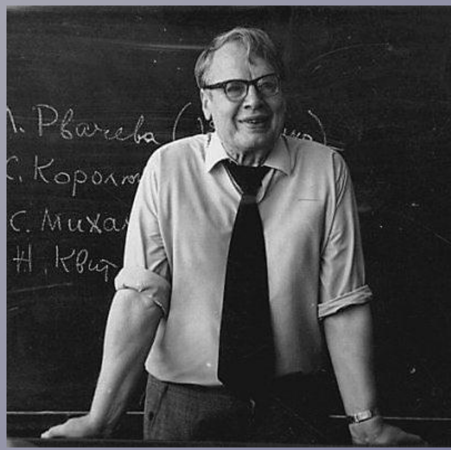
ELECTRONIC JOURNAL
OF INTERNATIONAL
GROUP ON RELIABILITY

Gnedenko Forum Publications





JOURNAL IS REGISTERED
IN THE LIBRARY
OF THE U.S. CONGRESS

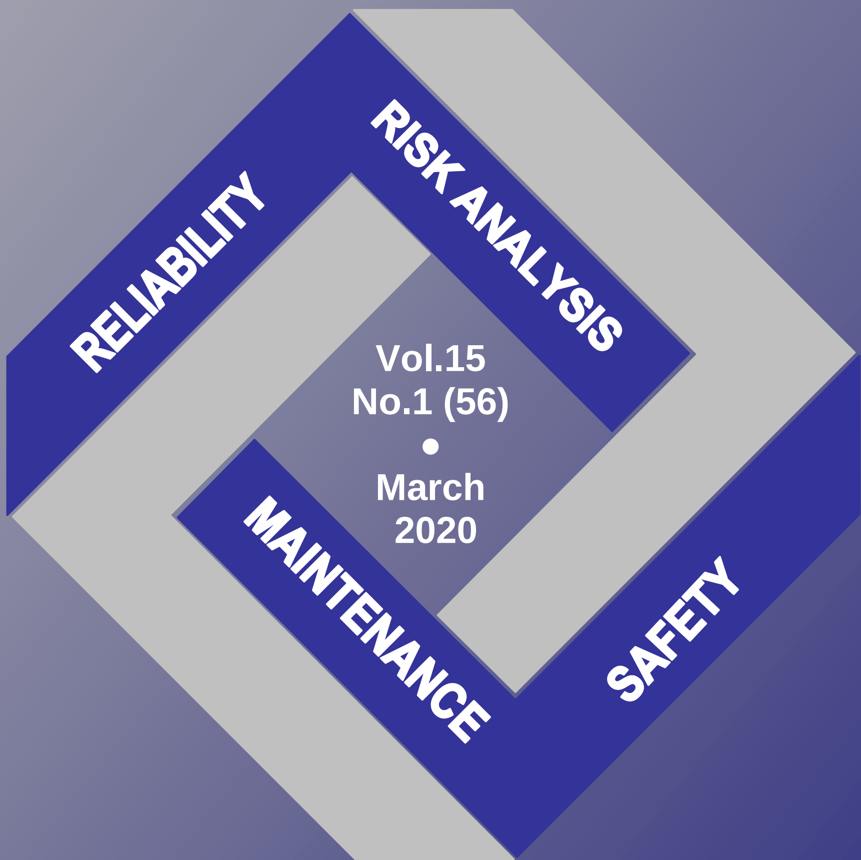
•

ISSN 1932-2321

•

VOL.15 NO.1 (56)
MARCH, 2020

RELIABILITY:
THEORY&APPLICATIONS



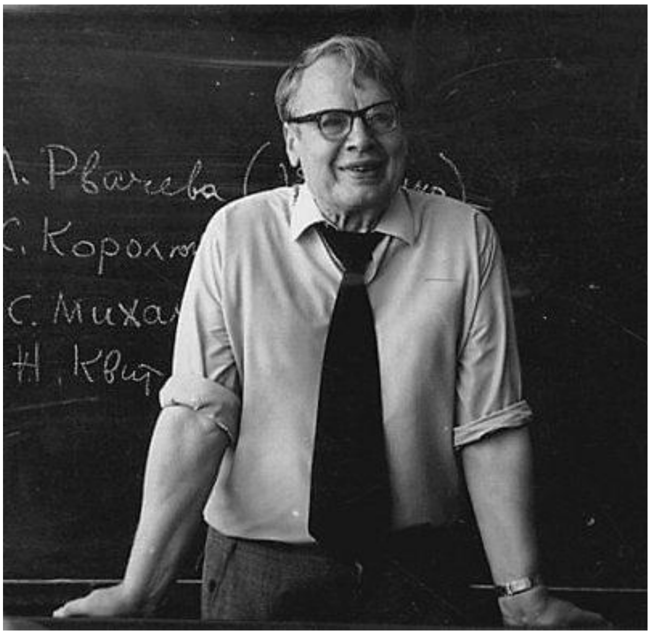
San Diego

ISSN 1932-2321

http://www.gnedenko.net/Journal/index.htm

Journal of International Group on Reliability





RELIABILITY:

THEORY & APPLICATIONS

Vol.15 No.1 (56),
March 2020

San Diego
2020

Editorial Board

e-Journal ***Reliability: Theory & Applications*** publishes papers, reviews, memoirs, and bibliographical materials on Reliability, Quality Control, Safety, Survivability and Maintenance.

Theoretical papers have to contain new problems, finger practical applications and should not be overloaded with clumsy formal solutions.

Priority is given to descriptions of case studies.
General requirements for presented papers

1. Papers have to be presented in English in MS Word or LaTeX format.
2. The total volume of the paper (with illustrations) can be up to 15 pages.
3. A presented paper has to be spell-checked.
4. For those whose language is not English, we kindly recommend using professional linguistic proofs before sending a paper to the journal.

The manuscripts complying with the scope of journal and accepted by the Editor are registered and sent for external review. The reviewed articles are emailed back to the authors for revision and improvement.

The decision to accept or reject a manuscript is made by the Editor considering the referees' opinion and taking into account scientific importance and novelty of the presented materials. Manuscripts are published in the author's edition. The Editorial Board are not responsible for possible typos in the original text. The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Publication in this e-Journal is equal to publication in other International scientific journals.

Papers directed by Members of the Editorial Boards are accepted without referring.
The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Send your papers to Alexander Bochkov, e-mail: a.bochkov@gmail.com

Table of Contents

Customer Orientated Indices and Reliability Evaluation of Meshed Power Distribution System 10

Aditya Tiwary

Reliability evaluation of a system or component or element is very important in order to predict its availability and other relevant indices. Reliability is the parameter, which tells about the availability of the system under proper working conditions for a given period. The study of different reliability indices are very important considering the complex and uncertain nature of the power system. In this paper reliability, evaluation of the meshed distribution system is presented. This paper also evaluates basic indices such as average failure rate, average outage time and average annual outage time. Along with basic indices, customer orientated indices such as system average interruption frequency index, system average interruption duration index and customer average interruption duration index of an electrical power distribution system is also evaluated. The electrical power distribution system taken for study is meshed distribution system in nature.

On Exponential-Weibull Distribution Useful in Reliability and Survival Analysis 20

Anakha K. K., Chacko V. M

In this paper, mixture of Exponential and Weibull distributions is considered for modelling real lifetime data. The basic mathematical properties including moments, generating functions, order statistics etc are derived. We obtain the reliability of stress-strength model. The maximum likelihood method is performed to estimate the parameters and a simulation study is conducted to validate the maximum likelihood estimators. The model is fitted to a real data set.

Building the Continuous Random Process Out of the Specified Sequence of Turning Points for Fatigue Testing 33

Irina Gadolina, Rashid Zaynetdinov

The method of regeneration of continuous process is intended for the first place for the calculation of spectral density of regenerated process. The main feature of this method is preserving the values and sequence of turning points (extremes) given in "saw-kind" realization. While doing so, the methods which is based on cycle-counting methods will give the exactly the same fatigue durability estimation, because the initial condition MAX-MIN-MAX … is guaranteed. To investigate the random process standard deviation (RMS) by the spectral density, the extrapolation of the original sequence is provided by continuous cosine functions. Continuality of the process and its first derivative is ensured by the condition of compatibility at the turning points. To decide on frequencies, the information from some sample realizations obtained in exploitation was employed. As one of te applications, the method is intended to be used in the analysis of comparability of two competing approaches in the evaluation of loading in the tasks of evaluation of durability, namely, those that apply cycle-counting methods and which are based on the spectral density of processes methods. Some other speculations made on the modelled process are performed.

