasma method (SPS) sintering, which includes pressing, sintering and, in most cases, thermal proce-ssing modes. The main advantage of this method is the combination of several technological opera-tions and a significant reduction in the time spent on the procession. In the process of sintering by the SPS method (Spark-Plasma-Sintering), the formed properties become higher and more durable than in other methods.

For several years now, they have been using the plasma method (SPS) big laser technology in powder metallurgy. With this sintering technology, it is possible to sintering powder products consisting of submicron and nano powder. Currently, very extensive literature can be found on the technology of sintering by the SPS method. Looking at these literature, it is possible to ac-company the production of high-quality powder steels of both simple shapes and complex confi-gurations. In general, in addition to plasma sintering technology, they also widely use FAST (Fi-eld Assisted Sintering Technology) sintering technology with electric heaters [2]. The main ad-vantages of FAST and SPS sintering technologies are the possibility of heating in a very short ti-me, the possibility of thermal processing in a short period of time and the possibility of automatic adjustment of obtaining a small-grain steel structure [3]. The strength and toughness of powder steels sintering by these methods is significantly higher than when sintering by conventional met-hods. Press-mold made of

graphite in the process of sintering with SPS and FAST technology creates conditions for obtaining such details and products from solid alloys by this method [4].

The Spark Plasma Sintering (SPS) method is an effective technique for the compaction of powder materials. A main characteristic of this method is the direct heating of the pressing tool and-or the sample by pulsed direct electrical current with low voltage. This results in high heating ra-tes and allows for short treatment times in order to obtain highly compacted sinter bodies. The material transport (e.g. by diffusion) occurring during the sintering process can also be used for performing chemical reactions. Especially the conditions during the SPS process allow the use of the method also as an alternative synthesis route for intermetallic compounds, of which, some can be obtained only with difficulties by other techniques.

The purpose of the work the MPI for Chemical Physics of Solids was the first institute in where an SPS setup had been installed. Two SPS apparatus are available, one of them being installed inside an Argon-filled glove box, which make it very useful for the compaction and/or synthesis of air/moisture sensitive samples. Both machines allow for external forces up to 50 kN, direct current up to 1500 A and a voltage limit of 25 V with typical pulse length of 2.5-3 ms.

II. Research Methodology

The brand sintering (FCT HP d 250/1) unit, which consists mainly of electric heaters, is used for high-speed sintering of powder steels. The working temperature of this unit is 2300-2500^{0}C and is equipped with a special vacuum chamber (Figure 1). For the implementation of the experiments, chromed steels were used, which are widely used in powder metallurgy and have high properties. The properties of steels sintering by this method (strength, hardness, resistance to crack formation, etc.) were obtained much higher than steels sintering by other methods. The diameter and height of 30 mm of chrome-plated cast steel was used in the sintering unit, and the sintering temperature was 1200-1300^{0}C. The sintering process was to heat the holding time was 2.4-2.6 minutes in total. Powder steels containing 0.5-1.0% and 1.5% Mo were used for the sinte-ring process. Steels containing 0.5-1% molybdenum were baked at 200-700^{0}C, steels containing 1.0% molybdenum at 500-1450^{0}C, steels containing 1-1.5% molybdenum at 1250-1300^{0}C. The sintering time was 3.0-3.5 minutes and 2.25 minutes, depending on the composition. The pressing pressure of the samples was 300 MPa, 500 MPa and 550 MPa, respectively.

III. Discussion of the results obtained

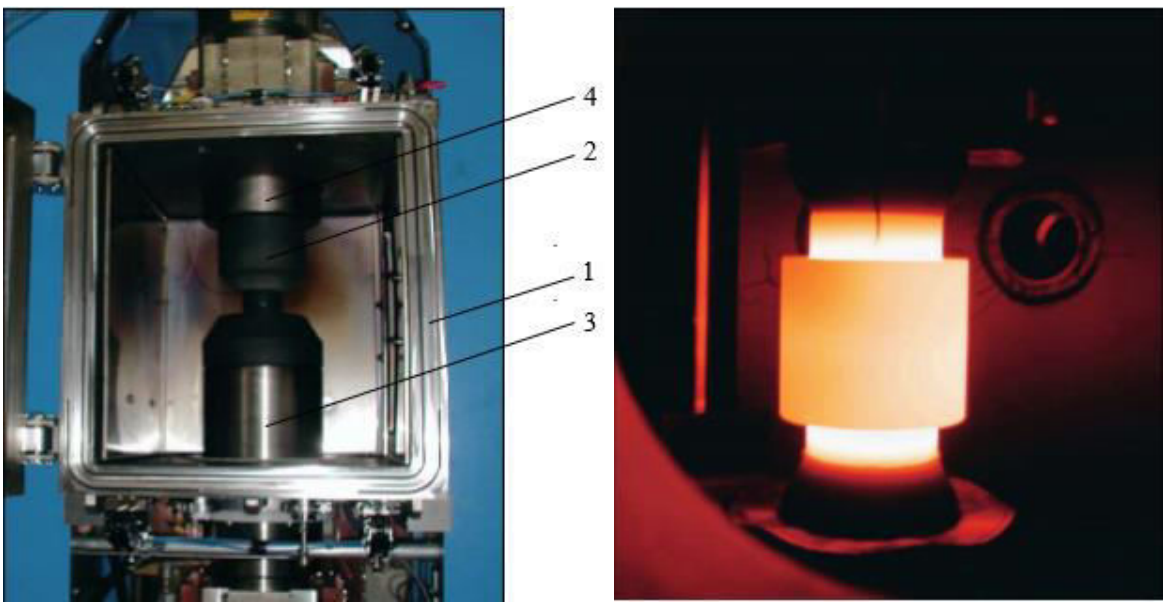
Tungsten and chromium are often present in the composition of molybdenum scrub steels, and, as a rule, the amount of molybdenum in such steels varies in the amount of 0.5-1.5%. When the molybdenum content of powder steels is more than 1.5%, the structure of the steels consists of ferrite, and the molybdenum in the composition, together with iron, forms a compound Fe_3Mo_2 and FeMo intermetallide, which, respectively, contains 53.2% Mo and 63.2% Mo [5]. Molybde-num in molybdenum scrub steels increases the concentration of carbon in the compound in perlite and points the S point in the Fe-C case diagram to the left. Molybdenum is a strong carbidifying element and can easily combine with carbon to form MoC and Mo_2C carbide. The uptake of these carbides occurs mainly when the molybdenum content in steels is 8-10% [6]. However, when the amount of molybdenum is greater than the amount of deyelled, the formation of 3C dicycarbides in cementite (Fe, Mo) becomes even more intenvivilized. It is possible to change the amount of the named carbides in the structure of molybdenum scrub steels with the help of thermal proce-ssing modes.

When the sintering temperature is set at 500^{0}C, Fe_3C carbide is first formed in molybdenum steels, and as the sintering time increases, the formation of Mo_2C carbide is also intensified, which

increases the dispersion of carbides mainly as the sintering time increases. In molybdenum scrub steels during sintering, the solubility of molybdenum in γ and α-iron is very low, and this is due to the fact that the α-iron rin area is extremely high in relation to the γ-iron area. The intensity and diffusion of carbon and molybdenum solubility in γ-iron occurs at a sintering temperature of 1000^0C. İs also found in cases where this diffusion occurs at 1000-1200^0C, and the diffusion coef-ficient of molybdenum intensifies even more during the recrystallization of iron [7]. As the initial temperature of martensite conversion increases in all molybdenum pofrets, the diffusion coeffici-ent of molybdenum rises significantly compared to that of carbon. However, this diffusion does not have a significant impact on the thermodynamics of the perlite structure. While the fragility of molybdenum scrub steels increases, their strength, corrosion resistance and inedible endurance increase. This is due to the fact that dispersed molybdenum carbides are formed in the structure, occupying the entire phase. The further dispersion of carbides, the increase in the temperature of tabulation and tabulation becomes even more crumbly per year. However, when molybdenum is added to some molybdenum scrub steels, the plasticity of steels rises [7]. The chrom-nickel-molyb-denum scrub steels of the brand П20ХН2М were baked at a temperature of 1250-1300^0C in argon mixture for 1.5-2 minutes.

Table 1: *By plasma method of powder steels containing 0,5-1,5% molybdenum sintering technology (SPS method)*

N	Argon pressure, Bar, 1bar =105 Pa	Sintering temperature, 0C	Sintering time, sec	Holding time, sec	Density of steel, qr/sm^3
1	0,1	1220	120	250	7,675
2	0,1	1230	125	270	7,694
3	0,1	1240	130	180	7,725
4	0,1	1350	135	65	7,757
5	0,1	1300	114	60	7,789



a) b)

Figure 1: *Vaccum of plasma sintering (SPS) device view of his camera (a)*
1-body of the vaccum chamber, 2-press-mold made of graphite, 3-lower score,
4-upper score. Sintering in a plasma method cooking (SPS) unit general view of the process (b)

After baking, the steels were subjected to steelmaking at 150-180^{0}C. Carbonyl iron powder with a size of 2.5-3 mkm, calloid graphite powder with a size of 7.6 mkm C-1 and molybdenum powder with a size of 0.75-0.9 mkm were used in the composition. In steels of this type during sintering in general, the coefficient of diffusion of carbon in γ-iron is higher than in molybdenum [8]. The presence of $Me_{23}C_6$ and Me_3C carbides in the structure of this type of steel leads to an increase in micromanagement. When baking a series of powder steel, its structure con-sists of perlite, similar to sorbite, and depending on the amount of molybdenum supplied to the composition, the degree of dispersibility of carbides in the structure is higher than that of pofrets with chrommolybdenum. The reason for this is the formation of the second carbides in the struc-ture, their distribution in the solid solution and the price of the cooling rate. This cooling tem-perature intensifies in the range of 400-500^{0}C, and as a result $(Fe, Mo)_2C$ carbide is formed, resulting in the hardness of steel 1080-1230 HV, and the hardness of molybdenum carbide is 800-900 HV [9].

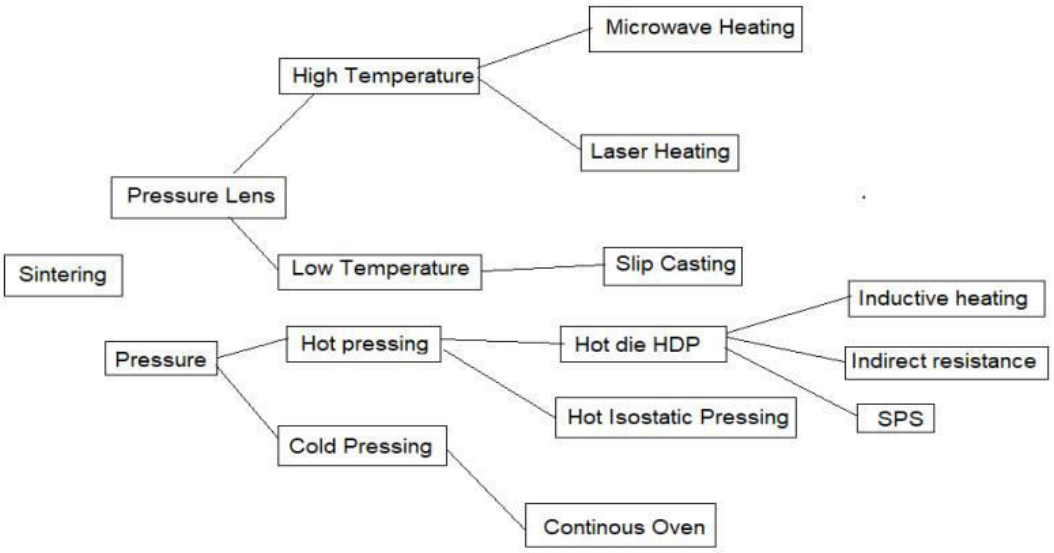


Figure 2: *Standard classification of sintered process*

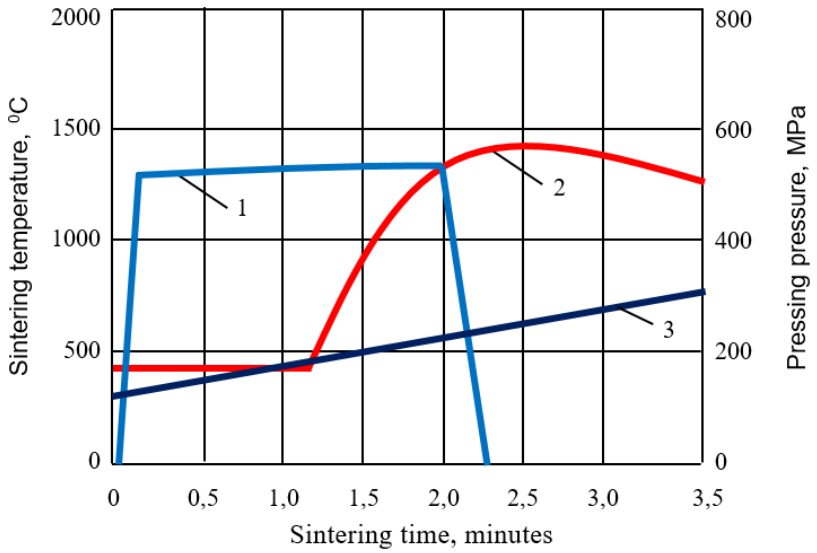


Figure 3: *Plasma method (SPS) sintering of powder steels containing different amounts of molybdenum 1-1,5% Mo, 2-1,0% Mo,3-0,5% Mo*

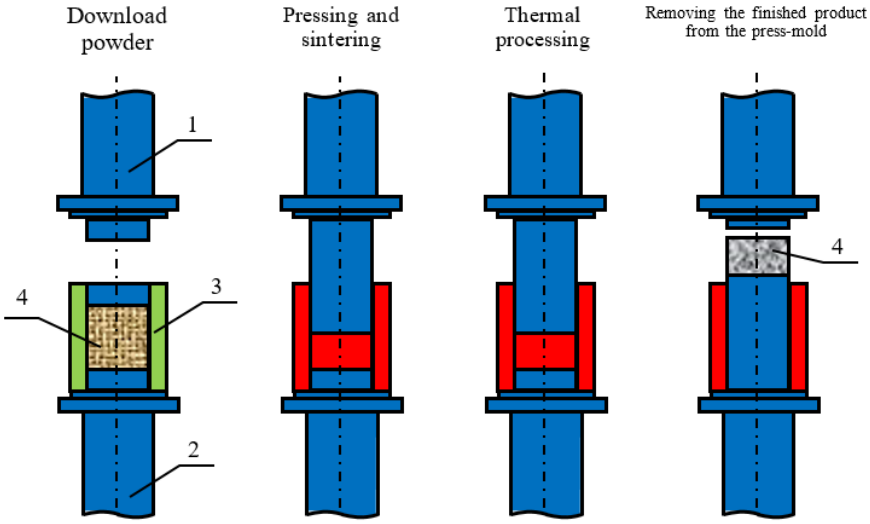


Figure 4: *With SPS technology of powder steels (by plasma method) model of sintering 1-top score, 2-bottom score, 3-special matrix, 4-powder glaze, 5-finished product*

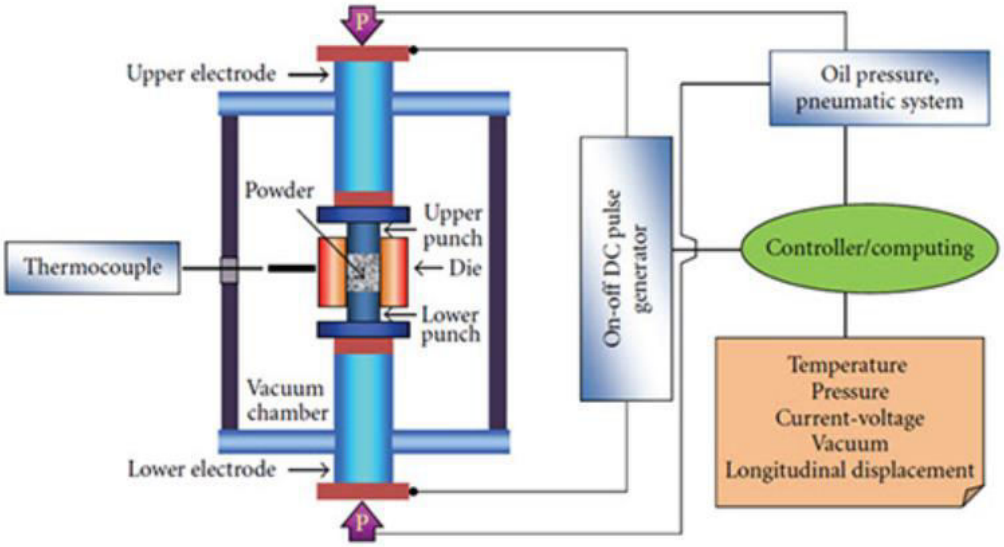


Figure 5: *Schematic of SPS process*

Control of sintering temperature is possible through setting the holding time, ramp rate, pulse duration, and pulse current and voltage. The DC pulse discharge could generate spark plasma, spark impact pressure, Joule heating, and an electrical field diffusion effect. In SPS, sintering is assisted by the on-off DC pulse voltage compared to conventional hot pressing as shown in Figure 5. The application of pressure helps plastic flow of the material. Figure 3 illustrates the flow of DC pulse current through the particles. Usually, SPS is carried out in four main stages. The first stage is performed to remove gases and create vacuum. Then pressure is applied in the second stage followed by resistance heating in the third stage and finally cooling in the fourth stage. When a spark discharge appears in a gap or at the contact point between the particles of a material, a local high-temperature state of several to ten thousands of degrees centigrade is generated mo-mentarily. This causes evaporation and melting on the surface of powder particles in the SPS pro-cess, and necks are formed around the area of contact between particles. The application of pres-sure and current, in addition to the high-localized temperatures generated through resistance pulse heating, improves heating rates and reduces sintering time and temperature leading to the console-dation of

nanopowders without excessive grain growth. On the other hand, SPS is not only a bin-derless process, but also does not require a precompaction step. The powder is directly filled into a graphite die through which current is passed and pressure is applied leading to a fully dense material with superior mechanical properties [10].

IV. Conclusion

1. For the sintering process, powder steels containing 0.5-1.0% and 1.0-1.5% Mo are used. Steels containing 0.5% molybdenum were baked at 250-750^{0}C, steels containing 1.0% molyb-denum at 500-1400^{0}C, steels containing 1.5% molybdenum at 1250-1300^{0}C. The sintering time was 2.0-2.5 minutes and 2.60 minutes, depending on the composition. The pressing pressure of the samples was 300 MPa, 500 MPa and 550 MPa, respectively.
2. The main advantages of fast and SPS sintering technologies are the possibility of heating pro-cess in a very short time, the possibility of thermal processing in a short period of time and the possibility of automatic adjustment of obtaining a small-grain steel structure.
3. During plasma sintering (SPS method), the strength and toughness of powder steels is signify-cantly higher than when sintering by conventional methods. The press-mold made of graphite in the process of sintering using SPS and FAST technology creates conditions for obtaining such details and products from solid alloys by this method.
4. The diameter and height of 30 mm of chrome-plated cast steel was used in the sintering unit, and the sintering temperature was 1250-1300^{0}C. The heating and storage time during the sin-tering process was 2.0-2.6 minutes in total.

References

[1] Kessel, H.U., Hennicke, J., Schmidt, J., Weissgarber, T.,Kieback, B.F.,Herrmann, M., Rat-hel, J.: Feldaktiviertes Sintern „FAST“-ein neues Verfahren zur Herstellung metallischer und keramischer Sinterwerkstoffe. Tagungsband 25. Pulvermetallurgisches Symposium, Hagen, 2006.

[2] J.,Kessel,H.U, Hennicke,.: Field Assisted Sintering Technology („FAST“) for the Consolidation of Innovative Materials. cfi/Ber.DKG 81 (2004) [11] E14–E16

[3] Van-Meensel,K.,Kandukuri,S.Y.,Hennicke, J.,Anne,G., Jiang,D.,Vleugels, J.,Van der Biest, O.: Spark Plasma Sintering of Nanometer Size ZrO2-Al2O3-TiC0.5N0,5 Composites. EMRS 2004, September 6-10, 2004, Poland

[4] Nygren, M., Shen, Z.: Novel Assemblies via Spark Plasma Sintering. Silic. Ind. Spec. 69 (2004) 211–218

[5] Echeberria, J.,Martinez,V., Sanchez, J.M., Bourgeois, L., Barbier, G.,Hennic-ke, J. : Sintering Behaviour of Low Co Content cBNWC/ Co Composites by Either GE-HIP or FAST. Proc.of 16th International Plansee Seminar 2005, Eds. G.Knerin-ger,P.Rodhammer and H.Wildner, Vol.2, HM23, 434–448,(2005)

[6] Alvarez,M., Sanchez, J. M.: Densification of Nanocrystalline Ti(C,N) Powders with Nickel Aluminide Binder Phases Using Field Assisted Sintering (FAST). Submitted for publication to J.Amer. Ceram. Soc.

[7] L., Hennicke, Van Meensel, K.,Echeberria, J.,Sanchez, J.M., Martinez, V., Bourgeois, J., Kessel, H.-U.,Harden,P.,Van der Biest,O., Vleugels, J. : Field Assisted Sintering of Cubic Boron Nit-ride Dispersed Cemented Carbide (CDCC) Composites.EuroPM 2006.

[8] Dorofeev Yu.G. Dynamic hot pressing in cermet. Publishing house "Metalurgia",1986, 177 p.

[9] Zhornyak A.F., Radomyselsky I.D. Obtaining parts with high properties from powders. GOSINTI, M., 1964, pp.4-18.

[10] https://www.totalmateria.com/en-us/articles/spark-plasma-sintering/

IMPROVING THE EFFICIENCY OF MACHINING OPPOSITELY DIRECTED CONICAL SURFACES BY MANAGING DYNAMIC TECHNOLOGICAL RELATIONSHIPS

Nariman Rasulov[1], Ugurlu Nadirov[2], Irada Abbasova[3]

[1]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan
[2]Department of Machine design and industrial technologies, Azerbaijan Technical University, Baku, Azerbaijan
[3]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan
nariman.rasulov@aztu.edu.az, ugurlu.nadirov@aztu.edu.az, i.abasova@aztu.edu.az

Abstract

This paper presents a technology for sequential-parallel machining of internal cylindrical and conical surfaces of oil field couplings, identifies a unique system of forces acting on the boring bar during parallel machining, and since all three corresponding components of the cutting forces are directed in opposite directions, a sharp decrease in the elastic deformations of the elements occurs in the technological system, due to the control of dynamic technological relationships and main angles in plan of the cutters during parallel machining, the values of the cutting force components change in a favorable direction, ensuring a decrease in the range of elastic deformations of the elements of the technological system, reducing their impact on the accuracy of machining, provides the results of studies of elastic deformations by modeling in ANSYS, it was recommended to apply the developed technology and technological measures that ensure the processing of responsible surfaces of couplings with high accuracy and productivity.

Keywords: internal, multidirectional, surface, coaxiality, boring bar, parallelism, machining, tool, dynamic, technological, relationships, deformation, accuracy.

I. Introduction

The multi-tool parallel machining of surfaces is widely used in solving various technological tasks and including for ensuring high processing efficiency. The application of similar technologies on modern CNC machines, as well as the execution of auxiliary tasks related to material removal with high precision and speed, further enhances the efficiency of the technological process [1-5]. The parallel machining of various internally multi-directional threaded and conical surfaces is recommended for a range of couplings in the oil and gas industry [3, 6]. The primary technological challenge is ensuring the precision of the machined conical surfaces, including coaxiality accuracy, as well as increasing productivity.

The accuracy of shape formation is primarily determined by dynamic technological relationships. This is because static interactions are relatively easier to ensure, and their impact on machining accuracy can even be reduced to zero through them control (e.g., errors in basing,

clamping-fixing errors, etc.) [2, 7-11]. However, dynamic relationships, due to the variability of cutting forces inherent to the machining process during material removal, induce elastic deformations of varying magnitude in the elements of the technological system (TS), which cannot be entirely eliminated.

Nevertheless, by controlling dynamic technological interactions, it is possible to reduce the range of elastic deformations in TS elements and include their expected minimum values within adjustment parameters, thereby ensuring high machining accuracy [4, 7, 12-14].

II. Statement of the Problem

Humanity's continuously growing demand for hydrocarbons necessitates drilling wells to greater depths. The conical threads of couplings, used to connect parts of drilling, pump-compressor, and casing pipes, as well as other equipment serving the transportation of extracted oil, are among the elements that bear the weight of downhole tools and equipment while ensuring the tightness of the connection [15-18]. Improving the parameters and coaxiality precision of multi-directional conical threads in couplings enhances their load-bearing and sealing capabilities, thereby ensuring better operational performance [3, 6, 18]. Increasing the load-bearing capacity of couplings, in turn, facilitates the efficient drilling of deeper wells. The high precision of the smooth conical surfaces between threads creates a foundation for improving both the parametric and coaxiality precision of the threads to be cut, through inherited accuracy. Additionally, the results obtained from the parallel machining of these surfaces serve as fundamental data and materials of high importance for developing parallel machining technologies for multi-directional conical threads.

Various machining technologies, implementation methods, and features of auxiliary equipment have been developed for the parallel machining of multi-directional internal conical surfaces and threads in couplings [3, 6]. To implement the proposed innovative technological process in practice, it is necessary to test it through simulation, theoretically determine the impact of the elastic deformations of TS (technological system) elements on machining accuracy, and ensure high machining precision by managing technological interactions.

The aim of the work is to analyze the influence of elastic deformations of TS elements on the accuracy of processing during parallel processing of internal conical surfaces in oil field grade couplings, using the control of dynamic technological relationships, to determine the directions for increasing accuracy and to test the process using simulation.

III. Methodology

The accuracy of parallel machining of surfaces in holes depends not only on the elastic deformations of the cantilever boring bar, but also on the accuracy of the mutual arrangement of the tools on it. During mechanical processing, the elastic deformations of the boring bar sections where the tool is located are different due to the effect of dynamic cutting forces. Although the tools are designed for the same cutting depth, their stability during machining is not guaranteed. Taking these different deformations into account in the dimensions of the tool positions on the boring bar ensures a reduction in the errors of sizes and shapes that occur on the under-thread surfaces and their negative impact on accuracy due to heredity in the threads.

The features of parallel machining developed for internal surfaces differ fundamentally from the features of parallel machining used in production for external surfaces. Since, machining of internal surfaces of the coupling in the last pass of the tool occurs as follows: first, the second tool sequentially machines the cylindrical and conical surfaces at the end of the coupling (Figure 1.1; in the figure: Lm- is the length of the coupling; φ- is the angle of inclination of the threaded conical

surfaces; φ_A and φ_B - are the main angles in plan of the cutting edges of the tools, in the middle and end of the boring bar, respectively; b- is the length of the cylindrical and conical surfaces at the end; a-is the length of the conical surfaces under the threads, taking into account the tool outlet;

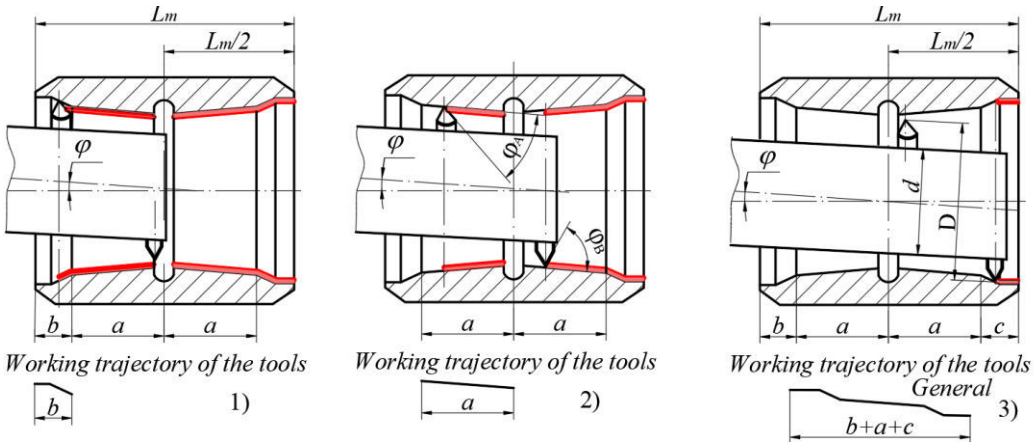


Figure 1: *Schemes of serial-parallel machining of internal surfaces of a coupling in one pass*

c - is the length of the conical and cylindrical surfaces of the coupling, taking into account the tool outlet; d- is the diameter of the boring bar; D- is the distance between the tools, in the direction perpendicular to the boring bar axis), then both cutting tools parallel machine differently directed conical surfaces under the threads (Figure 1.2). Finally, the first cutting tool is used to successively machine the transition conical and cylindrical surfaces in the inner part of the coupling (Figure 1.3). Thus, in this case, the cutting tools remove material from opposite, and from different sides, coaxial multidirectional conical and cylindrical surfaces. Therefore, the cutting force components acting on the cutting tool are directed opposite to each other, unlike traditional parallel processing schemes, and with well-organized operations even take the same values (Figure 2; the figure shows: P_z, P_y and P_x - vertical, radial and axial components of the cutting forces, respectively; Aa and Ba - installation dimensions of the second and first tools, respectively; l1 - axial distance between the

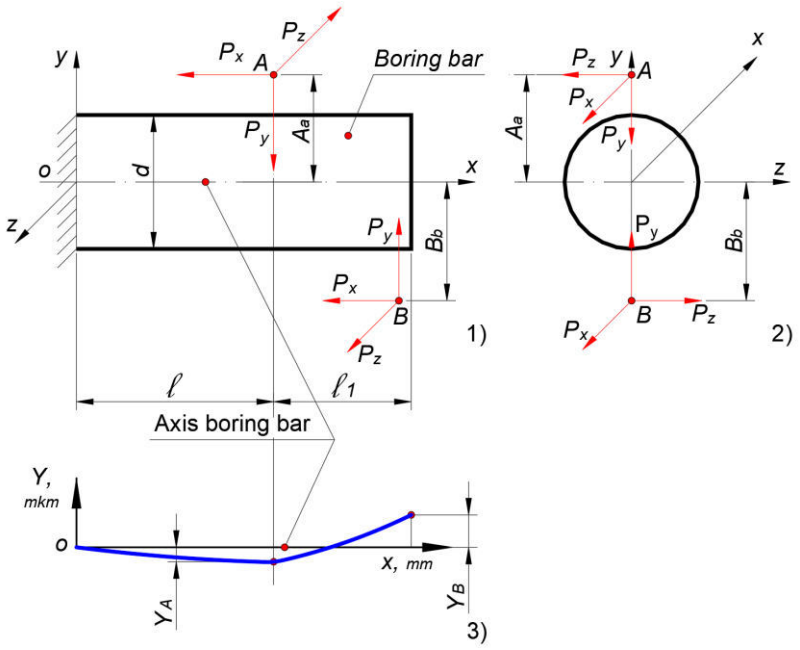


Figure 2: *Scheme of the system of cutting forces acting on the boring bar*

tools; l - distance from the second tool to the boring bar support). As a result, during processing, elastic deformations caused by different cutting force components on different sections of the second section of the boring bar occur in opposite directions, compared to the first section, and they compensate each other (Fig. 2.3). This ensures high processing accuracy.

The boring bar, and consequently the similar workpiece-workpiece system and the system of forces acting on the machine, is technologically very advantageous (Figures 2; 1 and 2). By controlling the geometric parameters of the cutting tools under given cutting conditions, in particular the main angles in the plan (φ_A and φ_B), it ensures both the optimal value of the cutting force components and favorable elastic deformations arising in different sections of the boring bar. This also facilitates self-adjustment of the boring bar to the most favorable cutting conditions.

To solve the stated problem, a coupling conforming to НКМ-60-ГОСТ 633-80 (according to the ANI standard) was selected as the research object. The choice of this size coupling is also related to the convenience of conducting experiments on natural samples under laboratory conditions. The cutting forces generated during the machining of conical surfaces and their components were determined analytically using a known methodology [2, 4, 7]. In this case, the cutting part of the tool is made of T15K6-grade hard alloy, while the boring bar material is 40XH steel with a hardness of HRCЭ 48-52. The cutting depth was set to t=3 mm, the feed rate to S=0.2 mm/rev, and the cutting speed to V=200 rev/min. Considering the special role of the principal cutting edge angles (φ_A and φ_B) in the formation of cutting forces, the components of the cutting force were determined and analyzed at various values of these angles ($\varphi_A=\varphi_B \Rightarrow 45^0; 60^0; 90^0$). The system of forces specific to the machining processes was determined (Figures 2.1 and 2), and based on the theories of "Material Strength" [18-21], the deformations ($\Delta=f(\varphi, P_z, P_y, P_x)$) of the boring bar along the direction of the machining dimensions of the tool's principal cutting edge angles ($45^0; 60^0; 90^0$) were determined, taking into account the mechanisms of their influence on the boring bar. The results were analyzed and generalized (Figure 2.3). According to the coupling design, the dimensions of the designed boring bar are as follows: A_a=30 mm; B_a=30 mm; ℓ_1=56 mm; ℓ=84 mm; d=45 mm; D=60 mm.

The deformations of the boring bar along the axis of its dimension were determined along the working length of the boring bar using the "EXCEL" software.

The technological operation, including the investigation of the elastic deformations of the boring bar and the stress state of its material, was simulated using the ANSYS software for all three variants provided in Figure 1. In the variant where the second cutter processes the cylindrical and conical surfaces at the output end of the coupling (Figure 1.1), the elastic deformation state of the boring bar is shown in Figure 3.

IV. Discussion

The system of forces affecting the boring bar during the parallel machining of surfaces in the processed machining scheme differs significantly from the force system formed during traditional multi-tool parallel machining, as described in the literature. This difference makes the process highly favorable for ensuring machining accuracy (Figure 1). Specifically, the fact that the cutting forces' three components acting on the boring bar and other elements of the technological system (TS) are

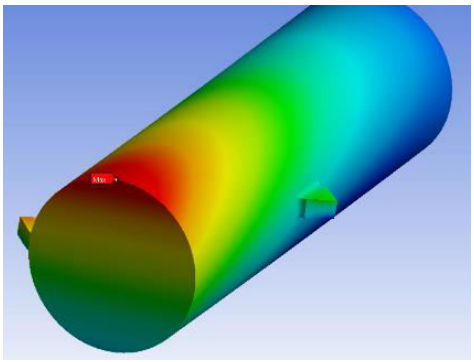


Figure 3: *Simulation of the elastic deformations of the boring bar*

oriented in opposite directions ensures a sharp reduction in elastic deformations on the XOY plane in various workpiece cross-sections, especially in the sections where the machining dimensions are formed (Figure 2).

It should be noted that, based on initial reports and studies, although the elastic deformations of the boring bar's axis under the influence of P_z forces may take relatively large values, their effect on machining accuracy is not technologically significant in this case, since they are oriented perpendicular to the machining dimensions.

The ratio of the forces P_z, P_y, and P_x in the system of forces during the parallel machining of surfaces, and consequently the bending behavior of the boring bar's axis, depends on the geometric parameters of the cutting tools, including the main angles in plan of the cutting edges of the tools (φ_A and φ_B). It is clear that from a technological point of view, the elastic deformations of the boring bar are important in the cross-sections where the vertices of the cutting tools are located.

Thus, the control of dynamic technological connections and the main angles in plan of the cutting edges of the tools during parallel processing of the internal surfaces of the coupling makes it possible to achieve a change in a favorable direction of the magnitude of the cutting force components and to ensure a reduction in the range of elastic deformations of the elements of the technological system and a reduction in their influence on the processing accuracy.

The dimensions of adjustment the tool setup determined taking into account the elastic deformations of axis of the boring bar in its sections, where the tool tips that form the machined surfaces are located (A_a and B_a). That is:

$$A_{af}=30+Y_a \text{ mm;} \qquad B_{af}=30+Y_b \text{ mm,}$$

Here, A_{af} and B_{af} are the adjustment dimensions for the position of the cutters in the boring bar.

- Y_a and Y_b are the elastic deformations of the boring bar's axis in the cross-sections where the cutters are located, respectively.

The adjustment tolerance is determined based on the required machining accuracy and technological capabilities.

The simulation of the elastic deformations of the boring bar using ANSYS software has confirmed the above-mentioned results.

V. Conclusions

1. The technology and technological measures for processing internal cylindrical and conical surfaces of oil field couplings are presented, providing sequential-parallel processing of its responsible surfaces with high accuracy and productivity,

2. A unique system of forces acting on the boring bar during parallel processing has been identified, in which all three corresponding components of the cutting forces are directed oppositely, which leads to a sharp decrease in the elastic deformations of the elements of the technological system,

3. By controlling the dynamic technological connections and the main angles in the plan of the

cutters during parallel processing, a change in the favorable direction of the magnitude of the components of the cutting forces is achieved, a decrease in the range of elastic deformations of the elements of the technological system and a decrease in their impact on the accuracy of processing are ensured.

References

[1] Bogatenkov, S.A., Sazonova, N.S., Yusubov, N.D. et al. Increasing the Productivity of Multitool Machining on Automated Lathes by Optimizing the Machining Plan. Russ. Engin. Res. 41, 1071–1074 (2021). https://doi.org/10.3103/S1068798X21110046

[2] DeGarmo's, Materials and Processes in Manufacturing, (2019). 13th Edition, J. T. Black, Ronald A. Kohser, Edition Wiley, ISBN: 978-1-119-49282-5, May 2019, 896.

[3] Rasulov Nariman Mogbil, Nadirov Ugurlu Muhammed, Abbasova Irada Aziz, Mammadov Arastun Salman, Shabiyev Elgun Tagi. (2020). Advanced Technology for Machining Tapered Threads of Pipe Coupling. American Journal of Engineering and Technology Management, 5(2), 35-40. https://doi.org/10.11648/j.ajetm.20200502.11

[4] Suslov A.G., Dalsky A.M. Scientific foundations of mechanical engineering technology, M.: Mashinostroenie, 2002. - 684 p. https://b.eruditor.link/file/895202/ (in Russian)

[5]. Yusubov, N., & Abbasova, H. (2020). Models for Machining Accuracy in Multi-Tool Adjustment. International Journal of Automotive and Mechanical Engineering, 17(3), 8067–8085. https://doi.org/10.15282/ijame.17.3.2020.01.0605)

[6]. Nariman Rasulov, Ugurlu Nadirov, Irada Abbasova, Improving of Machining Efficiency of Threads and Conical Surfaces of Diverse Directions by Managing Static Technological Relationships. / Advances in Science and Technology, ISSN: 1662-0356, Vol. 148, 2024, pp 97-102, https://doi.org/10.4028/p-4ozpgf

[7] Handbook of the technologist of the machine builder, Ed By: Dalsky A. M., Kosilova A. G., Meshcheryakova R. K. etc.: Mechanical Engineering, 2003. V. 1. 912 p. V. 2. 944 p.

[8] Lauro C.H., Brandão L.C., Filho S.M.R., Davim J.P. Quality in the machining: characteristics and techniques to obtain good results / In: Manufacturing Engineering: New Research. Ed. J.P. Davim. New York, Nova Publishers, 2014. P. 51–75, ISBN: 978-1-63463-378-9.

https://novapublishers.com/shop/manufacturing-engineering-new-research/

[9] Rasulov, N.M., Nadirov, U.M., Alekberov, M.Z. Generalized Assessment of Machined Surfaces Quality. Russ. Engin. Res. 40, 822-825 (2020), https://dio.org/10.3103/S1068798X20100202

[10] Rasulov N.M., Nadirov U.M., Alakbarov M. Z. Improving the efficiency of grinding teeth by copying with the control of dynamic technological connections. / SOCAR Proceedings. Special Issue 1 (2022) 029-035, DOI: 10.5510/OGP2022SI100697, http://dx.doi.org/10.5510/OGP2022SI100697

[11] Safarov D.T., Kondrashov A.G. Reduction of errors in the profile of the teeth of gears in the processes of gear milling with worm cutters. Fundamental and Applied Problems of Engineering and Technology. No. 3 (347) 2021, p. 63-72. (in Russian). DOI: 10.33979/2073-7408-2021-347-3-63-72

[12] V.I. Averchenkov, A.S. Vasiliev, M.L. Heifetz. Technological heredity at quality formation of parts manufactured. «Science intensive technologies in mechanical engineering», No10, 2018. No. 10 (88), p. 27-32, DOI: 10.30987/article_5bb4b1fa81a7f8.26650961. (in Russian).

https://naukaru.ru/ru/nauka/article/23024/view

[13] Rasulov N.M., Nadirov U.M. An approach to assessing the quality of manufacturing parts in device manufacturing // Scientific and Technical Bulletin of Information Technologies, Mechanics and Optics. 2019. Vol. 19. No. 4. P. 747–755. doi: 10.17586/2226-1494-2019-19-4-747-755 (in Russian)

[14] N.M. Rasulov, E.T. Shabiev, G.V. Damirova, M.Z. Alakbarov, Y.E. Huseynov, Increasing the Efficiency of Forming Complex Rotating Surfaces With the Controlling of Technological Connections. / Key Engineering Materials Vol. 979, april 2024, 55-62, Switzerland.

https://doi.org/10.4028/p-jc7OVU

[15] Aizuppe E.A., Polyachek D.N. Oil grade pipes. Volume 1 Drill pipes: Samara: LLC As Gard Publishing House, 2012, 284 p. ISBN 978-5-4259-0172-9. https://vdoc.pub/documents/1-5ceuebcmr420

[16] ISO 11960:2020(E). Petroleum and natural gas industries — Steel pipes for use as casing or tubing for wells. 2020, p. 267.

[17] Catalog of threaded connections of oil pipelines. OAO "Pipe Metallurgical Company". https://nekopipe.ru/downloadprices/tmk_katalog_rezbovih_soedineny_burilnie_trubi.pdf (in Russian)

[18] M.V. Pesin, V.F. Makarov, O.A. Khalturin, etc. Increasing the reliability of oil industry and drilling equipment by means of complex design and technological support for the manufacture and assembly of threaded connections. Herald of PNIPU. Mechanical engineering, materials science, 3 (2021) 91–100, DOI: 10.15593/2224-9877/2021.3.11 (in Russian).

[19] Beer, F.P., Johnston, E.R., DeWolf, J.T., Mazurek, D.F. Mechanics of Materials, McGraw-Hill Education, New York, 2014, p. 992.

https://students.aiu.edu/submissions/profiles/resources/onlineBook/Z5D4s5_Mechanics_of_Materials_7th.pdf

[20] Benediktov, Yu.A., Kharlamov, V.A. Fundamentals of the Mechanics of Materials and Structures, Bauman Moscow State Technical University, Moscow, 2016, p. 600. https://sme.njust.edu.cn/_upload/article/files/4a/fe/99e755274de8850b355ad4a7f6ed/8eb0d14b-8cf1-42fe-87d2-18095da6c68d.pdf

[21] Feodosyev V.I., Strength of Materials. - 10th ed., revised and enlarged. - M.: Publishing house of Bauman Moscow State Technical University, 1999. - 592 p.

https://books.google.az/books/about/Strength_of_Materials.html?id=gUs5zgEACAAJ&redir_esc=y

ISSUES OF INCREASING THE EFFICIENCY OF CYLINDRICAL GEAR GRINDING USING COPYING METHODS THROUGH A SYSTEMATIC APPROACH

Nariman Rasulov[1], Arastun Mammadov[2], Mursal Alakbarov[3],

Elgun Shabiyev[1], Yusif Huseynov[1]

•

[1]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan
[2]Dean of Faculty of German Engineering, Azerbaijan Technical University, Baku, Azerbaijan
[3]Department of Oil and Gas Production, BP Exploration (Caspian Sea) Limited, Baku, Azerbaijan

nariman.rasulov@aztu.edu.az, arastun.mammadov@aztu.edu.az, mursal.alakbarov82@gmail.com, elgun@aztu.edu.az, yusif.huseynov@aztu.edu.az

Abstract

The paper presents a system analysis of gear grinding with copying of cylindrical gears, subsystems associated with the gear grinding system, as well as their inputs and outputs connections; both direct and indirect ways of increasing the efficiency of gear grinding are identified, based on the management of connections by hierarchical sequence vertically and by sources of quality indicators; includes the results of an indirect increase in efficiency due to the adoption of a reasonable value for the allowance for grinding, a direct increase due to a reduction in the number of working passes while ensuring the required quality during gear grinding and also the use of a methodology for forming threads by plastic deformation; the developed methodology is recommended for use in solving similar problems when forming surfaces using other methods.

Keywords: Copying, gear grinding, system analysis, efficiency, increasing, parameter, indirect, direct, methode

I. Introduction

There is no area of industry where gear transmissions are not used. At the same time, the demand for various technical devices with high precision and operational performance continuously increases in society. To improve the quality indicators of the working surfaces of gear teeth, as well as the operational characteristics of their engagement, tooth grinding by the copying method is widely used [1-5].

It is obvious that when forming any surface by the method of material removal, the expected technological support and results depend on the parameters of its previous surfaces, functionally related to it from both the design and technological points of view [1, 6-9]. Therefore, the highest technological and economic results of the process of grinding teeth by the copying method can be achieved by implementing various innovative measures based on the system analysis of this operation. System analysis of tooth grinding is the most rational methodological mechanism for identifying and managing both direct technological measures characteristic of this operation and

indirect measures related to previous operations. In this context, increasing the quality and efficiency of tooth grinding by the copying method stands out for its relevance.

II. Statement of the Problem

Each technological operation, including tooth grinding with copying (TGC), special elements related to its implementation, other structural and technological elements related to it and preventing it, mechanisms of action and a complex of connections between them form a complex system [10-14]. Thus, in TGC, the material of the gear and the methods, tools, and workpiece used in the formation of its tooth profile, along with the processing of the tooth meshing surfaces, and the prevention of wear, are all part of a highly complex system, this includes the processing of rotational, keyseats, or spline surfaces, as well as thermal treatments, and involves a variety of constructive, technological, static, kinematic, and dynamic relationships, elements, and transformations. In the context of TGC, we can say that the upper system consists of the main operation, while the associated and preventive components form the subsystems. We can say that the upper system of TGC, and the associated or preventive components, are subsystems. Each technological parameter of the tooth grinding output is a manifestation of the transformations occurring in every component and structural element of the system. Therefore, to effectively address issues related to the quality and productivity of TGC, a systematic approach is required, considering all subsystems involved in ensuring its performance (Figure 1).

The purpose of the work is to conduct a systematic analysis of the formation of technological output parameters of the tooth grinding operation with copying, to develop and test direct and indirect methods for increasing the efficiency of the operation.

III. Methodology

A hierarchical diagram was developed based on the decomposition of the TGC technical production system, ensuring its integrity and emergent properties (Figure 1), when selecting key issues for increasing efficiency using hierarchical relationships, the system's constructiveness, orientation of component elements and their interrelationships, technological factors, and goal-directed characteristics were considered. These were studied at the micro level, taking into account the internal and interrelationships between lower and upper-level components (systems) and their interactions.

By means of system analysis it was established that the increase of efficiency of the TGC can be carried out both by direct and indirect measures (Figure 1). The essence of the indirect increasing of efficiency lies in the management of the impact on the current gear grinding process (referred to as the upper system) based on the inheritance principle of appropriate output parameters and technological results obtained in the precursor subsystems-technological processes for the gear wheel-workpliece to be ground. Thus, the positive results obtained in the upper system are only the cause and effect of the design and technological measures adopted in the subsystems.

It should be noted that in the existing literature [1, 6, 8], the impact of the quality of pre-processed surfaces on the quality parameters of the surface formed during the current operation or pass is associated with the inheritance principle and evaluated by a correction factor. Therefore, in this case, the management of the subsystems in the upper system, which is the TGC, to increase its efficiency through transformations in these subsystems should be systematically investigated, compared with direct efficiency improvement possibilities, evaluated, and a conclusion should be drawn.

The essence of directly increasing the efficiency of gear grinding for shaped surfaces involves increasing efficiency through changes, transformations, and management that are directly related to the upper system and executed within its elements. This includes increasing grinding quality and

productivity by managing the gear grinding technological process, along with the technical and technological measures related to it. Thus, system analysis determines that direct improvement of TGC efficiency in gear grinding involves enhancing parameters such as: - accuracy of the teeth, surface roughness, and surface layer quality, as well as productivity, through measures specifically related to gear grinding. This can be achieved by automating and mechanizing the technological system's component elements. The systematic analysis of increaing gear grinding efficiency E(↑) can be carried out in two directions:

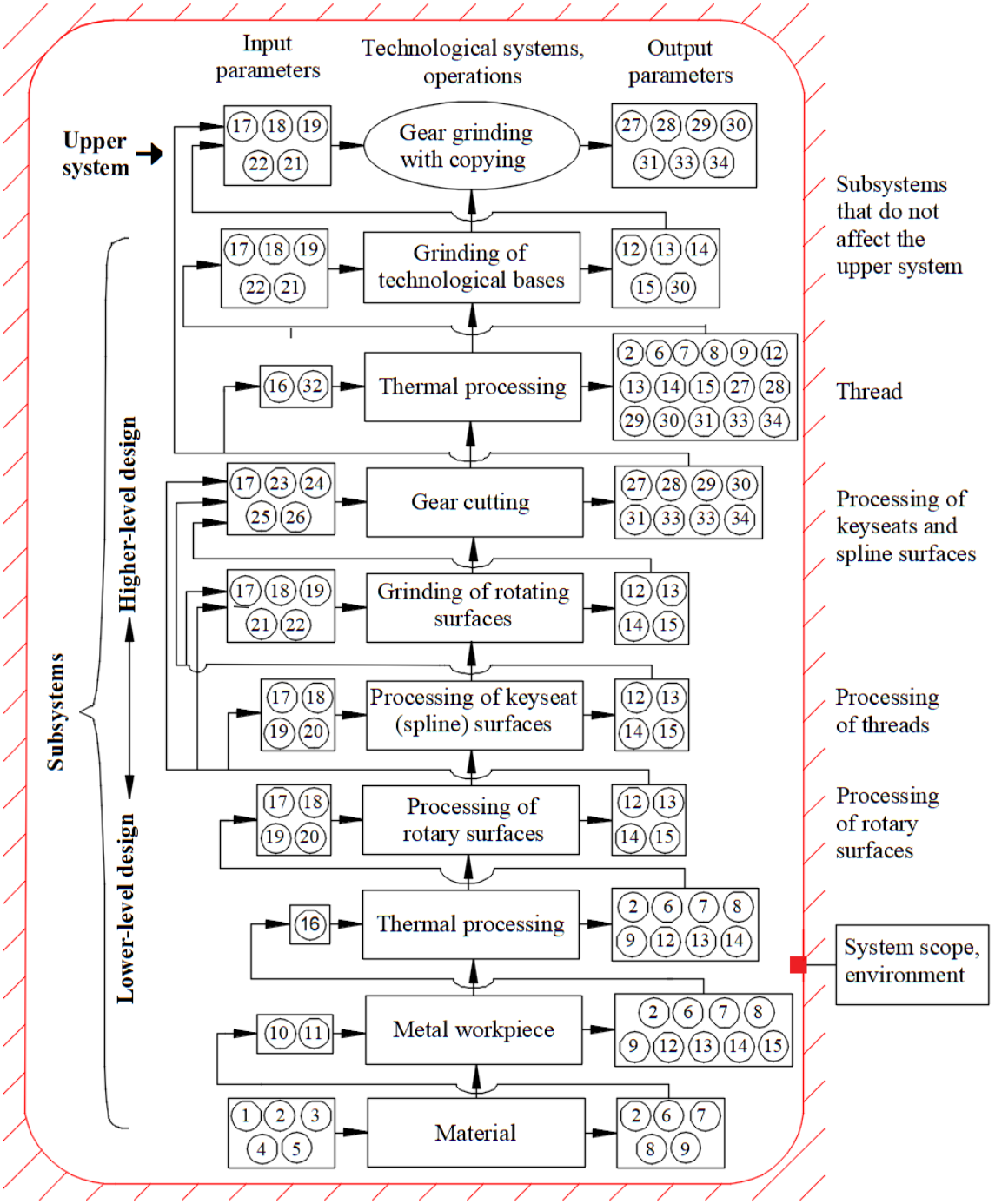


Figure 1: *System analysis of gear grinding in a hierarchical diagram*

1- Raw material, 2- Chemical elements (composition), 3- Formation method, 4- Forming conditions, 5- Environment, 6- Material structure, 7- Granularity, 8- Hardness of the material, 9- Material homogeneity, 10- Metal workpiece method, 11- Metal workpiece condition, 12- Dimensional and

shape accuracy, 13- Accuracy of relative surface positions, 14- Surface roughness quality, 15- Surface layer quality, 16- Thermal processing conditions, 17- Layout scheme, 18- Processing method, 19- Cutting mode elements, 20- Cutting conditions, 21- Grinding wheel, 22- Grinding conditions, 23- Tooth cutting method, 24- Tooth cutting tool, 25- Tooth cutting mode, 26- Tooth cutting conditions, 27- Accuracy of the parameters of the diametrical dimensions, 28- Accuracy of parameters throughout the circumference, 29- Accuracy of the profile, 30- Accuracy of the circle relative to the base positioning surface (eccentricity), 31- Allowance for tooth grinding, 32- Thermal processing method, 33- Surface roughness quality of the profile, 34- Surface layer quality of the profile.

- Increasing of efficiency $E(\uparrow)$ through the management of possible sources for increasing quality or productivity, or both can be expressed by the formula:

$$E(\uparrow) => f(M_i, S_i, C_i, O_i) = \begin{cases} < M_1, M_2, M_3, M_4 > \\ < S_1, S_2, S_3, S_4 > \\ < C_1, C_2, C_3, C_4 > \\ < O_1, O_2, O_3, O_4 > \end{cases} \quad (1)$$

this can be expressed as follows. Here, i- represents the sequence number of the efficiency improvement sources, which may vary for different sources:

$M_1, \dots, M_4$– quality of the material,

$S_1, \dots, S_4$ – static technological relationships,

$C_1, \dots, C_4$ – relationships related to changes in shape, size, etc. of the design,

$O_1, \dots, O_4$ – auxiliary motion relationships that serve to change the form are efficiency increasing through the management of these elements at various stages (subsystems).

- Increasing of efficiency $E(\uparrow)$ according to the sequence of the hierarchical formation of the tooth grinding system. In this case:

$$E(\uparrow) => F_j(f(M_i, S_i, C_i, O_i)) = \begin{cases} < M_1, S_1, C_1, O_1 > \\ < M_2, S_2, C_2, O_2 > \\ < M_3, S_3, C_3, O_3 > \\ < M_4, S_4, C_4, O_4 > \end{cases} \quad (2)$$

this can be expressed as follows. Here, j- represents the sequence of hierarchical formation, which may vary for different parameters.

In the last statement, the distinction of the lower stage of teeth cutting is related to its special significance in the formation of the grinding substage and the multi-parameter tooth surfaces. In cases where the latter expressions are required, each subsystem can be represented separately and with a larger number of signifiers.

As a result of the conducted research and the system analysis of the process, the following directions for increasing the TGC efficiency are accepted:

- indirect through the provision of minimum and maximum allowances for the grinding process, based on both theoretical and experimental foundations accepted in machine engineering;
- direct by reducing the number of passes through the reduction of the actual cutting depth in tooth grinding, using the method developed at Azerbaijan Technical University (AzTU) [3, 14].

IV. Discussion

The following results of the TGC efficiency increasing experiments, based on the application of the developed methodology and the expressions (1) and (2), are discussed:

1. Ensuring the theoretical values of allowance factors for the grinding of teeth of cylindrical gears. The solution to the problem is based on the methodology accepted in machine engineering, science, and practice and is solved by considering the shape and alignment errors formed in the initial cutting stage (gear cutting), the alignment errors in the grinding area of the gear teeth, and the mechanisms affecting the allowance [2, 14]. However, the complexity of the problem arises from the fact that the parameters (errors) of the gears obtained from the cutting process are multi-parameter and repetitive according to the relevant standards. It is necessary to select and consider those parameters that have a decisive effect on the allowance, along with other analogous

parameters. To do this, we categorize the standard parameter sets of the gear wheels in terms of their relationship with the allowance into four groups: (parameters with generalized relationships with the allowance; repetitive, complementary parameters specific to each gear for cutting and grinding, related to the allowance, mainly independent and random errors; parameters directly related to the formation of the allowance, non-repetitive individual indicators; parameters unrelated to the allowance).

After gear cutting and tooth grinding, analytical expressions have been derived to determine the minimum and maximum values of the grinding allowance, taking into account the characteristics of the formation of the teeth's quality parameters and the requirements for the quality of the teeth according to the standards.

Using the expressions developed and proposed for the grinding allowance (which are not presented here as they have been submitted to another journal), the grinding allowance for the teeth obtained with a 7th accuracy grade through gear cutting was calculated as follows: $2Z_{mini} = 0.28\ mm;\ 2Z_{maxi} = 0.31$ mm;

In experiments conducted at the Sumqayit Technology Park, the values $2Z_{mini} = 0.28\ mm;$ $2Z_{maxi} = 0.31$ mm; were accepted, and it was determined that the grinding quality has been ensured during the research.

2. Reducing the number of passes by decreasing the actual cutting depth using the method developed at AzTU. In gears with a modulus of m=4 mm, number of teeth z=40, a length of 25 mm, and made of steel 40XH, during the grinding of approximately vertically arranged teeth, it was determined that the actual total cutting depth was 1.75 times greater than the allowance (0.3 mm) at the top of the tooth, and 2.04 times greater at the start of the involute. When grinding the inclined tooth number 5, the actual total cutting depth decreases and is 1.36 and 1.47 times greater, respectively. Thus, grinding the of the inclined tooth ensures a reduction of the actual total cutting depth by $(2.04 - 1.47)Z = 0.57 \cdot 0.3 = 0.2\ mm$. In rough working passes, the radial feed is 0.1 mm. This allows for a reduction in the number of passes by one.

The increasing of the tooth grinding efficiency has been tested through experiments in two directions: grinding inclined teeth while maintaining the number of passes, and grinding inclined teeth while reducing the number of passes by one.

The grinding was performed on a Gleason Pfauter P400G model machine using the methodology accepted by the authors [12]. The three-pass grinding process consisted of two rough passes and one finish pass, with radial feeds of 0.18 mm for the first pass, 0.09 mm for the second pass, and 0.03 mm for the third pass.

The measurements and analysis showed that the accuracy parameters of the grinded surfaces, including tooth thickness, pitch, total normal length, diameter of the pitch circle, surface roughness parameters, microhardness, etc., were within the requirements for the part profiles in all cases and corresponded to the 7C accuracy grade according to GOST 1643-81.

For the tooth thicknesses, empirical distribution curves, constructed using the arithmetic average values obtained from at least three different measurements, are presented as m=F(fp) (where m is the number of measurements in each group on the empirical curve, plotted along the ordinate axis, and fp is the tooth thickness, plotted along the abscissa axis). The empirical distribution curves are shown in Figure 2. In the figure, the empirical distribution curves and histograms for the tooth thicknesses along the dividing circle of the grinded teeth are shown: with traditional grinding (curve 1), with the proposed method in a four-pass grinding process (curve 2), and with the proposed method in a three-pass grinding process (curve 3). The histograms and scatter centers of these curves have been marked as a1, a2 and a3, with their coordinates indicated.

Focusing only on the proposed method with a four-pass grinding process (curve 2), it can be noted that the smallest deviation in tooth thickness in the grinding process was $E_s = 29\ mkm,$ and the scatter area of the deviations in tooth thickness has been accepted within 52 mkm. For the conventional method, the smallest deviation in tooth thickness during grinding was $E_s = 32\ mkm,$ and the scatter area of the deviations in tooth thickness was 56 mkm. Overall, the closeness of the

results is attributed to the high precision of the machine tool and the minimization of external factors affecting the results during the experiments. Furthermore, the mathematical expectation of the pitch a in the proposed grinding methods (a3 = 6.221 mm for four-pass; ; a2 = mm for three-pass) is closer to the high precision compared to the traditional grinding method (a1 = 6.220 mm). It seems that these positive aspects are related to the fact that the cutting conditions of the grinding process with the proposed method are more stable compared to the traditional method.

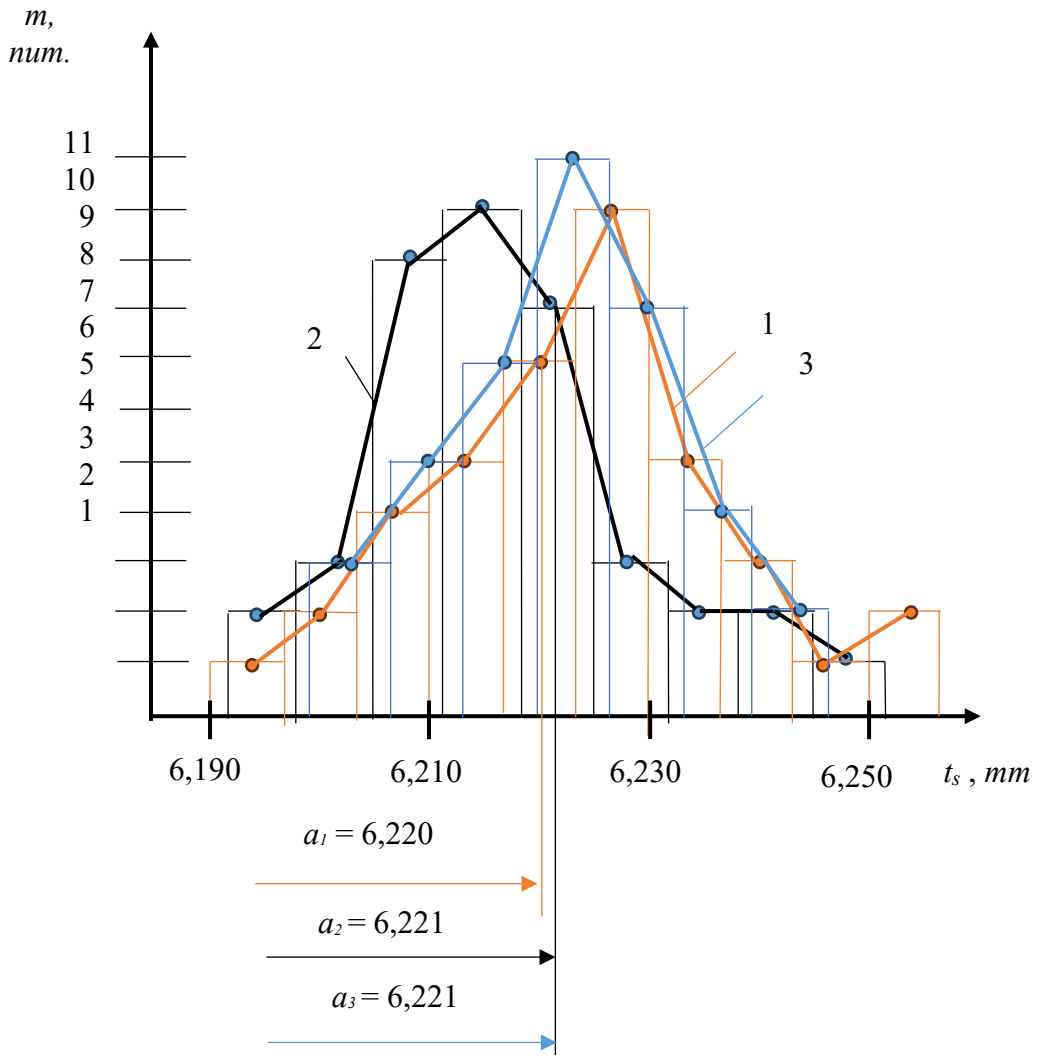


Figure 2: *Empirical curves of tooth thickness dispersion along the pitch circle for ground teeth: 1) traditional method, 2) proposed method (PM), four-pass processing, 3) PM, three-pass processing*

Thus, with the system analysis of the TGC process and the proposed systematic approach, the increasing of tooth grinding efficiency and the reduction of the tooth grinding time standard are achieved.

3. Increasing the efficiency of forming by plastic deformation of the material. By applying the methodology presented above, the processes of thread formation by plastic deformation of the material were studied, and it was determined that in the threading process with tangential rolling heads, the tool-workpiece contact area is large, resulting in higher rolling forces. As a result, problems arise in the threading of relatively long threads, as well as in the threading of pipes, when using tangential feed [15-17]. Through a systematic approach, the tool-workpiece contact area, and thus the rolling force, in tangential threading was controlled, which allows for increasing the efficiency of tangential threading and managing the technological capabilities of the tools, a method

and means for tangential threading, enabling control of these factors, has been developed at the patent level.

V. Conclusions

1. A system analysis of tooth grinding with copying has been conducted, a methodology for increasing its efficiency with a systematic approach has been developed, and the sources of efficiency increasing and their management in the hierarchical formation sequence of the tooth grinding system have been determined, the directions of increasing efficiency in direct and indirect ways have been determined.

2. When grinding a tooth located inclinely in the grinding zone, a significant reduction in the actual cutting depth is ensured, when grinding the tooth number 5, located inclinely on gears with a module of 4 mm, a number of teeth of 40, a length of 25 mm and steel material 40XH, when four working passes are replaced by three passes, the quality of gear grinding is ensured by all parameters as an analogue of four-pass grinding, efficiency in terms of productivity increases.

3. The increasing of tooth grinding efficiency by in directly managing the share of pre-grinding allowance, as well as enhancement of the efficiency of threads formation using tangential rolling heads through the control of actual rolling force, is presented.

References

[1] DeGarmo's, Materials and Processes in Manufacturing, (2019). 13th Edition, J. T. Black, Ronald A. Kohser, Edition Wiley, ISBN: 978-1-119-49282-5, May 2019, 896. https://www.wiley.com/enus/DeGarmo's+Materials+and+Processes+in+Manufacturing%2C+13th+Edition-p-9781119492825

[2] Lishchenko N.V. Optimization of profile gear grinding on a CNC machine and an allowance measurement system / N.V. Lishchenko, V.P. Larshin, V.V. Nezhebovsky // High technologies in mechanical engineering: zb. Sci. fuck off. - Kharkiv, NTU "KhPI", 2016. - Vip.1 (26). — P. 50 — 61 (in Russian). https://repository.kpi.kharkov.ua/handle/KhPI-Press/25662

[3] Nariman Rasulov, Elgun Shabiyev, Gulbaniz Damirova, Mursal Alakbarov, Yusif Huseynov. Increasing the Efficiency of Forming Complex Rotating Surfaces with the Controlling of Technological Connections, Key Engineering Materials, ISSN: 1662-9795, 2024, Vol. 979, pp 55-62, https://doi.org/10.4028/p-jc7ovu

[4] Rasulov, N.M., Alekberov, M.Z., Nadirov, U.M. More Efficient Copy Grinding of Complex Surface. Russ. Engin. Res. 41, 829-831 (2021).
https://link.springer.com/article/10.3103/S1068798X21090227

[5] Rikard Hjelm, Linus Everlid, Ellen Bergseth, Florian Reinle, Boris Brodmann, Minghui Tu, Lucas Bard, Jens Wahlström. A multi-perspective method for gear efficiency and contact analysis. Results in Engineering Vol. 20, December 2023, 101582. https://doi.org/10.1016/j.rineng.2023.101582

[6] V.I. Averchenkov, A.S. Vasiliev, M.L. Heifetz. Technological heredity at quality formation of parts manufactured. «Science intensive technologies in mechanical engineering», No10, 2018. No. 10 (88), p. 27-32, DOI: 10.30987/article_5bb4b1fa81a7f8.26650961.
https://naukaru.ru/ru/nauka/article/23024/view

[7] Rasulov, N.M., Nadirov, U.M., Alekberov, M.Z. Generalized Assessment of Machined Surfaces Quality. Russ. Engin. Res. 40, 822-825 (2020).
https://link.springer.com/article/10.3103/S1068798X20100202

[8] Safarov D.T., Kondrashov A.G. Reduction of errors in the profile of the teeth of gears in the processes of gear milling with worm cutters. Fundamental and Applied Problems of Engineering and Technology. No. 3 (347) 2021, p. 63-72. D.T. Safarov/ (in Russian). DOI: 10.33979/2073-7408-2021-347-3-63-72. https://www.researchgate.net/publication/353527897

[9] Shipulin, L.V., Yusubov, N.D., Frolov, A.A. (2023). Study of the Microrelief Obtained During Single Abrasive Grain Cutting. In: Radionov, A.A., Gasiyarov, V.R. (eds) Proceedings of the 8th International Conference on Industrial Engineering. ICIE 2022. Lecture Notes in Mechanical Engineering. Springer, Cham. https://doi.org/10.1007/978-3-031-14125-6_97

[10] Alan Dennis, Barbara Haley Wixom, Roberta M. Roth, Systems Analysis and Design, John Wiley & Sons, 2014 M11 11 - 448 p.

https://books.google.az/books/about/Systems_Analysis_and_Design.html?id=PxZ6oAEACAAJ&redir_esc=y

[11] Mahbubur Rahman Syed, Sharifun Nessa Syed, Handbook of Research on Modern Systems Analysis and Design Technologies and Applications, DOI: 10.4018/978-1-59904-887-1, 2009, 698 p. https://www.igi-global.com/book/handbook-research-modern-systems-analysis/490

[12] N.M. Rasulov, U.M. Nadirov, M. Z. Alakbarov, Improving the efficiency of grinding teeth by copying with the control of dynamic technological connections. / SOCAR Proceedings. Special Issue 1 (2022) 029-035. DOI: 10.5510/OGP2022SI100697. https://proceedings.socar.az/en/journal/85

[13] Dalsky A.M., Bazrov B.M., Vasiliev A.S. and others - Technological heredity in mechanical engineering // Ed. A.M. Dalsky. Scientific publication. - M.: MAI Publishing House, 2000. - 364 p. https://disk.yandex.ru/i/GlqimQ

[14] N.M. Rasulov, E.T. Shabiev. Increasing the efficiency of grinding gear teeth using the copying method based on controlling the depth of cut. News of Mechanical Engineering Universities, MSTU named after N. E. Bauman, 2017, No. 2, 90-97. DOI 10.18698/0536-1044-2017-2-90- 97. (in Russian). https://izvuzmash.bmstu.ru/eng/catalog/tpmachines/hidden/1398.html

[15] Afonin A.N. Increasing the efficiency of thread rolling. Auto. dis. on sois. uch. Art. Doctor of Technical Sciences according to special 02/05/07., Orel-2010.-34 p.
https://rusneb.ru/catalog/000199_000009_004610263/?ysclid=lqxxwjb1bl615980626

[16] N. M. Rasulov, G. V. Damirova, İ. A. Abbasova, Y. E. Huseynov. Improving the efficiency of thread rolling with management of technological connections. / SOCAR Proceedings. Special Issue 1 (2022) 011-015. DOI: 10.5510/OGP2022SI100694. https://proceedings.socar.az/en/journal/85

[17] S. Zhang, S. Fan, Q. Wang, et al. Deformation characteristics of self-infeed rolling process for thread shaft. Int J Adv Manuf Technol, 103 (5-8) (2019), pp. 2941-2951; https://link.springer.com/article/10.1007/s00170-019-03677-3

APPLICATION AREAS OF CURVES AND SURFACES IN ENGINEERING

Ahmed Imanov[1], Zakir Galandarov[1]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan
ehmed.imanov@aztu.edu.az, zakir.qelenderov@aztu.edu.az

Abstract

The general classification of curves widely used in constructing surfaces for rapid movements along curved surfaces in engineering is presented. Key parameters ensuring the smoothness of curved lines are analyzed. The potential positive or negative effects of the curvature of these lines on moving objects, profiled surfaces, and similar applications are examined. The study considers the fields of application of curved lines and surfaces in engineering. Surfaces obtained using the formula representing the curved line, resolved as an inverse problem based on curvature, are investigated. The formula derived from curvature and its modified version are applied in analyzing the relationship between smoothness and potential energy, further confirming the accuracy of this connection. Types of surfaces are shown, and examples are provided of surfaces generated based on guiding curves and frames. The efficiency of surface modeling using computer technologies is also examined.

Keywords: curvature, torsion, centrifugal force, speed, main road, transition curve, circle, tangent, transcendental.

I. Introduction

In descriptive geometry, a line can be understood as the trajectory of a moving point's successive positions. If the moving point changes direction during its motion, the resulting trajectory forms a curved line. Curved lines can also be derived by other methods. Curved lines whose points lie on the same plane are called plane curves, while those that do not belong to a single plane are known as space curves. Curves are divided into two groups: algebraic and transcendental. The maximum degree of an algebraic equation defines the order of a curve. For plane curves, the geometric order is determined by its maximum intersections with a straight line, while for space curves, it is defined by the maximum intersections with a plane. The direction at any point on a curve is defined by its tangent. Curved surfaces are primarily obtained from the trajectory traced by a line, known as the generating line, as it moves along another line, called the guiding line.

Reference [1] examined the general application areas of functional curves and surfaces used in engineering, highlighting the importance of smoothness and curvature of curves on profiled surfaces. Information was provided on the application of computer technologies in 2D and 3D modeling of curves, surfaces, aerodynamic surfaces, and technical design. In [2], the research generalized curves with monotonically varying curvature, known as super spirals. In [3], an expert evaluation was conducted on the aesthetic quality of curves with varying smoothness parameters. However, these studies did not explore the application of curvature as an inverse problem in deriving mathematical expressions for curves. Taking this into account, the presented article

extensively addresses the application of curvature as an inverse problem in the formation of curves and surfaces commonly used in engineering.

Nikolaos Eliou et al. [4] constructed the transition section of roads using a Symmetrically Projected Transition Curve (SPTC), derived from comparing clothoid and cubic curves. In the formulation of the SPTC, trigonometric expressions were approximated by polynomials. Esveld C. [5] applied cubic curves to railway tracks, while Kasper H. et al. [6] used clothoid curves for designing transition lines in highways. In [7], the application of various types of curves (clothoid, cubic curves, and Bernoulli lemniscate) to transition curves was discussed, with an attempt to justify their use. E.A. Gavrilenko et al. [8] proposed a classification of curves based on the dynamics of changes in their differential-geometric properties.

As is known, the movement of a point changing position along a curved line is related to the variation of two quantities:

- ds represents the displacement, which is the distance traveled from the initial position;
- dφ, the angle of rotation of the tangent relative to the initial position (Figure 1).

One of the key parameters of a curved line is the curvature coefficient (or simply curvature). Curvature is defined as the inverse of the radius of curvature at a given point on the curve and represents the ratio of the angle of rotation of the tangent to the distance traveled by the point. It is expressed by the differential equation as follows:

$$k = \frac{d\varphi}{ds} \tag{1}$$

For example, the curvature of a straight line is zero at all points.

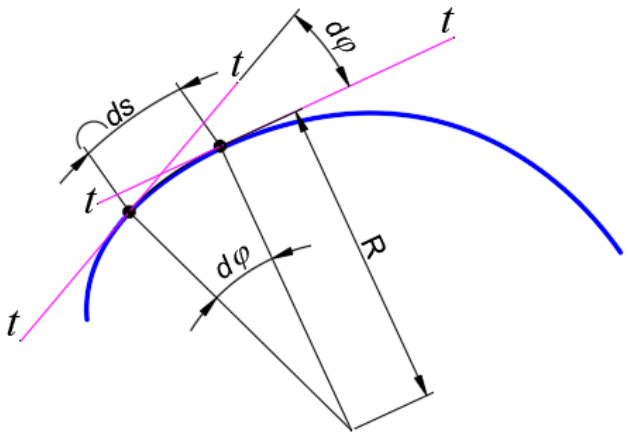


Figure 1: *Key Parameters of a Curved Line*

Application Areas of Curved Lines in Engineering. Curved lines used in engineering, whether in a plane or in space, express the functional characteristics of any object. Curved lines have found extensive applications during the design phase in engineering. Examples include:The profiling of blades' surfaces in turbo engines to enhance quality performance indicators.

- The design of transition curves that ensure comfortable and safe passageways on main roads, with the condition of providing maximum smoothness.
- The profiling of surfaces of pushers in distributors that ensure a gentle and impact-free motion of the valves that open and close timely in internal combustion engines.
- The profiling of surfaces of the stator to ensure that the plates located on the rotor of sliding (shutter) pumps move in contact with and without separation from the stator, among others.

Such curved lines can generally be classified as local convex (with negative curvature), local

concave (with positive curvature), and transition points (where the curvature changes). In space curves, torsion can also be added. Without considering the specific characteristics of the designed surfaces, the primary requirement imposed on all engineering curves is their smoothness. Smoothness refers to the differentiability of a function or geometric figure (curved line, surface, etc.) at all points in a given section. Different projects utilize curved lines with varying degrees of smoothness. For example:

- In classical design on main roads, clothoid curves are most commonly used for constructing transition curves, as their smoothness is ensured by a second-order derivative.
- In the profiling of the surfaces of pushers in distributors in internal combustion engines, the smoothness is generally maintained at a level not lower than three, which is why smooth curves that allow for derivatives of the third order are utilized.
- In the design of space curves, the smoothness of the curve must be of third order to ensure the continuity of the function's torsion.
- Regardless of the application area of the curved lines, higher-order smooth curves, such as transcendental curves, are also used for their superior smoothness.

One of the key parameters that ensures the smoothness of curved lines is the minimal number or complete absence of extremum points in the curvature graph over a given interval. For example:

$$y(t) = \frac{C_2}{p} \cdot \ell n\left(sin(p \cdot t) + \sqrt{B^2 - cos^2(p \cdot t)}\right) + C_1 \quad (2)$$

Here $C_1, \quad C_2, B, \quad p, a$ -is constant and $B = \frac{p}{a}$. Expression (2) is derived from the specific solution of the differential equation of curvature in Cartesian coordinates, with an additional constant included [9]. Expression (2) is a transcendental equation, and the smoothness is considered high. This equation can be used in the construction of all the planar engineering curves mentioned above. When the boundary conditions used to determine the constants in expression (2) are chosen correctly, there are no extremum points in the curvature graph, or at worst, there may be only one. The presence of excessive extremum points in curvature, for instance, in profiled technical surfaces and the design of objects, can lead to the following negative consequences:

- In a cam mechanism, it can cause premature wear of the cam and pusher due to the impact of the pusher.
- In aerodynamically profiled surfaces (e.g., blades), the flow of working fluids can become turbulent due to non-smooth surface irregularities, increasing the drag on the blade and causing undesirable local separations in the working fluid.
- On main roads, it can lead to hazardous movements and unnecessary braking or acceleration, increasing energy consumption during transition curves (10).
- It may cause premature wear of the stators and plates in sliding pumps.
- In computer graphics and CAD systems, it can result in incorrect visual perception of objects.

One of the main conditions imposed on profiled aerodynamic surfaces is that the maximum curvature should be minimized. If the substitution from $C_2 = 1$ [10,11] is applied in expression (2), then the curvature is defined by a sinusoidal curve. This ensures that the curvature is considered as a boundary condition in advance. That is, the minimum of the maximum value of curvature is directly used in determining the constants in the equation. For example, when a transition curve is connected to a circular road with a given radius on main roads, the radius of curvature at the endpoint of the transition curve must equal the radius of the circular road; otherwise, there will be an undesirable jump in the centripetal force.

Figure 2 (dimensions are given in meters), expression of k_2 curvation (2), k_1 is a curvation which is defined for the expression $C_2 = 1$. A seen from curvek_2, to ensure safe movement, the rate of change

of curvature is low at the entrance of the road (at the beginning) and increases as it approaches the exit.

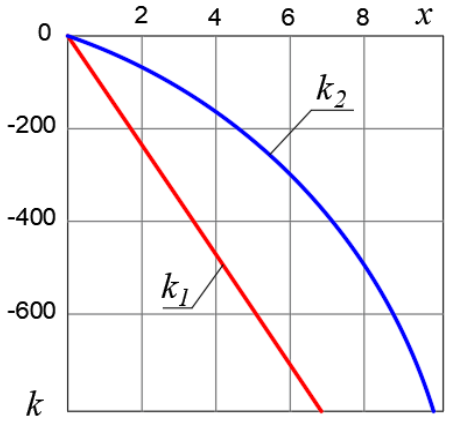


Figure 2: *Dependence of Curvature on the x Coordinate*

As is known, the smoothness of a curved line is directly related to its potential energy. The goal of selecting curves with low potential energy in high-speed curved trajectories is to ensure that the moving object behaves like an elastic body. It is well understood that less work is required to deform an elastic medium along a flow line with lower potential energy. The motion of an object along a convex curved trajectory with low potential energy will require relatively less work when taking friction into account.

The curved line with minimum potential energy from the given lines is referred to as elastic. The elasticity of the curved line defined by expression (2) can be expressed as follows:

$$E_{eyx} = \int_{l_1}^{l_2} k^2(s)\, ds = \int_{x_1}^{x_2} k^2(x) \cdot \sqrt{1 + \frac{{C^2}_2 \cos^2(px)}{\left(\frac{p}{a}\right)^2 - \cos^2(px)}}\, dx \tag{3}$$

where, E_{eyx}-is the potential energy of the curved line., x_1, x_2, ℓ_2, ℓ_2 beginning of the curve and its endpoints (interval), k - the curvature coefficient, defined by (4):

$$k(x) = \frac{-C_2\left(\frac{p}{a}\right)^2 \cdot p \sin(px)}{\sqrt{\left(\left(\frac{p}{a}\right)^2 + \cos^2(px)\left({C_2}^2 - 1\right)\right)^3}} \tag{4}$$

(3) the potential energy of the curved line obtained by the expression.
In expression (3) $C_2 = 1$ (then from (4) $k(x) = -a\, sin(px)$) is approximately 10% greater than the energy of the curved line obtained for the condition.

In addition to the properties that govern the quality of the smoothness of the aforementioned curve, there are also parameters that determine its technical aesthetics. When the surfaces present in real life are expressed mathematically, they tend to be more aesthetically pleasing (for example, the contour of a bird's wing). When modeling aesthetically pleasing surfaces, curves known as ℓog - aesthetic curves, which have a constant curvature represented as a straight line on a logarithmic scale, are widely used. As it can be seen from figure (2) $C_2 = 1$variant and ℓog in an expression (2) can be considered a special case of aesthetic curves.

II. Application of Curves in Engineering.

In descriptive geometry, there are various methods for generating surfaces using curved lines. Surfaces can be formed through the continuous motion of a guiding line, referred to as a generating line, or they can be created as a framework formed by multiple lines. Both methods are widely used in computer technologies.

For example, consider the modeling of a cam. In the kinematic analysis of cam mechanisms, it is essential to determine the motion law of the follower, which corresponds to the cam profile. Since the profile curve of the cam is complex, constructing it graphically in several positions can pose challenges, leading to inaccuracies in calculations.

Utilizing computational tools and advanced modeling techniques allows for more precise generation of these curves, enabling engineers to achieve higher accuracy and efficiency in the design and analysis of mechanical systems.

The application of modern computer technologies has made kinematic analyses more accurate and easier to perform. For this purpose, various software solutions in the fields of CAD (Computer-Aided Design) and CAE (Computer-Aided Engineering) are widely used around the world. For example, SolidWorks, produced by Dassault Systèmes, holds a leading position globally.

The profile of the cam is drawn in the SolidWorks program based on the displacement-rotation angle dependency graph obtained from (2) and using circles in the polar coordinate system (Figure 3). For example, the phase angles used as boundary conditions in the construction of the profile are as follows; $\varphi_u = 60°$ - divergence angle, $\varphi_{ud} = 40°$ - distant stopping angle, $\varphi_y = 60°$- approaching angle, $\varphi_{yd} = 200°$- the angle of stopping near has been accepted.[12].

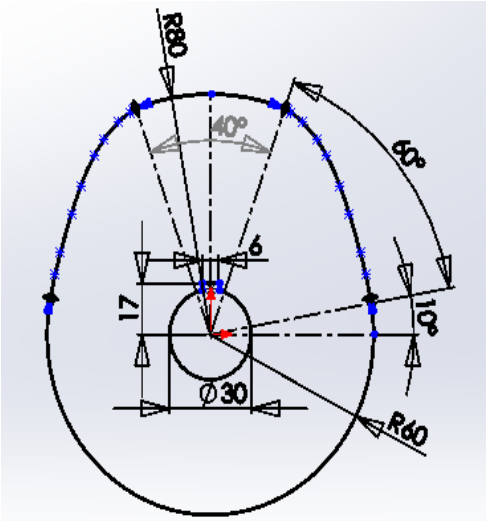


Figure 3: *The profile of the cam construction in SolidWORKS software.*

In Figure 3, other dimensions are taken arbitrarily and do not affect the intended calculations. The diameter of the cam's roller is assumed to be 25 mm. When constructing the contour in SolidWORKS, (2) is used directly. That is, the departure and approach lines are obtained using the mathematical expression (2) within the given interval. After the contour is established, it is extruded to obtain a 3D model.

To demonstrate the practical application of the cam, let's consider the assembly sequence of the designed cam mechanism. Just like the 3D model of the cam, 3D models of all parts related to the cam mechanism are also created. The 3D model of the parts included in the assembly is created and each is saved in a separate file. The 3D modeling is carried out based on the "bottom-up" principle. In Assembly mode, the 3D assembly of the cam mechanism is developed based on the 3D model of the parts (Figure 4).

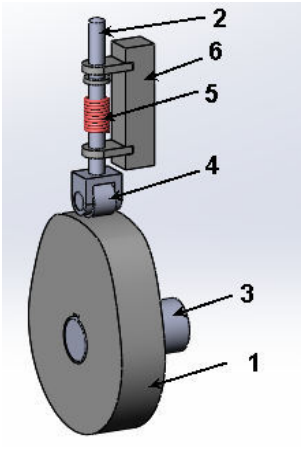


Figure 4: *3D model of the cam mechanism (1 - cam, 2 - pusher, 3 - shaft, 4 - roller, 5 - spring, 6 - support).*

Another example of a framework-type surface is blades. Aerodynamic blades are created in the form of a framework by twisting two-dimensional profiles located at varying distances along the height. The blade shown in Figure 5 is created based on three cross-sections (two-dimensional profiles). Using the expression (2) ($C_2 = 1$) he belly and back parts of the profile in SolidWorks are joined with the entry and exit circles to form a complete closed contour. In the height direction of the frame, the exit lines are assumed to be straight. In SolidWorks, using the Boundary Boss/Base command, the 3D model of the blade is created based on three profiles and the side straight lines (Figure 5).

For practical application, for example, to perform CFD (Computational Fluid Dynamics) analyses based on the 3D model of the blade, either SolidWorks is used, or the 3D model is converted into Parasolid (.x_t) format and transferred to the ANSYS program.

Figure 5: *3D Model of the Blade.*

III. Conclusion.

The conclusions and recommendations derived from the article are as follows:

• The smoothness of curves should be considered, and the appropriate selection should be made based on the application field of the curve;

• Smoothness should be ensured in torsion, which is one of the key properties of space curves;

• It is advisable to keep the minimum of the maximum curvature value and the rate of change of

curvature within a certain limit;

- Efforts should be made to minimize the potential energy of curvature;
- The design of aerodynamic surfaces should be based on intersecting lines, and the optimal variant should be selected;
- To ensure the effective use of surfaces in engineering and design, their 3D models should be created using computer technologies;
- Based on the presented data and methodology, it will be possible to select the appropriate applications of curves in engineering in the future.

References

[1] V. G. Muftiev, R.A. Ziatdinov. Functionality and Aesthetic of Curves in Industrial Design: A Multicriteria Approach to Quality Evaluation of Forms in Future CAD Systems. Published in *Vestnik Mashinostroyeniya*, Issue 7 (2018).
http://www.mashin.ru/eshop/journals/vestnik_mashinostroeniya/2039/18/.

[2] Ziatdinov R. Family of superspirals with completely monotonic curvature given in terms of Gauss hypergeometric function// Computer Aided Geometric Design. 2012. Vol. 29. No.7. P. 510-518

[3] Levien R. L. From Spiral to Spline: Optimal Techniques in Interactive Curve Design. PhD.thesis, University of California, Berkeley, 2009

[4] Nikolaos Eliou, Georgios Kaliabetsos, A new, simple and accurate transition curve type, for use in road and railway alignment design, European Transport Research Review An Open Access Journal, 29 September 2013.

[5] Esveld C., Modern railway track, 2nd edn. T.U. Delft Press, The Netherlands, 2001.

[6] Kasper H., Schuerba W., Lorenz H. The clothoid as an element of horizontal alignment. F. Dummlers, Publishing House, Bonn, 1954.