Transient Solution of a Heterogeneous Queuing System with Balking and Catastrophes 42

Rakesh Kumar, Sapana Sharma, Bhavneet Singh Soodan, P. Vijaya Laxmi, Bhupender Kumar Som

In this paper we consider a Markovian queuing system with heterogeneous servers, balking and catastrophes. The time-dependent behavior of the system is analyzed by using generating function technique. The expressions for mean and variance of the system are obtained in transient state. At last, some special cases of the model are derived and discussed.

A Hybrid Technique of Combining AES Algorithm with Block Permutation for Image Encryption 51

Ritu Shaktawat, Rajdeep Singh Shaktawat, N. Lakshmi, Avinash Panwar, Arun Vaishnav

*In this paper, a hybrid approach for image encryption is proposed by combining AES, a standard cryptography algorithm, along with splitting and block permutation. A standard image (Lena) is taken as an input and to enhance the security, the image is divided into 4*4 matrix followed by block permutation before encryption of the image with AES. Comparison of various parameters with and without block permutation confirms the superiority of the proposed method in terms of better results after using splitting and permutation functions.*

Variances of Linear Regression Coefficients for Safety Margin of Technical System 57

Gurami Tsitsiashvili, Marina Osipova

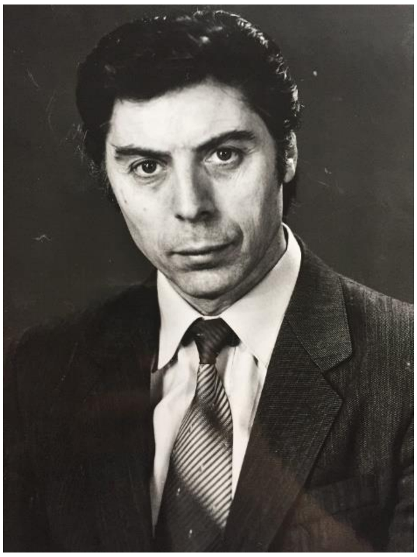
In this paper, we choose such a particular formulation of the problem of calculating linear regression coefficient, when the moments of observation form an arithmetic progression or depend in proportion to some degree of the observation number and variances of the estimation of the safety margin of the technical system decrease over time. It is proved that the variance of the trend estimation in this case decreases as a certain degree of the length of the series of observations. This makes it possible to evaluate the effectiveness of non-destructive testing for the safety margin of the technical system.

Probabilistic Comparisons of Systems Operation Quality for Uncertainty Conditions 63

Andrey Kostogryzov, Petr Kanygin, Andrey Nistratov

For systems, including systems of intelligent manufacturing, information and intellectual systems, an achievement of operation quality at admissible expenses is very important and needs prognostic comparisons. Here the probabilistic approach to compare an operation quality of functionally similar systems for uncertainty conditions is proposed. To be compared there may be: different systems for an one operation time period or for different time periods with identical duration; or the same system for different time periods on time line. The system operation outputs are considered in the forms of material products, information products and products, combined from material and information products. For the given time of prediction the main results of the approach application are: a relative part of functions executed with admissible quality, estimations of expenses considering inadmissible system operation quality, a relative part of system operation satisfaction connected with quality and cost. The approach is demonstrated by examples.

Friend's memory



On January 27 of this year, after a long illness, a well-known specialist in the field of reliability, Ph.D. Felix Izrailevich Fishbein. The total experience of his work in the field of reliability exceeds 50 years.

Felix Izrailevich was born on September 1, 1933. He began his career as a “reliable” at NII-101 (currently the Scientific Research Institute of Automatic Equipment named after Academician V.S. Semenikhin) in the unit, which was led by an outstanding scientist, doctor of technical sciences, Professor I.A. Ushakov. Under the leadership of the latter F.I. performed several studies in the field of planning control tests of military-technical systems and their components.

The results of these studies formed the basis of the dissertation and were published in a few collections of articles and books. In fact, in the 60s - 70s of the last century, F.I. He was part of a very small group of scientists who created and creatively developed the theory of reliability, which was still very young then, and it was F.I. was one of those who worked all his life at the intersection of theory and practice. His name is among the developers of several industry and national standards, including the fundamental standard for terms in the field of reliability. In the fundamental reference book “Reliability of technical systems”, published in 1985 under the editorship of I.A. Ushakova, who today remains one of the main Russian-language sources on the theory and practice of reliability, Felix Fishbein was the author of two chapters on the assessment of reliability indicators based on the results of operation and testing. Felix was an example of a specialist who combined in-depth knowledge of theory and an excellent ability to put it into practice. At the same time, he was always distinguished by the desire to often bring very complex scientific methods to a level understandable to a competent engineer, a non-specialist in the field of reliability. In this area, he made an outstanding contribution to the promotion and development of test planning methods based on solutions of the Clopper-Pearson equations using the nomogram of the cumulative binomial distribution (H. Larson. A Nomograph of the Cumulative Binomial Distribution. - Industrial Quality Control, 1966, Dec., 270-278). He wrote two brochures published by the Knowledge publishing house, with a detailed description of how to plan definitive and control reliability tests using this nomogram. F.I. He was a lecturer and consultant to the Cabinet of Quality and Reliability at the Polytechnic Museum. Much time F.I. devoted to the journal “Reliability and Quality Control”, from which the journal “Methods of Quality Management”, published today, “grew”. He was a member of the editorial board of the Reliability series from 1991 to 1995. (In those years, the journal “Reliability and Quality Control” was published in two series: “Reliability” and “Statistical Methods”).

He was a wonderful person, kind, sympathetic, cheerful, devoted friend, and fantastic father and grandfather. The world became poorer with the departure of Felix Izrailovich. His memory will forever remain in our hearts.

Group of mates

Customer Orientated Indices and Reliability Evaluation of Meshed Power Distribution System

Aditya Tiwary

•

Dept. of Fire Technology & Safety Engineering, IPS Academy, Institute of Engineering and science, Rajendra Nagar, Indore (M.P), India
raditya2002@gmail.com

Abstract

Reliability evaluation of a system or component or element is very important in order to predict its availability and other relevant indices. Reliability is the parameter, which tells about the availability of the system under proper working conditions for a given period of time. The study of different reliability indices are very important considering the complex and uncertain nature of the power system. In this paper reliability, evaluation of the meshed distribution system is presented. This paper also evaluates basic indices such as average failure rate, average outage time and average annual outage time. Along with basic indices, customer orientated indices such as system average interruption frequency index, system average interruption duration index and customer average interruption duration index of an electrical power distribution system is also evaluated. The electrical power distribution system taken for study is meshed distribution system in nature.

Keywords: Reliability; Availability; Meshed distribution system; average interruption frequency index; system average interruption duration index.

I. Introduction

Reliability evaluation of a system or component or element is very important in order to predict its availability and other relevant indices. Reliability is the parameter which tells about the availability of the system under proper working conditions for a given period of time. A Markov cut-set composite approach to the reliability evaluation of transmission and distribution systems involving dependent failures was proposed by Singh et al. [1]. The reliability indices have been determined at any point of composite system by conditional probability approach by Billinton et al. [2]. Wojczynski et al. [3] discussed distribution system simulation studies which investigate the effect of interruption duration distributions and cost curve shapes on interruption cost estimates. New indices to reflect the integration of probabilistic models and fuzzy concepts was proposed by Verma et al. [4].

Zheng et al. [5] developed a model for a single unit and derived expression for availability of a component accounting tolerable repair time. Distributions of reliability indices resulting from two sampling techniques are presented and analyzed along with those from MCS by Jirutitijaroen and Singh [6]. Dzobe et al. [7] investigated the use of probability distribution function in reliability worth analysis of electric power system. Bae and Kim [8] presented an analytical technique to evaluate the reliability of customers in a microgrid including distribution generations. Reliability network equivalent approach to distribution system reliability assessment is proposed by Billinton and Wang [9].