[7] Agata Basak, The Study of Geometry of the Selected Transition Curves in the Design of Circular RoadsAdvances in Science and Technology Research Journal 2022, 16(4), 270–278, https://doi.org/10.12913/22998624/152936, ISSN 2299–8624, License CC-BY 4.0

[8] E.A. Gavrilenko, YU.V. Kholodnyak. Forming of geometric characteristics of monotonous curves. Applied Geometry and Computer Technologies. Bulletin of Kharkiv National Technical University, Issue 3(58), 2016.

[9] A.S. İmanov, I.A. Khalilov 3D modeling and analysis of gas flow in the interblade channel. SOCAR Proceedings Special Issue 1 (2022) 001-005 http://proceedings.socar.az., 2022, 5 p., DOI: 10.5510/OGP2022SI100690.

[10] A.S. İmanov, I.A. Khalilov, A.G. Aliyev New approach to calculation of transition curves on curved roads. Proceedings of the International Conference on Problems of Logistics, Management and Operation in the East-West Transport Corridor (PLMO), Baku, Azerbaijan. Oktobmer 27-29, 2021, 6 p.

[11] A.S. Imanov, P.S. Abdullayev. Profiling of flat aviation blades based on the differential equation of curvature. *Vestnik of Engine Engineering*, Issue No. 2. Zaporizhzhia, JSC. «Мотор Сич», 2015, 6 ст, ISSN 1727-0219.

[12] A.S. İmanov, I.A. Khalilov Kinematic and Dynamic Calculation of profiled Cam Mechanism Based on New Equation. International Symposium on Unmanned Systems: AI, Design, and Efficiency was held in Baku, Azerbaijan on 22nd of May and 24th of May. ISBN : 978-9952-582-04-8 DOI : 10.30546/2224.978-9952-582-04-8.Springer.

MODELING AND SIMULATION OF DUCTILE-IRON BLANK CASTING PROCESSES FOR AN ELECTROHYDRAULIC POWER AMPLIFIER BODY

Vitaly Dubrovin[1], Boris Kulakov[1], Andrey Karpinsky[1], Dmitry Ardashev[2], Anastasiya Degtyareva-kashutina[2], Ramil Dadashov[3]

•

[1]Department of Pyrometallurgical and Foundry Technologies, South Ural State University, Russia, 454080, Chelyabinsk, Lenin st. 76

[2]Department of Automated Mechanical Engineering Technology, South Ural State University, Russia, 454080, Chelyabinsk, Lenin st. 76

[3]Department of Machine Building Technology, Azerbaijan Technical University, H.Javid avenue 25, Baku, Azerbaijan AZ 1073

dubrovinvk@susu.ru, kulakovba@susu.ru, karpinskiiav@susu.ru, ardashevdv@susu.ru, degtiareva-kashutinaas@susu.ru, dadashov@aztu.edu.az

Abstract

The paper describes the process of modeling a casting mold for a body blank of an electrohydraulic power amplifier and the technology of ductile iron melting in small furnaces.

Keywords: electrohydraulic power amplifier, casting, gating system modeling, melting technique.

I. Introduction

The body of the electrohydraulic power amplifier (EHPA) is a unique part; there is no information about the practice of manufacturing similar parts in the technical literature.

The body of EHPA is a parallelepiped, in the center of which there is a stepped hole for placing a spool-sleeve pair in it. A developed system of curved channels of small diameter diverges from the central hole throughout the body, designed for the flow of working fluid with the required pressure, flow rate, and force.

During operation the body is subjected to hydraulic loads under high pressure. The proposed material for manufacturing the body is high-strength cast iron (ductile iron with grades from VCh35 to VCh50 GOST 7293-85); the manufacturing method is casting [1, 2].

Intricate casting elements are internal cavities of developed configuration, which can be obtained using special elements of the casting mold, the so-called cores. Since the casting has small dimensions and is prepared in small-scale production, it is advisable to manufacture it in disposable resin sand molds obtained by 3D prototyping. This method does not require the manufacture of expensive model equipment.

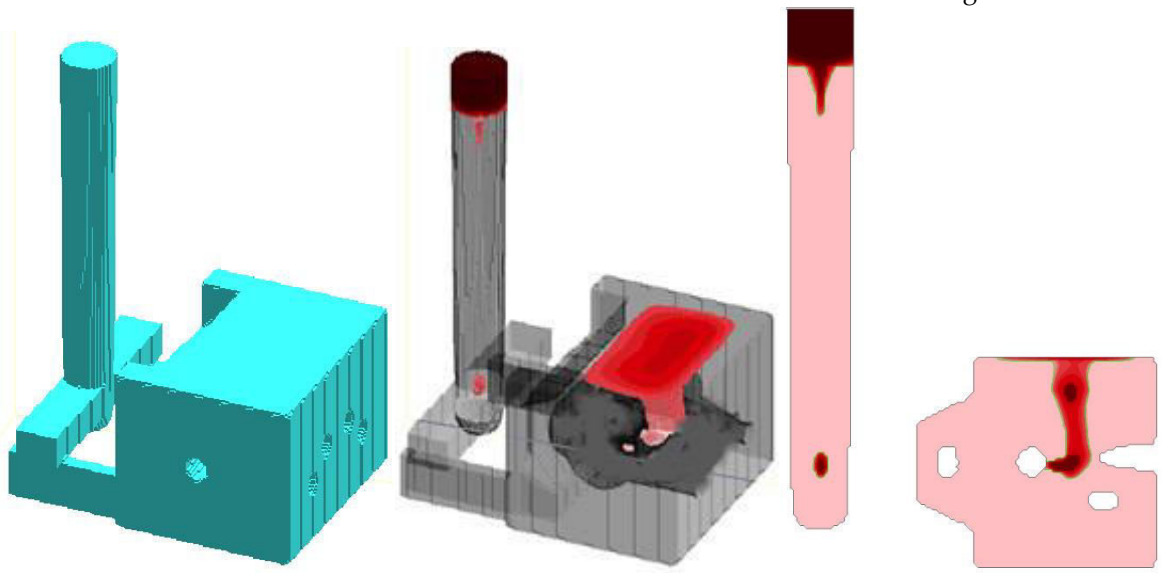
II. Development of the casting mold design

First, it is necessary to design the gating system (GS). Since the casting is small in size and made of cast iron, we select the classic GS type, consisting of a funnel, down gate, sump, slag trap, and feeders.

The calculation is based on determining the optimal time for filling the cavity of the casting mold and calculating the GS bottle neck, which determines this time [3-5].

Ductile iron has significant volumetric shrinkage. To prevent such shrinkage defects as cavities and porosity, it is necessary to provide for a riser, that is, a technological reservoir that feeds the casting with liquid metal during the solidification process.

To identify the features of shrinkage processes in the casting, computer modeling was carried out in the LVMFlow finite-difference system. The 3D model of the casting and the simulation results are shown in Figure 1. The dimensions of the sections of the GS elements were as follows: one down gate with a diameter of 19 mm, two slag trap branches of rectangular cross-section 14×18 mm each, and two feeders of rectangular cross-section 6×20 mm each. According to the calculation, each of the two risers was 166 mm high and 30 mm in diameter. The technological yield (TY) with this design and dimensions of the GS was 55.1%. The amount of metal in the mold was 5.525 kg for ductile iron.



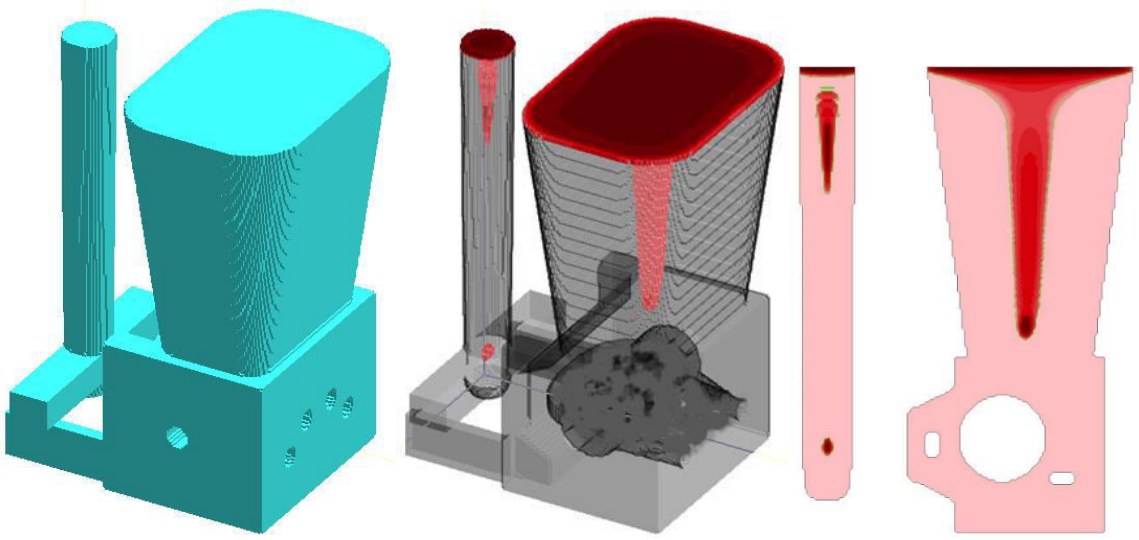
a) b) c)

a) 3D model; b) volumetric view; c) cross section of casting

Figure 1: *Casting with a gating system and the results of modeling the formation of shrinkage defects in the casting body*

The simulation results showed that these risers worked inefficiently, and shrinkage defects penetrated into the casting. The end of shrinkage defects in the casting was on the axis of its heating element. To improve the casting feeding process, we decided to use one central riser located above the center (on the axis) of the heating element. The simulation results showed (Fig. 2) that the place of probable formation of shrinkage defects was removed from the casting body. This allowed us to conclude that the volume of this open riser was sufficient to obtain a high-quality casting.

.



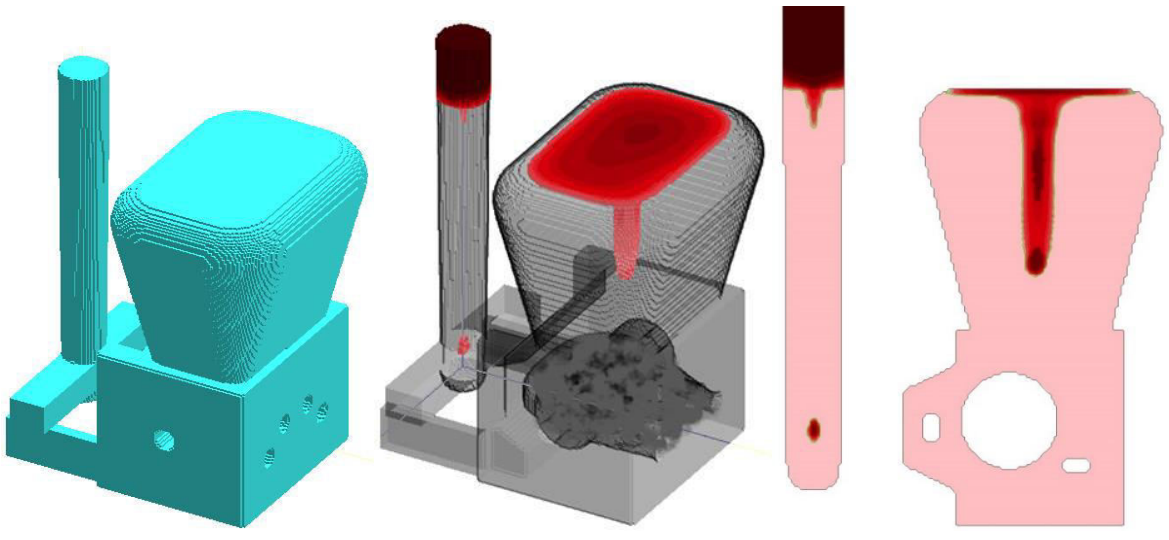
a) b) c)

a) 3D model; b) volumetric view; c) cross section of casting

Figure 2: *Results of modeling the formation of shrinkage defects in the casting body*

However, with such a riser weighing 5376.3 g, the technological yield decreased to 33.6%, and the metal weight in the mold was 9067.2 g. Such a level of TY when producing castings from ductile iron is considered very low. To reduce the weight of the riser, we switched to a closed type of risers to reduce heat loss by radiation and the rate of riser solidification. As a result, we changed the design of the riser: its height was reduced from 116 mm to 86 mm and all sharp corners of the riser, which acted as chillers during riser solidification, were rounded. As the simulation results showed (Fig. 3), the efficiency of such a riser became significantly higher than in the previous version. At the same time we managed to increase TY to 37.4% and reduce the amount of metal in the mold to 8146.8 g. We continued working in this direction.

.



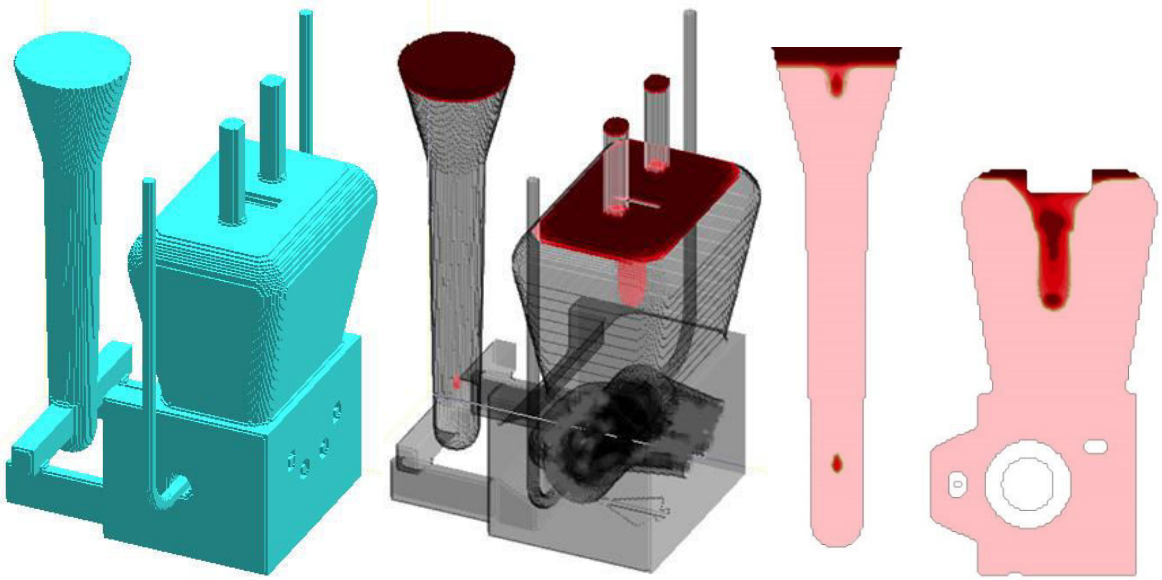
a) b) c)

a) 3D model; b) volumetric view; c) cross section of casting

Figure 3: *Results of modeling the formation of shrinkage defects in the casting body (a variant of the technology with one closed rectangular-sectioned riser with an expansion towards the top)*

At the final stages of computer modeling, we tested the variant of manufacturing a casting with empty spaces in those parts of the mold that play the role of cores and are intended to obtain internal

cavities in the casting. This will help to reduce the gas formation from those parts of the mold when obtaining experimental castings. In this version of manufacturing the casting in the mold, we also provided channels coming out to the surface and intended to improve the removal of gases formed from the mold. According to the results of modeling, no places of probable formation of shrinkage defects were found in the body of the casting (Fig. 4) and a calm laminar flow of the melt was noted when filling the casting cavity, and this variant of the technology is recommended for use in obtaining experimental castings.



a) b) c)

a) 3D model; b) volumetric view; c) cross section of casting

Figure 4: *Results of modeling the formation of shrinkage defects in the casting body with the pouring temperature TPOUR = 1420 °C*

Development of technology for melting high-strength cast iron in small-capacity furnaces

Melting was carried out in a PPI-0.02 induction crucible furnace with thyristor converters based on pig iron PL1. The load was 10 kg.

III. First version of melting

After melting the cast iron, ferromanganese was added to the furnace in the amount of 21 g. To obtain spheroidal graphite in the structure, a complex modifier of the following composition was used: a SferoMag 611 graphite spheroidizer in the amount of 200 g and a Sibar 4 graphitizer in the amount of 50 g, which was pre-loaded into the pouring ladle before heating it.

The ladle was heated to a temperature of 600 ° C together with the modifier. After pouring the metal from the furnace into the ladle, a significant flare indicated the processes of spheroidizing modification. Experimental rods were casted into sand-clay molds and samples for chemical composition were casted into a metal chill mold.

The results for chemical composition (Table 1) and structure (Fig. 5) showed the presence of spheroidal graphite of irregular round shape (allowed) and ferrite-pearlite base of the alloy with ferrite content of 50%. The manganese content in the alloy was slightly below the lower limit recommended by GOST 7293-85. We used the DFS-500 optical emission spectrometer to analyze chemical composition and the SIAMS complex to determine the structure.

Table 1: *Chemical composition of cast iron (melting no.1)*

Element	Fe	C	Si	Mn	Mg	Cr	Ni	Cu	S	P
Weight content, %	93.8	3.688	2.085	0.182	0.078	0.014	0.009	0.0026	0.017	0.0097

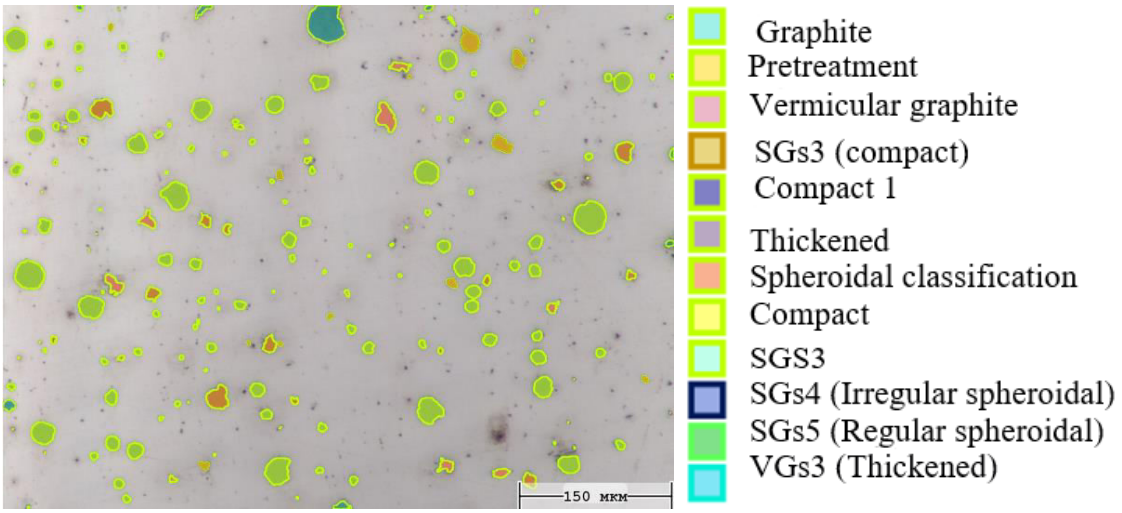


Figure 5: *Structure of cast iron (melting no.1)*

Determination of mechanical properties showed that the strength characteristics corresponded to the VCh50 grade of GOST 7293-85, and the relative elongation was slightly lower than recommended by GOST 7293-85.

We decided to conduct the second melting, adjusting the charge composition.

IV. Second version of melting

The loading of pig iron PL1 was also 10 kg. During the melting process, ferromanganese was added to the crucible in the amount of 25 g, ferrosilicon 24 g to the furnace and 25 g to the pouring ladle. The SferoMag 611 modifier in the amount of 180 g, the Sibar 4 in the amount of 120 g, and ferrosilicon in the amount of 25 g were placed in the pouring ladle before tapping the metal from the furnace. The metal tapping temperature was 1500 °C and the ladle temperature was 800 °C. Temperature measurements were made with a VR 20-5 tungsten-rhenium thermocouple. After the metal was poured from the furnace into the ladle, a significant flare occurred, indicating the processes of spheroidizing modification. The holding time in the ladle was 3 minutes and the mold pouring temperature was 1330 °C.

Experimental rods were casted into sand-clay molds and samples for chemical composition were casted into a metal chill mold. After pouring, the experimental rods were separated from the sprues and cleaned of burnt-on deposits. Dumbbell samples were cut from the rods to determine mechanical properties. The results for the chemical composition (Table 2) and structure (Fig. 6) showed the presence of spheroidal graphite of a regular round shape (the best option) and a ferrite-pearlite base of the alloy with a ferrite content of 88.4%, which is typical for cast iron of VCh50 grade.

Table 2: *Chemical composition of cast iron (melting no.2)*

Element	Fe	C	Si	Mn	Mg	Cr	Ni	Cu	S	P
Weight content, %	93.8	3.342	2.464	0.208	0.055	0.017	0.012	0.0074	0.019	0.025

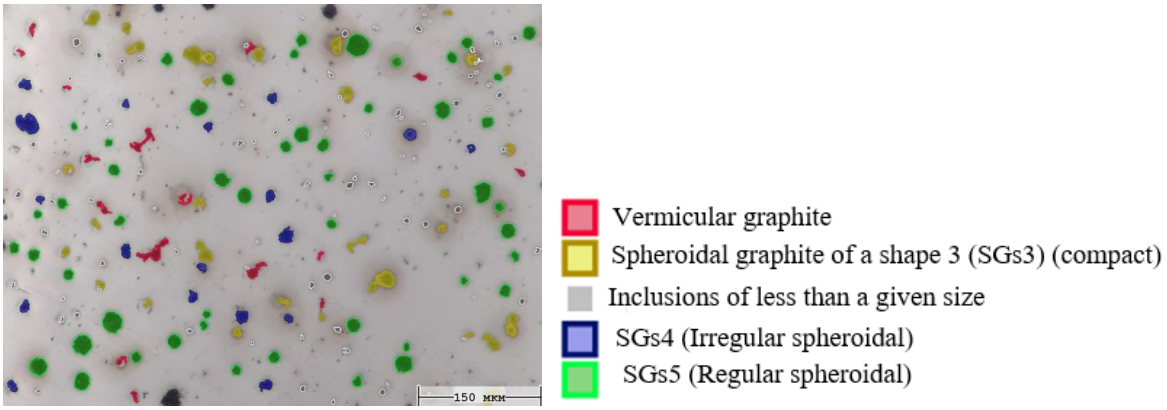


Figure 6: *Structure of cast iron (melting no.2)*

The mechanical behavior test showed that the characteristics averaged over three similar test results corresponded to the grade VCh50 GOST 7293-85 (Table 3).

Table 3: *Mechanical properties of cast iron (melt No. 2)*

Property	Tensile strength, MPa	Offset yield stress, MPa	Relative elongation, % (additional parameter)
Value	504	348	7

When melting and pouring small volumes of metal (10 kg), rapid cooling of the melt in the ladle occurs (approximately 50 ºC/min), which limits the time of graphitizing holding and complicates the production of cast iron with a fully ferritic matrix structure (grades VCh40 and VCh45).

High-strength cast iron VCh50 has a high level of strength properties (tensile strength and offset yield stress), sufficient plastic properties (relative elongation) and vibration resistance, for which there is no need to have high plastic properties. Therefore, for the production of the EHPA cast body, along with cast iron VCh45, high-strength cast iron VCh50 can also be recommended as the main material.

Two cast bodies of an electrohydraulic power amplifier were made of ductile cast iron in accordance with the developed technical solutions and successfully passed hydraulic tests under a pressure of 500 atm.

V. Acknowledgments

The work was carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation as part of the implementation of a high-tech production comprehensive project “Creation of a high-tech production of electrohydraulic power amplifiers with an electromechanical converter of an electrodynamic type with an extended frequency range” using measures of state support for the development of cooperation between Russian educational organizations of higher education, state scientific institutions and organizations of the real sector of the economy implementing comprehensive projects to create high-tech production, provided for by Decree of the Government of the Russian Federation of April 9, 2010 No. 218 under Agreement No.

075-11-2023-005 dated February 13, 2023. (State contract identifier 000000S407523Q8R0002) between the Ministry of Science and Higher Education of the Russian Federation and the Ural Engineering Center Limited Liability Company in cooperation with the Lead Contractor of the R&D - the Federal State Autonomous Educational Institution of Higher Education "South Ural State University (National Research University)". The R&D was performed in the organization of the Lead Contractor of the R&D.

References

[1] I.A. Savin, R.V. Gavariev. Design technological equipment for production of castings non-ferrous metal alloys. "Machine-building and Energy: New Concepts and Technologies" International Scientific-practical Conference materials, December 2-3, 2021, AzTU, Baku, Azerbaijan. Pp. 24-27.

[2] F.R. Rasulov, A.F. Rasulzade. Mathematical assessment of casting formation in contact within liquid metal-dispersed metallic surface of mold. "Machine-building and Energy: New Concepts and Technologies" International Scientific-practical Conference materials, December 2-3, 2021, AzTU, Baku, Azerbaijan. Pp. 148-151.

[3] V.K. Dubrovin, A.V. Karpinsky, O.M. Zaslavskaya. Technological processes of casting. Chelyabinsk: Publishing Center of SUSU, 2013. 194 p.

[4] B.S. Churkin Technology of foundry production. Yekaterinburg. Publishing State prof. ped. University. 2000. 662 p.

[5] V.M. Kolokoltsev, Ri Hosen Production of iron castings. Magnitogorsk, State Educational Institution of Higher Professional Education "MSTU", 2009. 521 p.

FLUID FLOW MODELING IN THE SPOOL AND SLEEVE OF AN ELECTRO-HYDRAULIC POWER AMPLIFIER

Darya Khabarova[1], Sergey Bitiutckikh[1], Alexander Ismagilov[1], Dmitry Ardashev[2], Mukhaddin Samadov[3], Heyran Abbasova[3]

•

[1]Hydraulic and Pneumatic Systems department, South Ural State University, Russia, 454080, Chelyabinsk, Lenin st. 76

[2]Department of Automated Mechanical Engineering Technology, South Ural State University, Russia, 454080, Chelyabinsk, Lenin st. 76

[3]Department of Machine Building Technology, Azerbaijan Technical University, H.Javid avenue 25, Baku, Azerbaijan AZ 1073

khabarovadf@susu.ru, bitiutckikhsi@susu.ru, ismagilovar@susu.ru, ardashevdv@susu.ru, m.semedov@aztu.edu.az , abbasova.heyran@aztu.edu.az

Abstract

This article deals with the initial data for modeling a spool valve. It describes the volumetric solid model of the flowing part of the high precision spool pair layouts in the electro-hydraulic power amplifier; the generation of the finite-element mesh of the solid model; the physical and mathematical model of the fluid flow process; and the estimations and their analysis. The pressure characteristics of the spool valve under different load are calculated. 3D modeling of fluid flow through the annular clearance between the spool and sleeve in the electro-hydraulic power amplifier (fluid leak modeling) is performed.

Keywords: electro-hydraulic power amplifier, spool and sleeve, fluid leak modeling, computational fluid dynamics

I. Introduction

The precision and efficiency of modern industrial hydraulic systems are due to the use of automation and servo control elements. One such element is the electro-hydraulic power amplifier (EHPA), which regulates the speed and power of the hydraulic drive. However, the increasing demands on the performance of hydraulic drives require an increase in the accuracy of design calculations, including those for spool-type EHPAs [1-4]. The simulation of fluid flow through the annular clearance between the sleeve and spool of an EHPA is effective for optimizing valve performance and estimating leak rates. Modeling results can be used to make informed decisions on valve design and tuning and to improve overall system performance.

II. Modeling of valve hydrodynamics

The EHPA model under study includes a housing and a high-precision spool and sleeve valve. The other valve components such as the linear motor, position sensor, fasteners and seals were

removed from the model as they do not contribute to the study of the fluid dynamics of the device. The high-precision spool and sleeve valve consists of a fixed sleeve and a spool; the latter is mounted inside the sleeve and can move along the longitudinal axis. The spool and sleeve are installed in the central groove of the housing. The spool is shaped to ensure the required connections of the hydraulic lines.

A volumetric solid model of the P-A and B-T valve flow paths was designed in Autodesk Inventor software from a 3D model of the valve. The P-B and A-T paths are of the same geometry as P-A and B-T because of the symmetry of the spool and valve housing. Therefore, it is reasonable to limit the analysis to the P-A and B-T paths only.

The 3D model of the hydraulic amplifier is simplified and adapted for hydrodynamic calculations. We made the following assumptions: there is no radial clearance between the sleeve and the rod, and the radial grooves and the unloading line of the spool end surfaces are excluded. The 3D model was formed at different spool positions: 5%, 10%, 25%, 50%, 75% and 100% of X_{max}. We applied the Ansys Meshing preprocessor to build the computational mesh [5, 6].

The modeling of the valve hydrodynamics revealed the pressure characteristics of a high-precision spool valve. We determined the volumetric flow rate through the valve and the pressure drop at different spool movements.

Figure 1 shows the characteristics of the EPHA, represented by the dependence of the hydraulic fluid flow rate on the spool movement at four pressure drops ΔPAB.

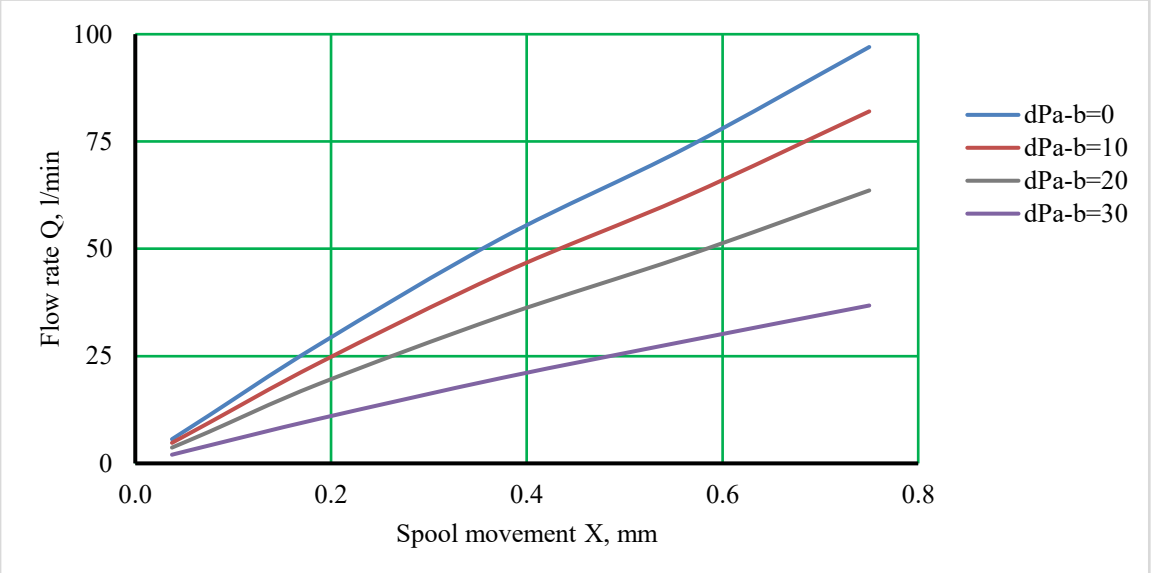


Figure 1: *Dependence of the spool movement on the hydraulic fluid flow rate through the valve (nominal flow rate 40 l/min)*

III. Calculation of the generalized characteristic of the EHPA.

The spool-sleeve system in the EHPA is a system of controlled restricting openings, combined in one structure (Fig. 2).

The fluid flow with flow rate Q_p and pressure p_p enters the sleeve, then the goes around the spool and enters the hydraulic motor operating cavities with flow rate Q_m and pressure p_1. The fluid flow goes back from the hydraulic motor to the operating cavity of the spool at the same flow rate Q_m and pressure p_2 and then to the output in the drainage pipeline at a flow rate Q_{dr} and pressure p_{dr}. Flow rates Q_3 and Q_4 must be taken into account due to the internal fluid leakage through the gaps.

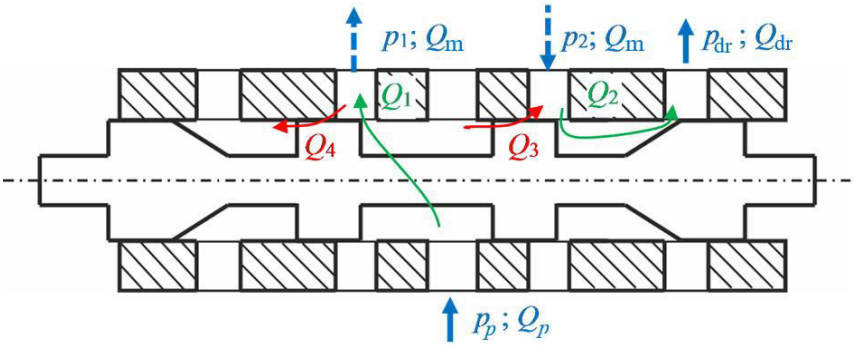


Figure 2: *Calculation of the sleeve-spool system*

The system of equations describing the physical fluid flow through the EHPA-hydraulic motor system, includes the following continuity equations and Bernoulli equations:

$$Q_1 = \mu_1 \cdot S_{or1} \cdot \sqrt{\frac{2\cdot(p_p - p_1)}{\rho}} \quad (1)$$

$$Q_2 = \mu_2 \cdot S_{or2} \cdot \sqrt{\frac{2\cdot(p_2 - p_{dr})}{\rho}} \quad (2)$$

$$Q_3 = \mu_3 \cdot S_{or3} \cdot \sqrt{\frac{2\cdot(p_p - p_2)}{\rho}} \quad (3)$$

$$Q_4 = \mu_4 \cdot S_{or4} \cdot \sqrt{\frac{2\cdot(p_2 - p_{dr})}{\rho}} \quad (4)$$

$$Q_m = Q_1 - Q_4 \quad (5)$$

$$Q_m = Q_2 - Q_3 \quad (6)$$

$$P_m = p_1 - p_2 \quad (7)$$

where p_m is pressure drop on the hydraulic motor; p_1, p_2 are pressures in the operating cavities of the hydraulic motor; Q_m is liquid flow rate entering the hydraulic motor; Q_i is liquid flow rate through the i-th restricting opening; μ_i is the flow coefficient of the liquid flow through the i-th restricting opening; $S_{or\text{-}i}$ is the cross-sectional area of the i-th restricting opening.
Flow coefficients μ_i are determined by the results of computational modeling presented in Table 1.

Table 1: *Flow coefficients*

Spool position, mm	Pressure drop ΔP_{AB}=0 MPa		Pressure drop ΔP_{AB}=10 MPa		Pressure drop ΔP_{AB}=20 MPa		Pressure drop ΔP_{AB}=30 MPa	
	Flow rate Q, l/min	Flow coefficient	Flow rate Q, l/min	Flow rate Q, l/min	Flow rate Q, l/min	Flow coefficient	Flow rate Q, l/min	Flow coefficient
0.04	5.4	1	4.6	2.0	2.0	1	2.0	1
0.08	11.1	0.96	9.4	4.2	4.2	0.96	4.2	0.96

0.19	27.8	0.92	23.5	10.4	10.4	0.93	10.4	0.91
0.38	52.7	0.86	44.5	19.9	19.9	0.86	19.9	0.86
0.56	73.9	0.8	62.4	28.6	28.6	0.8	28.6	0.82
0.75	97.5	0.79	82.4	36.8	36.8	0.79	36.8	0.79

The following formulae have been applied to calculate the areas:

$$S_i = \begin{cases} n \cdot b \cdot \sqrt{(X_з - G_i)^2 + \delta^2} \text{ при } (X_з - G_i) > 0 \\ n \cdot b \cdot \delta \text{ при } (X_з - G_i) \le 0 \end{cases}; \quad (8)$$

where n = 4 is the number of restricting openings; b = 5.2; 3.5; 1.8 mm are the opening widths; $X_з$ is spool movement; G_i are the overlaps of corresponding restricting openings.

When testing, the fluid flow is poured through the EHPA unloaded, when the hydraulic lines connecting the EHPA to the hydraulic motor are connected to each other. In this case, the pressures in the hydraulic lines controlling the hydraulic motor are equal to $p1 = p2$.

To determine the unified characteristics, Equations (1–8) were calculated. These establish the relationship between the spool X_{sp} positions, fluid flow and load, indirectly characterized by the pressure difference in the operating chambers of the hydraulic motor ($p_p = p_1 - p_2$) for the pressure drop at the valve $\Delta p = 35$ *MPa*.

The results of 3D modeling allow us to determine the flow coefficients μ at the restricting openings. For this purpose, the coefficients $\mu_1$and μ_2 are set using the MathCad computer program by successive approximations. The results of Equations (1–8) in the form of flow rates are compared with the results of 3D modeling. The coefficients of flow rates μ were determined with a relative error of flow rates Q_i not exceeding 1%.

Figs. 3 and 5 show characteristics for three EHPA models in dimensional and dimensionless coordinates. Here $dP = p_1 - p_2$ is the pressure drop in the operating chambers; $X = X_{sp} / X_{max}$ is a dimensionless coordinate of the spool position at X_{max} = 0.75 mm; $\overline{Q_m} = Q_m / Q_{max}$ is a dimensionless flow rate at maximum flow rate Q_{max}, when the spool takes position 0,75 mm, and $dP = 0$; $\overline{p_m} = dP / p_p - p_{or}$ is dimensionless pressure difference in operating chambers of the hydraulic motor at $p_p - p_{or} = = 35$ *MPa*.

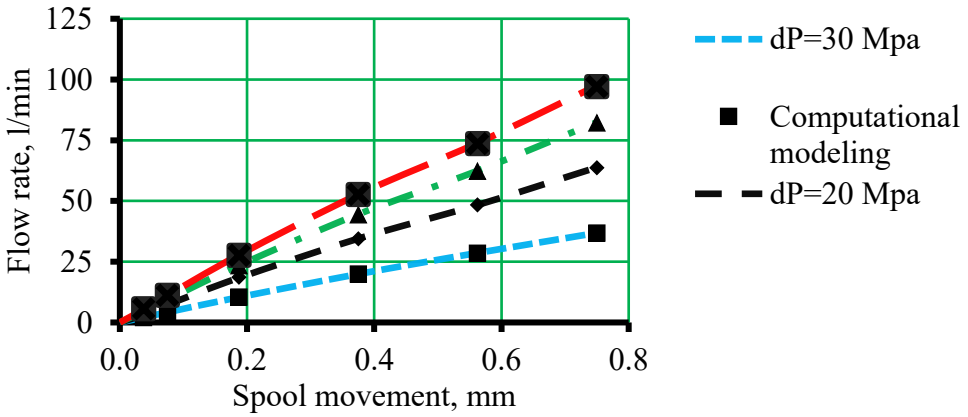


Figure 3: *Influence of spool position on the flow rate through EHPA with a nominal flow rate of 40 l/min at different pressure drops in channels dP = p1 - p2*

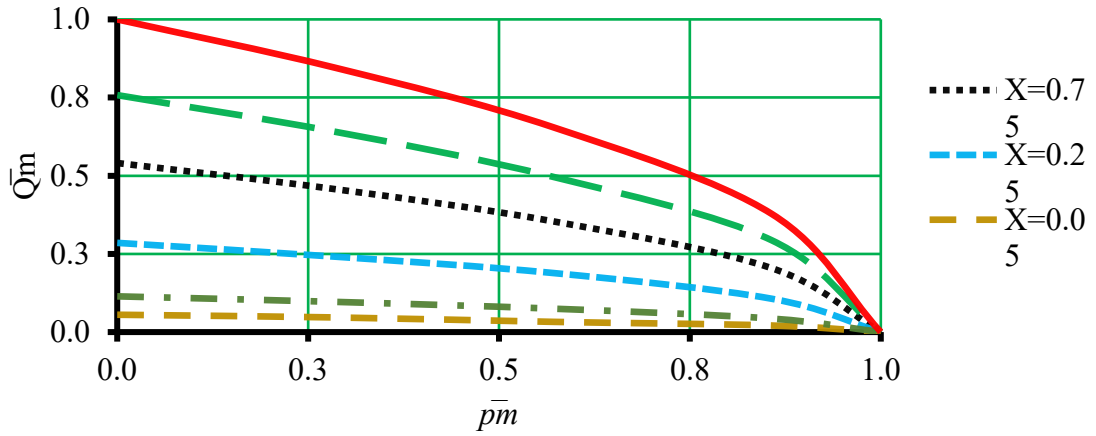


Figure 4: *Characteristics of the EHPA with a nominal flow rate of 40 l/min at Q_{max} = 97.5 l/min*

The pressure distributions show that an increase of the open flow area of the spool leads to increase in the liquid flow rate and the irregularity of the pressure distribution in the opening of the flowing part of the valve. This includes an increase in the irregularity of the velocity accompanied by local vortex and a decrease in hydrostatic pressure.

IV. Modeling fluid leaks in the valve

The operation of the EPHA results in fluid leak in the annular clearance. The annular clearance is the area between the sleeve and the spool that carries the fluid flow. Leaks depend on the fluid properties, pressure drop, and the clearance shape. The geometric parameters of the annular clearance have been determined according to the design documentation. The outer diameter of the ring is D = 8.058 mm, the inner diameter is d = 8.052 mm. The clearance widths for four spool positions 5%, 10%, 50%, and 100% of X_{max}, are 0.04 mm, 0.08 mm, 0.38 mm, and 0.75 mm, respectively.

The mathematical model of fluid flow through the valve annular clearance is based on the equations of the continuity of flow and impulse. We adopted the laminar model of fluid flow through the annular clearance. The inlet pressure is set on the inlet of the calculation area. The static pressure is accepted on the outlet. The wall condition is defined on the opening housing surfaces. We carried out calculations at three values of pressure drop on the annular clearance, respectively: 17.5 MPa, 7.5 MPa, and 2.5 MPa.

According to the results of hydrodynamics modeling through the annular clearance of the valve, we determined the values of the pressure drop and the flow rate of liquid (Fig. 5).

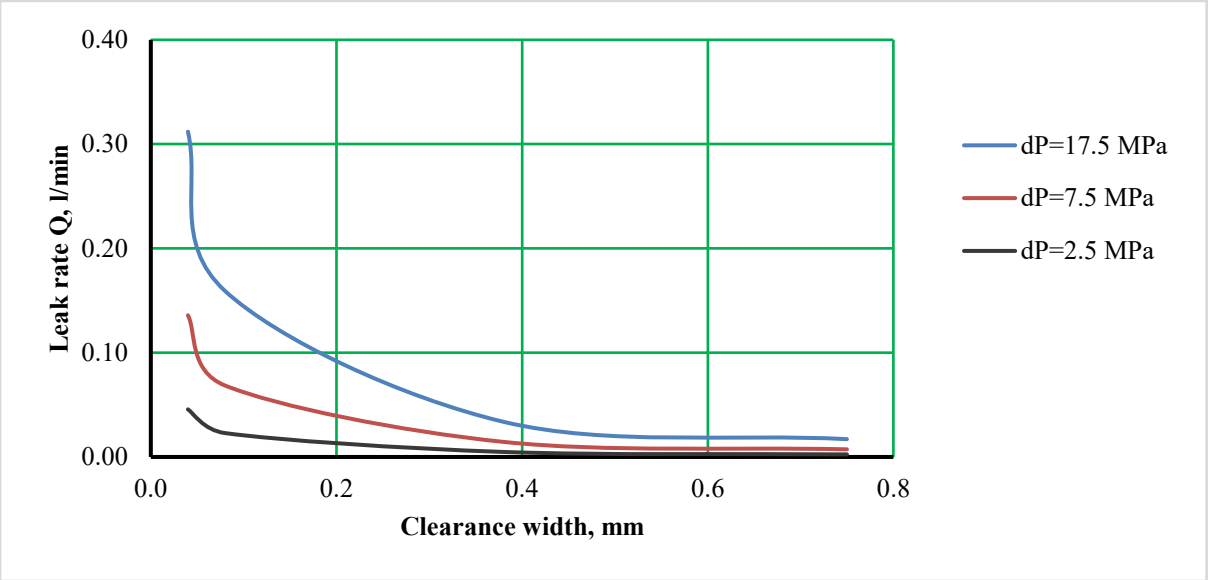


Figure 5: *Dependence of leak rate on the clearance width*

Figure 5 shows that an increase in the pressure drop across the spool and the width of the annular clearance leads to a significant increase in fluid leak.

The modeling provides a reasonable estimate of the leak rate through the annular clearance between the spool and sleeve. This helps to understand how flow affects the valve and to improve the valve performance.

V. Acknowledgments

The work was carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation as part of the implementation of a high-tech production comprehensive project "Creation of a high-tech production of electrohydraulic power amplifiers with an electromechanical converter of an electrodynamic type with an extended frequency range" using measures of state support for the development of cooperation between Russian educational organizations of higher education, state scientific institutions and organizations of the real sector of the economy implementing comprehensive projects to create high-tech production, provided for by Decree of the Government of the Russian Federation of April 9, 2010 No. 218 under Agreement No. 075-11-2023-005 dated February 13, 2023. (State contract identifier 000000S407523Q8R0002) between the Ministry of Science and Higher Education of the Russian Federation and the Ural Engineering Center Limited Liability Company in cooperation with the Lead Contractor of the R&D - the Federal State Autonomous Educational Institution of Higher Education "South Ural State University (National Research University)". The R&D was performed in the organization of the Lead Contractor of the R&D.

References

[1] E. Lisowski, W. Czyzycki, J. Rajda Three Dimensional CFD Analysis and Experimental Test of Flow Force Acting on the Spool of Solenoid Operated Directional Control Valve. Energy Conversion and Management. 2013. 70. P. 220–229. DOI: 10.1016/j.enconman.2013.02.016

[2] H. Chattopadhyay, A. Kundu, B.K. Saha, T. Gangopadhyay Analysis of flow structure inside a spool type pressure regulating valve. Energy Convers Manage, 2012, 53.

[3] J.R.Valdes, M.J. Miana, J.L. Nunez, T. Putz Reduced Order Model for Estimation of Fluid Flow and Flow Forces in Hydraulic Proportional Valves. Energy Conversion and Management. 2008.

49, 6. P. 1517–1529. DOI: 10.1016/j.enconman.2007.12.010.

[4] S.X. Karimov, F.S. Hajiyeva, S.V. Ismayilova. Calculation of the structure of the column of the rods of the deep pump installations for separate operation of two layers. "Machine-building and Energy: New Concepts and Technologies" International Scientific-practical Conference materials, December 2-3, 2021, AzTU, Baku, Azerbaijan. Pp. 140-143

[5] Ansys Fluent 19 Theory Guide. Ansys Inc.

[6] Ansys Fluent 19 User Guide. Ansys Inc.

INVESTIGATION OF THE INFLUENCES ON ENERGY CONSUMPTION DURING TURNING AND ITS MODELLING

Ilgar Abbasov[1], Rezo Aliyev[2], Arastun Mammadov[3], Mahabbat Suleymanov[1], Huseyn Mammadov[1]

•

[1]Brandenburg University of Technology Cottbus-Senftenberg, Platz der deutschen Einheit 1, 03046 Cottbus, Germany.

[2]TU Bergakademie Freiberg, Agricolastrasse 1,09599 Freiberg, Germany

[3]Azerbaijan Technical University, Huseyn Cavid pr.25, Baku, AZ1073

Ilgar.Abbasov@b-tu.de, rezo.aliyev@imkf.tu-freiberg.de, arastun.mammadov@aztu.edu.az

Mahabbat.Suleymanov@b-tu.de, huseyn.mammadov@b-tu.de

Abstract

This article examines the interrelationships between technological parameters, including cutting speed, feed rate, and depth of cut, and their influence on average power consumption. A statistical design of experiments was employed with the objective of developing a model that would enable a quantitative and qualitative description of the interactions in question. The insights yielded by these analyses facilitate the optimisation of energy consumption during the turning process, thereby reducing the environmental impact of the manufacturing process. In order to identify the optimal process parameters, preliminary experiments were conducted under three distinct conditions: dry, coolant lubricants, and idle. Furthermore, restrictions were defined with regard to rough and finish turning. At the same time the study examines the influence of minimum quantity lubrication on energy consumption, thus providing further insights into the domain of energy-efficient machining.

Keywords: energy efficient, turning, cutting parameter, restriction, design of experiments, coolant lubricants.

I. Introduction

Optimizing energy consumption in the engineering industry, as in other advanced manufacturing fields, is one of the important issues for increasing productivity and efficient use of resources. So far, various studies have been conducted to measure the effect of various parameters—for example, cutting speed, feed rate, and material removal rate—on energy consumption [1,2,3,4].

In order to develop effective strategies for improving energy efficiency, it is necessary to take a differentiated look at the interactions between the various process parameters and energy consumption. It is only through a comprehensive understanding of the interrelationships that targeted optimisation measures can be derived, which help to minimise energy consumption and enhance the sustainability of the turning process. It is of the utmost importance to adopt a systematic approach to the analysis of process parameters. It is well known in industry that the same process results can be achieved with different energy consumption [5]. This is made possible by varying process parameters and components. In order to successfully implement this constellation, it is

necessary to employ an energy-based process model. Nevertheless, the construction of an analytical model is a challenging endeavour, due to the multitude of systematic and random factors that exert influence. A practical solution to this problem is to use statistical experimental design methods. These methods make it possible to mathematically describe the relationship between the input and output variables of the process. Complex relationships between process parameters and energy consumption can be modelled using polynomials [6,7]. In this way, optimum constellations of process parameters can be identified in order to achieve the desired process results while minimising energy consumption.

The influences of different cooling lubricants on energy consumption were also investigated as part of this work. Although cooling lubricants fulfil important functions with regard to improving tribological phenomena in the cutting zone, they do not always lead to optimum results in terms of energy consumption [8].

II. Experimental setup and equipment description

The experimental investigations were conducted using a CTX 310 ecoline lathe from DMG Mori (Fig. 1). The DMG Mori CTX 310 ecoline is a computer numerical control (CNC) lathe that has been designed with the objective of facilitating precise and efficient turning operations. The machine is distinguished by its versatility, accuracy and reliability, rendering it suitable for processing a diverse range of workpieces within the manufacturing industry. The machine is equipped with a turret head, which enables the rapid and precise modification of the tools in use, thereby facilitating the overall process and allowing for greater efficiency. In addition, the CTX 310 ecoline offers options such as driven tools, programmable tailstock functions and advanced control to meet the needs of a wide range of turning operations.

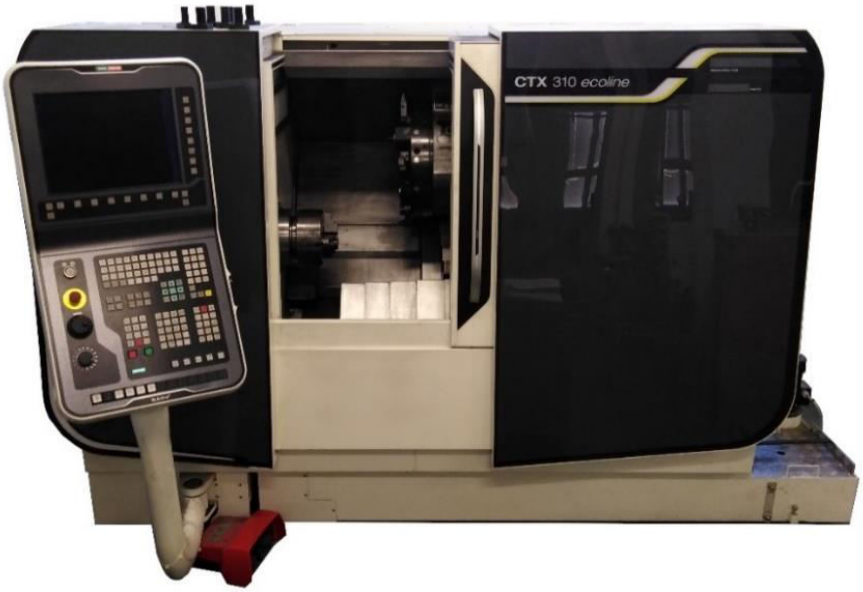


Figure 1: *DMG Mori CTX 310 ecoline lathe*

Fluke 434 Series II instrument was utilised for the purpose of measuring power. In accordance with the instructions provided, the Fluke 434 power quality analyser was connected to the three live conductors and the neutral conductor. These conductors and the neutral conductor supply power to the machine tool from the power grid. In order to facilitate the downloading and subsequent analysis of the data obtained from the Fluke 434 power quality analyser, the Power-Log software was employed as the appropriate tool for this purpose, given its suitability for the task.

III. Influence of cooling lubricants on energy requirements

Figure 2. illustrates the effects of various cooling lubricants on the average power. There are only minor differences among the cooling lubricants in terms of their impact on the average power. This suggests that the performance of the lubricants in preventing wear and adhesion may be similar.

Since there are only very small differences between these cooling lubricants, OZ 33 CBF was selected for comparison with dry machining due to its physical and chemical properties such as pH value, corrosion protection, and biocompatibility (Figure 3).

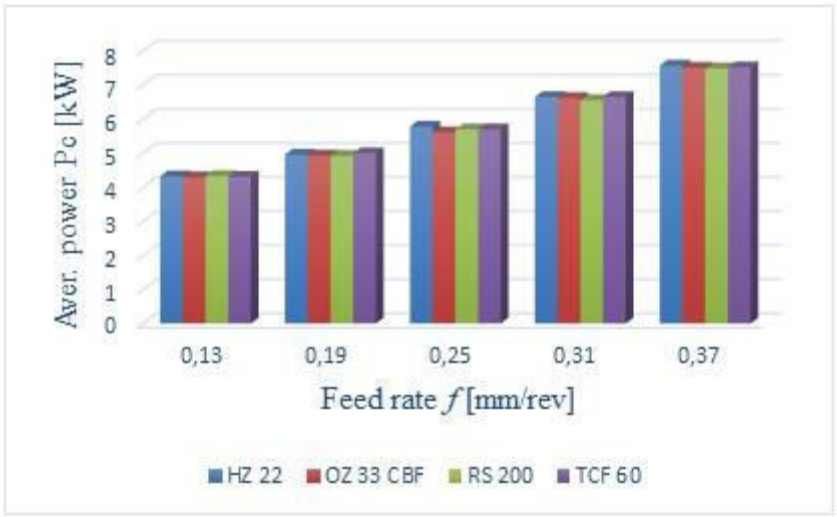


Figure 2: *The effect of feed rate on the average power at a cutting speed of V_c=200 m/min and a depth of cut a_p=2 mm in various coolants lubricants*

The observation from the experiments revealed that there was no significant difference in the average power between the use of cooling lubricants and dry machining (Figure 3).

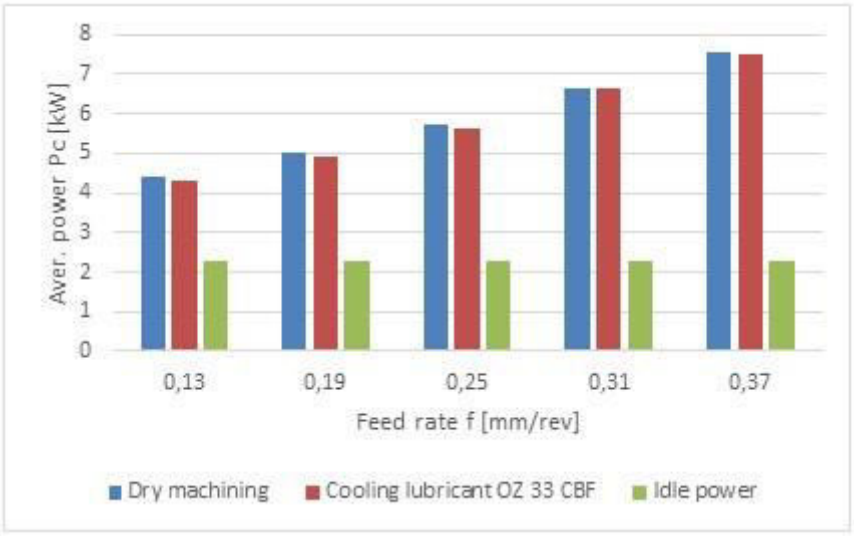


Figure 3: *The influence of feed rate on the average power at a cutting speed V_c=200 m/min and a depth of cut a_p=2 mm in dry machining, cooling lubricant and idle power*

In addition, there are very slight changes in the average idle power (no load power) when the cutting parameters change. Pumps for minimum quantity lubrication (MQL) systems typically require electrical energy to deliver and spray the lubricant into the cutting zone. The operation of pumps can therefore lead to additional energy consumption, impacting overall operating costs and environmental performance. In the experiments conducted, one liter of cooling lubricant costs nearly €18, making it relatively expensive. During the trials, the flow rate of cooling lubricants was also measured (Figure 4). In total, almost 1000 ml/s of cooling lubricant was used in each experiment. Typically, with minimum quantity lubrication (MQL), less than 50 ml of lubricant is used per hour, representing a significantly reduced amount compared to conventional methods such as flood cooling or spray cooling [9]. After each experiment, traces of cooling lubricant appeared on the tool and workpiece.

In conclusion, the minimum quantity lubrication (MQL) device may not be suitable for the MQL task, as the amount of cooling lubricant used significantly exceeded the typical value. The results also show that there was no significant difference between the use of cooling lubricants and dry machining. This suggests that, in this case, the minimum quantity lubrication was not effective and may not have provided the intended benefits.

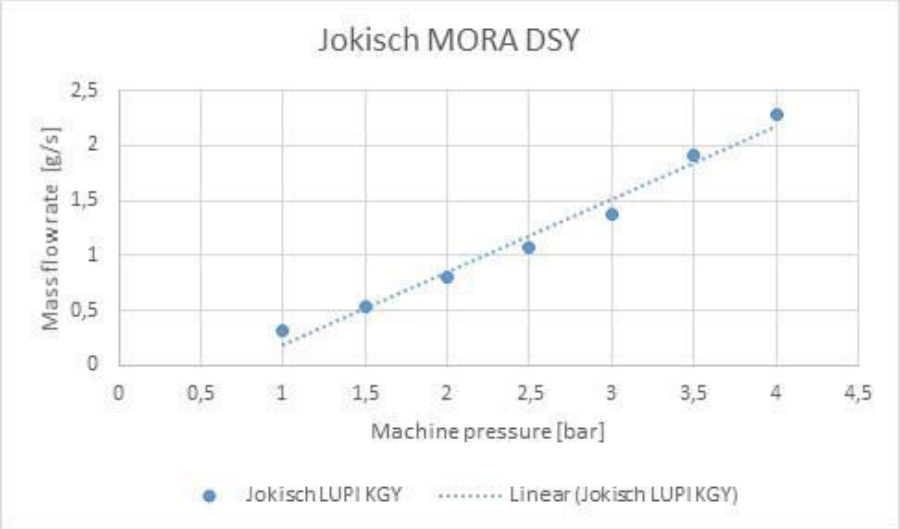


Figure 4: *Determination of the flow rate of cooling lubricants (OZ 33 CBF Jokisch)*

IV. Experimental design and methodology

Due to the non-linear nature of the parameter influences, employing a second-degree model is essential for accurate analysis. To achieve precise and reliable insights into the relationships defined in the statistical experiment design, centralized approaches such as the Central Composite Design (CCD plan) are recommended. This method offers a balance between efficiency and accuracy by requiring a limited number of test points while enabling clear and straightforward calculations to interpret the results effectively. These features contribute to its widespread adoption as a preferred second-order planning method [6,7,10]. The cutting parameter and other conditions values applied during the experiments are detailed in Table 1. The experimental design includes eight factorial points, six axial points, and ten center points. To assess the consistency and reliability of the process, repeated trials were conducted at the center points, ensuring robust conclusions regarding process stability.

Influencing variable		Dimension	-a	-1	0	+1	+a
1	Feed rate f	mm	0,0785	0,1	0,2	0,3	0,3215
2	Cutting speed V_c	mm	140,33	150	195	240	249,675
3	Depth of cut a_p	mm	1,34	1,5	2,25	3	3,16

V. Determination of restrictions for roughing

The process parameters in roughing are limited by machine and tool parameters. Roughing in machining is the process of removing large amounts of material from a workpiece in preparation for semi-finishing and finishing operations. High feed rates and large depths of cut are used to remove the excess material as quickly and efficiently as possible. On the other hand, as the feed rate and depth of cut increase, the mechanical load on the insert and the required power $P_c$increase. These process parameters are limited by the machine parameters. Therefore, power is one of the most important parameters for limiting the work value range.

When machining steel with an indexable insert, the maximum possible feed rate given by the technical limits of the machine tool can be utilized. Therefore, the technically maximum permissible cutting depths and cutting speeds were adopted to carry out the roughing process. Figure 5 shows a graphical representation of the restrictions when roughing with indexable inserts on the machine tool to determine the cutting parameters.

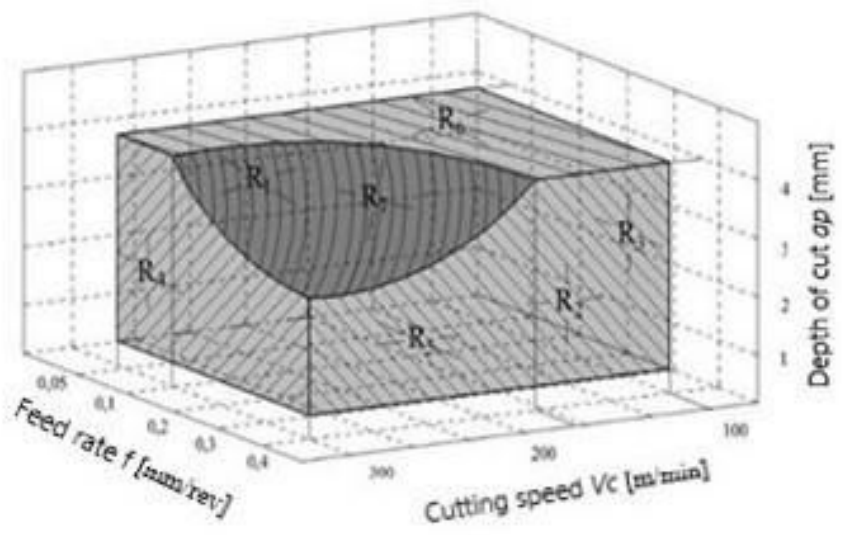


Figure 5: *Solution space for roughing with power restriction (P_c= 11 kW)*
The restrictions of the solution space presented here are determined according to the maximum possible parameter values: feed rate, depth of cut and cutting speed (Table 2)

Table 2: *Restrictions for roughing*

Restriction system		
Feed rate	$R_1: f \geq 0{,}05\ mm/rev$	$R_2: f \leq 0{,}4\ mm/rev$
Cutting speed	$R_3: V_c \geq 100\ m/min$	$R_4: V_c \leq 0{,}4\ mm/rev$
Depth of cut	$R_5: a_p \geq 1\ mm$	$R_6: a_p \leq 4\ mm$
Power	$R_7: f \cdot a_p \cdot V_c \leq (P_c \cdot 60000/k_c)$	

VI. Determination of restrictions for finishing

In the context of finish turning, permissible surface roughness is typically specified alongside limits for positional deviations of the workpiece. The working value limits, which must ensure the permissible roughness of the workpiece surface, already achieve satisfactory accuracy when the influence of the secondary cutting edge and the minimum chip thickness are taken into account [10, 11]. The primary cause of surface roughness lies in the imprint of the cutting edge or adjacent parts of the cutting edge on the workpiece surface. From this, geometric relationships can be derived, which are primarily dependent on the feed rate. In finishing operations, low feed rates are used to match the surface's form deviation. Regarding surface quality, the roughness limits of the feed rate must be considered during process optimization. The restriction on roughness in the feed direction can be derived based on geometric relationships [10].

The restriction of the feed rate is determined by the permissible roughness depth of the required workpiece surface. According to BAUER, the boundary equation for machining provides a good approximation for this purpose [11]:

$$f \leq \sqrt{8 \cdot r_\varepsilon \cdot R_{th}} \tag{1}$$

Here r_ε – is tool corner radius, which influences the surface finish and cutting stability and R_{th} is allowable surface roughness threshold, which defines the maximum roughness acceptable for the machined surface. The roughness restriction (Eq. 1) limits the solution space for optimal parameters. Figure 6 illustrates this using the example of machining with tool having r_ε = 0,4 mm and R_{th} = 0,02 mm. The calculation of the feed rate according to Eq. (1) results in a value of ffinish = 0,25 mm.

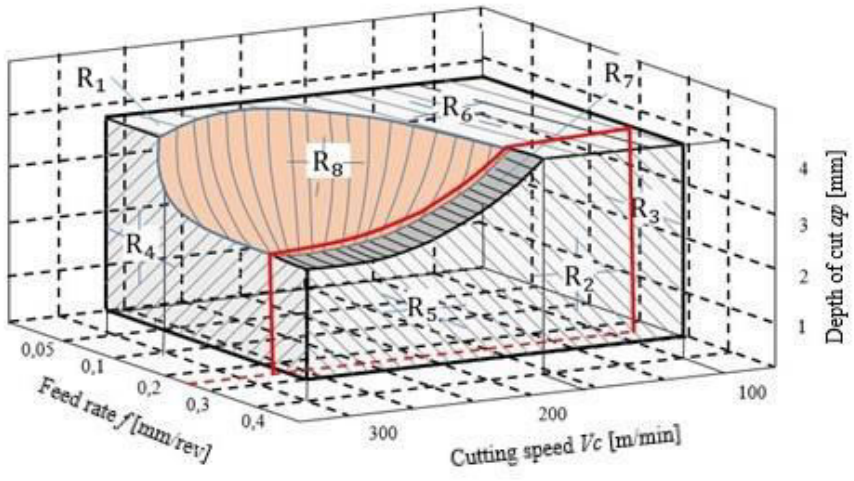


Figure 6: *Solution space for finishing with roughness restriction (red line)*

In this solution space, the relationship between the optimal feed rate and other machining parameters is shown to achieve the desired surface quality. The graphical representation of the solution space, based on previous systematic investigations, provides a visual depiction of potential settings and aids in selecting optimal parameters for the finishing process.

Another restriction in finishing is the deformation of the component. This requires defining the solution space in the optimization model while considering the permissible deflection of the workpiece. When machining workpieces with low stiffness, high cutting parameters can lead to elastic deformations, as these workpieces tend to deform more under load compared to stiffer ones. High cutting forces, particularly at high feed rates and/or cutting depths, can significantly increase the deformations, affecting the dimensional accuracy and surface quality of the machined part [10]. Using Eqs. (2), the restrictions due to elastic deformation of workpieces can be derived [10]:

$$f \leq \sqrt[1-m]{\frac{\Delta \cdot (\sin\gamma)^{m} \cdot 3E \cdot I}{a_p \cdot k_{c1.1} \cdot l^3}} \tag{2}$$

Here E – modulus of elasticity of the material, reflecting stiffness and I is moment of inertia of the workpiece cross-section, indicating resistance to bending. Besides that l is length of the workpiece and $k_{c1.1}$ is specific cutting force, often used to describe material resistance to cutting.

To keep deviations during finishing within the permissible range, the deflection of the shaft, which can lead to a form deviation, must be calculated. Since the feed rate is fixed, the deflection is calculated for various combinations of cutting depth and cutting speed according to Equation 2 and is presented in Figure 7.

As demonstrated in Figure 7, the deviations are less than 0.018 mm, which can be attributed to the high rigidity of the workpiece. This value is within the tolerance range for turning operations. Consequently, the restriction concerning positional deviation cannot be considered when searching for optimal cutting parameters in the examined case. However, this does not exclude its influence on process optimization. For machining less stable workpieces, considering this restriction is essential.

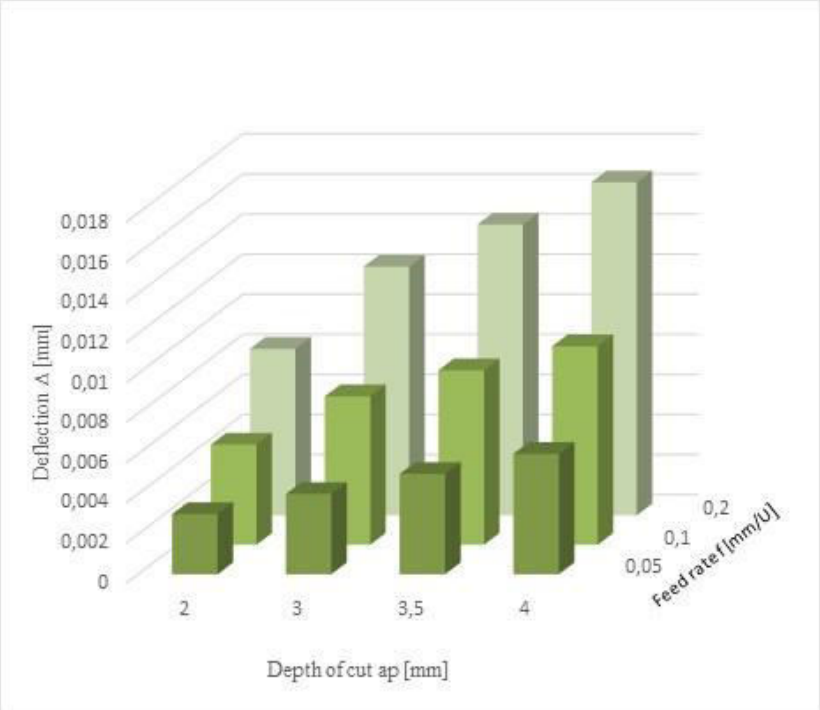


Figure 7: *Strength-related restriction of the feed rate*

Technological influences on power consumption

Using statistical experimental design, a model was created that describes the relationship between average power and the process parameters f, V_c and a_p in both numerical and analytical terms.

MODELL: $y = 0{,}89 - 20{,}04 \cdot f - 4{,}24 \cdot a_p + 0{,}0474 \cdot V_c + 8{,}533 \cdot f \cdot a_p + 0{,}08 \cdot f \cdot V_c + 0{,}0083 \cdot a_p \cdot V_c + 0{,}604 \cdot a_{p_2}^2 - 0{,}000158 \cdot V_{c_2}^2$

The significance of this polynomial is highlighted by displaying the results of the y-model as a response surface. The results show that the main quantitative factor influencing the average power is the cutting speed and its interaction with the feed rate and depth of cut [12]. Additionally, the required average power and other key optimization parameters, such as cutting speed, play a critical role in determining processing time. For this reason, it is essential to ensure that these parameters are carefully optimized and kept within an ideal range to achieve efficient machining while maintaining desired performance outcomes.

VII. Conclusions

The study investigates the impact of technological and non-technological factors on the energy consumption in machining processes, focusing specifically on turning operations. A comprehensive methodology was developed to quantitatively and qualitatively evaluate these influencing factors. Experimental verification was conducted to measure the effect of individual input parameters on process outcomes, laying the foundation for identifying technological and technical constraints within the machining system. Using statistical experimental design, systematic data collection facilitated the development of mathematical models that elucidate the relationship between process parameters and average power consumption. This approach ensures a robust understanding of parameter influence, enabling targeted optimization of the turning process for improved energy efficiency and reduced manufacturing costs.

Besides that the experiments revealed no significant differences in average power between the use of cooling lubricants and dry machining, indicating that the cooling lubricants tested offered no substantial advantage in this case. Furthermore, the minimum quantity lubrication (MQL) system used excessive amounts of lubricant, far exceeding typical values, which undermines its effectiveness and cost-efficiency. These findings suggest that MQL was not suitable for the task and did not deliver the expected benefits.

References

[1] Zhang, Y.; Liu, Q.; Zhou, Y.; Ying, B.: Integrated optimization of cutting parameters and scheduling for reducing carbon emissions. J Clean Prod 149, 886-895. (2017)

[2] Qianqian, Z.; Renzhong, T.; Tao, P.: Decision rules for energy consumption minimization during material removal process in turning. Journal of Cleaner Production 140, 1819-1827, Elsevier, (2017)

[3] Kant, G.; Sangwan, K.S.: Prediction and optimization of machining parameters for minimizing power consumptions and surface roughness in machining. J.Clean. Prod. 83, 151-164, (2014)

[4] Camposeco-Negrete, C.: Optimization of cutting parameters for minimizing energy consumption in turning of AISI 6061 T6 using Taguchi methodology and ANOVA. K. Clean. Prod. 53, 195-203. (2013)

[5] Aliyev R., Hentschel B.: HSC – Fräsen von stäubenden Werkstoffen. Strategien zur Auslegung von Prozess und Komponenten. ISBN 978-3-662-63692-3. https:/doi.org/10.1007/978-3-662-63693-0. Springer Vieweg Berlin, Heidelberg, (2021)

[6] Siebertz, K.; Bebber, D.V.;Hochkirchen, T.: "Statistische Versuchsplanung", 2. Auflage,

Achen: Springer Vieweg, (2017)

[7] E. Scheffler.: "Einführung in die Praxis der statistischen Versuchsplanung", VLN 152-915/49/74, (1974)

[8] Jersak, J.; Simon, S.: Influence of cooling lubricants on the surface roughness and energy efficiency of the cutting machine tools. Int. J. of Applied Mechanics and Engineering, vol 22, No.3, S.779-787, DOI: 10.1515. (2017)

[9] Ramazan, H.N.; Sadık, E.K.; Bahram, L.S.: Investigation of the effect of minimum quantity lubrication technique on cutting forces in rough machining of Ti-6Al-4V alloy. Machine design and manufacturing magazine, Volume 17, Issue 2, (2020)

[10] Jacobs, H.J.: Spanungsoptimierung. VEB Verlag Technik Berlin. (1977)

[11] Jacobs, H.J.; Dürr, H.: Entwicklung und Gestaltung von Fertigungsprozessen. Planung und Streuung der spanenden Teilefertigung. Fachbuchverlag Leipzig im Carl Hanser Verlag. (2002)

[12] Abbasov I.; Suleymanov I.; Mammadov A.: Modelling of turning process for prediction of energy consumption. Key Engineering Materials, ISSN: 1662-9795, Vol. 979, 19-25 (2024)

WEAR PROPERTIES OF CAMSHAFT CAMS AND IMPROVEMENT OF THEIR WEAR RESISTANCE

Vaqif Abbasov[1], Fariz Amirov[1], Azad Karimov[1]

•

[1]Azerbaijan Technical University, Huseyn Javid Ave.-25, Az-1073, Baku
vaqif.abbasov@aztu.edu.az, fariz.amirov@aztu.edu.az,azad.kerimov@aztu.edu.az

Abstract

In the article, the problems of increasing the durability and wear resistance of the cam of gas camshaft by using the ion implantation method were considered. In operation, since the cams of camshaft work the working profile under high pressure, the wear is fast. As a result, the regular operation of the engine is disturbed and the engine does not provide the necessary power and force at the output. It is known that the gas camshaft mechanism requires complex and quite economic costs in the repair of engines. In this regard, increasing the wear resistance of camshaft using the ion implantation method is appropriate. To conduct the research, samples of steel 40X, 50X and X6B3MTiC used in the production of camshaft were prepared and researches were conducted.

Keywords: internal combustion engines, camshaft, ion implantation, coating, wear resistance.

I. Introduction

Camshafts with cams used in internal combustion engines (ICE) of cars, ships and other transport vehicles have a complex construction, and during their operation, uneven wear occurs on the working surface of the cams as a result of moving friction of the valve pusher in their working profile. The uneven wear observed on the working surface of the pistons has a negative effect on the efficiency of the ICE work process, the cycle of supplying the fuel mixture to the cylinder and removing burnt gases from the cylinder. At the same time, the analysis of the construction of gas camshaft shafts with additional cams shows that gas distribution shafts with different constructions, including whole and differently assembled cams, are used in order to reduce cams wearing in ICE of different sizes and powers [1-4].

The wear of cam shafts during operation depends on a number of factors:

1. The gas distribution shaft should be constructed in such a way that the kinematic relationships placed in front of them, i.e., the nature of the contact between the cam and the pusher, work with a regular trajectory of movement corresponding to a sinusoidal smooth curve and minimize the impact of striking forces, etc..:
2. The kinematic connections must be designed so that the intake of the fuel flow of the valves and the expulsion of the burnt mixture from the cylinder can be performed with high precision.:
3. The material of the cam relative to the pusher should be selected in such a way that its wear is compatible with the physical and mechanical properties of the material that affects the wear character of the contact pair.

By examining the mentioned issues on the basis of scientific analysis, it is possible to reduce the wear of the working surfaces of camshafts and, as a result, to increase the operating efficiency of

internal combustion engines. For example, the problem of increasing the durability of cams against wearing should be solved by using modern coating technologies, the issue of using cams equipped with wear-resistant coatings, so that it is possible to extend their service life.

The theoretical and practical solutions to the above-mentioned technological issues were investigated and the development of new technologies was considered.

Purpose of work. Improving the operational quality of the available integral and assembled gas distribution shafts is aimed at reducing the intensity of wear of working profiles of cams and the linear size of wear of this surface. A new construction of the gas distribution mechanism was developed in order to make the repair of the gas distribution mechanism during the operation period without disassembling the engine and to achieve a high result. At the same time, coating is applied to the surface of more cams at the same time in the chamber of the coating device.

II. Research methods

The accuracy of the geometric dimensions of the prepared cams was studied with the help of a transfer polishing machine using special designs and watch-type indicators with an accuracy of 1 μm. The hardness of the surface after quenching was studied using the Portable Hardness Tester Tm 210 hardness measuring device. The thickness of the coating applied by diffusion was studied with the help of the Tm-8812C ultrasonic thickness measuring device. After polishing, the roughness of the surface of the cams was measured using a TR220 profilometer-profilograph. The resistance to wear of prepared cams was experimentally studied depending on the friction path of 10,000 km at the contact pressure of 250 kN at the base of the İİ5018 friction machine. The camshaft installed on a Kamaz truck for the purpose of testing under real operating conditions continues to be in working condition by completing 160,000 km mileage. The powder profile of the cams subjected to friction test was studied by dividing them into 5°- angles. The essence of the measurement methodology is that the profile of the cam is 3 mm apart, and the graph of the wear is made by measuring at 360 points, one every 5°-on the profile, in five cross-sections.

III. Discussion of research conducted.

The wear of the cam shafts of the internal combustion engines of KAMAZ 740 trucks, which is one of the widely used vehicles in the transport system, depending on the service life, was studied. In this machine, the role of the rocker in the kinematic relations was considered in order to reduce the cams wear. Also, ways of reducing the intensity of wear cams made of different materials are being investigated.

It has been determined from the researches that the nature of the contact between the cams and the pusher is caused by one of them, i.e. the rotation of the cams, and the other as a result of the back and forth movement of the pusher. Wear of the cam's profile depends greatly on its speed of rotational movement, as well as the force acting on the profile, the speed of back-and-forth movement, and shocks due to the increase in clearances caused by contact. Therefore, the appropriate material was selected, which has a positive effect on the reduction of cams wear in the first place. In the studies, cams were made from the following materials: steel 40X, 50X, X6B3MTiC.

The technology of making cams is based on their physical and mechanical properties. Thus, in the selection of the material, taking for example 45X, 50X from low-temperature annealed medium carbon alloy steels is related to the efficiency of their machining process and wear resistance. Their high hardness and resistance to wear make it possible to use instead of cemented steels. Thus, instead of the expensive and long-term cementing process, it is possible to obtain the required hardness on the surface of the cams made of steel 40X, 50X materials by using a low-labor and productive high-frequency induction heating method.

Taking these into account, a new assembly structure of the camshaft was created. [2]

The purpose of the new construction is aimed at increasing the service life of camshafts,

reducing the wear of cams and improving its production process, simplifying the repair process.

The new construction was solved in several ways as follows. The camshaft installed in the assembly construction consists of a stepped shaft with a diameter difference between the cam journal and cams heels, the cams fitted very tight above shaft key fastened with different method. That is, the purpose of the construction is to cancel the fastening scheme on the top of the cams and the central bolt, and the fastening of the cams on the cam heel is carried out according to the different schemes given below.

The cams is fitted on the cam heel in a form based on its side surfaces and is fixed by means of hole is opened on the side surface, and it is fixed by riveting to the hole. (Figure 1)

In the mounting scheme of the camshaft with the proposed assembly structure, the impact force acting on the cams has a relatively small effect on the pressed fastening axis. Since the external-side surfaces of the fixing axis are made in accordance with the profile of the cams, the contact with other parts that may occur during the rotation of the camshaft is smooth.

According to the second fastening scheme, the cams is fixed by means of a bolt with a threaded hole with a bottom in the cam heel, which facilitates the replacement of wearing cams in the repair areas. In both proposed fastening schemes, the normal and tangential forces on the cams, as well as other forces, are cross-distributed between the frame and the bolt, thus preventing the formation of gaps in the fitting surfaces of the cam.

In all proposed fastening options, the working profile of the cams, the top has a smooth geometric profile, and its change will be subject to wear during operation, depending on the material, construction and other factors.

The design of the cams made in the new versions allows the camshaft to be reused for a long time by changing them periodically after wearing. The production of cams with the proposed fastening scheme can be applied to the assembly of camshafts using sufficiently available technological processes: drilling, milling, cutting, heat treatment and polishing operations [1,5,6-8].

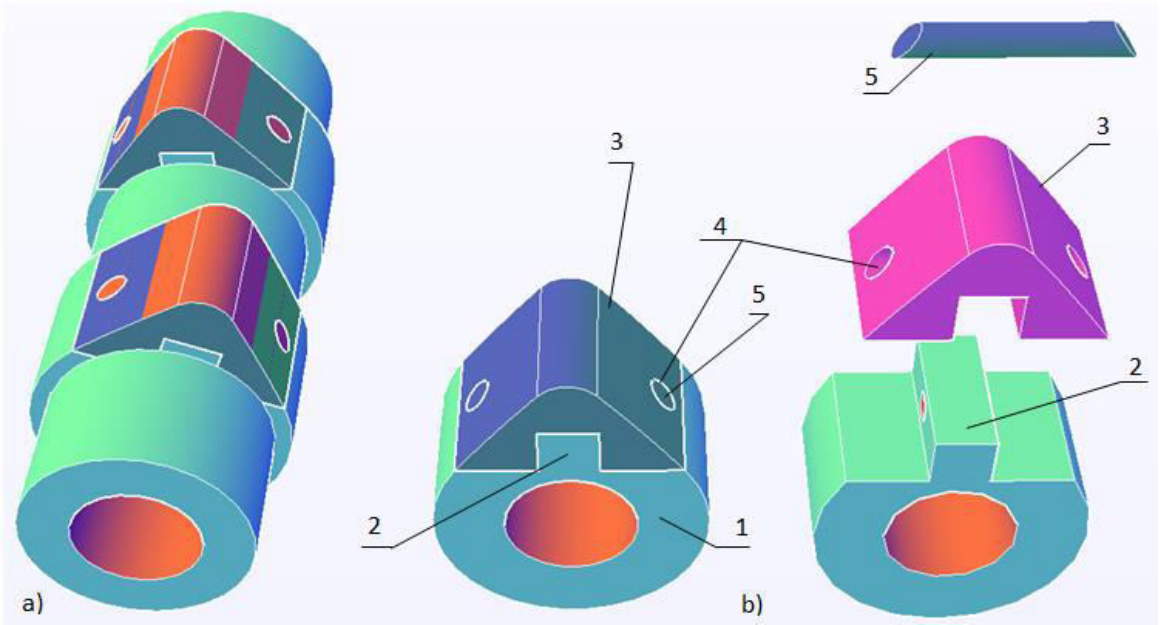


Figure 1: *a shows a schematic view of part of the camshaft. b shows the construction of camshaft 1 cam heel and its elements. Here, 1- cam heel, 2-shaped surface of shaft key, 3-cam, 4-hole opened in the direction perpendicular to the shape of the cam, 5-axis connecting cam with cam heel are extra*

According to the first fastening scheme (Fig. 1.), the cams (3) is fitted on the figure in the cam heel (1), then it is connected to the cam heel by press fit through the holes (4) opened in the cam and the shift key. The pressed axis (5) limits the displacement of the cam in the direction perpendicular

to the axis and axis of the camshaft during the working process, which leads to reduction of shocks and noises that may occur during operation of the engine.

It is known that the cams is loaded more than other structural elements of the camshaft. It is also suggested to make the cams from harder, wear-resistant, abrasive materials that withstand cyclic loads, for example, X6B3MTiC [1,5,6-8]

Making technology of the cams of row material made of X6B3MTiC powder material is as follows

1. Cold pressing to 45-50% density;
2. Pre-sintering in argon or nitrogen atmosphere at a temperature of 600-7000C;
3. Final sintering in a vacuum for 30-60 minutes at a temperature of 1320-13800C and a residual pressure of 1...0.1 MPa

The composition of the material produced by X6B3MTiC sintering technology is given in table 1:

Table 1

Content		HRC hardness				Strength limit, MPa		Density obtained as a result of sintering
Carbide steel		After sintering	After anealling	After quenching	After anealling	TM_{H}	$TM_{сж}$	%
X6B3M 90%	TiC 10%	47	15	57	56	1320	2500	98-99

With the continuous cooking technology, it is possible to obtain the necessary sizes and shapes of the cams lobe from the material of the press-forms, keeping a small amount of processing for polishing with a small labor capacity and high productivity. By applying the above mechanical properties, it allows to increase the service life of the cams.

Production of cams from steel 40X and 50X material was carried out by two methods. After the steels taken in the first method, the whole structure of the gas distribution shaft is prepared by forging the cams, after passing through the appropriate mechanical processing process, the surface is subjected to a surface treatment within the limits of HRC 55..58 with a high frequency current, and the profile is polished on the transfer polishing machine. The resistance of these cams to being wear is being studied. [7-9]

In the second case, after the ends of the cams made of steel 40X and 50X have undergone mechanical processing and the surface is subjected to high-frequency current treatment with a depth of 1.5...2 mm, the support surface is subjected to HRC 52..55 limit and diffusion chrome plating process in the CHB 1.3.1/6 vacuum oven. it is intended to increase the resistance to wear of the working hardness of the cam. The thickness of one layer of the applied coating varies between 20÷30 μm. If required, by increasing the number of coating layers, it is possible to increase the thickness of the coating layer and increase the life of the cam.

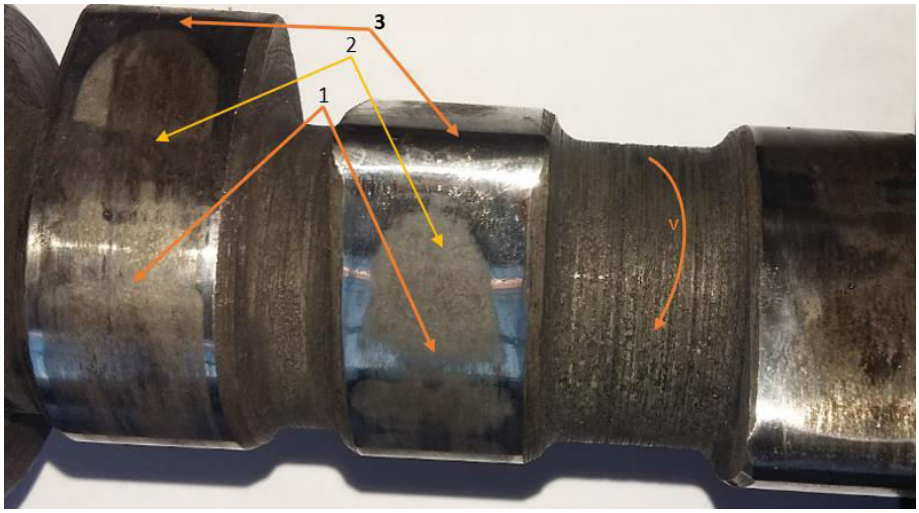


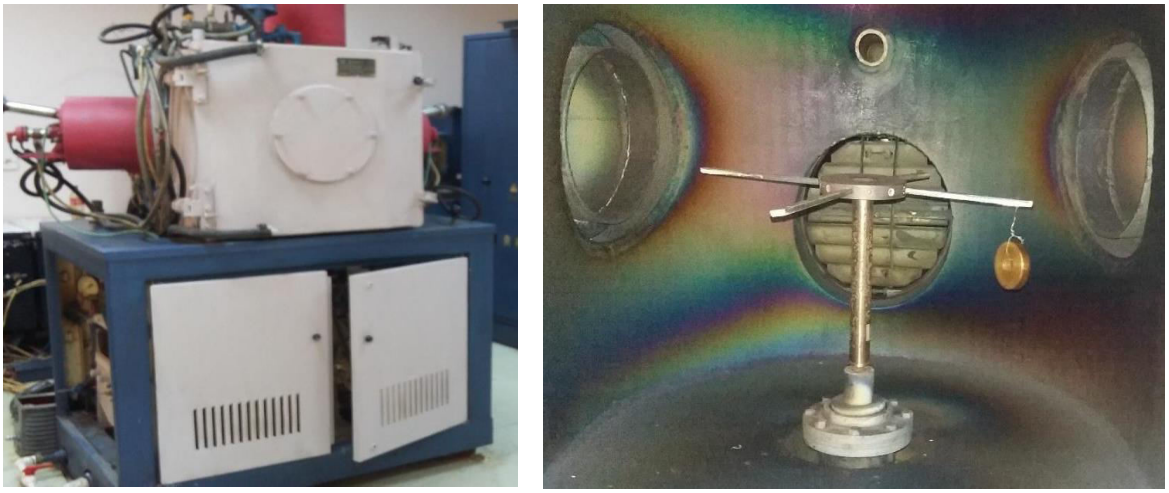
Figure 2: *Wear characteristics of the working surface of a camshaft with a single-piece design*

Wear was studied by taking the cams of the gas distribution shafts manufactured by the existing industrial method and used in the Kamaz 740 vehicle from different materials, for example, steel 40X, steel 45. After taking the steel 45 material in the martensite structure and undergoing thermal treatment, the punches with a hardness of 56÷58 HRC were examined for exploitation and experimental wearing in laboratory conditions. Figure 2 shows the behavior of these cams after they were used in real conditions on the Kamaz 740 vehicle.