Evaluation of Reliability indices accounting omission of random repair time for distribution systems using Monte Carlo simulation [10]. Determination of Optimum period between Inspections for Distribution system based on Availability Accounting Uncertainties in Inspection Time and Repair Time, Tiwary et al. [11]. Jirutitijaroen et al. [12] developed a comparison of simulation methods for power system reliability indexes and their distribution. Determination of reliability indices for distribution system using a state transition sampling technique accounting random down time omission Tiwary et al. [13]. Tiwary et al. [14] proposed a methodology based on inspection repair based availability optimization of distribution systems using Teaching Learning based Optimization. Bootstrapping based technique for evaluating reliability indices of RBTS distribution system neglecting random down time was evaluated [15].

Volkanavski et al. [16] proposed application of fault tree analysis for assessment of the power system reliability. Li et al. [17] studies the impact of covered overhead conductors on distribution reliability and safety. Reliability enhancement of distribution system using Teaching Learning based optimization considering customer and energy based indices was obtained in Tiwary et al. [18]. Self-Adaptive Multi-Population Jaya Algorithm based Reactive Power Reserve Optimization Considering Voltage Stability Margin Constraints was obtained in Tiwary et al. [19]. A smooth bootstrapping based technique for evaluating distribution system reliability indices neglecting random interruption duration is developed [20]. The impact of covered overhead conductors on distribution reliability and safety is discussed [21]. Sarantakos et al. [22] introduced a method to include component condition and substation reliability into distribution system reconfiguration. Battu et al. [23] discussed a method for reliability compliant distribution system planning using Monte Carlo simulation. Tiwary et al. [24] has discussed a methodology for reliability evaluation of an electrical power distribution system, which is radial in nature. Uspensky et al. [25] has developed a method for reliability assessment of the digital relay protection system. Sharma et al. [26] deals with the reliability analysis of a two identical unit system model with safe and unsafe failures, switching device and rebooting.

II. Reliability evaluation of electrical distribution system

Physically a system configuration will be a series reliability network if system fails even if a single component fails or system survives if all the components are working successfully.

The system is having a constant failure rate and therefore the reliability of the system having constant failure rate is evaluated by using the following relation.

$$R(t)=e^{(-\lambda t)} \tag{1}$$

Where R(t) represents the reliability of each distribution section. Λ represents the failure rate per year and t represents time period which is taken as one year.

If one assumes reliability of each component as r1,r2…rn, then reliability of series system (Rs) is given as

$$Rs=\prod ri \tag{2}$$

Where ri represents the reliability of components from i=1….n.

A system configuration will be a parallel reliability network, the system fails, if all components fail. System performs its function even if a single component is working.

The reliability of parallel system (Rp) is given as

$$R_p = 1 - \prod_{i=1}^{n}(1 - r_i) \tag{3}$$

Where ri represents the reliability of components from i=1….n.

III. Basic reliability indices evaluation of series and parallel system

When reliability studies are concerns, three basic reliability parameters are average failure rate, average outage time and average annual outage time, which are mentioned as follows for series system.

$$\lambda s = \sum \lambda i \tag{4}$$

$$Us = \sum 〚\lambda i.ri〛 \tag{5}$$

$$rs = (\sum \lambda iri)/(\sum \lambda i) \tag{6}$$

For parallel system the three basic reliability indices can be evaluated as follows

$$\lambda_{para} = \frac{\lambda_i \lambda_j (r_i + r_j)}{8760} \tag{7}$$

$$r_{para} = \frac{r_i r_j}{r_i + r_j} \tag{8}$$

$$U_{para} = \lambda_{para} \cdot r_{para} \tag{9}$$

Where λj is failure rate per year
ri , rj is average repair time, hours

IV. Evaluation of customer orientated indices

Customer orientated indices related to reliability studies are system average interruption frequency index, system average interruption duration index and customer average interruption duration index, which are mentioned as follows.

System average interruption frequency index (SAIFI)

$$SAIFI = \frac{total\ number\ of\ customer\ interruptions}{total\ number\ of\ customers\ served} \tag{10}$$

System average interruption duration index (SAIDI)

$$SAIDI = \frac{sum\ of\ customer\ interruption\ durations}{total\ number\ of\ customers} \tag{11}$$

Customer average interruption duration index (CAIDI)

$$CAIDI = \frac{sum\ of\ customer\ interruption\ durations}{total\ number\ of\ customer\ interruptions} \tag{12}$$

V. Results and Discussions

Meshed distribution system consists of 18 distributor segments and 4 load points from LP-1 to LP-4 Fig. 1, [10].

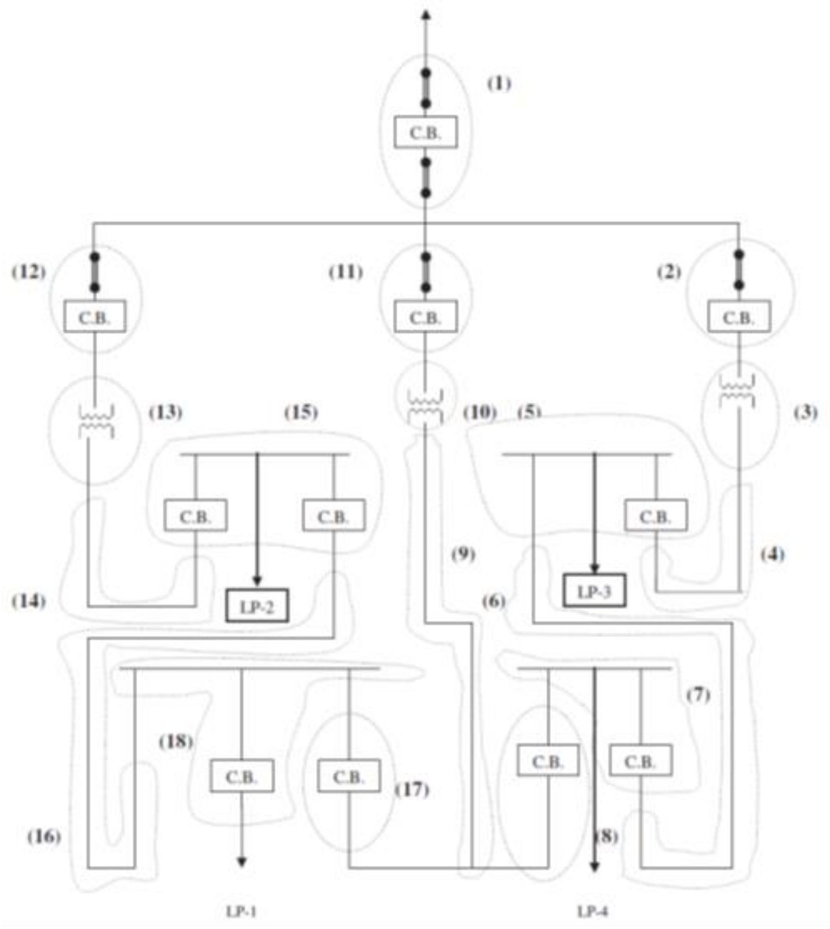


Figure 1 Sample Meshed Distribution System

Table I [10] provides the initial data for the meshed distribution system. Table I consists of failure rate per year and repair time in hours of each distribution section from 1 to 18 of the meshed distribution system. Table II consists of number of customers at each load point LP-1 to LP-4.
Evaluated Reliability at each distribution section is provided in Table III. The reliability at each distribution section is evaluated by using equation 1. Fig.2 provides magnitude of reliability at different distribution sections. Table IV gives the calculated value of the reliability at each and every load point of the meshed distribution system. The reliability at each and every load point is obtained by using the equation 2 and 3. Fig.3 provides magnitude of reliability at different load points of the distribution system.

Basic reliability indices at each load point i.e average failure rate, average outage time and average annual outage time are evaluated are presented in Table V. Fig.4 provides magnitude of average failure rate at different load points of the distribution system. Fig.5 and Fig.6 provides magnitude of average outage time and average annual outage time at different load points of the meshed distribution system.

Customer orientated indices such as system average interruption frequency index (SAIFI) is evaluated as 0.3965, system average interruption duration index (SAIDI) is obtained as 4.1057 and customer average interruption duration index (CAIDI) is evaluated as 10.3556 for the meshed distribution system.

Table 1 [10]

System data for sample meshed distribution system.

Distribution section	λ_i^0, failure/year	Average repair time r_i^0, h
#1	0.310400	10.280412
#2	0.127600	5.010658
#3	0.070000	33.985714
#4	0.013520	14.335503
#5	0.084600	10.557447
#6	0.017640	13.555102
#7	0.008460	10.557447
#8	0.078000	11.023077
#9	0.008460	15.800000
#10	0.069000	27.565217
#11	0.155200	6.865979
#12	0.155200	6.865979
#13	0.070000	33.985714
#14	0.013520	14.335503
#15	0.156600	10.714943
#16	0.017640	13.555100
#17	0.078000	11.023077
#18	0.084600	10.557447

Table 2 Initial data for the load points

Load Point	1	2	3	4
Number of customers	1000	800	600	700

Table 3 Evaluated Reliability at each distribution section

Distribution Section	Reliability
1	0.7331
2	0.8802
3	0.9324
4	0.9866
5	0.9189
6	0.9825
7	0.9916
8	0.9250
9	0.9916
10	0.9333
11	0.8562
12	0.8562
13	0.9324
14	0.9866
15	0.8550
16	0.9825
17	0.9250
18	0.9189

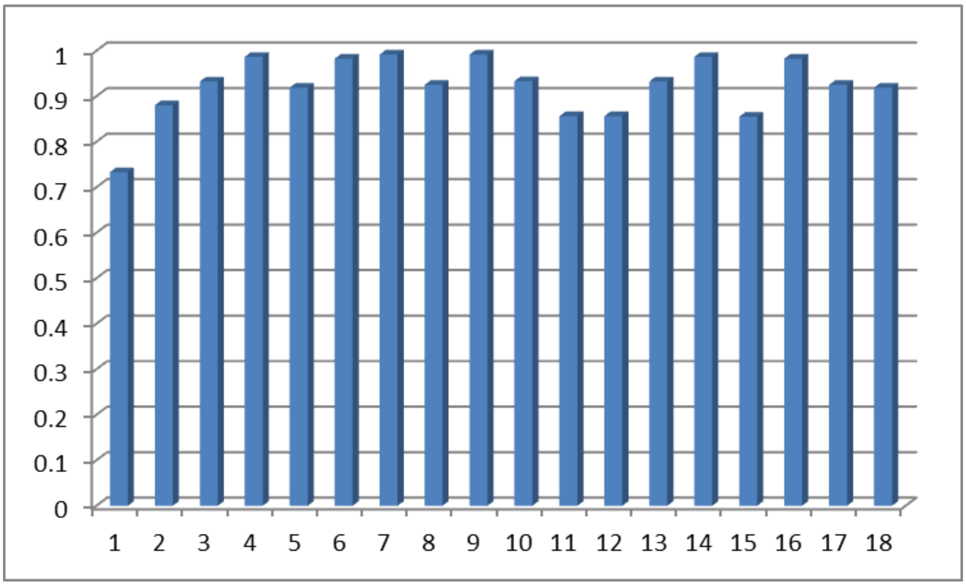


Figure 2 Magnitude of Reliability for distribution sections 1 to 18.

Table 4 Evaluated Reliability at each load points

Load point	Reliability
1	0.6421
2	0.5973
3	0.6504
4	0.6958

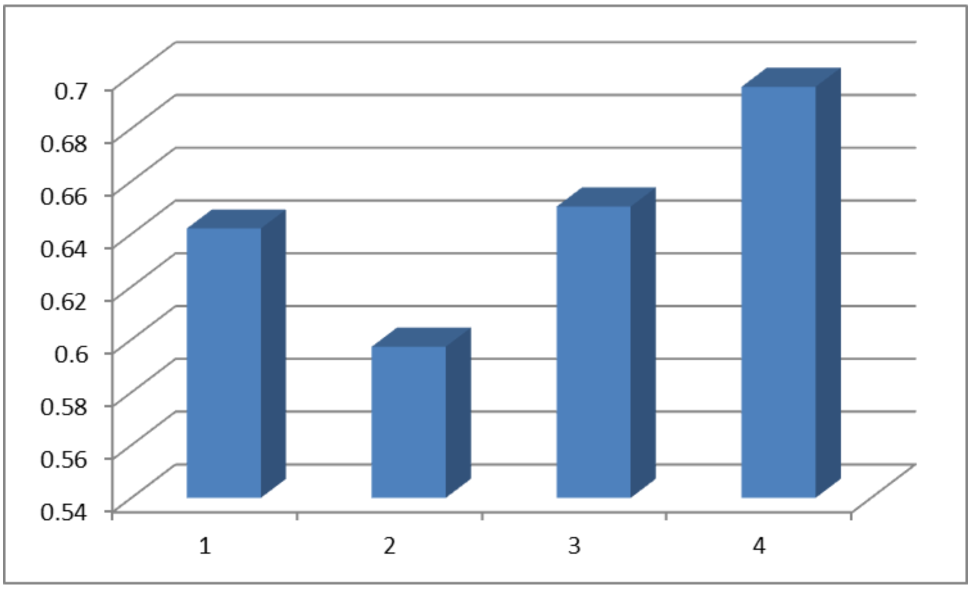


Figure3 Magnitude of Reliability at different Load Points 1 to 4.

Table 5 Evaluated Basic Reliability indices at each load points

Load point	1	2	3	4
average failure rate	0.3951	0.4671	0.3951	0.3189
average outage time	10.3384	10.4247	10.3381	10.2875
average annual outage time	4.0847	4.8697	4.0846	3.2807

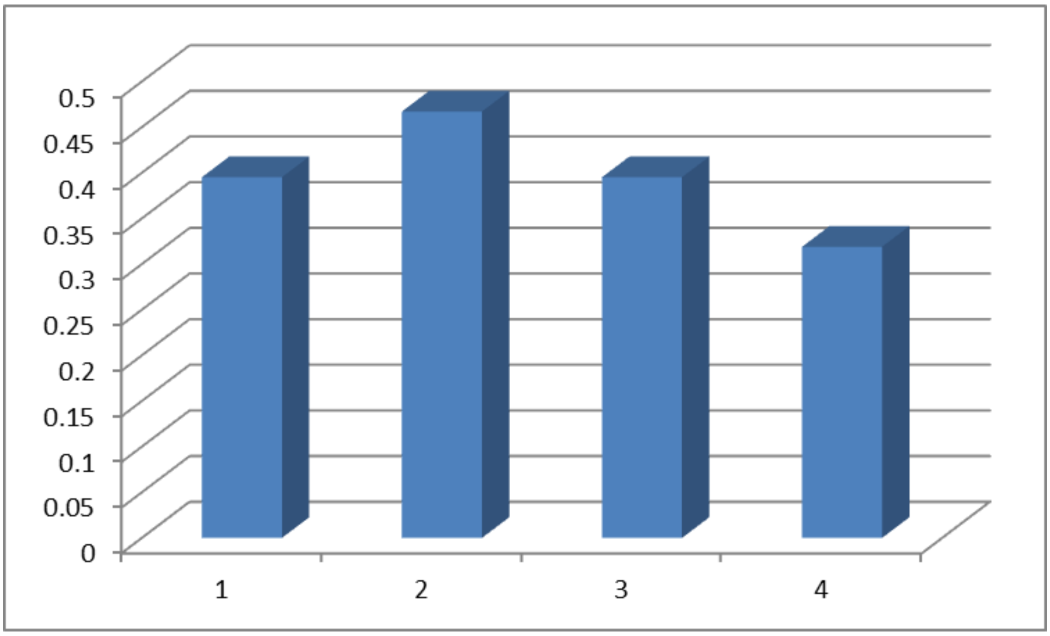


Figure 4 Magnitude of average failure rate at different Load Points 1 to 4.