From the study of the used distributor camshat, it can be seen that the smoothness of the transition from its surface 1 to the working surface of the cams to ensure the opening of the valve in the working process with the rotation speed v is disturbed due to wear. The study of the linear dimensions of the wear on the surface of 1 (from the point of creation of the tension medium between the valve pusher and the cam) in different areas shows that the depth of the wearing surface varies in the range of 0.15-0.25 mm. Due to the increase in the pressure caused by the effects of the return spring of the valve and the pusher compared to the 1st surface, the amount of wear of this area has increased by 35-40% compared to the 1st surface. Since the 1st and 2nd areas of the cam serve to open the valve, the wear occurring in this area initially reduces the opening distance of the valve and causes the formation of a stepped surface at the top of the cam in the 3rd 1.7° area, i.e. in the transition to the valve dwell time, which at the moment of full opening of the valve, under the influence of additional force and increased frictional moment, it causes sudden critical loading of the camshaft and the generation of noise in the engine. Regardless of the hardness obtained from the thermal treatment of the surface of the cams of the current construction, despite the fact that the pressure generated at the top of the cam during the operation of the cam is at its maximum value, the wear occurring in areas 1 and 2 of the profile of the cam is not less than the wear at the top of the cam. There are various factors that influence this: uneven distribution of hardness of the surface of the cam, irregularity of residual stresses generated on the working surface of the cam, roughness of the surface of the cam, etc. [1,4].

Depending on the thickness of the coating layer to be applied to the surface, the coating method is selected. During the application of each method, a set limit is determined for the quality and reliability of the coating to be applied.

Taking this into account, in order to increase the wear resistance of cam and ensure the surface hardness of its working profile, the following studies were conducted on the НИВ-6.6.И1УХЛ4 model Ion implantation device.



a) b)

Figure 3: *Ion implantation device model HHB-6.6.И1УХЛ4*
a) general view, b) working chamber of the device

1. Deposition of the XrN coating on the samples of the working surface of the cam in different modes;
2. Deposition of the ZrN coating on the working surface of the cam on the samples in different modes;
3. Deposition of the TiN coating on the working surface of the cam on the samples in different modes;

On the basis of the ion implantation device model HHB-6.6.И1УХЛ4, the purpose of carrying out the coating material on special samples in different modes is to conduct a comparative analysis of the surface hardness of the practically obtained coatings and the thicknesses of the obtained coating.

IV. Conculicion

1. Based on the research, it can be concluded that the cam shafts are subjected to uneven wear on the surface due to the maximum load at the initial start-up of the engine and high frequency friction at operating temperature.
2. As a result of the literature research, the maximum amount of wear on the surface of the cam can reach 300 μm. At the same time, there are deposits of 90÷150 μm in size on the surface of the cam, which ensures the time of staying away, which is caused by the compression of the pusher and the cam with the maximum contact force.
3. Taking into account the conditions of the impact of the gas camshaft cam, the new construction was studied in operation and made it possible to achieve a significant result in the direction of improving the efficiency of the repair.
4. New technological mode parameters and an improved oven design were developed for conducting ion implantation of the wear surfaces of the gas camshaft cam of ICE. The efficiency of the coating method carried out in new technological conditions and in an improved oven has been confirmed

References

[1] Karimov A.F. Improvement of construction and production of cam shafts. Innovative equipment and technologies in machine engineering. St. Petersburg Scientific Research Center for MS (St. Petersburg). May 2024. Volume: 17. p.151-154. DOI: 10.26160/2307-342X-2024-17-151-154

[2] Camshaft and its preparation method. Abbasov V.A., Karimov A.F., Sadıxov A.H, Amirov F.Q. Intellectual Property Agency of the Republic of Azerbaijan. Patent İ2024 0004 01.02.2024

[3] Camshaft and its preparation method. Mammadov A.M., Sadikhov A.I., Mammadov A.T., Musaev Y.B. Azerbaijan patent, N a20070261. - 2008.

[4] Vagif Abbasov, Fariz Amirov, Azad Karimov. Features of the New Design of the Camshaft of Internal Combustion Engines. International Symposium on Unmanned Systems: AI, Design, and Efficiency ISUDEF '24 Abstract Book// Abstract Book-National Aviation Academy. 2024.p 27. DOI: 10.30546/2224.978-9952-582-04-8

[5] Karimov A.F. Analysis of operating conditions of the main part of the gas distribution mechanism. Materials of the 52nd student scientific and technical conference dedicated to the 86th anniversary of Heydar Aliyev, AzTU, Baku, 2009.

[6] Vagif Abbasov, Fariz Amirov, Samir Amirli, Shalala Hasanli, Karel Frana, Formation of Shaft Accuracy during Mechanical Processing on CNC Machines //Advances in Science and Technology. Trans Tech Publications Ltd, Switzerland. 2024. Vol. 148, pp 81-86. doi:10.4028/p-N5uKb8

[7] Farhad Shirzadov, Azad Karimov, Ramin İsgandarov. Tribotechnical analysis of cams control mechanism and application of a wear-resistant layer on the surface of camshafts. Machine Science. September 2015.Vol4,1

[8] Xiaona Shi, Kelong Wang, Guochao Li, Chenghao Lyu, Lei Zhao*, Jianzhi Chen, Li Sun and Hengheng Wu.Study on Temperature Field Uniformity of Dynamic Induction Heating for Camshaft of Marine Diesel Engine//Machines. March 2024

[8] https://iopscience.iop.org/article/10.1088/1742-6596/1273/1/012066

[9] Irgashev A., Hamrouev R.K. Features of fuel economy of agricultural equipment and environmental influence on oil // Technical science and innovation. - Tashkent, 2021. - No. 1 (07). – P. 297–302

PROCESSING OF HIGH-SPEED STEELS BY PULSED LASER RADIATION

Igor Savin[1], Isag Khankishiyev[2], Asim Mirzayev[3], Jeyhun Rahimov[3], Rufat Abbasov[3], Gunay Dadashova[3]

•

[1]Department of Design and Technology of Machine-Building Industries, Kazan National Research Technical University named after A.N.Tupolev – KAI, Kazan, Russia,
[2]Department of Shipbuilding and Ship Repair, Azerbaijan State Marine Academy, Baku, Azerbaijan
[3]Azerbaijan Technical University, Baku, Azerbaijan
1705kzn@mail.ru, isaq.xankishiyev@asco.az, asimmi@aztu.edu.az, jeyhun.rahimov@aztu.edu.az, rufet.abbasov88@mail.ru, gunay.dadashova@aztu.edu.az

Abstract

The paper considers the issues of improving the performance of cutting tools made of high-speed steels. The analysis of modern methods of changing the properties of the surface layers of tools according to the specified parameters is carried out. Currently, about a hundred types of hardening technologies of cutting tools are known to have been developed. Almost all performance criteria are determined by the properties of the surface layers. It is shown that the achievement of a given set of properties of surface layers is possible not only by conventional methods, but also by laser treatment. Various options for laser surface treatment of cutting tools made of high-speed steel are considered. It is determined that pulsed laser radiation has a number of advantages for such processing compared to conventional processing. Based on the results of the study, optimal modes of pulsed laser processing of high-speed steels were established, carried out without melting the surface.

Keywords: High-speed steels, Laser, Hardening technologies, Laser radiation.

I. Introduction

Increasing the life of the cutting tool is an urgent task of modern surface treatment methods. The surfaces of the friction pairs work in conditions of high temperatures, aggressive environments, and at the same time experience cyclic, alternating and shock loads.

In this case, the surface layers of the material are subjected to the processes of deformation, deformation, oxidation and destruction. The strength and geometric properties of the surface are the key factors determining its wear resistance. High-speed steels are widely used for metal-cutting tools [1], due to their higher viscosity and strength compared to other tool materials. Surface treatment is an integral part of the technological process of hardening machine parts, equipment and machining tools, which allows you to increase durability and wear resistance by 2 to 5 times [2]. To increase the life of cutting tools, well-known technologies are used aimed at improving the performance of already known tool materials by modifying its surface layer by processing high-energy beams and applying wear-resistant coatings. With the help of such processing, it is possible to significantly change the mechanical, electrical, thermal and chemical properties of the initial tool material, its real surface.

II. Research

One of the most promising ways to improve the performance of a cutting tool is to strengthen its working surfaces, while achieving a combination of high strength properties of the base material with hardness and heat resistance of the surface layer. Currently, about one hundred types of hardening technologies of cutting tools have been developed. Almost all performance criteria are determined by the properties of the surface layers.

To increase wear resistance, traditional methods of hardening are widely used: plasma-arc methods of applying composite coatings, mechanical methods of surface treatment, thermal and chemical-thermal treatment [2]. Since the durability of the tool material is an integral property that depends not only on viscosity and strength, but also on hardness, heat resistance and wear resistance, attempts have been made to harden the surface layer of high-speed steels using the following methods: chemical (CVD) and physical (PVD) deposition from the vapor or plasma phase, as well as chemical-thermal and laser surface treatments. The greatest interest at the moment is caused by laser treatment, which can be used not only for hardening, but also for local restoration of worn surfaces of tools [3]. The following methods of using a laser for surface treatment are known: thermal, which ensures a high rate of phase transformations in the solid state and the formation of highly supersaturated solid solutions; surface melting followed by accelerated crystallization of the melt and the formation of fine crystalline nonequilibrium phases; surface alloying, for example, laser cladding or laser remelting. Today, laser treatment is widely used to harden tool steels and alloys [4].

To improve the performance of cutting tools made of high-speed steels, the use of laser technologies is a very promising direction, among which laser hardening occupies an important place. The advantages of laser hardening are due to the possibility of supplying a high concentration of energy to the processing zone, ensuring locality of impact on the hardened surface of the tool, environmental cleanliness, as well as great automation and process control capabilities. Laser quenching is carried out outdoors without the use of vacuum or quenching media [5].

High-speed local laser heating followed by rapid cooling into the volume of the material makes it possible to obtain a finely dispersed nonequilibrium structure with a wide range of properties. At the same time, there is no mechanical impact and warping of the processed product, which, combined with easy automation and fast payback of equipment, makes laser surface hardening technology a worthy competitor to traditional processing methods.

Research in this area is aimed at improving the adhesion of the applied layer, increasing the hardness of the coating and the modified zone, and reducing residual stresses in the modified layer [6]. However, despite the development of a number of hard ceramic coatings resistant to various types of wear, having very high hardness and low coefficients of friction, intensive destruction during plastic deformation of the tool base cannot be avoided. In most cases, the destruction of the coating-substrate system begins with plastic deformation of the substrate near the interface, when this system is subjected to severe stress. Thus, the load resistance in the coating-substrate system also depends on the properties of the substrate. In order to improve the performance of products with wear-resistant coatings, it is necessary to balance the difference between internal stresses in the boundary layers of the substrate and in the coating without significantly reducing hardness and wear resistance. For this purpose, combined tool processing is increasingly used [7].

To improve the performance of cutting tools with high-speed steel coatings, methods of complex surface treatment are used, combining the processes of ion-plasma coating and surface hardening treatment. Pre-surface treatment helps to reduce the tendency of the cutting wedge of the cutting tool to elastic deflections and loss of shape stability, which helps to improve the performance of cutting tools. At the same time, such a preliminary surface treatment does not significantly affect the strength of the coating material and the adhesive bond strength of the coating with the tool base, which, as mentioned above, affect the process of coating destruction. An increase in these characteristics can be achieved by additional hardening treatment after coating [8].

The creation of structurally nonequilibrium states in the surface layer of the tool during their

pulsed electron beam alloying seems to be a promising treatment of cutting tools as part of a wear-resistant complex before applying a wear-resistant coating.

In the last decade, the use of concentrated energy flows has become widespread, which make it possible to transfer high densities of absorbed energy (tens of J/cm2) to the material in relatively short (less than 100 microseconds) time intervals. In particular, research in the field of modification of the cermet surface by a pulsed electron beam has made serious progress. It has been established that with such processing, it is possible to obtain a multiple increase in operational characteristics due to a complex of modifying effects due to thermal, thermomechanical and diffusion phenomena [9]. The treatment is accompanied by a transformation of the phase composition and defective substructure of the surface layer, which consists in the formation of a submicron and nanoscale grain structure (100...200 nm). Due to the high temperature gradient and thermomechanical stresses, cascades of displaced atoms and a large concentration of structural defects occur, which, apparently, is the cause of the observed segregation phenomena and intense diffusion flows of atoms. At the same time, a variety of hardening mechanisms in a layer up to 10 microns was revealed.

The main factors of electron beam irradiation of materials that determine the temperature profiles of the surface layer heating zone, and, accordingly, the nature and kinetics of structural-phase transformations, are the energy density in the electron beam, the duration, number and frequency of irradiation pulses [10] [11]. An increase in the duration of the electron irradiation pulse leads to a decrease in the heating temperature of the irradiated surface with a noticeable increase in the depth of heating of the layers, which reduces the intensity of remelting of the composite structure. An increase in the energy density in the electron beam leads to an increase in the heating temperature of the irradiated surface and to a significant increase in the heating depth. An increase in the number of irradiation pulses also increases the depth of warming up of the instrument, but for each value of the energy density in the electron beam there is a limit value of the heating zone from the irradiated surface [12] [13].

The use of pulsed radiation for these purposes provides a greater effect than the use of continuous radiation, since the degree of hardening during pulsed laser hardening is greater than with continuous. This difference in the degree of hardening is explained by the fact that with pulsed laser hardening, the heating rate is 2 orders of magnitude higher than with continuous.

The influence of the power and duration of the laser radiation pulse on the microhardness of P6M5 steel coated with titanium nitride and uncoated has been studied. The surface laser treatment was applied to tool materials that had undergone heat treatment and final grinding. A Russian-made laser installation based on an IPG pulsed solid-state laser, widely used in industry, was used, made in the form of separate units: a machine with an optical-mechanical unit, a feed mechanism and a synchronizer; a modulator, which includes a current source, a storage device, a combined unit and a cooling device (Fig.1). The installation is used in surface treatment (hardening) of tool steels, as well as thermally non-hardening materials without disturbing the surface and geometry of the product. If it is necessary to protect the treatment area from oxidation, a protective gas blowing system is provided in the installation [2].

The principle of operation consists in pulsed laser generation of powerful light radiation. The optical system generates a laser beam of the required diameter and directs it to the workpiece [14] [15].

The processing process can be observed visually through the optical system of the installation.

The pulse duration (τ) was assumed to be 1.5; 2.5; 4 ms.

The pump voltage (Uh) varied from 380V to 800V.

Metallographic studies were carried out to study the structure of the surface layer treated with a laser beam and determine the depth of the zone of thermal influence along the section of the spot.

From the analysis of the results obtained, it was found that the change in microhardness significantly depends on the duration of the radiation pulse and the pump voltage (Fig. 2.). For example, with a pulse duration of $\tau = 1.5$ ms, the microhardness of the sample at a depth of 0.05 mm increases by 10% compared to the initial one and reaches 9.2 GPa, and with a pulse duration of

τ= 4ms, the microhardness increases by 18% and reaches (11 GPa).

When analyzing the dependences of the microhardness of P6M5 steel on the value of the pumping voltage, we see that the microhardness decreases and at a value of Up > 520V becomes lower than the initial one. This is explained by the melting of the surface layer, since the radiation energy density increases with increasing Uh.

Also, based on the results of the study, optimal modes of pulsed laser processing of high-speed steels were established, carried out without melting the surface.

Combined hardening treatment with the use of pulsed laser radiation as a hardening surface treatment was also studied. Two processing options were studied: - preliminary laser treatment of the contact pads of the RI with subsequent coating; - laser treatment of contact pads after coating.

Figure 1: *Example of the layout of a laser laboratory installation*

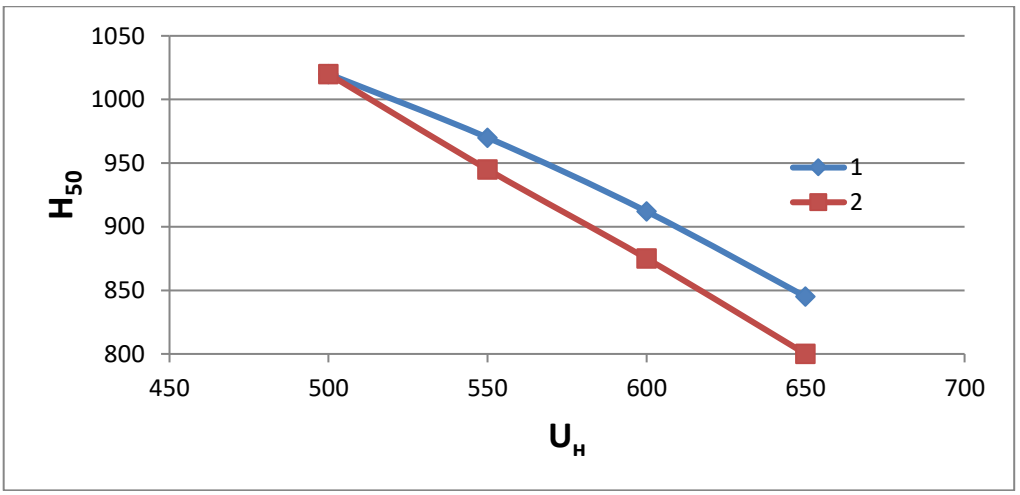


Figure 2: *Dependence of the microhardness of P6M5 steel on the pumping voltage (τ=4 ms, ΔF=18x4x, f=20 Hz, ν=3 mm/s) 1 – at a depth of 0.05 mm; 2 – at a depth of 0.1 mm*

R6M5K5 steel was used as a tool material. Wear-resistant TiN coatings were applied with a thickness of 3-8 microns. Laser treatment of coatings and high-speed substrate was carried out on a pulsed laser installation at a power density of q = 2,7*104 – 5,2*104 W/cm, radiation pulse duration τ = 3.5 ms, laser spot diameter equal to 1 mm. The overlap coefficient of the laser spot was determined for each combination of treatment modes according to the criterion of the minimum volume of the non-hardened zone formed during laser treatment. The efficiency of the cutting tool was assessed by the intensity of wear during 30 minutes of operation and during resistance tests for the period of resistance to wear on the back surface hz = 0.75 mm.

A decrease in the amount of residual stresses occurring in coatings as a result of the use of complex leads to a change in the mechanical properties of coatings: increases the microhardness and adhesion strength of the coating to the tool base, as evidenced by a decrease in the peeling coefficient. In the entire range of laser radiation power density, high-speed tools that have undergone complex processing are characterized by higher microhardness values and lower values of the peeling coefficient. At the same time, the dependence of the separation coefficient on the power density of the laser radiation is extreme.

When implementing complex hardening options for high-speed steel tools, it is important to determine the depth of the hardening zone formed as a result of laser exposure. It is proved that the coating composition and thickness do not significantly affect the depth of the hardening zone. For all coating compositions, the decrease in the depth of the hardening zone compared to the high-speed uncoated base did not exceed 6 %, which allows us to conclude that approximately the same depths of the hardening zone for uncoated and coated cutting tools.

As a result, it was found that for high-speed tools, during complex processing, it is possible to increase the coating thickness by up to 20% compared with coatings obtained using traditional technology by increasing the resistance of the cutting wedge to elastic-plastic deformations arising under the action of thermal force loads during cutting, as a result of laser exposure.

III. Conclusion

Surface treatment of complex shaped tools made of high-speed steel is a necessary process to improve the efficiency of operation of such a tool. This processing is well implemented when using robotic complexes based on laser equipment. Laser hardening of the tool makes it possible to increase its durability by 2-6 times while creating optimal properties of the surface layer at a given depth in the absence of modification of the main material of the tool. The creation of structurally nonequilibrium states in the surface layer of the tool during their pulsed electron beam alloying seems to be a promising treatment of the surface layer of cutting tools before applying a wear-resistant coating.

Also, based on the results of the study, optimal modes of pulsed laser processing of high-speed steels were established, carried out without melting the surface.

This work was supported by the Azerbaijan Science Foundation - **Grant № AEF-MGC-2024-2(50)-16/01/1-M-01**

References

[1] Geller Yu.A. Tool steels. M.: Metallurgy, 1975. – 584 p.

[2] Savin, I.A., Laser hardening of stamps in the conditions of a large engineering company. Defect and Diffusion Forum, 410, pp.450–455, 2021. DOI: 10.4028/www.scientific.net/DDF.410.450

[3] Ardashev, D., Yusubov, N., Shipulin, L., Degtyareva-Kashutina, A. Development of environmental protection system for installation of chrome plating of the internal surfaces of hydraulic drives with hydrostatic guideways. E3S Web of Conferences, 2020, 193, 02015

[4] Shaparev A.V., Savin I.A., Improving Technology of Manufacturing Preparations for Brackets of Heavy Truck Cars. Lecture Notes in Mechanical Engineering, pp. 248–255, 2021. DOI: 10.1007/978-3-030-54817-9_29

[5] Grigoryants A. G., Kazaryan M. A., Lyabin N. A. Laser precision micropro-cessing of materials. – M.: Fizmatlit, 2017. – 416 p.

[6] Rasulov, N.M., Nadirov, U.M. Ensuring the Quality of Taper Pipe Threads Rolled in Pipes with the Control of Kinematic Technological Connections. Russian Engineering Research, 2024, 44(1), pp. 20–22. DOI: 10.3103/S1068798X24010337

[7] Leushin, I.O., Leushina, L.I. Balabanov, I.P. Savin, I.A. Production of moulding cores and waterglass mixtures using "dry ice" for steel and iron casting. CIS Iron and Steel Review, 2021, 21, p. 34–37 DOI:10.17580/cisisr.2021.01.05

[8] Korobatov, D., Baturin, A., Ardashev, D., Abbasova, H. Requirements Definition, Modeling, and Simulation of Control Units of an Electrohydraulic Power Amplifier. Advances in Science and Technology, 2024, 148, pp. 179–186. DOI: 10.4028/p-c1JZF9

[9] Gavariev R.V. Savin I.A. Gavarieva, K.N. Leushin, I.O. Increasing the service life of metal injection molds. Chernye Metally., 2023(6). DOI:10.17580/chm.2023.06.02

[10] Bogatenkov, S.A., Sazonova, N.S., Guzeev, V.I., Yusubov, N.D., Abbasova, G.M. Increasing the Productivity of Multitool Machining on Automated Lathes by Optimizing the Tool Positions.

Russian Engineering Research, 2021, 41(11), pp. 1075–1079. DOI: 10.3103/S1068798X21110058

[11] Korovin V. A., Leushin I.O. Balabanov I.P. Savin I.A. Increase of resistance of steel moulds using the complex modifier INSTEEL-7. CIS Iron and Steel Review. – 2024. – Vol. 27. – P. 31-34. – DOI 10.17580/cisisr.2024.01.05

[12] Shaparev, A., Savin , I., Ptichkin, S., Khankishiyev, I., Mirzayev, A. Punches and Matrices Recovery for Hot Punching by Electric Arc Hardfacing. Advances in Science and Technology, 2024, 148, P. 65–71 DOI:10.4028/p-q4tfAa

[13] Ardashev, D., Yusubov, N., Shipulin, L., Degtyareva-Kashutina, A.Development of environmental protection system for installation of chrome plating of the internal surfaces of hydraulic drives with hydrostatic guideways. E3S Web of Conferences, 2020, 193, 02015. DOI: 10.1051/e3sconf/202019302015

[14] I.P. Balabanov, O.N. Balabanova and A.V. Groshev, "Formation of initial data of the workpiece batch in simulation modelling precision forming", IOP Conference Series: Materials Science and Engineering, vol. 86, no. 1, pp. 012011, 2015. doi:10.1088/1757-899X/86/1/012011

[15] Yusubov, N.D., Khankishiyev, I.A., Abbasova, H.M., Mammadov, E.D., Huseynov, R.A. Matrix models of machining errors in multi-tool multi-carriage adjustments. International Journal on Technical and Physical Problems of Engineering, 2023, 15(3), pp. 309–315.

THE INFLUENCE GAS JET ON THE QUALITY LASER CUTTING METALS

Alexander Shaparev[1], Asim Mirzayev[2], Malik Qarayev[2], Sadaqat Mehdiyeva[2], Rufat Abbasov[2], Yusif Huseynov[2]

•

[1]Kazan National Research Technical University named after A. N. Tupolev - KAI, Kazan, Russia

[2]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan

avshaparev@kai.ru, asimmi@aztu.edu.az, malik.qarayev@aztu.edu.az, sadaqatme@aztu.edu.az, rufet.abbasov88@mail.ru, yusif.huseynov@aztu.edu.az

Abstract

The methods using gas jet in laser cutting are considered. The criteria for the quality laser cutting, such as the absence craters, burrs or ridges on the cutting surface, depend not only on the power the laser radiation, but also on the optimal use the gas jet. An analysis the pressure losses the gas jet in the gap between the nozzle and the metal surface is performed. The passage the gas jet in the cutting zone is significantly affected by the shock wave formed in the gap between the nozzle and the surface the cut metal. Double-jet nozzles provide increased efficiency in removing the liquid phase from the surface the cutting zone.

Keywords: Gas jet, cutting nozzle, dross, shock wave, double-jet nozzle.

I. Introduction

For high-quality laser cutting, a system is required that can control the laser beam and direct it to the processing site with extreme precision. Before starting to cut the contour, the laser beam must first pass through the material at a certain point. The laser cutting process must be accompanied by the supply gas to the cutting zone, which affects the cutting results. The choice gas jet depends on the material being processed and the required quality the workpiece. Oxygen, nitrogen, argon or simply air are used as gas jets for laser cutting.

II. Advantages and Applications Laser Cutting

Compared to other methods separating materials, such as plasma cutting, punching, die cutting or electrical discharge machining, laser cutting has many advantages:

1) Non-contact processing the workpiece is possible.

2) Unlike punching and die cutting, contours almost any shape can be created without a single tool change.

3) Using a laser beam, it is possible to cut both very large contours any shape and small, filigree and complex contours. Geometric contours are processed especially quickly with only a few cuts.

4) The material separation process is carried out with great precision. The width the resulting slot is very small and can be kept almost unchanged. Tolerances up to 0.05 mm can be maintained even in series production.

5) The cutting speed is high. Due to this, it is possible to significantly accelerate the production process, for example, compared to electrical discharge machining.

6) Due to the high energy density, the heat-affected zone can be kept minimal: the depth the hardened layer from 0.1 to 0.2 mm is possible. An oxide layer is formed during oxygen cutting.

7) Slight heating the material minimizes deformation the workpiece.

8) The depth roughness the cutting surfaces is minimal: less than 100 μm. Additional processing the workpiece is not required.

9) Cutting the most frequently used steel grades is performed without the formation burrs, its subsequent removal is not required [1], [2]

III. Methods using gas jet

Today, the laser copes with various cutting tasks and ensures, for example, the formation slots in the thinnest semiconductor chips and high-quality cutting metal up to 30 mm thick. Various cutting methods are used to perform them.

Oxygen cutting: burning cutting. Burning cutting, which uses oxygen (gas purity 99.95, volume percentage 3.5) and creates a pressure maximum 6.0 bar, is used mainly for structural steel. The heated metal reacts with oxygen in the cutting zone, burns and oxidizes. The gas jet blows out the melt formed from the cutting zone together with iron oxides [3].

During the oxidation process, additional energy is generated (exothermic reaction), allowing cutting at a higher speed, as well as processing materials greater thick-ness than is possible when cutting using nitrogen.

The disadvantage this method is the formation an oxide layer on the cutting sur-faces. If the parts are subsequently coated with paint, the oxide layer must first be re-moved. If it protrudes outward, the part becomes unprotected against corrosion [4] [5].

Nitrogen cutting: melt cutting. For melt cutting, nitrogen or argon are used as a gas jet. In this method, the material is also first melted and then blown out the slot using a gas, usually nitrogen. There is no reaction with the molten metal, which allows for oxide-free cutting edges. This cutting method is practiced with a gas pressure 2 to 20 bar (high-pressure cutting) with a nitrogen purity 99.999 and a volume percentage 5.0 (Table 1).

Table 1: *Comparison pressure parameters for burning cutting and high-pressure cutting with nitrogen*

Metal thickness, mm	Oxygen pressure (1.2 mm nozzle), bar	Nitrogen pressure (2.3 mm nozzle), bar
1…3	4.0	20.0
5…30	0.3	20.0

Due to the high gas pressure, almost no burrs are formed on the cutting edges and no slag remains. No additional processing is required.

Cutting with compressed air. Compressed air with a pressure 5-6 bar is used to cut thin sheets metal and blow the melt out the resulting slot. Since air is 80% nitrogen, cutting with compressed air is cutting with melting. The air must be pre-compressed, dried and oil-free. The thickness the sheets being cut depends on the compressed air pressure and the laser power. With a laser power 5 kW and a pressure 6 bar, it is possible to cut a sheet 2 mm thick without burrs. The best results are achieved when cutting aluminum [6] [7].

Laser cutting with plasma support. During laser cutting with plasma support, a plasma cloud

is formed in the cut. It consists ionized metal vapors and ionized gas used for cutting. The plasma absorbs part the CO2 laser radiation and transfers additional energy to the cutting zone. As a result, the material melts faster, which helps to increase the cutting speed. The formation a plasma cloud is useful only when cutting thin metal sheets up to 3 mm thick. Thin metal sheets can be cut at a very high speed. With a sheet thickness 1 mm, a speed 40 meters per minute or more is possible. The cutting edge is rougher than with nitrogen cutting with melting. The maximum sheet thickness depends on the laser power. With a power 6 kW, this method can quickly cut aluminum sheets up to 4 mm thick [8] [9].

IV. Statement the research problem

Cutting quality criteria, such as the absence dimples, burrs or ridges on the cutting surface, may depend not only on the laser radiation power, but also on the optimal use the gas jet. The roughness the cut, the squareness, the width the cut, the absence burrs, the cutting speed largely depend on the gas jet control modes (Fig. 1).

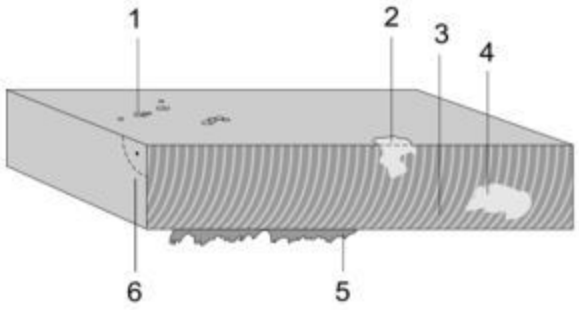


Figure 1: *Criteria for cutting edge quality: 1 – material deposition; 2 – dimple; 3 – ridges on the cutting surface; 4 – washout; 5 – burr; 6 – squareness*

When the laser beam hits the workpiece, the metal heats up to such an extent that it begins to melt or evaporate (Fig. 2).

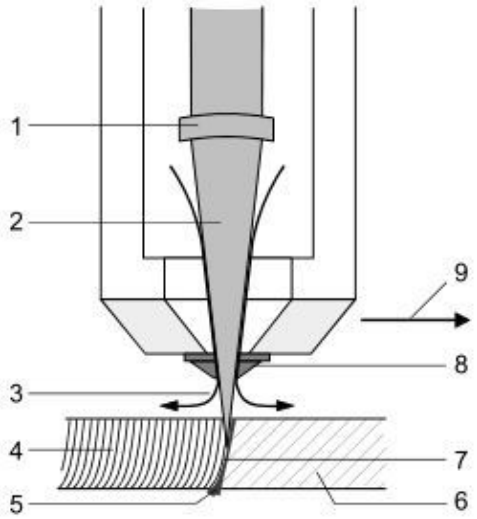


Figure 2: *Laser cutting diagram: 1 – focusing optics; 2 – laser beam; 3 – radial losses gas jet; 4 – cut roughness; 5 – dross; 6 – workpiece; 7 – cutting zone on the workpiece; 8 – nozzle; 9 – cutting direction*

When beam 2 passes completely through workpiece 6, the cutting process begins. Laser beam 2 moves along the contour the part and continuously melts the material. The gas jet comes out nozzle 8 the cutting head together with laser beam 2 and blows out molten metal 5. A narrow cut appears between part 4 and the remainder the workpiece 6 material. In this case, partial pressure losses gas jet 3 occur in the gap between the nozzle and the metal surface. The resulting liquid phase in the laser cutting zone reduces productivity and cutting depth, worsens the quality the side surface the cut and increases the depth the heat-affected zone under it, leading to the formation burrs at its outlet. In this regard, in all laser technological installations, in order to remove the melt from the cutting zone, compressed air under pressure $P_0 = 8 \ldots 6$ atm, or inert gas with a maximum pressure $P_0 \approx 20 \ldots 30$ atm is supplied through a special nozzle [10] [11].

The required gas jet pressure for removing the liquid phase from the cutting zone per unit area is determined by the ratio

$$P_f = 0{,}5\left(c_f \rho u^2\right) \tag{1}$$

where c_f is the coefficient characterizing the resistance to the gas flow on the side surfaces the cut, which is a function the Reynolds $R = \rho u b/\mu$;

ρ is the gas density;

u is the gas flow velocity;

b is the size the cutting width,

μ is the gas viscosity [5].

The gas flow velocity is great importance for the effective removal the liquid phase from the cutting zone [12].

The cutting width varies from 0.15 mm (material thickness 1…6 mm) to 0.5 mm (material thickness 20…30 mm). It must remain the same throughout the working area the unit, since otherwise the dimensions and contours the parts will not be observed [13].

Examples studying the control modes and gas jet flow are shown in Fig. 3 and 4 [2]. The consumption gas used for cutting depends on its pressure and the size the nozzle opening. The higher the pressure and the larger the nozzle opening, the higher the gas jet consumption.

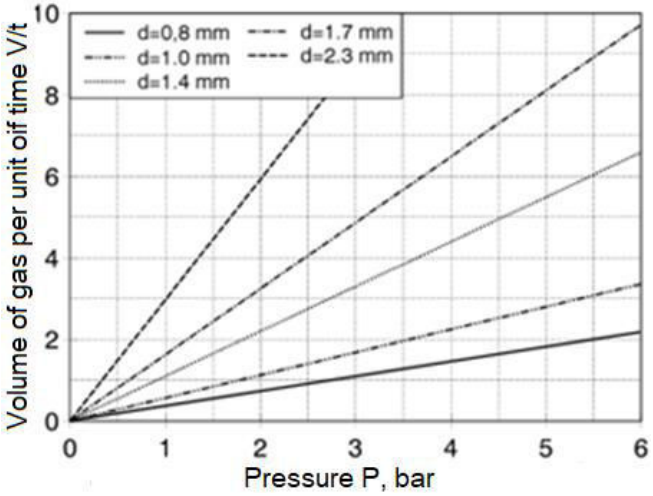


Figure 3: Maximum consumption oxygen used for cutting per hour during continuous cutting [2]

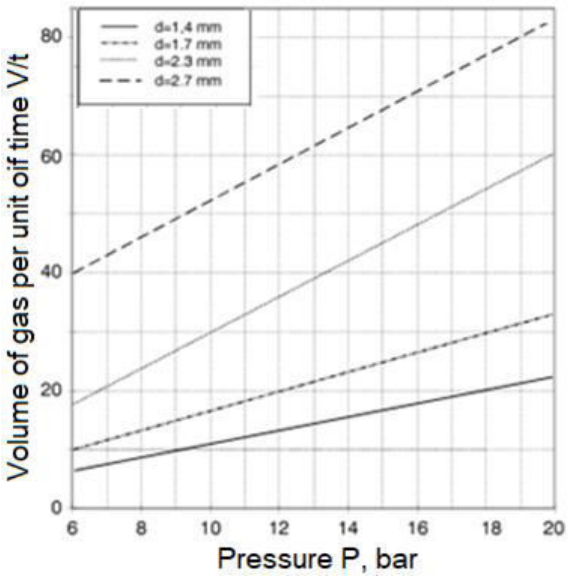


Figure 4: Maximum consumption nitrogen used for cutting per hour during continuous cutting at gas pressure up to 20 bar [2]

During laser cutting, gas jet pressure losses occur, which worsen the laser cutting performance. Therefore, it is necessary to identify the causes and calculate the required gas jet pressure to prevent defects during laser cutting (see Fig. 1). To do this, it is necessary to consider the gas jet flow distribution patterns outside the nozzle.

V. Analysis gas jet pressure losses in the gap between the nozzle and the metal surface

Gas jet pressure losses can reduce the gas jet flow in the laser cutting zone and worsen the quality the cut. Let us consider the gas jet nozzle flow distribution pattern (Fig. 5).

The total pressure the gas flow along the axis its outflow is constant and equal to the sum the static and dynamic components. The shock wave formed in the gap between the nozzle and the surface the cut metal has a significant effect on the passage the gas jet in the cutting zone.

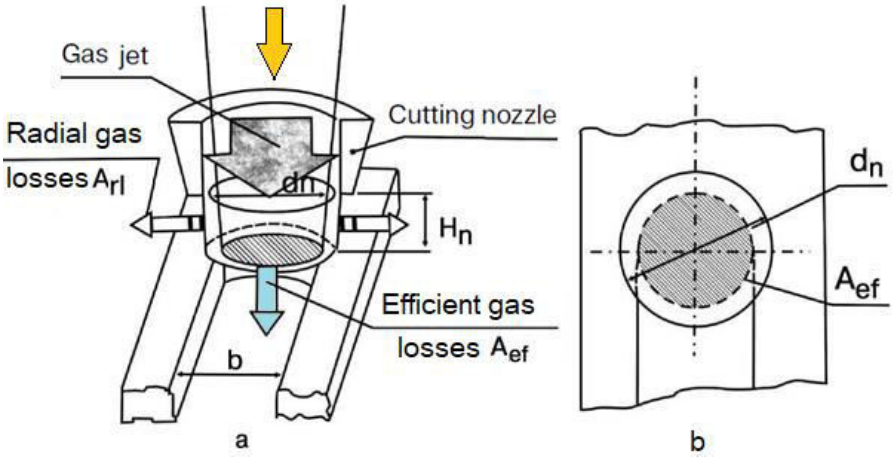


Figure 5: *Nozzle operation diagram [3, 4]*

At the outlet the nozzle with a flow diameter d_n, the gas accelerates to a local supersonic velocity. Behind the nozzle, the gas undergoes adiabatic expansion, the pressure which p_c is determined from the relation

$$p_c = P_0[(2/\gamma_{ad}) + 1]^{\gamma(\gamma+1)} \quad (2)$$

where P_0 is the gas pressure inside the nozzle;

γ_{ad} is the adiabatic index, equal to 1.4 for air.

The pressure loss during gas expansion outside the nozzle is estimated taking into account the area part its flow passing into a cut width b, and the area the other part its flow flowing out in the radial direction, limited by the working distance H_n between the nozzle and the cut [4]. The area the effective part the flow passing into the cut (A_{eff}) is shown in Fig. 4,b by the shaded area. It is limited by the dimensions the cutting zone and the projection the part the circle whose diameter is equal to the outlet diameter the nozzle [14]. The area the cylindrical surface through which the other part the flow flows out in the radial direction is equal to

$$A_{rl} = \pi d_n H_n \quad (3)$$

The value pressure in the effective part the flow, determined from the flow continuity equation, is equal to

$$p_{eff} = p_c A_{eff}/\left(A_{eff} + A_{rl}\right) = f(P_0) \quad (4)$$

Together, relations (2) and (3) show that in order to increase the pressure in the effective part the gas jet flow p_{eff}, it is necessary to reduce the nozzle outlet diameter d_n and its working distance H_n. However, in practice, when $d_n < 1\,mm$ and $H_n < 1\,mm$, the nozzle outlet quickly becomes clogged with products removed from the cutting zone. The rate clogging depends on the cutting depth and the thermophysical properties the material being processed.

It is technically possible to repeatedly increase the pressure P_0 gas jet inside the nozzle and thereby increase the pressure in the effective part the flow p_{eff}. However, here too there is a limitation physical origin. When $P_0 > 0{,}3\,MPa$, the gas flow velocity begins to exceed the speed sound. In this case, a reflected shock wave departs from the surface the material being cut, which limits the gas velocity inside the cut [15]. Therefore, it makes no sense to increase the pressure inside the nozzle more than 0,3 MPa.

A two-jet nozzle, the geometry which is shown in Fig. 6, provides increased efficiency in removing the liquid phase from the surface the cutting zone.

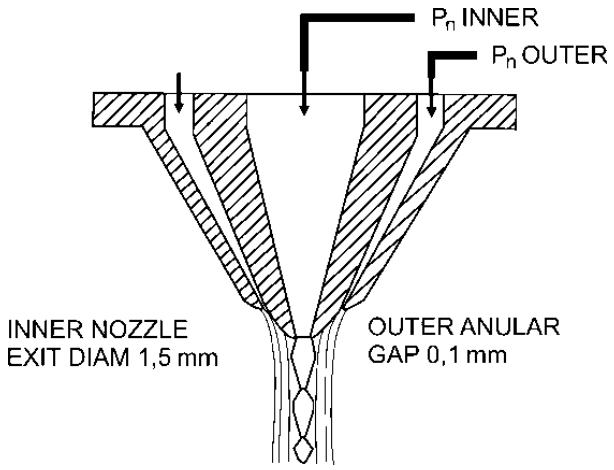


Figure 6: *Geometry a two-jet nozzle [4]*

As can be seen from the examples shown in Fig. 7, the use such a nozzle radically improves the cleanliness the side surface the cut.

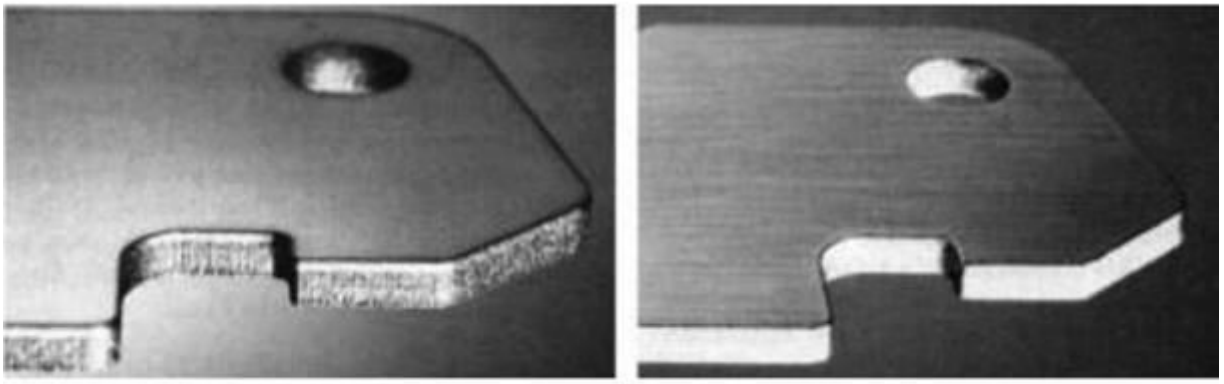


Figure 7: *Quality the side surface the cut in stainless steel 2 mm thick: a – the cut is made using a single-jet nozzle; b – the cut is made using a dual-jet nozzle [4]*

This result is obtained due to the fact that the gas flow exiting the central passage opening with a diameter this nozzle equal to, for example, 1.5 mm (inner nozzle exit diam.), is surrounded by a gas flow flowing out the slotted annular contour (outer anuler gap). Its diameter exceeds the diameter the central flow by several millimeters. The component its expansion velocity is directed radially, including away from the nozzle axis. This limits the pressure losses arising due to the expansion the central flow. Gas is supplied to the inner part the nozzle (inner) under a pressure $0{,}7\ MPa$, to the outer part (outer) – under a pressure $0{,}35\ MPa$. The nozzle is installed at a distance $4…5$ mm above the surface the part [7].

VII. Analysis the dependence the gas jet flow on the shock wave structure

If the ratio the pressure P_0 inside the nozzle to the surrounding atmospheric pressure P_a exceeds $P_0/P_a > 1{,}89$, then the gas jet flow from the nozzle will flow out at a supersonic speed [7]. At a supersonic speed the gas jet, a shock wave is formed as a result its reflections from the atmospheric air. his occurs immediately behind the edge the nozzle, and then the shock wave develops in the form a "barrel", at the end which a gas compaction zone is formed, called the Mach disk (Fig. 8).

Due to the reflections the lateral shock wave from the surrounding air, a second "barrel" is formed behind the first Mach disk, similar in shape to the first, and a second Mach disk. And then a similar structure is repeated several times (Fig. 8). The structure the shock wave changes with a change in the ratio P_0/P_a [7].

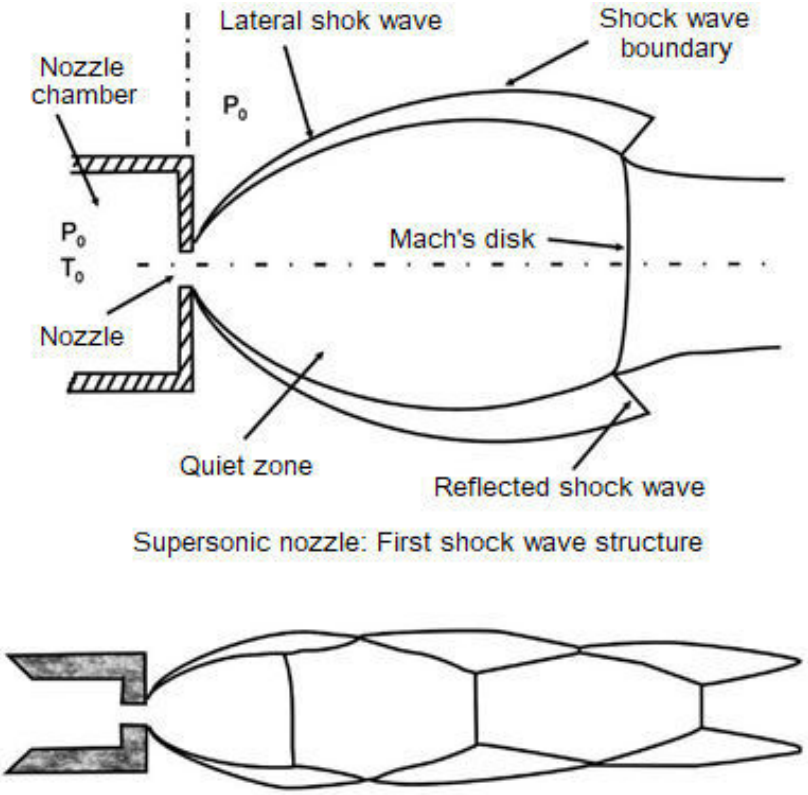


Figure 8: *Formation a shock wave in a supersonic flow*

The effect a shock wave on the pressure value on the surface a part is cyclical and depends on the distance between the part and the nozzle (Fig. 9).

This cyclicity corresponds to the shock wave structure shown in Fig. 8, formed at the specified pressure in the nozzle. When the surface the part is in its calm zone, the pressure on it increases. If it enters the Mach disk zone, the pressure on it decreases sharply, since a smaller amount gas jet passes through the nozzle.

To prevent clogging the liquid phase the molten metal by evaporation products, the nozzle must be placed at a distance 5 mm or 7.5 mm from the surface the part. In this case, the gas jet pressure will decrease by approximately 10% the maximum value, while it is necessary to ensure effective removal the liquid phase from the cutting zone.

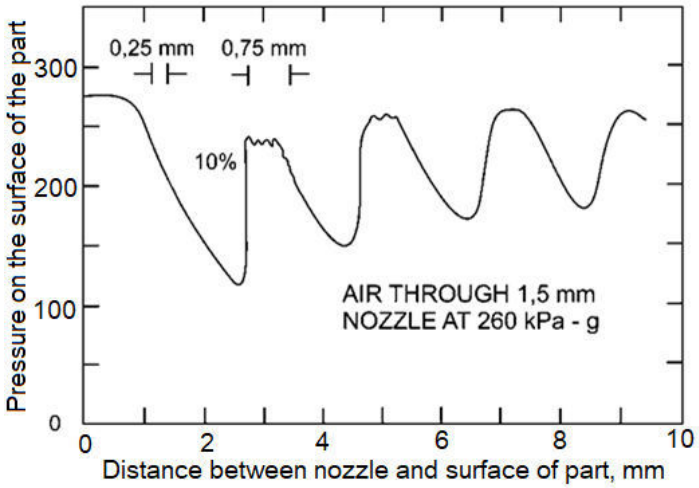


Figure 9: *The pressure on the surface the part depending on the distance between the part and the nozzle*

The best cutting performance to a depth 2 mm in air or inert gas at a pressure no more than 0.5 MPa is achieved by experimentally selecting the distance between the nozzle and the surface the part. When laser cutting to a great depth with a nozzle pressure up to 2...3 MPa, it is necessary to select the optimal values for both the nozzle diameter and its location relative to the surface the part.

VIII. Summary

During gas cutting, when the gas jet is supplied under pressure to a small-diameter nozzle (from 0.7 mm to 1.5 mm), in a narrow gap between the nozzle and the surface the workpiece (up to 2 mm), an interaction the shock wave and its reflection from the metal surface occurs. At supersonic speed the gas jet, as a result its reflection from the atmospheric air, a shock wave is formed. This happens immediately behind the edge the nozzle, and then the shock wave develops in the form a "barrel", at the end which a gas compaction zone is formed, called the Mach disk

The mechanism this process: the gas jet, supplied under pressure up to 2.7 atm, forms a supersonic flow in the nozzle. When exiting the nozzle, this flow collides with the surface the part, forming a shock wave. This wave, in fact, is an abrupt change in the gas parameters, i.e. pressure, temperature and speed. The shock wave, reflecting from the surface the part, interacts with the initial gas flow. This interaction causes additional pressure losses in the gap, which leads to a decrease in the speed the gas flow passing into the hole being processed.

Effect on the cutting process: Pressure losses in the gap can have a significant impact on the efficiency the gas cutting process:

1. Reduced cutting speed: Reducing the gas flow leads to a decrease in the cutting speed, since the pressure required to melt the metal is reduced.

2. Deterioration in cut quality: Insufficient gas pressure can lead to uneven cut edges, burrs and other surface defects.

3. Increased gas consumption: To achieve the required pressure and cutting speed, it may be necessary to increase the gas supply, which leads to increased costs.

4. How to minimize the impact shock waves:

Optimization nozzle geometry: A properly selected nozzle shape, taking into account the gas characteristics and cutting mode, can reduce the intensity shock waves. Use special nozzles: There are nozzles that facilitate a smoother transition the supersonic flow from the nozzle to the gap, which reduces the intensity shock waves.

Use gas mixtures: The use certain gas mixtures can change the characteristics shock waves and reduce their impact on the cutting process. The interaction shock waves with the surface the part affects the efficiency laser cutting, leading to pressure losses, reduced cutting speed and deterioration in cutting quality. To improve the efficiency laser cutting, it is necessary to take this factor into account and apply appropriate measures to minimize the impact shock waves.

This work was supported by the Azerbaijan Science Foundation-**Grant № AEF-MGC-2024-2(50)-16/01/1-M-01**

References

[1] D. K. Y. Low, L. Li, P. J. Byrd, Hydrodynamic Physical Modeling of Laser Drilling, Transaction of the ASME, 852/ Vol. 124. November 2002.

[2] Savin, I.A., Laser hardening of stamps in the conditions of a large engineering company. Defect and Diffusion Forum, 410, pp.450–455, 2021. DOI: 10.4028/www.scientific.net/DDF.410.450

[3] E. D. Vaks, M. N. Milenkiy, L. G. Saprykin. Practice of Precision Laser Processing, Technosfera, Moscow, 2013. – 696 p.

[4] Shaparev A.V., Savin I. A. , Ptichkin S.N. Application of the Polymeric Material RIMAMID for Production of Machine Parts. IOP Conference Series: Materials Science and Engineering, 969,

pp.012021, 2020. DOI: 10.1088/1757-899X/969/1/012021

[5] Abbasov V., Amirov F., Amirli S., Hasanli S., Frana K., Formation of Shaft Accuracy during Mechanical Processing on CNC Machines. Advances in Science and Technology, 2024, 148, pp. 81 - 86, DOI: 10.4028/p-N5uKb8

[6] Savin, I.A., Akhmedeev, M. V., Connection of the Steel Pipes Having a Polymeric Covering on Internal and External Surfaces. Solid State Phenomena, 299, pp. 766–771, 2020. DOI: 10.4028/www.scientific.net/SSP.299.766

[7] Simon, S., Yusubov, N., Amirli, S. Formation of Geometric Parameters of the Surfaces of Cylindrical Parts during Waterjet Cutting. Advances in Science and Technology, 2024, 148, pp. 59–64. DOI:10.4028/p-l5wSEu

[8] Grigoryants A. G., Kazaryan M. A., Lyabin N. A. Laser precision micropro-cessing of materials. – M.: Fizmatlit, 2017. – 416 p.

[9] Leushin, I.O., Leushina, L.I. Balabanov, I.P. Savin, I.A. Production of moulding cores and waterglass mixtures using “dry ice” for steel and iron casting. CIS Iron and Steel Review, 2021, 21, p. 34–37 DOI:10.17580/cisisr.2021.01.05

[10] Rasulov, N.M., Nadirov, U.M. Ensuring the Quality of Taper Pipe Threads Rolled in Pipes with the Control of Kinematic Technological Connections. Russian Engineering Research, 2024, 44(1), pp. 20–22. DOI: 10.3103/S1068798X24010337

[11] Korovin V. A., Leushin I.O. Balabanov I.P. Savin I.A. Increase of resistance of steel moulds using the complex modifier INSTEEL-7. CIS Iron and Steel Review. – 2024. – Vol. 27. – P. 31-34. – DOI 10.17580/cisisr.2024.01.05

[12] Balabanov, I. P., and F. F. Davletshin. "Implementation of Iso 9001, Iso 14001, Iso 45001 Requirements with the Systems of Electronic Document Turnover." International Journal of Engineering and Technology (UAE), vol. 7, no. 4, 2018, pp. 78-81. doi:10.14419/ijet.v7i4.7.20388

[13] Abbasov, I.T., Simon, S., Fritzsche, P.D., Yusubov, N.D. Study on reducing energy consumption in rough turning operations. SOCAR Proceedings, 2022, pp. 23–28. DOI: 10.5510/OGP2022SI100696

[14] Shaparev, A., Savin , I., Ptichkin, S., Khankishiyev, I., Mirzayev, A. Punches and Matrices Recovery for Hot Punching by Electric Arc Hardfacing. Advances in Science and Technology, 2024, 148, P. 65–71 DOI:10.4028/p-q4tfAa

[15] Korobatov, D., Baturin, A., Ardashev, D., Abbasova, H. Requirements Definition, Modeling, and Simulation of Control Units of an Electrohydraulic Power Amplifier. Advances in Science and Technology, 2024, 148, pp. 179–186. DOI:10.4028/p-c1JZF

TECHNOLOGICAL FEATURES LASER CUTTING COPPER AND BRASS

Alexander Shaparev[1], Iliya Avvakumov[1], Vagif Movlazade[2], Ugurlu Nadirov[2], Jeyhun Rahimov[2], Lachin Babayev[3]

•

[1]Kazan National Research Technical University named after A. N. Tupolev - KAI, Kazan, Russia
[2]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan
[3]Department of Shipbuilding and Ship Repair, Azerbaijan State Marine Academy, Baku, Azerbaijan
avshaparev@kai.ru, iiavvakumov@kai.ru, movlazade.vaqif@aztu.edu.az, ugurlu.nadirov@aztu.edu.az, jeyhun.rahimov@aztu.edu.az, lachin.babayev@asco.az

Abstract

The article discusses the features laser cutting copper and brass, certain requirements for equipment and technology. The advantages using fiber lasers for cutting copper and brass are formulated, such as better absorption laser radiation, high precision cutting small parts, clean and smooth cut without burrs, high cutting speed. The features cutting copper and brass blanks with lasers with a power 1 and 2 kW are studied. Recommendations for laser cutting copper and brass are developed. Copper and brass laser cutting is possible, but much more difficult than other metals. This is partly due to the fact that copper is a highly reflective material. Copper's reflective properties make it difficult for the laser's infrared light to be absorbed, slowing down the cutting process. To get the most out laser cutting copper, you need to consider speed, power, reflectivity, and focal point.

Keywords: laser cutting, copper, brass, cutting quality, fiber lasers.

I. Introduction

Laser cutting copper and brass has its own characteristics due to the high ther-mal conductivity the material. Copper also has a high heat capacity coefficient. This imposes certain requirements on the equipment and technology:

1) Low absorption infrared laser radiation makes cutting these metals difficult;

2) Copper and brass are good reflectors (and therefore poor absorbers) infrared laser radiation in the solid state;

3) Pure copper in the solid state reflects approximately 95% laser radiation with a wavelength ~ 1 μm;

4) The reflectivity copper and brass decreases when the metal is heated, and drops sharply when the metal melts (for example, up to 70% for copper in the molten state). These metals in the molten state absorb significantly more laser energy;

5) Laser cutting copper and brass is more difficult, the greater the thickness the material being processed;

6) The laser spot size should be as small as possible, and the laser power should be high.

II. Statement the research problem

With the right choice laser, optics and cutting process, the laser beam quickly melts the surface reflective materials, then interacts with the more absorbent molten metal and initiates an efficient, stable cutting process. Incorrect choice laser/optics set-up or use suboptimal process parameters can result in the laser coming too close to the solid metal, resulting in excessive light reflection [1]. Too much reflection, in turn, results in an inefficient cutting process and potential damage to the optics.

Fiber lasers are widely used in the industry for laser cutting copper and brass. They differ from CO2-lasers in the way the laser beam is generated. In fiber lasers, the beam is generated inside a thin light guide, which is a glass fiber. Due to this, fiber lasers are more compact, more efficient and more durable than CO2-lasers and are considered preferable for cutting copper and brass [2] [3].

Advantages fiber lasers for cutting copper and brass:

1. Material absorption: Fiber lasers emit a beam with a shorter wavelength, which is more easily absorbed by reflective materials such as copper and brass com-pared to CO2-lasers.

2. Accuracy: The diameter the laser spot fiber lasers is significantly smaller than that CO2-lasers. This allows for higher cutting accuracy, which is especially important when working with small parts.

3. Cut quality: Fiber lasers provide a cleaner and smoother cut, without the for-mation burrs.

4. Processing speed: Due to higher efficiency, fiber lasers allow metal cutting at a higher speed, which increases processing productivity.

In conclusion, fiber lasers are currently a more popular and sought-after solution for metal cutting due to their high efficiency, accuracy and processing speed. However, fiber lasers also have their limitations. They are most effective when cutting thin sheets metal. The maximum material thickness that can be cut with a fiber laser is limited to 10 mm for copper and brass. The limit the material thickness that can be processed depends on the power the emitter and the model the laser machine [4][5] .

The plants manufacture a wide range products from sheet copper. The use laser cutting technology using powerful fiber lasers for this would bring significant ad-vantages in terms productivity and savings on the rather expensive metal. However, the developers such lasers strongly recommend careful development the technological processes for laser cutting copper and brass in order to protect the laser system from damage due to slag and splashes molten metal sticking to the surface the quartz lens during burning and cutting operations [6] [7].

The formation these black spots on the lens can significantly reduce the perfor-mance and quality the cut [6]. As the laser continues to operate, these spots become localized heat absorption centers, which leads to a number cascading problems:

1) Thermal Effect the Lens: Uneven heating caused by the black spots creates a thermal gradient across the lens, changing its refractive properties. This leads to focus drift, where the actual focal point shifts from the intended position, reducing cutting accuracy.

2) Reduced Transmittance: Black spots obscure the laser beam, reducing the overall power transmitted to the workpiece. This can lead to uneven cutting depth and quality across the material.

3) Lens Degradation: Intense heating in these areas can cause microcracks or even thermal degradation the lens, significantly reducing its service life.

4) Focal Length Change: As the central area the lens overheats due to black dots, it can cause a local change in the curvature the lens, effectively reducing the focal length. This interferes with normal cutting operations, especially in precision work.

5) Beam Distortion: The presence black dots can introduce aberrations into the laser beam prile, resulting in uneven energy distribution and may result in wider kerfs or poor edge quality.

III. Research cutting copper and brass with fiber lasers

Features cutting copper with a 1 kW laser. Table 1 shows the maximum possible cutting speed copper with a thickness 1…3 mm, produced by a 1 kW laser.

Table 1. *Maximum speed laser cutting copper [3]*

Sheet thickness h, mm	1.0	1.5	2.0	3.0
Cutting speed v, m/min	1.00	0.70	0.50	0.15

The copper surface was installed in the focal plane the lens with a focal length $f = 200\ mm$ and a light spot size $d_1 \approx 350$ μm. Cutting was carried out in a compressed air environment supplied to a single-jet nozzle with a through-hole diameter 1.8 mm under a pressure 8 atm. Since the air contains 20% oxygen, the temperature the liquid phase layer increases along with the heating from the laser beam according to the reaction [3] [8]:

$$Cu + 0{,}5O_2 \rightarrow CuO + 150\ kJ/mol$$

With the specified settings, the copper thickness $h = 3.0\ mm$ is the maximum possible for cutting with a fiber laser with a power 1 kW. The width the laser cut when processing sheets with a thickness $h = 1.5 \dots 3\ mm$ is 400 μm. The burr height at all thicknesses at the cut exit does not exceed 20 μm.

Using a focusing lens with $f = 145\ mm$ and setting the metal surface below the focal plane at a distance $\Delta f = 2\ mm$ reduces the cutting width, since the diameter the light spot radiation localization does not exceed $d_1 \approx 200\ \mu m$. Using a higher-power laser increases the cutting speed and depth. There is no burr on the surface the cutting zone, the roughness index $R_a = 0.8\ \mu m$, the cutting quality is satisfactory.

The surface the cutting zone for copper sheets 2 mm thick is shown in Fig 1 [4] [8].

Figure 1:. *Side surface a cut in a copper workpiece 2 mm thick*

The absence burrs and grooves on the surface the cutting zone is explained as follows. The laser radiation power density in the focal plane with a beam diameter $d_1 \approx 350$ μm is $W = 1{,}04 \cdot 10^6\ W/cm^2$. Since this value power density W is close to the evaporation threshold, an increased layer liquid phase is formed on the side surface the cut, which cannot be completely removed from the cutting zone. As a result, grooves are not formed on the side surface [9].

Cutting L63 brass with a 1 kW laser. The maximum possible speed laser cutting brass quickly

decreases with an increase in the thickness the workpiece. As can be seen from the graph (Fig. 2), the thickness the L63 brass workpiece, equal to 5 mm, is the limit for cutting with a 1 kW fiber laser. In this case, the maximum possible cutting speed will decrease to approximately 0.4...0.5 m/min.

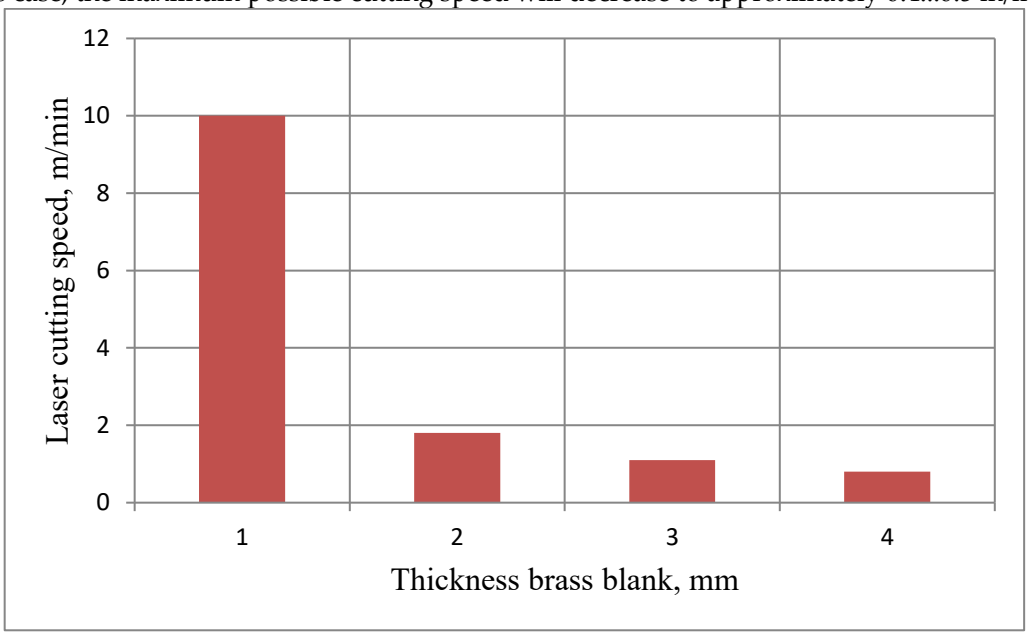


Figure 2: *Maximum speed laser cutting brass L63 depending on the thickness the brass blank*

The cutting width in a 3 mm thick brass blank is 250 μm, while in a 4 mm thick blank it is 270 μm. The cutting width in a 3 mm thick brass blank is 150 μm less than the cutting width in copper the same thickness. The cutting speed a brass blank is 10 times greater than the cutting speed a copper blank. This is explained by the fact that the absorption coefficient on the copper surface changes periodically [9]. First, the absorption coefficient on the copper surface decreases and then increases again. This process is associated with the periodicity the process removing the liquid phase from the surface in the cutting zone and the subsequent restoration its layer, which requires additional energy. In laser cutting brass, a similar dependence is manifested to a much lesser extent.

From the data shown in Fig. 2 it follows that under constant cutting conditions, the maximum cutting speed and depth a brass blank are greater than on copper, which is explained by the presence 30% zinc in brass [8] [10].

Figure 3: *Side surface the cut in 3 mm thick brass*

When zinc interacts with pure oxygen, its oxidation reaction begins. In the cutting zone, heat is released from two oxidation reactions copper and zinc [11], so the total heat is approximately 3 times

greater than when cutting copper.

With a brass blank thickness 1 mm, there is no burr at the exit from the cutting zone. The burr height at thicknesses 2.5 mm, 3 mm and 4 mm does not exceed 50 μm. On the side surface the cut in a brass blank 3 mm thick (Fig. 3) there are no grooves, the cut surface looks rougher.

Cutting brass L63 with a 2 kW laser. Parts made brass 2 mm thick are shown as an example in Fig. 4 [8].

Figure 4: *Example contour cutting a part made 63 brass with a thickness 2 mm*

Laser cutting was carried out under the following conditions. The radiation was focused by an objective with $f = 145\ mm$. The surface the brass was installed below the caustic constriction at a distance equal to $\Delta f = 3\ mm$. To remove the liquid phase, a two-jet nozzle NK15-15 was used, providing a cutting gas pressure 16 atm. The cutting width was 0.35 mm, which ensured improved removal the liquid phase. The cutting speed was 3.5 m/min.

To cut parts with a more complex configuration from brass blanks with a thickness 3 mm and 5 mm, an objective with $f = 200\ mm$ was used. Compressed air or nitrogen was used as the cutting gas, supplied to the two-jet nozzle NK25-20 under a pressure 16 atm [8] [12].

The maximum possible cutting speed 5 mm thick brass did not exceed $v = 1\ m/min$. At this speed, the cutting width at the entrance decreased to 250 μm. Due to the need to cut a complex contour, including rounded sections with a radius 10...15 mm and internal cutouts with an acute angle at their apex, the cutting speed was reduced relative to the maximum possible to 0.7 m/min. At this speed, the cutting width was equal to 150 μm. Removal the liquid phase from the cutting zone with a width 150 μm becomes impossible even at a compressed air pressure 16 atm. Part the liquid phase remaining on the surface the cutting front accumulates in its lower zone [13]. Slowly flowing out it, it welds to the surface the brass and, solidifying, forms a burr. The height the burr on 3 mm thick brass parts did not exceed 0.05 mm, and on 5 mm thick brass it did not exceed 0.15 μm. In both cases the burr is easily removed.

Fig. 5 shows a photograph the side surface a 5 mm thick cut made in a compressed air environment. In a compressed air environment the side surface the cut has a dark gray color and corresponds to $R_a = 3.2\ \mu m$.

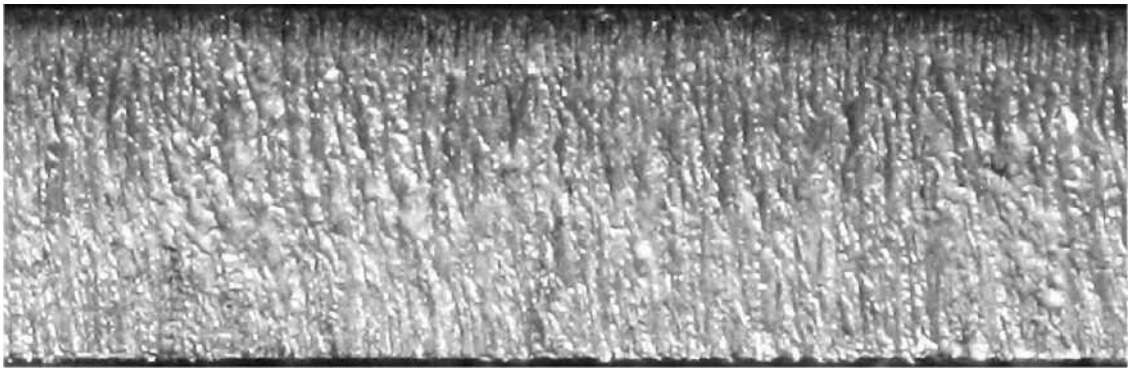


Figure 5: *Side surface the cut in 5 mm thick brass [8]*

IV. Recommendations for laser cutting copper and brass

When cutting materials such as copper and brass with fiber lasers, it is recommended to slightly reduce the feed rate - by about 10…15% to ensure a higher-quality burn-through the material. For a quick burn-through the material, it is recommended to use maximum power. This is important because the metal being processed has the greatest probability reflecting laser radiation at the beginning the cutting process, because as the metal heats up, its reflectivity decreases [8] [14].

Below are approximate values power required for cutting copper sheets different thicknesses:

Table 2: *Required power required for cutting copper sheets*

Thickness copper sheets, mm	1.0	2.0	3.0	4.0	6.0
Required minimum power, W	1000	1500	2000	3000	4000

In addition, such a factor as the position the focus is quite important. The optimal focal length is determined for each specific material separately. It is necessary to select a focal distance so that it is as close as possible to the surface being processed, but not to such an extent that the cutting quality suffers. Correct selection the focal length ensures maximum efficiency the cutting process.

You can also improve the quality cutting metal materials by using an gas jet, such as air, oxygen, nitrogen and argon [15]. These gases can perform various functions during the cutting process: removal molten material and smoke from the cutting zone, cooling, protective function (nitrogen is used to prevent oxidation when cutting copper blanks), protection optical elements from combustion products, stabilization the cutting process, etc. For example, the use oxygen leads to the formation copper oxide in the processing zone, which reduces its reflectivity.

V. Summary

Copper and brass laser cutting is possible, but much more difficult than other metals. This is partly due to the fact that copper is a highly reflective material. Copper's reflective properties make it difficult for the laser's infrared light to be absorbed, slowing down the cutting process.

To get the most out laser cutting copper, you need to consider speed, power, reflectivity, and focal point. All these factors help make laser cutting copper and brass easier. There are several factors to consider when using fiber lasers to cut copper and brass:

1. Laser Power. When using a fiber laser for cutting, power is a consideration. This is one the most important factors when cutting copper and brass. The more powerful the fiber

lasers, the better the quality the copper and brass cutting will be. Pulsed CO_2-lasers have significantly higher power density, but a fiber laser is best suited for cutting copper and brass because it has a wavelength $\lambda = 1.06\ \mu m$ and can absorb more energy quickly.

2. Power Setting: Ideally, you want to use the highest possible peak power to reduce the amount time the material is in its most reflective state.

3. Cutting Speed: To use fiber lasers and achieve excellent copper cutting, you need to consider the cutting speed. To optimize the speed, you need to consider the thickness the workpiece and the power the laser machine. It is always a good idea to start with a slower speed to ensure that you can get the laser through the piercing hole before you start laser cutting.

4. Gas Jet: One the main factors when cutting copper and brass is the gas jet, as it moves the compressed gas into the cutting area. This gas also protects the lens from the vaporized metal from the cutting area when the metal is in a liquid state. This also helps ensure the required quality, productivity and speed metal cutting.

5. Reflection Detector: When laser cutting metals such as copper and brass, special attention should be paid to the laser beam. If too much light is reflected from the copper cut by the laser, it can cause damage to the machine. The reflection detector monitors the infrared laser light emitted by the fiber laser and the radiation it produces. If too much radiation hits the fiber laser lens, the detector will turn f the machine.

6. Focusing Position. When cutting copper and brass, it is necessary to set the focus the laser beam as close to the top the surface as possible, which will reduce the amount evaporated material coming into contact with the laser head, increase the specific power and accelerate the melting the metal.

This work was supported by the Azerbaijan Science Foundation-**Grant № AEF-MGC-2024-2(50)-16/01/1-M-01**

References

[1] D. K. Y. Low, L. Li, P. J. Byrd, Hydrodynamic Physical Modeling of Laser Drilling, Transaction of the ASME, 852/ Vol. 124. November 2002.

[2] Savin, I.A., Laser hardening of stamps in the conditions of a large engineering company. Defect and Diffusion Forum, 410, pp.450–455, 2021. DOI: 10.4028/www.scientific.net/DDF.410.450

[3] E. D. Vaks, M. N. Milenkiy, L. G. Saprykin. Practice of Precision Laser Processing, Technosfera, Moscow, 2013. – 696 p.

[4] Shaparev A.V., Savin I. A. , Ptichkin S.N. Application of the Polymeric Material RIMAMID for Production of Machine Parts. IOP Conference Series: Materials Science and Engineering, 969, pp.012021, 2020. DOI: 10.1088/1757-899X/969/1/012021

[5] Ardashev, D., Yusubov, N., Shipulin, L., Degtyareva-Kashutina, A. Research on chrome plating quality of the internal surfaces in the hydraulic drives with hydrostatic guideways. Materials Science Forum, 2021, 1037 MSF, pp. 417–422. DOI: 10.4028/www.scientific.net/MSF.1037.417

[6] Savin, I.A., Akhmedeev, M, V., Connection of the Steel Pipes Having a Polymeric Covering on Internal and External Surfaces. Solid State Phenomena, 299, pp. 766–771, 2020. DOI: 10.4028/www.scientific.net/SSP.299.766

[7] Simon, S., Yusubov, N., Amirli, S. Formation of Geometric Parameters of the Surfaces of Cylindrical Parts during Waterjet Cutting. Advances in Science and Technology, 2024, 148, pp. 59–64. DOI: 10.4028/p-l5wSEu

[8] Grigoryants A. G., Kazaryan M. A., Lyabin N. A. Laser precision micro-processing of materials. – M.: Fizmatlit, 2017. – 416 p.

[9] Leushin, I.O., Leushina, L.I. Balabanov, I.P. Savin, I.A. Production of moulding cores and waterglass mixtures using “dry ice” for steel and iron casting. CIS Iron and Steel Review, 2021, 21, p. 34–37 DOI:10.17580/cisisr.2021.01.05

[10] Frana K., Attia S.H., Otremba F., A numerical simulation of the filling process in the pressure bottle (2018) International Journal of Mechanical Engineering and Robotics Research, 7 (5), pp. 558 – 563. DOI: 10.18178/ijmerr.7.5.558-563

[11] Korovin V. A., Leushin I.O. Balabanov I.P. Savin I.A. Increase of resistance of steel moulds using the complex modifier INSTEEL-7. CIS Iron and Steel Review. – 2024. – Vol. 27. – P. 31-34. – DOI 10.17580/cisisr.2024.01.05

[12] Balabanov I P, Balabanova O N, Gilman V N. Development of a parametric model for calculating cutting forces for external cylindrical turning of steel 20CrMNTi. IOP Conference Series: Materials Science and Engineering, 2020, 915(1), 012005 DOI: 10.1088/1757-899X/915/1/012005

[13] Yusubov, N., Abbasova, H. Models of Cutting Forces in The Matrix Theory of Multitool Machining Accuracy. Key Engineering Materials, 2024, 979, pp. 27–38. DOI: 10.4028/p-bW48Sb

[14] Shaparev, A., Savin , I., Ptichkin, S., Khankishiyev, I., Mirzayev, A. Punches and Matrices Recovery for Hot Punching by Electric Arc Hardfacing. Advances in Science and Technology, 2024, 148, P. 65–71 DOI:10.4028/p-q4tfAa

[15] Ardashev, D., Yusubov, N., Shipulin, L., Degtyareva-Kashutina, A. Development of environmental protection system for installation of chrome plating of the internal surfaces of hydraulic drives with hydrostatic guideways. E3S Web of Conferences, 2020, 193, 02015. DOI: 10.1051/e3sconf/202019302015

ON THE ISSUE OF ALLOY CRYSTALLIZATION DURING CASTING INTO METAL MOLDS

Igor Savin[1], Renat Gavariev[1], Mukhaddin Samadov[2], Elgun Shabiyev[2], Fazil Orujov[3]

•

[1]Department of Design and Technology of Machine-Building Industries, Kazan National Research Technical University named after A.N.Tupolev – KAI, Kazan, Russia
[2]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan
[3]Department of Shipbuilding and Ship Repair, Azerbaijan State Marine Academy, Baku, Azerbaijan
IASavin@kai.ru, RVGavariev@kai.ru, muxeddin@mail.ru, elgun@aztu.edu.az, fazil.orujov@asco.az

Abstract

The modern scientific and technological revolution, along with social progress, has ensured tremendous scales of production development in our country. When obtaining many machine parts, instruments, and structures, cast blanks are the most cost-effective option, and in the case of particularly complex parts or when using alloys that cannot be processed by pressure, casting technology becomes the only possible method. At the same time, a significant number of specific defects occur in castings; surface roughness and dimensional accuracy often do not meet the requirements set for products; the properties of alloys in their cast state are generally lower than those in their deformed state. Therefore, the problem of further improving the quality of cast blanks has become critically important today. To address this issue, a specialist must understand the mechanisms of casting processes and be able to control them.
The article discusses the physical and mechanical processes that occur during the pouring of molten metal into a metal mold. Various scenarios of premature formation of heterogeneous nucleation sites are considered. The features of nucleus formation on the solid surface of the mold are described depending on its geometric characteristics. Additionally, physical models of crystallization of heterogeneous nuclei are presented, describing the mechanism of appearance of a widespread defect in castings in the form of shrinkage cavities.

Keywords: casting, crystallization, shrinkage.

I. Introduction

Casting in metal molds is a common method for producing castings. At the same time, modern trends in market development dictate the need for constant improvement in the operational and quality characteristics of manufactured products. Traditionally, enhancing the operational properties of components made from metallic alloys is achieved by increasing the chemical and physical homogeneity of the metal and minimizing defects of various scales, the majority of which arise during the formation of blanks—ingots and castings. Currently, defects in the internal structure of ingots and castings are often associated solely with the thermal processes of solidification. However, there is evidence that both thermal and mechanical processes significantly influence the

formation of alloy structures and the occurrence of defects [1].

The primary cause of defects in cast blanks and parts is the phase transition of the alloy from a liquid to a solid state, which is accompanied by a series of physicochemical phenomena, including shrinkage—reduction in volume during solidification. Due to shrinkage, mechanical movement of the alloy occurs during the formation of blanks to compensate for the shrinkage of the already solidified part of the metal. Insufficient feeding of the ingots and castings at this moment leads to the formation of defects such as shrinkage cavities, dents, and porosity.

This article investigates the patterns of alloy mechanics during the crystallization period. Understanding these patterns provides the opportunity to develop and optimize new and existing casting technologies.

II. The main part

When pouring the melt into the mold cavity, it comes into contact with the metallic surface, which determines the formation of a solid shell. This moment is crucial in terms of the formation of crystallization nuclei, which can negatively affect the strength of the resulting casting. The further cooling process depends on several factors, including the superheat of the melt relative to the crystallization temperature, the physicochemical properties of the mold, the flow rate, and the mechanical characteristics of the gating system and the mold cavity [2].

The degree of superheat is the main factor influencing the subsequent cooling and solidification process of the melt. Depending on the level of superheat, there are three possible pouring scenarios: without superheat, with slight superheat, and with significant superheat [3], [4].

In the first scenario of pouring without superheat (Figure 1a), the contact of the metallic mold surface with the melt immediately leads to the formation of a shell, over which a new shell is subsequently formed when the following streams of melt reach the surface of the mold.

When pouring the melt with slight superheat (Figure 1b), a solid shell will also form at the point of contact between the melt and the metallic surface of the mold. However, the subsequent flow of the melt will deliver a heat transfer that slows the growth of the shell near the gate and impedes the solidification process at the initial stage. Once the heat supply ends, solidification will continue similarly to the first case.

In the case of pouring the melt with significant superheat (Figure 1c), a solid shell will form upon contact with the metallic surface of the mold, just like in the first two cases. However, the heat flow from the incoming stream of overheated alloy will melt the formed shell, leading to solidification occurring some distance away from the gate, following the mechanism of the second case. At some distance from the gate, the further solidification process will proceed similarly to the case of slight superheat.

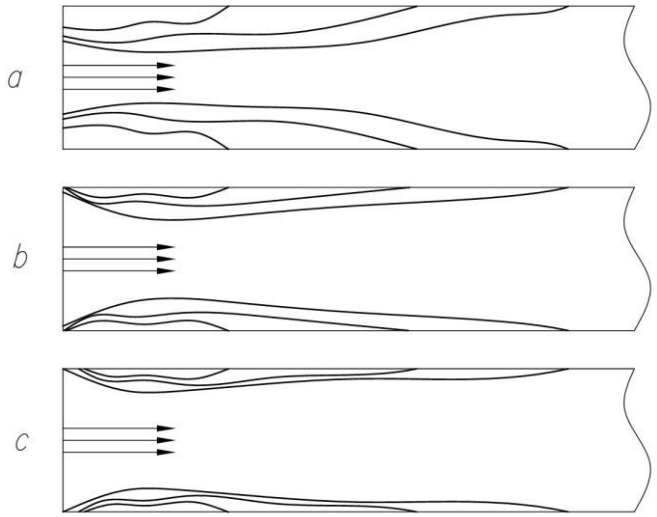


Figure 1: *Diagram of the interaction between the metal and the mold during the filling of the channel when pouring the melt: a) without superheat; b) with slight superheat; c) with significant superheat*

In the areas near the gate, where the stream of the poured melt comes into contact with the walls of the metallic mold, enhanced heat transfer occurs, leading to the erosion of the mold wall or the previously formed shell. This results in localized heating and, consequently, slows down the solidification process [5], [6]. After the melt completely fills the mold cavity, solidification begins approximately uniformly from all surfaces except for those where the initial stage of solidification has been inhibited. In the vicinity of the gate, there is a likelihood of the formation of some volume of unsolidified melt, which may lead to the development of shrinkage cavities. The extent of this defect depends on the level of melt superheat and the duration of flow through the gate [7], [8].

In cases where the nucleus forms not in the space but on a solid surface or substrate, it may take the shape of a spherical segment with a radius of curvature that provides stability with significantly fewer atoms or volume (Figure 2a). Therefore, the formation of such a heterogeneous nucleus will require less undercooling than for a homogeneous nucleus of the same volume [9], [10].

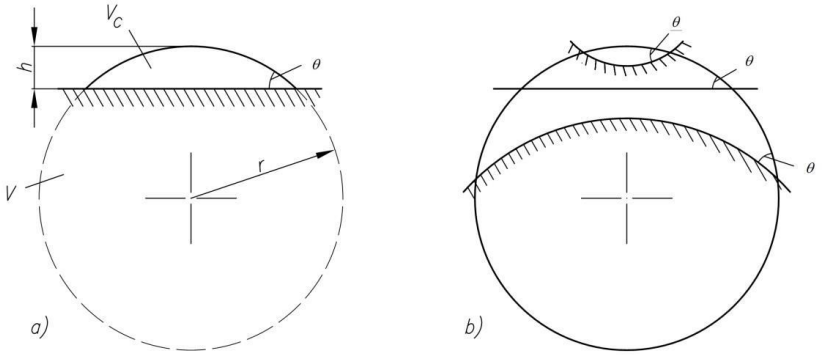


Figure 2: *Formation of a nucleus on flat (a) and curved (b) surfaces of the mold; V - volume of the sphere, VC - volume of the segment, h - height of the segment, θ - edge angle*

Most often, such surfaces will be the forming surfaces of the mold, and sometimes small solid particles of impurities in the melt. The conditions of interaction between the materials of the mold, the poured melt, and the forming solid phase have the greatest influence on the formation of the embryo [11], [12].

It is worth noting that the shape of the solid surface affects the stability of the embryo if the surface curvatures are comparable to the dimensions of the embryo. Various depressions and indentations contribute to the formation of stable embryos at smaller sizes even under slight undercooling (Figure 2 b), and vice versa.

In addition, the blurred particles of the metal mold or shell, when further introduced into the flow of the melt, will act as nucleation centers, starting their growth slightly earlier than the main volume of the poured melt. The grain size depends on the rate of crystallization and the rate of nucleation. If we assume that in a certain plane, the filling with solid phase occurs such that the rates of crystallization and nucleation are constant, and that after a certain time, two nucleation centers will form at random locations, then as the crystals grow in the shape of squares around each of them, there will be a growth of the crystal in the form of layers of constant thickness (see Figure 3).

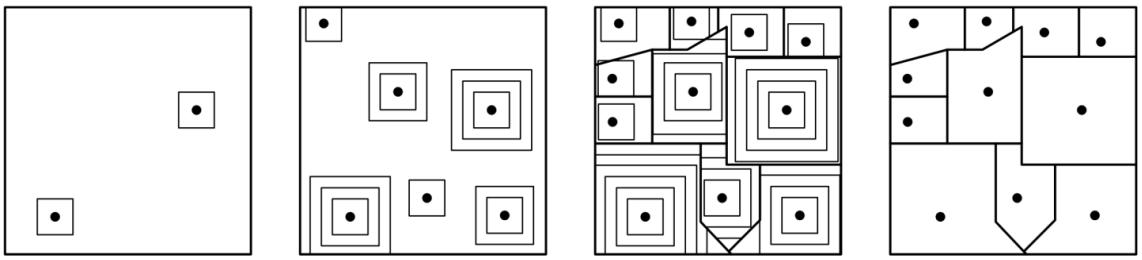


Figure 3: *Crystallization Model on a Plane*

The solution to the problem of covering a solid phase of such a plane made it possible to formulate the relationship that grain size increases with an increase in the rate of crystallization and a decrease in the rate of nucleation, and vice versa [13], [14]. An important point from a practical perspective is that the nucleation rate in pure metals is lower than that in their alloys.

The specified planar model is simplified and does not take into account certain negative aspects associated with defects during shrinkage [15]. When crystallizing castings with equiaxed structures, defects are most easily formed in a high-concentration suspension when the casting transitions from a liquid state to a solid. This occurs when the average amount of solid phase exceeds 50% and depends on the size and shape of the grains. At this stage, the risk of microcrack formation, porosity, and other defects that impair key performance characteristics increases. Such defects arise as a result of grain movement under the influence of pressure, vibration, or stresses due to hindered shrinkage.