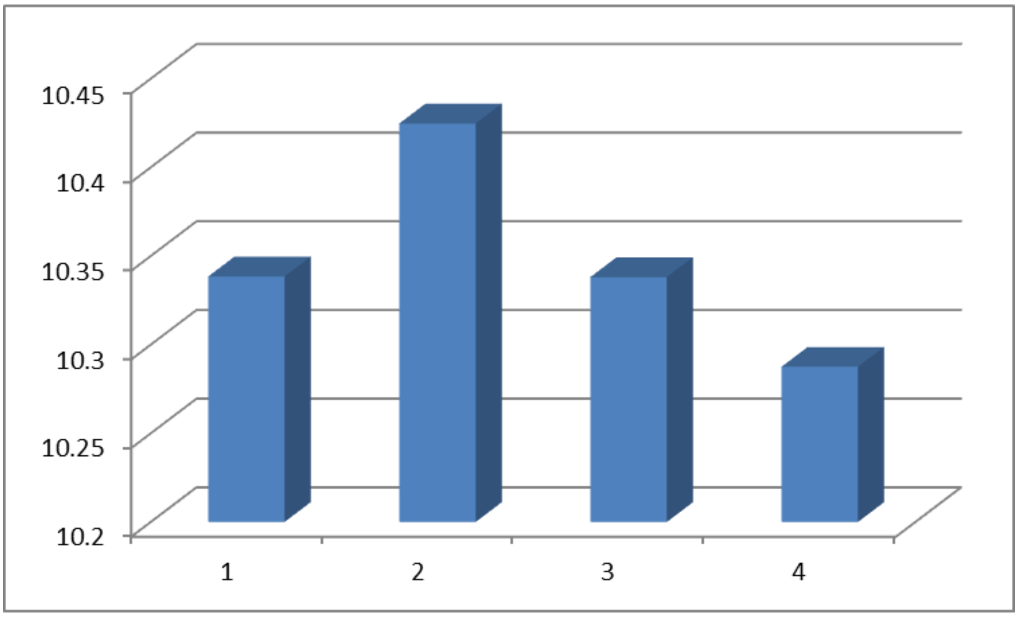


Figure 5 Magnitude of average outage time at different Load Points 1 to 4.

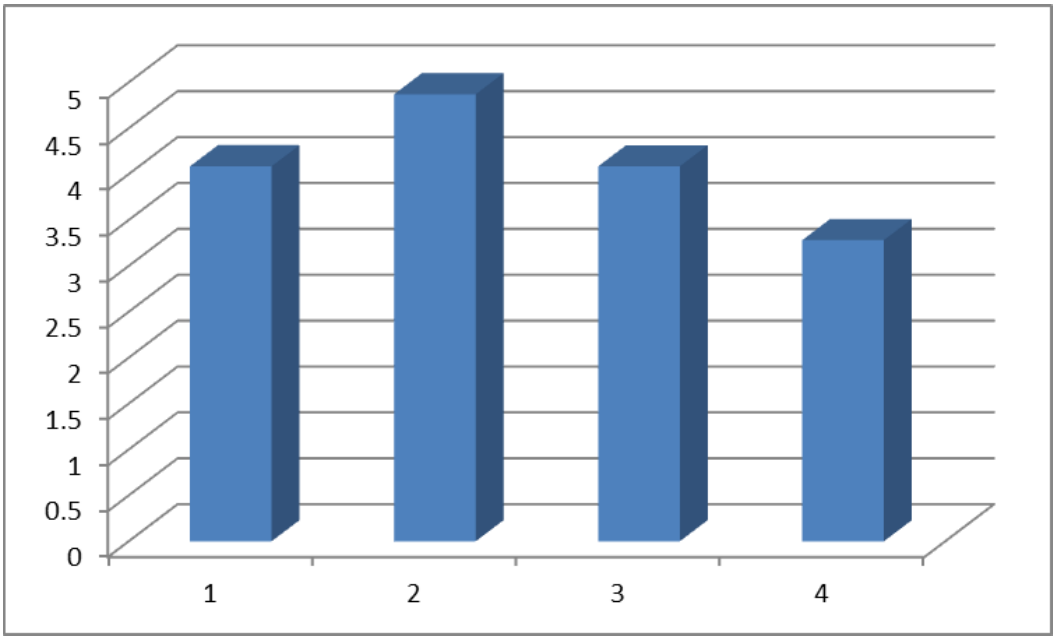


Figure 6 Magnitude of average annual outage time at different Load Points 1 to 4.

VI. Conclusion

Evaluation of reliability for a power distribution system is very essential. In this paper a meshed distribution system is taken in consideration. Reliability of each and every distribution section is calculated. Load point reliability is also obtained for the meshed distribution system for each and every load point. The three basic reliability parameters of importance average failure rate, average outage time and average annual outage time are also obtained for the load points considered. Important customer orientated indices such as system average interruption frequency index, system average interruption duration index and customer average interruption duration index are also evaluated for the meshed distribution system.

References

[1] C. Singh. (1981). Markov cut-set approach for the reliability evaluation of transmission and distribution systems. IEEE Trans. on Power Apparatus and Systems, 100: 2719-2725.

[2] R. Billinton. (1969). Composite system reliability evaluation. IEEE Trans. on Power Apparatus and Systems, 88:276-281.

[3] E. Wojczynski, R. Billinton. (1985). Effects of distribution system reliability index distributions upon interruption cost/reliability worth estimates. IEEE Trans. on Power Apparatus and Systems, 11:3229-3235.

[4] A. K. Verma, A. Srividya, H. M. R. Kumar. (2002). A framework using uncertainties in the composite power system reliability evaluation. Electric Power Components and Systems, 30:679-691.

[5] Z. Zheng, L. Cui, Alan G. Hawkes. (2006). A study on a single-unit Markov repairable system with repair time omission. IEEE Trans. on Reliability, 55:182-188.

[6] P. Jirutitijaroen, C. singh. (2008). Comparison of simulation methods for power system

reliability indexes and their distributions. IEEE Trans. on Power Systems, 23:486-493.
[7] O. Dzobe, C. T. Gaunt, R. Herman. (2012). Investigating the use of probability distribution functions in reliability-worth analysis of electric power systems. Int. J. of Electrical Power and Energy Systems, 37:110-116.
[8] I. S. Bae, J. O. Kim. (2008). Reliability evaluation of customers in a microgrid. IEEE Trans. on Power Systems, 23:1416-1422.
[9] R. Billinton, P. Wang. (1998). Reliability-network-equivalent approach to distribution-system-reliability evaluation. IEE Proc. generation, transmission and distribution, 145:149-153.
[10] L. D. Arya, S. C. Choube, R. Arya and Aditya Tiwary. (2012). Evaluation of Reliability indices accounting omission of random repair time for distribution systems using Monte Carlo simulation. Int. J. of Electrical Power and Energy System, (ELSEVIER), 42:533-541.
[11] Aditya Tiwary, R. Arya, S. C. Choube and L. D. Arya. (2012). Determination of Optimum period between Inspections for Distribution system based on Availability Accounting Uncertainties in Inspection Time and Repair Time. Journal of The Institution of Engineers (India): series B (Springer), 93:67-72.
[12] Jirutitijaroen P, singh C. (2008). Comparison of simulation methods for power system reliability indexes and their distribution. IEEE Trans Power Syst, 23:486–92.
[13] Aditya Tiwary, R. Arya, S. C. Choube and L. D. Arya. (2013). Determination of reliability indices for distribution system using a state transition sampling technique accounting random down time omission. Journal of The Institution of Engineers (India): series B (Springer), 94:71-83.
[14] Aditya Tiwary, L. D. Arya, R. Arya and S. C. Choube. (2016). Inspection repair based availability optimization of distribution systems using Teaching Learning based Optimization. Journal of The Institution of Engineers (India): series B (Springer), 97:355-365.
[15] Aditya Tiwary, R. Arya, L. D. Arya and S. C. Choube. (2017). Bootstrapping based technique for evaluating reliability indices of RBTS distribution system neglecting random down time. The IUP Journal of Electrical and Electronics Engineering, X:48-57.
[16] Volkanavski, Cepin M, Mavko B. (2009). Application of fault tree analysis for assessment of the power system reliability. Reliab Eng Syst Safety, 94:1116–27.
[17] Li BM, Su CT, Shen CL. (2010). The impact of covered overhead conductors on distribution reliability and safety. Int J Electr Power Energy Syst, 32:281–9.
[18] Aditya Tiwary. (2017). Reliability enhancement of distribution system using Teaching Learning based optimization considering customer and energy based indices. International Journal on Future Revolution in Computer Science & Communication Engineering, 3:58-62.
[19] Aditya Tiwary. (2018). Self-Adaptive Multi-Population Jaya Algorithm based Reactive Power Reserve Optimization Considering Voltage Stability Margin Constraints. International Journal on Future Revolution in Computer Science & Communication Engineering, 4:341-345.
[20] R. Arya , Aditya Tiwary , S. C. Choube and L. D. Arya. (2013). A smooth bootstrapping based technique for evaluating distribution system reliability indices neglecting random interruption duration. Int. J. of Electrical Power and Energy System, (ELSEVIER), 51:307-310.
[21] M. BinLi, C. TzongSu, C. LungShen. (2010). The impact of covered overhead conductors on distribution reliability and safety. Int. J. of Electrical Power and Energy System, (ELSEVIER), 32:281-289.
[22] I. Sarantakos, D. M.Greenwood, J.Yi, S. R. Blake, P. C. Taylor. (2019). A method to include component condition and substation reliability into distribution system reconfiguration. Int. J. of Electrical Power and Energy System, (ELSEVIER), 109:122-138.
[23] N. R. Battu, A. R. Abhyankar, N. Senroy. (2019). Reliability Compliant Distribution System Planning Using Monte Carlo Simulation. Electric power components and systems, 47:985-997.
[24] Aditya Tiwary. (2019). Reliability evaluation of radial distribution system – A case study. Int. J. of Reliability: Theory and Applications, 14,4(55):9-13.

[25] Michael Uspensky. (2019). Reliability assessment of the digital relay protection system. Int. J. of Reliability: Theory and Applications, 14:10-17.

[26] Akshita Sharma, Pawan Kumar. (2019). Analysis of reliability measures of two identical unit system with on switching device and imperfect coverage. Int. J. of Reliability: Theory and Applications, 14:44-52.

On Exponential-Weibull Distribution Useful in Reliability and Survival Analysis

Anakha K. K & Chacko V. M

•

Department of Statistics, St. Thomas' College (Autonomous), Thrissur
Kerala-680 001, India
Email: anakhaappu@gmail.com, chackovm@gmail.com

Abstract

In this paper, mixture of Exponential and Weibull distributions is considered for modelling real lifetime data. The basic mathematical properties including moments, generating functions, order statistics etc are derived. We obtain the reliability of stress-strength model. The maximum likelihood method is performed to estimate the parameters and a simulation study is conducted to validate the maximum likelihood estimators. The model is fitted to a real data set.

Keywords: Failure rate function, Reliability, Survival Analysis, Stress-Strength

I. Introduction

Lifetime distributions have a significant role in Reliability theory and survival analysis. Exponential, Gamma, Weibull and log-Normal distributions are some of the distributions commonly used for modeling lifetime data. Exponential and Weibull distributions are more popular than Gamma and log-Normal because the survival function of Gamma and log-Normal distributions doesn't have a closed form.

Mixture distributions have been used widely in reliability and survival analysis studies recently. It has been getting great attention, since mixture models are more appropriate and multiple causes of failure can be simultaneously modeled. Due to high flexibility, survival mixture models are better choice to analyze the reliability or survival data in situations when the data are believed to be heterogeneous and a single parametric distribution may not be sufficient to analyze the data.

There are several mixture distributions and generalizations of existing distributions, available in literature. A study on Generalized Lindley distribution useful in reliability study is given by Nadarajah et.al (2011). Nadarajah and Gupta (2007) studied on Exponentiated Gamma distribution with application to drought data. Mustafa et.al (2016) proposed Weibull Generalized Exponential Distribution. Gupta and Kundu (2001) studied on Exponentiated Exponential family as an alternative to gamma and Weibull. A detailed study on Statistical Models and Methods for Lifetime Data can be seen in Lawless (2003). Chacko et. al (2018) proposed Weibull-Lindly Distribution for modeling a bathtub shaped failure rate data.

In section 2, the mixture of Exponential and Weibull distributions is considered. The failure rate or hazard rate function is given in section 3. Moments are given in section 4 and generating functions are given in section 5. Conditional moments are given in section 6 and Quantile function is given in section 7. Mean deviation is given in section 8 and distributions of order statistics are given in section 9. Bonferroni and Lorenz Curves are given in section 10. Reliability in stress-strength model is given in section 11. Estimation of parameters using maximum likelihood estimation method is described in section 12. Simulation study and real data analysis are given in section 13 and 14 respectively. Conclusions are given in last section.

2. Exponential-Weibull distribution

Here we consider the mixture of two lifetime distributions, namely Exponential and Weibull distributions. The cumulative distribution function (cdf) of mixture of Exponential and Weibull distribution can be represented as

$$F(x) = \theta F_E(x) + (1-\theta)F_W(x),$$

where $\theta = \lambda/(1+\lambda)$, $\lambda > 0$, $F_E(x) = 1 - e^{-\lambda x}, \lambda > 0, x > 0$, the cdf of Exponential distribution with scale parameter λ and $F_W(x) = 1 - e^{-(\lambda x)^\alpha}, \lambda > 0,$, $\alpha > 0$, x>0, the cdf of Weibull distribution with scale parameter λ and shape parameter α. The mixture of Exponential and Weibull distributions can be written as,

$$F(x) = 1 - \frac{\lambda e^{-\lambda x} + e^{-(\lambda x)^\alpha}}{1+\lambda}. \tag{1}$$

The corresponding probability density function(pdf) is

$$f(x) = \frac{\lambda^2 e^{-\lambda x} + \alpha \lambda^\alpha x^{\alpha-1} e^{-(\lambda x)^\alpha}}{1+\lambda}. \tag{2}$$

Let X be a random variable, then we say that X has a 'Exponential-Weibull distribution' (EW(α, λ)) with scale parameter α and shape parameter λ, if it has the pdf (2).

The reliability function of the EW(α, λ) distribution is

$$\overline{F}(x) = \frac{\lambda e^{-\lambda x} + e^{-(\lambda x)^\alpha}}{1+\lambda},$$

and corresponding failure rate or hazard rate function is

$$h(x) = \frac{\lambda^2 e^{-\lambda x} + \alpha \lambda^\alpha x^{\alpha-1} e^{-(\lambda x)^\alpha}}{\lambda e^{-\lambda x} + e^{-(\lambda x)^\alpha}}. \tag{3}$$

The mode for the mixture model EW(α, λ) can be found by solving the derivative of the (2)

$$f'(x) = \frac{-\lambda^3 e^{-\lambda x} - \alpha^2 \lambda^{2\alpha} x^{2\alpha-2} e^{-(\lambda x)^\alpha} + \alpha(\alpha-1)\lambda^\alpha x^{\alpha-2} e^{-(\lambda x)\alpha}}{1+\lambda} \tag{4}$$

By solving (4), we observe that the mixture model EW(α, λ) is unimodal. Figure 1 shows the pdf of EW(α, λ) for various choices of parameters.

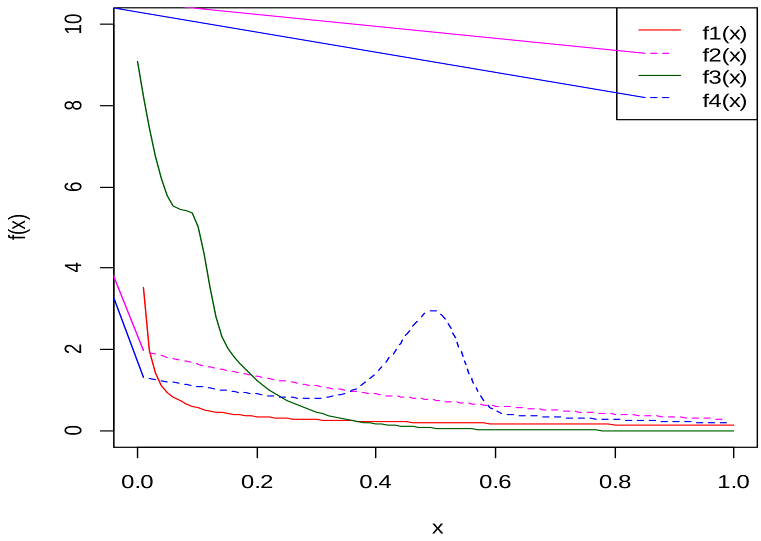


Figure 1: pdf f(α, λ) of EW(α, λ) for values of parameters

$f_1(x) = f(0.2, 0.4)$, $f_2(x) = f(1, 2)$, $f_3(x) = f(5, 10)$ and $f_4(x) = f(10, 2)$

3. Failure Rate Function

The failure rate or hazard rate function of the mixture EW(α, λ) distribution is given as follows:

$$h(x)=\frac{\lambda^2 e^{-\lambda x}+\alpha\lambda^{\alpha}x^{\alpha-1}e^{-(\lambda x)^{\alpha}}}{\lambda e^{-\lambda x}+e^{-(\lambda x)^{\alpha}}},\lambda>0,\alpha>0,x>0.$$

EW(α, λ) distribution has increasing, decreasing, upside-down bathtub shape behaviors for its failure rate. When $\alpha = 1$; $h(x) = \lambda$, a constant, i.e. it has Exponential distribution with lack of memory property. From (3) $\lim_{x\to 0} h(x)=\frac{\lambda^2}{\lambda+1}$, a constant and $\lim_{x\to\infty} h(x)=0$. Figure 2 shows the failure function of EW(α, λ) distribution with various choice of parameters. These shapes of failure function show that EW(α, λ) distribution fit in with both monotonic and non-monotonic behaviors which are more likely to be come across when dealing with lifetime data.