This circumstance becomes particularly important during the crystallization of castings with a two-phase and multi-layer structure, as well as for alloys prone to dendritic liquation. In such alloys, which exhibit selective solidification, the remaining liquid will have a composition corresponding to the eutectic or a second phase with a lower crystallization temperature. During this period, the crystallization process may slow down or even halt entirely, despite continuous heat removal, until the alloy reaches the solidification temperature of the eutectic or second phase. This stage poses a certain danger, as intergranular cracks may form that are not always easily welded. This largely depends on the volume of residual liquid and its distribution among solid grains. When there is a sufficient amount of eutectic or a second phase, effective filling of the formed micropores is possible, making such an alloy less hazardous, as the resulting microcracks and tears can be successfully filled and welded when large volumes of liquid move. However, the most significant danger arises from a small amount of eutectic or a second phase (less than 15%), which may lead to a slowdown in crystallization at such residual liquid content. The tendency to form the described defects also depends on the distribution of the residual liquid phase. A structure in which grains are completely surrounded by thin shells of the liquid phase will not withstand significant stresses, leading to typical intercrystalline destruction. Microcracks and tears that occur during shrinkage compensation or vibrations are not always repairable and remain within the solid material, which may cause leaks during pressure testing and reduce mechanical properties. The fracture characteristics of samples from such alloy resemble a break under tension within the solidification range of the alloy, further confirming the link between defect formation and the specified temperature range.

The situation changes when a small amount of residual liquid alloy of eutectic composition or

a second phase does not surround the solid grains but exists between them as separate isolated areas. In this case, grains touch each other at their edges, forming a rigid structure that significantly better withstands stresses from shrinkage, vibrations, metallostatic, and other pressures. Shrinkage during the crystallization of the last portions of the liquid alloy, in this case, manifests as discrete, isolated pores. Such pores (usually oval in shape) have a negligible effect on the occurrence of leaks and the mechanical properties of the alloy.

Primarily, the shape and position of the liquid phase are influenced by the interfacial tension at the boundary between the crystal and the residual liquid. At low interfacial tension, the liquid, even in a small volume, envelops the grain from all sides. Conversely, at high interfacial tension, the residual liquid forms a significant angle with the face of the forming crystal, promoting the formation of separate areas of liquid phase.

Figure 4 shows a schematic representation of the layer-by-layer growth process of three crystal nuclei, resulting in the formation of a residual liquid phase in a three-grain junction. Subsequently, crystallization — the filling of this liquid phase region — can occur primarily through one of the following mechanisms:

1. Crystallization mainly occurs along the faces of the growing crystals, while this process slows down at the edges forming the angles of the crystals. The slowdown of crystallization in the two-grain junction is due to the accumulation of liquid with an increased concentration of the alloying component. This enrichment in the two-grain junction arises from selective solidification at the faces of the two crystals, where the two growing faces push the impurity into the narrow space between them. When the wetting angle is small, interfacial tension also facilitates this distribution. The liquid enriched with the alloying component has a lower freezing temperature (according to the normal phase diagram), which further slows down crystallization in the two-grain junction.

2. Crystallization occurs mainly in the two-grain junction, meaning that it happens at a greater rate along the edges than on the faces of the growing crystals. In this case, due to selective solidification, the liquid becomes enriched with a component that increases interfacial tension. Here, the alloying element or the alloy enriched by capillary forces (surface tension) is continually expelled from the two-grain angles. An alloy with a lower impurity content from the central part of the liquid enters the void of the two-grain junction, facilitating further filling of the two-grain junction during continuous circulation of the liquid phase. The driving force behind this process is the potential difference arising from the surface tension.

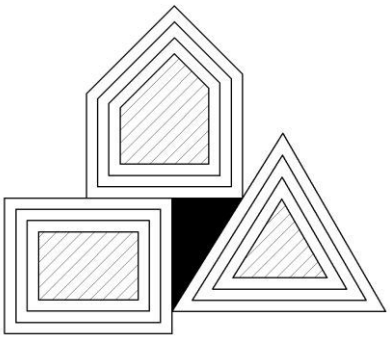


Figure 4: *Planar Model of Crystallization Considering the Residual Liquid Phase*

If dendritic liquation is often observed in alloys, micro-voids can also form in them. Alloying and modifying additions can significantly alter the surface tension of liquid metals and, consequently, the interfacial tension at the boundary between the solid crystal and the residual liquid. This opens up the possibility for controlling the feeding mechanism of the solidifying alloy. This change can be illustrated using the example of the tin bronze OC10-2. The bronze, by its structure, is two-phase: the eutectoid, which represents a more fusible fraction, is located along the grain boundaries of the α-solid solution of tin in copper. Depending on the amount of added

elements (such as Ni, Ti, Ce, and others), the shape and position of the eutectoid in the alloy's structure, as well as its density and mechanical properties, change [1].

Considering the studied relationship between interfacial tension and the location of the fusible component, the dihedral angle formed by the faces of two α-solid solution grains in the region of the eutectoid inclusion can serve as a measure of the magnitude of interfacial tension. If the additions increase the angle of the eutectoid component formed by the faces of two solid solution crystals, thus increasing interfacial tension, it leads to a compact and isolated arrangement of the eutectoid. Conversely, if the additives decrease the dihedral angle and reduce interfacial tension, favorable conditions are created for an elongated position of the eutectoid along the grain boundaries [2].

When studying the microstructure of samples on microsections, it is important to consider that the plane of the section is usually not perpendicular to the faces of the oriented crystals, leading to distortion of the visible angles in the section. Nevertheless, it can be statistically demonstrated that the most frequently observed angles in the planar section represent the true magnitude of the spatial angle. This implies the necessity of conducting several measurements for each structure and subsequent statistical processing of the obtained data.

Based on structural analysis, it has been found that various additions can increase, decrease, or have no effect on the wetting angle of solid grains by the residual liquid at the final stages of crystallization. This, in turn, affects the shape and position of the eutectoid within the structure. At the very final stages of solidification, solid crystals are surrounded by the liquid and are not rigidly connected to each other, creating conditions for the movement of the entire two-phase mass. If the structure forms with a large dihedral angle at the final stages of solidification, the crystals become rigidly interconnected, and the liquid occupies isolated compact regions, which enhances the alloy's resistance to shear stresses and complicates two-phase feeding. Changes in the shape and position of the eutectoid in bronze, with a constant relative amount of the eutectoid phase, predetermine changes in the dominant feeding mechanism and, consequently, influence the ratio of isolated to transitive porosity of the alloy, which, in turn, determines its tightness, mechanical properties, and other characteristics.

The cooling rate of castings affects defect formation in the following ways. At very high cooling rates, dendritic liquation is reduced due to the suppression of segregative diffusion. On the other hand, at very low cooling rates, castings also show reduced dendritic liquation as a result of the homogenizing diffusion process. At moderate cooling rates, there is an increased tendency for dendritic liquation and the formation of micro-voids. High cooling rates, as well as sharp temperature gradients, contribute to minimizing the suspended and solid-liquid states. Additionally, the nearest zone of fully solidified metal is subjected to all stresses. At very slow cooling rates, when a polycrystalline structure with randomly oriented grains is formed, the reduction of defects is associated with the increased time available for their healing due to the prolonged solidification process.

III. Summary

Thus, the improvement of the operational properties of castings obtained in metal molds depends on the formation of the casting as it transitions from a liquid state to a solid one. The degree of superheating is one of the most important factors influencing the cooling and solidification process of the melt. In this case, the most optimal variant involves significant superheating, which helps eliminate the negative effect of premature crystal formation on the surface of the metal mold. However, the greatest impact is exerted by the volume of the liquid phase and its distribution among the solid phases. With a sufficient amount of eutectic or a secondary phase, it is possible to effectively fill the formed microvoids, making such an alloy less hazardous because the resulting microcracks and tears can be successfully filled and welded when large volumes of liquid are moved.

This work was supported by the Azerbaijan Science Foundation-**Grant № AEF-MGC-2024-2(50)-16/01/1-M-01**

References

[1] Gulyaev BB (1976) Teoriya liteinih processov L.: Mashinostroenie p 216

[2] Gavariev RV, Gavarieva KN, Soldatkina EN (2020) Features of Design of Chill Molds for Casting of Non-Ferrous Metal. IOP Conference Series: Materials Science and Engineering 969(1). 012068 doi:10.1088/1757-899X/969/1/012068

[3] Korovin V. A., Leushin I.O. Balabanov I.P. Savin I.A. Increase of resistance of steel moulds using the complex modifier INSTEEL-7. CIS Iron and Steel Review. – 2024. – Vol. 27. – P. 31-34. – DOI 10.17580/cisisr.2024.01.05

[4] Ardashev, D., Yusubov, N., Shipulin, L., Degtyareva-Kashutina, A. Development of environmental protection system for installation of chrome plating of the internal surfaces of hydraulic drives with hydrostatic guideways. E3S Web of Conferences, 2020, 193, 02015. DOI: 10.1051/e3sconf/202019302015

[5] Gavariev RV, Savin IA, Leushin IO, Gavarieva KN (2021) On the question of porosity of aluminum castings obtained by injection molding. TsvetnyeMetally 8:76-82. DOI: 10.17580/tsm.2021.08.12

[6] Korobatov, D., Baturin, A., Ardashev, D., Abbasova, H. Requirements Definition, Modeling, and Simulation of Control Units of an Electrohydraulic Power Amplifier. Advances in Science and Technology, 2024, 148, pp. 179–186. DOI: 10.4028/p-c1JZF9

[7] I.A. Savin, M. Akhmedeev, Connection of the Steel Pipes Having a Polymeric Covering on Internal and External Surfaces. Solid State Phenomena, 299, pp. 766–771, 2020 DOI: 10.4028/www.scientific.net/SSP.299.766

[8] Ardashev, D., Yusubov, N., Shipulin, L., Degtyareva-Kashutina, A. Research on chrome plating quality of the internal surfaces in the hydraulic drives with hydrostatic guideways. Materials Science Forum, 2021, 1037 MSF, pp. 417–422. DOI: 10.4028/www.scientific.net/MSF.1037.417

[9] Gavariev RV, Gavarieva KN (2020) On the Issue of Heat Balance of Molds for Injection Molding of Non-Ferrous Metal Alloys. IOP Conference Series: Materials Science and Engineering 969(1). doi:10.1088/1757-899X/969/1/012069

[10] Shaparev A.V., Savin I. A. , Ptichkin S.N. Application of the Polymeric Material RIMAMID for Production of Machine Parts. IOP Conference Series: Materials Science and Engineering, 969, pp.012021, 2020. DOI: 10.1088/1757-899X/969/1/012021

[11] Rasulov, N., Nadirov, U., Abbasova, I. Improving of Machining Efficiency of Threads and Conical Surfaces of Diverse Directions by Managing Static Technological Relationships. Advances in Science and Technology, 2024, 148, pp. 97–102. DOI: 10.4028/p-4ozPgF

[12] Leushin, I.O., Leushina, L.I. Balabanov, I.P. Savin, I.A. Production of moulding cores and waterglass mixtures using “dry ice” for steel and iron casting. CIS Iron and Steel Review, 2021, 21, p. 34–37 DOI:10.17580/cisisr.2021.01.05

[13] Gilman V N, Balabanov I P, Faskhutdinov A I. Improving the efficiency of shaving through the use of wear-resistant coatings. IOP Conference Series: Materials Science and Engineering 2019 DOI: 10.1088/1757-899X/570/1/012024

[14] Savin, I.A., Laser hardening of stamps in the conditions of a large engineering company. Defect and Diffusion Forum, 410, pp.450–455, 2021. DOI: 10.4028/www.scientific.net/DDF.410.450

[15] Bogatenkov, S.A., Sazonova, N.S., Guzeev V.I., Yusubov, N.D., Abbasova, G.M. Increasing the Productivity of Multitool Machining on Automated Lathes by Optimizing the Tool Positions. Russian Engineering Research, 2021, 41(11), pp. 1075–1079. DOI:10.3103/S1068798X21110058

DEVELOPMENT OF A PARAMETRIC MODEL FOR CALCULATING CUTTING FORCES IN EXTERNAL CYLINDRICAL TURNING OF 20CRMN STEEL (1.7147) USING AN SNMG 15 06 16-PR 4425 INSERT

Igor Balabanov[1,2], Vagif Movlazade[3], Nizami Yusubov[3], Heyran Abbasova[3], Ramil Dadashov[3], Rasul Huseynov[4]

•

[1]Kazan National Research Technical University named after A. N. Tupolev – KAI, K.Marx Street 10, Kazan, Tatarstan Republic, 420111, Russian Federation

[2]Kazan Federal University, Kremlyovskaya str. 18, Kazan, Tatarstan Republic, 420008, Russian Federation

[3]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan

[4]Department of Shipbuilding and Ship Repair, Azerbaijan State Marine Academy, Baku, Azerbaijan

balabanovip@mail.ru, movlazade.vaqif@aztu.edu.az, nizami.yusubov@aztu.edu.az, abbasova.heyran@aztu.edu.az, dadashov@aztu.edu.az, rasul.huseynov1966@gmail.com

Abstract

This article develops and presents a mathematical model for calculating cutting forces during the machining of 20CrMn steel (1.7147) using an SNMG 15 06 16-PR 4425 T-Max® P insert for turning. We conducted experimental research on a specially designed test rig based on the 16D25 lathe. This setup measures spindle speed, feed rate, cutting depth, and the cutting forces generated during the machining process with high precision. We used the LTR-EU-8 workstation for data acquisition and analysis, equipped with galvanic isolated modules and a synchronized data transmission interface to ensure accurate measurements. The system transmitted real-time data to a computer for further processing, which helped verify the theoretical model. The results showed a high correlation with actual measurements: the deviation between calculated and experimental values did not exceed 5.68%, proving the model's accuracy in predicting cutting forces. This accuracy plays a key role in optimizing machining processes, reducing tool wear, and lowering energy consumption. The study also found that cutting forces provided by major tool manufacturers are often overestimated. In some cases, the discrepancies between calculated and actual forces reached 17.8%, potentially affecting the accuracy of process planning and the choice of optimal cutting parameters. Additionally, the study revealed that the cutting forces typically provided in calculations by leading tool manufacturers are often overestimated. In some cases, discrepancies between calculated and actual force values reached up to 17.8%, which can impact the accuracy of process planning and the selection of optimal cutting parameters.

Keywords: Modeling, cutting force, cutting speed, cutting force model.

I. Introduction

Cutting force is an integral parameter that determines the surface quality during machining. It reflects the interaction between the cutting tool and the workpiece material during material removal, directly impacting the final surface of the part [1]. Both the magnitude and direction of the cutting force significantly affect characteristics such as surface roughness, accuracy, and geometric deviations of the machined surface [1], [2].

An increase in cutting forces can lead to adverse effects, including higher surface roughness, accelerated tool wear, and potential workpiece deformation [3]. Conversely, reducing cutting forces typically improves surface quality and provides more accurate and stable part dimensions, though it may reduce productivity and increase the final product cost. This highlights the importance of controlling, optimizing, and predicting cutting forces as key factors in achieving the desired machining characteristics. In this context, selecting optimal cutting conditions, tool geometry, and machining strategies aimed at minimizing applied forces [4] is crucial to meeting the required quality standards for finished products.

Various methods and mathematical models are used in machining to predict cutting forces [4], including analytical, empirical approaches, and numerical modeling techniques. Analytical models include Merchant's circle diagram and the orthogonal cutting model, which based on the tool's geometric parameters, the material properties of the workpiece, and the cutting conditions [5]. In contrast, empirical models rely on experimental data to establish the relationship between cutting force and process input parameters.

Modern machining technologies increasingly employ numerical modeling to predict cutting forces, including the use of Computer-Aided Design (CAD) [6] and Computer-Aided Manufacturing (CAM) [7] software. However, the accuracy of these models often depends on several factors, including the quality of input data and assumptions made during modeling. Finite Element Analysis (FEA) and Finite Element Method (FEM) simulations allow for a more detailed examination of the cutting process, accounting for complex interactions between the tool, workpiece, and cutting conditions [6]. Despite their utility, these methods require extensive input data for accurate calculations. While numerical models play a key role in optimizing machining processes, improving surface quality, extending tool life, and enhancing overall manufacturing efficiency, they still require further refinement.

Empirical models offer a practical and accessible way to analyze and predict cutting forces during machining, allowing engineers to avoid complex mathematical calculations or laborious equations. Based on experimental data, such models establish relationships between input parameters and cutting forces [8], making them a valuable tool for quick assessments in real production environments. By using these models, engineers and technicians can make informed decisions, optimize cutting conditions, select tools [8], and ensure the required surface quality without needing complex computations or extensive practical research.

In recent years, there has been a noticeable simplification in approaches to modeling cutting forces, and many tool manufacturers often avoid using them. Instead, they prefer to rely on derived metrics such as energy consumption and tool life [9]. For instance, Walter (https://www.walter-tools.com), a leading cutting tool manufacturer, recommends using a specific formula to calculate cutting forces, where A is the cross-sectional area of the chip (mm2); h is the thickness of the chip (mm); kc is the specific cutting force (H/mm2); mc is the correction coefficient [10].

$$F_c = A \times k_c \times h^{-mc} \ (N) \tag{1}$$

It is important to note that cutting force depends on a number of additional factors, including the hardness of the workpiece material, the presence of scale on the surface, as well as the stability and accuracy of the workpiece. The coefficients kc and mc represent average values specific to certain materials and tools, which limits the precision of predictions and makes the model suitable

only for preliminary assessments [11].

With the advancement of computer modeling technologies, Finite Element Method (FEM) and Finite Element Analysis (FEA) are becoming increasingly significant in modern manufacturing [12]. These approaches allow for detailed analysis of cutting processes at relatively low costs [13]. However, approximate models often yield conditional results, which are unacceptable in automated design environments. This highlights the need for more accurate and reliable models for predicting cutting forces.

An example of such models is the Taylor empirical equations (2) [14], also known as the Taylor-Bühl model. These equations are widely used to estimate cutting forces and tool wear based on machining conditions. The Taylor model enables engineers to predict tool life, evaluate cutting forces, and analyze the impact of various machining parameters on tool wear. By adjusting variables such as cutting speed, feed rate, and depth of cut, it is possible to optimize the process, extend tool life, and improve the quality of the machined surface. It is important to emphasize that the constants in this model must be determined experimentally for specific materials and tools, which underscores its empirical nature [15].

$$P = 10 \times Cp \times t^x \times s^y \times V^n \times K \quad (2)$$

Where t, S, and V represent cutting parameters; K adjusts for cutting conditions; Cp, x, y, and n serve as empirical coefficients and exponents.

Kosilova and Baranovsky developed methods based on empirical models and experimental data [13]. These approaches rely on practical measurements and observations of machining processes to create equations and dependencies that describe cutting forces in various operations. Their models consider key factors such as cutting speed, feed rate, and depth of cut, tool geometry, material properties, and tool wear. Kosilova and Baranovsky focused on analyzing real cutting conditions and identifying relationships between input parameters and cutting forces. Through experiments and data collection, they built mathematical models that predict cutting forces with high accuracy [13]. Engineers and machinists use these empirical models to optimize machining processes, reduce tool wear, and achieve the desired surface quality in production.

This study follows an empirical approach, but the coefficient tables for different steel types have remained unchanged for over 40 years. These outdated tables fail to reflect modern advances in tool design, manufacturing, and new materials.

The primary goal of this research aims to develop a new empirical cutting model based on the Taylor equation for machining 20CrMn-1.7147 steel with a specific cutting tool. The objective focuses on creating a methodology that allows quick generation of cutting force models for specific materials and tools [14]. These models will support precise simulations of machining processes using computer tools.

II. Methods

The experimental setup used a universal lathe model 16D25, which was equipped with numerical control (CNC). The main technical specifications of this machine include a maximum spindle speed of 2000 revolutions per minute (RPM), a maximum feed rate of 2 millimeters per revolution, and a maximum workpiece diameter over the lathe bed of up to 500 millimeters. To accurately measure the actual cutting speed, an inductive rotational speed sensor, the Balluff BES M12MI-PSC40B-BV03, installed on the spindle pulley (see Table 1). Four grooves machined into the pulley to ensure that the sensor would trigger four pulses per revolution, allowing for the implementation of an error-checking system to verify the accuracy of the measurements taken [15].

Table 1: *Characteristics of Balluff BES M12MI-PSC40B-BV03*

Characteristic	Malue
Operating temperature, °C:	-25..70
Thread size of the housing:	M12
Tripping distance, mm:	4
Supply voltage, V:	10-30 DC
Switching frequency, Hz:	2500

To monitor actual movements, linear displacement sensors mounted on brackets were used, allowing measurements of up to 1000 mm in the longitudinal direction and 100 mm in the transverse direction relative to the spindle axis, with high precision of 0.01 mm. Additionally, a three-coordinate force sensor, installed on a DCLNR 2525 M15 tool holder, was utilized, having undergone preliminary calibration. The cutting tool employed was a new SNMG 15 06 16-PR 4425 T-Max® P insert for turning. All sensors and devices connected to the LTR-EU-8 workstation via galvanically isolated LTR modules and a synchronization interface. The force sensor linked to a TR212M, equipped with a 24-bit analog-to-digital converter (ADC) with a frequency of 7.6 kHz and designed to connect up to eight strain gauges with resistance from 100 Ω to one kΩ [16]. The other sensors connected to LTR24 modules (with four parallel channels, a 24-bit ADC, and a frequency of 117 kHz) and LTR11, which supports multi-channel data acquisition (up to 32 channels for single-channel signals with a 14-bit ADC and a frequency of 400 kHz) [20].

To determine the coefficients of the empirical cutting model, the team developed a method for finding the optimal solution through constraint enumeration, demonstrating high accuracy and effectively utilizing specialized software[17]. In this case, the team used the "Solver" function in MS Excel for calculations [18]. To compute the model coefficients, they measured cutting forces under various combinations of cutting speed, feed rate, and depth of cut [19]. The cutting speed range varied from 230 to 330 m/min in increments not exceeding 50 m/min. The machine featured a stepped spindle speed adjustment system, which defined the relationship between cutting speed, workpiece diameter, and spindle RPM. During the experiments, they recorded the actual spindle speeds from the Balluff BES M12MI-PSC40B-BV03 sensor [20].

$$V = \frac{\pi \times d \times n}{1000} \tag{3}$$

The thickness of the removed allowance varied from 0.5 to 2.5 mm in increments not exceeding 0.5 mm, which depended on the capabilities of the transmission. It is important to note that the normal operating thickness for the SNMG 15 06 16-PR 4425 insert is ap = 5 mm (1.5 - 8). To eliminate errors in longitudinal feed caused by equipment wear, the team measured the actual feed rate using displacement sensors. They set the feed values at 0.11, 0.13, 0.22, and 0.25 mm/rev. The machining occurred on workpieces from the same batch, which had a hardness of 193 HB after prior measurement. The team conducted the machining process without cutting fluids, due to the presence of sensors and the inability to measure the temperature of the cutting insert and tool holder using a laser-based sensor.

III. Results

Table 2 presents a data sample reflecting the dependencies of cutting speed (V), feed rate (s), thickness (t), and pressure (P). Adjusting parameters to achieve desired results through changes in input values represents a fundamental method in various fields of science and engineering. This approach involves systematically optimizing variables to meet specific goals. Such techniques

enable the optimization of a system or process, ensuring their most efficient and productive operation.

Table 2: *Selective Experimental Results*

	V (m/min)	fa (mm/rev)	a (mm)	P (H)
1	232,1	0,505	1,51	324,9
2	250,9	0,807	1,51	468,1
3	264	0,751	1,02	316,6
4	282,3	0,251	0,74	205,2
5	299,8	0,251	1,02	256,9
6	332,2	0,251	1,02	268,3

For instance, in mechanical engineering, engineers can vary parameters such as cutting speed, feed rate, and tool geometrical characteristics to ensure the required accuracy and quality of machining. Similarly, in data analysis, adjusting model parameters can enhance the accuracy of predictions and results. By carefully selecting and modifying input parameters, one can fine-tune the process to meet specific criteria and achieve the desired outcome. This methodological approach serves as a universal tool for specialists across various disciplines, allowing them to optimize performance and accomplish their objectives.

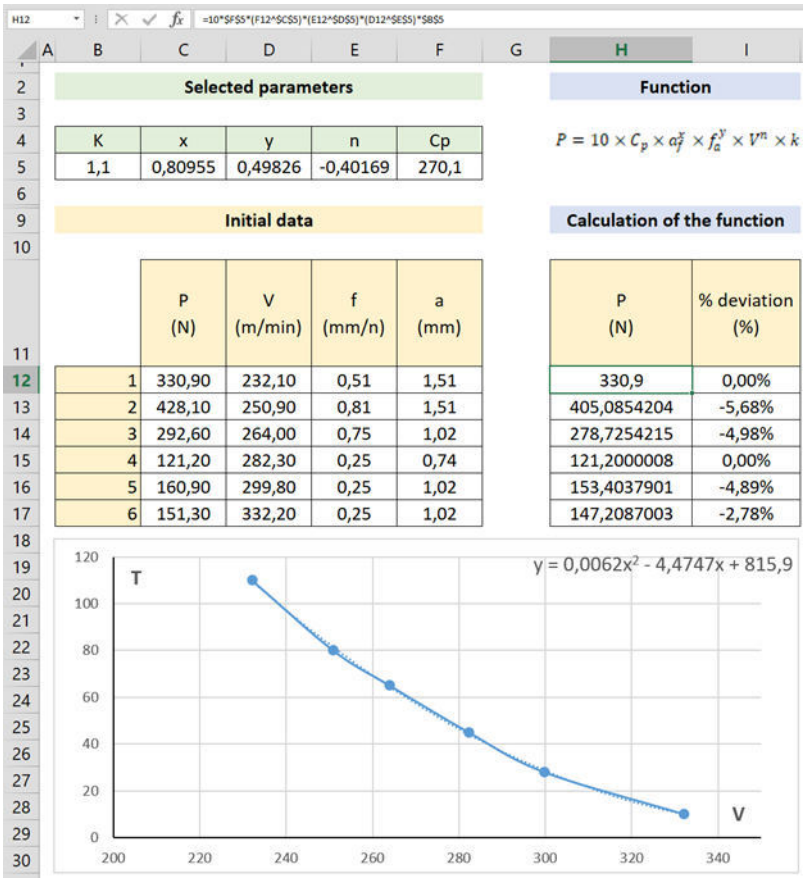


Figure 1: *Interface for Coefficient Selection in Cutting Model Construction*

Figure 1 presents the software interface designed for calculating the coefficients of the empirical cutting model, which implements a complete enumeration method considering constraints, utilizing built-in functions of MS Excel. Here the data obtained from the experiment, located in cells C12-C17. Cells B5-F5 indicate the values derived using the "Solver" function in MS Excel based on the specified constraints (2). Cells H12-H17 present the values calculated using Taylor's empirical model (1).

The calculation of the cutting force deviation in percentage occurred as follows: first, the absolute difference between the actual cutting force value and the theoretical (expected) value was calculated. The obtained difference divided by the theoretical cutting force value, and then the result multiplied by 100 to represent the deviation as a percentage. This percentage calculation allows for a standardized assessment of the difference between the expected and actual cutting force values. It provides a convenient comparison, enabling a clearer understanding of how much the actual results differ from the predicted ones, thus offering a better insight into the accuracy and precision of the forecasts.

The discrepancies between the experimental data and the calculated cutting force values, determined based on the fitted coefficients, appear to be minor and do not exceed 5.68%.

IV. Conclusions

As a result of the experiment on machining the 20CrMn material using the SNMG 15 06 16-PR 4425 T-Max® P insert for turning, the team developed a model for calculating the cutting force (4), where a, f, and V represent the cutting parameters:

$$P = 10 \times 271.289 \times a^{0.95624} \times f^{0.5119} \times V^{-0.32398} \times 1.2797 \quad (4)$$

As the calculations in Table 3 show, the deviation of the theoretical model did not exceed 5.68% from the results obtained on the test stand.

Table 3: *Analysis of model deviation*

№	Cutting force measured on the stand (H)	Cutting force measured on the stand (H)	Cutting force measured on the stand %
1	330,90	330,9	0,00%
2	428,10	405,0854204	-5,68%
3	292,60	278,7254215	-4,98%
4	121,20	121,2000008	0,00%
5	160,90	153,4037901	-4,89%
6	151,30	147,2087003	-2,78%

Based on the results of preliminary modeling of cutting forces for a similar material using a tool with analogous geometry and cutting parameters (feed rate s = 0.51 mm/rev, depth of cut t = 1.5 mm, cutting speed V = 238 m/min), the following data obtained:

- According to the source https://www.walter-tools.com, the cutting force will be 402.56 Fc/N at a specific material removal rate of 246.21 cm³/min and a power of 6.75 Pmot/kW.
- According to the source https://www.sandvik.coromant.com, the specific material removal rate will be 246.00 cm³/min, and the cutting power will be 6.70 Pmot/kW. This will result in approximately 400.0 Fc/N.

- Due to the use of new materials and increased cutting speeds, the cutting force calculations based on the methods of Baranovsky and Kosilova are no longer relevant.

As shown, the results presented by online services predictably inflated since their primary goal is to ensure the claimed tool life. However, actual tests demonstrate that the cutting forces are 17.8% lower than indicated by these services. Cutting tool manufacturers intentionally increase the calculated cutting force values to create a safety margin that guarantees the functionality of their tools even under higher loads. This practice aims to prevent breakage and extend tool life, which is critically important for customer satisfaction and reducing the likelihood of unexpected failures during production.

Nonetheless, it is essential for users to recognize the difference between theoretical data and actual performance metrics. Understanding that actual cutting forces may differ from specifications enables operators to make more informed decisions and adjustments during machining. This knowledge contributes to optimizing performance, reducing tool wear, and enhancing overall efficiency in production processes.

Recognizing that tool manufacturers inflate calculated cutting forces underscores the importance of designing machining processes based on empirical data. By conducting real tests and collecting actual results, operators and engineers can accurately adapt their processes to practical requirements. This approach not only ensures the safety and reliability of tools but also increases the overall efficiency of the production process.

In conclusion, the critical importance of designing and optimizing machining processes based on real data stands out. This enables manufacturers to make informed decisions, reduce production costs, enhance product quality, and maximize tool life, while effectively responding to the demands of specific machining tasks.

This work was supported by the Azerbaijan Science Foundation-**Grant № AEF-MGC-2024-2(50)-16/01/1-M-01**

References

[1] S. A. Tobias and W. Fishwick, "A theory of Regenerative chatter," The Engineer-London, 1958.

[2] J. Tlusty and M. Polacek, "The Stability of Machine Tools against Self Excited Vibrations in Machining," International research in production engineering, ASME, 1963, pp. 465-474.

[3] D. Montgomery and Y. Altintas, "Mechanism of Cutting Force and Surface Generation in Dynamic Milling," ASME, Journal of Engineering for Industry, Vol. 113, No. 2, 1991, pp. 160-168.

[4] Y. Altintas, D. Montgomery and E. Budak, "Dynamic Peripheral Milling of Flexible Structures," ASME Journal of Engineering for Industry, Vol. 114, No. 2, 1992, pp. 137-145.

[5] E. Budak and A. Altintas, "Modelling and Avoidance of Static form Errors in Peripheral Milling of Plates," International Journal of Machine Tools and Manufacture, Vol. 35, No. 3, 1993, pp. 459-476.

[6] Balabanov I.P., Balabanova O.N., Gilman V.N. Development of a parametric model for calculating cutting forces for external cylindrical turning of steel 20CrMNTi. IOP Conference Series: Materials Science and Engineering; 2020. 915(1), 012005 DOI: 10.1088/1757-899X/915/1/012005

[7] Bashegurov, S. V., Nasybullin, F. F., Khusainov, R. M., & Faskhieva, Z. R. "Ensuring the safety of operation of the truck with the semi-trailer". Paper presented at the IOP Conference Series: Materials Science and Engineering, 2019, 632(1) doi:10.1088/1757-899X/632/1/012017

[8] Leushin, I.O., Leushina, L.I. Balabanov, I.P. Savin, I.A. Production of moulding cores and waterglass mixtures using "dry ice" for steel and iron casting. CIS Iron and Steel Review, 2021, 21, p. 34–37 DOI:10.17580/cisisr.2021.01.05

[9] Krastyaninov, P. M., & Khusainov, R. M. "Selection of equipment for machining processing

of parts using NX and TEAMCENTER programs", Paper presented at the IOP Conference Series: Materials Science and Engineering, 2016, 134(1) doi:10.1088/1757-899X/134/1/012041

[10] Shaparev, A., Savin , I., Ptichkin, S., Khankishiyev, I., Mirzayev, A. Punches and Matrices Recovery for Hot Punching by Electric Arc Hardfacing. Advances in Science and Technology, 2024, 148, P. 65–71 DOI:10.4028/p-q4tfAa

[11] Gilman V.N., Faskhutdinov A.I., Balabanov I.P. Increase effectiveness of shaving by using wear-resistant coatings and preliminary modeling cutting (2020) Solid State Phenomena, 299 SSP, pp. 839-844. doi: 10.4028/www.scientific.net/SSP.299.839

[12] W. T. Corpus and W. J. Endres, "A High Order Solution for the Added Stability Lobes in Intermittent Machining," in Proceeding of the Symposium on Machining Processes, No. MED-11, 2000, pp. 871-878. Paper number DETC97 /VIB-4021.

[13] Korovin V. A., Leushin I.O. Balabanov I.P. Savin I.A. Increase of resistance of steel moulds using the complex modifier INSTEEL-7. CIS Iron and Steel Review. – 2024. – Vol. 27. – P. 31-34. – DOI 10.17580/cisisr.2024.01.05

[14] M. A. Davies, J. R. Pratt, B. Dutterer and T. J. Burns, "Stability Prediction for Low Radial Immersion Milling," Journal of Manufacturing Science and Engineering, Vol. 124, No. 2, 2002, pp. 217-225.

[15] Yusubov N.D., Khankishiyev I.A., Abbasova H.M., Mammadov E.D., Huseynov R.A. Matrix models of machining errors in multi-tool multi-carriage adjustments (2023). International Journal on Technical and Physical Problems of Engineering, 15(3), pp. 309-315

[16] Yusubov, N.D. Matrix models of the accuracy in multitool two-support setup. Russian Engineering Research, 2009, 29(3), pp. 268–271. DOI: 10.3103/S1068798X09030125

[17] Abbasov V., Amirov F., Amirli S., Hasanli S., Frana K., Formation of Shaft Accuracy during Mechanical Processing on CNC Machines (2024). Advances in Science and Technology, 148, pp. 81 - 86, DOI: 10.4028/p-N5uKb8

[18] Yusubov N., Abbasova H. Models of Cutting Forces in The Matrix Theory of Multitool Machining Accuracy. Key Engineering Materials, 2024, 979, pp. 27–38

[19] Rasulov, N.M., Nadirov, U.M., Alekberov, M.Z. Generalized Assessment of Machined Surfaces Quality. Russian Engineering Research, 2020, 40(10), pp. 822–825. DOI: 10.3103/S1068798X20100202

[20] A. Savin Laser hardening of stamps in the conditions of a large engineering company. Diffusion and Defect Data. Pt A Defect and Diffusion Forum, vol. 410 DDF, pp. 450–455, 2021. DOI: 10.4028/www.scientific.net/DDF.410.450.

DEVELOPMENT OF AN ALGORITHM FOR AUTOMATIC CUTTING TOOL SELECTION

Leonid Shipulin[1], Egor Shulezhko[2], Sadaqat Mehdiyeva[3], Konul Shammadova[3]

[1]Head of UNID, SUSU, Russia, Chelyabinsk
[2]Postgraduate student at SUSU, Russia, Chelyabinsk
[3]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan
shipulinlv@susu.ac.ru, schuleshko21@mail.ru, sadaqatme@aztu.edu.az
konul.shammadova@aztu.edu.az

Abstract

An inexperienced technologist will have to spend a lot of time to select a cutting tool, so a neural network for cutting tool selection is needed. To create a neural network, we need to understand the algorithm, how to select the tool now. The purpose of this paper is to analyse domestic and foreign sources in the selection of cutting tools. The methodology in this study is to find an algorithm from all possible sources. Each source found is analysed to find the algorithm. A block diagram has been developed, and this is the algorithm itself to create an automatic selection of cutting tools. These publications have shown that work in the direction of automatic selection of cutting tools is ongoing and are of scientific and practical interest. But the majority of works have declarative character with the absence of such important information as: criteria of criteria of cutting tool selection, results of approbation, influence of selection results on the cost of technological operation. on the cost of technological operation.

Keywords: cutting tool, algorithm, automation, cutting tool selection

I. Introduction

At machine-building enterprises, workpieces are processed on metal-cutting machines with cutting tools. Nowadays there is a large number of manufacturers of cutting tools and each has a wide range of foreign cutting tools produced and supplied. The total nomenclature of modern cutting tools in some cases numbers in the thousands for specific types of machining. Selection of a suitable cutting tool by an engineer is a regular technical task and is carried out on the basis of advertising materials or production experience. The biggest problems in selection of cutting tools can be experienced by small machine-building enterprises due to the absence of a highly qualified specialist or lack of production experience.

Traditionally, enterprises approach the issue of tool selection in different ways: they select tools from the catalog of a tool manufacturer, turn to an engineering company, and sometimes use what is purchased by the purchasing service without a technical specification. Such approaches can lead to the fact that in reality an unsuitable tool may be used, with low durability or high cost. Selecting the optimum, or at least rational cutting tool, is a widely demanded task that, if done correctly, will result in machining with the required productivity, cutting tool consumption and acceptable cost

Artificial neural network solutions are becoming more and more advanced and popular, so we can assume that in the future artificial neural networks will be widely used due to a better understanding of their underlying principles. In our goals and objectives, we need a neural network to assist the technologist and select the right cutting tool. To properly train artificial intelligence, we need to understand how cutting tools are selected for certain operations in general.

Nomenclature of cutting tools is determined on the basis of analysis of shapes, dimensions, required accuracy and roughness of main and additional surfaces, taking into account the type of selected workpiece. The basis for selecting the tool nomenclature is the rules of machining the main and additional surfaces of workpieces (surface transitions, tool paths, modes, etc.).

The choice of cutting tools lies in the experience of the technologist, but there are also reference books for calculating cutting modes [1,3,5,6], which include the choice of cutting tools, as well as in training manuals [2,4] are given criteria for the choice of cutting tools. Today, the choice of cutting tools can be made by catalogues [7]. The selection criteria differ, but not significantly.

Despite the advantages of automatic selection of cutting tools in catalogues, it is not in wide demand among not large productions, yes we can say large productions rarely use this system because it does not show the economic benefits and durability of the tool.

The purpose of this paper is to analyses domestic and foreign sources in cutting tool selection, to determine the cutting tool selection algorithm.

A more in-depth study of this issue will give an understanding of what criteria are important for the selection of cutting tools to create a programmer and facilitate the work of inexperienced technologists.

II. Methods

The choice of cutting tools is in limbo because in manufacturing plants the choice lies in the expertise of the technologist. Today there are a large number of tools and a large number of manufacturers, and it is not clear whether a given tool will be effective. Technologists spend a lot of time experimenting with this or that tool. And it should not be excluded that there may not be an experienced technologist. An inexperienced technologist will have to conduct experiments again, as there are no records on the tool. To summarize, it will be necessary to spend a lot of time for the inexperienced technologist to conduct experiments and it will be permanent. It is possible that some data is kept by the companies, but this is not disclosed. The purpose of this paper is to analyses domestic and foreign sources in the selection of cutting tools.

Based on the objectives of the tasks, in order to determine the methodology it is necessary to analyses all possible sources for this research. The first method is to analyses manuals for the selection of cutting tools [1, 2, 3, 4, 5, 6]. Usually guides on the calculation of cutting modes and in them the tool selection is determined by maps.

In the reference book of Guzeyev V.I., Batuev V.A., Surkov I.V. [1] the tool selection starts with the choice of the machine tool. Next, the material of the cutting part of the tool and the method of fixing the insert are selected. Based on the machining conditions, the angles in plan view are selected. Other geometric parameters of the cutting part are determined (back angle, front angle, shape of the front surface, chamfer width along the main cutting blade, radius of cutting edge rounding, radius of the cutter tip). The normative period of durability of the selected cutting tool is given.

In the textbook by Pozdnyakova I.V. [2] the choice of tools begins with the equipment. Next, the material of the cutting part and the material of the holder are selected. Based on the equipment, the cross-section of the holder is selected, which is suitable for the machine. The type of cutter, cutter design, insert angle, structural dimensions of the cutter, insert number and dimensions, shape of the front surface and sharpening of the cutter are selected. Finally, the geometric parameters of the

cutter, the permissible wear on the back surface and the resistance are selected.

In the reference book on metal cutting processing [3], tool selection is done sequentially by selecting the type of cutter, the shape of the front surface, the angle of sharpening of the cutter, and the geometric parameters of the cutting part of the cutter.

In the textbook by M.A. Bolotov, A.N. Zhidyaev, N.D. Pronichev and A.I. Khaimovich [4], the cutting tool is selected by sequentially choosing the insert mounting system, the size of the holder and the shape of the insert, the geometry of the insert and the material grade of the cutting part, the dimensions of the insert, the value of the radius at the top of the insert.

In the reference book of Baranovsky Yu. [5] the type and design of the cutter, material of the cutting part, material of the holder and its geometrical parameters, angular parameters of the cutting part of the cutter and durability period are selected.

In the reference book of machine-building technologist Kosilova A.G. and Mesheryakov R.K. [6], the material of the cutting part is selected, and then the assortment of cutting tools according to GOST is presented. The tool itself is selected based on the qualification of the process engineer.

The second method is the analysis of catalogues of different companies. The analysis of modern catalogs of cutting tools of various manufacturers has been carried out: Sandvik [7], ISCAR [8], Mitsubishi [9], Kennametall [10], Walter [11], Dormer Tools [12], SEKO [13], Korloy [14], OKE [15], VIRIAL [16], KZTS [17]. The methodology of tool selection is generally identical: the choice starts with holders (dimensions in cross-section, length), inserts (shape, dimensions) and the way of their fastening. Next, the material of the cutting part and chipbreaker is selected based on the material of the workpiece. Then the cutting pattern is selected. Further selection of suitable tools is left to the technologist. Some manufacturers have the option of selecting tools via online services. The third method is to analyses all possible articles on tool selection [18, 19, 20, 21, 22].

The authors [18] proposed an approach for automatic selection of a special cutting tool for machining complex surfaces, based on a 2D drawing. This approach was implemented by creating a program using a self-learning network (ResNet), which is able, from the available database of drawings of complex surfaces, to select the tool Fig. 39. The working principle of the developed software was also presented.

The novelty of the work is the emergence of a new method of automatic cutting tool selection, as well as the creation of software using a self-learning network. It significantly increases the efficiency of production preparation and facilitates the work of the technologist. However, the decision on the final selection of cutting tools is left to the technologist, and the main criteria for the selection of cutting tools and the economic justification of this choice are not specified.

The author's work [19] analyzed publications that showed various methods of providing automated selection of cutting tools, namely: the creation of an algorithm with an implemented database containing all the necessary information about cutting tools; introduction of a self-learning network (neural network) into the selection algorithm; algorithm of automated selection of cutting tools on 2D drawings; application of mathematical models in algorithms.

The authors [20] present a method for automating the technological preparation of production, including the calculation of cutting modes and the selection of optimal cutting tools. An algorithm is proposed in the form of a block diagram, on the basis of which specialized software is developed that allows creating a database of cutting tools, machined materials, taking into account the parameters of the cutting part of the tool, the criteria of tool operability, surface quality, productivity and cost-effectiveness in calculations. However, the work has a declarative character and the main criteria for the selection of cutting tools, the validity of their selection and practical confirmation of the performance of this software are not disclosed.

The authors [21] present an algorithm that recognizes the topography of the part and determines the best set of tool diameters for cavity milling. Since a relatively large tool diameter reduces the milling time but cannot cover all the corners, on the other hand, a small tool that can

cover all the corners requires more time for milling. As a result, this algorithm showed a reduction of about 17% time for roughing the workpiece. In addition, the algorithm saves time spent in analyzing the geometry and deciding which tools to use. However, the paper does not provide an economic justification for cutting tool selection and technological criteria for tool selection.

The authors' publication [22] presents a database containing all necessary information on tools, as well as on quick-change non-transferable inserts for them and recommended cutting modes. The database is based on Mitsubishi Carbide catalog. An application in C# programming language was developed for convenient work with the database and quick retrieval of necessary information in the form of a table. The algorithm of the application was described. However, from the conclusions of the work it is not clear what level of automation of this system, the main criteria for the selection of cutting tools and its economic justification.

Allocation by method it was necessary to understand the cutting tool selection algorithm itself, more reasonable and working algorithm on selection for further programmer creation. The programmer, which will be used by a non-experienced technologist, will program the algorithm that will be selected in this study.

Many sources have been analyses for tool selection and exactly justified in the selection very few and almost everywhere the economics and durability of the tool is not considered and this study needs to be investigated further. From the analysis at the moment the reference books and catalogues which adequately select the tool are indicated.

III. Results

Since the developed algorithm is aimed at solving the problem of automatic tool selection, the reduction of labor intensity of technological design occurs at this stage. Economic efficiency of design automation consists of four factors: increase in labor productivity; decrease in production costs; reduction of preparation and production time of products; increase in the quality of products and processes.

Labor productivity is ensured by:

- Elimination of the stage of studying catalogs and manual determination of tool parameters.

Reduction of production costs is ensured by:

- Exemption of skilled workers from non-creative labor;
- Increasing the efficiency of CNC equipment utilization due to the use of the most productive tools.

Reduction of terms of preparation and production of products is provided by:

- Acceleration of development of technological documentation.
- Reducing the cycle of technological preparation of production.

The results are presented in the form of a flow chart (see Figure 1).

Using the reference book for assigning cutting modes, the initial data for the initial stage of algorithm development were determined, the initial data ensuring correctness of cutting tool selection, such as: shape; machining stage; workpiece qualification; cutting tool qualification were determined.

An algorithm has been developed that takes into account the principle of CAM systems and the procedure for selecting a cutting tool from a directory. This algorithm consists of eight stages, which contain about twenty-two steps that provide automatic selection of cutting tool in CAM-system.

IV. Discussion of results

Analysis of publications has shown various methods of providing automated selection of cutting tools, namely: creation of an algorithm with a embedded database containing all necessary information about cutting tools; introduction of a self-learning network (neural network) into the selection algorithm; introduction of an algorithm with an automated cutting tool selection system. tools; introduction of a self-learning network (neural network) into the selection algorithm; algorithm for automated selection of cutting tools based on 2D drawings; application of mathematical models in the algorithms.

These publications have shown that work in the direction of automatic selection of cutting tools is ongoing and are of scientific and practical interest. But the majority of works have declarative character with the absence of such important information as: criteria of criteria of cutting tool selection, results of approbation, influence of selection results on the cost of technological operation. on the cost of technological operation.

Thus, it can be concluded that the topic of automatic selection of cutting tools in automated production systems is promising. The topic of automatic cutting tool selection in automated production systems is a promising one and there is a need for a holistic approach to its development, based on clear technological criteria of cutting tool selection, economic justification of this choice and the possibility of its application in automated production preparation systems. The created automated system of cutting tool selection allows to significantly reduce the time and quality of work related to the selection of cutting tools. The software system recognizes the design and manufacturing model of the part from the drawing and 3D model, selects a suitable cutting tool according to this model and then determines the optimal tool parameters.

In addition, the use of automated tool selection system allows to increase the efficiency of equipment utilization by selecting the most productive tool. the gain in productivity is from 5 to 50%. It is not possible to give an exact estimate for this parameter, as it strongly depends on the human factor and is random in nature.

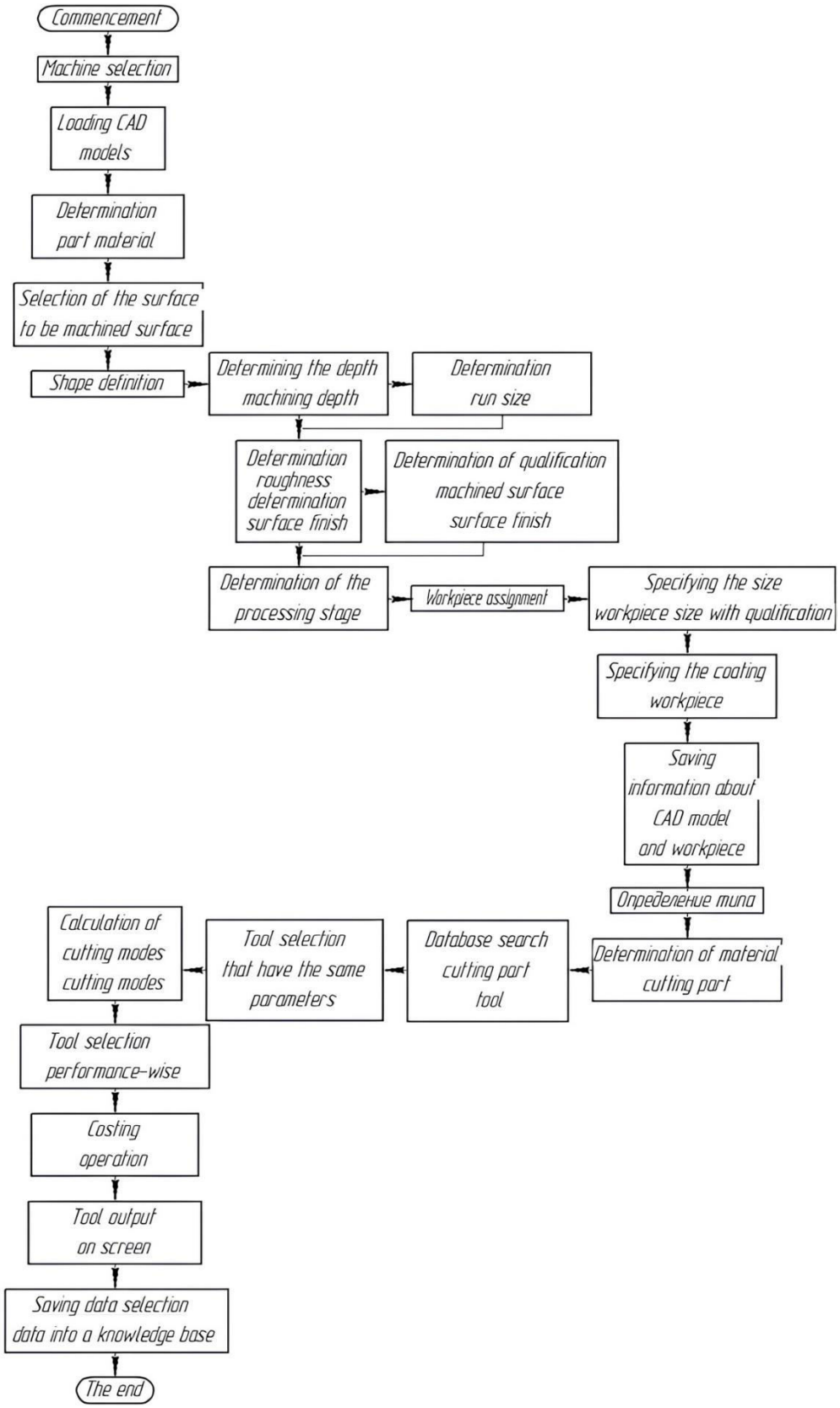


Figure 1: *Block diagram of the algorithm*

References

[1] Guzeyev V.I., Batuev V.A., Surkov I.V., Cutting modes for lathes and boring-milling-boring machines with numerical programme control: Reference Book 2nd ed./Edited by V.I. Guzeyev. M.: Mashinostroenie, 2007, 368 p.

[2] Pozdnyakova, I.V. Calculation of the cutting modes at turning on the computer numerical control system: a textbook for course and diploma design / I.V. Pozdnyakova, T.V. Shostak. - Chelyabinsk: CPI, 1989. - Ч. 2.

[3] Reference book on machining metals by cutting / F.N. Abramov, V.V. Kovalenko, V.E. Kovalenko, V.E. Lyubimov et al. - K. Technics, 1983. - 239 p., ill. - Bibliography: pp. 235-237.

[4] M.A. Bolotova, A.N. Zhidyaeva, N.D. Pronicheva, A.I. Khaimovicha Recommendations on the Assignment of Cutting Modes and Tool Selection: Methodical Instructions - Samara: Izd-vo Samar.gos. aero-cosm. un-ta, 2010.-85 p.: ill.

[5] Baranovsky Y.V., Metal cutting modes. Reference book. Izd. 3-th, worked and supplemented. M., Mashinostroenie, 1972.

[6] Reference book of technologist-machinostroitechnik. In 2 Vol. 2/ Edited by A.G. Kosilova and R.K. Mesheryakov. - 4th ed., rev. and supplement. - M.: Mashinostroenie, 1985, 496 pp., ill.

[7] Based on the materials of the site: www.coromant.sandvik.com/ru

[8] Based on the materials of the site: www.iscar.ru

[9] Based on the materials of the site: www.mitsubishicarbide.ru

[10] Based on the materials of the site: www.kennametal.com

[11] Based on the materials of the site: www.walter.com

[12] Based on the materials of the site: www.dormertools.ru

[13] Based on the materials of the site: www.secotools.com

[14] Based on the materials of the site: www.korloy.com

[15] Based on the materials of the site: oke-russia.ru

[16] Based on the materials of the site: www.virial.ru

[17] Based on the materials of the site: www.kzts.ru

[18] Guanghui Zhouab Xiongjun Yangb Chao Zhangb Zhi Lib Zhong dong Xiaoc.«Deep learning enabled cutting tool selection for special-shaped machining features of complex products» 1-11s.

Electronic Source: https://www.sciencedirect.com/science/article/abs/pii/S0965997819300225

[19] Batuev, V. V. Development of the algorithm of automatic selection of milling cutters for CAM systems / V. V. Batuev, A. S. Onishchenko // Innovative ideas in mechanical engineering : Collection of scientific papers of the All-Russian Scientific and Practical Conference of Young Scientists, St. Petersburg, May 24-26, 2022 / Edited by A. A. Popovich, D. P. Gasyuk. - St. Petersburg: POLYTEKH-PRESS, 2022. - C. 109-114. - EDN NBACTZ.

[20] Smirnova D.A., Rychkov D.A. “Automation of cutting tool selection for mechanical processing of materials taking into account production criteria”. Bratsk State University, Bratsk, Russia. 255-260 p. Electronic source: https://yadi.sk/d/qRqyWqtGjBhb9w

[21] Nícolas de Lima, Rossini Severino, Rodrigo Berretta Käsemodel, Adriano Fagalide Souza. “New algorithm identifies the best set of cutting tools to mill cavities”. 51-53c. Electronic source: https://www.sciencedirect.com/science/article/pii/S2212827121006545

[22] Golod I.A., Murashko V.S. “Automation of cutting tool selection in turning with longitudinal feed”. Gomel State Technical University named after P.O. Sukhoi. P.O. Sukhoi, Gomel, Belarus. 32 c. Electronic source: https://yadi.sk/d/YRAFZFCBzqXaBA

STUDY OF DYNAMIC CHARACTERISTICS OF THE ROTARY HONING PROCESS IN THE PROCESSING OF NON-RIGID THIN-WALLED PARTS

Aydin Gafarov[1], Isag Khankishiyev[1], Alihuseyn Haziyev[1], Irada Abbasova[2]

[1]Department of Shipbuilding and Ship Repair, Azerbaijan State Marine Academy, Baku, Azerbaijan
[2]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan
aydin.qafarov@hotmail.com, isaq.xankishiyev@asco.az, alihuseyn.haziyev@asco.az, i.abbasova@aztu.edu.az

Abstract

The article studies the dynamic characteristics of the rotary honing process when machining high-precision non-rigid thin-walled parts. The process is modeled and optimized to determine its rational parameters, providing the lowest cutting forces.

Keywords: rotary honing, process, dynamics, forces, deformation, cutting, modulation, optimization, factors.

I. Introduction

In modern mechanical engineering, due to high requirements for the accuracy and quality of machine parts processing, there is often a need to create new progressive processing methods and special cutting tools that provide high quality of the processed surface, along with increased tool life and process productivity [1-5]. One of these processing methods is the rotary honing method, which provides a roughness of the processed surface within 0.63-2.5 μm, accuracy of 9-10-th quality, and an increase in productivity by 1.5-1.9 times. The widespread use of this process is currently limited by a number of circumstances, in particular, by the relatively small amount of study of the process. The recommendations in the technical literature on choosing the optimal parameters for rotary honing are too general and cannot be extended to all types of processing. The possibilities of rotary honing in the processing of low-rigidity thin-walled parts such as bushings have not been studied at all.

Rotary honing has a number of advantages over other processing methods. These include the corrective ability of the process, high productivity, ensuring high quality and accuracy of processed surfaces, stability of the operational properties of parts, tool durability, etc.

The state of the problem of rotary honing, along with the known advantages, is also characterized by increased cutting forces, which can negatively affect the geometric accuracy of non-rigid thin-walled parts. In this regard, there is an obvious need to study the influence of the main design and technological parameters of the process on the cutting forces, with the aim of optimizing these parameters in relation to ensuring processing with extremely small deformations of parts.

II. Purpose of the work

Research of dynamic characteristics of rotary honing process during processing of high-precision non-rigid thin-walled parts. Modeling and optimization of the process for determination of its rational parameters, providing the least components of cutting forces.

III. Solution tasks

The total error of mechanical processing, as is known, is determined by the combined effect of a number of factors that generate primary elementary errors.

When processing thin-walled, relatively low-rigid parts such as bushings, the accuracy of processing is significantly affected by elastic movements of the elements of the technological system and, in particular, the workpiece under the action of cutting forces. In addition, the accuracy of the parameters of thin-walled parts can be affected by other force factors.

The article examines some issues of calculating the forces acting on the inner surface of bushings, as well as the influence of the main parameters ($P_{sp.}$, 3, $V_{rec.}$, $V_{per.}$) of rotary honing on the components of the cutting force.

To solve the specified problem, thin-walled bushings are considered as shells.

We calculate the concentrated forces affecting the inner surface (Fig. 1).

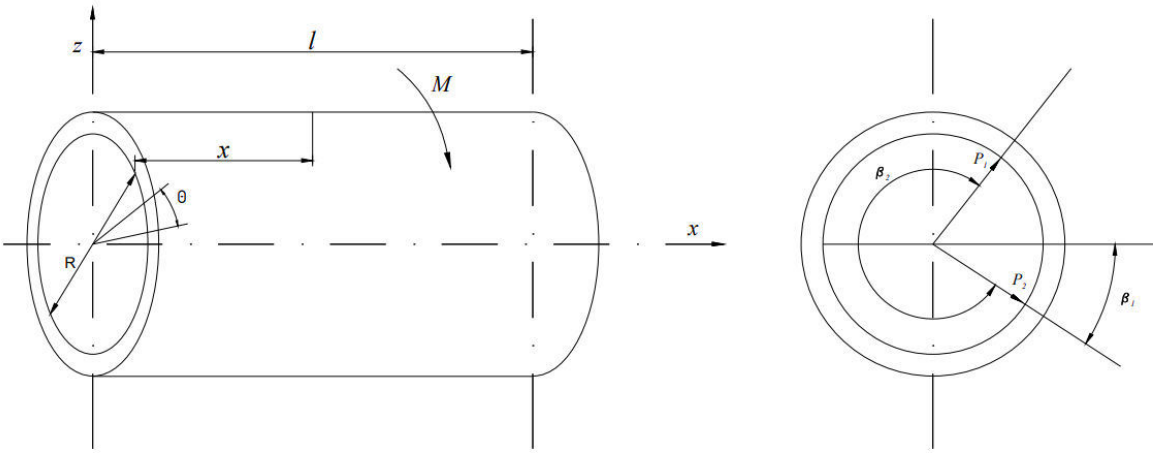


Figure 1: *Scheme for determining concentrated forces acting on an elementary section of the shell surface*

We assume that one end of the shells is fixed. In this case, a concentration of force is applied to a certain area.

We will determine the stress and strain state of the shell for the specified case (2).

In this case, the deformation equations will have the form

$$\frac{\partial^4\varphi}{\partial x^4}+2\frac{\partial^4\varphi}{\partial x^2\partial y^2}+\frac{\partial^4\varphi}{\partial y^4}=\frac{E\delta}{R}\frac{\partial^2\omega}{\partial x_2^2}\frac{1}{R}\frac{\partial^2\varphi}{\partial x^2}+D\left(\frac{\partial^4\omega}{\partial x^4}+2\frac{\partial^4\varphi}{\partial x^2\partial y^2}+\frac{\partial^4\varphi}{\partial y^4}\right)=qz \quad (1)$$

We will seek the solution of these equations in the form of double trigonometric series

$$\omega_1=\sum\sum A_{1mn}\sin\frac{m\pi x}{l}\sin n\theta$$

$$\varphi_1=\sum\sum B_{1mn}\sin\frac{m\pi x}{l}\sin n\theta$$

These functions satisfy the following boundary conditions:

$$\left.\begin{matrix}\omega = 0\\ M_x = 0\end{matrix}\right\} \quad \begin{matrix}x = 0\\ x = l\end{matrix}$$

$$\left.\begin{matrix}V = 0\\ N_x = \dfrac{\partial^2 \varphi}{R\partial\theta^2} = 0\end{matrix}\right\} \quad \begin{matrix}x = 0\\ x = l\end{matrix}$$

In addition, at $\theta = 0$, the deflection and stress functions, as follows from the nature of the loading, vanish and are odd functions of the angle θ.

To determine the coefficients A_{1mn} and B_{1mn}, it is necessary to substitute the expressions ω_1 and φ_1 into equations (1). But first, we will expand the external acting load into a double trigonometric series by the sought functions. In this case, the external load is represented as a concentrated moment, which is statically equivalent to a pair of forces with a shoulder.

$$d = [2\pi - \beta_2] + \beta_1 R$$

Therefore, we can write

$$P_1 = \frac{0{,}5M}{R\beta_1},\ P_2 = \frac{0{,}5M}{(2\pi-\beta_2)^2}$$

Thus, in the case under consideration, the external load is represented in the form of two concentrated forces

$$P_1 = P_2$$

To reduce this load to the dimension of distributed pressure q_z, it is necessary to represent it in the form

$$q_z = \frac{P_1}{\Delta F}$$

Where $\Delta F = \Delta S_1 \Delta X_1$ is a small area on which the force is applied;
ΔS, ΔX - are the dimensions of this area in the circumferential and axial directions.
Then

$$q_z = \frac{P_1}{\Delta S_1 \Delta X_1} = \sum_n \sum_m C_{mn} \sin\frac{m\pi x}{l} \sin n\theta$$

Let's multiply the right and left sides of this expression by

$$\sin\frac{m\pi x}{l} \sin n\theta d\theta dx$$

and integrate the right side over X from 0 to l and over θ from 0 to 2π, and the left side over X from the value of X to $X + \Delta X$ and from the value of β_1 to $\beta_1 + \Delta\beta_1$. Solving the equation C_{mn} obtained after this integration, we determine that

$$C_{mn} = \frac{2\dfrac{P_1}{\Delta S_1 \Delta X_1}}{\pi R l} \frac{l}{\pi m}\left[\cos\frac{\pi m(X_1 + \Delta X)}{l} - \cos\frac{m\pi x_1}{l}\right] \cdot\cdot \frac{R}{n}[\cos n\,(\beta_1 + \Delta\beta_1) - \cos n\beta_1]$$

Passing to the limit as $\Delta S \to 0$, $\Delta X_1 \to 0$
(where $\Delta\beta_1 = \frac{\Delta S_1}{R}$, $\beta_1 = \frac{S_1}{R}$) we obtain

$$C_{mn} = \frac{2P_1}{\pi^2_{mn}} \cdot \lim_{\Delta x_1 \to 0} \frac{\cos\frac{\pi m(\Delta X + X_1)}{l} - \cos\frac{m\pi x_1}{l}}{\Delta X_1} \cdot \lim_{\Delta S_1 \to 0} \frac{\cos\frac{n(S_1+\Delta S_1)}{R} - \cos\frac{nS}{R}}{\Delta S_1}$$

or

$$C_{mn} = \frac{2P_1}{\pi Rl} \sin\frac{m\pi x_1}{l} \sin n\beta_1$$

and

$$q_z = \frac{2P_1}{\pi Rl} \sum_n \sum_m \sin\frac{m\pi x_1}{l} \sin n\beta_1 \sin\frac{m\pi x}{l} \sin n\theta \qquad (2)$$

Here X_1, β_1 - are the coordinates of the point of application of force P_1, taking into account the corresponding derivatives of the function φ_1 and ω_1 and the value of the external load q_z, we obtain:

$$B_{1mn}\left[\left(\frac{m\pi R^2}{l}\right) + n^2\right] = -E\delta R A_{1mn}\left(\frac{m\pi R}{l}\right)^2 -$$
$$-\frac{B_{1mn}}{R^3}\left(\frac{m\pi R}{l}\right)^2 + \frac{DA_{1mn}}{R^4}\left[\left(\frac{m\pi R^2}{l}\right) + n^2\right]^2 = \frac{2P_1}{\pi Rl} \sin n\beta_1.$$

Solving the obtained equations with respect to the parameters A_{1mn} and B_{1mn}, we determine

$$A_{1mn} = \frac{2P_1 R\left[\left(\frac{m\pi R^2}{l}\right) + n^2\right] \sin\frac{mpx_1}{l} \sin n\beta_1}{\pi E\delta l\left\{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4\right\}}$$

$$B_{1mn} = \frac{2P_1 R^2\left(\frac{m\pi R}{l}\right)^2 \sin\frac{m\pi x_1}{l} \sin n\beta_1}{\pi l\left\{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4\right\}}$$

By entering the force values into these expressions using the formula $p_1 = \frac{0.5\mathcal{M}}{R\beta_1}$, we obtain

$$A_{1mn} = \frac{\mathcal{M}\left[\left(\frac{m\pi R^2}{l}\right) + n^2\right]^2 \sin\frac{m\pi x_1}{l} \frac{\sin n\beta_1}{\beta_1}}{\pi E\delta l\left\{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4\right\}}$$

$$B_{1mn} = \frac{MR\left(\frac{m\pi R}{l}\right)^2 \sin\frac{m\pi x_1}{l} \frac{\sin n\beta_1}{\beta_1}}{\pi l\left\{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right) + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4\right\}}$$

Where A_{1mn} - is the amplitude coefficient of the sine wave when solving trigonometric series;
B_{1mn} - is the amplitude coefficient of the cosine series when solving trigonometric series.

For the force P_2 we obtain similar expressions, but with opposite signs and replacing β_1 with β_2 and $n\beta_1$ with $n(2\pi - \beta_2)$, i.e.

$$A_{2mn} = \frac{M\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^2 \sin\frac{m\pi x_1}{l_1}\frac{\sin n\beta_2}{2\pi - \beta_2}}{\pi E\delta l\left\{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4\right\}}$$

$$B_{2mn} = \frac{MR\left(\frac{m\pi R}{l}\right)^2 \sin\frac{m\pi x_1}{l}\frac{\sin n\beta_2}{2\pi - \beta_2}}{\pi l\left\{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4\right\}}$$

Passing in the obtained expressions for A_{1mn}, B_{1mn}, A_{2mn}, B_{2mn} to the limit $\beta_1 \to 0$ and $\beta_2 \to 2\pi$, we have

$$A_{1mn} = \frac{Mn\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^2 \sin\frac{m\pi x}{l}}{\pi E\delta\left\{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4\right\}}$$

$$A_{2mn} = \frac{Mn\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^2 \sin\frac{m\pi x_1}{l}}{\pi E\delta l\left\{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R^2}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4\right\}}$$

$$B_{1mn} = -\frac{MRn\left(\frac{m\pi R}{l}\right)^2 \sin\frac{m\pi x_1}{l}}{\pi l\left\{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R^2}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4\right\}}$$

$$B_{2mn} = -\frac{MRn\left(\frac{m\pi R}{l}\right)^2 \sin\frac{m\pi x_1}{l}}{\pi l\left\{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R^2}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)\right\}}$$

The complete solution to the problem can be found by adding up the solutions obtained.

Let us calculate the distributed forces acting on the inner surface of the shell. Let us consider the shell (Fig. 2) loaded with a flow of tangential forces transmitted as a result of the action of the torque $M_{to.}$, in section "b", in order to determine the value of P, taking into account the elementary arc d_s, allocated on the cross-section of the shell

$$\omega = \omega_1 + \omega_2 = \frac{2M}{\pi E\delta l}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{n\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^2 \sin\frac{m\pi x_1}{l}\sin\frac{m\pi x}{l}\sin n\theta}{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4} \quad (3)$$

$$\varphi = \varphi_1 + \varphi_2 = \frac{2M}{\pi l}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{n\left(\frac{m\pi R}{l}\right)^2 \sin\frac{m\pi x_1}{l}\sin\frac{m\pi x}{l}\sin n\theta}{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right)^2 + n^2\right]^4 + \left(\frac{m\pi R}{l}\right)^4} \quad (4)$$

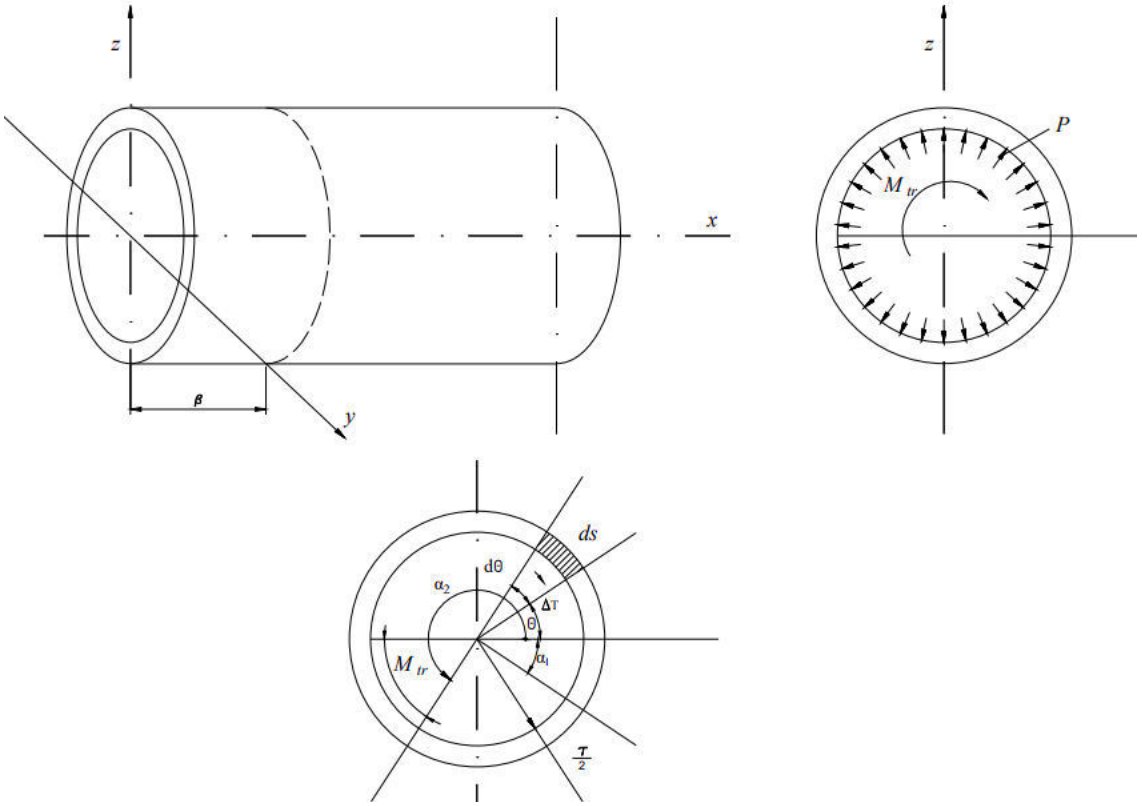


Figure 2: *Scheme for determining concentrated forces acting on the inner surface of shells*

The elementary force dT, acting along the arc d_s, will be equal to

$$dT = pd_s b = prd\theta b(x) \tag{5}$$

Then

$$T = \int_{\alpha_1}^{\alpha_2} prb(x)d\theta = prb(x)(\alpha_2 + \alpha_1) \tag{6}$$

The moment causing the force is equal to

$$M_{torq.} = T \cdot r = pr^2 b(x)(\alpha_2 + \alpha_1) \tag{7}$$

Where

$$P = \frac{M_{to.}}{r^2 b(x)(\alpha_2+\alpha_1)} \tag{8}$$

Substituting (8) into (4) and multiplying by d, we obtain the elementary moment of the tangential load

$$dM = dT \cdot r = \frac{M_{to.} \cdot r \cdot b(x)d\theta}{r \cdot b(x)(\alpha_2+\alpha_1)} = \frac{M_{to.} \cdot d\theta}{(\alpha_2+\alpha_1)} \tag{9}$$

Or section $\theta E[-\beta_1, 2\pi - \beta_2]$

$$M = \int_{-\beta_1}^{2\pi-\beta_2} \frac{M_{to.}d\theta}{(\alpha_2+\alpha_1)} = \frac{M_{to.}}{(\alpha_2+\alpha_1)}(2\pi - \beta_2 + \beta_1). \tag{10}$$

Using the solutions obtained in (2) and (3), taking into account equation (10), we obtain the value of the displacement,

$$\omega = \omega_1 + \omega_2 = \frac{2M_{to.}(2\pi-\beta_2+\beta_1)}{\pi E\delta l(\alpha_2+\alpha_1)} \cdot \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{n\left[\left(\frac{m\pi R}{l}\right)^2+n^2\right]\sin\frac{m\pi x_1}{l}\sin\frac{m\pi x}{l}\sin\theta}{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right)^2+n^2\right]^4+\left(\frac{m\pi R}{l}\right)^4} \tag{11}$$

$$\varphi = \varphi_1 + \varphi_2 = -\frac{2RM_{to.}(2\pi-\beta_2+\beta_1)}{\pi l(\alpha_2+\alpha_1)} \times \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{n\left(\frac{m\pi R}{l}\right)^2\sin\frac{m\pi x_1}{l}\sin\frac{m\pi x}{l}\sin n\theta}{\frac{D}{E\delta R^2}\left[\left(\frac{m\pi R}{l}\right)^2+n^2\right]^4+\left(\frac{m\pi R}{l}\right)^4} \tag{12}$$

Having the expressions of displacement and stress functions, it is possible to obtain all the internal force factors arising in the shell from

$$M_x = \frac{\partial^2\varphi}{R^2\partial\theta^2},\ N_y = \frac{\partial^2\varphi}{\partial x^2},\ N_{x\,y} = \frac{\partial^2\varphi}{R\partial\theta\partial x}$$

$$M_x = D\left(\chi_x + \mu\chi_y\right)$$

$$M_y = D\left(\chi_y + \mu\chi_x\right)$$

As is known, the accuracy of the characteristics of machine parts, along with other force factors, largely depend on the components of the cutting forces that arise during their mechanical processing. This circumstance is especially evident when force technological processes are used in the processing of non-rigid thin-walled parts, such as rotary honing.

Rotary honing, along with its known advantages, is also characterized by increased cutting forces, which can negatively affect the geometric accuracy of non-rigid thin-walled parts. In this regard, there is an obvious need to study the influence of the main design and technological parameters of rotary honing on the components of the cutting force, with the aim of optimizing these parameters in relation to ensuring processing with extremely low forces.

The study was conducted according to the methodology described in work (1) using second-order orthogonal planning. As a result of calculations for coded values of factors, the following mathematical model was obtained, characterizing the influence of specific pressure $P_{sp.}$, grain size of diamond stones 3, reciprocating speed $V_{rec.}$ and peripheral speed $V_{per.}$ on the peripheral $P_{per.}$ cutting force.

$$\begin{gathered} Y_{P_{per.}} = 117{,}41 + 5{,}42X_1 + 3{,}22X_2 - 1{,}91X_3 - 1{,}51X_4 - \\ -0{,}08X_1X_2 - 0{,}07X_1X_3 + 0{,}44X_1X_4 - 0{,}21X_2X_3 - \\ -0{,}07X_2X_4 + 0{,}21X_3X_4 + 0{,}47X_1^2 + 0{,}99X_2^2 + 0{,}46X_3^2 + 0{,}42X_4^2 \end{gathered} \tag{13}$$

The reproducibility of the experiments was checked using the Cochran criterion, the significance of the regression coefficients using the Student criterion, and the adequacy of the model for the significance level $\alpha = 0{,}05$ ($F_P < F_T$) using the Fisher criterion.

For natural values of factors, equation (13) has the form:

$$\begin{gathered} P_{per.} = 97{,}51 + 25{,}41P_{sp.} + 6{,}31\,3 - 4{,}23V_{rec.} - 2{,}11V_{per.} + \\ +1{,}21P_{sp.}3 + 1{,}02P_{sp.}V_{rec.} - 1{,}21P_{sp.}V_{per.} + 0{,}91\,3V_{rec.} - \\ -1{,}01\,3V_{per.} - 2{,}1V_{rec.}V_{per.} + 12{,}41P_{sp.}^2 + 4{,}04\,3^2 - 3{,}11V_{rec.}^2 - 1{,}22V_{per.}^2 \end{gathered} \tag{14}$$

As a result of the calculations, the minimum value of the circumferential cutting force was determined, $P_{per.} = 121{,}12N$, with optimal values of the rotary honing parameters: $P_{sp.} = 0{,}6MPa$, $3 = 160/125\ mkm$; $V_{rec.} = 0{,}16\ m/s$; $V_{per.} = 0{,}51\ m/s$.

The mathematical model of the rotary honing process, characterizing the influence of $P_{sp.}$, 3, $V_{rec.}$, $V_{per.}$ on the axial cutting force $P_{a.f.}$ for coded values of the factors is written in the form

$$Y_{P_{a.f.}} = 75{,}62 + 6{,}11X_1 + 3{,}27X_2 - 2{,}67X_3 - 2{,}05X_4 - \\ -1{,}21X_1X_2 - 2{,}14X_1X_3 + 1{,}23X_1X_4 + 1{,}41X_2X_3 - \\ -2{,}12X_2X_4 - 1{,}13X_3X_4 - 3{,}14X_1^2 - 3{,}21X_2^2 + 2{,}16X_3^2 + 1{,}02X_4^2 \quad (15)$$

For natural values of factors, the mathematical model (15) has the following form

$$P_{a.f.} = 61{,}43 + 21{,}12P_{sp.} + 3{,}24\,3 - 4{,}14V_{rec.} - 2{,}18V_{per.} - \\ -3{,}21P_{sp.}3 - 2{,}41P_{sp.}V_{rec.} - 2{,}56P_{sp.}V_{per.} + 1{,}27\,3V_{rec.} - \\ -3{,}11\,3V_{per.} - 2{,}46V_{rec.}V_{per.} - 4{,}21P_{sp.}^2 - 4{,}56\,3^2 + 3{,}71V_{rec.}^2 + 2{,}16V_{per.}^2 \quad (16)$$

The minimum value of the axial cutting force $P_{a.f.}$ at optimal values of the parameters of the rotary honing process ($P_{sp.} = 0{,}6MPa,\ 3 = 160/125\ mkm;\ V_{rec.} = 0{,}16\ m/s;\ V_{per.} = 0{,}51\ m/s.$) is $P_{a.f.} = 87{,}42N$.

IV. Conclusions

Using the developed mathematical models, it is possible to obtain different graphical dependencies characterizing the influence of the parameters of the rotary honing process on the components of the cutting force $P_{per.}$ and $P_{a.f.}$.

This work was supported by the Azerbaijan Science Foundation - **Grant № AEF-MGC-2024-2(50)-16/01/1-M-01**

References

[1] Gafarov, A.M. Rotary turning /A.M. Gafarov. Baku: - Science, - 2000. - p. 128

[2] Gafarov, A.M. Rotary honing /A.M. Gafarov, G.M. Babayev. Baku: - Science, - 1999. - p. 132

[3] Gafarov, A.M. Study of kinematic features of the honing process when machining external surfaces of high-precision cylindrical parts of marine machines and mechanisms / A.M. Gafarov, I.A. Khankishiyev, S.G. Pashazade // Bulletin of mechanical engineering, - Moscow: - 2023. - N6, - pp. 486 - 490

[4] Gafarov, A.M. Study of dynamic features of the external honing process when machining high-precision parts of marine machines and mechanisms / A.M. Gafarov, I.A. Khankishiyev, S.G. Pashazade // Technology of mechanical engineering, - Moscow: -2023. -N4, - pp.10 - 14

[5] Suleymanov, P.G. Improving the reliability of machines and equipment operated in extreme conditions / P.G. Suleymanov. Baku: - Science, - 2018. - p.308.

RESEARCH OF MAIN DIMENSIONS OF NEW GENERATION SUBSEA CONSTRUCTION VESSELS AND INVESTIGATION OF INFLUENCE OF CHOICE OF DIVING COMPLEX AND REMOTELY OPERATION VEHICLES ON CONCEPT DESIGN STAGE

Rasim Bashirov[1], Alexander Egorov[1], Oyrad Abdullayev[2], Zaur Jafarov[3], Rahim Abdullayev[4]

•

[1]Azerbaijan Technical University, Huseyn Javid Ave.-25, Az-1073, Baku
[1]Marine Engineering Bureau, GEO, Ukraine, Odessa
[2]ASCO Engineering Ltd. Ziya Usifzade Street 48, Az-1003, Baku
[3]SOCAR, Diving department, Director, Azerbaijan, Baku
[4]Maritime University of Szszecin, Student, Poland, Szczecin
rasim_agma@aztu.edu.az, egorovag@meb.com.ua, oyrad.abdullayev@asco.az, Zaur.A.Jafarov@socar.az, r.o.abdullayev@gmail.com.

Abstract

The mobile diving complexes and remotely operated vehicles available to companies operating offshore in the Caspian Sea are described, along with their proposed installation on subsea construction vessels. Based on research, methods for selecting the main dimensions of subsea construction vessels are presented, utilizing a database of vessels with similar functions and taking into account the installation of the described mobile diving complexes and equipment. The optimization of the main dimensions of subsea construction vessels is carried out using various methods, considering the parameters of the diving complexes and remotely operated vehicles. The main dimensions of the proposed subsea construction vessel are determined according to the parameters of the installed diving complexes and remotely operated vehicles.

Keywords: Subsea construction vessel (SSCV), Remote Operation Vehicle (ROV), Diving complex, The vessel main dimension, Optimization model, The vessel modeling

I. İntroduction

This study describes the mobile diving complexes and remotely operated vehicles (ROVs) available to companies working offshore in the Caspian Sea, as well as their proposed installation on Subsea Construction Vessels (SSCVs). Based on previous research [1 - 6], methods for selecting the main dimensions of SSCVs are proposed, utilizing a database of vessels with similar functions and considering the installation of the described mobile diving complexes and ROVs. The optimization of SSCV dimensions is carried out using various methods that take into account the parameters of the diving complexes and ROVs. The main dimensions of the proposed SSCV, based on the parameters of the installed diving complexes and ROVs, are simulated. Additionally, the

study outlines future research opportunities and potential applications. For example, the designed vessel could serve as a carrier for manned submersibles, as described by researchers such as Rahul Bharti, Bhaskaran Pranesh, Dharmaraj Sathianarayanan, Manickavasagam Palaniappan, & Gidugu Ananda Ramadass [7] in relation to offshore operations in the Caspian Sea.

Types of the installation mobile diving complexes and ROV on SSCV.

In the development of a sea oil field, diving complexes and ROVs are widely used for deep-water and technical operations. Considering their application offshore in the Caspian Sea, companies require vessels optimally designed to carry mobile diving complexes and ROVs. The operators working offshore in the Caspian Sea have access to the mobile diving complexes and ROVs illustrated in Figure 1 and described in Tables 1 and 2.

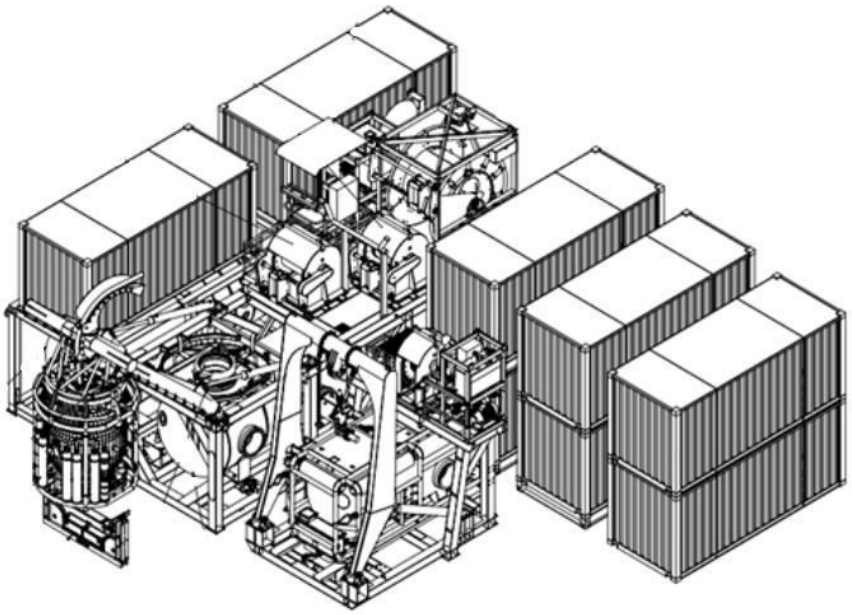


Figure 1: *Mobile diving complex MDDK-200*

For the safe and efficient operation of mobile diving complexes and equipment, it is necessary to design and construct SSCVs according to the requirements of the described mobile technology. In particular, the main dimensions, ship structures, devices, and systems must ensure safe operation at sea. To accommodate mobile complexes and equipment, it is proposed to design a subsea vessel that can also perform technological operations by installing the appropriate mobile technology (such as a mobile drilling complex, pipe-laying, or cable-laying equipment). Thus, an SSCV equipped with mobile technology will be capable of performing the full range of operations typically conducted by specialized vessels, such as diving, drilling, pipe laying, or cable-laying vessels.

Table 1: *The main dimension of the mobile diving complexes installation on SSCV*

MDDK-200		**MDK-60**	
The deep-water diving module with approximate dimension, m	14×16	The davits dimension, with diving basket, m	2,5×6
Reclaim Bag Container, m	2,5×6	Chamber, м	2,5×6
Power module (diesel generator, air compressor) – 2piec., m	2×2,5×6	Power module (diesel generator, air compressor), m	2,5×6
Compressed air cylinders packed	25×50*l*. 10piec.	Compressed air cylinders	16×50*l*. 6piec.
Workshop, m	2,5×6	Workshop, m	2,5×6
Store, m	4×2,5×6	Store, m	2,5×6
Staff, man.	56	Staff, man.	15

Table 2: *The main dimensions of the mobile ROV, installation on SSCV*

ROV complexes «COUGAR XT»	
ROV module, m	2,5x6
Control post module, m	2,5x6
Workshop module, m	2,5x6
Store container, m	2,5x6

Methods of the select of the main dimension of vessels in parameters of mobile diving complexes and ROV.

Considering the above, the main dimensions projected for the described complexes and SSCV hardware are determined using the nomogram developed by the author (Abdullayev O.M., 2024), as shown in Figure 2. At the same time, the main dimensions of the projected SSCV are defined as the sum of the areas of the open deck, the area occupied by mobile technology equipment, and the area necessary for servicing and performing technological operations, as expressed by Equations (1), (2), and (3). As a result, the main dimensions obtained according to the nomogram in Figure 2 and the formulas (1), (2), and (3) lead to the selection of optimal dimensions based on the optimization models (4), (5), (6), and Table 3.