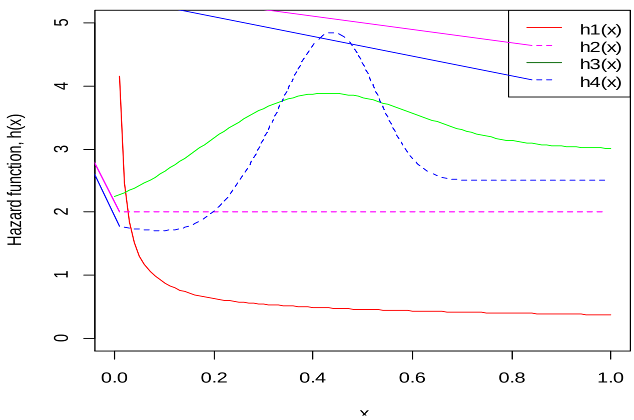


Figure 2: Failure function of EW(α, λ) distribution for various choice of parameters

h1(x) = h(0.2, 0.6), h2(x) = h(1, 2), h3(x) = (2, 3) and h4(x) = (4, 2.5) for h(α, λ)

4. Moments

The r[th] raw moment of the EW(α, λ) distribution with pdf in (2) is given by,

$$\mu_r = E(X^r) = \int_0^\infty x^r f(x)dx$$

$$= \int_0^\infty x^r \left(\frac{\lambda^2 e^{-\lambda x} + \alpha\lambda^\alpha x^{\alpha-1} e^{-(\lambda x)^\alpha}}{1+\lambda} \right) dx$$

$$= \frac{\lambda r! + \Gamma\left(1+\frac{r}{\alpha}\right)}{(1+\lambda)\lambda^r}.$$

The first four raw moments are,

$$E(X) = \frac{\lambda + \Gamma\left(1+\frac{1}{\alpha}\right)}{(1+\lambda)\lambda},\ E(X^2) = \frac{2\lambda + \Gamma\left(1+\frac{2}{\alpha}\right)}{(1+\lambda)\lambda^2},\ E(X^3) = \frac{6\lambda + \Gamma\left(1+\frac{3}{\alpha}\right)}{(1+\lambda)\lambda^3}$$ and

$$E(X^4) = \frac{24\lambda + \Gamma\left(1+\frac{4}{\alpha}\right)}{(1+\lambda)\lambda^4}.$$

The variances of EW(α, λ) distribution is given by;

$$Var(X) = \frac{(1+\lambda)\left(2\lambda + \Gamma\left(1+\frac{2}{\alpha}\right)\right) - \left(\lambda + \Gamma\left(1+\frac{1}{\alpha}\right)\right)^2}{(1+\lambda)^2\lambda^2}.$$

Central moments can be obtained using raw moments.

5. Generating Functions

Let X be a random variable with probability density function (2). Its moment generating function (mgf) is given by,

$$E(e^{tx}) = \int_0^\infty e^{itx} \left(\frac{\lambda^2 e^{-\lambda x} + \alpha\lambda^\alpha x^{\alpha-1} e^{-(\lambda x)^\alpha}}{1+\lambda} \right) dx$$

$$= \frac{\lambda^2}{(1+\lambda)(\lambda - t)} + \frac{1}{1+\lambda}\sum_{n=0}^\infty \frac{t^n}{\lambda^n n!}\Gamma\left(1+\frac{n}{\alpha}\right), \alpha \geq 1,\ t < \lambda.$$

The characteristic function (cf) of X is $\phi(t) = E\left(e^{itx}\right)$, which is

$$E(e^{itx}) = \int_0^\infty e^{itx} \left(\frac{\lambda^2 e^{-\lambda x} + \alpha\lambda^\alpha x^{\alpha-1} e^{-(\lambda x)^\alpha}}{1+\lambda} \right) dx$$

$$= \frac{\lambda^2}{(1+\lambda)(\lambda - it)} + \frac{1}{1+\lambda}\sum_{n=0}^\infty \frac{(it)^n}{\lambda^n n!}\Gamma\left(1+\frac{n}{\alpha}\right).$$

The cumulant generating function (cgf) of X is given by,

$$K_X(t) = \log\phi_X(t)$$

$$= \log\left(\frac{1}{1+\lambda}\right) + \log\left(\frac{\lambda^2}{\lambda - it} + \sum_{n=0}^{\infty}\frac{(it)^n}{\lambda^n n!}\Gamma\left(1+\frac{n}{\alpha}\right)\right).$$

6. Conditional Moments

The conditional expectation for the EW(α, λ) distribution is given by,

$$E(X^n / X > x) = \frac{\int_x^\infty x^n \left\{\frac{\lambda^2 e^{-\lambda x} + \alpha\lambda^\alpha x^{\alpha-1}e^{-(\lambda x)^\alpha}}{1+\lambda}\right\}dx}{\overline{F}(X)}$$

$$= \frac{1+\lambda}{\lambda e^{-\lambda x} + e^{-(\lambda x)^\alpha}}\left\{\frac{\lambda^2\Gamma(n+1,\lambda x)}{(1+\lambda)\lambda^{n+1}} + \frac{\Gamma\left(\frac{n}{\alpha}+1,(\lambda x)^\alpha\right)}{(1+\lambda)\lambda^n}\right\}$$

$$= \frac{1}{\lambda e^{-\lambda x} + e^{-(\lambda x)^\alpha}}\left\{\frac{\Gamma(n+1,\lambda x)}{\lambda^{n-1}} + \frac{\Gamma\left(\frac{n}{\alpha}+1,(\lambda x)^\alpha\right)}{\lambda^n}\right\}.$$

In particular,

$$E(X / X > x) = \frac{1}{\lambda e^{-\lambda x} + e^{-(\lambda x)^\alpha}}\left\{\Gamma(2,\lambda x) + \frac{\Gamma\left(\frac{1}{\alpha}+1,(\lambda x)^\alpha\right)}{\lambda}\right\},$$

$$E(X^2 / X > x) = \frac{1}{\lambda e^{-\lambda x} + e^{-(\lambda x)^\alpha}}\left\{\frac{\Gamma(3,\lambda x)}{\lambda} + \frac{\Gamma\left(\frac{2}{\alpha}+1,(\lambda x)^\alpha\right)}{\lambda^2}\right\},$$

$$E(X^3 / X > x) = \frac{1}{\lambda e^{-\lambda x} + e^{-(\lambda x)^\alpha}}\left\{\frac{\Gamma(4,\lambda x)}{\lambda^2} + \frac{\Gamma\left(\frac{3}{\alpha}+1,(\lambda x)^\alpha\right)}{\lambda^3}\right\}, \quad and$$

$$E(X^4 / X > x) = \frac{1}{\lambda e^{-\lambda x} + e^{-(\lambda x)^\alpha}}\left\{\frac{\Gamma(5,\lambda x)}{\lambda^3} + \frac{\Gamma\left(\frac{4}{\alpha}+1,(\lambda x)^\alpha\right)}{\lambda^4}\right\}.$$

7. Quantile Function

The p^th quantile say Q(p), p ϵ (0,1) is defined by Q(p) = p. Let X be a EW random variable with pdf (2), then its quantile function is the root of the equation,

$$1-\frac{\lambda e^{-\lambda Q(p)}+e^{-(\lambda Q(p))^{\alpha}}}{1+\lambda}=p$$

$$(1-p)(1+\lambda)=\lambda e^{-\lambda Q(p)}+e^{-(\lambda Q(p))^{\alpha}}$$

8. Mean Deviation

The average amount of scatter in a population from either the mean or the median is termed as mean deviation. The mean deviation about mean and, mean deviation about median are defined by,

$$\delta_1(X)=\int_0^{\infty}|x-\mu|\ f(x)dx$$ and

$$\delta_2(X)=\int_0^{\infty}|x-M|\ f(x)dx,$$

respectively, where μ and M are mean and median respectively. $\delta_1(X)$ and $\delta_2(X)$ can be calculated using the formulae,

$$\delta_1(X)=2\mu F(\mu)-2\mu+2\int_{\mu}^{\infty}xf(x)dx$$

and $\delta_2(X)=-\mu+2\int_{M}^{\infty}xf(x)dx$ respectively.