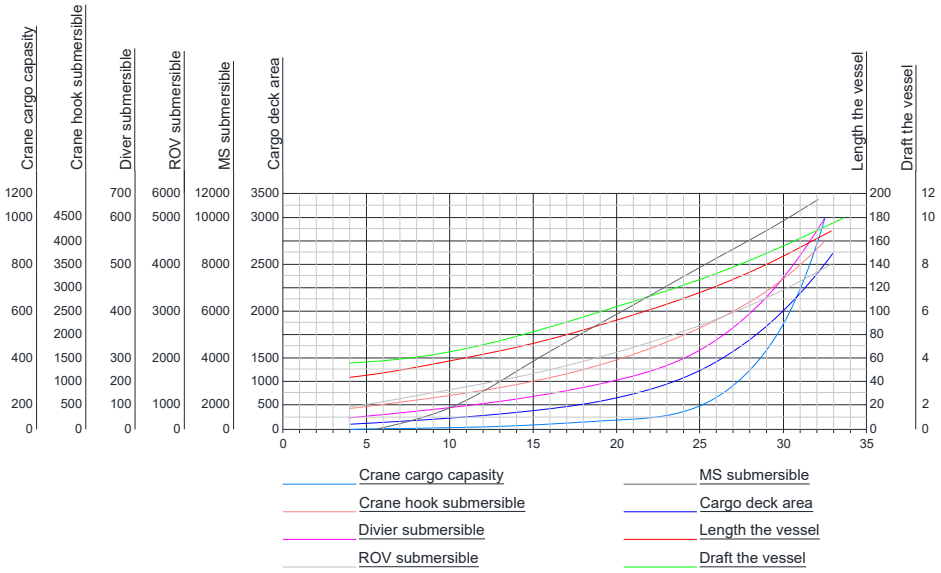


Figure 2: *Nomogram of definition of the main dimension of vessels of definition of subsea and technical operations*

Determination of the main dimensions according to the nomogram is carried out in the following sequence:

1. The vessel width is selected based on the required technological and design parameters reflected in the vertical scales on the left side of the nomogram:

– the required value of the technological or design parameter specified in the vessel design conditions (specified on the left side of the nomogram) is selected;

– a horizontal line is drawn until it intersects with the curve of the corresponding name;

– a vertical line is drawn downwards from the point of intersection of the horizontal line and the corresponding curve until it intersects with the horizontal scale;

– the point of intersection of the vertical line and the horizontal scale indicates the desired vessel width.

2. The largest value of the vessel width is selected:

– having carried out actions for all the technological and design parameters required for a specific vessel design, independently of each other, we obtain a group of width values for the designed vessel;

– from the obtained width values, the largest one is selected for the designed vessel.

3. The length of the vessel is selected:

– from the selected value of the vessel's width on the horizontal scale, draw a vertical line to the intersection with the vessel's length curve marked in red;

– from the point of intersection of the vertical line and the vessel length curve, a horizontal line is drawn to the right to the vertical scale on the right side of the nomogram reflecting the vessel length;

– the point of intersection of the horizontal line and the vertical scale of the vessel length reflects the length of the designed vessel.

4. The vessel draft is selected:

– from the selected value of the vessel width on the horizontal scale, a vertical line is drawn to the intersection with the vessel draft curve;

– from the point of intersection of the vertical line and the vessel draft curve, a horizontal line is drawn to the right to the vertical scale on the right side of the nomogram reflecting the vessel draft;

– the point of intersection of the horizontal line and the vertical scale of the vessel draft reflects the draft of the designed vessel.

Width of SSCV is determined by a formula (1) that is the sum of width of a diving complex, ROV width (at estimated joint installation) and width of passes for service.

$$B = b_{D.C} + b_{SERV.} + b_{ROV} \quad (1)$$

Here:
$b_{D.C}$ – width of a mobile diving complex, (Table 1);
$b_{SERV.}$ – pass width for service (~ 1-2m);
b_{ROV} – width of a mobile ROV complex (Table 2).

Length of the cargo deck for installation of diving complexes and apparatus, is determined by a formula (2) that is the sum of lengths of a diving complex, ROV and the necessary serving part of the deck.

$$L_{C.D.} = l_{D.C} + l_{ROV} + l_{SERV.} \quad (2)$$

Here:
$L_{C.D.}$ – length of a cargo deck;
$l_{D.C}$ – length of a mobile diving complex (Table 1);
l_{ROV} – length of a mobile ROV complex (Table 2);
$l_{SERV.}$ – length of a service part of the cargo deck (~Σ $l_{D.C}$, l_{ROV}).

Square of the cargo deck for installation of diving and ROV complexes, is determined by a formula (3) that is the sum of the spaces occupied by diving complexes, ROV and the serving part of the deck. At the same time the received sum of the areas has to be approximately equal to the work of length of the cargo deck and width of the designed vessel.

$$S_{C.D.} = (S_{D.C} + S_{ROV}) + S_{P.SERV.} \approx L_{C.D.} \times B \quad (3)$$

Here:

$S_{C.D.}$ – square of the cargo deck;

$S_{D.C}$ – square occupied by the mobile diving complex;

S_{ROV} – square occupied by the mobile ROV complex;

$S_{P.SERV.}$ – squre of the service part of the cargo deck (~Σ $S_{D.C}$, S_{ROV}).

Optimization of the main dimension of Subsea Construction Vessel got in parameters of diving complexes and ROV.

Counting dimension of SSCV by types and structure of the established diving complexes and apparatus, with use of models of optimization (4), (5) and (6), we form criteria of optimization:

$$L_{(X)} = Cf(\sum_{n=1}^{n} X_L) \rightarrow L_{mid} \quad (4)$$

$$B_{(X)} = Cf(\sum_{n=1}^{n} X_B) \rightarrow B_{max} \quad (5)$$

$$d_{(X)} = (C + B)f(\sum_{n=1}^{n} X_d) \rightarrow d_{min} \quad (6)$$

here:

C – constant values of a problem of optimization;

L – rated length of the vessel, m;

B – rated width of the vessel, m;

d – rated draft of the vessel, m;

X_L, X_B, X_d – the varied elements participating by optimization of overall dimensions of the vessel respectively;

L_{mid} – average dimension of length of the vessel;

B_{max} – maximum dimension of width of the vessel;

d_{min} – minimum dimension of the draft of the vessel.

Definition of optimum main dimension it is reflected in an optimization matrix, (see table 3). The specified main dimension the immersions of divers and the area of the cargo deck received on a formula (2) determined by formulas (1), (2), (3) and received according to the nomogram (fig. 2) taking into account depth, we define optimum main dimension of SSCV for safe and quality operation of diving complexes and ROV.

Table 3: *Matrix of optimization of the main dimension of SSCV*

Methods of the received dimension	**L**	**B**	**d**
On formulas (1), (2), (3)	110	23	4
On depth of immersion of divers (fig.2)	90	20	3
On square cargo deck. On formulas (3) and fig.2	85	18	3
The received main dimension	**90**	**23**	**3**

Formation of the main dimension of the Subsea Construction Vessel, received in parameters of diving and ROV complexes.

Relying on the received values of the main dimension of SSCV, length, width and draft, with use of dependence of the cubic module (which is the work of length, width and draft of the vessel) and displacement, received from values of the database of relatives to destination of vessels, we define required for operation of diving and ROV complexes, the displacement of the designed vessel.

Key parameters of the offered SSCV and parameters of the diving and ROV complexes installed on the vessel are reflected in table 4.

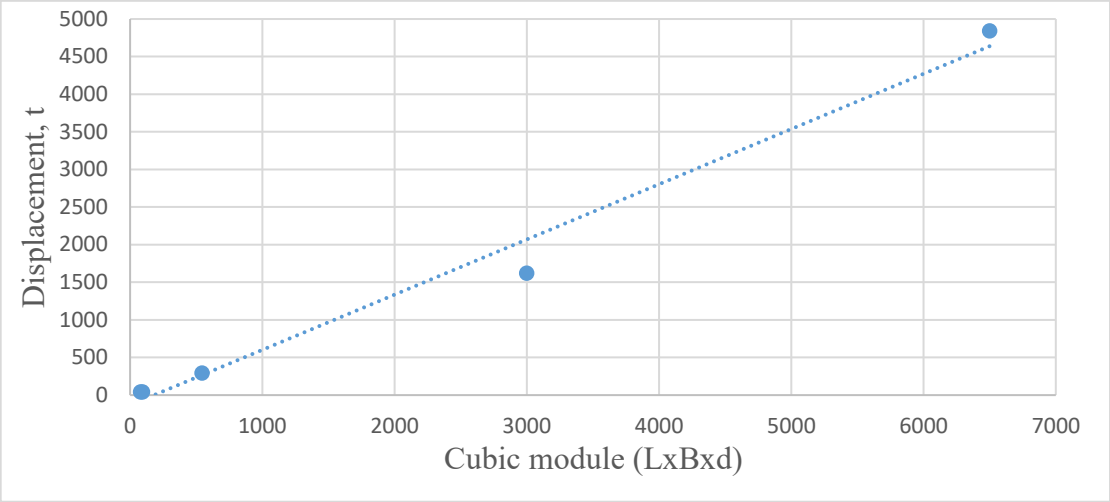


Figure 3: *Ratio cubic module (L×B×d) and displacement of diving vessels.*

Table 4: *Main parameters of the offered SSCV, in parameters of a diving and ROV complex*

Main dimensions of offered SSCV	
Length, m	90
Width, m	23
Draft, m	3
Square of cargo deck, m^3	440
Displacement, t	4500
Block coefficient, C_b	0.72
Bath for service staff	90
Main parameters of mobile diving complex MDDK-200	
Working depth, m	200
Type of respiratory mix	Helium-oxygen
Volume of the respiratory mix, m^3	8000
Power of the power station, kVt	800
Service staff, man.	60
Main parameters of mobile diving complex MDK-60	
Working depth, m	60
Type of respiratory mix	Compressed air
The sum volume compressed air, m^3	600
Power of the power station, kVt	400
Service staff, man.	15
Main parameters the mobile ROV complex	
The ROV class	Working
Working depth, m	2000
Type of the power	Hydraulic
Service staff, man.	10

Considering that the kind of work of SSCV, consists in collaboration of technology equipment installed on it, each of them, postpones the print when forming the main dimension of the offered vessel. At the subsequent stages of a research, it is supposed to consider influence of the installed mobile technology equipment on the main dimension of the Subse Construction Vessel and to

remove optimum values of the main dimension. It is separately possible to consider the possibility of perspective use of the offered vessel as the carrier of manned submersibles, for a research of a sea side of the Caspian Sea.

II. Conclusion

The types of mobile diving and ROV complexes available to companies operating offshore in the Caspian Sea are described. The method for selecting the main dimensions of SSCVs based on the parameters of the applied diving and ROV complexes is presented. A model for optimizing the main dimensions of SSCVs has been created. Using the optimization model, the main dimensions of SSCVs have been determined, and a type of vessel designed to ensure deep-water and technical operations for the safe and effective operation of the described mobile diving and ROV complexes is proposed. Expected subsequent research on the subject is also outlined.

Reference

[1] Abdullayev O.M., & Sadigov V.B. (2022). Prospects of improvement of scientific bases of design of Subsea Construction Vessels. XVII International scientific-technical conferece on "Water Transport Problems". Baku, 10 – 12.

[2] Abdullayev O.M. (2023). The analysis of functions of the support vessel underwater tekhyoniyocheskikh works taking into account processing equipment and vliyoyaniye assessment on the sizes and technical characteristics of the vessel. Proceedings of Azerbaijan State Marine Academy. Baku, 1(37), 11 – 21.

[3] Egorov A.G. & Абдуллаев O.M. (2023). Risk-based analysis of operational design restrictions and main design characteristics of subsea construction VESSELS. Вестник Одесского Национального Университета. Одесса. 1(68), 7 – 26.
https://doi.org/10.47049/2226-1893-2023-1-7-26

[4] Abdullayev O.M. (2023). Method for the formation of algorithm elements for designing a vessel for the subsea technical operations support using a database of similar vessels. Vestnik Gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S.O. Makarova. Saint-Petersburg. 15 (4), 680 – 691.
https://doi.org/10.21821/2309-5180-2023-15-4-680-691

[5] Lebedeva M.P. & Abdullayev O.M. (2023). Algorithm for designing vessel for underwater technical operations using database method. Russian Journal of Water Transport. Volga State University of Water Transport. Nijniy-Novgorod. 76(3), 32 – 49.
https://doi.org/10.37890/jwt.vi76.386

[6] Bashirov R.D. & Abdullayev O.M. (2024). Classification of subsea construction vessels and evaluation of study effectiveness. 2nd international scientific-practical conference "Machine Building and Energy: New Concepts and Technologies". Baku. (148). 217 – 223.

Abdullayev O.M. (2024). Development of the nomogram of the choice of the main dimension of Subsea Construction Vessels. XIX International scientific-technical conferece on "WATER TRANSPORT PROBLEMS". Баку. 15 – 17.

[7] Rahul Bharti, Bhaskaran Pranesh, Dharmaraj Sathianarayanan, Manickavasagam Palaniappan & Gidugu Ananda Ramadass. (2024). Added mass analysis of submersible using computational fluid dynamics. Maritime Technology and Research. 6(3). 2 - 23. https://doi.org/10.33175/mtr.2024.267954

DEVELOPMENT OF HIGH-STRENGTH DEEP-WELL PUMP RODS FOR OIL PRODUCTION BASED ON INNOVATIVE METALLURGICAL TECHNOLOGIES

Rahim Shukyurov[1], Naila Mirbabayeva[1], Lala Azimova[1]

•

[1]Azerbaijan Technical University, Azerbaijan, Baku
rahim33@mail.ru, nailamirababayeva@outlook.com, lala.azımova.77@mail.ru

Abstract

In the world, including the Republic of Azerbaijan, oil production from 75-85% of wells is carried out by deep-well pumps. A deep-well pump located at the bottom of the well is connected to a pumping unit at the wellhead with rods. The pump brings oil to the surface of the wells either due to the reciprocating motion of the rod string or due to its rotation in one direction. In the first case, the rod string operates under a repetitive, variable tensile load with an asymmetric cycle, and in the second - under a torsional load. The chemical composition and design of the rods used in both cases are the same. Pump rods (Fig. 1) are long (8 m long) products of small diameter (16-25 mm) and are considered the weakest link in oil production by pumping. Shutdown of oil wells in most cases occurs as a result of rod breakage. Therefore, the level of oil production by this method depends on the strength and reliability of the pump rod. The aim of the research work is to obtain high-strength rods for deep-well pumps by using innovative metallurgical technologies. This is a pressing scientific and technical problem.

Keywords: sucker rod, steel, rolling, structure, sorbite, ingot, thermomechanical treatment, hardenability.

I. Introduction

The main factor determining the reliability and durability of deep well pump rods is the structural state of the rod metal after all operations. Unfortunately, many researchers have not paid due attention to this serious factor. For this reason, there is no significant information in the literature on the influence of steel structure on the operational properties of rods. Research shows that the most optimal working structure of the metal of pump rods operating in very difficult conditions should be sorbite (Fig. 6) [1].

In most cases, pump rods used in oil fields are rolled hot and cooled in air, i.e., normalized. In this case, excess ferrite grains appear in the metal structure. Even after heating the rod with high-frequency current, excess ferrite remains in the internal structure of the steel. To remove excess ferrite grains from the steel structure, prepare its structure for operating conditions and increase the steel's ability to elastic deformation, the pump rod must be hardened and tempered. For this purpose, the plant must have a workshop equipped with special equipment. In addition, the chemical composition of the rod steel should be selected so that the depth of hardening penetration (hardenability) fully ensures the cross-section of the product. However, many years of research show

that this method is not suitable, since it causes problems when hardening long parts of small diameter (bending during hardening and cracking in the rods during the subsequent straightening process, etc.) [1].

II. Main part

It is known from the studies conducted on broken rods that the main reason for the failure of sucker rods after some time is the presence in their structure of excess ferrite grains with low mechanical properties (Fig. 2). Ferrite grains are subject to plastic deformation under stresses even below the yield strength ($\sigma_{0.2}$) of rod steel, soften and break under the action of forces periodically changing in direction (Fig.3). Since the yield strength of steel corresponds to the plastic and elastic region, irreversible plastic deformation develops significantly in the structure during operation. For this reason, the rod quickly breaks and fails. Therefore, the yield strength cannot fully characterize the quality of rod steel. Thus, steels used in the production of sucker rods should be characterized by the elastic limit at low residual deformation during measurement, that is, the limit of resistance at low deformations ($\sigma_{0.005}$).

Figure 1: *Head and transition zone of pump rods ShG-19 and ShG-22*

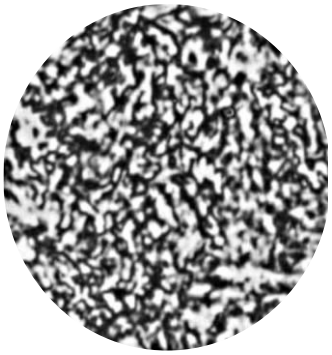


Figure 2: *Structure of rods made of 20NM steel after normalization ferrite + pearlite. x 400*

Figure 3: *Destroyed ferrite. x 2000*

In the Azerbaijan Republic and foreign literature, it is indicated that 77% of rods break in the body, including the transition zone at a distance of 120-160 mm from the rod head. In wells with a depth of 2000-2800 meters, this figure is 92%. Rod breakage: 42% occurs at the top, 35% in the middle, 23% at the bottom. According to literature, 62% of rods in the USA break in the upper part of the rod. The reason for this is that there is strong stress in the upper part of the rod. The reason for the breakage of rods in the lower part of the column is their bending and vibration along the length [1].

The steels tested and used in the Azerbaijan Republic and abroad for the manufacture of deep-well pump rods can be conditionally divided into three groups.

The first group includes low- and medium-carbon alloy steels 15NM, 20Cr, 20NM, 20CrGR, 35G, 35G2, 35GN, 35GCrM, 20Cr2M, 30CrMA, 40. Rods made of these steels are used mainly after normalization, the mechanical properties are at a fairly high level: $\sigma_в$=550-750 MPa; $\sigma_{0.2}$=300-480 MPa; δ=40-70%; φ=14-30%; $a_н$=0.6-1.6 MJ/m2. However, due to the presence of a large amount of structurally free ferrite in the structure, their elastic limit is very low ($\sigma_{0.005}$=150-220 MPa). Therefore, the use of some of them can be justified only if the deep well pump rods were hardened with sufficient stability of supercooled austenite and hardenability of steel.

The second group includes high-alloy steels 10Cr2GN4, 25CrGN3MD, 40CrGN3M, 35CrGN3M. Rods made of steels of this group are also delivered to the fields after normalization and tempering. They have high elastic mechanical properties: $\sigma_в$=600-900 MPa; $\sigma_{0.2}$=600-750 MPa; $\sigma_{0.005}$=350-450 MPa; δ=12-22%.

The third group includes low-carbon high-alloy nickel steels 03N4, 04N3M, 08N4M ($\sigma_в$=400-

500 MPa; σ_{02}=330-370 MPa). It should be noted that steels of the second and third groups are difficult to machine due to their high viscosity and contain (3-4%) scarce nickel. It is not advisable to use these steels for the production of rods.

In connection with the above, it was proposed [1] to use self-hardening steels for the manufacture of deep-well pump rods in such a way that duppring the cooling process after hot rolling or normalization, the rods are hardened in air along their entire length and cross-section. Subsequent tempering at optimal temperatures should ensure the receipt of a sorbitol structure and high values of the elastic limit. In this connection, 15X2NMFAl steel was developed, which after rolling or normalization with the corresponding rod diameter (16-22 mm) self-hardens in air along their entire cross-section to a bainitic structure. The chemical composition of this steel is given in table 1.

Table 1: *Chemical composition of steel*

Brand steel	Chemical composition of steel, %									
	C	Si	Mn	Cr	Mo	V	Ni	Al	S	P
15Cr2NMFAl	0,13-0,17	0,17-0,37	0,5-0,7	1,85 2,20	0,2-0,3	0,08-0,13	0,9-1,2	0,01-0,05	0,015	0,015

The steel was smelted in a 60-ton electric arc furnace. Deoxidation and alloying were performed during tapping into a steel-pouring ladle using liquid ligature and synthetic slag [2].

Aluminum for final deoxidation was introduced through a chute in small pieces weighing 0.5-1.0 kg in an amount of 0.6-1.0 kg/ton. The temperature in the ladle after tapping the metal was 1600-1620 °C. The smelted steel was poured in a continuous metal casting unit into billets with a cross-section of 280 x 280 mm. Ingots heated to 1150-1200 °C were rolled into intermediate billets with a diameter of Ø110 mm. The temperature of the metal at the beginning of rolling was 1120-1170 °C and at the end of rolling it was 850-950 °C. Intermediate blanks were rolled on a small-section mill 350 to a profile of Ø 19- and 22-mm. Steel grade 15Cr2NMFAl has high hardenability (Fig. 5), which makes it possible to use the heat of rolling heating during HTMT. This allows the strengthening of rods to be carried out in two ways:

1. after leaving the last stand with a deformation of 15%, the rods are cooled (table 2) in the air in the shop to obtain hardened rolled products with a bainite structure. Then the profile rolled products were cut to standard sizes and subjected to heating along the entire length in a furnace to a temperature of 630-650 ° C and cooled in air - tempering to obtain a sorbite structure. It was found that a rod with a hot rolling end temperature of 850-950 ° C before the onset of bainitic transformation (480 ° C) under shop conditions is cooled along the entire length in 90-100 sec., i.e., the cooling rate of the rods is 3-5 degrees / sec. These cooling rates of the rods under shop conditions correspond to the stability of supercooled austenite according to the diagram of isothermal and thermokinetic transformation in the range of 700-500 ° C (Fig. 4). The mechanical properties of deep-well pump rods made of 15Cr2NMFAl steel depending on the cooling rate from 860-920 °C are given in table 2.

Table 2: *Cooling rate after rolling and normalizing*

Cooling rate after rolling and normalizing, deg/sec.	G 0,2, MPa	HRC/HB
0,1	540/260	25/248
1	720/760	29/248
3	740/680	32/248

6	820/780	34/248
500	110/740	41/248
Note: the numerator shows the properties of steel in the untampered state, the denominator shows the properties after tempering at 660-680 °C		

2. The presence of more than 3% alloying elements in the composition does not allow rapid cooling with water after deformation to the workshop temperature (i.e., quenching) during HTMT. For this reason, we manufactured pump rods from 15Cr2NMFAl steel using the high-temperature isothermal thermomechanical treatment method.

The 15Cr2NMFAl steel blanks were subjected to deformation by 15% upon exiting the last stand of the rolling mill and cooled according to the thermokinetic diagram (Fig. 4). After the rolled products with a diameter of 22 mm exited the last stand at a speed of 9 meters per second, they were cooled on 10 devices installed along the rolling with air under a pressure of 6-7 atm. to a temperature of 650-630 °C. In this case, the metal surface was cooled at a speed of 250 °C/sec. After the cooling by inertia ceases, the surface temperature of the rolled product decreases to 550-570°C. Then, due to the internal heat, the temperature of the rolled product rises to 630-650°C, i.e., it was subjected to isothermal thermomechanical treatment [3]. After that, the rolled product is cooled in the air in the shop conditions, i.e., the tempering process takes place. The deformed and supercooled austenite is transformed into a thin ferrite-cementite mixture – sorbite (Fig. 6). The structure of 20NM steel consists of pearlite, which is associated with low hardenability. The resulting profile was cut in a hot state and cooled in the air, obtaining rolled product hardened for sorbite.

The hardenability of steels was studied by the standard method [5,6] and is shown in fig. 5. As can be seen from the figure, the hardenability of steel grades 20NM, 15Cr2NMFAl, respectively, is 6, 34 mm.

The following mode of upsetting rod heads on a horizontal forging machine was developed: heating to 1150-1200 ° C upsetting, finishing upsetting at 1020-1060 ° C. The fibers on the rod head are located in the direction of deformation. No overheating was detected in the metal of the rod head (the fracture is matte).

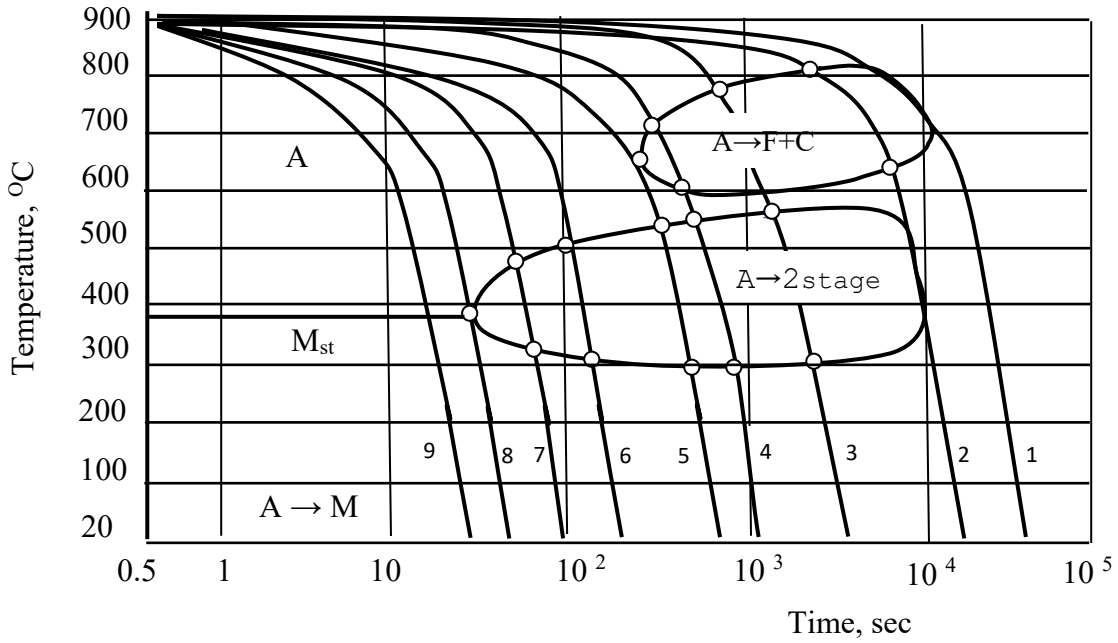


Figure 4: *Thermokinetic diagram of 15X2NMFAl steel. 1-0.015/HB 201; 2-0.01/HB229; 3-0.15/HRC 27; 4-05/31; 5-1/33; 6-9.5/36; 7-6.5/38; 8-13/41; 9-25/41. In the numerator - cooling rate in deg/sec. In the denominator - hardness - HB, HRC*

The results of the study of mechanical properties show that the transition zone in the steel under study after upsetting and tempering at 650 °C is practically absent. After upsetting the heads, one part of the rods was tempered, the other – normalized and tempered. In both cases, i.e., after hot rolling or normalization, the resulting structure in the deep-well pump rods was uniform and consisted of l ower baite. The hardness along the entire length of the rods was HB 235-240, the difference in hardness of the center of the section and the surface was HRC 0.5-0.8, which did not exceed the measurement error

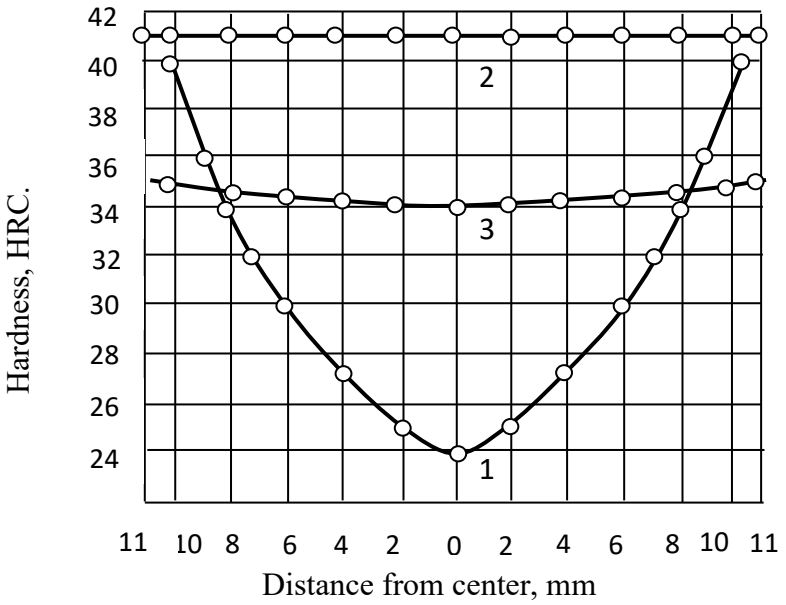


Figure 5: *Hardenability by cross-section of ShG-22 pump rods made of steels: 1-20NM and 2-15Cr2NMFAl after quenching in water, 3- after hot rolling, air cooling 15Cr2NMFAl*

.

Average data on mechanical properties of the transition zone and the rod body are presented in table 3. It is evident from the table that mechanical properties of deep-well pump rods made of 15Cr2NMFAl steel are higher than those of rods made of 20NM steel. And mechanical properties in the transition zone of rods made of 20NM steel are significantly lower than those in the rod body. The elastic limit of this steel (σ0.005) is within 53-57% of the yield strength (σ0.2), which is due to low hardenability of steel, the presence of ferrite in the structure and the coarse-grained transition zone.

Table 3: *Mechanical properties of the rods on the body and in the transition zone*

Brand steel	$\sigma_в$, MPa	$\sigma_{0,2}$, MPa	δ, %	φ, %	KV, J/cm2	$\sigma_{0,005}$, MPa
20NM*	600/530	390/320	21/17	56/42	120/80	210/178
15Cr2NMFAl *	810/760	730/685	21/18	76/70	165/152	643/618
* In the numerator the data in the body, in the denominator the data in the transition zone						

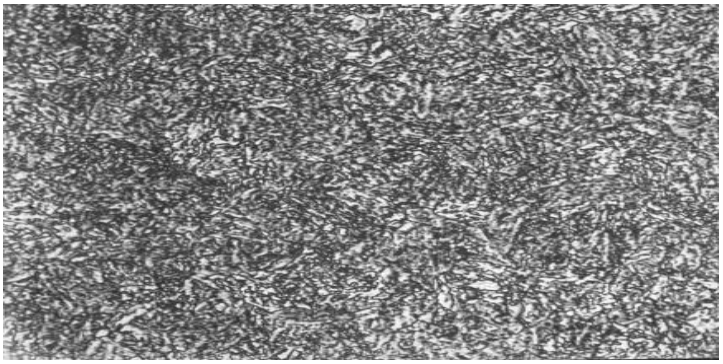


Figure 6: *Structure of steel rods 15Cr2NMFAl after VTMO and self- vacation. Sorbitol. x 400*

The presence of a small amount of molybdenum and vanadium in the composition of steel grade 15Cr2NMFAl contributes to increased hardenability and the formation of a fine-grained structure of the transition zone. Therefore, the mechanical properties of the transition zone and rod bodies made of steel grades 15Cr2NMFAl are high. The elastic limit of this steel ($\sigma_{0.005}$) is within 87-89% of the yield strength ($\sigma_{0.2}$). The mechanical properties of the transition zone of rods made of this steel are much higher than the indicators in the body of rods made of 20NM steels, provided for in GOSTs of the CIS and Russia [4].

Tests of pilot rods of all pilot heats were carried out in oil wells. Depending on the depth of the wells, 86-250 rods were used to assemble each column. The chemical composition of the pilot steel heats and test data are given in tables 4 and 5.

Table 4: *Chemical composition of experimental melts*

Number of compositions		Chemical composition of experimental melts, %									
		C	Cr	Mo	V	Ni	Al	Si	Mn	S	P
1		0,13	1,85	0,25	0,13	1,2	0,02	0,25	0,50	0,016	0,016
2		0,15	2,20	0,20	0,10	0,9	0,03	0,30	0,60	0,015	0,015
3		0,17	2,15	0,30	0,08	1,1	0,05	0,32	0,58	0,015	0,016

Table 5: Test data are given

Number of compositions	Tested bars, pcs.	Number of rod breaks		Durability of the bars	
		pcs	%	in hours	in work cycles x10-6
Rods ShG -19					
1	1086	3	0,3	16025	10,5
2	1672	3	02	18213	10,9
3	1893	6	0,2	19074	13,1
Rods ShG -22					
1	1242	4	0,3		10,3
2	1514	2	0,1	19008	11,6
3	1897	4	02	20716	12,6

III. Conclusion

According to the results of comparative tests, the probability of failure-free operation of rods made of 20NM and 15Cr2NMFAl steel is 0.974 and 0.995, respectively.

Tests of deep well pump rods made of 15Cr2NMFAl steel have shown that their use in

oil production allows for an increase in the durability of the rod string by approximately 1.4-1.7 times and a decrease in the number of rod breaks by 6-8 times, thereby reducing well downtime for emergency rod replacements, reducing repair work and increasing oil production.

References

[1]. Shukyurov R.I. Metallurgy and thermal processing, Baku: Sabah, 2020. 507.

[2]. Shukyurov R.I., Rakhimov M.M. Metallurgy, Baku: Maarif, 1985. 344

[3]. Bernstein M.L. Thermomechanical processing of steel, M. Metallurgy. Volumes 1-2, 1968. 1171.

[4]. GOST (State Standart) 31825, Moscow, Standardinf, 2013.

[5]. Gulyaev A.P. Metallurgy, Moscow: Metallurgy, 1986.

[6]. Mirbabayeva N.R., Hasanova S.S., Shukyurov R.I. Investigation of the penetration depth in thermomechanical processing // Mechanical Engineering. 2003; 1: 35-37.

THE GENERAL REGULARITY OF SURFACE LAYER WORK HARDENING IN THE HONING OPERATION OF MEDIUM CARBON STEELS

Sarvan Aziz Shirvan
•
Azerbaijan Technical University, H. Javid avenue 25, Baku, Azerbaijan AZ 1073
sarvan_e@hotmail.com

Abstract

This paper discusses the general issues of controlling the formation of work hardening in the internal cylindrical surfaces during the honing process using technological methods. Based on the general characteristics of the honing operation of medium carbon steel materials, and considering surface quality and productivity, the technological regularity of work hardening has been identified. Experimental research results on the influence of the main processing parameters-such as the forward-backward (Vb-f) and rotational (Vr) speeds of the honing head, specific pressure (Ps.p), abrasive grit size (Z), and processing time (T) on the degree of work hardening are presented. The functional dependencies of the main input parameters of the process on work hardening, which is taken as the main output parameter for surface quality, have been graphically determined. As a result, the technological foundations for selecting optimal processing regimes have been established.

Keywords: abrasive grain, work hardening, honing process, hardness, special pressure, surface roughness, accuracy.

I. Introduction

The application of the honing operation in the production of machine parts, as in other machining methods, is accompanied by plastic deformation -work hardening of the material's surface layer. The degree of work hardening (H) and the thickness of the hardening layer (a) are primarily proportional to the external force and processing time. The heat generated in the cutting zone increases the value of plastic deformation on the working surface of the material, which results in a change in the frictional force and also leads to structural transformations of the metal. During the honing process of the internal cylindrical surfaces of parts made of medium carbon steel materials, which operate under high pressure-temperature conditions, it becomes necessary to study the dependence of the degree and depth of work hardening on cutting regime parameters. Experimental research results show that, during surface plastic deformation of steels, in addition to the crushing of blocks, the hardening effect plays a crucial role in the dispersion hardening, changes in dislocation density, and phase transformations [1, 2]. An increase in dislocation density and an increase in the dispersion of carbides, which restrict movement on the slip plane, are also observed to lead to an increase in hardness. A similar situation is observed when the amount of martensite deformation and the transformed residual austenite increases.

The formation of work hardening in the surface layer during honing is a sufficiently complex technological issue, as a number of input parameters of the process, as well as the influence of

external factors, can cause significant changes, all of which have a substantial effect on the degree of surface work hardening.

The effect of various factors on the honing process was investigated using the method of planning multifactorial experiments. As the research object, medium carbon steels used in the production of high-precision parts in the special machine engineering and instrument-making industries were selected. Based on the calculations, a mathematical model of the honing process was developed, characterizing the dependence of the surface layer work hardening (H) on the factors such as specific pressure ($P_{s.p}$), abrasive grit size (Z), processing time (T), forward-backward ($V_{b\text{-}f}$) and rotational (V_r) speeds of the honing head, for both the coded and natural values of these factors. Relevant checks were performed for the necessary criteria for second-order orthogonal planning. The obtained models are fully applicable for calculating the values of the work hardening degree (H) depending on the factors $P_{s\text{-}p}$, Z, T, $V_{b\text{-}f}$, and V_r at the $\alpha = 0.05$ significance level.

Let's write the equation characterizing the effect of the coded values of the factors $P_{s.p}$, Z, T, $V_{b\text{-}f}$, and V_r:

$$\begin{gathered}\hat{y}_H = 7343{,}23 - 164{,}56x_1 - 321{,}23x_2 + 191x_3 - 167{,}23x_4 + \\ +134{,}14x_5 - 74{,}26x_1x_2 - 61{,}95x_1x_3 - 62{,}38x_1x_4 + 81{,}12x_1x_5 - \\ -58{,}87x_2x_3 - 73{,}67x_2x_4 + 85{,}45x_2x_5 - 73{,}14x_3x_4 + 83{,}21x_3x_5 + \\ +81{,}37x_4x_5 + 32{,}79x_1^2 - 187{,}23x_2^2 + 53{,}51x_3^2 + 54{,}52x_4^2 + 51{,}57x_5^2\end{gathered} \quad (1)$$

For the natural values of the factors, after the appropriate transformations, equation (1) takes the following form:

$$\begin{gathered}H = 8121{,}22 - 951{,}73V_{b-f} - 71V_r - 4{,}14P_{s.p} - 0{,}782Z - \\ -12{,}42T - 8517{,}67V_{b-f}{:}\,V_r - 28{,}51V_{b-f}{:}\,P_{s.p} - 134{,}41V_{b-f}\cdot Z + \\ +23{,}24V_{b-f}T - 10{,}91V_r\cdot P_{s.p} - 52{,}32V_r\cdot Z + 58{,}45V_r\cdot T - \\ -0{,}31P_{s.p}\cdot Z + 0{,}4P_{s.p}\cdot T + 0{,}32ZT + 998{,}63V_{b-f}^2 - \\ 8723{,}45V_r^2 + 0{,}11P_{s.p}^2 + 0{,}62Z^2 + 0{,}03T^2\end{gathered} \quad (2)$$

By applying equation (2) to the optimization, the following values of the friction process parameters, which ensure the maximum surface layer work hardening, are obtained: - $V_{b\text{–}f}$=0.22 m/sec, V_r=0.41 m/sec, $P_{s.p}$=71 kPa, Z=160/125 μm, T=95 sec. Using the obtained equations, the surface layer work hardening degree (H) has been calculated based on the input parameters of the honing process ($P_{s.p}$, Z, T, $V_{b\text{–}f}$ and V_r), and the corresponding functional dependencies have been established in the following graph.

Technological factors that cause an increase in temperature in the honing zone lead to a rise in temperature and cause a decrease in the hardness of the surface layer material. Examples of such factors include the forward-backward ($V_{b\text{–}f}$) and rotational (V_r) speeds of the honing head. With an increase in the values of these parameters, the temperature in the cutting zone rises proportionally, and the degree of work hardening in the surface layer starts to decrease (Fig. 1). The results of the studies [3 and 4] show that if the plastic deformation process occurs at relatively high temperatures (where atomic displacements increase and the elimination of distortions in the crystalline lattice becomes easier), then both relaxation and hardness reduction occur more rapidly.

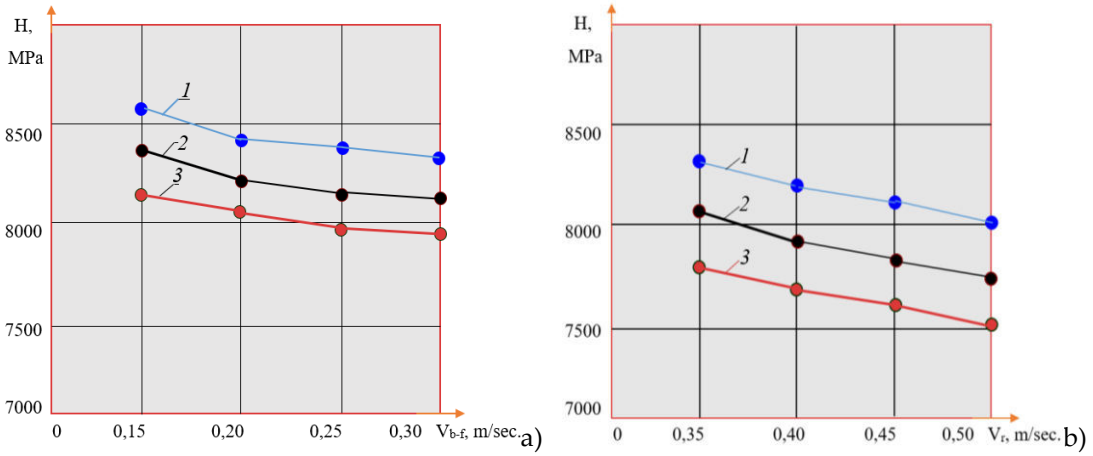


Figure 1: *Dependence of work hardening (H) on forward-backward $V_{b\text{-}f}$ (a) and rotational $V_{r\,(b)}$ speeds: 1, 2, 3 - corresponding to abrasive grit sizes: Z=125/100Z,160/125, and 230/200 μm*

As seen in Fig. 2a, an increase in specific pressure ($P_{s.p}$) reduces the surface layer work hardening degree. As is known from [5, 6], the work hardening of the metal caused by plastic deformation leads to a state of tension and structural instability. Consequently, upon the completion of the plastic deformation process, relaxation occurs directly in the metal, which results in the metal attempting to return to its initial, stable state. Even under normal processing conditions, i.e., with low and medium values of specific pressure, atomic displacements can alter the structure of the atomic crystal lattice and reduce the work hardening of the surface layer.

With an increase in the abrasive grit size (Z), the work hardening value decreases (Fig. 2b). This regularity can be explained by the fact that, under a certain balance of abrasive action, the absolute value of plastic deformation increases with the increase in grit size. Therefore, alongside the removal of material, a decrease in the surface layer work hardening degree also occurs.

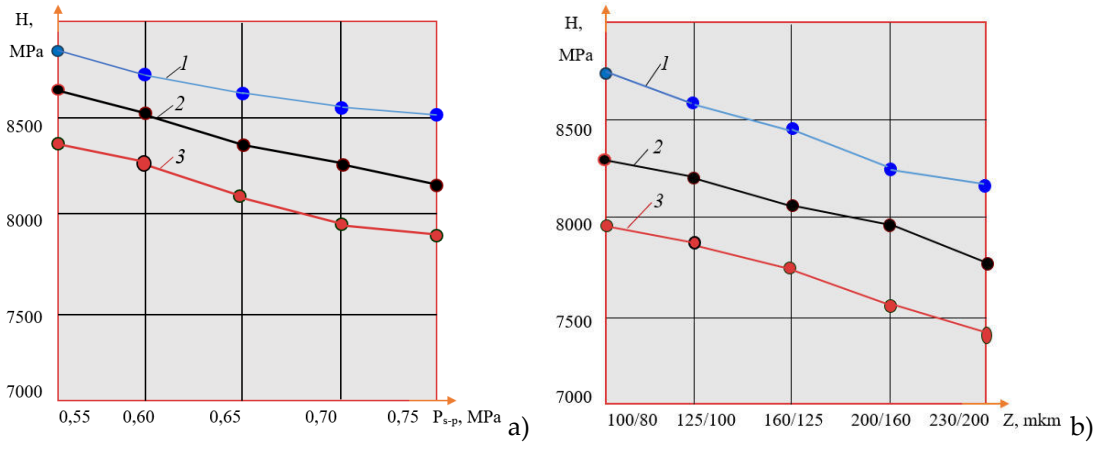


Figure 2: *Dependence of work hardening (H) on specific pressure ($P_{s.p}$) and abrasive granularity size (Z): 1, 2, 3 - corresponding to V_r=0.40, 0.45 and 0,50 m/s.*

The graphical dependencies obtained through experimental research show that with an increase in processing time (T), the surface layer work hardening degree (H) increases proportionally (Fig. 3). This can be explained by the increase in the duration of abrasive action on the processed surface [7, 8]. However, during the course of the processing, the number of cutting grains decreases, and due to the clustering of the main cutting edges of the abrasive grains, more favorable conditions

for surface layer work hardening are created, which also leads to an intensive increase in the work hardening depth (*a*).

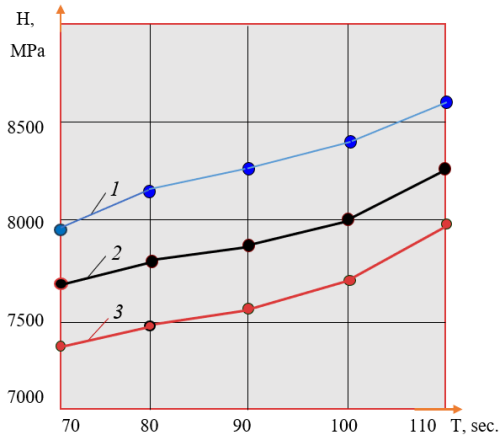


Figure 3: *Dependence of work hardening (H) on processing time (T): 1, 2, 3 – corresponding to abrasive granularity sizes: Z=125/100, 160/125 and 230/200 μm*

It is well known that the machining process of metal materials by any cutting method is accompanied by plastic deformation of the surface layer. Although in abrasive machining methods, especially in the honing process, this is observed at relatively small values compared to traditional cutting methods, it is not negligible. During the honing of medium-carbon steel materials, the work hardening degree (D) of the surface layer can increase by 35-40% within a depth range of 4-24 μm. It is particularly important to note that, as with other machining methods, in the internal cylindrical honing process, the values of work hardening degree and work hardening depth can be controlled (regulated) by changing the main input parameters of the process.

The graphical dependence of the work hardening degree (H) on the depth (a) of the hardening surface layer during the honing process is shown in Fig. 4. Here, the results obtained from experiments for three different values of the rotational speed of the honing head (V_{r1}=35, $Vr2$=40, V_{r3}=45 m/sec) under the same processing conditions are presented.

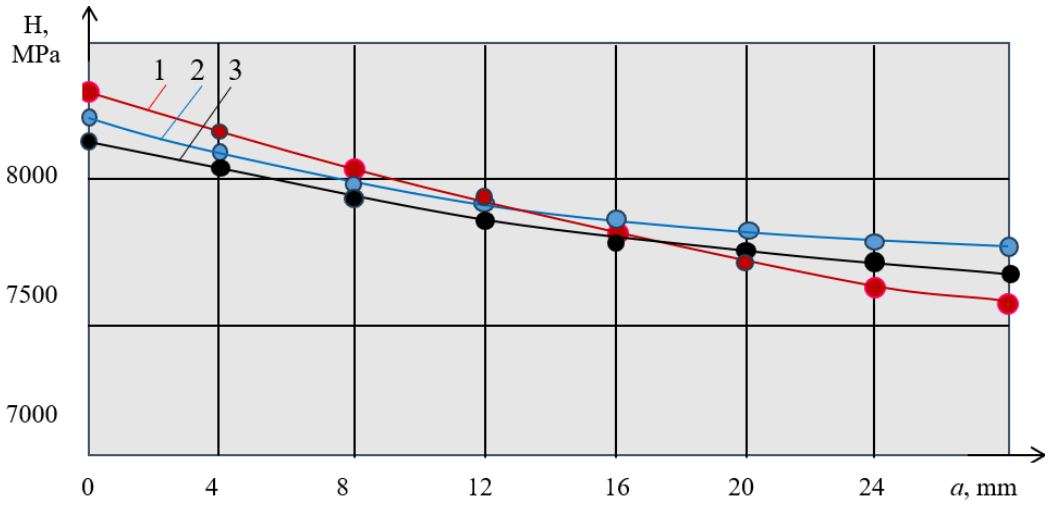


Figure 4: *Distribution of work hardening (H) on the depth (a) of the surface layer during the honing process (1-V_r=35, 2-V_r=40 and 3-V_r=45 m/sec)*

When considering the general regularity of the influence of each parameter of the honing process on the surface layer work hardening, it becomes clear that any technological process involving cutting is accompanied by plastic deformation- work hardening of the surface layer, which

significantly changes in depth direction from the metal surface. Therefore, depending on the values of processing parameters ($V_{b.f}$, V_r, $P_{s.p}$, Z, T) in the friction process, it is possible to achieve a wide range of work hardening values in the surface layer. Based on the purpose of the machine parts and the technological requirements set for the main working surface, the optimal values of work hardening degree and depth are determined.

In the honing operation, an increase in grit size is accompanied by a proportional decrease in the work hardening degree. This regularity can be related to the ability of abrasive particles to cause plastic deformation on the surface. The combined rotary and forward-backward motion of the honing head leaves fine tracks of the abrasive particles on the machined surface. The photograph and topography of the honed surface are shown in Fig. 5a and b, respectively. The plateaus of the honed surface constitute a large apparent contact area, which results in higher load-carrying capacity. The cross-hatch patterns retain the lubricant, which helps reduce the friction coefficient and running-in wear [9].

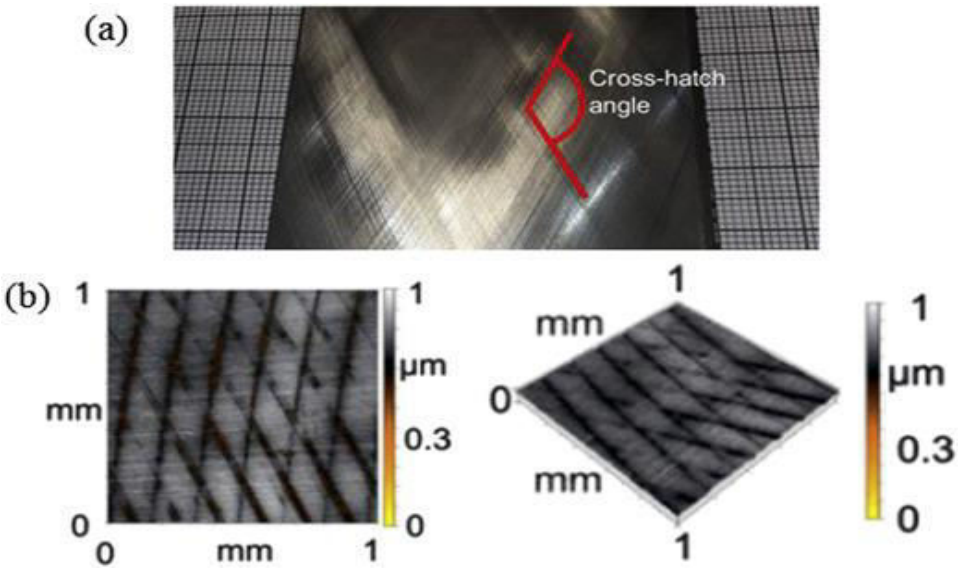


Figure 5: *Photograph (a) and 3D topography (b) of honed surface*

With an increase in the forward-backward and rotational motion speeds, the work hardening of the surface layer decreases. This can be related to the rise in temperature in the contact zone, which results in the softening of the material. In this regard, under specific processing conditions, the work hardening of the surface layer can only decrease to a certain minimum value. Based on the results of experimental studies, it can be stated that during the honing operation, a temperature field is generated in the surface layer that exceeds the initial value, which leads to a reduction in the plasticity and work hardening.

Increasing the specific pressure ($P_{s.p}$) from 0,55 MPa to 0,73 MPa results in a decrease in the surface layer work hardening (H) from 8500 MPa to 7700 MPa. Increasing the value of Ps.p enhances the material removal from the surface layer, and the honing process transitions from diamond polishing to micro-cutting. A decrease in the value of $P_{s.p}$ reduces the value of plastic deformation, and the abrasive grains only participate in cutting the tips of the protrusions on the surface. After that, the intensity of the cutting process sharply decreases, the apparent surface area increases, and the value of specific pressure becomes insufficient to generate plastic deformation in the surface layer.

When the granularity size is increased from Z_{min}=100/80Z to Z_{max}=230/220Z, the initial state of the surface continuously changes from a highly hardened layer to a less hardening one, and this difference can sometimes reach 35-40%. This trend can be explained by the fact that, at the beginning of the process, the cutting grains come into contact with the tips of individual protrusions on the

surface's micro-relief. As the grit size increases, the abraded surface smoothens, the number of engaged grains increases, and the intensity of material removal from the surface layer rises, which leads to a decrease in the work hardening degree of the surface layer.

The abrasive grains of the honing wheel, in their cutting area, are irregular in shape and size, with an arbitrary arrangement and varying cutting angles, thus creating micro-cutting. In this regard, each abrasive grain penetrates the surface to a different depth-larger grains are more loaded, while smaller grains are less loaded [10]. Some abrasive grains are unable to perform cutting work; they only cause plastic deformation of the surface, while very small grains generally do not make contact at all. Microscopic observation of the chips obtained from honing shows that their size and shape are also varied.

The formation of different processing conditions for each grain individually naturally leads to the creation of varying cutting temperatures in the contact zone. Therefore, it is completely logical that different levels of work hardening appear on the individual micro-areas of the surface. When designing the technological process, processing conditions should be created in such a way that the process can transition from micro-cutting to the polishing process without the need for additional conditions [11, 12]. At this point, the optimal value of processing time plays a significant role. Specifically, the maximum (T_{max} = 110 sec) and minimum (T_{min} = 70 sec) values of processing time do not provide the required surface quality, and are also considered economically inefficient. However, when the processing time is in the range of T =90-95 sec, the surface layer's work hardening degree increases intensively, while at longer times, it increases gradually. This progressive increase in work hardening with increasing processing time is considered normal. This is because, as processing time increases, the honing head's contact time with the surface also increases, resulting in additional deformation and strengthening of the surface layer.

II. Conclusion

In order to achieve high surface quality during the honing operation of internal cylindrical surfaces of high-precision parts, and especially to control the surface layer’s work hardening in a systematic way, the following key technological issues must be addressed:

- the optimal processing conditions should be selected, taking into account the main characteristics of the honing process, which is the final technological operation, in order to enhance the product quality and durability;

- the mechanism of the effect of the main honing parameters on the work hardening degree should be investigated through both theoretical and experimental research, considering the production efficiency and quality;

- the characteristics of the work hardening formed on the surface layer, as well as the influence of these characteristics on the friction and wear processes during operation, should be considered as a key factor, with particular regard to the physical-mechanical properties of medium-carbon steel, especially its hardness.

References

[1] Eziz, S. Sh. Kinematic features of the lapping process and determination of its basic parameters // − Moskow: Computational Nanotechnology, − 2020. №3 (Vol.7), − p. 11-16 https://urvak.org/articles/compu-5261-vypusk-3-kinematicheskie-osobennosti-pr/

[2] Eziz, S. Sh. Theoretical studies of the dynamic characteristics of the internal lapping process // − Magnitogorsk: Vestnik of Nosov Magnitogorsk State Technical University, − 2020. №2 (Vol.18), − p. 30-37 https://vestnik.magtu.ru/en/archive/86-archive/no-2-2020/1095-30.html

[3] Aziz S. Sh. Characterıstıcs of qualıty formatıon ın processıng of 32CrMoV12-10 steel by honıng, 2nd International Scientific-Practical Conference "Machine Building and Energy: New

Concepts and Technologies", December 4-5, 2023, AzTU -Baku, Azerbaijan
https://www.scientific.net/KEM.979.39

[4] Yusubov N.D., Abbasova H.M. Full-factor matrix model of accuracy of dimensions performed on multi-purpose CNC machines. Obrabotka metallov (tekhnologiya, oborudovanie, instrumenty) = Metal Working and Material Science, 2021, vol. 23, no. 4, pp. 6–20. DOI: 10.17212/1994-6309-2021-23.4-6-20. (In Russian).

[5] Yusubov N.D., Abbasova H.M. Model of machining process control on multi-tool single-carriage adjustments // MACHINE SCIENCE. – 2023. – Vol. 12, No. 1. – P. 22-27. – EDN FUCHCT.

[6] Yusubov N.D., Abbasova H.M. Models of Machining Accuracy in Multi-Tool Adjustments. International Journal of Automotive and Mechanical Engineering (IJAME). vol. 17, no. 3, pp. 8067-8085, Oct. 2020. Malaysia. https://doi.org/10.15282/ijame.17.3.2020.01.0605

[7] Yusubov, N. Theoretical basis for the development of an algorithmic unified complex of mathematical models of cutting forces / N. Yusubov, H. Abbasova, R. Dadashov // Machine Science. – 2023. – Vol. 12, No. 1. – P. 55-60. – EDN SJWPFT.

[8] Mingyang Wu, Jianyu Zhang and others, Study on the Work Hardening and Metamorphic Layer Characteristics of Milling Contour Bevel Gears, Szymon Wojciechowski Antoine Ferreira swears Krzysztof Talaska, 20 p. https://pubmed.ncbi.nlm.nih.gov/36431463/

[9] Gourhari Ghosh, Mayank Kumar, Ajay M. Sidpara, P.P. Bandyopadhyay, Tribological aspects of different machining processes in Machining and Tribology, Available online 14 January 2022, Version of Record 14 January 2022. https://www.sciencedirect.com/topics/engineering/honing

[10] Gafarov, A.M., Sharifov, Z.Z., Gafarov, V.A. Surface Microhardness of Parts Machined by Asymmetric Uninterrupted Lapping: [Electronic resource] / Russian Engineering Research, 39(9), − Moscow, 2019. URL: https://sci-hub.se/https://doi.org/10.3103/S1068798X19090065

[11] Ozturk, S., Kayabasi, E., Celik, E., Kurt, H. Determination of lapping parameters for silicon wafer using an artificial neural network: [Electronic resource] / Journal of Materials Science: Materials in Electronics, 29(1), − Izmir, 2017. URL: https://sci-hub.se/https://doi.org/10.1007/s10854-017-7912-4

[12] Rasulov, N.M., Alekberov, M.Z., Nadirov, U.M. More Efficient Copy Grinding of Complex Surface. Russ. Engin. Res. 41, 829-831 (2021).
https://link.springer.com/article/10.3103/S1068798X21090227

DETERMINATION OF DEFORMATION AND MACHINING ALLOWANCE OF PRECISION PARTS HARDENED BY LASER METHOD

Alakbar Huseynov[1], Ilgar Nazarov[2], Farid Huseynli[1], Mirzabay Safarov[3]

[1]Azerbaijan Technical University, Huseyn Javid Ave.-25, Az-1073, Baku
[2]Azersilah Defense Industry Holding CJSC, Matbuat Avenue, 40, Az-1141, Baku
[3]National Aerospace Agency, Binagadi district, Az-1115, Baku
h_alakbar@aztu.edu.az, ilqar.nezerov@mail.ru, farid.huseynli@aztu.edu.az,
ali.rahman.7534@gmail.com

Abstract

In laser diffusion metallization, precision parts of machines and equipment are heated to a temperature of 1100-1250 °C. Deformation of parts occurs at this temperature. Another reason for the deformation of parts when increasing the strength of the surface with laser technology can be the normalization of the internal stresses of the parts. Deformation of products under conditions of exposure to high temperatures can also arise from its own weight, which is not possible during diffusion processes in densely packed powder environment, but is possible with other saturation methods (gas, vapor vacuum, liquid). The characteristic of the variation of the value of the deformation depending on the thickness of the diffusion zone. The value of the deformation changes with the change of Poisson's ratio. The article is devoted to the determination of the change value of bending during nanodiffusion metallization.

Keywords: stresses, crystallograph, machining, deformation, metallization, laser, nanodiffusion.

I. Introduction

Literature analysis shows that increasing the surface hardness of precision parts by laser has not been fully investigated. When increasing the surface strength of precision parts with a laser, temperature deformations (bending) inevitably occur [1 - 4].

The main reason for deformation is the change of thermal regimes in surface strengthening during nanodiffusion metallization of parts. The higher the cooling rate and the heating temperature, the higher the resulting thermal stresses and their induced deformation [5 - 8]. Therefore, the deformation is inevitable and irreversible, but the change of the geometrical parameters of the deformation can be affected by the heating temperature and the cooling mode [9 - 14].

Another reason for the deformation of parts when increasing the strength of the surface with laser technology can be the normalization of the internal stresses of the parts. Thus, according to Yu.M. Lakhtin [15], compressive stresses in the surface layer after nitriding are 600...800 MPa, and tensile stresses in the depeer layers - in the transition zone - are 200...300 MPa.

Ya.E. Geguzin shows two main reasons for the occurrence of these stresses [16]. The first of

them is that there is a crystallographic mismatch of existing phases at the "layer-core" boundary, which is characterized by the value of the dimensionless ratio $\Delta a/\bar{\alpha}$, where Δa is the difference in the parameters of the contact phases; $\bar{\alpha} = \frac{a_1 + a_2}{2}$ is the arithmetic mean of the phases. Crystallographic mismatch causes tangential stresses that can reach a maximum value at the plane of the interfacial boundary.

$$\sigma_{\tau}^{\Delta a} \approx G\frac{\Delta a}{\bar{\alpha}} \quad (1)$$

where G - is the displacement modulus. $\Delta a/\bar{\alpha} = 5 \cdot 10^{-2}$, $G = 5 \cdot 10^{6}$ N/sm^2, $\sigma_{\tau}^{\Delta a} = 2{,}5\text{-}10^{5}$ N/sm^2 [16].

Full normalization of internal stresses does not occur. The size of the residual elastic stresses should not exceed Payerls limit $\sigma p = (10^{-4} .. 10^{-5})$ G, that is, the value at which dislocations will shift relative to each other – $\sigma \leq \sigma_p$. Due to these stresses, the growth of the bending increases with the increase in the thickness of the diffusion layer.

"The second type of stress in the "layer-core" system occurs when the precessed workpiece, which has different coefficients of linear expansion of the layer and the core, cools. The value of these stress is as follows [16].

$$\sigma_{\tau}^{\Delta T} = G\Delta a T \quad (2)$$

where $G = 5 \cdot 10^{6}$ N/sm^2; $\Delta a = 10^{-5}$ °C^{-1}; $\Delta T = 10^{3}$°C; $\sigma_{\tau}^{\Delta T} = 10^{4}$ N/sm^2.

Deformation of products under conditions of exposure to high temperatures can also arise from its own weight, which is not possible during diffusion processes in densely packed powder environment, but is possible with other saturation methods (gas, vapor vacuum, liquid).

The purpose of the work is to investigate the main cause and theoretical determination of the deformation caused by the technological process of laser surface hardening. In addition, it is necessary to study the influence of the deformation parameters on the geometrical parameters of precision parts.

II. Research Methodology

It is almost impossible to avoid bending of the part during surface hardening with laser technology. Therefore, in order to meet the technical requirements for hardening the surface of precision parts of machines and devices with laser, it eliminates bending by increasing the thickness of the layer and by machining. In this regard, research on the possibility of increasing the surface strength of such high-precision parts with laser technology is required. Let's explain this with Fig. 1.

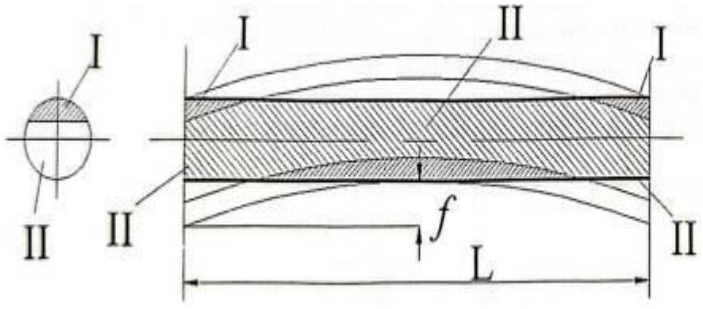


Figure 1: *Machining scheme of the bending of the precision part by laser surface-hardened*

It can be seen from Fig. 1, that it is possible to eliminate the bending (f) by machining. During machining, the thickness of the layer appears in II. A limit layer is a must to ensure surface hardening after processing. This is shown in Fig. 2.

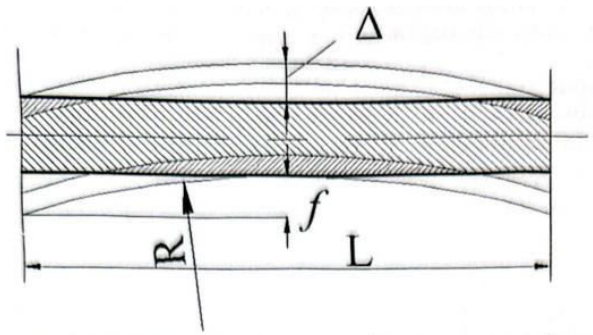


Figure 2: *Determining the limit value of bending*

From Fig. 2, it can be seen that the maximum value of the diffusion layer is greater than the geometrical dimensions of the bending, which allows to meet the mechanical demand implemented on precision details by machining. Thus, the condition that the deformation of parts in the diffusion zone does not occur can be written:

$$\Delta > f \tag{3}$$

where, Δ - is the thickness of the diffusion layer.

It is known from the theory of elasticity that the bending value of the precision part in the elastic deformation zone [17]:

$$f = \frac{q * L^4}{48\, E\, I_x} \tag{4}$$

where,

q – the stress in the part causing the bending moment, MPa;

L – part's lentgh, mm;

E – modulus of elasticity of the material, MPa;

Ix – moment of inertia of the cross section, mm^3.

As a result of laser surface hardening, the plasticity of the part increases. At this time, a sharp decrease in the modulus of elasticity (for example, 42.5% for iron at 1000°C) occurs, which leads to the normalization of the internal stresses present in the workpiece [6]. Eq. 2 would be valid within the limits of Hooke's law. It follows that in the presence of a deformation moment, the bending moment is directly proportional to the value of the stresses, to the fourth degree to the length of the section, and inversely proportional to the modulus of elasticity and the moment of inertia of the curve.

The value of the deformation of the precision part can be expressed by the thickness of the diffusion layer and the stresses present there [4]:

$$f \approx \frac{L^2}{8R} = \frac{3 * d_n * L^2 * (1-\mu) * \Delta * \delta_c}{4(d_n + \Delta)^3} \tag{5}$$

where,

R - is the radius of curvature of the bend;

d_n - diameter of the precision part, mm;

μ - Poisson's ratio;

δ_c - the value of the relative change in the dimensions of the precision part along the line of action of the stress,

$$\delta_c \approx \frac{2\Delta}{d_n}\, 100\% \tag{6}$$

From this dependence, it follows that the value of bending is directly proportional to the square of the length of the precision part. Due to the increase in the thickness of the diffusion layer, while compensating the bending during the machining of the precision part, it allows the surface layer to remain at a thickness that ensures the wear resistance of the part.

The characteristic of the variation of the value of the deformation depending on the thickness of the diffusion zone is shown in Fig. 3. The curves were constructed according to the Eq. 5, assuming d_n = 9 mm, L = 51 mm, geometric dimensions of the plunger of the fuel device: μ_1 = 0.5, μ_2 = 0.2. By changing the value of Δ, the graph of f = F(Δ) is constructed. It follows from Eq. 5 that the value of the deformation changes with the change of Poisson's ratio. It can be seen from the figure that as the thickness of the diffusion layer increases and the coefficient (μ) decreases, the deformation increases.

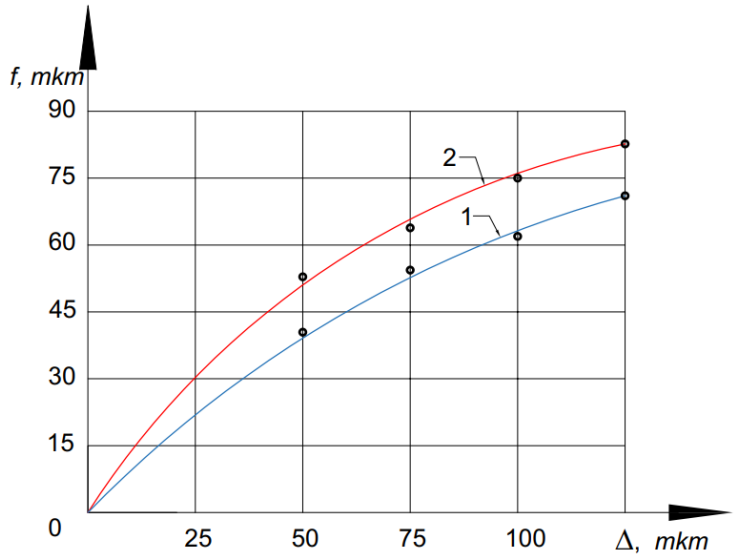


Figure 3: *Dependence of the value of bending (f) on the thickness of the diffusion zone (1-μ=0.5; 2- μ=0.2)*

III. Discussion of the results obtained

In laser surface hardening, the nanodiffusion layer is evenly distributed on the surface of precision parts. In laser surface hardening, the change of the geometric shape of the layer in the precision part after the technological process depends on the quality of the previously processed parts, metallization modes, etc. it depends. This will affect the allowance of the nanodiffusion layer removed during machining.

The calculation of the minimum required diffusion layer for machining of precision parts was developed based on the analytical calculation method proposed by Professor B. P. Kovan [18].

When processing internal and external surfaces, the minimum allowance for barel and shaft type parts according to V.M. Kovan's formula is calculated by the following formula [18].

$$2z_{min} = 2\left[\left(R_{z_{i-1}} + T_{i-1}\right) + \sqrt{p_{i-1}^2 - \varepsilon_i^2}\right] \quad (7)$$

where,

$R_{z_{i-1}}$ - the height of the disparity achieved after the previous operation;

T_{i-1} - the depth of the defective surface layer after the previous operation;

p_{i-1} - spatial deviation of the processed surface relative to the main surfaces of the workpieces obtained in the previous operation;

ε_i - installation error in the placement of precision parts when performing the operation.

In the finishing operation, the displacement of the axis of the precision parts relative to the outer

surface is not corrected, i.e. p_{i-1}= 0. The installation error is zero, because the tool used in finishing and precision parts act as a base for each other during processing, i.e. $\varepsilon_i = 0$.

When processing a nanodiffusion layer in laser surface hardening, there is no defective surface layer, because the properties of the layer in depth are practically the same and do not change during the machining $T_{i-1} = 0$.

The main indicator of quality in the technological process of finishing is the roughness of the surface. It is clear from the Eq. 7 that the minimum allowance for the delivery of the diffusion layer is: $z_{min} = R_{z_{i-1}}$.

Eq. 7 does not take into account the effect of the initial geometrical failure on the allowance.

Geometrical failure (deformation, ovality, conicity, barreling, etc.) obtained after the laser surface hardening process are 5...6 times greater than the value of form deviations allowed from previous operations [8]. This determines the allowance during the finishing operation. The deformation after the laser technological process is 6 μm on average [19].

Therefore, the Eq. 7 for calculating the allowance in the finishing operation is expressed by the following formula:

$$Z_{min} = \kappa_i + \sum_{i=1}^{n} R_{z_i} \tag{8}$$

where,

κ_i – the largest failure of the part after laser surface hardening;

n – number of finishing operation.

After the laser surface hardening process, the main variable of precision parts is deformation, which includes other failurer.

From here:

$$\kappa_i = f_{кр} \tag{9}$$

where, $f_{кр}$ – deformation of precision parts.

$f_{кр}$ - quantity can be found using the equation proposed in Eq. 5 [20].

$$f_{кр} \approx \frac{L^2}{8R} = \frac{3 * d_n * L^2 * (1-\mu) * \Delta * \delta_e}{4(d_n + \Delta)^3} \tag{10}$$

where,

R - is the radius of curvature of the deformation, mm;

L - the length of the precision section, mm;

d_n- diameter of the precision part, мм;

μ - Poisson's ratio;

δe - the value of the relative change in the dimensions of the precision part along the line of action of the stress;

Δ - thickness of the nanodiffusion layer.

$$\delta_e \approx \frac{2\Delta}{d_n} \tag{11}$$

The final formula for the machining allowance (z_{min}) during machining of precision parts with laser surface hardening is as follows.

$$z_{min} = \frac{3}{4} \frac{d_n * L^2 * (1-\mu) * \Delta * \delta_e}{(d_n + \Delta)^3} + \sum_{i=1}^{n} R_{z_i} \tag{12}$$

During the machining of precision parts with a surface hardened by laser, a relationship was established along the processing allowance, the deformation in the parts, and the total height of the micro-roughness in the previous operation.

By knowing the thickness of the layer in laser surface hardening, we can determine the minimum required allowance for processing.

Determination of the minimum allowance after the laser surface hardening of the plunger of the fuel device in the YaMZ engine: the diameter of the plunger - 9 mm; plunger length - 51 mm; according to the experiment, the total height of the microroughness is 0.0016 mm; the value of δ_e is equal to 0.008 for the thickness of the diffusion layer - 0.125 mm; Poisson's ratio - 0.5 [21];

$$z_{min} = \frac{3 \bullet 9 \bullet 51^2 \bullet (1-0{,}5) \bullet 0{,}125 \bullet 0{,}008}{4 \bullet (9+0{,}125)^2} + 0{,}0016 = 0{,}107$$

According to the experimental data, the thickness of the layer for machining is in the range of 0.1-0.110 mm, which is in good match with the theoretical data determined by the Eq. 12.

IV. Conclusion

Thus, it follows from the above that in increasing the surface strength of precision parts with a laser, the deformation of the parts, as well as the allowance of machining, should be taken into account.

References

[1] F.M. Castro Cerda, C. Goulas, D. Jones, A. Kamyabi, D. Hamre, P. Méndez, G. Wood, Modelling the laser surface hardening process in a steel with a spheroidized initial microstructure, Journal of Manufacturing Processes, Volume 125, 2024, Pages 364-373, ISSN 1526-6125.

[2] HS Dewi, J Volpp, AFH Kaplan, Short thermal cycle treatment with laser of vanadium microalloyed steels, Journal of Manufacturing Processes, 2020 Elsevier

[3] Balasubramanian, K.R., Babu, P.D., Buvanashekaran, G., 2015. Laser transformation hardening of steel, Second ed., Manufacturing engineering: new research. Elsevier Ltd. doi:https://doi.org/10.1016/b978-0-08-101252-9.00011-x.

[4] A.G. Guseinov Mechanical processing of precision parts with diffusion coatings. Machine engineering, Moscow, 2002, No. 11, pp. 51-57.

[5] Zhiping Zhou, Jinlong Lv, Maolei Gui, Weiqi Yang, New insights into annealing induced hardening and deformation mechanisms in a selective laser melting austenitic stainless steel 316L, International Journal of Plasticity, Volume 178, 2024, 104008, ISSN 0749-6419, https://doi.org/10.1016/j.ijplas.2024.104008.

[6] K.-K. Yoon, W.-B. Kim, S.-J. Na, Shape deformation of a piston ring groove by laser surface hardening, Surface and Coatings Technology, Volume 78, Issues 1–3, 1996, Pages 157-167, ISSN 0257-8972, https://doi.org/10.1016/0257-8972(94)02403-0.

[7] Philipp Schüßler, James Damon, Fabian Mühl, Stefan Dietrich, Volker Schulze, Laser surface hardening: A simulative study of tempering mechanisms on hardness and residual stress, Computational Materials Science, Volume 221, 2023, 112079, ISSN 0927-0256, https://doi.org/10.1016/j.commatsci.2023.112079.

[8] Lyakhovich L.S., Voroshnin L.G., Panich G.G., Shcherbakov E.D. Multicomponent diffusion coatings. - Mn.: Science and Technology, 1974, 288 p.

[9] A.G. Huseinov. Change in increase in linear sizes of parts during diffusion metallization. Mechanical engineering. Baku, No. 2, 2002, pp. 52-55.

[10] Xinlei Pan, Liucheng Zhou, Dianyin Hu, Weifeng He, Ping Liu, Zichuan Yu, Xiaoqing Liang, Superior wear resistance in cast aluminum alloy via femtosecond laser induced periodic surface structures and surface hardening layer, Applied Surface Science, Volume 636, 2023, 157866, ISSN 0169-4332, https://doi.org/10.1016/j.apsusc.2023.157866.

[11] A.G. Huseynov Determination of the thickness of the diffusion coating during restoration and strengthening of machine parts. Вестник машиностроения, Moscow, 2003, №12, p.82-83.

[12] Filonenko B.A. Complex diffusion coatings, M.: Mashinostroenie, 1981, 16 p.

[13] Eberschbach G., Men E. Thermodynamics and kinetics of the formation and deposition of titanium carbide layers on metallic substrates, -Die Technik, 1976, y 31, N7, p 467-470.

[14] Shcherbedinsky G.V., Kondratenkoonnykh L.A. Diffusion growth of phases in three-component systems in the presence of mutual influence of elements. In: Protective coatings on metals. Kyiv, 1971, Vol.5, pp. 23-30.

[15] Lakhtin Yu.M. Metal science and heat treatment of metals. Moscow: Metallurgy, 198, 60 p.

[16] Geguzin Ya.E. Diffusion zone, M., Nauka, 1979, 344 p.

[17] Fedosyev V.I. Strength of materials. Moscow: Nauka, 1979, 560 p.

[18] Kovan V.M. Calculation of allowances for processing in mechanical engineering. - M.: Mashgiz, 1968. - 208 p.

[19] Guseinov A.G. Analysis of plunger deflection of the UTN-5 fuel pump in the process of diffusion metallization. Repair and reliability of agricultural machinery. Collection of scientific papers of MIISP. - M., MIISP. - 1985. - P. 14-15.

[20] Yuryev S.F. Deformation of steel during chemical-thermal treatment. Cementation and nitriding. - M.-].; - 1950. - 307 p.

[21] Belyaev N.M. Strength of materials. - M.: Nauka, 1976. - 608 p.

INVESTIGATING THE APPLICATIONS AND IMPACTS OF LARGE LANGUAGE MODELS IN CHEMISTRY EDUCATION

Canan Kocak Altundag[1], Sencer Yucel[1], Faxraddin Yusubov[2]

•

[1]Turkey, Ankara, Hacettepe University, Beytepe
[2]Azerbaijan Technical University, Baku, H. Cavid avenue, 25
canan.kocak@hacettepe.edu.tr, senceryucel@hacettepe.edu.tr, yusubov.faxraddin@aztu.edu.az

Abstract

Organic chemistry, which is an experimental science, deals with the acquisition and characterization of pure organic compounds from natural or synthetic pathways by appropriate methods. In addition to a good textbook in chemistry education, the laboratory of this course is also of great importance. For the realization of the processes, equipment made of various materials and their efficient use are required. This study has been prepared in order to make this laboratory more understandable and more enjoyable for students to recognize organic synthesis and see its products. In this article, Iodoform Synthesis experiment is taken as the basic experiment to explain the system. A virtual experimental environment powered by a Large Language Model (LLM) was created. The system presents students with a multi-step experiment which they are asked to complete correctly. The innovative side of this environment is how it interacts with the student when they make a mistake. This study was also conducted in an Organic Laboratory class, and thus created a research question: How did the idea of using Large Language Model in organic chemistry course affect student's mental workload?

Keywords: Laboratory instruction, large language model, chemistry education.

I. Introduction

Organic chemistry is a sub discipline of chemistry that investigates the synthesis, reactions and properties of compounds containing carbon in their structure Organic chemistry is basically an experimental field, the final decisions on the identity and structure of reaction products, their thermodynamic properties, and spectroscopic appearance are based on observation and measurement [1]. The laboratory is the place where students learn and apply the synthesis, separation, purification and analysis techniques of organic compounds and thus become acquainted with the experimental aspect of organic chemistry. However, the laboratory environment can be very complex for students. Students may have difficulty in performing experiments in organic chemistry laboratories. This study investigates the role of virtual experiment environment strengthened by an LLM in chemistry education. Artificial intelligence (AI) tools are spurring rapid developments in the chemical industry and academia [2]. Traditional learning methods in chemistry education, often limited to static textbooks and one-way knowledge transfer, have been changing lately to interactive AI tools.

II. The applications in organic chemistry laboratory

Students often use worksheets while conducting experiments in organic chemistry laboratory applications. With worksheets, known laws, principles, etc. can be revealed (induction) or verified (deduction) [3]. In the worksheet, the diagram of the experimental setup, the steps of the experiment, the collection of data and the evaluation of the data are given. Sometimes students can be expected to find the result by asking partially guiding questions [4]. These students who completed the experiment in accordance with the worksheet were not mentally challenged and did not succeed in becoming active participants of the experiment [5]. Because it is seen that a printed scenario can neither present specific laboratory techniques ready for reproduction nor prepare students for laboratory work well enough [6].

In order to provide students with a successful laboratory experience, it is necessary to use authentic organic chemistry applications that reflect organic chemistry practices and include modern techniques [7]. It is thought that especially prospective chemistry teachers who are not successful enough in experiments will react positively to such an innovative application in organic chemistry laboratory. Because learning can be more meaningful when the cognitive load caused by experimental work can be reduced [8].

One of the unique problems of students during organic laboratory applications is mental workload. Mental workload is defined as the amount of mental work required for a person to complete a task over a certain period of time. The fact that organic chemistry laboratory courses are conducted via experiments creates a mental workload rather than a physical workload. Issues such as the temporal pressures on students caused by long lectures at the laboratory and experimental reports that must be completed by the due date, and coordinating with the courses by showing intense effort create mental workload on students.

III. The Large Language Models in Organic Chemistry Laboratory Development

Organic Chemistry Laboratory I and II courses include a multifaceted education in which the basic concepts of organic chemistry and the necessary scientific skills are acquired. For this study, a different laboratory day was prepared for 16 students. The students were not given laboratory worksheets to perform the iodoform synthesis experiment. The students were given only materials required to synthesis the iodoform. A special application was prepared for students that they could follow on their mobile phones. We created a virtual experiment environment strengthened by an LLM. As the LLM, OpenAI's ChatGPT-3.5-Turbo was used. The system presents students with a multi-step experiment which they are asked to complete correctly.

The Iodoform Synthesis experiment is taken as the base experiment to explain the system. The student is expected to complete the experiment in the correct order among the mixed steps. There are some locked steps as can be seen (Figure 1). Even in this relatively short experiment consisting of only 8 steps, a mistake in one critical step could lead to many different outcomes and confuse the student. Thus, we created locked steps and did not allow the students to make mistakes when they came to the locked step and forced them to find the correct step without prompting any question to the LLM. Locked steps change their state once they have been selected in the correct order.

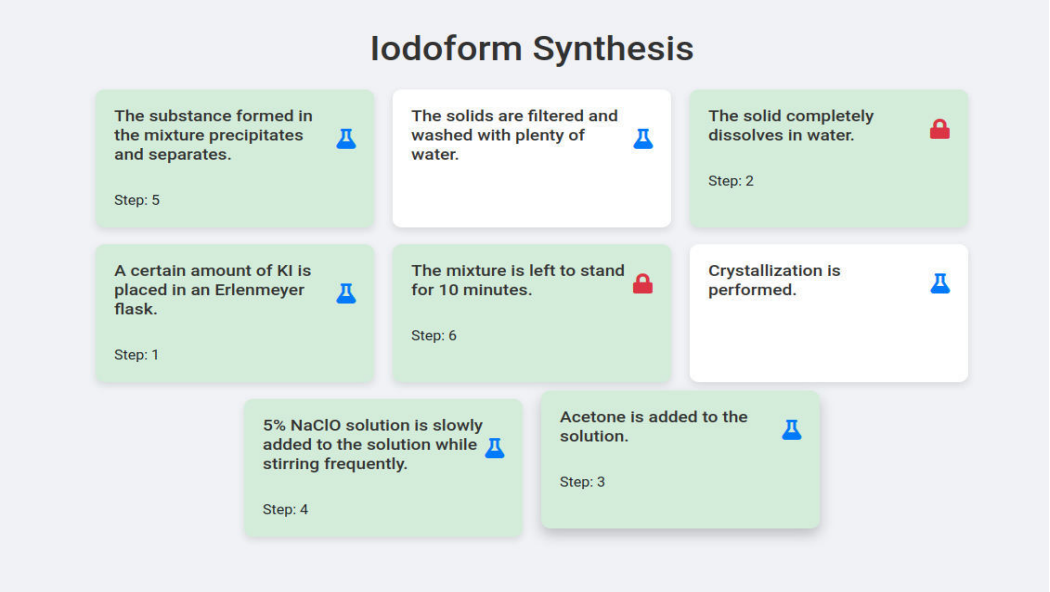


Figure 1: *The user interface seen by a student who has successfully reached step 7*

Below is an example of how the system feedback the student once they selected the step "The mixture is left to stand for 10 minutes." instead of "The substance formed in the mixture precipitates and separates.":

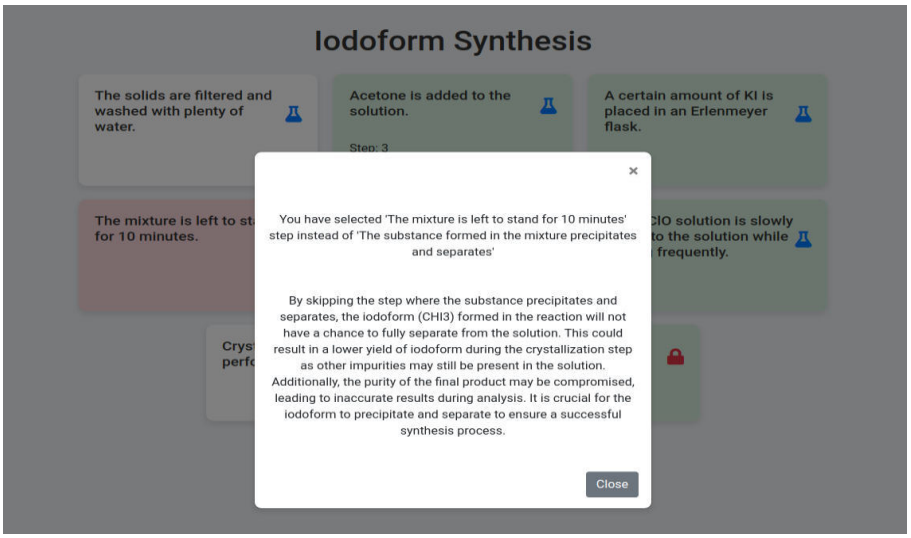


Figure 2: *LLM's response to a student who made a mistake in step 5 of the experiment*

Thanks to the working principle of LLMs, the system does not give the same answer when the same mistake is made again. It explains the same result with different words and expressions. This is a serious advantage as a student is likely to grasp something that he/she could not fully grasp before, using different expressions. Furthermore, there is no chance for the system to remain silent against a mistake.

IV. Data Collection Tool

This study, conducted in an Organic Laboratory Course through Large Language Model, had a research question. How did the idea of using Large Language Model in organic chemistry course affect student's workload? Several types of uni- and multidimensional subjective scales exist.

However, results from various studies have shown that NASA-TLX [9] is superior to SWAT in terms of sensitivity, especially for low mental workloads [10, 11]. Therefore, it was decided to use the NASA-TLX method to determine the mental workload of students (jobs required intense mental demands, physical requirements, and time constraints imposed by long-term analyses, etc.).

NASA-TLX is a multidimensional scale for which the overall mental workload is a function of mental demand, physical demand, temporal demand, performance, effort, and frustration dimensions, with each of these dimensions on a continuum. In the method, the workload score is determined in stages: scoring and ranking. Students in the first stage; considering the work they did, scored the workload they felt for the 6 dimensions by valuing a scale between 0 (very low) and 100 (very high) (divided into 5-point intervals) [12]. The determined scores are considered at five levels: very low (0-20), low (21-40), medium (41-59), high (60-79), and very high (80-100) [13].

The workload score was created by multiplying the scores given to the six criteria by the percentages of repetitions of the choices prioritized in the pairwise comparison. The NASA-TLX and mobile application solutions are evaluated according to the levels determined in the NASA-TLX solution.

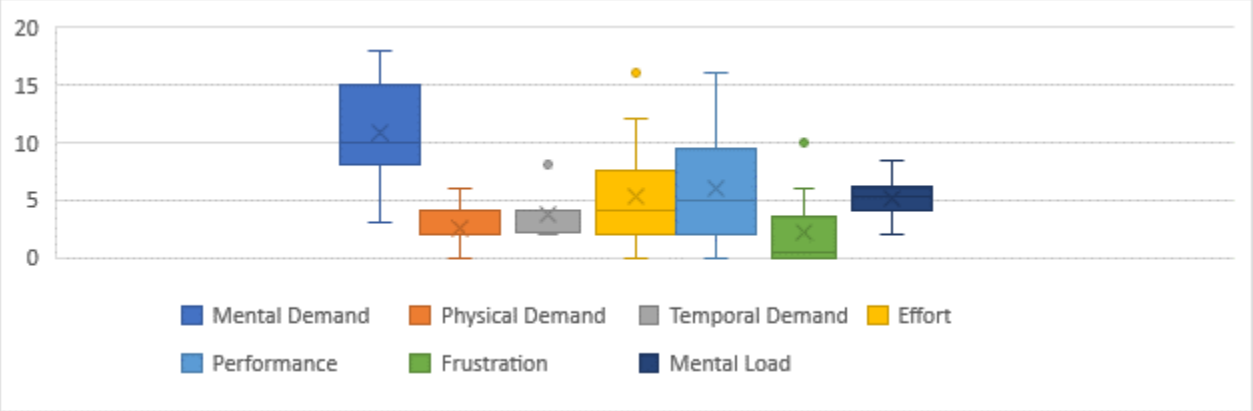


Figure 3: *NASA-TLX solution*

It was determined that the mental workload of 5 students was between 0-20 points, which is considered to be a very low level. It was determined that the mental workload of 11 students was between 21-40 points, which is considered to be at a low level.

IV. Student Views of the LLM

This study investigates student responses of virtual experiment environment strengthened by an LLM in an organic laboratory chemistry course. At the end of the study, the performance of application asked to the students. An 18-app survey questions, with scores ranging from 0 to 10, was administered to assess the students' experience. Student responses are shown in Table 1.

Table 1: *Students' responses by survey statement.*

Item		1	2	3	4	5	6	7	8	9	10
							%				
1	How do you rate your overall experience with our app?							18	37	12	33

2	How difficult is it to read the characters?	75	20	5						
3	How useful is our product for you?					12	18	33	37	
4	How would you rate the user-friendliness of the product interface?					26	18		56	
5	Considering that you use our interface extensively, how likely are you to recommend it to your friends and colleagues?					20	25	5	50	
6	If you were to review the product, how would you rate it out of 10?					18	32	32	18	
7	Overall, how easy did you find it to use?					8	10	32	50	
8	How would you rate the usability of our app?						32	43	25	
9	How would you rate your experience?				44			37	19	
10	How would you rate the reliability of app?			1	18	18	10	12	37	
11	How do you evaluate the quality of the content presented on the screen?			3	3	10	10	32	32	
12	How would you rate the speed of our application?		3	3	10	10	32	32		
13	How would you evaluate the interface design of our application?			12	15	9	43	12	9	
14	How useful are the displayed error messages?					13	13	13	61	
15	How well does our app keep you informed about the progress of a task?						25	25	50	
16	How difficult is it to read the characters on the screen?						19	19	62	
17	How consistent is the use of terms across the application?					25	25	7	43	
18	How much do you like the look of this app?			1	18	18	10	12	37	

Student responses are shown in Table 1. Students engaged with LLM to follow experiment steps, revealing overall positive perceptions toward their usefulness. The responses in general were found to have high agreement in terms of playability, usability, and content and demonstrated that the LLM had positive features. Only item 12 showed that students were indecisive in speed of the application. This might stem from the fact that this was the first time these students were participating in a LLM about organic chemistry laboratory. In addition, they used the LLM with the slow internet.

VI. Results

Chemistry teacher candidates are given theoretical knowledge in organic chemistry I and II courses at the university, while experimental organic chemistry knowledge is explained comparatively in organic chemistry laboratory I and II courses. Students can learn both the application of theories and models in the interpretation and explanation of experimental results and the chemistry of organic compounds comparatively. Students derive their knowledge from the analysis of data collected by scientific methods. Therefore, laboratory applications are very important for chemistry education [14].

This study created a virtual experiment environment strengthened by an LLM. The LLM application includes the application steps of the 'Iodoform Experiment' that students should do in the organic chemistry laboratory. This application was a route that helped students to perform the steps of the experiment in order and was drawn to iodoform synthesis. The application aims to introduce students to modern chemistry through the use of artificial intelligence in organic chemistry laboratory. To date, despite the interesting progress made in artificial intelligence, artificial intelligence applications have not been common in chemistry teaching laboratories. The existence of an error rate should not be ignored in any AI system, but this is still not an obstacle to systems being able to use AI. With the development of artificial intelligence, it is inevitable that some of the traditional education methods we know will disappear, while others will change radically. In this first phase of LLMs, we see that learning and education can be significantly personalized. The use of LLMs in chemistry education offers a different learning style by providing a real-time feedback system in a virtual experimental environment.

As of the day this paper was written, one of the most advanced LLMs is ChatGPT from OpenAI. The most ideal way is to fine-tune an LLM for this specific purpose. Due to its requirements of serious resources, we did not go into fine-tuning, instead we built the system on ChatGPT in a way that can be considered a proof of concept. Overall, students in this study have a positive view of the utility of generative virtual experiment environment strengthened by an LLM in the organic chemistry laboratory course.

The minimum number of errors a student can possibly make in a virtual experiment consisting of 8 steps in total, 2 of which are locked, is 15. Similarly, for a 15-step experiment with 3 steps locked, the minimum number of errors that can be made is 66. It takes an unprecedented knowledge for a teacher to know what the consequences of all the different combinations of errors will be in all the different chemistry experiments. However, LLMs have already reached this level of knowledge, because the datasets they were trained on already contain the vast majority of errors that can be possibly made. One important application is the use of Large Language Models (LLMs) to provide instant feedback in virtual experimental environments. In this paper, we explored the role of LLMs in a virtual chemistry experimental environment and investigated their applications and impact on improving student learning outcomes. The innovative side of this environment is how it interacts with the student when they make a mistake. Rather than simply marking the answer as incorrect, the system stops the experiment and prompts an LLM to analyze the error and provide a technically detailed explanation. Research data were obtained from NASA-TLX. LLM application has enabled

them to do the experiment in the remaining time of a course, and it has not burdened them with an intense study task and a very high mental workload. With LLMs, the effects of such a stressful workload could be alleviated and flexibility could be provided.

References

[1] Carey, F. A. & Sundberg, R. J. (2007). Advanced organic chemistry: part A: structure and mechanisms; Springer Science & Business Media.

[2] Baum, Z. J., Yu, X., Ayala, P. Y., Zhao, Y., Watkins, S. P. & Zhou, Q. (2021). Artificial Intelligence in Chemistry: Current Trends and Future Directions. J. Chem. Inf. Model, 61 (7), 3197−3212.

[3] Hand, B., & Treagust, D. F. (1991). Student achievement and science curriculum development using a constructive framework. School Science and Mathematics, 91(4), 172-176.

[4] Redfield, D. L., & Rousseau, E. W. (1981). A meta-analysis of experimental research on teacher questioning behaviour. Review of Educational Research, 51(2), 237-245.

[5] Sands, M. & Özçelik, D.A. (1997). Application studies in schools secondary education. Ankara: Agency.

[6] Pölloth, B., Teikmane, I., Schwarzer, S., & Zipse, H. (2019). Development of a modular online video library for the introductory organic chemistry laboratory. Journal of Chemical Education, 97(2), 338-343.

[7] Slade, M. C., Raker, J. R., Kobilka, B., & Pohl, N. L. (2014). A research module for the organic chemistry laboratory: Multistep synthesis of a fluorous dye molecule. Journal of Chemical Education, 91(1), 126-130.

[8] Bretz, S. L. (2001). Novak's theory of education: human constructivism and meaningful learning. Journal of Chemical Education, 78 (8), 1107.

[9] Hart, S. G., Wickens, C. D. (1990). Workload Assessment and Prediction. In MANPRINT: An Approach to Systems Integration. Dordrecht: Springer Netherlands, pp. 257-296.

[10] Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. Human Factors, 33,17- 33.

[11] Reid, G. B.; Nygren, T. E (1988). The Subjective Workload Assessment Technique: A Scaling Proce-Dure for Measuring Mental Workload. In P. A. Hancock & N. Meshkati (Eds.), Human Mental Workload. Amsterdam: Elsevier, 1988; pp. 185–218.

[12] Kurtgün, M. Determination of mental workload with the fuzzy vikor method. Msc thesis, Balikesir University Institute of Science Industrial Engineering. Balıkesir, Turkey.

[13] Riono, R., Suparno, S.& Bandono, A. (2018). Analysis of mental workload with integrating NASA TLX and Fuzzy method. Journal Asro, 9(1), 37-45.

[14] Cooper, M. M., & Kerns, T. S. (2006). Changing the laboratory: Effects of a laboratory course on students' attitudes and perceptions. Journal of Chemical Education, 83(9), 1356.

MATRIX MODEL OF ACCURACY IN MACHINING CONICAL SURFACES ON CNC LATHES

Nizami Yusubov[1], Heyran Abbasova[1], Ramil Dadashov[1]

•

[1]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan
nizami.yusubov@aztu.edu.az, abbasova.heyran@aztu.edu.az, dadashov@aztu.edu.az

Abstract

The article presents the development of a matrix model for accuracy in machining conical surfaces on CNC lathes. Spatial computational schemes of elastic displacements of technological subsystems were constructed based on the balance of force interaction between the cutting tool and the workpiece. These schemes cover both single-tool setups with a rotating carriage and single-tool two-coordinate setups. Additionally, models of dimensional distortions were developed, taking into account the characteristics of the setups, and were analyzed based on cutting conditions, cone angle, and the comprehensive compliance characteristics of the technological systems. It is also possible to calculate setup parameters using the proposed models. The developed models can be applied in computer-aided design (CAD) systems for machining conical surfaces on CNC lathes, as well as for studying and predicting the influence of various parameters on the dimensional accuracy achieved during conical surface machining..

Keywords: conical surface machining, single-tool setups with rotating carriage, single-tool two-coordinate setups, matrix model of accuracy, CNC lathes, dimensional distortions, combined compliance matrix of technological system subsystems.

I. Introduction

In machining processes such as turning, milling, and grinding, the dimensional accuracy and form error of the produced surface are determined by the dynamics of the interaction (i.e., the force characteristics of the process) between the tool and the workpiece. The interaction of the main elements of the technological system is significantly influenced by auxiliary elements, such as fixtures, handles, machine carriages, and others [1-10].

From the perspective of the mechanism of error formation in machining, the technological system is considered a feedback system [3-4]. The application of the laws of analytical mechanics allows for the description of the error formation mechanism in machining, but it leads to nonlinear systems of equations.

Let us examine how the developed general matrix models of force interaction among the subsystems of the technological system [3-4] during the machining of conical surfaces are transformed into models of machining errors.

II. Single-tool setups with a rotating carriage

On conventional vertical multi-spindle semi-automatic lathes [1], as well as on several modern CNC machines [2-4], a rotating carriage is provided, which allows for the machining of conical surfaces using a contour cutting tool. The guideways of the carriage are inclined at an angle corresponding to the cone's generatrix, enabling the feed along the contour of the workpiece being machined. In such a setup, the cutting tool is oriented normal to the machined surface, meaning that its elastic displacements caused by the force interaction between the elements of the technological system are also measured relative to this normal. At the same time, the resultant dimension (diameter) is measured perpendicular to the axis of the machined surface. The misalignment of the cutting tool's orientation and the resultant dimension leads to a change in the degree of influence of each component of the cutting force on the distortion of the resultant dimension compared to setups using longitudinal or transverse carriages aligned with the axes of the workpiece and the directions of the dimensions being machined. As a result, the model of dimensional distortion [3, 4] developed for machining a workpiece with a carriage aligned along its axes is not applicable in this case. The situation is further complicated by the fact that the formulas for the components of the cutting forces are developed in the tool's coordinate system: the y-axis is directed normal to the feed vector, the x-axis is along the feed vector, and the z-axis is perpendicular to the first two axes. As a result, when determining the resultant force, it is necessary to account for the angle of rotation of the tool's coordinate system in relation to the workpiece's coordinate system. Consequently, the traditional scheme for calculating the coordinates of a single-tool setup [3, 4] is transformed into the form presented in Figure 1 [11].The situation is further complicated by the fact that the formulas for the components of the cutting forces are developed in the tool's coordinate system: the y-axis is directed normal to the feed vector, the x-axis is along the feed vector, and the z-axis is perpendicular to the first two axes. As a result, when determining the resultant force, it is necessary to account for the angle of rotation of the tool's coordinate system in relation to the workpiece's coordinate system. Consequently, the traditional scheme for calculating the coordinates of a single-tool setup [3, 4] is transformed into the form presented in Figure 1 [11].

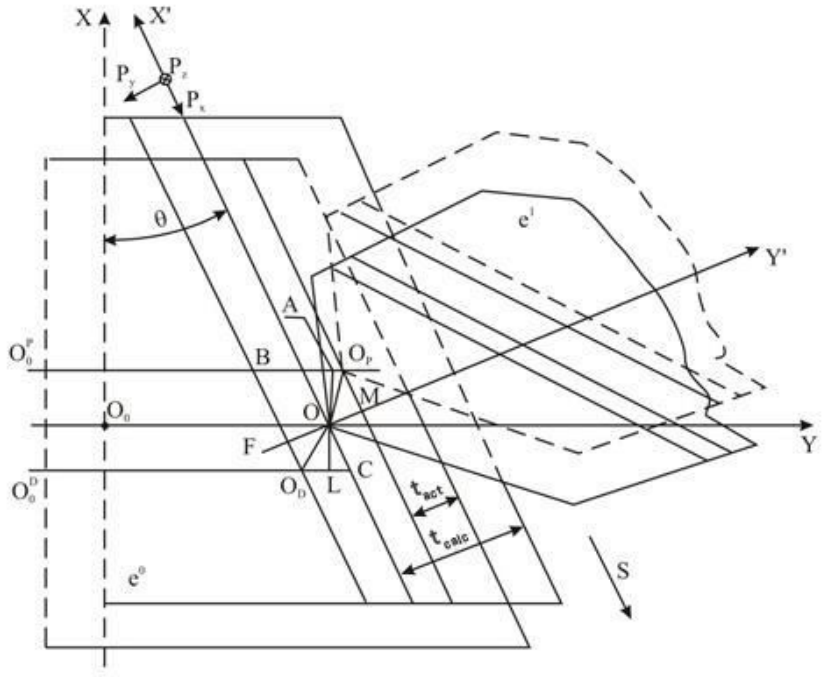


Figure 1: *Spatial computational scheme of elastic displacements in technological subsystems during single-tool machining with a rotating carriage [11]*

In Figure 1, the following designations are presented: ODL - movement of the "spindle-workpiece" subsystem along the OY axis, OL - movement of the "spindle-workpiece" subsystem along the OX axis, movement of the "carriage-tool" subsystem along the OY' axis, OpM - movement of the "carriage-tool" subsystem along the OX' axis, BO_P – dimensional distortion in the diametral direction, t_{calc} – calculated depth of cut, t_{act} – actual depth of cut after processing, S- feed rate along the contour, e^0, e^1 - compliance matrices of the "spindle-workpiece" and "carriage-tool" subsystems, respectively, θ – inclination angle of the cone's generatrix relative to the axis of the processed workpiece.

As can be seen, the considered setup allows the decomposition of the technological system into two subsystems: O_0XYZ– the coordinate system of the "spindle-workpiece" subsystem, and $OX'Y'Z'$- the coordinate system of the "carriage-tool" subsystem [3, 4]. As a result, the problem of elastic displacements in the technological system with this setup is reduced to studying a two-body system, for which a matrix model of the contact point's displacements has been developed [3]. Under the influence of cutting forces, the workpiece undergoes displacement ($O_0^D O_D$) due to elastic deformations in the "spindle-workpiece" subsystem. In response to the cutting forces, the tool also undergoes displacement ($O_0^P O_P$) due to elastic deformations in the "carriage-tool" subsystem, but in the opposite direction.

The basis for the model of dimensional distortions in the studied setup can be taken as the matrix model [3-4], developed for the setup with the carriage oriented along the coordinate axes of the workpiece:

$$\overline{g^{01}} = e^{01}\,\bar{P} \tag{1}$$

where $\overline{g^{01}}$ is the vector of dimensional distortions, e^{01} is the combined compliance matrix of the subsystems of the technological system, and $\overline{P}$ is the resulting cutting force applied to the workpiece.

Equation (1), written in matrix form, is developed for the workpiece coordinate system. For this reason, to adapt it to the studied setup, the vector $\overline{P}$ must also be expressed in the same coordinate system. Using known coordinate transformation relations, the vector $\overline{P}$ is obtained as:

$$\bar{P} = \begin{pmatrix} P_x \cos\theta + P_y \sin\theta \\ P_y \cos\theta - P_x \sin\theta \\ P_z \end{pmatrix} \tag{2}$$

In formula (2), P_x, P_y, and P_z are the components of the cutting force in the tool's coordinate system. The following well-known formulas from cutting theory apply to them: $p_i = c_i t^{x_i} s^{y_i} v^{z_i}$ $i = x; y; z$ [12], where c_i– are coefficients depending on the type of material being processed and other cutting conditions; t, s, v – represent the depth of cut, feed rate, and cutting speed, respectively; x y z – are exponents corresponding to the respective variables.

Considering these formulas and applying relation (2), the matrix model for the case of single-tool processing (1) can be represented in expanded form as follows [11]:

$$\begin{pmatrix} g_x \\ g_y \\ g_z \end{pmatrix} = \begin{pmatrix} e_{xx}^{01} & e_{xy}^{01} & e_{xz}^{01} \\ e_{yx}^{01} & e_{yy}^{01} & e_{yz}^{01} \\ e_{zx}^{01} & e_{zy}^{01} & e_{zz}^{01} \end{pmatrix} = \begin{pmatrix} c_x t^{x_x} s^{y_x} v^{z_x} \cos\theta + c_y t^{x_y} s^{y_y} v^{z_y} \sin\theta \\ c_y t^{x_y} s^{y_y} v^{z_y} \cos\theta - c_x t^{x_x} s^{y_y} v^{z_x} \sin\theta \\ c_z t^{x_z} s^{y_z} v^{z_z} \end{pmatrix} \tag{3}$$

Considering the projections of the displacement vector $\overline{g}$ and the dimensional distortions in the directions of interest: g_y - diametral dimension (along the Y axis); g_x - linear dimension (along

the X axis). The magnitude of the distortion of the diametral dimension g_y will be determined by the element-wise multiplication of the second row of the compliance matrix of the technological system e^{01} and the force vector.

The obtained model describes the distortions of the performed dimensions and the distortion of the shape of the processed surface. It can be used to calculate the setup dimensions and serves as the basis for determining other indicators of processing accuracy.

III. Single-tool two-coordinate setups

CNC lathes allow for the machining of conical and even complex contoured surfaces without the need to rotate the carriage [13-16]. The formation of the contoured surface is achieved by superimposing two coordinate movements of the tool that are performed simultaneously: the controlled movement along the Z coordinate, referred to as feed Sz, and the controlled movement along the X coordinate, referred to as feed Sx. The main difference between such setups and those discussed in [13-19] lies in the control over two coordinates. Therefore, it is reasonable to refer to these setups as two-coordinate setups.

Since the tool's trajectory in a single-tool two-coordinate setup is formed by two simultaneous feeds, the scheme for forming the processing error becomes significantly more complicated. In the single-tool single-coordinate setup [16] and in the single-tool setup with a rotating carriage [5], the tool was oriented relative to the processed surface (perpendicular to it); however, in the current case, the tool is oriented relative to the axis of the processed surface. Therefore, when determining the distortion of the performed dimension, it is necessary to account for the transition from the tool's coordinate movements to movements relative to the processed surface.

It is evident that for single-tool two-coordinate setups, the decomposition of the technological system into two subsystems is quite acceptable [13-19]. Therefore, as a methodological basis for the model of dimensional distortions in such setups, one can use the analytical model of movements in a two-body system with elastically deformable connections [4,11].

To identify the fundamental features of the influence of two-coordinate control on the distortion of the performed dimension, let us first consider the simplest case of machining—a conical surface. Taking into account the specifics of error formation during two-coordinate control of the machining process, the overall calculation scheme of the deformational interaction of the two-body system is transformed for the technological subsystems to the form presented in Figure 2.

Here, matrices e^0 and e^1 characterize the compliances of the technological subsystems—subsystem 0 (spindle – chuck – workpiece) and subsystem 1 (carriage – tool holder – tool), respectively. The line G describes the contour of the workpiece in its initial state (without any force acting on the technological system). The line D describes the contour of the part, i.e., the contour after the allowance has been removed (its calculated position). Point M is the calculated position of the tool tip (which determines the setup dimension).

After applying cutting forces $\bar{P}$, the technological subsystems undergo elastic displacements g, and accordingly, the previously described elements of the calculation model also shift. As a result of the displacement of subsystem 0, the contour of the workpiece will move to line G0, and line D (the calculated contour of the part) will take the position of line D0. Point M is the contact point between the workpiece and the tool; therefore, we can consider that there are two overlapping points M here: one belongs to the workpiece, and the other belongs to the tool tip. As a result of the force applied, both technological subsystems experience elastic displacements, causing point M on the workpiece (subsystem 0) to move to position M0, and point M on the tool tip (subsystem 1) to move to position M1. Thus, the actual contour of the part formed in the force interaction moves to line D1.

Thus, the tool tip, instead of the setup radius Rs, will form the actual radius Ract, thereby determining the distortion of the performed dimension.

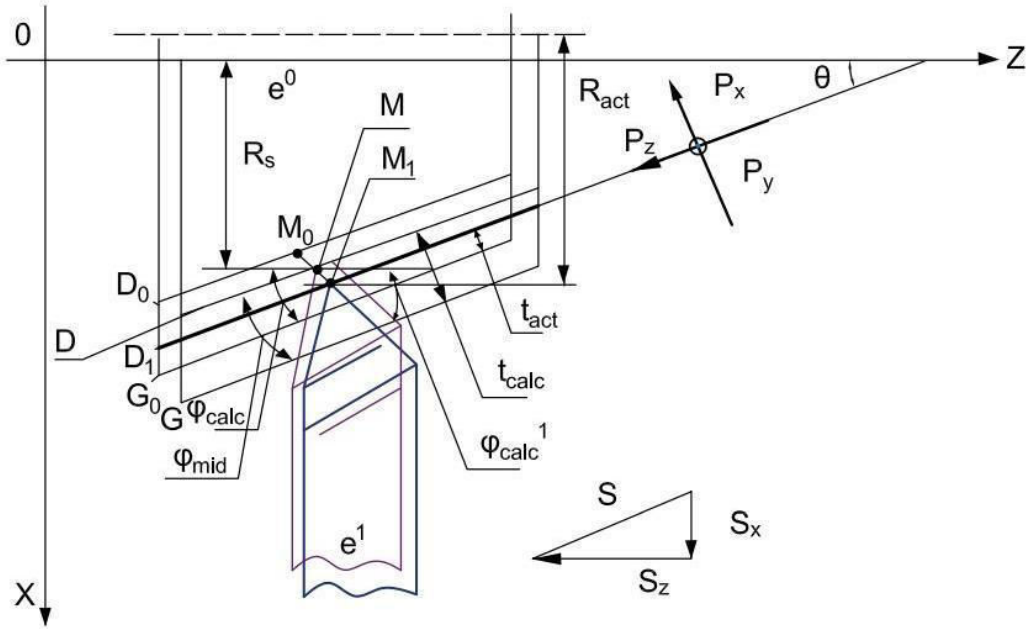


Figure 2: *The calculation scheme of elastic displacements of technological subsystems during single-tool two-coordinate machining of an external cone*

The calculation scheme in Figure 2 illustrates the force interaction between two contacting subsystems. In analytical mechanics, such a scheme is reduced to the interaction of a two-body system with elastic connections and, as shown in [4,11], is described by the matrix model (1), which in this case takes the following form:

$$\overline{MM_0} = g_0 = e^0\bar{P} \tag{4}$$

$$\overline{MM_1} = g_1 = -e^1\bar{P} \tag{5}$$

$$G = g_0 - g_1 = (e_0 + e_1)\bar{P} = e_{01}\bar{P} \tag{6}$$

The resulting model is presented in a compact matrix form. To identify distortions in the measured dimensions, it is necessary to expand it and derive expressions for all coordinate projections of the vectors $\overline{P}$ and g.

Naturally, for the desired vector g, only the notations of its coordinate projections can be written: gz, gx, gy.

The situation is more complex for the given vector $\overline{P}$, which defines the force loading of the technological system. As mentioned earlier, the control of the tool's movement is two-parameter: the coordinate feeds Sz and Sx are set so that, when executed simultaneously, the tool tip moves along the vector $\overline{S}$, i.e., along the formed contour (line D1). All known formulas for cutting forces (for example $p_i = c_i t^{x_i} s^{y_i} v^{z_i}$ $i = x; y; z$ [12] or formulas in [20-21]) define the components of cutting forces in a coordinate system associated with the formed surface: the X-axis is directed normal to the formed surface, the Z-axis follows the feed vector of the tool, i.e., along the formed surface, and the Y-axis is normal to the ZX plane. It should be noted that, according to ISO 841–74 and GOST 23597–79 standards, the information about the Z, X, and Y axes provided in the article should be interpreted for CNC machines as follows: Z ⇒X, X ⇒Y, and Y ⇒ Z. The point is that on CNC machines, the Z-axis runs along the spindle axis, while the transverse movement of the tool is along the X-axis. Accordingly, the feeds are labeled in the same way. When the tool is oriented normal to the formed surface, in single-tool, single-coordinate setups, the tool's coordinate system aligns with this

kinematic coordinate system, and applying the formulas of cutting theory presents no difficulties [6, 20-21]. In our case, the misalignment of these systems results in the force vector applied to subsystem 0 taking the following form:

$$\bar{P} = \begin{pmatrix} P_x\, cos\theta - P_z\ sin\theta \\ P_x\ \sin\theta + P_z\ \cos\theta \\ P_y \end{pmatrix} \tag{7}$$

Taking into account the introduced notations and the transformations performed, the general model (1) in its expanded form will be represented as follows:

$$\begin{pmatrix} g_z \\ g_x \\ g_y \end{pmatrix} = \begin{pmatrix} e_{zz}^{01} & e_{zx}^{01} & e_{zy}^{01} \\ e_{xz}^{01} & e_{xx}^{01} & e_{xy}^{01} \\ e_{yz}^{01} & e_{yz}^{01} & e_{yy}^{01} \end{pmatrix} = \begin{pmatrix} c_y t^{x_y} \left(\frac{s_z}{\sin\theta}\right)^{y_y} v^{z_y} \cos\theta - c_x t^{x_x} \left(\frac{s_z}{\sin\theta}\right)^{y_x} v^{z_x} \sin\theta \\ c_y t^{x_y} \left(\frac{s_z}{\sin\theta}\right)^{y_y} v^{z_y} \sin\theta + c_x t^{x_x} \left(\frac{s_z}{\sin\theta}\right)^{y_x} v^{z_x} \cos\theta \\ c_y t^{x_z} \left(\frac{s_z}{\sin\theta}\right)^{y_z} v^{z_z} \end{pmatrix} \tag{5}$$

Here, the contour feed s is defined through the coordinate feeds s_z and s_x. Since these feeds are functionally related (ensuring the tool follows the specified trajectory), s_z is more often taken as the independent variable. As a result, the contour feed is given as:

$$s = \frac{s_z}{\sin\theta} \tag{6}$$

It should also be noted that the depth of cut t in formula (5) is understood as the actual depth tf, which differs from the calculated depth. The calculated depth of cut tcalc in Figure 2 is the distance between lines G and D. The actual depth of cut tact in this scheme is the distance between lines G1 and D0. The relationship between these quantities is given by the formula:

$$t_{act} = t_{calc} - (g_z sin\theta + g_x cos\theta) \tag{7}$$

As shown in the works of A.A. Koshin [22] and V.I. Guzeyev [23], the stiffness of modern lathe machines is quite high, and the force levels during operation at standard cutting modes are such that the elastic deformations of the components of the technological system are much smaller than the specified depths of cut. Therefore, the terms in parentheses in formula (7) can often be neglected.

The derived formula (5) describes the coordinate components of the vector of the total displacement of the contact point. These components in the previously discussed setups directly determined the distortions in the measured dimensions in the specified directions [3-4, 13-16, 19]. In a two-coordinate setup, the scheme for determining the distortion of the performed dimension becomes more complex. The point is that the actual radius of the part Rf, due to the displacement of the workpiece and the tool along the Z-axis, no longer relates to the calculated cross-section of the part, but rather to a cross-section that is shifted from the calculated one by the amount gz. In this cross-section, the part should have a different radius, and the distortion in dimension must be measured from it. Therefore, an additional formula applies for the distortion of the measured diametrical dimension in a two-coordinate setup:

$$\Delta R = g_x + g_z \operatorname{tg}\theta \tag{8}$$

Thus, the overall analytical matrix model of the force interaction of a two-body system with elastic connections (1) for two-coordinate setups in cone machining is transformed into the matrix equation (5) and the additional relationship (8).

IV. Conclusion

1. Matrix models of machining errors have been developed for single-tool setups on a rotating carriage and for single-tool two-coordinate setups, taking into account the combined effects of cutting forces and elastic deformations of the technological system in all coordinate directions.
2. The resulting matrix models of machining accuracy for conical surfaces reflect dimensional distortions depending on cutting conditions, cone angle, and the comprehensive compliance characteristics of the technological system. These models can also be used to calculate setup dimensions.
3. The use of the developed accuracy models enhances the potential for automated design of operations on automatic lathes.

V. Financing

This work was supported by the Azerbaijan Sciense Foundation - **Grant № AEF – MGC – 2024 - 2(50) – 16/01/1 – M – 01**

References

[1] Sazonova Natalia Stalyevna. Increasing productivity and accuracy of processing on vertical multi-spindle turning machines by controlling technological parameters: Abstract of a dissertation for the degree of candidate of technical sciences, Specialty 05.02.08 - "Mechanical Engineering Technology", Chelyabinsk 1996, 21 p.

[2] Feldshteyn, E.E. Processing of parts on CNC machines: [a textbook for universities in mechanical engineering specialties] / E.E. Feldshteyn , Mikhail Anisimovich Kornievich . - 3rd ed., suppl. - M.; Mn.: New knowledge, 2008. - 298 p., ISBN 978-985-475-280-8

[3] Yusubov N.D. Increasing the efficiency of multi-tool processing on automatic lathes of the turning group: Abstract dissertation for the degree of doctor of technical sciences, Baku, 2009, 45 p.

[4] Yusubov N. Multi-tool machining on automatic lathes (Ma trix theory of multi-tool machining accuracy on modern CNC la thes). AV Akademikerverlag / LAP LAMBERT, Saarbrücken, 2013, 256 p.

[5] N. M. Rasulov, U. M. Nadirov, M. Z. Alakbarov. IMPROVING THE EFFICIENCY OF GRINDING TEETH BY COPYING WITH THE CONTROL OF DYNAMIC TECHNOLOGICAL CONNECTIONS. SOCAR Proceedings Special Issue 1 (2022) 029-035

[6] N. M. Rasulov, G. V. Damirova, I. A. Abbasova, Y. E. Huseynov IMPROVING THE EFFICIENCY OF THREAD ROLLING WITH MANAGEMENT OF TECHNOLOGICAL CONNECTIONS. SOCAR Proceedings Special Issue 1 (2022) 011-015

[7] Korovin V. A., Leushin I.O. Balabanov I.P. Savin I.A. Increase of resistance of steel moulds using the complex modifier INSTEEL-7. CIS Iron and Steel Review. – 2024. – Vol. 27. – P. 31-34. – DOI 10.17580/cisisr.2024.01.05

[8] Leushin, I.O., Leushina, L.I. Balabanov, I.P. Savin, I.A. Production of moulding cores and waterglass mixtures using “dry ice” for steel and iron casting. CIS Iron and Steel Review, 2021, 21, p. 34–37 DOI:10.17580/cisisr.2021.01.05

[9] Aziz, S. (2024). Characteristics of Quality Formation in Processing of 32CrMoV12-10 Steel by Honing. In Key Engineering Materials (Vol. 979, pp. 39–45). Trans Tech Publications, Ltd. https://doi.org/10.4028/p-8bwr6u

[10] Balabanov, I. P., Balabanova, Y., & Agayev, A. (2024). Development of a Parametric Model for Calculating Cutting Forces in External Cylindrical Turning of 16MNCR5 Steel. In Key Engineering Materials (Vol. 979, pp. 11–18). Trans Tech Publications, Ltd. https://doi.org/10.4028/p-

4xsctu

[11] Modeling of distortion of performed dimensions in singletool machining from rotary carriage / N. D. Yusubov, S. A. Bogatenkov, N. S. Sazonova [etc..] // Automated design in mechanical engineering. – 2023. – № 15. – pp. 42-46. – DOI 10.26160/2309-8864-2023-15-42-46. – EDN ETKOGA.

[12] Dal'skii A.M., Kosilova A.G., Mesheryakov R.K., Suslov A.G., eds. Handbook of a mechanical engineer. In 2 vol. Vol. 2. Moscow, Mashinostroenie Publ., 2003. 944 p. ISBN 5-94275-013-9.

[13] Yusubov, N., Abbasova, H. & Khankishiyev, İ. Entwicklung einer Projektierungstheorie für die Mehrwerkzeugbearbeitung mit den Möglichkeiten der modernen CNC Werkzeugmaschinen. Forsch Ingenieurwes 85, 661–678 (2021). https://doi.org/10.1007/s10010-021-00478-7

[14] N. Yusubov and H. Abbasova, "Models for Machining Accuracy in Multi-Tool Adjustment", Int. J. Automot. Mech. Eng., vol. 17, no. 3, pp. 8067–8085, Oct. 2020.

[15] Yusubov N.D., Abbasova H.M. Full-factor matrix model of accuracy of dimensions performed on multi-purpose CNC machines. Obrabotka metallov (tekhnologiya, oborudovanie, instrumenty) = Metal Working and Material Science, 2021, vol. 23, no. 4, pp. 6–20. DOI: 10.17212/1994-6309-2021-23.4-6-20. (In Russian).

[16] Yusubov, N.D. Matrix models of Machining errors in multi-tool multi-carriage adjustments / N.D. Yusubov, İ.A. Khankishiev, H.M. Abbasova, E.D. Mammadov, R.A. Huseynov // İnternational Journal on "Technical and Physical Problems of Engineering" (İJTPE),- September 2023, Issue 56, Volume 15, Number 3- p. 309-315.

[17] Bogatenkov, S.A., Sazonova, N.S., Guzeev, V.I. et al. Increasing the Productivity of Multitool Machining on Automated Lathes by Optimizing the Tool Positions. Russ. Engin. Res. 41, 1075–1079 (2021). https://doi.org/10.3103/S1068798X21110058

[18] Bogatenkov, S.A., Sazonova, N.S., Yusubov, N.D. et al. Increasing the Productivity of Multitool Machining on Automated Lathes by Optimizing the Machining Plan. Russ. Engin. Res. 41, 1071–1074 (2021). https://doi.org/10.3103/S1068798X21110046

[19] Yusubov, N.D. Matrix models of the accuracy in multitool two-support setup. Russ. Engin. Res. 29, 268–271 (2009). https://doi.org/10.3103/S1068798X09030125

[20] Koshin A.A. Processing on lathes: adjustment, cutting conditions. Handbook, Chelyabinsk, Siti-Print, 2012. 744 p.

[21] Yusubov N, Abbasova H. Models of Cutting Forces in the Matrix Theory of Multitool Machining Accuracy. KEM 2024; 979: 27–38. https://doi.org/10.4028/p-bw48sb.

[22] A. A Koshin. Precision theory and optimization of multi-tool turning. Abstract of a dissertation for the degree of Doctor of Technical Sciences, Specialty 05.02.08 - "Mechanical Engineering Technology", Chelyabinsk – 1997, 290 p.

[23] V.I. Guzeev. Theory and method of calculating the productivity of contour processing of parts of various accuracy on CNC turning and milling machines: Abstract dissertation for the degree of doctor of technical sciences, Chelyabinsk 1994, 517p.

INVESTIGATION OF THE POSSIBILITY OF CONTROLLING THE DEFORMATION OF THE CENTER AXIS OF A WORKPIECE PROCESSED BY TURNING DUE TO THE CUTTING FORCE WITH A DIGITAL PROGRAM

Agasi Agayev[1]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan
agasig@aztu.edu.az

Abstract

The article considers the issue of compensating the axial deformation caused by the cutting force during the cantilevered machining of rod-shaped and thin-walled parts processed by lathe operation with a digital program. The deformation of the axis of the part due to the cutting force causes a change in the cutting depth, as a result, the machining accuracy is not ensured, and after machining, shape errors appear in the cross section and also in the longitudinal section. This deformation takes a maximum value at the end of the cantilevered part, and as a result, the shape of the part is obtained according to the hyperbolic curve as it approaches the end. As a solution to this, the article proposes to use an equidistant to compensate for this deformation. The contour of this equidistant was taken into account in the digital program and the deformation of the center axis of the cantilevered parts due to the cutting force during the machining of experimental parts was minimized. The application of the equidistant in the numerical control program was performed in the turning operation of cantilever-mounted rod-shaped parts with a diameter of 10 mm and hollow parts with a diameter of 18 mm.

Keywords: turning operation, cutting force, cantilevered part, thin-walled part, center axis deformation, equidistant, numerical control, machining accuracy

I. Introduction

The introduction of CNC machines in agile manufacturing increases the need for machining accuracy control. The rigid control of accuracy by software gradually leads to the need to switch to software control of mechanical machining accuracy as a result of the improvement of modern numerical control devices [1].

In this regard, it is important to investigate the factors affecting the accuracy of mechanical processing, to study the causes of errors and the laws of variation. This is especially characteristic of cantilevered parts with low rigidity.

It is a very important and very urgent issue to study the law of variation of errors caused by the effect of cutting force during the turning of rod-shaped and thin-walled cantilevered parts, to determine the equidistant of the cutting tool's motion trajectory according to the law of variation of error, and to study the control of the resulting error with a numerical control program.

The main goal of the article is to investigate the possibilities of programmatically reducing the elastic deformation of the spindle axis caused by the radial component of the cutting force (P_y), as one of the factors affecting the machining accuracy of rod-shaped and thin-walled cylindrical parts on CNC lathes, and to determine ways to reduce errors caused by the influence of this factor.

II. Issues to be investigated

The study of the effect of deformations arising in the technological system under the influence of cutting force on the machining accuracy and its reduction is of great interest as a research object. Research in this area has been carried out mainly with the help of adaptive and rigid control systems [2]. The study of the possibility of program control of this problem in numerically controlled machine tools is of both theoretical and experimental importance, especially for parts with low rigidity.

In this regard, the article proposes to solve the following research questions in order to study the possibility of software control of errors arising in the technological system due to the effect of cutting force during the processing of hollow cylindrical parts with low hardness on numerically controlled lathes:

1.Investigation of errors caused by the effect of cutting force during the turning of cantilevered rod-shaped and thin-walled cylindrical parts;

2. Investigation of the possibility of controlling machining errors caused by the effect of cutting force in cantilevered rod-shaped and thin-walled parts using software on numerically controlled machine tools;

3. Designing variants of the compensating equidistant control program to compensate for the axial deformation of the part due to the effect of the cutting force;

4. Experimental study of software control of center axis deformation under shear force in cantilevered rod-shaped and thin-walled parts.

III. Errors caused by the effect of cutting force

One of the factors affecting the accuracy during the mechanical processing of thin-walled cylindrical parts is the displacements in the technological system caused by the action of cutting forces. The displacements caused by the cutting forces in the technological system cause errors in the longitudinal and transverse sections of the part and the violation of the shape accuracy [3]. Therefore, the investigation of the errors caused by the action of cutting forces in the longitudinal and transverse sections of the part is of great importance.

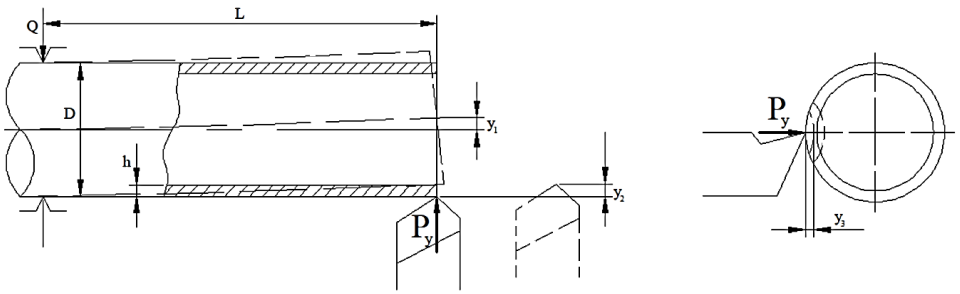


Figure 1: *Deformations caused by cutting force in the technological system*

As we know [3], in the simplest machining scheme, the deformations caused by the cutting force (P_y) in the technological system (figure 1) have the following designations: displacement of

the axis of the workpiece - y_1, displacement of the cutting tool under the influence of the cutting force - y_2, and radial displacement in the cross section under the influence of the cutting force in thin-walled workpieces - y_3. All these deformations mentioned cause a change in the intended cutting depth, as a result of which the machining accuracy is not ensured, and after machining, shape errors appear in the cross section as well as in the longitudinal section [4].

İt is known that the deformation under the influence of the cutting force at the end of the cantilevered part takes on a maximum value and, as a result, the resulting part corresponds to a hyperbolic curve that increases towards the end [5]. In this regard, it is important to theoretically and experimentally study the law of change of the axial deformation of the workpiece. Studying the law of change of the deformation of the central axis of the cantilevered part will allow to design a program of the movement trajectory of the cutting tool in numerically controlled machines corresponding to the equidistant, which can take into account the value of this deformation.

Let us examine the displacements caused by the action of cutting force in cylindrical parts fixed to a thin-walled cantilever with low stiffness. The deformation y_1 that can occur under the action of the cutting force P_y, which is the sum of the shear forces, varies according to expression (1):

$$y_1 = \frac{P_y \cdot L^3}{3EJ} \tag{1}$$

where: P_y is the sum of the cutting force normal to the axis of the beam and is theoretically calculated using expression (2):

$$P_y = C_{P_y} \cdot t^x \cdot S^y \cdot V^n \cdot K_y \tag{2}$$

L- length of the cantilever part of the workpiece;
E- modulus of elasticity of the material, for steel $E=2{,}1 \cdot 10^6$ kq/sm^2;
J- is the moment of inertia.
For hollow workpieces:

$$J = \frac{\pi}{64}(D^4 - d^4)$$

For stick workpieces:

$$J = \frac{\pi}{64}D^4$$

where, D and d are the outer and inner diameters of the workpiece, respectively.

IV. Designing a control program for the required equidistant

One of the initial stages in designing an equidistant control program is the selection of a program control tool. A complex and fundamental solution to this problem can be achieved by developing an intelligent, adaptive, or logical control system [6].

Although intellectual control systems allow to solve the problem completely, their application in solving simple problems does not justify itself. That is, it is necessary to put the problem in a very fundamental form. Here it is necessary to create such control algorithms that after processing the information coming from the transmitters, the control system should be able to make a decision

on which control algorithm to choose. Since modern digital software control devices are equipped with various transmitters, this will not create such a problem. However, various problems may arise in the means of obtaining the required information directly or indirectly based on electromechanical, etc. transformations. However, the control tool can also be solved through the control program of machines controlled by digital software. Two options are possible for this.

In the first variant, the direction of the solution is determined by the development of new types of interpolation algorithms. It is known that in modern digital program control systems, linear-circular interpolation algorithms are fully implemented. In order to construct an equidistant, theoretically any curve can be divided into elementary parts such as line-arc intersections. However, these elementary parts are not always equivalent to the exact line-arc intersection, they are replaced only within a certain error. On the other hand, sometimes the radius of curvature of the arc, even when it has very large values, does not correspond to the technical indicators of the existing linear-circular interpolation [7]. Therefore, the selection and development of new types of interpolation methods can be considered relevant in this sense. However, it should also be noted that in this case the software support system of the digital program control system must be changed.

The second option is a simpler, but relatively less accurate solution. The essence of this method is to replace the required equidistant with an equivalent equidistant that can reduce errors, using the technological and auxiliary commands provided by the existing control device and control program.

When building a logical control system, the main principle is to develop control algorithms for a specific situation based on the logical analysis of feedback signals received from the transmitters. The programmable controllers required for this control system are available in most CNC (Computer Numerical Control) machine tools. When building a logical control system, first of all, input-output programs are built, logic equations are compiled and minimized, and then converted into a programming language. The construction of mathematical equations is done with the help of "AND", "OR", "NEGATIVE" and other schemes.

For example:

IF L=50 mm; AND (X_1)

D=12 mm; AND (X_2)

t=1 mm if, THEN (X_3)

ON THE OTHER HAND S=0,3 mm/rev.

The text of the program is $\bar{y} = X_1 \cdot X_2 \cdot X_3$ or $y = \overline{X}_1 \cdot \overline{X}_2 \cdot \overline{X}_3$ will be in the form of.

During processing, the deformation due to the shear force at the end of the cantilevered part takes on a maximum value, and as a result, the shape of the part is taken to correspond to a hyperbolic curve as it approaches the end. Taking this into account, the scheme of the shape error resulting from processing and the compensating equidistant can be shown as follows:

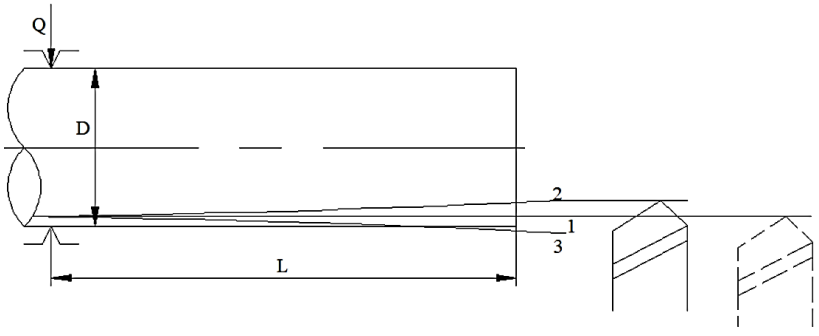


Figure 2: *Diagram of the error and the compensating equidistant during the processing of a cantilevered part*

In this figure, 1 is the desired surface, 2 is the real surface, and 3 is the compensating equidistant. This equidistant is obtained graphically by subtracting curves 2 and 1. As can be seen from figure 2, curves 2 and 3 are symmetrical with respect to surface 1 and will differ in the direction of the center.

As we know [7], three typical joining methods are used to join the support points of an equidistant: straight line-straight line, straight line-circle and circle-circle method. Since it is not possible to program an arbitrary curve directly as a hyperbolic curve in current numerically controlled machine tools, it is considered appropriate to program the equidistant as a line-circle junction [8].

To program the equidistant compensation for the error that occurs during the mechanical processing of the cantilevered part shown in figure 2, the part is first divided into five equal parts along the length. Then, a scheme for programming the equidistant is constructed. This scheme is shown in figure 3.

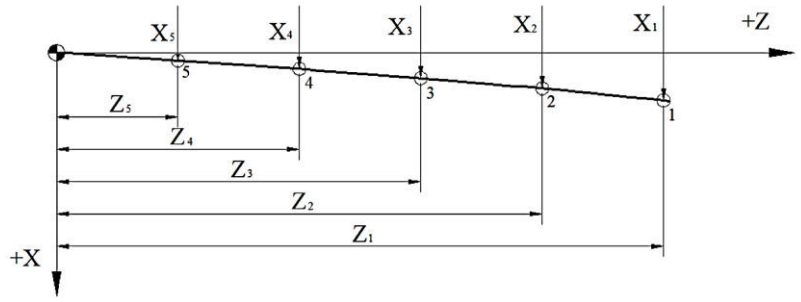


Figure 3: *Schematic of programming the compensating equidistant*

In this figure, the values of the coordinates on the Z axis are denoted by Z_1, Z_2, Z_3, Z_4, Z_5, and the values of the coordinates on the X axis are denoted by X_1, X_2, X_3, X_4, X_5. Three new methods can be proposed for programming this compensating equidistance. In order to conduct a comparative analysis of these three methods that we have proposed, it is necessary to refer to the text of the programs based on them.

In the first method, the most distorted part of the equidistant 0÷5 is approximated by chords that are broken line segments, and the processing share for each broken line is assigned in separate frames. Here, choosing the number of broken lines is an optimization issue. Thus, by increasing the number of these chords, the approximation accuracy of the equidistant will increase. On the other hand, increasing the number of chords will also lead to an increase in the number of intermediate support points arranged along the equidistant, which will reduce the cleanliness of the surface processed along the equidistant. The text of the program based on this method will be as follows:

N_i X0 Z0 E
N_{i+1} X1 Z1 F
N_{i+2} X2 Z2 F
N_{i+3} X3 Z3 F
N_{i+4} X4 Z4 F
N_{i+5} X5 Z5 F

This method is considered a relatively simple method, but the assignment of absolute coordinates of points limits its application, because in other details, it is enough to change the dimensions alone to make this program useless.

The second method is more universal than the first. Since, in the program text compiled by this method, displacements are assigned instead of coordinates. That is, regardless of the dimensions of the detail, such displacements can be programmed as many as the number of chords determined during the approximation. However, in this method, both the points on the X coordinate are still assigned with absolute coordinates, and the relatively large number of intermediate points at relatively large lengths will lead to an increase in the number of frames. The text of the program created based on this method will be as follows:

N_i X0 Z0 E

N_{i+1} W-A X1 U-a
N_{i+2} W-A X2 U-a
N_{i+3} W-A X3 U-a
N_{i+4} W-A X4 U-a
N_{i+5} W-A X5 U-a
where, W- axial displacement along the Z axis;
A- equal length of relative displacements along the Z axis;
a- the equal length of the relative displacements along the X-axis is.
The values of A and a are determined as follows:

$$A = \frac{Z_0 - Z_4}{n}$$
$$a = \frac{X_0 - X_4}{n}$$

It is known that in numerically controlled machines, the numerical time consists of processing and preparation-completion times. One of the processes that determines the preparation-completion time norm is the process of setting up the machine, which, together with other operations, includes the compilation of the control program. Therefore, minimizing the time spent on the compilation of the control program as much as possible is one of the main factors affecting the increase in processing productivity.

Herefore, in the third method, we use the commands for programming cycles. Here, both the program text is reduced many times, and relative coordinates or displacements are assigned to both coordinates, which makes the program text as universal as possible. The text of the program created based on this method will be as follows:

N_i X0 Z0 E
N_{i+1} W-A U-a
N_{i+2} M18
N_{i+3} L11 B N_{i+1} HK
N_{i+4} X0 Z0 E
where, L11- recycling cycle;
B- the frame at the beginning of the repeat;
K=n-1 is the number of repetitions.

The beginning of the repeated program part is the frame N_{i+1} specified at address B, and the end is the frame with M18.

Thus, the first proposed method is suitable for a specific part, the second method is suitable for all parts with the same diametrical dimensions, and the third method is suitable for typical parts with different diametrical and longitudinal dimensions. However, in any case, the optimization problem related to the number of intermediate support points remains unresolved.

V. Experimental study of software control of the deformation of the center axis due to the cutting force of a cantilevered part

To implement the problem, first the value of the radial component of the shear force (Py) was determined using expression (2), and the values of the displacements of the central axis for the rod and thin-walled rods in five equal sections along the length were calculated and recorded in table 1 (for rod parts) and table 2 (for thin-walled cylindrical parts). Then, according to these values, the dependence graphs y=f(L) were constructed for the rod parts (figure 4) and for the thin-walled cylindrical parts (figure 5).

Table 1: *Displacement of the axis of the rod-shaped workpieces*

Current cross-sections of the span of length L (x), mm	Axis displacement-y, mm J=0,05sm⁴ D=10mm				Cutting force Py, N	t, mm	S, mm/rev	n, rev/min
	Theoretical	Experimental	Without program	With the program				
10	0,0063	0,0055	0,0032	0,0022	198	0,5	0,3	500
20	0,0051	0,0043	0,0021	0,0012				
30	0,017	0,014	0,008	0,004				
40	0,04	0,035	0,016	0,009				
50	0,078	0,067	0,03	0,02				

Then, the displacements of the axis in five equal sections of the beam were measured experimentally and the values of these displacements were added to table 1 (for rod parts) and table 2 (for thin-walled cylindrical parts).

Table 2: *Displacement of the axis of a thin-walled workpieces*

Current cross-sections of the span of length L (x), mm	Axis displacement-y, mm J=0,41sm⁴ D=18mm d=15mm				Cutting force Py, N	t, mm	S, mm/rev	n, rev/min
	Theoretical	Experimental	Without program	With the program				
10	0,00007	0,00006	0,00003	0,00001	198	0,5	0,3	500
20	0,00061	0,00053	0,00027	0,00015				
30	0,0021	0,0017	0,0012	0,0007				
40	0,0049	0,0042	0,0028	0,0015				
50	0,0095	0,0083	0,006	0,003				

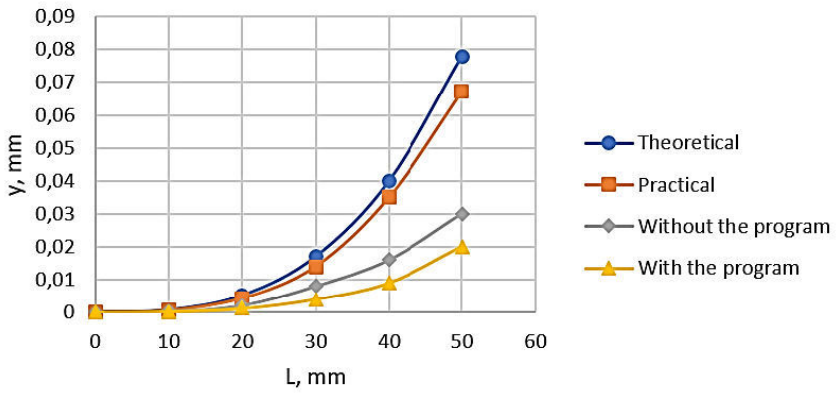


Figure 4: Displacement of the axes of rod workpieces (d=10 mm) depending on the length

Then, the workpieces were machined without a program in the machining modes t=0.5 mm, S=0.3 mm/cycle and n=500 cycles/min, the displacement values were measured and added to

Tables 1 and 2. The curve constructed based on these values was added to figures 4 and 5, respectively.

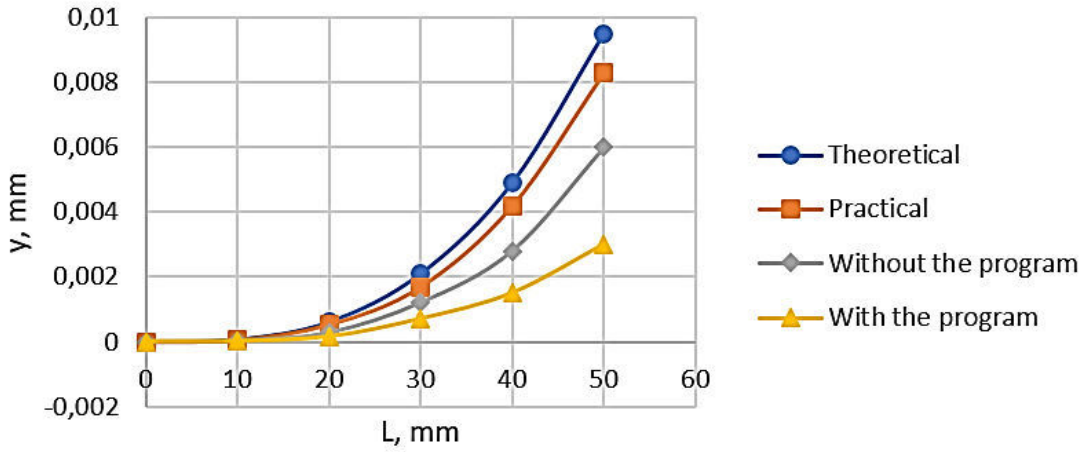


Figure 5: *Displacement of the axes of thin walled workpieces (D=18 mm, d=14 mm) depending on the length*

As a final result, the processing of rods and thin-walled plates was carried out using the developed control program in the same processing modes, and the results were added to the tables and figures. In all these experiments, the range of cutting force variation was 157...256 N.

VI. Results

1. Elastic displacements caused by the action of cutting force during the turning of cantilevered non-rigid parts have been studied theoretically and experimentally;

2. The possibility of software control of machining errors caused by the effect of cutting force in cantilevered rod-shaped and thin-walled parts on CNC machines has been investigated;

3. An equidistant was determined to compensate for the error caused by the bending of the axis due to the shear force of the cantilevered parts, a corresponding program was created and experimentally verified;

4. When comparing graphs constructed with known dependencies for determining the deflection of the center axis of cantilevered parts, it was found that the error during processing with and without programming the equidistant is reduced by 12-15%.

References

[1] Yusubov N, Abbasova H, Khankışhıyev İ. “Development of a planning theory for multi-tool machining with the possibilities of modern cnc machine tools”, Forschung ım ıngenıeurwesen, 2021, 85, №2, pp 661-678.

[2] Yusubov N, Abbasova H. “Models of Cutting Forces in the Matrix Theory of Multitool Machining Accuracy”. KEM 2024;979:27–38. https://doi.org/10.4028/p-bw48sb.

[3] Yusubov N, Abbasova H, Dadashov R. “Theoretical basis for the development of an algorithmic unified complex of mathematical models of cutting forces”. Machine Science 2023, №1. pp 55-60.

[4] Gafarov, A.M., Gafarov, V.A., Aziz, S.S., F.M. Kelbiev. “Surface roughness in lapping with dosed removal of surface layer”. Russ. Engin. 2011, Res. 31, 1147–1150.

[5] Sarvan Aziz "Characterıstics of quality formation in processing of 32CrMoV12-10 steel by honing", Key Engineering Materials (Volume 979), Pp 39-45.

[6] İbragimova N.A., İbragimov Z.Z. "Analysis of the programming stage to determine errors in the process of processing parts with numerical control", Enigma, 2020, №25, Pp 137-142.

[7] İ.V.Epifanov., L.V.Vinogradov. "The equidistant of fraction rational parabola". RUDN Bulletin, Engineering Research Series, 2012, No. 2, Pp 50-56.

[8] Valishin A.A., Tumanov I.A., Akhund-zade M.R. "Computer construction of a network of equidistant complex non-smooth curves on the ground". Mathematical modeling and numerical methods. 2020. No. 2. pp. 95–106.

STUDY OF MECHANICAL PROPERTIES OF THE COMPOSITION BASED ON LOW-DENSITY POLYETHYLENE MODIFIED WITH CARBON NANOTUBES

Matanat Mehrabova[1], Elbay Babayev[2], Farhad Kerimov[1], Musa Asadov[1], Niyazi Hasanov[3], Fail Shamilov[2], Yashar Musayev[4]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan
[2]State Oil Company of the Azerbaijan Republic, Baku, Azerbaijan
[3]Baku State University, Baku, Azerbaijan
[4]BalticNet-PlasmaNet, Nuremberg, Germany
metanet.mehrabova@aztu.edu.az, elbey.babayev@socar.az, ferhad.kerimov@aztu.edu.az, musa.asadov1@gmail.com, n.h.hasanov@rambler.ru, failshamilov@gmail.com, mussa_98@hotmail.de

Abstract

The composition and concentration of carbon nanotubes leading to a significant improvement in the electro-physical properties of low-density polyethylene (LDPE) of 10803-020 grade have been experimentally determined. Modified multi-wall carbon nanotubes (MWCN) were used as modifying additives. The difference of developed LDPE nanocomposite was that for the first time a small amount of carbon nanotubes was introduced into the LDPE, which varied within the range of 0.01–0.1 % wt. It was found that the content of 0.05 % wt. of carbon nanotubes in the composition of low-density polyethylene (LDPE + 0.05 % wt. of NC) significantly increased its electric strength.

The temperature dependences of dielectric loss tangent (tgσ), specific volumetric resistivity (ρV) and permittivity (ε), as well as the kinetics of physical and mechanical changes in them under the influence of electric discharges in the air and UV irradiation were studied. Specific volumetric resistivity significantly increases in the optimal LDPE nanocomposite and dielectric loss tangent is reduced, i.e. a good correlation is observed between these characteristics. It is shown that the addition of carbon nanotubes in an amount of 0.05 % wt. significantly increases its resistance to the effects of electric discharges and UV irradiation.

Based on the experiments, it was concluded that the observed improvement in the dielectric properties of the LDPE film after the introduction of 0.05 % wt. of carbon nanotubes into its composition is associated with a change in the supramolecular structure of the LDPE.

Keywords: LDPE, carbon nanotubes, permittivity, UV irradiation, specific volumetric resistivity

I. Introduction

The development of science and technology requires the creation of new materials, which is one of the most important areas of modern physics [1-6]. In recent decades, the use of various types

of nanomaterials has opened up new opportunities in many industries, with the oil industry being particularly noteworthy [7-10]. Polymer nanocomposites, due to their high characteristics, low weight, low cost and ease of processing, can be used as structural and insulating materials, as sealants, thermal insulation, binders in temperature-regulating coatings, seals, and reinforcing materials. Currently, polyolefin-based nanocomposites are used in the aviation and shipbuilding industries, as well as in mechanical engineering, energy, electronics, electrical engineering, radio engineering, transport, construction and other industries [11-13]. Various methods are known for creating compact, cheaper and light devices using polymer composites based on carbon nanotubes.

Reducing the filler size to the nano-range allows significantly decreasing its content in the polymer volume, and at the same time to achieve a significant improvement in operational characteristics due to the appearance of additional mechanisms of interaction of the polymer matrix with nanoparticles. Polyethylene differs from other thermoplastics by a very valuable set of properties.

Polyethylene products have high strength, resistance to aggressive environments and radiation, they are non-toxic, and have good mechanical and dielectric properties. Polyethylene is processed by all known methods for thermoplastics [14, 15].

It is known [16-18] that the effect of partial discharges in the air on polymers at room temperature is accompanied by the processes of oxidation, destruction and cross-linking of macromolecules, an increase in dielectric losses, a decrease in breakdown voltage, electrical and mechanical durability, as well as erosion of the polymer surface.

By using fillers based on various types of nanoparticles, it is possible to obtain polymers that combine traditional and new quality characteristics that at first glance seem mutually exclusive. This is extremely useful in cases when it is necessary to simultaneously provide transparency and flexibility of the material, a certain degree of impact resistance and rigidity, physical, insulating characteristics and conductivity.

The aim of this study is to develop a polymer nanocomposite based on LDPE with the additives of carbon nanotubes to obtain materials and products with the improved dielectric and mechanical properties, good heat resistance, as well as to minimize the rate of film ageing during operation.

II. Materials and experimental methods

Low-density polyethylene (LDPE) of 10803 grade was chosen as the object of the study, and modified multi-wall carbon nanotubes (MWCNT) were used as additives. Additives in the amount of 0.01-0.1% wt. were introduced into the original raw LDPE by mechanical melt-blending, which is widespread due to its obvious simplicity and ease of industrial development. To prepare films by hot pressing from blanks obtained in a casting machine, a manual electrically heated hydraulic press PG-60 was used. The calculated amount of MWCNTs was loaded into a polished flat mold in the form of 0.1-mm thick gasket. A fluoroplastic film (foil) was used to prevent adhesion. The mold was heated to 140-150°C. After reaching the required temperature, the mold was compressed under a pressure of 150 atm. After cooling the mold, the films were removed. Samples were made from the resulting 60-80 μm thick films to determine the physical, mechanical, electrical and optical properties.

The homogeneity of the film was determined by measuring the film thickness over the entire area. The film thickness was determined using an N3B-2 optical thickness gauge and a micrometer. The arithmetic mean of 10 measurements was taken as the sample thickness.

The results of multiple experimental studies of the mechanical, dielectric, optical characteristics and structural properties of the developed nanocomposite confirm the homogeneity of the MWCNT dispersion in the LDPE matrix.

Installation for the determination of the short-term electrical strength Eпр and lifetime τE, i.e. the time elapsed from the moment of application of an electric field to the sample until its breakdown, at different values of the voltage E was used.

The impact of electric discharges on polymer dielectrics was carried out in a test cell of asymmetric type. The cell is a system of metal electrodes, between which the studied polymer film was placed. A smooth stainless steel plate measuring 180x130 mm (the plate surface was nickel-plated on one side) was taken as the bottom electrode. The studied polymer film, before and after preliminary stretching according to the applied method, was tightly stretched onto this plate, which served as a grounded electrode. To obtain an air gap of constant thickness between the upper electrode and the polymer film, 1.5-mm thick glass spacers were placed along the edges. High voltage of industrial frequency from the AII-70 device was applied to the cell electrodes.

UV radiation was supplied by a DRSh-500 lamp. The DRSh-500 mercury-quartz lamp is a spherical flask made of quartz glass. The DRSh-500 lamp is a powerful, concentrated source of radiation in the visible and ultraviolet parts of the spectrum. The lamp operates in a limited volume (casing), provided that the dimensions of the casing and its ventilation conditions are such that the air temperature at a distance of 60 mm from the walls does not exceed 2500C

A 50 μm thick sample, attached to a duralumin frame with a holder, is fixed on a tripod. UV rays are directed to the center. The distance from the source to the sample is 250 mm. A distance of 250 - 200 mm increases the effect of UV irradiation (ageing).

UV rays hit the sample at a right angle. The experiment was conducted at room temperature (200C). The experiment mode is as follows: the lamp is clamped using an inductor with a sparking distance 15-20 mm. The nominal voltage on the lamp is 70 V, the current - 7.5 A. In this case, the resulting nominal luminous flux is 22500 lm. The duration of the sample' irradiation is 15 and 30 hours.

III. Results and discussion

Fig. 1 shows the dependence of the short-term electrical strength of LDPE on the content of MWCNTs: LDPE + 0.05 % wt. MWCNTs, LDPE + 0.07 % wt. MWCNTs and LDPE + 0.1% wt. MWCNTs at room temperature.

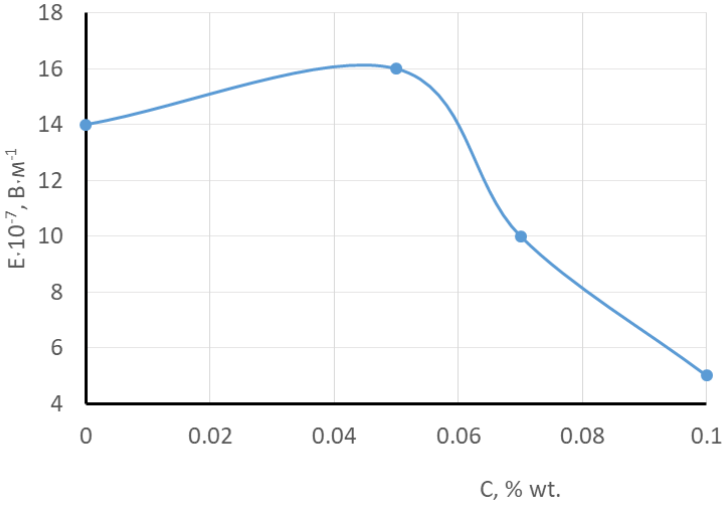


Figure 1: *Change in the electrical strength of the LDPE nanocomposite depending on the mass percentage of the MWCNT additive*

It is evident from fig. 1 that with an increase in the weight % of MWCNT, electrical strength of polymer increases, and the maximum value of *E*pr is achieved with the introduction of 0.05 % wt. of

MWCNT into the composition of LDPE. With a further increase in the MWCNT content to 0.1 weight %, a sharp decrease in electrical strength is observed. Experiments have shown that with the introduction of the optimal (0.05 % wt.) content of MWCNT into the composition of LDPE, its electrical strength increases from $14 \cdot 10^{-7}$ to $16 \cdot 10^{-7}$ $V \bullet m^{-1}$, i.e. by an order of 20%.

A good correlation is observed when studying the dependence of the specific volumetric resistivity (ρ_v), dielectric loss tangent (tgσ) and permittivity (ε) on the nanocarbon content at room temperature, which is shown in the table 1. Values of the electrophysical properties of LDPE and its nanocomposites are given in table 1.

It is evident from the table that ρ_v, tgσ and ε of LDPE are also sensitive to the introduction of the MWCNT additive, and the content of MWCNT - 0.05 % wt. is also optimal here, since it ensures the greatest stability of electrical properties in comparison with both the original LDPE and LDPE with other additive content.

Table 1

№	Materials	Properties			
		$E_d \cdot 10^{-7}$, $V \cdot m^{-1}$	tgδ f=kHs, t=20^{0}C	ρ_v (Ohm·m) t=20 ^{0}C	ε
1	LDPE-108 (original)	14	$4{,}5 \cdot 10^{-4}$	$1 \cdot 10^{15}$	2.4
2	LDPE-108+0.05% wt. of MWCN	16	$3 \cdot 10^{-4}$	$1 \cdot 10^{16}$	2.1
3	LDPE-108+0.07% wt. of MWCN	10	$4 \cdot 10^{-4}$	$1 \cdot 10^{12}$	2.3
4	LDPE-108+0.1 % wt. of MWCN	5	$5 \cdot 10^{-4}$	$1 \cdot 10^{8}$	2.6

Thus, the detected increase in the electrical strength and stability of the LDPE nanocomposite can be related to the structural features of the specified additive, which ensures dense packing of macromolecules during film formation [19, 20].

It is known that in order to determine the possibility of using a polymer as insulation and to obtain information about the thermal motion and molecular structure of the polymer, it is necessary to know not only the value of the dielectric characteristics (tgσ, ρ_v and ε), but also the nature and quantitative characteristics of their dependence on temperature, electric field, UV irradiation and other factors [21-23].

Figure 2 shows the temperature dependences of the dielectric loss tangent (tgσ) of LDPE (without additives) and nanocomposites on its base before and after electrical ageing under the influence of discharges in air.

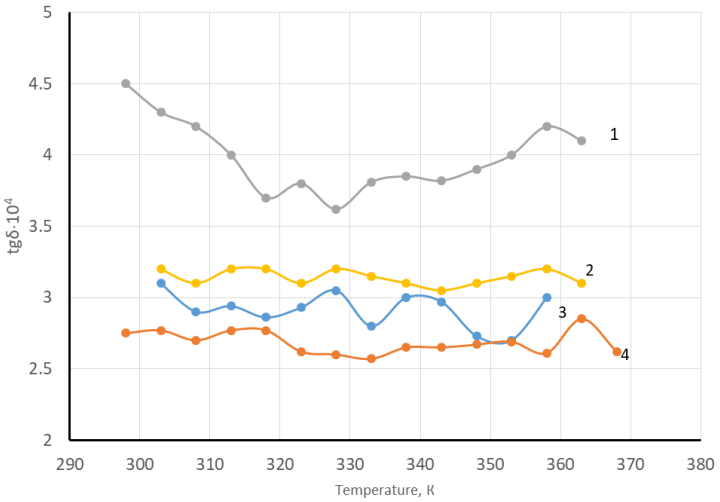


Figure 2: *Temperature dependences of the dielectric loss tangent of the initial LDPE film and its optimal nanocomposite before and after electrical ageing under the action of discharges in air at U_{age} = 7 kV and t_{age} = 5 hrs on the electric field strength: 1.2 - LDPE + 0.05 % wt. MWCNT before (1) and after (2) ageing; 3.4 - LDPE before (3) and after (4) ageing*

From the obtained experimental data (fig. 2) it is evident that in all temperature ranges of the test the introduction of the MWCNT additive in the amount of 0.05% wt. into LDPE leads to a decrease in the dielectric loss tangent (curve 1) comparing to LDPE without additive (curve 3). From the temperature dependence of the dielectric loss tangent it follows that under the action of discharges on LDPE films without the additive there is a significant increase in the maximum of dipole-segmental losses (in the region of 323-353K) (curve 4). At the same time, the introduction of the proposed MWCNT additive in the specified amount into LDPE prevents processes leading to its electrical ageing, and thus reduces the maximum of dielectric losses in this temperature range [curve 2].

It is known that under the influence of electric discharges in the air, oxidative destruction occurs [22]. As a rule, oxidative destruction in polymers is usually accompanied by the formation of polar groups in them, which is an independent movement of chain segments. The maximum value of tgσ is determined by the number of monomer links (segments) and polar groups formed in the polymer [24, 25].

In the temperature dependence of *tgσ* of the LDPE film and its nanocomposite after UV irradiation in air, the experimental data of which are shown in fig. 3, an insignificant change in the value of dielectric losses in the original LDPE film is observed (curve 2).

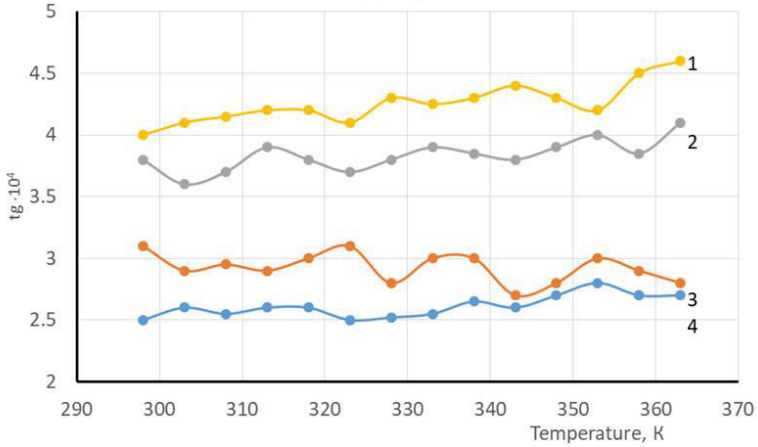


Figure 3: *Temperature dependences of the dielectric loss tangent of the LDPE film and its optimal nanocomposite before and after UV irradiation n air at t_{irr} = 15 hrs. 1, 2 - LDPE before (1) and after (2) irradiation; 3, 4 - LDPE + 0.05 % wt. of % NC before (3) and after (4) irradiation*

However, when introducing the optimal amount of nanocarbon into the LDPE composition, the value of *tgσ* after UV irradiation remains practically unchanged (curve 2). The experimental results allow us to assume that the addition of MWCNT reduces the change in *tgσ* of the LDPE film under the influence of electrical discharges and UV irradiation [26]

For real polymeric materials used as dielectrics for electrical and radio insulation, the specific volumetric resistivity depends on their composition, as well as on molecular structure and supramolecular structure of the polymers. Temperature, electric field, the effect of ionizing radiation and various additives also have a significant effect on the specific volumetric resistivity of polymeric dielectrics.

Figure 4 shows the results of study of the temperature dependence of the specific volumetric resistivity of the LDPE film and its optimal modification before and after electrical ageing under the influence of discharges in air.

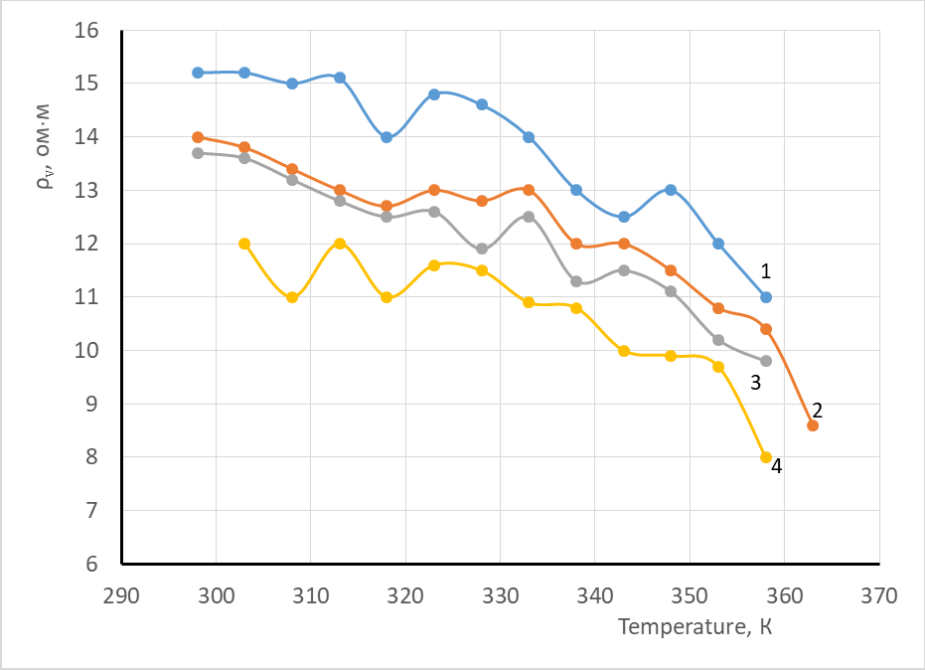


Figure 4: *Temperature dependence of logarithm of the specific volumetric resistivity of the LDPE film and its optimal modification before and after electrical ageing under the influence of discharges in air at V_{age} = 7 kV, t_{age} = 5 hrs. 1, 2 - LDPE + 0.05 % wt. of NC before (1) and after (2) ageing; 3, 4 – LDPE before (3) and after (4) ageing*

As can be seen from fig. 4, ϱ_v of the original sample of LDPE film decreases monotonously with temperature increase. After the introduction of the specified additive, ϱ_v value of the LDPE film increases (curve 1).

After the impact of electric discharges in air, the dependence $lg\ \varrho_v = f\ (T)$ undergoes significant changes for both the original LDPE and the nanocomposite. However, the addition of MWCNT in an optimal amount (0.05 % wt.) significantly slows down the intensity of their decrease during electrical ageing (curve 2) more than for the initial LDPE (curve 4).

Fig. 5 shows temperature dependence of the specific volumetric resistivity of the LDPE film and its optimal nanocomposite before and after UV irradiation.

As can be seen from Fig. 5, the specific volumetric resistivity ϱ_v of the modified LDPE after UV irradiation changes in the same way as before irradiation.

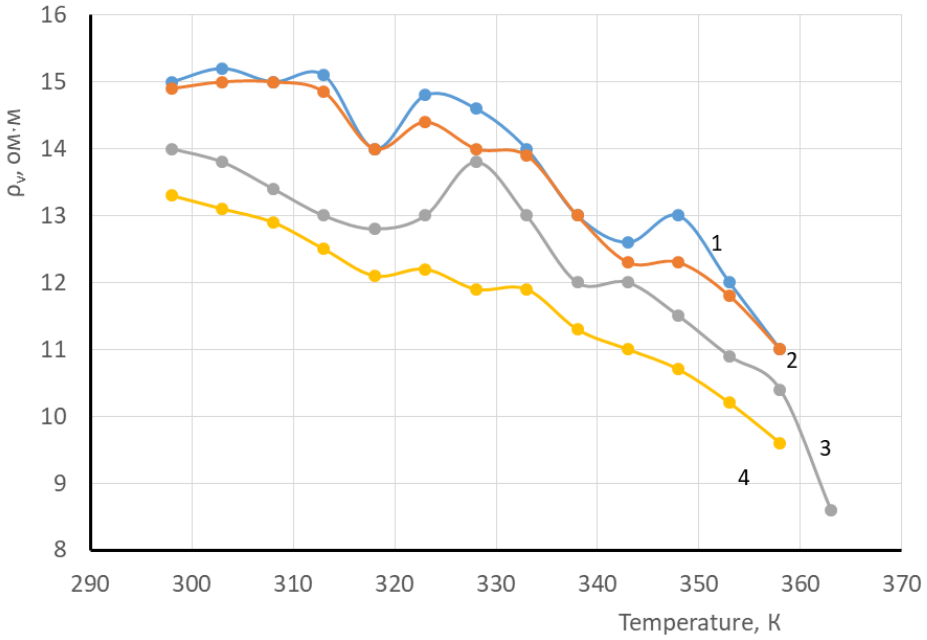


Figure 5: *Temperature dependence of logarithm of the specific volumetric resistivity of the LDPE film and its optimal modification before and after UV irradiation in air at* t_{irr} *= 15 hrs. 1, 2 - LDPE +0.05 % wt. of NC before (1) and after (2) of the irradiation; 3, 4 – LDPE before (3) and after (4) the irradiation*

However, as can be seen from Fig. 5, as well as from the dielectric loss tangent, with prolonged exposure to UV irradiation, the specific volumetric resistivity of the LDPE-based nanocomposite decreases slightly (curve 2), while for the LDPE without additives it decreases significantly (curve 4).

Probably, the temperature trend of electrical resistance and the dielectric loss tangent before and after electrical ageing and UV irradiation in the air, chemically active substances (ozone, nitrogen oxides, etc.) interacting with macromolecules cause the appearance of a number of new low-molecular groups and compounds, aldehydes, acids, alcohols, ethers, etc. According to [26], these groups can diffuse into the volume of the sample and thereby contribute to the improvement of the dielectric characteristics of the polymer material.

Thus, the improvement of ρ_v and $tg\sigma$ established by us before and after electrical ageing and UV irradiation with the introduction of the proposed MWCNT additive in the optimal amount is the result of formation of optimal structures in the polymer.

In the same samples, the temperature dependences of the dielectric permittivity (ε) were studied, the results of which are shown in Fig. 6.

The results of the studies showed that in the entire studied temperature range (283-373k), the dielectric permittivity of LDPE and its nanocomposite (LDPE + 0.05 % wt.) practically does not change with temperature change. From Fig. 6 it is evident that ε increases approximately equally in the entire studied temperature range, therefore, to clarify the dependence on the test time, it is sufficient to study the change in ε at only one temperature (300 K).

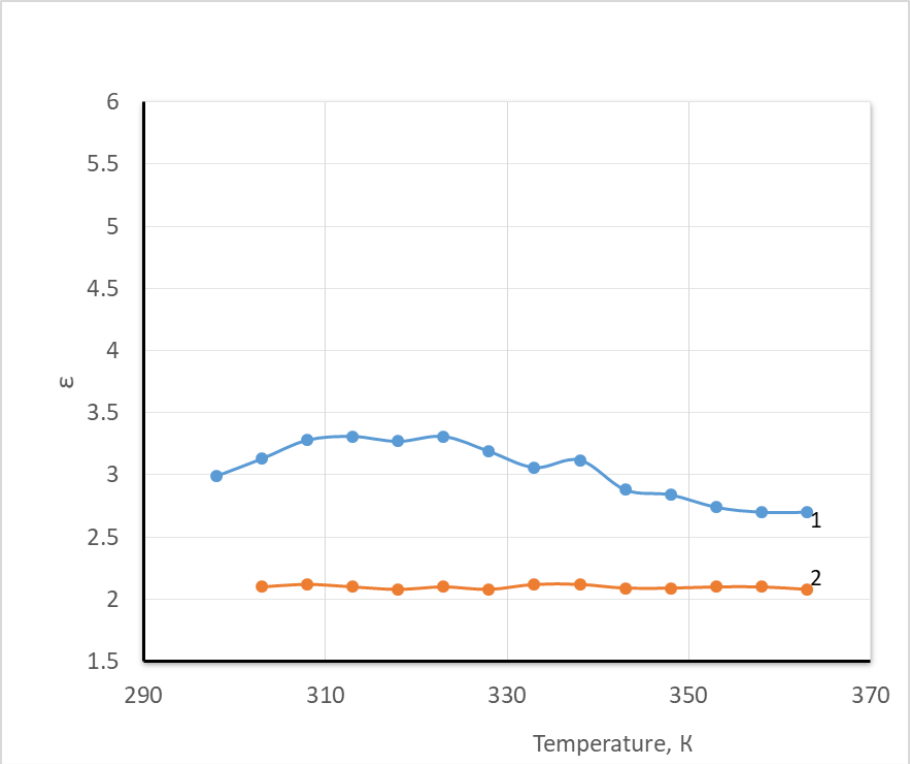


Figure 6: *Temperature dependence of dielectric permittivity* ***(ε)*** *of LDPE film and its optimal nanocomposite. 1 – PE (without additive), 2 - LDPE + 0.05 % wt. of MWCNT*

Thus, the detected increase in the dielectric properties and stability of LDPE can be associated with the structural features of said nano-additive, which ensures dense packing of macromolecules during film formation [26]. Dense packing of macromolecules in the structure of the developed nanocomposition leads to increased intermolecular interactions, which contribute to a significant slowdown in the rate of oxidation-destructive processes during electrical ageing and UV irradiation.

IV. Conclusion

The possibility of obtaining a nanocomposite for electrical insulation purposes based on the low-density polyethylene of grade 108 03 - 020 with MWCNT additives was established experimentally. The range of change in the filler's amount was experimentally substantiated, i.e. the optimal composition of the studied nanocomposite for improving operational characteristics was determined. The developed LDPE nanocomposite is distinguished by the fact that for the first time a significantly small amount of carbon nanotube additives was introduced into the LDPE composition, which varied within the range of 0.01-0.1 % wt. It was found that with an increase of MWCNT additives content in the LDPE composition, the electrical strength significantly decreases, and a positive effect is observed only at 0.05 % wt. of the said additives. It was established that the developed LDPE nanocomposite has a relatively increased resistance to the effects of electrical discharges and UV radiation, which opens up wide areas of application in various industries, especially in the oil industry.

References

[1] M.A. Mehrabova, H.S. Orujov, N.H. Hasanov, A.I. Kazimova, A.A. Abdullayeva. Ab initio calculations of defects in CdMnSe semimagnetic semiconductors. Mechanics of Solids, 2020, 55, p. 108-113, https://link.springer.com/article/10.3103/S0025654420010021

[2] I.R. Nuriyev, M.A. Mehrabova, A.M. Nazarov, R.M. Sadigov, N.G. Hasanov. On the growth, structure, and surface morphology of epitaxial CdTe films. Semiconductors. 2017, 51, p. 34-37, https://link.springer.com/article/10.1134/S1063782617010183

[3] I.R. Nuriyev, M.A. Mehrabova, A.M. Nazarov, N.H. Hasanov, R.M. Sadigov, S.S. Farzaliyev, N.V. Farajov. Structure and Surface Morphology of Cd1−x(Mn, Fe)xSe Epitaxial Films. 2019, 13, p. 1083-1085, https://link.springer.com/article/10.1134/S1027451019060168

[4] M.A. Mehrabova, H.R.Nuriyev, H.S. Orujov, N.H. Hasanov, A.A. Abdullayeva, Z.I. Suleymanov. Ab-initio calculations of electronic structure of CdFeTe and optical properties. Conference proceedings Modern Trends in Physics. Baku. 2019, p.1-3.
https://www.researchgate.net/profile/ElmiraKhanmamedova/publication/361788732_BSU_100/links/62c53617721b9c41cc329ee8/BSU-100.pdf

[5] M.A.Mehrabova A.O. Mekhrabov, Effect of gamma irradiation on electrical and photoelectrical properties of Cd1-xMnxTe thin films. Machine science, 2023, 2, p.70-77, http://dx.doi.org/10.61413/GDKV8772

[6] A.O. Mekhrabov, M.A.Mehrabova. Effect of gamma irradiation on optical properties of CdMnTe thin films. Scientific.net, Materials Science and Engineering. Online periodicals, Materials Science Forum, 2024, 1119, p.3-12, https://doi.org/10.4028/v-lTJp83

[7] V.M. Shamilov, E.R. Babayev, E.S. Kalbaliyeva, F.V.Shamilov Polymer nanocomposites for enhanced oil recovery// Materials Today: Proceedings. 4, 2017, S70–S74
https://www.sciencedirect.com/science/article/pii/S2214785317319028

[8] V.M. Shamilov, E.R. Babaev Polymer nanocomposites based on carboxymethyl cellulose and Al and Cu nanoparticles to increase oil production// Gas Territory, 2017ç № 3ç p. 34-38

[9] V.M. Shamilov, E.R. Babayev Development of multifunctional composite mixtures based on watersoluble surfactant, polymer and metallic nanopowder as agents of oil displacement// Oil and Gas Territory 2016 № 6 p.60-63

[10] V.M. Shamilov, E.R. Babaev, P.Sh. Mammadova, I.G. Ayubov, E.G. Gadzhiev. Some aspects of carbon nanotubes application for increasing oil recovery factor // SOCAR Proceedings Special Issue. № 1, 2023, 115-120

[11] V. Shevchenko. Fundamentals of Physics of Polymer Composite Materials Moscow. 2010, 99 p. https://f.eruditor.link/file/307099/

[12] Shitov, K.S. Babina, A.N. Pachino, T.P. Kravchenko Moscow. UDC 678, 6.046. Advances in chemistry and chemical technology. v.28, 2014, №3
https://www.muctr.ru/upload/iblock/c61/c61d634fb3d0bccae0f3a2b26b3d48ac.pdf

[13] A.Kopmalkov, S.Barinov, M.Alymov Fundamentals of technologies and application of nanomaterials. M.: Nauka, 2016, 208 p, https://biblioclub.ru/index.php?page=book&id=457666

[14] Ch. Wu, F. Xu, H. Wang, H. Liu, F. Yan, Ch. Ma. Manufacturing Technologies of Polymer Composites. A Review. Polymers 2023, 15,3, 712. https://doi.org/10.3390/polym15030712

[15] A.I. Slutsker Effect of mechanical loading on the kinetics of electrical destruction of polymers. Journal of Technical Physics. 2008, 78(ll) p. 60-63,
https://cyberleninka.ru/article/n/mehanizmy-elementarnyh-aktov-v-kinetike-elektricheskogo-razrusheniya-polimerov/viewer

[16] M.Pandey, G.M. Joshi, N.N. Ghosh. Ionic conductivity and diffusion coefficient of barium-chloride-based polymer electrolyte with poly(vinyl alcohol)–poly(4-styrenesulphonic acid) polymer complex. Bull Mater Sci. 2017, 40, p.655–666, https://doi.org/10.1007/s12034-017-1430-0

[17] N. Abduljabbar, Sh. Al-Busaltan, A. Dulaimi, R. Al-Yasari, M. Sadique, H. Al Nageim. The effect of waste low-density polyethylene on the mechanical properties of thin asphalt overlay. 2021. 125722, https://doi.org/10.1016/j.conbuildmat

[18] Z.R.Makamov, D.K. Yakubova Impact of ultraviolet rays on the mechanism of formation of transverse spatial bonds of polyethylene compositions, Tashkent, 2018.
https://www.researchgate.net/publication/339843416

[19] Z. Han, A.Fina, Thermal conductivity of carbon nanotubes and their polymer nanocomposites: A review. Progress in Polymer Science. 2011. 36. P.914–944.
https://doi.org/10.1016/j.progpolymsci.2010.11.004

[20] T. Kuila, S.Bose, C.E. Hong, M.E. Uddin, P. Khanra, N.M. Kim, J.H. Lee, Preparation of functionalized graphene/linear low density polyethylene composites by a solution mixing method. 2011. 49, p.1033–1037. http://dx.doi.org/10.1016/j.carbon.2010.10.031

[21] T. Chang, X. Zhang, H.-L. Cui. Terahertz Dielectric Spectroscopic Analysis of Polypropylene Aging Caused by Exposure to Ultraviolet Radiation. Polymers. 2019, 11, 12, p.2001.

[22] A.R. Sadygova, I.I. Abbasov, E.S. Safiev, A.A. Hadiyeva, Kh.O. Sadig, V.A. Alekperov. Molecular Processes in Electric Destruction of Nanocomposites Polyethylene + Nanoclay after Exposure to Electric Discharge and Orientation. Nanosystems, nanomaterials, nanotechnologies 2021, v.19, №1, p.103–110 Printed in Ukraine.
https://www.imp.kiev.ua/nanosys/media/pdf/2021/1/nano_vol19_iss1_p0103p0110_2021.pdf

[23] A.M. Magerramov, A.A. Shukyurova, M.A. Nuriev Electrical Conductivity of Nanocomposites Based on Low Density Polyethylene and Cu2S Nanoparticles, Surface Engineering and Applied Electrochemistry, 2018, 54, 1, p. 32–37.
https://link.springer.com/article/10.3103/S1068375518010118

[24] S. Therias, G. Rapp, C. Masso, J.L. Gardette. Limits of UV-Light Acceleration on the Photooxidation of Low-Density Polyethylene. Polymer Degradation and Stability. 2020, 183, 109443. doi: 10.1016/j.polymdegradstab

[25] S.Zdenko, T. Dimitrios, P.Konstantinos, G.Costas. Carbon nanotube–polymer compo sites: Chemistry, processing, mechanical and electrical properties. Journal of Progress in Polymer Science. 2010. 35, p.357–401. DOI: 10.1016/j.progpolymsci.2009.09.003

[26] M.A. Mehrabova, S.I. Mammadova, F.Sh. Kerimov, S.I. Safarova, K.J. Gulmamedov, I.H. Hamdillayeva. Influence of discharges and uv irradiation on the electrical properties of high pressure polyethylene and compositions on its base. Polymer-Plastics Technology and Materials. 63, 2024, 16, pp. 2237-2245. https://doi.org/10.1080/25740881.2024.2369677

DESIGN AND PRODUCTION TECHNOLOGY OF SPECIAL FRICTION CLUTCH INSIDE INNOVATIVE REDUCERS OF RAILROAD SWITCHES

Ayaz Abdullaev[1], Isa Khalilov[2], Goshgar Rasulov[2]

•

[1]Baku, Azerbaijan
[2]Department of Machine Design and Industrial Technologies, Azerbaijan Technical University, Baku, Azerbaijan
ayaz.abdullayev.40@mail.ru, khalilov@aztu.edu.az, qoshqarrasul@aztu.edu.az

Abstract

The article discusses the development of innovative technologies aimed at increasing the speed, efficiency, productivity and safety of railway transport systems based on ensuring modern requirements for dimensions, the number of components and the reliability of manufactured switch devices. The design of a special three-disk friction clutch is designed, which is placed between the gears of a double-crown gear block of the second stage of a package innovative reducer intended for mechanical transmission systems of railroad switch devices, based on their design and functional features. The design of the friction clutch is quite compact, provides for a decrease in the dimensions and weight of the entire transmission mechanism, an increase in its reliability and technical indicators, and protection of the system from overloads. Considering the advantages and technical level of the reducer kit, with a specially designed friction clutch, it is recommended for use on switch drives of the "СП" brand in the railway transport systems of the CIS countries.

Keywords: Railroad, switch drive, innovative, package reducer, friction clutch, construction, technology.

I. Introduction

Thanks to the new opportunities presented by the last industrial revolution, enterprises are striving to improve the quality of their products as much as possible to gain a competitive edge in the market. Over time, the goal of human industry has been to make each product more ergonomic, aesthetic, efficient, productive, and reliable. Accordingly, the size of the products should be smaller, the number of components as minimal and reliable as possible, and the mechanical systems used are continuously improved. Various research is being conducted to create more modern, compact, and safe transmission mechanisms [1-4].

In the modern era, one of the fields directed towards the application of continuously developing and more advanced technologies, particularly due to the importance of speed and safety, is the railway transportation system [5]. It also plays a significant role in the political activities of countries worldwide. The transit corridors and logistics of this system constantly contribute to the economic relationships between countries. Ensuring the safety, productivity, and sustainable development of railway systems, which transport passengers and large volumes of cargo over long distances efficiently, is always a priority. To ensure the proper functioning of railway infrastructure, the

coordination and safety issues between its various mechanisms are of special importance and must be efficiently resolved.

The railway system's structures and equipment are continually adapted to meet the modern structural and technological requirements, thus improving the overall performance of the system. One of the directions for increasing the productivity of the railway system is the railway switch, which plays a crucial role in regulating the routes of the rolling stock (trains, locomotives, wagons, etc.). In the railways of Azerbaijan and other CIS countries, the main switches used to facilitate the transition of rolling stock from one track to another are primarily of the "СП" brand (refer to Figure 1, Figure 2). The speed and reliability of the switch operation depend on the mechanical transmission mechanism, particularly the working condition of the protective friction clutch, which is a key element of the latter [6]. Therefore, the rational design of the friction clutch in the double-crown gear block of the package reducer and the development of its manufacturing technology is a pressing issue.

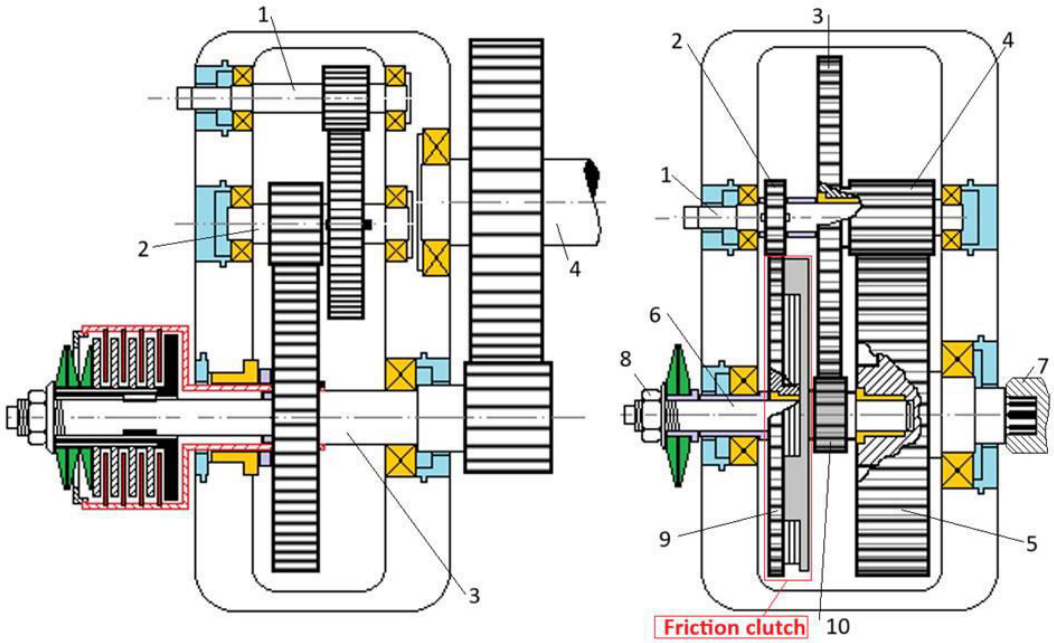


Figure 1: a) *Existing reducer of the "СП" brand railroad switch; b) Protective clutch designed inside of package brand reducer*

II. Statement of the problem

The increasing speed and safety requirements in railway transport, as well as the identification of the shortcomings of the components constituting the system, and the need for their elimination to ensure more reliable operation and productivity, are critical issues. Therefore, the functionality and reliability of "СП" brand switches have been investigated, and it has been determined that the reliability of the mechanical transmission system, consisting of a two-stage reducer and an open gear stage, is lower due to open gear drive stage [7, 8]. In order to make the gear transmission mechanism fully enclosed, an innovative mechanical transmission for the railway switch has been developed by replacing the existing reducer and open gear transmission with a compact three-stage package reducer [9]. The proposal of a three-stage package reducer instead of the existing reducer and open gear transmission for the railway switch has ensured the improvement of the technical level indicators of the transmission and compliance of the mechanical system with modern requirements [10]. The package reducer, which differs from traditional reducers in its structural and functional

characteristics and has its main components consisting of double-crown gear blocks, is an interesting issue in terms of ensuring the compactness and reliability of the protective clutch designed inside the mechanism. Its positive solution is an essential part of the efficient operation of the switch. Therefore, when designing the clutch for the package reducer, the features of modern designs and technologies compatible with this field are utilized [11, 12]. It is clear that the role of the protective clutch is to prevent the overload of the motor by preventing foreign objects (such as stones, iron pieces, etc.) from falling between the switch blade and the rail, which would otherwise cause the transmission motor to fail.

Taking into account the above-mentioned factors and the operating conditions of the clutches, the need arises to develop the construction of an optimal compact protective clutch in the double-crown gear transmissions of the proposed innovative transmission for "СП" brand switches, as well as the manufacturing technologies of its components.

It should be noted that some of the reducers marked with "СП" (for example, "СП-2", "СП-8", "СП-10") are lubricated with solid grease on the friction discs, while others (such as "СП-6", "СП-6М", "СП-6МГ") are lubricated with liquid oil inside the reducer [6, 13, 14].

The aim of the work is to rationally place the friction clutch in the reducer designed for the railroad switches, design its construction, and develop advanced processing technologies for its components.

III. Methodology

To achieve the set goal, the following constructive and technological issues must be addressed:

Constructive Issue: Ensuring the minimum number of parts, along with the friction clutch, while determining the dimensional measurements of the gearbox and ensuring its reliability.

Technological Issue: Developing advanced methods for the machining, assembly, and repair technologies of the friction clutch components.

The existing transmission mechanism of the "СП" brand switch used in the railway system of the CIS consists of three stages and four shafts (1-4) (refer to Fig. 1, a). Since sufficient information is provided in the literature regarding the mechanical transmission mechanism of the "СП" brand switches and the features of the package-type reducers, these issues are not discussed in detail in this work [6, 10].

To ensure the high operational performance of the protective clutch, the following provisions and requirements are considered when designing it within the package type reducer:

1. Minimum mass and dimensional measurements: If the friction clutch is placed inside the reducer housing, no additional housing is required;
2. Minimal number of components: The number of key connections and friction discs in the clutch must be minimized.;
3. Simplicity of maintenance and repair: The lubrication of the friction discs should be performed together with the lubrication of the reducer's gear, and the replacement of the friction discs should be simplified by the ease of disassembling and assembling the reducer;
4. Ease of adjustment: The nut for adjusting the friction disc compression force must be located in a convenient place;
5. The manufacturability of the manufacturing and assembly of the clutch parts must be ensured..

Taking into account the set requirements, existing conditions, and given data (such as the dimensions of the reducer, kinematic and structural configuration, the position within the switch box, internal space of the box, etc.), an analysis of the system was carried out. It was determined that since the driving shaft (1) of the first stage is connected to the motor shaft, it is not possible to install the friction clutch between the driving gear (2) and the driving shaft. In the last (III) stage, i.e.,

between the gear wheels (3-4) of the second two-crown block or between the driven gear (5) and the driven shaft (6), the design of the friction clutch also appears inefficient, because (refer to Figure 1, b):

1. This stage is heavily loaded, and since the transmitted torque is large, the number of friction discs must be increased or their diameters must be larger.;

2. The electric-automation device installed in the output direction of the reducer and the main shaft of the transmission (7) make it impossible to place the adjustment nut (8) (refer to Figure 1, b).

Therefore, it is possible and efficient to design the protective friction clutch between the gear wheels (9, 10) of the two-crown block, which is mounted on the driven shaft, i.e., between the driven gear of the first stage and the driving gear of the second stage (refer to Figure 1, b).

Since the working parts of the clutch are placed between the driven gear (8) and the driving gear (9), each of which also serves as a half-clutch, its dimensions must be coordinated with the dimensions of the gear wheels.

The half-clutch are part of the driven gear (8) and the driving gear (9) and are in static contact only with the discs. Based on the operational experience of friction clutches, materials such as hardened steel or metal-ceramics are chosen for the friction disks.

Since the gear wheels of the block separately perform the function of half-clutch, the dimensions of the friction discs must also be coordinated with the sizes of these gear wheels. Therefore, the friction discs are accepted to have the largest possible areas, and their number is determined functionally through calculation. The clutch is designed for "СП" brand railway switches. Thus, the outer and inner diameters of the compressing discs are determined based on the dimensions of the driven gear of the stepped gear. As a result of the calculation for the package reducer designed for "СП" brand railway switches, the dimensions for the driven gear are as follows: Pitch circle diameter: $d_2 = 150\ mm$; Outside circle diameter: $d_{a2} = 154\ mm$; Root circle diameter: $d_{f2} = 145\ mm$; Gear width: $b_2 = 5\ mm$; Gear hub diameter: $d_{t2} \approx 34\ mm$; Gear hub length: $l_{t2} \approx 12\ mm$.

The two pressing discs of the friction clutch (with thickness initially accepted as b_S=1,5 mm, and later verified through calculation) are mounted on the hub of the driven gear (9), which also serves as a half-clutch, and rotate together with it. Therefore, the outer surface of the hub and the inner surface of the discs are designed as cylindrical surfaces cut by symmetrical sectors. Thus, the outer diameters of the pressing discs are taken as D_{Sx}=133 mm, and the inner diameters as d_{Sd}=34,5 mm (refer to Figure 2, a-b).

Between the pressing discs, a driven disc (initially accepted with a thickness of b_A=3 mm) is placed, which has three key projections on its external cylindrical surface. This disc, with its key projections-teeth, is inserted into the three keyseat grooves of the driving gear's (10) cylindrical structural element (referred to as the cover here), and is mounted on its internal cylindrical surface. The outer and inner diameters of the driven disc are D_{Ax}=134 mm and d_{Ad}= 67 mm, respectively; the height of the teeth and the length along the circumference are h_t=3,5 mm and L=8 mm (refer to Figure 2, b).

The outer and encompassing inner diameters of the cover, the height, length, and depth of the tooth grooves, the thickness of the cover, and the depth of its encompassing inner surface, as well as the diameter of the cover's hub and the diameter and depth of the groove into which the hub of the driven gear of the first gear stage can fit during the compression of the discs, are determined based on the intended function and operating conditions. These values are calculated using known formulas according to the methodologies for determining the structural elements of machine components. The remaining dimensions are taken as design choices (refer to Figure 2, a-b; in the figure, only nominal dimensions are provided, surface qualities, etc., are not specified) [8, 15].

The design of the friction clutch for the "СП" brand railway switches was carried out using the well-known methodologies provided in the literature, based on the operating conditions and parameters [16, 17].

The force exerted by the pressing discs: F_a= 8289,6 *N*;

The number of friction surface pairs in the clutch:

$$i=\frac{K\cdot T}{f\cdot F_a\cdot R_{or}}=\frac{1{,}3\cdot 14{,}06\cdot 10^3}{0{,}06\cdot 8289{,}6\cdot 50}=\frac{18278}{24868{,}8}=0{,}73\ \approx 2 \quad (1)$$

Where: *K*=1,3 – operating mode coefficient;

T=14,06 *Nm;* - the maximum transmitted torque of the clutch,

$F_a = 8289{,}6\ N$ – the force exerted on the pressing discs;

$R_{or} = 50\ mm$ - the average radius of the friction surface pairs;

f = 0,06 – the friction coefficient between the treated steel surface.

The projections-teeth of the driven disc in the clutch have been checked for crushing stress:

$\sigma_c = 14{,}1\ N/mm^2$ <$[\sigma_c]$=(25÷ 35) MPa (condition is met)

To facilitate the repair process and ensure the longevity of the main parts, the number of friction pairs is assumed to be 2 (refer to Eq. 1) and the working surfaces of the friction discs are assumed to be the friction surfaces. The dimensions of the nut, which will compress the half-clutches by being connected to the intermediate shaft, have been selected according to the existing "СП" reducer's clutch. Additionally, the threads have been checked for shear stress.

The designs of the friction clutch parts are presented in a simplified form (refer to Figure 2, Figure 3).

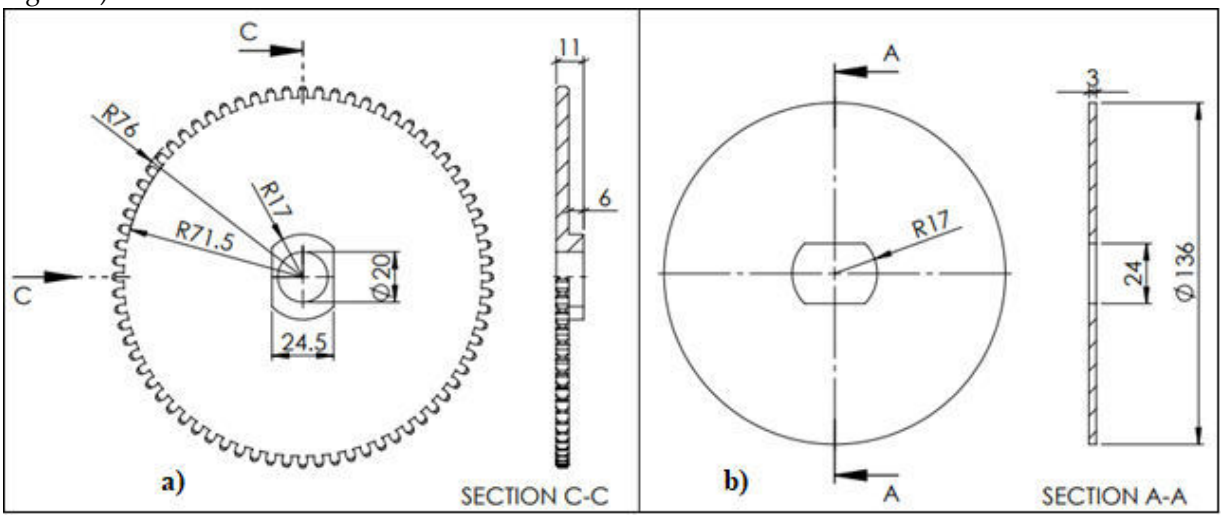


Figure 2: *Left half-clutch: a) driven gear of the first stage; b) pressure discs of the friction clutch*

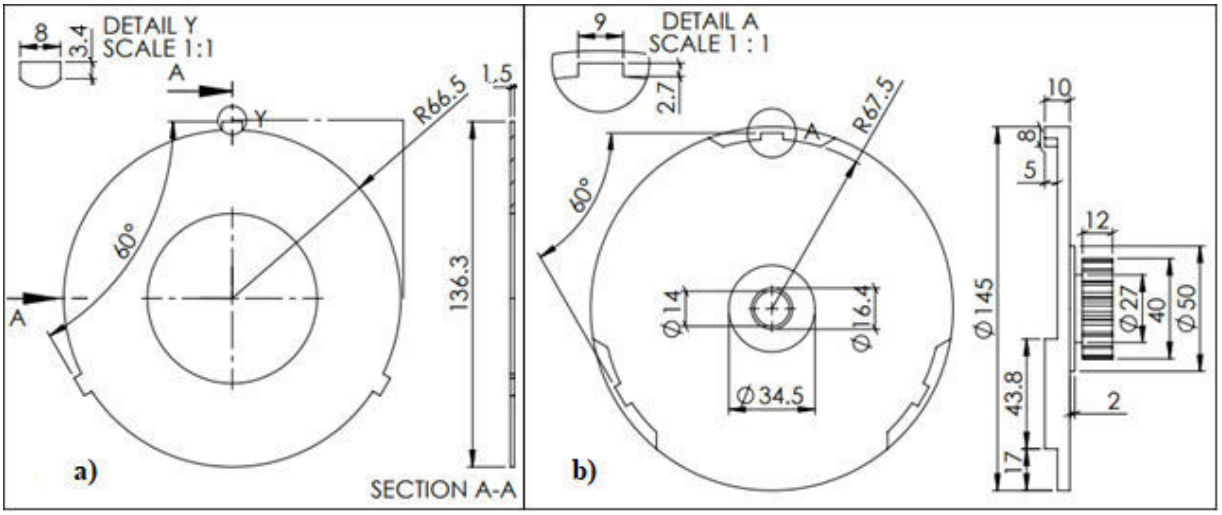


Figure 3: *Right half-clutch: a) driven disc of the friction clutch; b) drive gear of the second stage*

Manufacturing technologies of the friction clutch components. When developing technological processes for the components of the clutch, typical technologies for the components have been used [18, 19]. However, in this case, the structural features of the half-clutchs and discs, as well as the issues related to their base processing during machining, are new and different. Therefore, the creation of advanced technologies is required.

The main components of the clutch, which are different from traditional designs, consist of 5 working parts: three friction discs and two half-clutch gear wheels. Their manufacturing is carried out using two groups of technologies: for the gear wheels-half- clutchs and friction discs:

Manufacturing of half- clutch gear wheels**:**

Mass Production: 1) Turning of the rotation surfaces according to the typical gear wheel technology, 2) Rough polishing of high-precision rotation surfaces, 3) Milling of flat surfaces on the hub using a milling machine, 4) Gear cutting, 5) Thermal processing, 6) Cleaning, 7) Fine polishing;

Individual, small-scale production: 1) Turning of rotation surfaces in two placements on RPI machines, milling flat surfaces on the hub in sequence; 2) Rough polishing of high-precision rotation surfaces; 3) Gear cutting; 4) Thermal processing; 5) Cleaning; 6) Fine polishing.

Manufacturing of discs**:**

In mass production for steel discs: 1) Stamping along the contour from sheet material; 2) Thermal treatment; 3) Cleaning;

For steel discs, individual, small series production: 1) Cutting along the contour from sheet material using laser (or water jet); 2) Cleaning; 3) Thermal treatment; 4) Cleaning.

For metal-ceramic discs: 1) Selection and mixing of materials (metal parts: stainless steel and alloyed steels; ceramic parts: aluminum oxide and silicon carbide); 2) Addition of binder materials to obtain a homogeneous mixture; 3) Pressing the mixture into a disk shape; 4) Heat treatment (sintering at 900-1400°C) to achieve a hard crystal structure; 5) Surface treatment - applying coatings (titanium nitride or boron nitride) to the working surface of the disk.

IV. Discussion

The driven gear of the fist stage moves together with the shaft. Since the gear also serves as a half-clutch, its hub is designed according to the width of the driven discs, with the shape and dimensions shown in Figure 2. The dimensions of the clutch components have been determined using known calculation methodologies and are provided in the figures [16, 17]. The internal cavity

of the two pressing discs, which are rigidly mounted on the hub of the gear, is designed in the form of a profile joint to match the hub of the gear, ensuring its movement together with the gear.

Between the pressing discs, in case of excessive loading, a driven disc is placed, which can move freely, independent of the driven gear and the pressing discs. This disc has three protruding teeth positioned at a 120° angle relative to each other on its outer circumference. Its internal diameter is taken as 0,5037 times the external diameter of the friction surface, in accordance with the allowable limit (refer to Figure 3).

To drive the intermediate stage's driving gear by meshing with the teeth of the compressed driven gear, a cylindrical cap is designed with corresponding holding cavities for the teeth, serving as a second crown for the gear, with dimensions matching those of the teeth (refer to Figure 3).

Since the gears are tightened with a nut, there is no need for a key connection. Within the specified limit, the driving moment is transmitted, and the gears rotate together with the shaft.

V. Conclusions

1. A protective friction clutch has been designed within the proposed three-stage, two-shaft package-type reducer, which serves as the mechanical transmission mechanism of the railway switch. The clutch is installed between the gear wheels of the intermediate stage's two-crown gear block, ensuring the minimization of component count, along with the dimensions of the reducer, and guaranteeing its reliability together with the friction clutch.

2. Based on the functional design parameters and the structural features of the two-crown gear blocks, the main functional components of the mechanical system were calculated, and their constructive dimensions were determined. The working drawings have been developed accordingly.

3. Taking into account the types of production, progressive technologies are proposed for the manufacturing of the main working components of the friction clutch.

References

[1] Abdullaev, A.I., Rasulov, G.N., and Ismailov, O.F. (2020). Mathematical Modeling of the Difference in Tooth Direction Angles in the Meshing Zone and the Contact Ratio in Gear Transmissions. *Scientific and Technical Bulletin of Information Technologies, Mechanics, and Optics*, 1(1), 110–117. https://doi.org/10.17586/2226-1494-2020-20-1-110-117.

[2] Abdullaev, A.I., and Rasulov, G.N. (2022). Optimization of 'AN' Type of Three-Stage Package Reducers in Accordance with Gear Ratios. *Proceedings of the 8th International Conference on Control and Optimization with Industrial Applications*, 30–32. Baku, Azerbaijan, 24–26 August 2022. https://www.researchgate.net/publication/386148569_.

[3] Lustenkov, M., Khalilov, I., and Moiseenko, A. (2024). Comparative Analysis of Dynamic Characteristics of Spherical and Eccentric Transmissions with a Double-Ring Satellite. *Advances in Science and Technology*, 148, 103–110. https://doi.org/10.4028/p-msJjS7.

[4] Chalabi, I.G., Ahmedov, B.B., Charkasov, Sh.B., and Yusubov, Sh.T. (2022). Comparative Service Life Analysis for Novikov Gears of Mechanical Drives of Sucker-Rod Pumps. SOCAR Proceedings, Special Issue 2, 53–59. file:///C:/Users/admin/Downloads/OGP2022SI200776.pdf

[5] Elyazov, I.S., Huseynov, I.D., and Ravlyuk, V. (2022). Determination of Technical Features of the New Generation Electric Carriers of the Azerbaijani Railway. *Collection of Scientific Works of the Ukrainian State University of Railway Transport*, 198, 151–158. https://doi.org/10.18664/1994-7852.198.2021.256651.

[6] Soroko, V.I., Kainov, V.M., and Kaziev, G.D. (2006). *Automation, Telemechanics, Communication, and Computing Equipment on the Railways of Russia: Encyclopedia*. Vol. 1. Moscow: NPF "Planeta".

[7] Ahmadov, B.B. (2024). *Machine Design: A Textbook*. Baku: AzTU. https://www.researchgate.net/publication/386292517.

[8] Khalilov, I.A., Huseynov, A.S., and Yusubov, S.T. (2009). *Reliability of Machines: A Textbook*. Baku: AzTU.

[9] Abdullaev, A.I. (2022). Switch Electromechanical Drive for Railways, Eurasian Patent No. 040109. *Eurasian Patent Office*. Edited by A.A. Alifov, B.B. Ahmadov, I.G. Chalabi, G.N. Rasulov, and A.B. Hadjiev, 21 Apr. 2022. https://www.researchgate.net/publication/386135915.

[10] Abdullaev, A.I., Rasulov, G.N., Huseynov, I.D., and Ismayilov, O.F. (2022). Innovative Reducer for Railroad Switch Drives and Evaluation of Friction Work on Double Sliding Bearings. *SOCAR Proceedings*, Special Issue 1, 1–6. https://doi.org/10.5510/OGP2022SI100700.

[11] Khalilov, I.A., and Sofiyev, A.H. (2024). Dynamic Behavior of Shafts, Couplings and Working Body of the Machine under Torsional Impact Moment. *Journal of Applied and Computational Mechanics*, 10(4), 842–852. https://doi.org/10.22055/jacm.2024.46217.4482.

[12] Khalilov, I.A. (2010). Influence of the Gap in Clutches on the Dynamic Properties of Drives. *Russian Engineering Research*, 30(3), 206–212. https://doi.org/10.3103/S1068798X10030032.

[13] "Electric Drive Type СП-6." (n.d.). *Studfile.net*. Accessed [11.11.2024]. https://studfile.net/preview/2913663/page:5/.

[14] Buryak, S.Y., Gavryliuk, V.I., Gololobova, O.A., et al. (2014). Study of Diagnostic Features of Alternating Current Switchgears. *Science and Progress of Transport: Bulletin of the Dnipropetrovsk National University of Railway Transport*, 4(52), 7–22. ISSN 2307-3489.

[15] Nadjafov, A.M., Chalabi, I.G., Amanov, Y.A., and others. (2018). *Fundamentals of Machine Parts and Construction: Designing Mechanical Transmissions*. Baku: AzTU.

[16] "Calculate Clutch Torque." (n.d.). *X-Engineer.org*. Accessed [12.11.2024]. https://x-engineer.org/calculate-clutch-torque/.

[17] "Clutches." (2015). *wordpress.com*. Chapter 24. Accessed [16.11.2024]. https://sedyono.wordpress.com/wp-content/uploads/2015/10/ch-24.pdf.

[18] Suslov, A.G., and Dalsky, A.M. (2002). *Scientific Foundations of Mechanical Engineering Technology*. Moscow: Mashinostroenie. http://dx.doi.org/10.1063/5.0119565

[19] Asthana, R., Kumar, A., and Dahotre, N.B. (2006). Materials Science in Manufacturing. In *Materials Processing and Manufacturing Science*, Chapter 3: Powder Metallurgy and Ceramic Forming, 167–245. https://doi.org/10.1016/B978-075067716-5/50005-4

STRESS DISTRIBUTION IN THE MATRICES OF END FRICTION SEALS UNDER LOADING

Afet Jafarova[1], Farid Jafarov[2], Fuad Jafarli[3]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan, [2]Karabuk University, Turkey,
[3]Columbia University, USA,
afetceferova8@gmail.com, jafarovfarid3@gmail.com, jafarlyfj@gmail.com

Abstract

The distribution of stresses generated in antifriction seal materials is considered. A photo of isochromes is presented when loading an isotropic material and a material containing solid inclusions. Isochromes when loading a material consist of solid inclusions of various shapes under a load of 120 kg. Photo drawings of the distribution in depth from a concentrated load of tangential stresses in an isotropic material are also presented. In a material consisting of solid inclusions of a round shape, the distribution of tangential stresses depends on the introduction of a "stamp" into the material. Isochromes are obtained when loading a material containing many inclusions.

Keywords: isochrome, shear stress, sealant, isotropic material, solid inclusion, concentrated load, matrix, particle.

I. Introduction

High-pressure end-face friction seals operate under very difficult conditions of high friction and wear loads. The development of new effective tribotechnical materials and the selection of friction pairs are associated with the improvement of testing methods and measuring output characteristics, as well as methods for assessing the performance of friction pairs. At the same time, the development of rational test cycles is one of the necessary conditions for the optimization and selection of friction pairs.

The development of tribology and tribotechnics has led to the need to develop models for optimizing complex systems. Successful solutions to complex problems that cannot be reduced to a single known algorithm are possible using various modeling, primarily based on system analysis. For some problems, physical modeling is still of great importance, allowing one to clarify the nature of phenomena, their mathematical description, and perform experiments with friction units that have no analogues. The greatest correspondence between the model and the original can be achieved using the basic provisions of adaptability, and with the accumulation of new statistically reliable data on the correspondence of the model to the original, it is necessary to consistently adjust the model parameters.

For friction and wear problems, methods of assessing the determining processes and phenomena through a model, rather than a natural experiment, the description and results of which are presented in criterial form, are of particular importance. The use of modeling methods with a correct and objective approach to their reliability and accuracy allows for a significant acceleration of the implementation of new technical solutions and a reduction in labor costs and expenses for scarce materials.

The purpose of this work is to create and select materials for end friction seals and implement the most rational design of this type of friction, which is possible only on the basis of a deep study of the processes of stress distribution in the matrices of materials under its loading.

II. Problem statement

There is a correlation between wear tests of composite materials, as well as the amount and distribution pattern of stresses in the surface layers [1, 2]. In works [3-5], an analytical definition of the amount of solid inclusions in composite materials of the filled matrix type is given, depending on the condition of transfer of contact loads to the matrix, the size of the areas of the solid component on the surface of the composition is calculated, depending on the contact conditions of the surface of the composite material and the counterbody, and an analysis of the dependence of the friction force on the composition and structure of the composite material is carried out using a friction surface model. It was found that in order to ensure high antifriction properties and load-bearing capacity of the material, it is necessary to fill it with wear-resistant solid inclusions exceeding the contact patch in size. Among these, inclusions with sizes from 0.35 to 2.0 mm are considered more optimal. However, this information is insufficient for the design of this type of antifriction materials. It is likely that not only the size, but also the shape and arrangement of the filler particles have a significant impact on the antifriction properties, wear resistance and strength characteristics of the composite. Depending on the listed factors, the solid filler inclusions take up the load caused by friction, and the composite matrix promotes a favorable distribution of stresses or, on the contrary, ensures the concentration of stresses in local areas, thereby reducing the fatigue strength of the material [6].

To select the optimal variant of the "structural design", it is interesting to study the stress state of an iron-cast iron composite containing cast iron inclusions of various shapes. These inclusions should be oriented relative to the direction of the load and significantly exceed (several times) the material matrix in strength characteristics.

III. Experimental methodology

Based on the methodology developed at the Kyiv State University named after T.G. Shevchenko, the stress state of structural anisotropic models was investigated using the polarization-optical method under conditions of static loading by normal tangential forces, which are formed depending on the shape and mutual arrangement of the structural elements. We made models from optically active material ED in the form of 5 mm thick plates. The technology for producing such a polymer was investigated in [7].

However, obtaining an optically sensitive polymer applicable to this class differs from generally accepted methods by the presence of high-modulus inclusions that differ in quantity and shape, as well as mutual arrangement in the material. Due to the need to acquire structures of non-uniform plates with different modules, the developed method consisted of the sequential implementation of two processes. The high-modulus polymer intended for inclusions was produced by thermal curing of epoxy resin FD (100 wt/h) with methyltetrahydrotalin anhydride (60 wt/h) in a stepwise heat treatment, the initial temperature of the polymer was 70 ^{0}C and the final temperature was 120 ^{0}C.

To obtain a relatively low-modulus structure of the model, ED-20 or ED-5 epoxy resin was used, fixed with L-20 polyamide in a ratio of 100:140 or 100:150. Before mixing, we heated the original components to 700 ^{0}C, and then mixed them at this temperature. The composition was kept for 5 minutes to evaporate the bubbles and poured into a mold, into which pre-prepared inclusions of high-modulus polymer were placed. The shape, location and number of inclusions varied.

The mould consisted of metal plates coated with organosilicon liquid K-21. Hardening was carried out at room temperature for 24 hours. The optics – mechanical properties (modulus of elasticity E, punch coefficient μ, relative optical coefficient C, size of the material strip) accumulated by forces along the diameter in compacted disks, as well as elementary tension of the rod were determined. Photographs of the strips were taken with a photoelastimeter of the GMB-57 type (Czech Republic), and the number of strips was designated by the letter n.

IV. Solution to the problem

We encrypted the obtained isochrome images using the strip method, which consisted of direct optical measurement of the order of the stripes at the points under study and determination of the difference in the main equations using Bertheim's formulas

$$\sigma_1 - \sigma_2 \tag{1}$$

Considering the moderation of the stress state at critical points, it is necessary to check the strength under contact stresses according to the third theory of strength.

$$\sigma_{ucl} = \sigma_1 - \sigma_2 \leq [\sigma] \tag{2}$$

As a strength criterion, we adopted the value of the greatest shear stress. It was assumed that, in general, the limit state occurs when the greatest shear stress τ_{max} reaches a critical value. According to the equation

$$\tau_{max} = \frac{1}{2}(\sigma_1 - \sigma_2) \tag{3}$$

the failure condition and strength can be expressed in terms of principal stresses.

Considering that dividing the principal stresses and obtaining individual components σ_x, σ_y, τ_{xy} is difficult, we presented the results in the form of distribution graphs for individual sections $\sigma_1 - \sigma_2 = 2\tau_{max}$ in the depth of the model, obtained from interference images of the bands (isochrome). The graph shows the distribution of maximum shear stresses in horizontal parts along the depth of the model, starting from the contour of load application. The influence of the particle strength shape and its location in the matrix in relation to the place of application of the normal load was studied using the example of single inclusions (Fig. 1 and Fig. 2). We selected similar linear dimensions of particles of different shapes. At a distance from the surface equal to 0.1; 0.3; 0.5; 1 to the linear parameter of particle a in the images of stripes, we selected the concentration coefficient τ_{max}. The distance between the axis of force application and the particle symmetry axis was 0.5; 1.2 according to the linear parameters of particle a.

Figure 1: *Isochromes under loading of an isotropic material (left) and a material containing solid inclusions (right). Load 150 kg; E_{ucl} = 46000; E_{max} = 16000*

a) *b*) *c*)

Figure 2: *Isochromes under loading of a material consisting of solid inclusions of various shapes under a load of 120 kg. a - E_{ucl} = 45000, E_{max} = 16000; b - E_{ucl} = 45000, E_{max} = 16000; c - E_{ucl} = 46000, E_{max} = 16000*

At a distance of approximately 1–1.2 times exceeding the characteristic linear size of the particle, the magnitude of the stress state depends significantly on the particle shape. At acute angles, in places of abrupt changes in the direction of the geometric shape of the particles, i.e. where the derivative of the geometric shape is prone to destruction, a sharp concentration of stress areas is observed. Probably, the stress function at the boundaries of the area under study has its own characteristics [8]. Large local deformation, naturally, leads to some relaxation of stresses and, at the same time, it was not possible to note the maximum jump in concentration. According to the images of the bands, a local stress concentration was noted that was 2–3.5 times greater than neighboring areas. According to the circumferential shape of the inclusions, the transition from the matrix to the inclusions occurs smoothly, with an insensitive concentration (Figs. 1–3).

At a distance of 1.3 times exceeding the characteristic linear dimension a from the particle surface, the stresses in the matrix are 0.5 – 0.8 stresses on the same area in an isotropic material. If the load is applied to inclusions at a distance of 0.5*a* from the symmetry axis, the particle itself takes the main load, reducing the stress area in the matrix by 5 – 5.1 times compared to local areas corresponding to an isotropic low-modulus plate. This is especially evident in the uniform distribution of the apparent load - "stamp" (Fig. 4). The difference in local coefficients of comparative concentration τ_{max} for the total force is 3 – 1.2.

In this case, inclusions are subject to loads at a distance of more than 1*a* from the axis of symmetry, the influence of the particle gradually decreases, and at a distance of more than 1.5*a* this influence becomes insignificant.

Thus, a solid inclusion affects the distribution of the stress field in the matrix along a radius equal to approximately 1.2 – 2*a*. To obtain a composite material with high load-bearing capacity and fatigue strength, it is necessary that inclusions resistant to strong wear are located in the matrix at a distance of at least 1.5 – 2*a*.

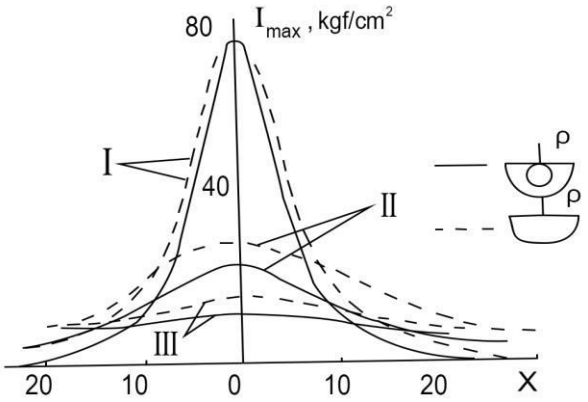


Figure 3: *Distribution in depth from a concentrated load of tangential stresses in an isotropic material with round solid inclusions. i..I-5; II-15; III-30 mm, p=150 kg*

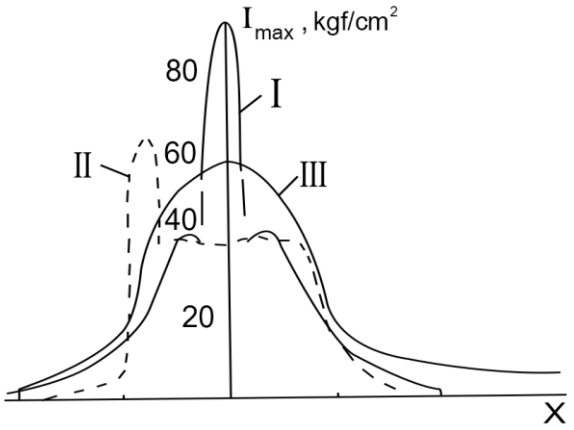
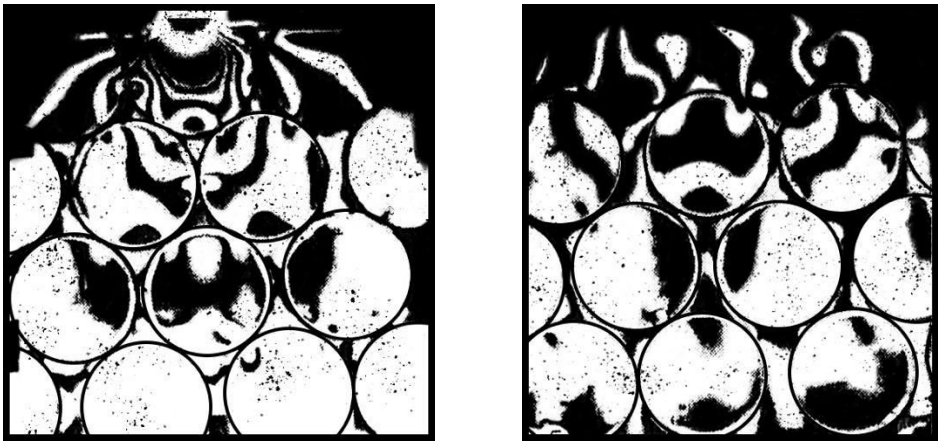


Figure 4: *Distribution of tangential stresses depending on the penetration of the "stamp" into the material containing a solid inclusion. p=120 kg, i=1.6 mm*

The study of the stress state of the model in most inclusions qualitatively complements the results obtained in the study of the stress state of models with homogeneous solid inclusions (Fig. 5).

Under the total load, the matrix particles (Fig. 5, a) are combined and experience the action of stresses, and it is clear that it is closed between two adjacent particles and creates greater stresses compared to the case of applying the total load of a homogeneous inclusion.

In the case of application of the "stamping" load in the surface layer of the matrix material, it is significantly lower than in the surface layer of inclusions. This occurs because the "stamping" load deforms both the matrix and the inclusions of the solid filler, but the elastic modulus of the matrix is three times lower than the elastic modulus of the inclusions, so the stress in the matrix is significantly lower [9].



a) *b*)

Figure 5: *Isochromes under loading of a material containing many inclusions. a) - total load 120 kg, E_{ucl} = 46000, E_{max} = 16000; b - "stamp", load 120 kg, E_{ucl} = 46000, E_{max} = 16000*

If we consider the conditions of contact of real composite surfaces with solid contact surfaces when analyzing images of models of most inclusions, then it is necessary to distinguish two qualitatively different variants of "die" load. One or another model corresponds to the conditions of loading of a single contact spot (in this case, solid inclusions have dimensions of about 1 – 10 microns) and the conditions of the outer contour of the nominal surface (for solid inclusions with linear dimensions of more than 30 microns). In both cases, the pattern of stress distribution is the same, but in the first case, the stresses are significantly greater, since contact loads, as a rule, exceed circumferential stresses by 100 times [10].

Thus, the concentration of stresses at the points of contact of solid inclusions is significantly less dangerous for compositions with large inclusions, since the load level is 100 times less.

III. Conclusions

1. The nature of the distribution and accumulation of stresses in antifriction powder materials of seals has been studied and the dependence of inclusions on their location and the dependence of their quantity on the shape, size of inclusions and matrix have been determined. This study has made it possible to select the optimal version of the "structural design" of the material. It has been established that in the compositions of composites containing large inclusions, the concentration of stresses at the points of contact of inclusions is significantly less dangerous, since in this case the level of load on the inclusion is significantly lower.

2. Regardless of the magnitude of the load and the method of its application to the surface, the optimal shape of solid inclusions in the composite material should be spheroidal or an ellipsoid of revolution. The characteristic linear size of the solid inclusion should be greater than the calculated diameter of a single contact spot.

3. To ensure high fatigue strength of the composite material, inclusions resistant to severe wear must be located at a distance of at least 1.6 – 2a (a is the characteristic linear dimension of the inclusion).

Financing

This work was supported by the Azerbaijan Sciense Fondation Grant **№ AEF – MGC – 2024 - 2(50) – 16/01/1 – M – 01.**

References

[1] Solonin S.M., Solonin M.D. Dependence of pressing pressure on the ratio of pressing area and friction surface against the matrix walls.//Powder Metallurgy, 1985, No. 11, pp. 15-18.

[2] Jafarova A. A. Compaction of Elastoplastic Iron Batch in Pressing. Journal //Russian Engineering Research, 2023, Vol. 43, No. 9, pp. 1095–1100.

[3] Maltsev M.V., Nikolaev A.N. Study of external friction of metal powders in the low pressure region. // Powder Metallurgy, 1989, No. 7, pp. 73 – 77.

[4]Mamedov A.T. and Mamedov V.A. New tecnological approach to fabrication of high density PM parts by cold pressing sintering// Powder metallurgy, 2004, London, Vol.47, №3. pp.278-283.

[5] Mamedov A.T. Structural and antifriction powder materials. Baku, Elm, 2005. - 458 p.

[6] Mamedov A.T., Mamedov V.A. Pnenomenological approach to the plastik deformation of metallik powder in confined cavity// 5Th international powder metallurgy conference. Ankara-Türkiye, Oktober 08-12, 2008, s.59-68.

[7] Berkovich I.I., Vinogradov G.A., Katashinsky V.P. Study of friction and shear resistance of iron and aluminum powders. // Powder Metallurgy, 1981, No. 11, pp. 84 – 88.

[8] Masanori Y., Toshio A., Katuji O. Studies on compacting of spherial cast iron powders. –Bul. ISME, 2006, 19, №132, p.699-706.

[9] Karafiath L.L., Morh G. Effect of ultrahigh vakuum on friction betwen metals and granular soils.- J. Vacuum science and technolog, 2009,6, №1, p.198-205

[10] Mammadov Arif, Mammadov Vugar. Comparative Analysis of Diffusion Metallization Coatings Applied on Steel Parts. *Metallurgical and Materials Engineering.31.10.2023, p.43-54.https://metallmatereng.com/index.php/home/article/view/1020/523*

FEATURES OF OPTIMIZATION OF PRESSING MODES OF POWDER MATERIALS FOR PARTS OF SHIP MECHANISMS

Nizami Ismayilov[1], Fazil Orujov[1], Elkhan Mammadov[1], Lachin Babayev[1], Aga Shixseyidov[2]

•

[1]Department of Shipbuilding and Ship Repair, Azerbaijan State Marine Academy, Baku, Azerbaijan

[2]Department of Machine Building Technology, Azerbaijan Technical University, Baku, Azerbaijan

nizami.ismayilov@asco.az, fazil.orujov@asco.az, elxan.mammadov@asco.az, lachin.babayev@asco.az, agashixseyidov@aztu.edu.az

Abstract

The article considers the features of optimization of modes of pressing powder materials for ship machinery parts using mathematical models. It is established that in order to obtain an adequate mathematical model, it is necessary to more accurately specify the rheological properties of powder materials, which largely determine the kinetics of compaction during HIP.

It is indicated that the solution of this problem by the finite element method for products of complex configuration with nonlinearity of the used relations and non-stationary nature of deformation requires the use of iterative procedures in the presence of a large number of finite elements at each time step.

It is determined that the upper-level model does not allow solving the problems of optimization and optimal control of the HIP pressing process. To overcome these difficulties, it is possible to use the lower level of modeling, including zero-dimensional and one-dimensional HIP models.

A system of technological modes of hot isostatic pressing of powder materials is proposed, which provides for discrete and continuous mathematical modeling of HIP. Discrete and continuous representation of the HIP technological process design system structurally includes the interaction of mathematical models of the upper and lower levels.

The use of lower-level models with the use of an optimization apparatus allows us to seriously narrow the search area for technological solutions, thereby obtaining the most reliable and accelerated information on the modes of pressing powder materials for parts of ship mechanisms.

Keywords: optimization, pressing modes, powder materials, mathematical models

I. Introduction

It is indicated that HIP is usually used for powder materials of relatively high cost, i.e. alloy steels, therefore, it is important to determine the optimal pressing conditions that provide the required level of physical and mechanical properties at minimal costs [4-6].

In this regard, the most important tasks are the design of a press mold for HIP in order to obtain a semi-finished product that is closest to the required product, as well as the determination of HIP modes to achieve the required level of properties in semi-finished products [3-5].

One of the ways to solve this problem is to develop a system for designing GIP process modes based on an adequate mathematical description of the object under study [6-8].

Analysis of literature data allows us to conclude that there are two main directions in mathematical modeling of GIP of powder materials: discrete and continuous. The first direction [3-5] considers the deformable body as a set of individual incompressible particles with a pore volume between them that changes as they compact.

The second direction [1-4], based on the phenomenological approach, considers the deformable body as a whole, endowing it with the ability to plastically change not only its shape but also its volume. In discrete theories, the system of equations for describing one-dimensional distributions of density, temperature, etc. can be generally represented as follows:

$$\frac{dD}{dt} = K_D \cdot f(D) \tag{1}$$

kinetic equation of compaction ($D-$ relative density; K_D- kinetic constant of compaction);

$$\rho_S \cdot D \cdot C_p \frac{d\theta}{dt} = div(\lambda qrad\theta) \tag{2}$$

heat conductivity equation (ρ_S- density in compact state, C-heat capacity, $\theta-$ temperature, $\lambda-$ thermal conductivity coefficient; $\lambda = \lambda(\theta D)-$ given function);

$$div\, v = -\frac{1}{D}\frac{dD}{dt} - \tag{3}$$

continuity equation ($v-$ speed).

This system is supplemented by a boundary condition

$$v_r /_{r=0} = 0 \tag{4}$$

and also the ratios

$$\rho_S = \rho_e(t) \text{ and } \theta_S = \theta_e(t) \tag{4'}$$

where $\rho_e(t)$ and $\theta_e(t)-$ are the functions specified by the pressing cycle.

Equation (1), depending on the functions included in it, describes various compaction mechanisms; diffusion through the grain body and along their boundaries to isolated pores, and a number of others.

Compaction from instantaneous plastic deformation can be calculated using fairly simple relationships (3) and obtain starting density values, from which the compaction process itself begins, depending on time. In this case, the overall compaction rate is the sum of the components from its individual mechanisms.

II. Purpose of the work

The aim of the work is to identify the features of technological modes of hot isostatic pressing of powder materials used for the manufacture of parts of ship mechanisms.

III. The object of the study

The object of the study is the technological modes of hot isostatic pressing of powder materials, and the subject of the study is mathematical models obtained using the finite element method for isostatic pressing of powder products of ship mechanisms.

IV. Equations of the process of pressing powder materials

It should be noted that the discrete theories used to describe the processes of deformation of powder products have a significant drawback. They make it difficult to describe the process of pressing at arbitrary points of bodies of complex configuration, which significantly distorts the uniformity of compaction. However, these problems can be successfully solved by using the apparatus of continuum mechanics.

Let us consider in general terms a system of equations describing the non-isothermal flow of a viscoplastic isotropic compressible material with strain hardening.

Equilibrium equation [7]:

$$div T_\sigma = 0 \tag{5}$$

where $T_\sigma -$ is the stress tensor;

$$T_\xi = \frac{1}{2}(qrad\ \vec{v} + qrad^T \vec{v}) \tag{6}$$

Kinematic relations:

$$\frac{d \ln \rho}{dt} + div\, \vec{v} = 0 \tag{7}$$

where $T_\xi -$ is the strain rate tensor;

Continuity equation, defining relations [8]:

$$D_\sigma = 2\mu D, \quad \sigma_0 = 3R\xi_0 \tag{8}$$

where $D_\sigma -$is the stress deviator; $D_\xi -$ the strain rate deviator; $\sigma_0 -$ the average stress; $\xi_0 -$the average strain rate, $\mu, R -$ the coefficients of shear and bulk viscosity.

Heat equation [9]:

$$c \cdot \rho \cdot \frac{d\theta}{dt} = div(\pi \cdot qrad\theta) + \nu \cdot T_{\sigma} \cdot T_{\xi} \quad (9)$$

where ν - is the coefficient of conversion to mechanical work.

Note that the coefficients of shear 2μ and bulk 3R viscosity included in the defining relations can be established experimentally. In this case, the concepts of creep potential Φ and loading surface F are used. In our case, we assume

$$F = F(T, \sigma_0, \sigma_{\rho}, x_i) = 0 \quad (10)$$

where T - is the intensity of tangential stresses, a and x_i refers to the hardening parameters, the environments of which can be deformation and speed.

The solution of equations (5) - (9) with boundary conditions (4') makes it possible to construct kinematic, force and temperature distributions in a deformable body. Based on the use of discrete and continuous representations, a design system for HIP technology is proposed, structurally including the interaction of mathematical models of the upper and lower levels.

The top-level model is a solution using the finite element method (FEM) of the system of equations (5) - (9). To obtain a system of nonlinear equations dependent on time, the weighted residual method [9-10] was used.

V. Algorithm for solving the problem

Let us consider the algorithm for solving the problem when creating a top-level model. In this case, the time axis is divided into intervals or time steps. In each such interval, the quantities that depend on time are assumed to be constant. The problem is reduced to solving a nonlinear system of equations at each time step. For this purpose, the method of simple iterations is used, which reduces the solution of a nonlinear system to a multiple solution of a linear system of equations. When the convergence of the iteration cycle is achieved, the period to the next time interval passes.

It should be noted that the applied solution algorithm is traditional and is described in detail in a number of fundamental works [3,5]. Let us dwell in more detail on some of the details that distinguish this implementation from a number of similar studies.

To divide the time axis into intervals, an algorithm for automatic adaptation of the time step is used. The need to use such an algorithm follows from the duration of the modeled process and the desire to obtain sufficiently accurate solutions with reasonable costs of machine time.

The essence of the algorithm is that the time step value is selected depending on the speed of change of the solution over time. This algorithm limits the change of the solution in the time interval, allowing to consider the moments of the fastest changes of the process parameters. When the process reaches the stationary stage, the time step increases sharply, thereby increasing the overall efficiency of the calculation.

VI. Discussion of results

For the model under consideration to function, it is necessary to specify the rheological properties of the deformed medium and the shell material. At the first stages of the HIP process, the main influence on the onset of plastic deformations of the mold-powder system is exerted by the mold itself, which is determined by the level of the mechanical properties of the material. Accurate knowledge of the rheological properties of the shell material is of great importance for an adequate description of the HIP process at its initial and subsequent stages.

Rheological properties of stainless steel 12X18H10T, widely used for manufacturing press molds, were determined. Experiments were conducted on pressing cylindrical samples $d = 9$ and $l = 14\ mm$ in the following temperature-speed modes: mm/min. The specified values for the given sample sizes gave an initial deformation rate in the range of $6{,}2\cdot10^{-5}$-$1{,}2\cdot10^{-1}$ c^{-1}, which corresponds to the real parameters of the HIP.

Fig. 1 shows the dependences of the yield strength on the test temperature at different initial deformation rates, and Fig. 2 and 3 show the hardening curves at different values of temperature and deformation rates. The general course of the curves is quite traditional, but their value is determined by the comparatively low values of the speed parameters.

Mathematical processing of the test results allowed them to be presented in the form

$$T = \tau_S(1 + \alpha H^{\beta} + \gamma L^{\delta}) \cdot e^{R\theta} \tag{11}$$

where H- is the intensity of deformation rates; L- is the degree of deformation; τ_S -is the yield strength for shear. This equation was later used in the complex of programs for the HIP process. For stainless steel grade 12X18H10T, the coefficients in formula (11) have the following values: $\tau_S = 802{,}5$ MPa; $\alpha = 4{,}8$; $\beta = 0{,}5$; $\gamma = 11{,}1$; $\delta = 0{,}75$; $R = -0{,}003$.

The upper-level model describes the rheological properties of a deformable medium under the ellipsoidal plasticity condition and will be discussed in more detail in the next publication when compiling the lower-level model.

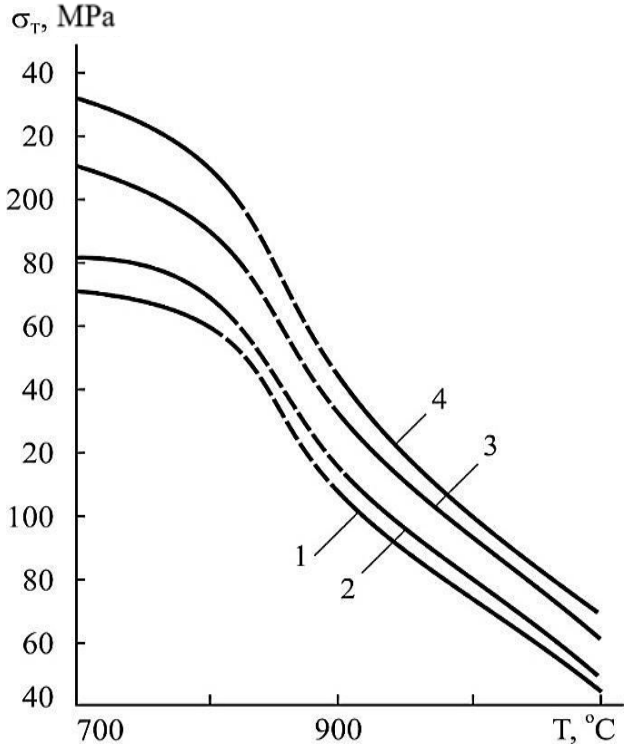


Figure 1: *Dependence of the yield strength of 12X18H10T steel on temperature at deformation rate $6{,}2\cdot10^{-5}$ (1), $1{,}2\cdot10^{-4}$ (2), $6{,}2\cdot10^{-4}$ (3), $1{,}2\cdot10^{-3}c^{-1}$ (4)*

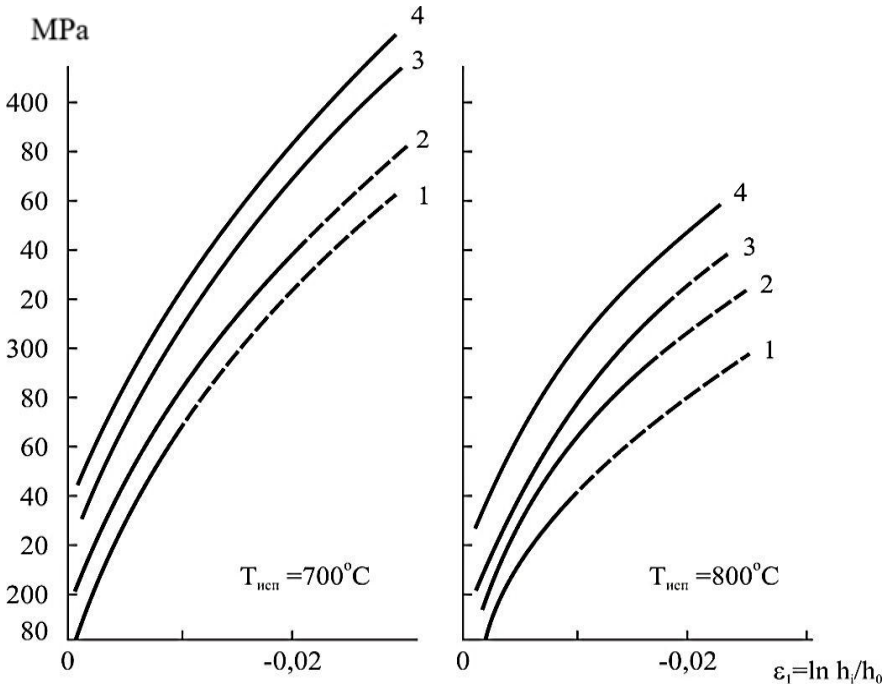


Figure 2: *Hardening curves of 12X18H10T steel at a deformation rate of 0.05 mm/min (1), 0,1 (2), 0,5 (3), 1,0 mm/min (4)*

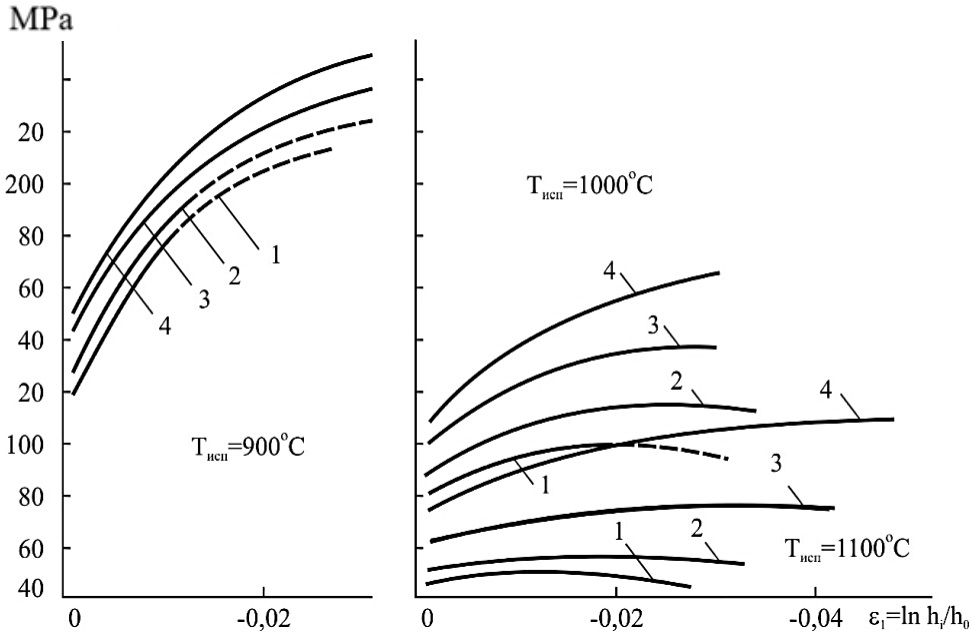


Figure 3: *Hardening curves of 12X18H10T steel at deformation rates of 0,05 (1), 0,1 (2), 0,5 (3), 1,0 mm/min (4)*

VII. Conclusion

Thus, to obtain an adequate mathematical model, it is necessary to more accurately specify the rheological properties of powder materials, which largely determine the compaction kinetics during HIP. The solution of this problem by the finite element method for products of complex configuration with nonlinearity of the used relationships and non-stationary nature of deformation

requires the use of iterative procedures in the presence of a large number of finite elements at each time step.

Therefore, the upper-level model does not allow solving the problems of optimization and optimal control of the HIP pressing process. To overcome these difficulties, it is possible to use the lower modeling level, including zero-dimensional and one-dimensional. HIP models.

A system of technological modes of hot isostatic pressing of powder materials is proposed, which provides for discrete and continuous mathematical modeling of HIP. Discrete and continuous representation of the HIP technological process design system structurally includes the interaction of upper and lower level mathematical models.

When using the upper-level model to obtain an adequate mathematical description, it is necessary to know the rheological properties as accurately as possible, mainly determining the compaction kinetics during HIP. This problem can be solved by the finite element method for products of complex configuration.

To overcome the difficulties associated with the use of the upper-level model, it is necessary to apply the lower modeling level, including zero-dimensional and one-dimensional HIP models. In this case, under conditions of all-round compression, the stress, strain and strain rate tensors are spherical tensors.

Within the framework of the proposed system, the upper-level model can be used in a dialog mode, allowing, by specifying the initial data of the HIP process, to obtain complete information on its progress and results, and, if necessary, to correct one's actions. The use of lower-level models with the use of the optimization apparatus allows one to significantly narrow the search area for technological solutions, thereby obtaining the most reliable and accelerated information.

This work was supported by the Azerbaijan Science Foundation - **Grant № AEF-MGC-2024-2(50)-16/01/1-M-01**

References

[1] Vydrin V.N., Smolin A.P., Kraynov V.I. et al. Installation for studying the deformation resistance of metals and alloys during rolling // Steel, 2010, No. 2, pp. 85-87.

[2] Druyanov B.A. Applied theory of plasticity of porous bodies. Moscow: Mechanical Engineering, 2009, 168 p.

[3] Zenkevich O., Mergan K. Finite elements and approximation. Moscow: Mir, 2006, 318 p.

[4] Laptev A.M. Calculation of compaction processes of porous materials based on creep theory // Technology of light alloys, 2012, No. 7, pp. 16-20.

[5] Spraying high-speed steel with water // A.F. Sanin, D.M.Karikos, V.I.Kalinichenko, V.A.Domoratsky // Powder Metallurgy. - 2003, No.12, pp.1-3.

[6] Skorokhod V.V., Shtern M.B., Martynova I.F. Theory of nonlinear-viscous and plastic behavior of porous materials // Powder Metallurgy, 2007, No.8, pp.23-30.

[7] Abonaf M., J.L.Chenot, G.Raisson, P.Bauduin. Finite element Simulation of hot isostatic pressing of metal powder. 2010, I bid, p.191-212/.

[8] Zienriewiez O.C. The fimite element method. London: Me Grw. Hill, 2007, 787p.

[9] Microstructural development and densitication during hipping of ceramics and metals/W.A.Kaysser, M.Aslan, E.Arztet al // Powd. Met., 2008, 31, I bid, №1, p.63-69.

[10] Shima S., Oyane M. Plastisity theory for parous metals//Int. J.Mech,Sci, 2006, 18, №6, p.285-291.

SENSOR-INTEGRATED MACHINE ELEMENTS - A BRIEF OVERVIEW OF PRODUCTS AND CURRENT TECHNICAL DEVELOPMENTS

Andreas Bürger[1], Sylvio Simon[1],
Pascal Dirk Fritzsche[1], Shalala Hasanli[1]

•

[1] BTU Cottbus - Senftenberg, Universitätsplatz 1, 01968 Senftenberg, Germany
andreas.buerger@b-tu.de, sylvio.simon@b-tu.de,
pascaldirk.fritzsche@b-tu.de, hasansha@b-tu.de

Abstract

Digitalisation offers enormous potential for significant innovations. In order to utilise this potential, comprehensive and reliable data on the operating status of machines is required. Almost every machine contains standardised machine elements (screws, bearings, gears, seals, etc.) in the immediate vicinity of the process and thus offer the possibility of obtaining and evaluating process-related measurement data by integrating sensor systems into these machine elements. A brief overview of commercially available sensor-integrated machine elements and a presentation of some current developments will be given. In addition, solutions for the power supply and data transmission of sensor-integrating machine elements are also briefly presented.

Keywords: Machine elements, energy harvesting, bearings, seals, screws, couplings, splined shafts, idlers

I. Introduction

A machine element is defined as the smallest component in technical applications that can no longer be meaningfully dismantled. These are components or construction principles that fulfil the same or similar functions in different machines and devices and therefore always occur in the same or similar form. They fulfil certain functions and may be moving or stationary parts. They are made of different materials such as metal, plastic or composite materials, depending on the load-bearing capacity and durability requirements. Machine elements include, for example, screws, nuts, seals, couplings, shafts or bearings. [1]

Modern technologies make it possible to integrate sensors and other electronic components into machine elements without restricting their handling or function. This turns the machine element into a *sensor-integrating machine element*, or *SiME* for short [2].

Electronic components require electrical energy and the use of a SiME only makes sense if the data obtained can be transmitted for further use. Cable connections for energy and data transmission are technically simple, but more complex to install. Wireless data transmission and the generation of the required power by the SiME itself is technically possible and greatly simplifies the installation, but is technically more complex.

II. Measured variables and sensors

Suitable and usually highly miniaturised sensors are available for most of the measured variables relevant to machine elements, e.g. speeds, forces, pressures, strains, accelerations and temperatures. Further information, e.g. frequencies, spectra etc., can be easily derived from the measured variables. For this reason, the acquisition and processing of measured variables will not be discussed further here. The measurement principles and the algorithms for processing and deriving other measurements differ little or not at all from those used in conventional measurement systems.

III. Data transmission

As wired data transmission is often associated with installation work in the machine, only wireless data transmission options are discussed here.

Machines usually consist to a large extent of metallic structures, covers, etc. which makes the transmission of data by radio difficult. Ensuring reliable communication may require complex modifications, e.g. to the machine housing.

Suitable radio transmission protocols include WiFi [3], LoRaWAN (Long Range Wide Area Network) [4], Bluetooth Low Energy (BLE) [5], Zigbee [6], Mioty [7], Sigfox [8] or NB-IOT [9].

Which protocol is the most suitable for a specific application depends on various criteria, which may have different priorities depending on the application.

- Possible criteria can be, for example
- the required range
- the amount of data to be transmitted
- Transmission frequency
- the type of technical infrastructure into which the SiME must be integrated

Ultrasonic waves can penetrate metal with low attenuation. This approach is used successfully in some applications. [10]

Optical data transmission is unsuitable in the vast majority of cases due to the required line of sight and sensitivity to contamination.

IV. Energy supply

The SiME should fulfil its function in the machine reliably and maintenance-free, at least over the service life of the integrating machine element. A secure power supply is essential for this. A cable connection requires increased installation effort and the power supply via battery or accumulator requires regular battery changes or recharging and is therefore not maintenance-free.

An elegant option for maintenance-free energy supply is energy harvesting, the ‘harvesting’ of small amounts of energy and the operation of the SiME when sufficient energy has been ‘harvested’. [11]

The possible energy sources are

- thermal energy, to be collected via thermocouples, so-called thermogenerators
- mechanical energy, to be collected via piezo elements, unbalance generators or similar inductive systems
- Optical radiant energy, to be collected via photoelectric systems (solar cells)

V. Selection of commercial SiME

PiezoBolt load cells (Consenses GmbH). The data of the sensor bolts from Consenses GmbH correspond to the usual bolt standards, but are fully-fledged force sensors and are available in

sizes M12, M16 and M20. The screws are supplied with energy and the measured values are transmitted via the plug connection in the screw head. The force is measured in real time as long as the plug connection is in place, which also enables dynamic force curves to be recorded. However, the need for a cable connection restricts the field of application. [12]

Figure 1: *Force measuring screw PiezoBolt PB12 from Consenses GmbH with plug [12]*

Elastomer components with integrated DELTA-C® sensors. DELTA-C® technology works according to the capacitive principle and enables force measurement in conventional elastomer components such as elastomer bearings, couplings or seals. The sensors can be integrated into components that are already required and loads can be measured directly in the force path. The data and energy transmission is cable-bound and therefore restricts the field of application. These are prototypes that are further developed on behalf of customers. [13]

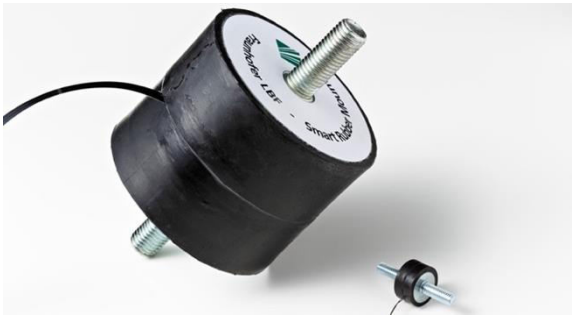


Figure 2: *Elastomeric bearings with integrated force sensors [13]*

Intelligent coupling from R+W Antriebselemente. The company R+W Antriebselemente has further developed its couplings into intelligent couplings in which the complete measuring, processing and transmission electronics for measuring torques, speeds or temperatures are integrated. Power can be supplied by rechargeable batteries or externally by inductive coupling. Data is transmitted wirelessly and can be sent simultaneously to a mobile device or via a gateway to the machine control system. [14]

Figure 3: *Schematic of the intelligent coupling of R+W Antriebselemente [14]*

VI. The DFG-Focus programme 2305 "Sensor-integrating machine elements"

As part of the focus programme 2305 of the Deutsche Forschungsgesellschaft (DFG), the scientific basis for sensor-integrating machine elements and their methodically supported conceptual design and system integration are to be researched. The focus is on 'ordinary' machine elements as standardised basic elements of mechanical engineering with a defined shape and design, which generally cannot be dismantled non-destructively without losing their primary function. [15]

Some of the projects in the priority programme are briefly presented below. The projects were selected and ranked without prioritising them.

I²G - Integral instrumentation of gas foil bearings. A gas foil bearing is being researched and developed as a sensor-integrated machine element that determines the physical measured variables of temperature, acceleration, structure-borne noise and airborne noise in-situ without restricting the primary functionality. In addition to these directly measured variables, the speed is estimated via a Kalman filter and the lift-off and friction conditions are determined. This makes it possible to monitor the condition of the bearing and to record data for monitoring the operation of the bearing-mounted rotor system. The wireless sensor network consists of a base station and a sensor node that is as energy self-sufficient as possible and is integrated into the gas foil bearing. For energy self-sufficient operation, the different energy harvesting approaches - solar cell, thermoelectric generator, piezoelectric foil and piezoelectric vibration generator - are being analysed and, if necessary, combined. For some selected sensor concepts, a first concrete implementation in a prototype with an experimental proof of function in the bearing test bench will be carried out. [16, 17]

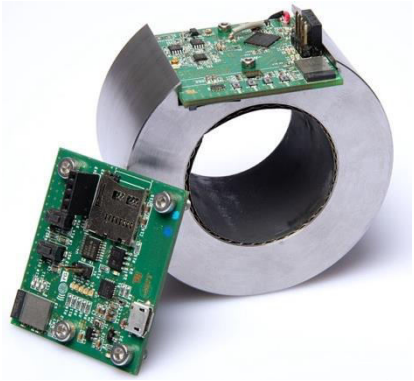


Figure 4: *Gas foil bearing with the first version of the sensor system [16]*

Load-sensitive splined shaft with sensory material. The principle of the sensory material is based on a structural transformation from metastable austenite to martensite when stressed above a limit stress which can be detected by eddy current testing. The sensory material stores the load information, does not require an electrical power supply and can be read out at any interval. As a measuring point, the sensory material can be individually adjusted locally by means of laser heat treatment. The research focuses on the methodical investigation of the constructive integration, the analysis of the effects of sensor integration on the machine element and the qualification of the sensory material. An energy-efficient eddy current readout unit that can be operated autonomously by means of energy harvesting is used to provide proof of function on splined shaft connections in the test rig. [18, 19]

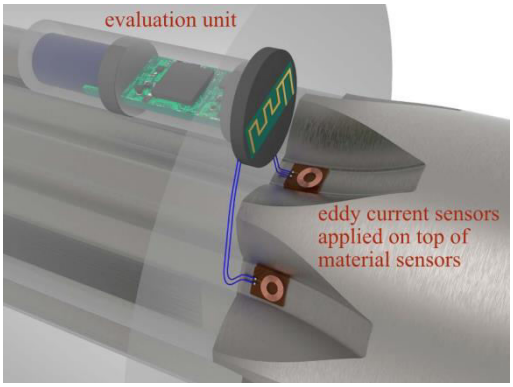


Figure 5: *Concept of the load-sensitive splined shaft with sensory material [19]*

SiSmaK - Sensor-integrating screws for multi-axial force measurement and derivation of a design methodology for sensor integration in closed cylindrical machine elements. Fasteners such as screw connections are particularly suitable for measurement at process-relevant points in and on the machine, as these are located directly in the force flow. Multi-axial force measurement is indispensable for recording loads from different directions and utilising them for process monitoring with early fault detection. In this project, a design methodology for cylindrical machine elements is being researched that addresses the aspects of sensor integration, energy management and signal transmission. The aim of this project is to design and solve the interdisciplinary research questions of a sensor-integrating screw with multi-axial force measurement that fulfils the requirements of installation space neutrality and a self-sufficient energy supply with hermetic sealing. At the same time, the primary function of load-bearing capacity should be affected as little as possible. The procedure for realising this goal should also be used to expand existing design methods for mechatronic systems, especially for the development of sensor-integrating machine elements. [20, 21]

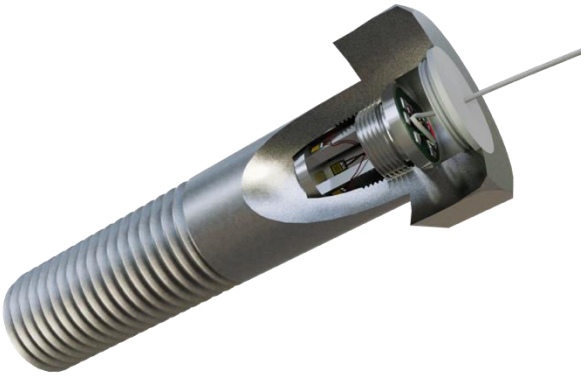


Figure 6: *Sensor-integrating screw of the project SiSmaK [21]*

The sensor-integrating labyrinth seal. As passive support and guide elements of the conveyor belt, idlers are an essential component of belt conveyor systems. An idler consists of an axle, an idler shell, which is rotatably mounted on the axle by two bearings, and two labyrinth seals that protect and seal the inside of the idler and the bearings against dirt and moisture from the outside. Regular visual and acoustic checks with and without measuring devices are standard. Direct monitoring of one or more parameters of each individual idler roller is technically possible, but very time-consuming.

The authors at the BTU Cottbus - Senftenberg are developing a system that records measured values inside an idler roller on the labyrinth seal and transmits them wirelessly to an external receiver module. The required energy is obtained from an inductive system by means of energy harvesting. [22, 23]

The system was used for initial testing in an idler roller in the idler roller test rig (Fig. 9) to measure the heating of a labyrinth seal in the idler roller. The measuring module installed inside the idler roller is shown in Fig. 7.

It is able to measure 4 temperatures on the inner labyrinth sealing ring using PT100 sensors (Fig. 8). At the same time, a further 4 temperatures are measured on the outer, stationary labyrinth sealing ring. All measured values are recorded and analysed on a PC using MATLAB/Simulink. (Fig. 10)

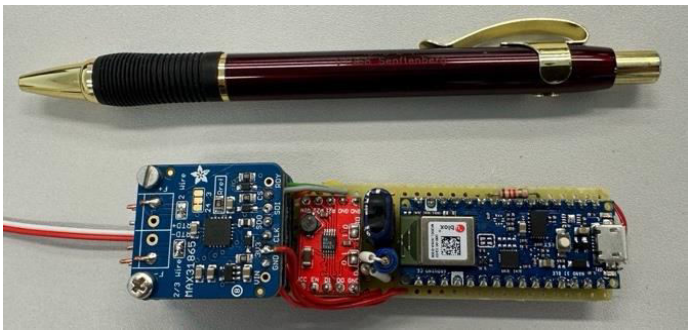


Figure 7: *Measuring module for measuring inside the roller in size comparison with a pencil [24]*

Figure 8: *PT100 temperature sensors idler in size comparison with a match [24]*

Figure 9: *Test rig for the Investigations*

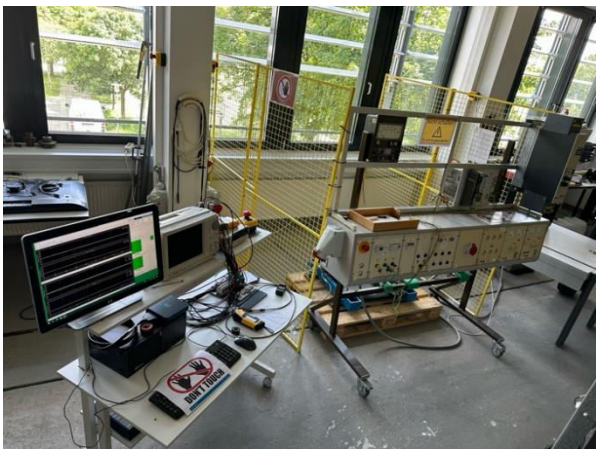


Figure 10: *Complete test stand [24]*

The results of two measurements are shown below. Fig. 11 shows the temperature curve when the labyrinth seal is operated without grease. The temperature increase is only slight but the sealing effect is not fulfilled. Fig. 12 shows the temperature curve when the labyrinth seal is filled with grease. With grease, the labyrinth seal can fulfill its task of protecting the bearing from contamination. The heating is greater due to the friction in the moving grease, but is far below the dropping point of the grease.

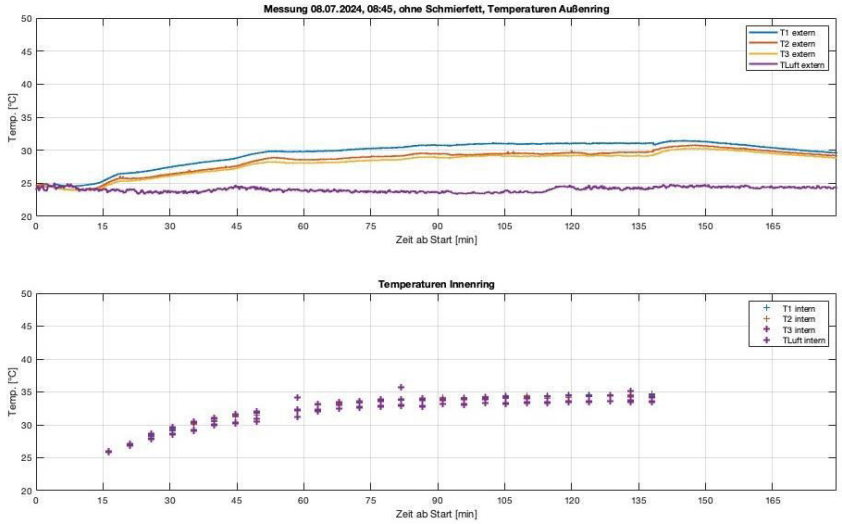


Figure 11: *Temperature curve when operating the labyrinth seal without grease [24]*

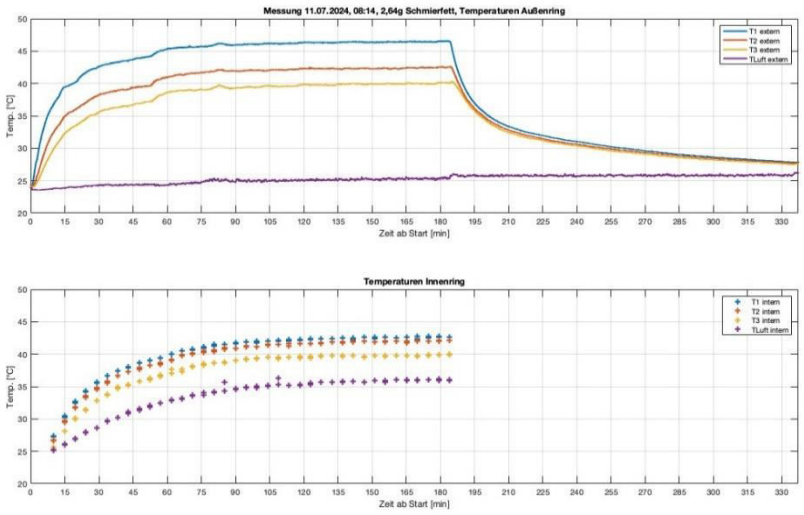


Figure 12: *Temperature curve during operation of the labyrinth seal with grease [24]*

VI. Summary and outlook

The article provides a brief overview of the current market offering of sensor-integrating machine elements without claiming to be exhaustive, and some examples of current research results for SiME are presented.

It can be assumed that with the strong developments in the field of electronics, the field of application of SiME will also grow. New mechanical production methods (coating technologies, additive manufacturing, new joining processes) and modern materials (composite materials, material production) will also accelerate development in the field of SiME.

References

[1] Herbert Wittel, Christian Spura, Dieter Jannasch: Roloff/Matek Machinenelemente. 25. Auflage. Springer, Wiesbaden 2021

[2] Information on https://www.spp2305.de/, retrieved on 04.11.2024

[3] Information on https://en.wikipedia.org/wiki/Wi-Fi, retrieved on 04.11.2024

[4] Information on https://en.wikipedia.org/wiki/LoRa, retrieved on 04.11.2024

[5] Information on https://en.wikipedia.org/wiki/Bluetooth_Low_Energy, retrieved on 04.11.2024

[6] Information on https://en.wikipedia.org/wiki/ZigBee, retrieved on 04.11.2024

[7] Information on https://en.wikipedia.org/wiki/MIoTy, retrieved on 04.11.2024

[8] Information on https://en.wikipedia.org/wiki/Sigfox, retrieved on 04.11.2024

[9] Information on https://en.wikipedia.org/wiki/Narrowband_IoT, retrieved on 04.11.2024

[10] Information on https://en.wikipedia.org/wiki/Ultrasound, retrieved on 04.11.2024

[11] Information on https://en.wikipedia.org/wiki/Energy_Harvesting, retrieved on 04.11.2024

[12] Information on https://consenses.de/en/products/sensors/force-and-vibrationsensors, retrieved on 04.11.2024

[13] Information on https://www.lbf.fraunhofer.de/de/projekte/elastomerbauteile-sensorintegration-delta-c.html, retrieved on 04.11.2024

[14] Information on https://www.rw-couplings.com/sensor-technology/, retrieved on 04.11.2024

[15] Information on www.spp2305.de retrieved on 04.11.2024

[16] Information on https://www.tu.berlin/en/mdt/research/projects/i2g-standardized-integral-instrumentation-of-gas-foil-bearings-for-condition-and-operation-monitoring
retrieved on 04.11.2024

[17] Information on https://gepris.dfg.de/gepris/projekt/466782279?language=en, retrieved on 04.11.2024

[18] Information on https://gepris.dfg.de/gepris/projekt/466760574 retrieved on 04.11.2024

[19] Heinrich, Gansel, Schäfer, Barton, Lohrengel, Maier: An approach to interpreting metastable austenitic material sensors for fatigue analysis, Smart Materials and Structures, June 2024, doi.org/10.1088/1361-665X/ad4f38

[20] Information on https://www.etit.tu-darmstadt.de/must/research_must/projects_must/current_projects_must/index.de.jsp, retrieved on 04.11.2024

[21] Information on https://gepris.dfg.de/gepris/projekt/466650813, retrieved on 04.11.2024

[22] Bürger, A.; Simon, S.; Hernschier, S.: A self-powered sensor with wireless data transmission for use on rotating parts; Tagungsband Energieautonome Sensorsysteme 2024, 12. GMM-Fachtagung, VDE Verlag Berlin, 2024, ISBN 978-3-8007-6299-6

[23] Bürger, A.; Simon, S.; Fritzsche, P.; Hernschier, S.; Fritsch, T.; Magister, J.: Untersuchungen zur Eigenerwärmung von Labyrinthdichtringen in Tragrollen, researchgate.net, 2024, doi.org/10.5281/zenodo.11057426

[24] Bürger, A.; Simon, S.; Fritzsche, P.; Hernschier, S.; Schneider, R.; Magister, J.; Frana, K.: Sensorintegrierte Labyrinthdichtung - Zustandsdiagnose der Lager und Labyrinthdichtungen in einer Tragrolle, September 2024, 28. Fachtagung Schüttgutfördertechnik 2024: TU München

SURFACE TOPOGRAPHY IMPROVEMENT OF 18CRNIMO7-6 STEEL USING THE TAGUCHI TECHNIQUE

Andrzej Dzierwa[1], Anita Ptak[2], Ahmedov Beyali[3], Anar Hajiyev[3]

•

[1]Rzeszow University of Technology, Department of Manufacturing Technology and Production Engineering, Powstancow Warszawy 8, 35-959 Rzeszow, Poland
[2]Wroclaw University of Science and Technology, Department of Fundamentals of Machine Design and Mechatronic Systems, I.Lukasiewicza 5, 50-371 Wroclaw, Poland
[3]Azerbaijan Technical University, Department of Machine Design and Industrial Technologies, H.Javid Avenue 25, AZ 1073, Baku, Azerbaijan
adzierwa@prz.edu.pl, anita.ptak@pwr.edu.pl, ahmedov.beyali@aztu.edu.az
anar.hajiyev@aztu.edu.az

Abstract

The ball burnishing process provides a fast, cost-effective, and straightforward method to enhance the physical-mechanical properties and surface integrity of industrially manufactured parts. In this study, ball burnishing was applied to improve the surface topography of 18CrNiMo7-6 steel, with milling used as a pre-burnishing treatment. A Taguchi L9 orthogonal array was employed to conduct hydrostatic ball burnishing investigations and optimal values of the process parameters have been identified. The positive effect of the burnishing process on a number of surface topography parameters was also demonstrated.

Keywords: ball burnishing, surface topography, roughness, Taguchi technique.

I. Introduction

Surface topography plays a crucial role in mechanical engineering. It considerably affects the friction and wear properties of interacting materials, thereby influencing their tribological behavior [1]. Additionally, surface topography governs surface properties such as adhesion and stiffness, which are very important for material performance [2]. Moreover, variations in surface structure can lead to differences in other mechanical properties, such as fatigue strength and lubrication efficiency [3].

The condition of surface topography can be influenced by most elements of the manufacturing process; however, the finishing treatment has a decisive influence. Depending on the main purpose of the finishing treatment, it can be smoothing or strengthening [4]. One of the types of finishing treatment is the burnishing process (plastic deformation treatment of the surface) used, among other purposes, to achieve low values of surface roughness parameters. The authors of the work [5] applied the ball burnishing process to improve the final quality of form tools (moulds and dies). They reported that by using the optimal input parameters it is possible to reduce the arithmetical mean roughness Ra from 3.01 μm to approximately 0.30 μm, while an initial hardness (HRA) of about 66.35 can be increased to 71.33. The purpose of the research by Rodriguez et al. [6] was to improve of surface topography properties of 2050 aluminum

components with two different heat treatments. The authors concluded that ball burnishing enhanced surface finish, hardness, and fatigue life of components. They also identified optimal parameters for burnishing aluminum alloy components. The positive effect of ball burnishing process on reducing surface roughness parameters and increasing hardness was also noted by the authors of [7, 8]. Dzionk et al. [9] analyzed the changes in surface topography of the shafts after ceramic ball burnishing. They revealed that the application of ball burnishing allowed for the achievement of surfaces suitable for high-performance applications, such as rolling bearings and hub joints. Livatyali [10] indicated that ball burnishing improves surface integrity of thin Ti6Al4V flat sheets by inducing compressive residual stresses. The advantages of compressive residual stresses in the surface layer of components were also highlighted by the authors of [11].

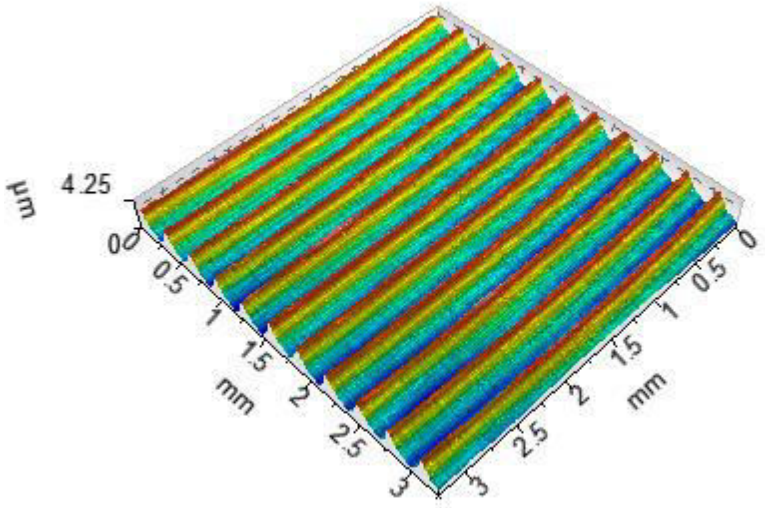
Currently, the surface burnishing process dominates in the machining of rotational surfaces - such as on lathes, boring machines - sometimes effectively replacing grinding or honing. Dynamic burnishing is often used for non-rotary surfaces [12, 13], providing good strengthening effects, but significantly poorer dimensional and smoothness effects. Therefore, the aim of this study was to analyze the effect of hydrostatic ball burnishing parameters on selected geometric structure parameters of flat surfaces after the milling process.

II. Experimental procedure

In the present article, 18CrNiMo7-6 steel with a hardness of 45±2 HRC was chosen to investigate the influence of hydrostatic ball burnishing on surface topography parameters. The burnishing tool with a 10 mm ceramic Al2O3 ball was installed in Haas CNC Vertical Mill Center VF-3. The test samples were shaft sections with a diameter of 25 mm and a height of 8 mm. The medium used for the process was a water-and-oil emulsion, Hysol. Hydrostatic ball burnishing was carried out using the following parameters:

- Burnishing pressure: 10; 20 and 30 [MPa],
- Burnishing velocity: 750; 1500 and 2250 [mm/min],
- Burnishing width: 0.04; 0.08 and 1.2 [mm].

Milling was performed as a pre-burnishing treatment. All samples were milled to achieve an arithmetical mean height of the surface (Sa) of 0.72 μm. The isometric view of the milled surface, along with various surface topography parameters, is presented in Figure 1. A detailed description of these surface topography parameters – skewness (Ssk), kurtosis (Sku), maximum height of the surface (Sz), texture aspect ratio (Str), peak density (Spd), developed area ratio (Sdr), core roughness depth (Sk), reduced peak height (Spk) and reduced valley depth (Svk) – is available in [14, 15]. Surfaces after the milling process were characterized by positive values of skewness which may indicate a more bumpy nature of the surface. We can reach similar conclusions considering the ratio of Spk to Svk – value of reduced peak height is higher then reduced valley depth. Texture aspect ratio (Str) is at the level of 0.0369, which indicates a clear anisotropy of the surface.



Sa	0.72 [µm]
Ssk	0.544
Sku	2.20
Sz	4.28 [µm]
Str	0.0369
Sdr	0.105 [%]
Spd	103 [1/mm²]
Sk	1.72 [µm]
Spk	1.21 [µm]
Svk	0.0999 [µm]

Figure 1: *Isometric view of the base surface and selected surface topography parameters*

Before and after hydrostatic ball burnishing process, the surface topography of all samples was measured using a white light interferometer Talysurf CCI Lite. A 5*x* lens was used in the measurements to obtain a measurement area of 3.3 mm x 3.3 mm. After measurement surfaces were leveled using the TalyMap 6.0 software – digital filtration was not used.

Taguchi L9 orthogonal array was used to perform investigations (Table 1). In G. Taguchi's plans, during the implementation and mathematical analysis of research results, the signal, noise and control

Table 1: *Design of experiment*

Exp. number	Input burnishing parameters and their levels		
	Pressure [MPa]	Velocity [mm/min]	Width [mm]
1	10	750	0.04
2	10	1500	0.08
3	10	2250	1.2
4	20	750	0.08
5	20	1500	1.2
6	20	2250	0.04
7	30	750	1.2
8	30	1500	0.04
9	30	2250	0.08

factors are considered. Noise (process disturbances) takes into account the influence of factors that are beyond the operator's control, while the operator sets the control factors during machining. In ideal conditions, the output signal (surface roughness after burnishing) will only react to the operator's signals and will not react to random changes during machining. In Taguchi technique [16, 17, 18], a loss function is used to calculate the deviation between the experimental value and the desired value. This loss function is then converted into a signal-to-noise (S/N) ratio. Various types of S/N ratios are available, depending on the nature of the characteristics involved:

- lower is better,
- nominal is better,
- higher is better.

In the case of "higher is better" and "lower is better" ratio, the definitions of the loss function for hydrostatic ball burnishing process are as follows:

$$Lower\ is\ better = -10log\left[\frac{1}{n}\right]\sum_{i=1}^{n} y_i^2 \quad (1)$$

$$Higher\ is\ better = -10log\left[\frac{1}{n}\right]\sum_{i=1}^{n} y_i^{-2} \quad (2)$$

where n is the number of measurements and y_i is the measured value. In our study "lower is better" ratio was applied.

III. Results and discussion

Table 2 presents selected surface topography parameters achieved after hydrostatic ball burnishing. The arithmetical mean height of the surface (Sa) decreased significantly as a result of the

Table 2: *Results of the experiments*

Exp. number	Sa [μm]	Sz [μm]	Sdr [%]	Ssk	Sk [μm]	Spk [μm]	Svk [μm]
1	0.253	1.45	0.0239	-0.502	0.732	0.0264	0.31
2	0.378	2.39	0.0373	-0.165	1.23	0.0875	0.262
3	0.43	3.27	0.0511	0.00843	1.66	0.206	0.253
4	0.291	1.79	0.0268	-0.297	0.954	0.111	0.296
5	0.307	2	0.0301	-0.201	0.902	0.182	0.274
6	0.166	1.79	0.0159	-1.07	0.372	0.0947	0.36
7	0.261	1.75	0.0219	-0.123	0.785	0.178	0.287
8	0.051	0.596	0.00402	-0.117	0.145	0.0421	0.0522
9	0.181	1.38	0.0141	-0.123	0.589	0.111	0.164

process with values ranging from 0.051 μm (sample 8) to 0.43 μm (sample 3). A similar trend was observed for the maximum height of the surface (Sz), which, compared to the milling process (Sz = 4.28 μm), was reduced to between 0.056 and 3.27 μm. Skewness (Ssk), which measures the symmetry of surface variation about its mean plane, also decreased and showed negative values in almost all samples, indicating a predominance of valleys. The only exception was sample 3, where Ssk was 0.008. Surfaces after burnishing, like those after milling, were characterized by high anisotropy, with an Str parameter below 0.1. Analysis of changes in the Sk family group showed a visible reduction in the Spk parameter value, while changes in the Svk parameter were significantly smaller. This suggests that modifications to the surface texture occurred primarily in the peak and core areas, with less change in the valley regions. The developed interfacial area ratio Sdr decreased distinctly compared to the milling process reaching values close to zero. Isometric views of selected surface topography after ball burnishing process are presented in Figure 5.

To assess the effect of burnishing process input parameters on surface topography, the signal-to-noise ratio and means were calculated (Tables 3-5). These factors are presented graphically in Figures 2-4. The parameters Sa, Sz and Sdr were selected for detailed analysis. Response tables for S/N ratio and means of Sa, Sz and Sdr indicate the optimal levels of input parameters for minimizing surface topography values. Analysis of the data shows that burnishing pressure and burnishing width play significant roles in reducing surface roughness across all analyzed parameters. It was confirmed for all analyzed parameters. Although the influence of burnishing velocity was smaller, it should not be considered insignificant. The optimal hydrostatic ball burnishing performance for surface topography

Table 3: *Response table for S/N ratio and means of Sa*

S/N ratio				Means			
Level	**Pressure [MPa]**	**Velocity [mm/min]**	**Width [mm]**	**Level**	**Pressure [MPa]**	**Velocity [mm/min]**	**Width [mm]**
1	8.982	11.442	17.795	1	0.3670	0.2683	0.1567
2	12.192	14.852	11.340	2	0.2547	0.2453	0.2833
3	17.454	12.334	9.494	3	0.1643	0.2723	0.3460
Delta	8.472	3.314	8.301	Delta	0.2027	0.0270	0.1893
Rank	1	3	2	Rank	1	3	2

Table 4: *Response table for S/N ratio and means of Sz*

S/N ratio				Means			
Level	**Pressure [MPa]**	**Velocity [mm/min]**	**Width [mm]**	**Level**	**Pressure [MPa]**	**Velocity [mm/min]**	**Width [mm]**
1	-7.029.	-4.382	-1.263	1	2.370	1.663	1.276
2	-5.378	-3.031	-5.141	2	1.860	1.662	1.853
3	-1.054	-6.049	-7.057	3	1.242	2.147	2.340
Delta	5.974	3.017	5.794	Delta	1.128	0.485	1.061
Rank	1	3	2	Rank	1	3	2

Table 5: *Response table for S/N ratio and means of Sdr*

S/N ratio				Means			
Level	**Pressure [MPa]**	**Velocity [mm/min]**	**Width [mm]**	**Level**	**Pressure [MPa]**	**Velocity [mm/min]**	**Width [mm]**
1	28.94	32.35	38.77	1	0.037	0.024	0.015
2	32.61	35.64	32.34	2	0.024	0.024	0.026
3	39.37	32.94	29.62	3	0.013	0.027	0.034
Delta	10.43	3.28	8.96	Delta	0.024	0.003	0.020
Rank	1	3	2	Rank	1	3	2

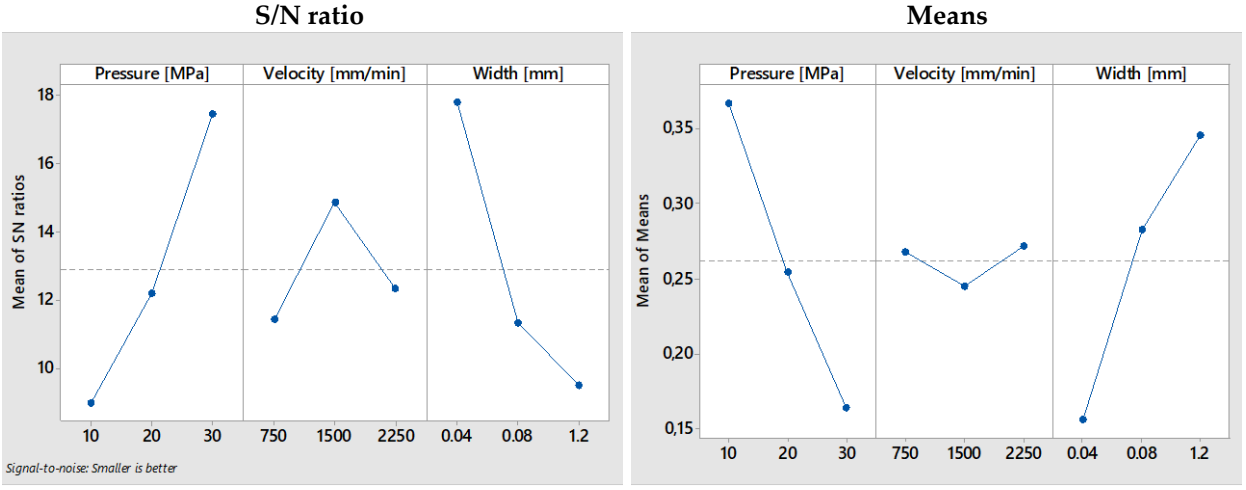


Figure 2: *Main effect plots for the S/N ratio and means of Sa parameter*

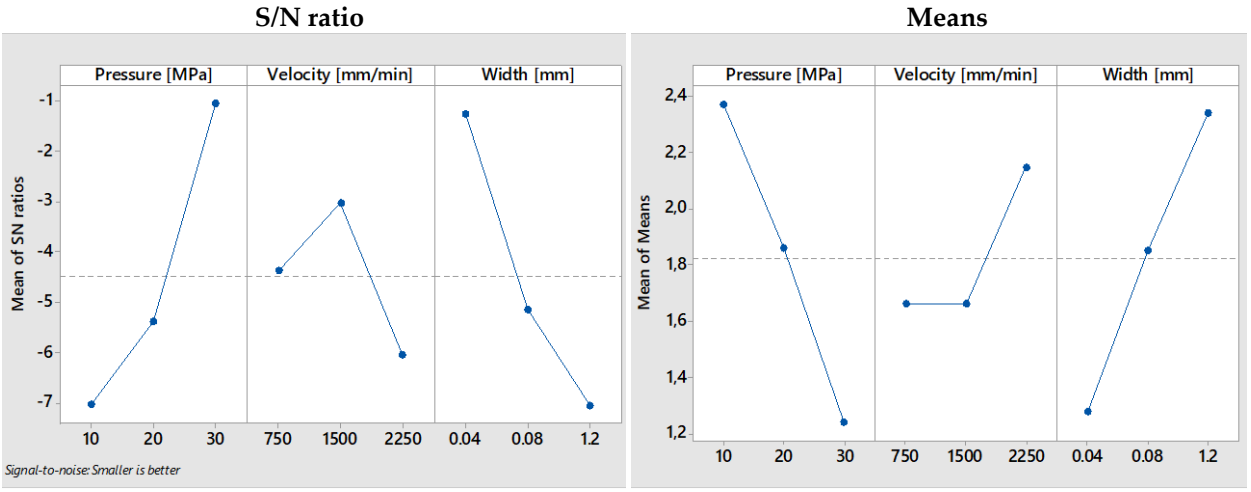


Figure 3: *Main effect plots for the S/N ratio and means of Sz parameter*

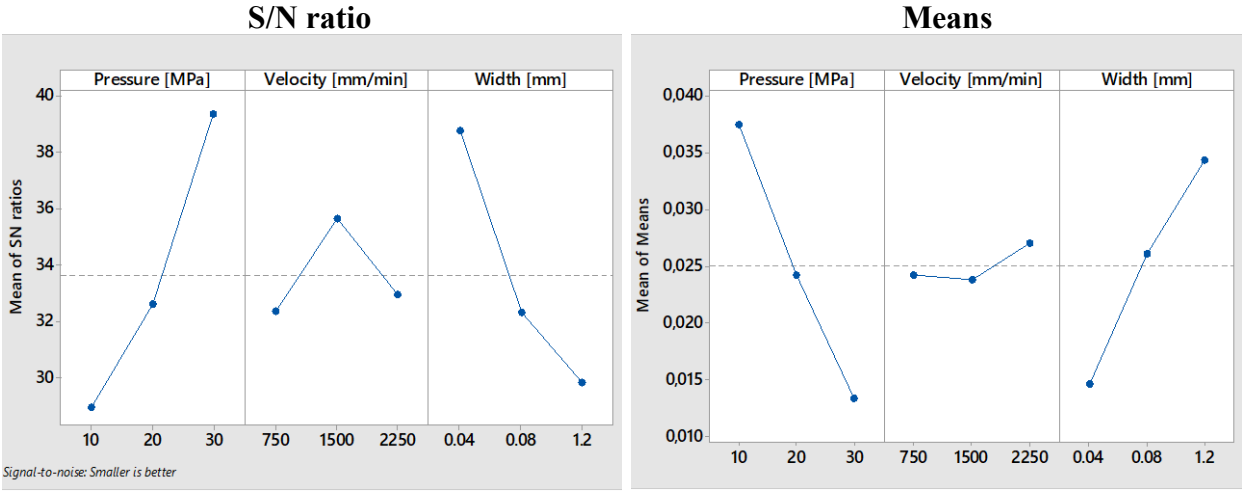
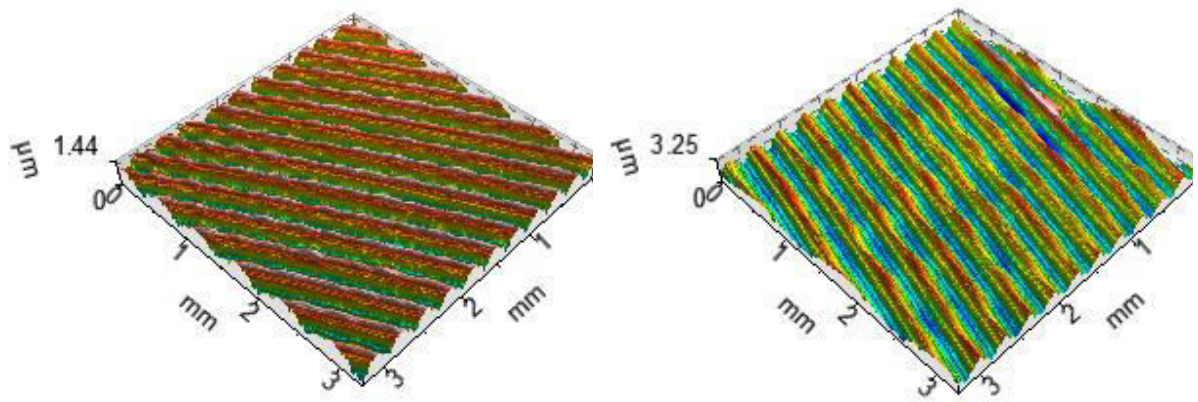


Figure 4: *Main effect plots for the S/N ratio and means of Sdr parameter*
parameters were achieved with the following settings: burnishing pressure – 30 [MPa]; burnishing velocity – 1500 [mm/min] and burnishing width – 0.04 [mm]

Low burnishing pressure causes the burnishing ball to penetrate only shallowly into the metal surface, resulting in plastic flow of the metal. This action reduces surface roughness parameters. In the tests, surface roughness was observed to decrease as burnishing force increased.

a) b)



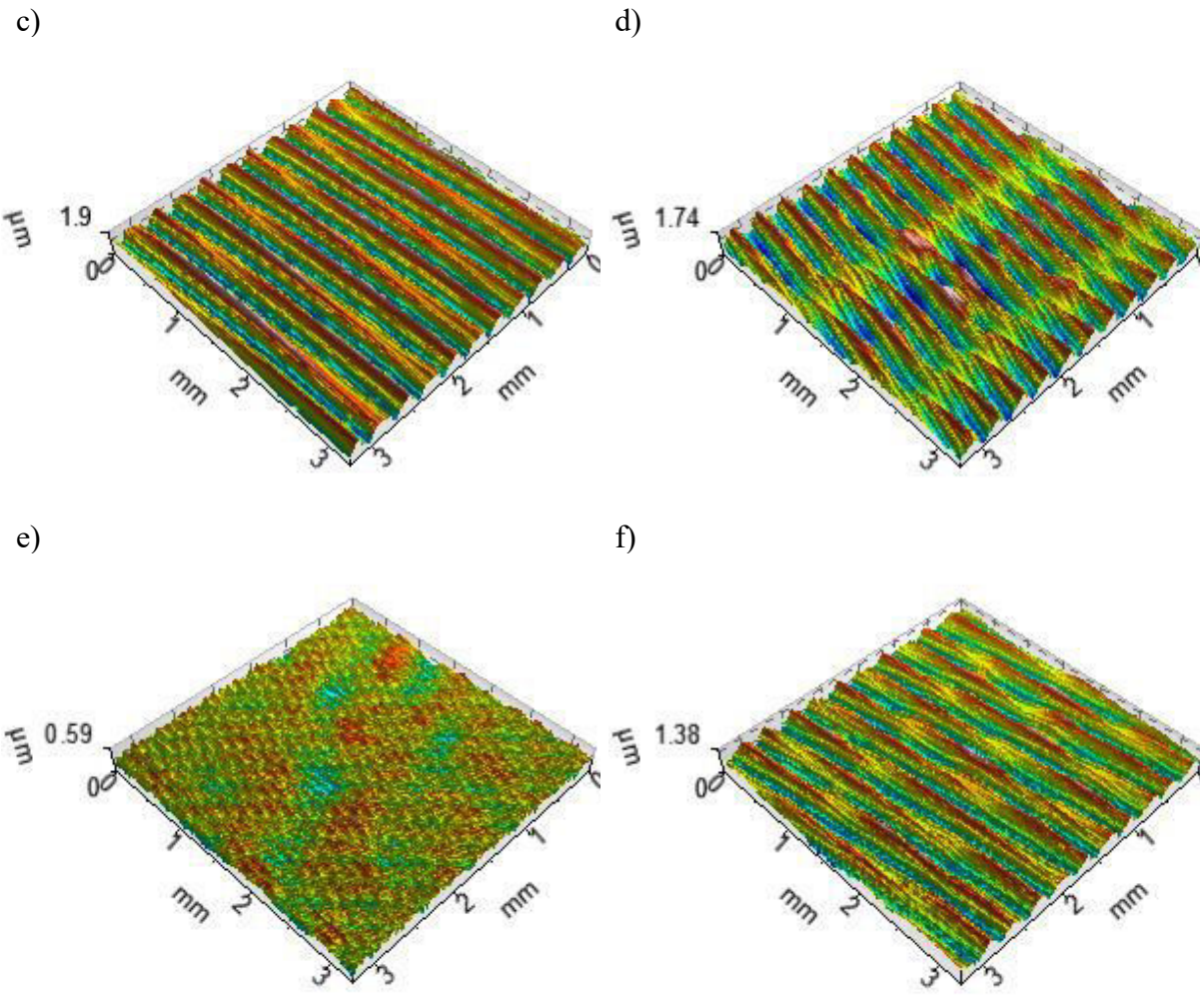


Figure 5: *Isometric views of the selected surface topography after the burnishing process: sample 1 (a), sample 3 (b), sample 4 (c), sample 7 (d), sample 8 (e), sample 9 (f)*

This effect can be attributed to the higher pressure exerted by the ball on the workpiece surface, which compresses most surface irregularities and enhances metal flow. This process fills more of the empty spaces or valleys created in the subsurface layer during processing. It is well known that each material has a specific cold-working capability and a certain work hardening limit. Exceeding this limit would likely lead to an increase in surface roughness. However, based on the tests conducted, it can be concluded that this limit has not yet been reached. Burnishing width was also one of the very important parameters that affect the results of the process. The most favorable width of the machining traces was found to be 0.04 mm. Increasing this width led to higher values in the analyzed roughness parameters. It appears that at the smallest trace width, the distance between machining traces is minimal allowing the tool has sufficient time to smooth the surface effectively. Although the influence of burnishing velocity on the analyzed roughness parameters was smaller than other input parameters, it seems that the most favorable velocity is 1500 mm/min.

IV. Conclusions

This study investigates the hydrostatic ball burnishing process of 18CrNiMo7-6 steel using the Taguchi technique. Based on the analysis of experimental results, the following conclusions can be drawn:

1. The experiments conducted showed that two input factors in the burnishing process – burnishing pressure and burnishing width – significantly influenced changes in surface topography parameters: arithmetical mean height of the surface (Sa), maximum height of the surface (Sz), and developed area ratio (Sdr). In contrast, the effect of burnishing velocity on these parameters was less pronounced.

2. Signal-to-noise ratio analysis revealed that the best surface roughness (minimum values of Sa, Sz and Sdr parameters) was obtained at 30 MPa burnishing pressure, 0.04 mm burnishing width, and 1500 mm/min burnishing velocity.

3. The ball burnishing process applied to the 18CrNiMo7-6 steel produced a new, refined surface layer that improved surface topography. All height parameters were significantly reduced, with the Sa parameter reaching of 0.051 μm in the most favorable variant (sample 6). A positive effect of the burnishing process was also observed in parameters from the Sk family, with the most significant changes occurring in the peak and core parts. All of these beneficial effects were achieved with a single tool pass.

References

[1] A.A.A, P. Onu, A.S.A, M. Stella I, An overview on impact of surface roughness on tribological behaviour, in: 2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), Omu-Aran, Nigeria, 2024, pp. 1-10.

[2] T.D.B. Jacobs, L. Pastewka, Surface topography as a material parameter, MRS Bulletin. 47 (2022) 1205–1210.

[3] W. Bai, Y. Gao, R. Sun, Surface topography and roughness in vibration assisted machining. In: Vibration assisted machining. research on intelligent manufacturing, Springer, Singapore, 2023, pp 113–133.

[4] A. Dzierwa, L. Galda, M. Tupaj, K. Dudek, Investigation of wear resistance of selected materials after slide burnishing process, Eksploatacja i Niezawodność – Maintenance and Reliability. 22 (2020) 432-439.

[5] H. Hamadache, Z. Zemouri, L. Laouar, S. Dominiak, Improvement of surface conditions of 36 Cr Ni Mo 6 steel by ball burnishing process, J. Mech. Sci. Technol. 28 (2014) 1491-1498.

[6] A. Rodriguez, A. Calleja, L.N.L. de Lacalle, O. Pereira, H. Gonzalez, G. Urbikain, J. Laye, Burnishing of FSW aluminum Al–Cu–Li components, Metals. 9 (2019) 260.

[7] T. Li, M. Uddin, G. Stevens, Effect of ball burnishing on surface roughness, wettability, mechanical integrity and antifouling resistance of UHMWPE, Surf. Coat. Technol. 476 (2024) 130237.

[8] S. Swirad, D. Wydrzynski, P. Nieslony, G.M. Krolczyk, Influence of hydrostatic burnishing strategy on the surface topography of martensitic steel, Measurement. 138 (2019) 590-601.

[9] S. Dzionk, B. Scibiorski, W. Przybylski, Surface texture analysis of hardened shafts after ceramic ball burnishing, Materials. 12 (2019) 204.

[10] H. Livatyali, A novel force control strategy for improved surface integrity in low plasticity burnishing. Journal of Engineering Sciences (Ukraine). Vol. 10 (2023) A18–A26.

[11] R. Teimouri, S. Amini, A.B. Bami, Evaluation of optimized surface properties and residual stress in ultrasonic assisted ball burnishing of AA6061-T6. Measurement. 116 (2018) 129–139.

[12] M. Bucior, R. Kluz, T. Trzepiecinski, K. Jurczak, A. Kubit, K. Ochał, The effect of shot peening on residual stress and surface roughness of AMS 5504 stainless steel joints welded using the TIG method, Materials. 15 (2022) 8835.

[13] M. Korzynski, A. Dzierwa, A. Pacana, J. Cwanek, Fatigue strength of chromium coated elements and possibility of its improvement with ball peening, Surf. Coat. Technol. 204 (2009) 615–620.

[14] A. Dzierwa, R. Reizer, P. Pawlus, W. Grabon, Variability of areal surface topography parameters due to the change in surface orientation to measurement direction, Scanning.36 (2014) 170–183.

[15] PN-EN ISO 25178-2:2012; Geometrical Product Specifications (GPS)—Surface texture: Areal—Part 2: Terms, Definitions and Surface Texture Parameters. International Organization for Standardization: Geneva, Switzerland, 2012.

[16] D.C. Montgomery, Design and analysis of experiments, 8th ed.; John Wiley & Sons, Inc. Hoboken, NJ, USA, 2012.

[17] T. Kıvak, Optimization of surface roughness and flank wear using the Taguchi method in milling of Hadfield steel with PVD and CVD coated inserts, Measurement. 50 (2014) 19–28.

[18] M.H. El-Axir, A.A. Ibrahim, Some surface characteristics due to center rest ball burnishing, J. Mater. Process. Technol. 167 (2005) 47-53.

MATHEMATICAL MODEL OF THE SINTERING PROCESS OF A MIXTURE OF MESOSCALE WC-CO POWDERS

Arif Mamedov[1], Aqil Babayev[1], Mukhtar Huseynov[1], Beture Musurzayeva[1]

•

[1]Azerbaijan Technical University
arif.memmedov@aztu.edu.az, aqil.babayev@aztu.edu.az, muxtar.huseynov@aztu.az, beture.musurzayeva@aztu.az

Abstract

The paper proposes models of the non-isothermal (solid-phase) sintering stage and isothermal (liquid-phase) sintering. When constructing the sintering process model, assumptions were made about the structure of the solid alloy, the mechanisms of its compaction and grain growth.
Based on literature data, a system of equations for the mathematical description of non-isothermal solid-phase sintering of WC-Co was constructed. The equations of the compaction kinetics at the j-m temperature stage and at the first 3.5% shrinkage, as well as the equations of the grain growth kinetics and temperature increase, are presented. The coefficients of volume grain boundary and surface diffusion were calculated.
A system of equations for the mathematical description of liquid-phase (isothermal) sintering of WC-Co under pressure in a vacuum is proposed. The proposed system considers the following equations: compaction kinetics and rheological model of a porous material, as well as the equation of the alloy grain growth kinetics. The relationships of the quality index which depends on the residual porosity, average grain size and density of the sintered hard alloy are derived.
The general equation of compaction kinetics j-m in the temperature mode, as well as the control equation of the product quality index and the characteristics of the WC-Co synthesis process with a mesostructure are obtained.

Keywords: hard alloy, mathematical model, solid-phase sintering, hot-phase sintering, rheology, compaction.

I. Introduction

In general, sintering is divided into three stages [1,2]: 1) temperature increase – heating (non-isothermal sintering); 2) holding at a constant temperature (isothermal sintering); 3) temperature decrease – cooling.

At the heating stage, sintering is carried out by a diffusion (solid-phase) mechanism. Solid-phase sintering (SPS) of a powder body occurs without the formation of a liquid phase, with the following main processes occurring: volume and surface diffusion of atoms, shrinkage; recrystallization of a metal powder body (growth of some grains at the expense of others of the same phase); transfer of atoms of a substance through the gas phase due to volume and surface diffusion, viscous flow and flow caused by external loads during sintering under pressure. This type of sintering is accompanied by the effect and development of bonds between particles, the

formation and growth of contacts (necks), the "healing" of pores (closing of through porosity) with their enlargement and spheroidization, compaction of the workpiece and its shrinkage, which occurs during heating mainly due to the volumetric deformation of particles, carried out by volumetric self-diffusion of atoms and adsorption of atoms (adatoms) on the surface of the substance/surface diffusion).

During liquid-phase sintering (LPS), an alloy is formed, one of the components of which is in the liquid state. Therefore, the question of the mechanism of diffusion displacement, or interaction, of the solid and liquid phases is of fundamental importance. The diffusion of atoms from the liquid phase to the solid at the first stage of their interaction should cause an increase in the volume of the component particles forming the basis of the powder body, the movement of their geometric centers and, consequently, an increase in the dimensions of the entire body. The subsequent dissolution of the solid phase particles in the liquid should be accompanied by a decrease in their volume, the convergence of their centers as a result of the action of capillary forces and, as a consequence, shrinkage of the powder body. Thus, after the formation of the liquid phase, the powder body during sintering should, in general, first exhibit growth and then shrinkage [1].

However, it should be borne in mind that the indicated successive stages of growth and shrinkage of the powder body must take place with low solubility of the liquid phase component in the solid or significant content of the liquid phase. Otherwise, the liquid phase quickly disappears, since due to the finely dispersed solid phase, the area of its surface through which diffusion occurs becomes very large and the powder body will only experience growth.

Since the solubility of tungsten carbide in cobalt in the solid state is about 10%, the dissolution process must be completed before the appearance of a eutectic liquid with a melting point below the melting point of cobalt. Therefore, sintering of the WC - Co alloy is characterized by significant compaction even before reaching the temperatures of the liquid phase appearance and rapid (within a few minutes) almost complete compaction after the liquid phase appears. Compaction before the liquid phase appears is carried out in solid phases (Co and WC) and is accompanied by diffusion of WC into Co, with the formation of solid solutions (intermetallics) [3]. The process of "liquid-viscous" flow leads to almost complete compaction of the sintered body, which is true in the presence of liquid of at least 20 ... 35% (of the total volume); liquid cobalt melt flows into the pores and pulls together the WC grains.

At a sintering temperature exceeding the melting point of the eutectic in the WC-Co system, a density close to the theoretical one is achieved in a very short time (1...4 min); during sintering in the absence of a liquid phase (below 1300°C), the final density is lower than the theoretical one even with a sintering duration of 30 min and its value depends on the temperature. As the results of experimental studies show, the degree of shrinkage for fine-grained powder is higher both during sintering in the solid phase and in the presence of a liquid one. The mechanism of WC grain growth during sintering of the WC-Co alloy has several variants [3]. Firstly, due to the precipitation of WC during cooling from the Co ◎ solid solution. An examination of the data on the interaction of tungsten carbide with cobalt shows that during sintering, significant amounts of tungsten are first dissolved and the resulting liquid phase contains about 40% WC. Saturation of the liquid occurs due to the dissolution of some of the grains. When cooling, almost all of the WC contained in the liquid precipitates on undissolved tungsten carbide crystals as crystallization centers. However, as a result of this process, one cannot expect a significant increase in the size of the WC crystals.

If the content of Co in the WC-Co alloy is, for example, 10%, the latter is capable of dissolving about 6% of the WC present in the alloy during sintering. With uniform distribution of this amount of WC (when it precipitates from the liquid) on the remaining undissolved WC crystals, the size of the latter may increase slightly. Secondly, the growth of WC grains occurs due to recrystallization through the liquid phase. This is the main mechanism of WC phase growth. With an increase in the

cobalt content in the alloy, a more noticeable increase in the size of WC grains is observed, their maximum growth is achieved with an increase in the cobalt content from 1 to 6%, and with a large amount of cobalt, the growth becomes less due to an increase in the path of crystal transfer through the liquid due to an increase in the thickness of the interlayers. Some growth of the WC phase is also possible due to the coalescence mechanism, similar to the collective recrystallization of grains.

II. Mathematical model of the sintering process and quality control of WC-Co hard alloy from a mixture of meso-sized powders

As stated above, the sintering process is divided into three stages. Let us construct mathematical models of the first two stages: non-isothermal (solid-phase) sintering (i=1) and isothermal (liquid-phase) sintering (i=2). For solid-phase sintering (SPS), we will use the general equation (1) as the differential equation for porosity, which is valid throughout the entire solid-phase sintering stage, and for the first 3.5% of shrinkage at this stage, we can use equation (2).

$$\frac{dP_c}{dt} = \frac{3(1-P_c)}{1-P_0}\frac{d_y}{dt} \tag{1}$$

$$\frac{dP_c}{dt} = \frac{3(1-P_c)^2}{1-P_0}\cdot\frac{1}{\left[\frac{P_0-P_c}{3(1-P_c)}\right]^{2,1}} \quad \left\{\frac{2\sigma\Omega D_v}{RTr_g^2}\cdot\frac{P_0-P_c}{3(1-P_c)}+\frac{\varphi\Omega bD_b}{2KTr_g^a}\right\} \tag{2}$$

For liquid-phase sintering (LPS), the kinetics of compaction will be analyzed based on the rheological description of the deformation of the porous structure, considering the diffusion-viscous flow and compaction of a porous body under a uniform stress state under conditions of all-round compression [4]. This formulation of the problem is directly related to the sintering process under the simultaneous action of both the Laplace (capillary) pressure N_k and the hydrostatic pressure N_g applied from the outside (which is summed with the Laplace:

$$N=N_k+N_g$$

Dimensionless volumetric deformation under all-round compression is described by the relation [5]:

$$\Delta V/V = \varepsilon_{ii} = -N/\chi \tag{3}$$

where V- is the volume of the sintered porous body; ε_{ii} - is the volume of the body after sintering; χ - is the sum of the diagonal terms of the strain tensor; is the coefficient of bulk viscosity. Based on the hydrodynamic analogy of the theory of elasticity (it is on this analogy that the modern version of the rheological theory of sintering is based [6]), the following equation for the viscous flow of a compressible body can be written [4]:

$$V = \varepsilon_{ii} = 3\varepsilon_{rr} = -N/\chi \tag{4}$$

since in the above formulation of the problem the strain rate tensor is spherically symmetric and is reduced to one radial component $\dot{\varepsilon}_{rr}$. Taking into account the obvious connection $V = V_b(1-P)$ between the volume of a porous body and the volume of the substance enclosed in it (without pores) V_b, we obtain:

$$\dot{V} = \dot{p}(1-p) \tag{5}$$

where p - is the porosity of the body (dimensionless quantity).

The hydrodynamic analogy of the theory of elasticity gives grounds to believe that, when applied to an isotropic medium, the viscosity tensor has two components: χ – the coefficient of bulk viscosity and η – the coefficient of shear viscosity; they are analogous to the modulus of uniform compression and shear, respectively.

In the rheological description of the porous structure (χ,η), the Mackenzie-Shuttleworth model is adopted [7,8], according to which the void phase can be represented as localized in an ensemble of non-interacting pores of equal size, spaced so far apart that it is possible to introduce into consideration an element of the porous body consisting of a pore of radius R, surrounded by a layer of an incompressible body with radius R°, while $p = (R/R^{\circ})^3$.

By hydrodynamic analogy, the relationship between χ and η has the form

$$\chi = 4\eta\ (1-p)/3p \tag{6}$$

Taking into account the dependence of the shear viscosity coefficient η on porosity, studied in works [6,8]

$$\eta = \eta_0\,(1-p)^{5/3}\left[pa\cdot c\right] \tag{7}$$

and using the relation (6), (7), assuming $(1-p)^{5/3} \approx 1-(5/3)\cdot p$, we obtain the following law of compaction of a porous body:

$$\frac{dP}{dT} = \frac{1}{(1-5p/3)\cdot p}\cdot\frac{3(N_k+N_g)}{4\eta_0} \tag{8}$$

When constructing a model of the sintering process, we will accept the following assumptions from [9] about the structure of the material, the mechanisms of its compaction and grain growth:

1) the grains of the material are single-crystal and the structural parameter of the material is the average grain size (L);

2) compaction of the material at the sintering stage in a vacuum is carried out by the mechanism of diffusion-viscous flow, accompanied by thermally activated (with activation energy Q_b) slip along grain boundaries with a decrease in pore volume due to the absorption of voids by vacancy sinks, where the role of sinks is played by intergranular boundaries;

3) compaction (ρ) at the sintering stage under the pressure of uniform compression is carried out by the mechanism of viscous flow with volume (χ) and shear (η) viscosities;

4) the Mackenzie - Shuttleworth model is used for the rheological description of the porous structure (χ·η);

5) the growth of grains of the material occurs due to thermally activated (with activation

energy Q_s) diffusion coalescence of dispersed particles, consisting in the redistribution of the substance of small particles over the surface of large ones under the action of surface self-diffusion with the coefficient D_s;

6) the temperatures T_j at the j-stages (j=1,...n) of surface sintering (SSS) and the temperatures T_c of liquid-phase sintering (LPS) do not change over the volume of the furnace space;

7) the temperature of the material and the furnace space are equal.

Let us denote by time $\Delta_{1,j} = \left(t_{1,j=1}, t_{1,j}\right] - j - e(j = 1,...,n$ intervals of non-isothermal sintering with temperature $T_{1,j}$ maintained over an interval $\Delta_{1,j}$ of duration $\tau_{1,j}$;

$$t_{1,j} = \sum_{i=1}^{j} \tau_{1,l} (j = 1,...n); \;\; \tau_1 = t_{1,n}.$$

Through we denote $\Delta_2 = \left(\tau_1, \tau_1 + \tau_2\right]$ the time interval of isothermal sintering of duration τ_2, during which a constant temperature is maintained. We denote the porosity and grain sizes $\Delta_{1,j} (j = 1,...n)$ on through $p_{1,j}$ and L1,j, and Δ_2 on through p_2 and L_2.

The system of equations for the mathematical description of non-isothermal solid-phase sintering includes:

- the equation for the kinetics of compaction at the j-m temperature stage (see equation (2)):

$$\frac{dp_{1,j}}{dt} = -\frac{3(1-p_{1,j})^2}{1-p_0} \frac{1}{\left[\frac{p_0 - p_{1,j}}{3(1-p_{1,j})}\right]^{2,1}} \times \left\{ \frac{2\sigma\Omega D_{V,j}}{kT_{1,\upsilon} r_g^3} \frac{p_0 - p_{1,j}}{3(1-p_{1,j})} + \frac{\sigma\Omega b D_{b,j}}{2kT_{1,\upsilon} r_g^4} \right\}, \; t\epsilon\Delta_j, \; j=1 \quad (9,a)$$

for the first 3.5% of shrinkage, the equation of compaction kinetics is written as:

$$\frac{dp_{1,j}}{dt} = -\frac{3(1-p_{1,j})^2}{1-p_0} \frac{1}{\left[\frac{p_0 - p_{1,j}}{3(1-p_{1,j})}\right]^{2,06}} \times \left\{ \frac{2,63\sigma\Omega D_{V,j}}{kT_{1,\upsilon} r_g^3} \left[\frac{p_0 - p_{1,j}}{3(1-p_{1,j})}\right]^{1,03} + \frac{0,70\sigma\Omega b D_{b,j}}{kT_{1,\upsilon} r_g^4} \right\} \; t\epsilon\Delta j,=1,\ldots n \quad (9,b)$$

- the equation of the kinetics of grain growth of the material (from equation (10)):

$$\frac{dL}{dt} = \frac{B_1 \sigma D_s}{4L^3} \frac{\delta^4}{kT} \quad (10)$$

will

$$\frac{dL_{1,j}}{dt} = \frac{B_1 \sigma D_{s,j} \delta^4}{4L_{1,j}^3 \cdot kT1,j} \quad (11)$$

- temperature increase equation

$$\frac{dT_{1,j}}{dt} = w_{T,j}, w_{T,j} = (T_{1,j} - T_{1,j-1}) / \tau_{1,j}, \; t\epsilon\Delta j. \quad (12)$$

The coefficients of volume grain boundary and surface diffusion are calculated using form (13):

$$D = D_0 \cdot \exp\left(-\frac{Q}{kT}\right) \tag{13}$$

$$D_{V,j} = D_{V,0} \cdot \exp\left(-\frac{Q_V}{kT_{1,j}}\right), \qquad D_{b,j} = D_{b,0} \cdot \exp\left(-\frac{Q_b}{kT_{1,j}}\right), \quad D_{s,j} = D_{S,0} \cdot \exp\left(-\frac{Q_S}{kT_{1,j}}\right) \tag{14}$$

- initial conditions:

$$p_{1,1}\big|_{t=0} = p_0, \; L_{1,1}\big|_{t=0} = L_0, \; T_{1,1}\big|_{t=0} = T_0 \tag{15}$$

- conditions for conjugation of temperature stages in time:

$$p_{1,j}\big|_{t=t_{1-j}} = p_{1,j-1}, \; L_{1,j}\big|_{t=0} = L_0, \; T_{1,1}\big|_{t=0} = T_0 \tag{16}$$

The system of equations for the mathematical description of the stage of isothermal (liquid-phase) sintering under pressure includes:

- the equation of the kinetics of compaction of the material (see equation (8)):

$$\frac{dp_2}{dt} = \frac{1}{\left(1-\frac{5}{3}p_2\right)} \cdot \frac{3(N_k + N_g)}{4\eta_0}, \; t \epsilon \Delta 2, \tag{17}$$

- rheological model of a porous material (see equations (6) and (7)):

$$\chi = \frac{4\eta(1-p_2)}{3p_2}, \qquad \eta = \eta_0(1-p_2)^{5/3} \tag{18}$$

- equation of the kinetics of grain growth of the material:

$$\frac{dL_2}{dt} = \frac{B_1 \sigma D_{s,n} \delta^4}{4L_2^3 \cdot kTc}, \; t \epsilon \Delta 2 \tag{19}$$

where

$$D_{S,n} = D_{S,0} \cdot \exp\left(-\frac{Q_S}{kT_S}\right) \tag{20}$$

The capillary pressure Pk applied to the pore surface and the current pore radius Rp are determined by the equations [9]:

$$P_k = 2p_2 \cdot \sigma / R_p, \qquad R_p = R_p \left\{ \frac{\left[1 - p_{1,n}\right]}{p_{1,n}(1-p_2)} \right\}^{1/3} \tag{21}$$

- initial conditions:

$$p_2\big|_{t=t_n} = p_{1,n},\ L_2\big|_{t=t_n} = L_{1,n},\ T_c = T_{1,n} \tag{22}$$

The density of the material at the stages of the process is calculated depending on its current porosity using the equations

$$\rho_{1,j} = (1-p_{1,j})\rho_0, \quad j = 1,...,n, \qquad \rho_2 = (1-p_{1,n})\rho_0 \tag{23}$$

Volume shrinkage at stages is calculated using the equations:

$$\frac{\Delta V_{1,j}}{V_0} = \left(V_0 - \frac{m}{\rho_{1,j}}\right)/V_0,\ (j = 1,...,n); \frac{\Delta V_2}{V\Delta_0} = \left(V_0 - \frac{m}{\rho_2}\right)/V_0, \tag{24}$$

where $\Delta V_{1,j} = V_0 - V_{1,j},\ \Delta V_2 = V_0 - V_2$, m - sample mass.

The quality index (residual porosity pρ and average grain size Lρ and density ρρ) of a hard alloy is determined by the following relationships:

$$p_p = p_2\big|_{t=\tau_1+\tau_2}; \qquad L_p = L_2\big|_{t=\tau_1+\tau_2}; \qquad \rho_\rho = \rho_2\big|_{t=\tau_1+\tau_2} \tag{25}$$

Discussion. Since at low and medium annealing temperatures used in non-isothermal sintering, $D_b >> D_V$, namely $D_b / D_b = 10^3...10^5$, then in approximate calculations using formulas (9) we can assume $D_V \approx 10^{-4} D_b$. Instead of such an approximation, we can use the kinetic equation of linear shrinkage, applicable in the case of dominance of grain-boundary diffusion with respect to volume diffusion. This leads to the following equation of compaction kinetics in the j-m temperature regime instead of equation (9a):

$$\frac{dp_{1,j}}{dt} = \frac{-3(1-p_{1,j})^2}{1-p_0} \cdot 0{,}33 \left(\frac{2{,}14 \cdot \sigma\Omega b D_V}{kT_{1,j} r_g^4}\right)^{0{,}33} \cdot t^{-0{,}67}$$

which can be applied to all t∈Δj, starting from j=2 to j=n.

Let us denote by and the volume shrinkages at the end of the TFS stage (i=1) and at the end of the LFS stage (i=2) and the corresponding rates $W_i = \frac{d\gamma_i}{dt}\ (i = 1,2)$, where $\Delta V_{1,j} = V_0 - V_{1,n},\ \Delta V_2 = V_0 - V_2$.

The indicators Y_1 and Y_2 are calculated using the equation

$$Y_1 = \left(V_0 - \frac{m}{\rho_{1,n}}\right)/V_0, \quad Y_2 = \left(V_0 - \frac{m}{\rho_2}\right)/V_0$$

Product quality management consists of selecting such control actions $U_1 = \{T_{1,j}, \tau_{1,j},\ j = 1,...,n\},\ U_2 = \{T_C, \tau_2\}$ with elements that satisfy the constraints

$U_{k,i}^{\min} \le U_{k,i} \le U_{k,i}^{\max}$ that ensure the specified (desired) values of the quality indicators of the obtained material

$$p_\rho \le p_\rho^{\max}, \quad L_\rho \le L_\rho^{\max}, \quad p_\rho^{\min}$$

and characteristics of the synthesis process

$$Y_i^{\min} \le Y_i \le Y_i^{\max}, \quad W_i^{\min} \le W_i \le W_i^{\max} \quad (i = 1,2)$$

Where $p_\rho^{\max}$, $L_\rho^{\max}$, $p_\rho^{\min}$, – ultimate residual porosity, average grain size and alloy density, $Y_i^{\min}$, $Y_i^{\max}$ – hreshold values of volume shrinkage at the stage i, $W_i^{\min}$, $W_i^{\max}$ – threshold values of shrinkage rate of the synthesized material, $U_{k,i}^{\min}$, $U_{k,i}^{\max}$ – regulatory values of control actions from the set Ui, depending on the equipment and material.

III. Conclusions: 1

1. A mathematical model of the non-isothermal stage of sintering of a WC-Co hard alloy from a mixture of meso-sized powders has been obtained. It consists of the equation of compaction kinetics at the j-m temperature stage, at the first 3.5% shrinkage, and the equation of grain growth kinetics and temperature increase. The coefficients of volume grain boundary and surface diffusion have been calculated.

2. Equations for the mathematical description of isothermal sintering of a WC-Co hard alloy under pressure in a vacuum have been proposed. They consider the following equations: compaction kinetics equations, a rheological model, and a grain growth equation in the alloy structure. These equations allow one to control the product quality indicator and the characteristics of the WC-Co synthesis process with a meso structure.

IV. Conflict of interest

The authors declare that they have no conflict on interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper

Financing

This work was supported by the Azerbaijan Science Foundation - Grant № AEF-MCG-2023-1(43)-13(01)1-M-01

Data avoailability

The manuscript has no associated data

References

[1] J. Musil. Hard nanocomposite coatings: Thermal stability, oxidation resistance and toughness// Surface and Coatings Technology. Volume 207, 25 August 2012, Pages 50-65. https://doi.org/10.1016/j.surfcoat.2012.05.073.https://www.sciencedirect.com/science/article/abs/pii/S0257897212004793.

[2] Lan Sun, Cheng-Chang Jia, Min Xian. A research on the grain growth of WC–Co cemented carbide// International Journal of Refractory Metals and Hard Materials. Volume 25, Issue 2, March 2007,Pages121-124.https://doi.org/10.1016/j.ijrmhm.2006.03.002.
https://www.sciencedirect.com/science/article/abs/pii/S0263436806000102.

[3] I. Konyashin a b, S. Farag a, B. Ries a, B. Roebuck.WC-Co-Re cemented carbides: Structure, properties and potential applications// International Journal of Refractory Metals and Hard Materials.https://doi.org/10.1016/j.ijrmhm.2018.10.001.
https://www.sciencedirect.com/science/article/abs/pii/S0263436818304876.

[4] SS Gill, J Singh, H Singh, R Singh. Metallurgical and mechanical characteristics of cryogenically treated tungsten carbide (WC–Co)// The International Journal of Advanced …, 2012 – Springer.DOI 10.1007/s00170-011-3369-4.https://link.springer.com/article/10.1007/s00170-011-3369-4

[5] Hongbo Nie & Taiquan Zhang. Development of manufacturing technology on WC–Co hardmetals.// Tungsten 1, 198–212 (2019). https://doi.org/10.1007/s42864-019-00025-6. https://link.springer.com/article/10.1007/s42864-019-00025-6#citeas.

[6] Panov V.S., Chuvalin A.M. Technology and properties of sintered hard alloys and products made from them. Textbook for universities.-M: "MISIS", 2001, 428 p.

[7] Savitsky A.P. Liquid-phase sintering of systems with interacting components. - Novosibirsk: Science, Siberian Branch, 1991.-184p.

[8] Skorokhod V.V. Rheological foundations of the theory of sintering.-Kyiv: Naukova Dumka., 1976.-151p

[9] Kornienko I.G., Chistyakova T.B., Polosin A.I. and others. Sat.tr.scientific conference dedicated to the 185th anniversary of the formation of St. Petersburg State Technological Institute (TU), November 27, 2013. G.-SPb.: SPbGTI (TU), 2013.-P.396.-397.

INTELLIGENT ADAPTIVE SYSTEMS FOR PERSONALIZED EDUCATION: A NEURO-FUZZY APPROACH

Tokhirov Ezozbek[1], Subhan Namazov[2], Bakhtiyar Badalov[2]

•

[1]Tashkent State Transport University
[2]Azerbaijan Technical University, Baku, Azerbaijan Republic
etokhirov@yahoo.com, subhan.namazov@aztu.edu.az, bakhtiyar.badalov@aztu.edu.az

Abstract

Neuro-fuzzy models, integrating the adaptive learning capabilities of neural networks with the interpretability of fuzzy logic systems, have emerged as powerful tools in educational data mining. This research explores the application of neuro-fuzzy models in education, focusing on their role in predicting student performance, classifying academic outcomes, and enhancing personalized learning experiences. By analyzing various case studies and methodologies, this study highlights the effectiveness of neuro-fuzzy systems in handling the inherent uncertainties and complexities of educational data.

Keywords: Neuro-fuzzy models, adaptive neuro-fuzzy inference system (ANFIS), educational data mining, student performance prediction, fuzzy logic, neural networks, academic classification, personalized learning.

I. Introduction

The integration of artificial intelligence in education has led to the development of intelligent systems capable of analyzing and predicting student outcomes. Among these, neuro-fuzzy models stand out due to their hybrid nature, combining the learning ability of neural networks with the reasoning capability of fuzzy logic. These models are particularly adept at managing the imprecision and vagueness inherent in educational data, such as student behavior and performance metrics.

Recent studies have demonstrated the efficacy of neuro-fuzzy systems in various educational contexts. The university employed a neuro-fuzzy classifier to categorise students based on their academic performance, utilising inputs like exam results and socioeconomic factors. The model achieved high accuracy, outperforming traditional classification methods such as support vector machines and decision trees.

II. Literature review

This literature review presents general descriptions of educational processes in our country and abroad, analyses of conducted research and their effective methods, as well as the content and essence of scientific work and their new directions in eliminating shortcomings while supplementing educational conditions with new innovative equipment.

Mehdi & Nachouki (2023) developed a neuro-fuzzy model (ANFIS) to predict student graduation performance in IT programs. It achieved high accuracy (RMSE = 0.28) and outperformed traditional models. Key predictors included high school GPA and core courses like Data Structures and Software Engineering. Cortez & Silva (2008) applied decision trees, neural networks, and support vector machines to Portuguese student data. Random forest and neural networks showed the highest accuracy, identifying alcohol consumption and past grades as strong performance predictors. Macfadyen & Dawson (2010) demonstrated that LMS activity metrics (e.g., forum participation, assignment views) could reliably predict academic outcomes, paving the way for real-time learning analytics systems. Kotsiantis et al. (2004) found that parental education and previous grades were significant predictors of high school students' performance using Naive Bayes and decision trees. Nghe et al. (2007) compared various data mining models to predict student performance in a Vietnamese university. Decision trees offered the best trade-off between accuracy and interpretability. Thai-Nghe et al. (2011) employed support vector machines (SVMs) and matrix factorization on course data to predict whether students would pass or fail, achieving high classification accuracy. Romero et al. (2013) used association rule mining on Moodle data to find patterns of student behavior that correlated with academic success or failure. Baker & Yacef (2009) outlined differences between educational data mining (EDM) and learning analytics (LA), emphasizing how both fields contribute to student performance prediction through complementary lenses. Al-Barrak & Al-Razgan (2016) showed artificial neural networks outperform logistic regression in predicting student academic success in Saudi universities based on GPA and course performance. Zafra & Ventura (2009) utilized evolutionary algorithms combined with classification techniques to predict student dropouts in e-learning environments, demonstrating high performance and generalizability. Vapnik (1998) introduced the statistical theory behind SVMs, which have since been adapted effectively in educational settings for student performance classification tasks. Jayaprakash et al. (2014) developed an early warning system using LMS data and logistic regression, showing that early predictions of risk enabled successful student interventions. Gray et al. (2014) analyzed behavioral engagement patterns in online environments and linked them with GPA outcomes, reinforcing the importance of tracking online learning behaviors. Zhang & Rangwala (2018) used graph-based models to analyze course sequences and student trajectories, improving accuracy in long-term GPA predictions. Kabra & Bichkar (2011) applied decision tree classifiers on engineering student data in India, highlighting attendance, test scores, and subject difficulty as key predictors. Binns et al. (2018) explored fairness and accountability in predictive educational models, emphasizing that unchecked bias in input data can lead to discriminatory outcomes. Aher & Lobo (2013) used collaborative filtering and content-based filtering for student recommendation systems, indirectly aiding performance by aligning learning resources with student needs. Temraz (2020) evaluated ensemble learning methods like bagging and boosting for performance prediction, showing superior accuracy compared to standalone algorithms. Papamitsiou & Economides (2014) provided a systematic review of learning analytics tools and found that predictive models using multimodal data offered greater insights into student behavior and outcomes. You (2016) developed a model using temporal learning analytics to predict weekly performance in MOOCs, showing the effectiveness of time-series features in forecasting final grades.

III. Methods

ANFIS is a hybrid intelligent system that combines the benefits of neural networks and fuzzy logic. It uses a learning algorithm to tune the parameters of a Takagi-Sugeno fuzzy inference system. The structure of ANFIS consists of five layers:

1. **Layer 1 (Fuzzification):** Each node generates a membership grade for the input variables.

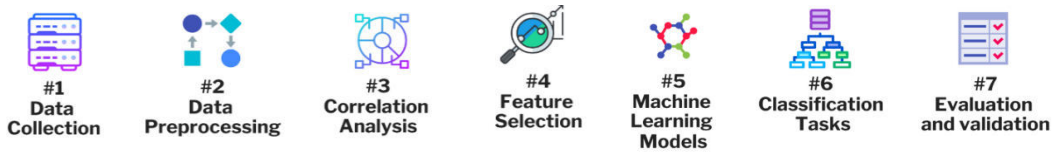
2. **Layer 2 (Rule Layer):** Nodes represent fuzzy rules, and each node's output is the product of the input membership grades.
3. **Layer 3 (Normalization Layer):** Nodes calculate the normalized firing strengths of the rules.
4. **Layer 4 (Defuzzification Layer):** Nodes compute the output of each rule.
5. **Layer 5 (Output Layer):** Nodes compute the overall output as the summation of all rule outputs.

The learning process involves adjusting the parameters of the membership functions and the consequent parameters of the fuzzy rules to minimize the error between the predicted and actual outputs.

B. Case Study: Student Performance Prediction

In a study by Mehdi and Nachouki (2023), ANFIS was utilized to predict the graduation grade point average (GPA) of students in an information technology program. Input variables included high school GPA and grades in core IT courses. The model demonstrated a high degree of accuracy, with 77% of predictions falling within one root mean square error of the actual GPA.

This image illustrates a 7-step process for Predictive Academic Performance Analysis Workflow, which is a structured methodology used to analyze and model student academic performance data.



Picture 1: *Steps of performance analysis in education*

Step 1: Data Collection
-Source: Academic records from a specific dataset (e.g., DPAES University, 2019–2021).
-What is collected: Grades, courses, and semesters of students.
Step 2: Data Preprocessing
-Step 1: Filter to include only first course attempts.
-Step 2: Convert results to binary or multi-class labels (e.g., Pass/Fail, Grade categories).
-Step 3: Organize data — for instance, by student or semester.
Step 3: Correlation Analysis
-Identify courses that correlate highly with each other.
-Use a threshold (e.g., correlation > 0.3) to decide which features (courses) are related.
-Goal: Understand which courses are strongly related and may impact performance.
Step 4: Feature Selection
-Choose only the most relevant, highly correlated features.
-Goal: Eliminate noisy or irrelevant data, which helps improve model performance.
Step 5: Machine Learning Models
-Algorithms used:
*k-Nearest Neighbors (k-NN)
*Random Forests
*Logistic Regression
*Neural Networks
-Purpose: Use these models to predict academic performance.
-Each model is fine-tuned for the best results.
Step 6: Classification Tasks
-Predict whether students will pass or fail, or what grade range they'll fall into.
-Types:

*Binary Classification: Pass vs. Fail

*Multi-Class Classification: Grades (e.g., Fail = 0–4.9, Average = 5–6.9, Excellent = 7–10)

Step 7: Evaluation and Validation

-Metrics used: Accuracy, F1 Score, Precision, Recall.

-Cross-validation is applied to make sure the results are generalizable to new data.

Table 1: *Sample: Student Performance Dataset*

Student ID	**Gender**	**Age**	Study **Hours**	Parental **Education**	Lunch **Type**	Test **Prep**	Math **Score**	Reading **Score**	Writing **Score**	Final **Grade**
1	Male	17	3.5	Bachelor's	Standard	Completed	78	72	74	Average
2	Female	16	5.0	Master's	Free	None	88	90	92	Excellent
3	Female	18	2.0	High School	Standard	Completed	64	68	70	Average
4	Male	17	1.0	Associate Degree	Free	None	45	40	42	Fail
5	Female	16	4.5	Master's	Standard	Completed	92	95	93	Excellent
6	Male	18	2.5	High School	Free	None	58	62	60	Average
7	Female	17	3.0	Bachelor's	Standard	Completed	75	80	78	Average
8	Male	17	0.5	Some College	Free	None	38	35	40	Fail
9	Female	16	6.0	Bachelor's	Standard	Completed	95	97	96	Excellent
10	Male	18	1.5	High School	Free	None	50	52		

IV. Results

The application of neuro-fuzzy models in educational settings has yielded promising results:

•Student Classification: Neuro-fuzzy classifiers have been effective in categorizing students into performance groups, aiding in targeted interventions.

•Performance Prediction: Models have accurately predicted student outcomes, facilitating early identification of at-risk students.

•Course Analysis: Sensitivity analysis within ANFIS models has highlighted key courses influencing student success, guiding curriculum improvements.

Table 2: *Student Performance Dataset*

Name		Grade	
First name	Last Name	Gender	Average Score
Jasur	Aliyeva	Male	74.67
Sevinch	Sobirova	Female	90.00
Muxsin	Olimov	Male	41.00
Axrora	Tolipova	Female	81.00

The application of neuro-fuzzy models in educational settings has yielded promising results:

- **Student Classification:** Neuro-fuzzy classifiers have been effective in categorizing students into performance groups, aiding in targeted interventions.
- **Performance Prediction:** Models have accurately predicted student outcomes, facilitating early identification of at-risk students.

Course Analysis: Sensitivity analysis within ANFIS models has highlighted key courses influencing student success, guiding curriculum improvements.

V. Discussion

I. Subsection One

Advantages of Neuro-Fuzzy Models

Neuro-fuzzy models offer several advantages in educational applications:

-Interpretability: The fuzzy rules provide a transparent understanding of the decision-making process.

-Adaptability: The neural network component allows the model to learn from data, adapting to new trends.

-Handling Uncertainty: Fuzzy logic effectively manages the ambiguity and vagueness in educational data.

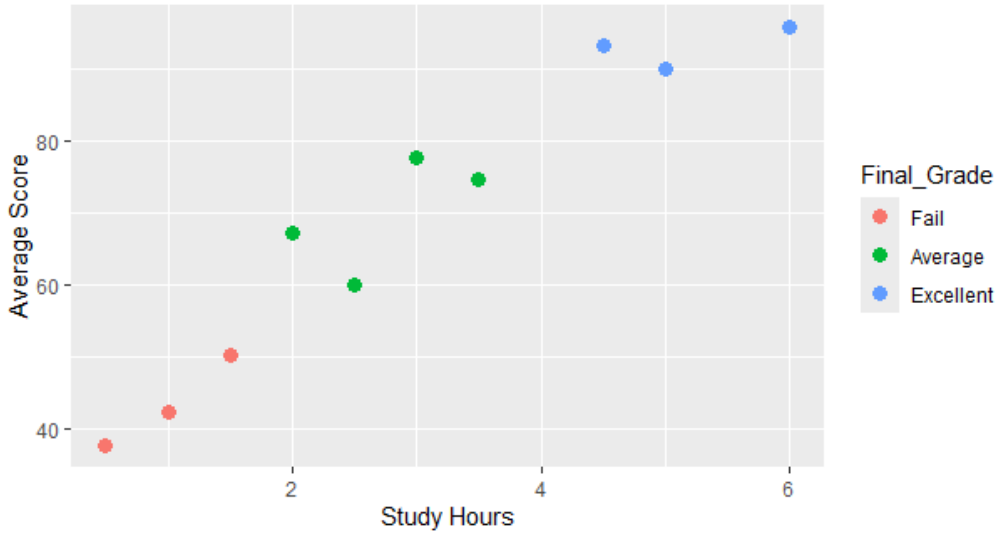


Figure 1: *Figure comparison of between caption study hours and average score*

II. Subsection Two

Challenges and Limitations

Despite their benefits, neuro-fuzzy models face certain challenges:

•Data Quality: The accuracy of predictions is highly dependent on the quality and completeness of the input data.

•Complexity: Designing and tuning neuro-fuzzy models can be complex and time-consuming.

•Scalability: Applying these models to large-scale educational systems may require significant computational resources.

VI. Conclusion

In this study, a binary logistic regression model was employed to examine the impact of study hours and test preparation on student academic performance, categorized as either pass or fail. The model included two independent variables: daily study hours and participation in a test preparation course.

Although the regression coefficients indicated that increased study time and completion of test preparation were associated with a higher likelihood of passing, these relationships were found to be statistically insignificant. Specifically, the p-values for both predictors were equal to 1, and the standard errors were extremely large, suggesting that the model estimates were unstable. This instability likely resulted from the small sample size (n = 10) and potential data issues such as perfect or quasi-perfect separation, where outcomes may be perfectly predicted by one or more variables.

Given these limitations, the model lacks sufficient statistical power to draw reliable conclusions about the predictors' influence on academic success. As such, while the direction of the coefficients aligns with established educational theories—that more study time and structured preparation enhance performance—the current analysis does not provide statistically robust evidence to support these claims.

VII. Recommendations for Future Research

To obtain more reliable and generalizable results, future research should be conducted using a substantially larger dataset. Moreover, employing additional variables such as socioeconomic background, school engagement, and classroom environment could improve model accuracy. Alternative machine learning models like decision trees or ensemble methods may also offer better performance with small or imbalanced datasets.

References

[1] Do, Q. H., & Chen, J. F. (2013). A neuro-fuzzy approach in the classification of students' academic performance. Computational Intelligence and Neuroscience, 2013, 179097. https://doi.org/10.1155/2013/179097

[2] Mehdi, R., & Nachouki, M. (2023). A neuro-fuzzy model for predicting and analyzing student graduation performance in computing programs. Education and Information Technologies, 28, 2455–2484. https://doi.org/10.1007/s10639-022-11205-2

[3] Stathacopoulou, R., Grigoriadou, M., Magoulas, G. D., & Mitropoulos, D. (2003). A neuro-fuzzy approach in student modeling. In User Modeling 2003 (pp. 337–341). Springer. https://doi.org/10.1007/3-540-44963-9_46

[4] Cortez, P., & Silva, A. M. G. (2008). Using data mining to predict secondary school student performance. In Proceedings of 5th FUture BUsiness TEChnology Conference (FUBUTEC), 5, 5–12.

[5] Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. Computers & Education, 54(2), 588–599.

[6] Kotsiantis, S. B., Pierrakeas, C., & Pintelas, P. (2004). Predicting students' performance in distance learning using machine learning techniques. Applied Artificial Intelligence, 18(5), 411–426.

[7] Nghe, N. T., Janecek, P., & Haddawy, P. (2007). A comparative analysis of techniques for predicting academic performance. In Proceedings of the 37th Annual Frontiers in Education Conference (pp. T2G-7–T2G-12).

[8] Thai-Nghe, N., Drumond, L., Horváth, T., Krohn-Grimberghe, A., & Schmidt-Thieme, L. (2011). Matrix factorization vs. classification models: A case study in predicting student performance. In Proceedings of the 12th International Conference on Educational Data Mining (EDM), 1–10.

[9] Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. S. J. D. (2013). Handbook of educational data mining. CRC Press.

[10] Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining, 1(1), 3–17.

[11] Al-Barrak, M. A., & Al-Razgan, M. S. (2016). Predicting students' performance through classification: A case study. Journal of Theoretical and Applied Information Technology, 88(1), 1–8.

[12] Zafra, A., & Ventura, S. (2009). Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data. Journal of Information Sciences, 180(6), 1106–1124.

[13] Vapnik, V. (1998). *Statistical learning theory*. Wiley.

[14] Jayaprakash, S. M., Moody, E. W., Lauría, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early

alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics, 1*(1), 6–47.

[15] Gray, G., McGuinness, C., & Owende, P. (2014). An application of classification models to predict learner progression in tertiary education. In *Proceedings of the 2014 International Conference on Educational Technologies* (pp. 172–177).

[16] Zhang, Y., & Rangwala, H. (2018). Grade prediction with temporal course-wise influence. In *Proceedings of the 9th International Conference on Educational Data Mining (EDM)* (pp. 39–48).

[17] Kabra, R., & Bichkar, R. (2011). Performance prediction of engineering students using decision trees. *International Journal of Computer Applications, 36*(11), 8–12.

[18] Binns, R., Veale, M., Van Kleek, M., & Shadbolt, N. (2018). 'It's reducing a human being to a percentage': Perceptions of justice in algorithmic decisions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–14).

[19] Aher, S. B., & Lobo, L. M. R. J. (2013). Combination of machine learning algorithms for recommendation of courses in E-learning system based on historical data. *International Journal of Computer Applications, 39*(1), 48–52.

[20] Temraz, M. (2020). Comparative analysis of ensemble learning techniques for student performance prediction. *International Journal of Advanced Computer Science and Applications, 11*(9), 167–174.

[21] Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society, 17*(4), 49–64.

[22] You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education, 29, 23–30.

THE ROLE OF ARTIFICIAL INTELLIGENCE IN ENGINEERING STUDIES: A COMPARATIVE PERSPECTIVE FROM GLOBAL UNIVERSITIES AND AZERBAIJAN TECHNICAL UNIVERSITY

Bakhtiyar Badalov[1], Parvana Movsumova[1]

•

[1]Azerbaijan Technical University, Baku, Azerbaijan Republic
bakhtiyar.badalov@aztu.edu.az, parvana.movsumova@aztu.edu.az

Abstract

The advance of AI creates a field for innovation, efficiency, and creativity in many fields, not passing beyond engineering education. This study examines the integration of AI technologies into engineering curricula, highlighting practices at globally leading universities, including Stanford University, the Technical University of Munich (TUM), and the National University of Singapore (NUS), and comparing them with initiatives undertaken at Azerbaijan Technical University (AzTU). Using the narrative review approach, the article analyzes the role of AI in modeling, virtual laboratories, generative design and personalized learning. In addition, it discusses the critically important skills that modern engineers must acquire, and considers the ethical issues associated with the implementation of AI. The results show that although AI significantly enhances the educational experience and expands students' technical skills, strategic planning and oversight are critical to its evaluation. The article concludes with recommendations for future AI advocates in engineering education, particularly in new academic contexts.

Keywords: Artificial intelligence, engineering education, higher education, simulation, personalized learning, comparative study, digital technology, ethical AI

I. Introduction

In a rapidly evolving technological landscape, artificial intelligence (AI) is emerging as a transformative force, reshaping educational paradigms across sectors, with engineering being a particular focus. As a discipline focused on innovation, precision, and problem solving, engineering education can benefit greatly from the integration of AI. The introduction of intelligent technologies promises to facilitate more effective teaching methods, personalized educational trajectories and increased student engagement. Leading global institutions, from MIT to Stanford University, have quickly adopted AI-powered strategies, giving students access to cutting-edge tools and techniques that will make them resilient to the digital future.

The AI's role in education is versatile, and recent research increasingly highlights its reframing on engineering curricula. For example, Alimisis highlights how AI-based robotics education enhances critical thinking, problem solving, and creativity among engineering students, particularly in disciplines such as robotics and mechatronics [1].

Similarly, Khan et al. note that integrating AI into civil engineering technology programs not only enhances students' understanding of predictive maintenance concepts but also greatly enriches their training in smart city design by providing tools to address such urban challenges [14]. In turn, the study by Saldivar et al. [19]. demonstrates how the application of AI in electrical engineering – in particular the use of intelligent tutoring systems and adaptive educational platforms – contributes to improving learning outcomes by providing personalized support. Real-time feedback creates a more responsive and engaging learning environment with the help of this systematic approach.

According to a systematic review by Zawacki-Richter et al. there is compelling evidence that artificial intelligence technologies, including machine learning, natural language processing, and data analytics, are rapidly picking up steam in personalized learning environments, particularly in STEM disciplines [20].

This shift to AI-powered education not only improves the efficiency and effectiveness of teaching, but also allows for a more personalized approach to learning that takes into account students' different backgrounds and learning styles. Technology-enhanced learning and quantitative prediction can assist to explore gaps in students' knowledge, allowing teachers to actively intervene and adapt educational content accordingly. These AI tools are really important in engineering studies which mostly focuses on multifaced diagnostic and continuous innovation.

Furthermore, the future possibilities of AI in science and technology education exceed educational institutions. As industry demands continue to change, university curricula are increasingly integrating with AI technologies to ensure they align with current technological advances and societal needs, thereby equipping graduates with the competencies needed to succeed in an increasingly digital and automated world. The continual integration of AI tools into education also provides a floor for all students without any diversity. With AI tools they gain equity regardless of geographical location or socio-economic background, with high-quality learning experiences.

Succinctly, AI is balanced to reform engineering education by the quality and equity enhancement, encouraging creativity, and equipping students with the skills needed to manage the complexities of today's innovative environment. As educational institutions continue to leverage these advances, the opportunities for using AI at the forefront of engineering education become immense and inspiring.

II. Methodology

This study adopts a non-systematic comparative review approach to examine the integration of AI in engineering education across four universities: Stanford University, Technical University of Munich (TUM), National University of Singapore (NUS), and Azerbaijan Technical University (AzTU). Data were collected from academic literature, institutional reports, and publicly available documents, focusing on AI applications in engineering curricula, including adaptive learning systems, virtual labs, and AI-driven simulations.

A qualitative thematic analysis was employed to identify key trends, challenges, and opportunities in AI integration across these institutions. The analysis also addressed ethical considerations, such as algorithmic bias and data privacy, and explored the role of AI in preparing students for the digital future of engineering. The study also considered the varying institutional capacities and approaches, particularly in emerging economies.

III. Comparative Analysis

The fusion of neural network to education system created a paramount movement in last decades. With the advance of AI into the engineering disciplines which is always full of innovation and precision, a great deal of transformative pedagogical approaches was made. Aligning with industry many leading universities are actively embedding AI-driven methods into their curricula

in order to prepare graduates for increasingly digital and automated industries. However, these experiences are not the same for every region, they can vary across educational systems. This paper aims to explore the role of AI in engineering studies by examining experiences at Stanford University, TUM, NUS, and Azerbaijan Technical University (AzTU), offering insights and recommendations for future development. This study uses a non-systematic review integrating insights from academic literature, institutional documents and personal experiences of different universities. AI applications in engineering programs, including modeling tools, virtual labs, generative design, and personalized learning environments are in the spotlight of data collection.

Stanford University has been a pioneer in the implementation of AI in engineering education. Through programs such as the "Design Thinking AI Labs" and "AI for Mechanical Innovation", the university is harnessing the power of AI to facilitate advanced computer simulation, predictive modeling, and advanced industrial techniques. The goal of these initiatives is to provide students with the tools and frameworks needed to solve real-world engineering problems in a highly automated world. [13].

At TUM, the integration of AI into engineering is deeply interdisciplinary, fostering innovation through its Center for Digital Technology and Management (CDTM), which blends AI with entrepreneurial skills, particularly in the domains of aerospace and robotics. This emphasis on combining technical knowledge with business acumen prepares students for success in high-tech industries [4]. The university has introduced AI-powered simulation labs, especially for smart energy systems and sustainable design projects. Additionally, students participate in hands-on workshops and research projects focused on robotics and machine learning, equipping them with practical skills essential for the modern engineering landscape.

AI in engineering, particularly in urban systems and civil engineering is a paramount movement in the National University of Singapore (NUS). The use of AI-driven simulations to predict and model urban growth and environmental impacts, supporting the broader goal of creating smarter, more sustainable cities is called the "Smart Nation Initiative" [7].

Artificial intelligence is being increasingly integrated into engineering education and research to foster innovation and improve learning outcomes at Azerbaijan Technical University (AzTU). Departments such as Computer Engineering and Automation and Control Systems offer AI-focused courses, while specialized laboratories support projects in machine learning, robotics, and smart automation [3]. AzTU collaborates with industry partners to incorporate real-world AI applications into student projects, enhancing practical skills [2]. Research initiatives, supported by national and international grants, focus on smart manufacturing, predictive maintenance, and intelligent infrastructure development (AzTU, 2024). Additionally, Smart Campus projects apply AI in areas such as energy management, campus security, and digital services, contributing to a more sustainable and efficient learning environment [17]. Recognizing the ethical implications of AI, AzTU is embedding AI ethics, fairness, and transparency into its engineering curricula, aiming to develop graduates who are both technically skilled and socially responsible [2].

Table: *Here is a comparative overview of AI integration in engineering education at different universities*

University	AI Applications in Engineering Education	Notable Initiatives
Stanford University	AI in mechanical design, intelligent systems	Design Thinking AI Labs, AI for Mechanical Innovation program
Technical University of Munich (TUM)	AI-enhanced aerospace simulations, robotics	TUM Center for Digital Technology and Management (CDTM)
National University of Singapore (NUS)	Smart urban systems, machine learning in civil and environmental engineering	Smart Nation Initiative, AI for Smart Cities program

Azerbaijan Technical University (AzTU)	Simulation software, robotics, sustainable design using AI	Establishment of AI-based research projects in smart energy and sustainability

AI offers transformative benefits for engineering education, including enhanced simulation capabilities, access to virtual laboratories, and opportunities for personalized learning experiences. Despite the growing integration of AI, emerging economies continue to face challenges such as inadequate infrastructure, imperfect teacher training, ethical issues, and difficulties in adapting curricula. All of these require targeted solutions. AzTU's initiatives demonstrate promising progress but highlight the need for sustained investment, international collaboration, and strategic capacity building. To ensure responsible AI use ethical considerations such as algorithmic transparency, data privacy, and bias, must be incorporated into engineering curricula [9] [12]. In order to prepare engineers for future challenges the development of critical thinking, creativity, and interdisciplinary knowledge have to be considered carefully before the implementation [18] [21].

IV. Challenges and Ethical Concerns

On the one hand AI brings its benefits for change, on the other hand the integration of AI into engineering education presents several complex challenges. One significant issue is algorithmic bias, where AI systems may reinforce existing inequalities in learning opportunities or outcomes if their training data is unbalanced or non-representative [12]. Data privacy is another critical concern, as AI often requires access to extensive datasets containing personal, academic, and behavioral information, raising ethical questions about consent, security, and responsible data stewardship [9]. Moreover, many AI models function as "black boxes," with decision-making processes that are difficult for educators and students to interpret or contest, undermining transparency and trust [8]. Accountability also becomes blurred when AI-driven systems influence critical educational outcomes, necessitating clear frameworks to assign responsibility among software developers, faculty members, and institutional policymakers [20]. Additionally, there is a growing risk of over-reliance on AI tools, which may discourage students from developing essential cognitive skills such as critical thinking, problem-solving, and independent research, if these tools are used uncritically or excessively.

Accessibility remains a persistent issue; resource-constrained institutions, particularly in developing regions, may face difficulties in acquiring and maintaining advanced AI infrastructure, thereby exacerbating global educational inequalities [11]. Beyond infrastructural challenges, cultural and linguistic biases embedded in AI systems can disadvantage non-dominant language speakers and diverse learner populations. Furthermore, embedding AI ethics education into engineering curricula is increasingly recognized as essential to prepare future engineers to anticipate, mitigate, and respond to ethical dilemmas posed by intelligent systems [12]. Lastly, ongoing stakeholder engagement — involving students, educators, technologists, and ethicists — is crucial to ensuring that AI deployment in education remains aligned with human-centered values, promoting inclusivity, fairness, and social responsibility.

V. Future Trends in AI and Engineering Education

Looking ahead, several key trends are likely to define the next decade of AI-driven engineering education. First, adaptive learning platforms powered by AI will become more sophisticated, offering customized learning pathways based on individual performance, learning styles, and personal interests [15]. These systems will dynamically adjust content and pacing, optimizing educational experiences to meet the diverse needs of students. Second, AI-enhanced virtual labs and simulations will provide students with more opportunities to experiment in risk-free, immersive

environments, fostering creativity and problem-solving skills without the constraints of physical resources or safety concerns [10]. These tools will be particularly beneficial in engineering disciplines like robotics, automation, and aerospace, where physical testing can be costly and dangerous.

Furthermore, AI will likely play a crucial role in continuous curriculum updates, ensuring that engineering programs stay aligned with industry demands in real time [5]. By leveraging data from various sources — including industry trends, academic performance, and evolving technologies — AI systems will help educational institutions rapidly adapt their curricula, keeping pace with the ever-changing needs of the job market. This real-time responsiveness will enhance the relevance of engineering education and improve the employability of graduates.

Ethical AI education will also become a core component of engineering programs, ensuring that future engineers are not only technologically proficient but also socially responsible [12]. As AI technologies evolve, their applications raise important societal concerns, such as fairness, transparency, and accountability. As a result, curricula will increasingly focus on teaching ethical frameworks and decision-making processes for AI deployment, enabling students to navigate complex moral issues in their future careers.

AI is poised to redefine engineering education, fostering greater creativity, precision, and interdisciplinary collaboration. While leading universities such as MIT and Stanford demonstrate the vast possibilities AI offers, institutions like AzTU are making impressive advancements despite resource constraints. Future endeavors should focus on:

- Expanding AI-focused curricula across engineering disciplines to ensure a broader understanding of AI's potential applications in fields such as civil, mechanical, electrical, and environmental engineering.
- Strengthening capacity building projects that equip educators with the skills needed to integrate AI into academic and empirical approaches.
- Encouraging interdisciplinary research initiatives that combine AI with other advanced technologies such as biotechnology, nanotechnology and environmental sustainability to address complex problems of global concern.
- Create a comprehensive ethical framework for AI in education programs to ensure the sustainable design, implementation, and governance of AI technologies in real-world engineering applications.

Moreover, AI will enable universities to develop more effective and personalized learning assessment systems, providing students with real-time feedback and recommendations to enhance their academic progress [6]. These AI-driven assessments could help identify learning gaps early, offering timely interventions to improve student outcomes. However, challenges related to data privacy, security, and algorithmic bias will need to be carefully managed to avoid unintended consequences.

Further analysis should examine the lasting impact of AI integration on learning outcomes, focusing on how AI tools impact not only technical knowledge but also the development of critical reflection, ingenuity, and interpersonal skills. Empirical research using grounded theory approach will be crucial in evaluating the effectiveness of AI-powered learning environments and curriculum innovations in diverse academic contexts.

VI. Conclusion

The integration of artificial intelligence (AI) into engineering education has brought about transformative changes across various institutions, offering new opportunities for enhancing teaching, learning, and real-world applications. Experiences of the institutions like Stanford University, Technical University of Munich (TUM), National University of Singapore (NUS), and Azerbaijan Technical University (AzTU) are at the forefront of utilizing AI to prepare students for

the rapidly evolving engineering landscape. By incorporating AI-powered tools, simulations, and interdisciplinary projects, these universities are equipping students with the skills needed to address the complex challenges of the modern world.

However, while the potential of AI in education is vast, several challenges must be addressed to fully realize its benefits. Issues such as algorithmic bias, data privacy concerns, and the need for better infrastructure and teacher training remain critical obstacles, particularly for universities in developing regions. Institutions like Stanford and TUM, with their established resources and infrastructure, are better positioned to navigate these challenges, while universities like AzTU are making progress but require further investment and capacity-building efforts.

Looking ahead, the future of AI in engineering education holds great promise, particularly in areas like personalized learning, interdisciplinary collaboration, and real-world applications. As AI continues to evolve, educational institutions will need to adapt their curricula, invest in state-of-the-art technologies, and focus on ethical considerations to ensure that AI serves as a force for good in preparing the next generation of engineers. Collaboration between institutions, industries, and governments will be key to overcoming existing challenges and ensuring that AI contributes to the development of a sustainable, innovative, and inclusive global engineering community.

References

[1] Alimisis, D. (2021). Robotics and Artificial Intelligence: Enhancing Engineering Education. Journal of Engineering Education Transformations, 34(2), 97–105.

[2] Aliyev, T., & Huseynov, S. (2023). "Artificial Intelligence Applications in Engineering Education: The Case of Azerbaijan Technical University," Baku Journal of Engineering Education, 5(2), 45–58.

[3] Azerbaijan Technical University (AzTU). (2024). Innovation and Research Annual Report.

[4] Bauer, J., et al. (2018). Innovation in Engineering Education through Interdisciplinary Learning. TUM Innovation Report.

[5] Brynjolfsson, E., & McAfee, A. (2017). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W.W. Norton & Company.

[6] Chen, X., Zhang, L., & Li, H. (2020). "Intelligent Educational Systems: A Comprehensive Review and Future Prospects." Journal of Educational Technology & Society, 23(4), 1-12.

[7] Chowdhury, S., & Chan, P. (2020). Urban Analytics for Smart Cities. National University of Singapore Press.

[8] Doshi-Velez, F., & Kim, B. (2017). Towards A Rigorous Science of Interpretable Machine Learning. arXiv preprint arXiv:1702.08608.

[9] Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. Harvard Data Science Review.

[10] Guerra, P., & Camacho, D. (2021). "Virtual Labs for Engineering Education: Opportunities and Challenges." International Journal of Engineering Education, 37(3), 768-777.

[11] Holmes, W., Bialik, M., & Fadel, C. (2021). Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign.

[12] Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399.

[13] Jordan, M. I., & Mitchell, T. M. (2015). Machine Learning: Trends, Perspectives, and Prospects. Science, 349(6245), 255–260.

[14] Khan, R., Siddiqui, F., & Javed, Y. (2022). AI and Civil Engineering Education: A New Paradigm. International Journal of Civil Engineering Education, 48(1), 15–34.

[15] Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence Unleashed: An Argument for AI in Education. Pearson Education.

[16] Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A Survey on Bias and Fairness in Machine Learning. ACM Computing Surveys (CSUR), 54(6), 1–35.

[17] Ministry of Education, Azerbaijan. (2023). Higher Education Digitalization Strategy 2022–2030.

[18] National Academy of Engineering. (2017). *The engineer of 2020: Visions of engineering in the new century*. National Academies Press.

[19] Saldivar, J., et al. (2019). Intelligent Tutoring Systems in Electrical Engineering Education. Journal of Intelligent Learning Systems and Applications, 11(3), 43–54.

[20] Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. Learning, Media and Technology, 45(3), 223–235.

[21] World Economic Forum. (2020). *The future of jobs report 2020*. Retrieved from weforum.org

[22] Zawacki-Richter, O., Marín, V. I., & Bond, M. (2019). Systematic Review of the Impact of AI in STEM Education: A Focus on Personalized Learning. Computers & Education, 141, 103604.