The mean deviation about mean of the EW(α, λ) distribution is,

$$\delta_1(X)=2\mu\left(1-\frac{\lambda e^{-\lambda\mu}+e^{-(\lambda\mu)^{\alpha}}}{1+\lambda}\right)-2\mu+2\int_{\mu}^{\infty}x\left(\frac{\lambda^2e^{-\lambda x}+\alpha\lambda^{\alpha}x^{\alpha-1}e^{-(\lambda x)^{\alpha}}}{1+\lambda}\right)dx$$

$$=-\frac{2\left(\lambda^2\mu e^{-\lambda\mu}+\lambda\mu e^{-(\lambda\mu)^{\alpha}}\right)}{\lambda(1+\lambda)}+\frac{2\left(\lambda\mu e^{-\lambda\mu}+e^{-(\lambda\mu)}\right)}{1+\lambda}+\frac{2\Gamma\left(\frac{1}{\alpha}+1,(\lambda\mu)^{\alpha}\right)}{\lambda(1+\lambda)}$$

$$=\frac{2\left(e^{-\lambda\mu}-\mu e^{-(\lambda\mu)^{\alpha}}\right)}{1+\lambda}+\frac{2\Gamma\left(\frac{1}{\alpha}+1,(\lambda\mu)^{\alpha}\right)}{\lambda(1+\lambda)},$$

where $\gamma(a,x)=\int_0^x t^{a-1}e^{-t}dt$ and $\Gamma(a,x)=\int_x^\infty t^{a-1}e^{-t}dt$ are denotes the incomplete gamma functions. The mean deviation about median of the EW(α, λ) distribution is,

$$\delta_2(X)=-\mu+2\int_M^\infty x\left(\frac{\lambda^2 e^{-\lambda x}+\alpha\lambda^\alpha x^{\alpha-1}e^{-(\lambda z)^\alpha}}{1+\lambda}\right)$$
$$=-\mu+\frac{2}{\lambda(1+\lambda)}\left(\lambda^2 Me^{-\lambda M}+\lambda e^{-\lambda M}+\Gamma\left(\frac{1}{\alpha}+1,(M\lambda)^\alpha\right)\right).$$

9. Order Statistics

Let $X_1, X_2, \ldots X_n$ be a random sample of EW(α, λ) distribution with cdf and pdf as in (1) and (2) respectively. Their corresponding order statistics is denoted by $X_{(1)} < X_{(2)} < X_{(3)} \cdots < X_{(n)}$. The pdf and cdf of the rth order statistic are,

$$f_{X_{(r)}}(x;\alpha,\lambda)=\frac{n!}{(r-1)!(n-r)!}f(x;\alpha,\lambda)F^{r-1}(x;\alpha,\lambda)\overline{F}^{n-r}(x;\alpha,\lambda)$$
$$=\frac{n!}{(r-1)!(n-r)!}\sum_{l=0}^{n-r}\binom{n-r}{l}(-1)^l F^{r-1+l}(x;\alpha,\lambda)f(x;\alpha,\lambda)$$

and

$$F_{X_{(r)}}=\sum_{j=r}^{n}\binom{n}{j}F^j(x;\alpha,\lambda)\{1-F(x;\alpha,\lambda)\}^{n-j}$$
$$=\sum_{j=r}^{n}\sum_{l=0}^{n-j}\binom{n}{j}\binom{n-j}{l}(-1)^l F^{j+l}(x;\alpha,\lambda),$$ respectively for r = 1,2,3..n.

The pdf of $X_{(r)}$ of EW(α, λ) distribution is ,

$$f_{X_{(r)}}=\frac{n!\left(\lambda^2 e^{-\lambda x}+\alpha\lambda^\alpha x^{\alpha-1}e^{-(\lambda x)^\alpha}\right)^{n-r}}{(r-1)!(1+\lambda)}\sum_{l=0}^{n-r}\frac{(-1)^l}{l!(n-r-l)!}\left\{1-\frac{\lambda e^{-\lambda x}+e^{-(\lambda x)^\alpha}}{1+\lambda}\right\}^{r-1+l}$$

and corresponding cdf is

$$F_{X_{(r)}}(x;\alpha,\lambda)=\sum_{j=r}^{n}\sum_{l=o}^{n-j}\binom{n}{j}\binom{n-j}{l}(-1)^l\left\{1-\frac{\lambda e^{-\lambda x}+e^{-(\lambda x)^\alpha}}{1+\lambda}\right\}^{j+l}.$$

The pdf of 1st order statistic is

$$f_{X_{(1)}}=\frac{n!}{(1+\lambda)}\left(\lambda^2 e^{-\lambda x}+\alpha\lambda^\alpha x^{\alpha-1}e^{-(\lambda x)^\alpha}\right)\sum_{l=0}^{n-1}\frac{(-1)^l}{l!(n-1-r)!}\left\{1-\frac{\lambda e^{-\lambda x}+e^{-(\lambda x)^\alpha}}{1+\lambda}\right\}^l$$

and the pdf of nth order statistic is

$$f_{X_{(n)}}=\frac{n}{(1+\lambda)^n}\left(\lambda^2 e^{-\lambda x}+\alpha\lambda^\alpha x^{\alpha-1}e^{-(\lambda x)^\alpha}\right)\left(1+\lambda-\lambda e^{-\lambda x}-e^{-(\lambda x)^\alpha}\right).$$

The distribution of order statistics can be used for obtaining reliability of series or parallel system.

10. Bonferroni & Lorenz Curves

Bonferroni and Lorenz curves are the popular tools for analyzing data emerging in Economics and Reliability. They are the fundamental tool for income analysis. The Bonferroni and Lorenz curves are defined by

$$B(p) = \frac{1}{p\mu}\int_0^q xf(x)dx$$

and

$$L(p) = \frac{1}{\mu}\int_0^q xf(x)dx$$

respectively, or correspondingly by

$$B(p) = \frac{1}{p\mu}\int_0^p F^{-1}(x)dx$$

and

$$L(p) = \frac{1}{\mu}\int_0^p F^{-1}(x)dx$$

respectively, where $\mu = E(X)$ and $q = F^{-1}(p)$. For a random variable X with pdf (2), the Bonferroni and Lorenz curves are

$$B(p) = \frac{1}{p\mu(1+\lambda)}\left\{\left(1-(\lambda q+1)e^{-\lambda q}\right)+\frac{\Gamma\left(\frac{1}{\alpha}+1,(\lambda q)^{\alpha}\right)}{\lambda}\right\}$$

and

$$L(p) = \frac{1}{\mu(1+\lambda)}\left\{\left(1-(\lambda q+1)e^{-\lambda q}\right)+\frac{\Gamma\left(\frac{1}{\alpha}+1,(\lambda q)^{\alpha}\right)}{\lambda}\right\}$$

respectively.

11. Stress-Strength Reliability

Let X_1 and X_2 be two independent random variables, where X_1 represents the strength and X_2 represents the stress. Suppose X_1 and X_2 follows EW distribution, with parameters (α_1, λ_1) and (α_2, λ_2) respectively. Then the system reliability is

$$R = P[X_1 > X_2] = \int_0^\infty f_1(x)F_2(x)dx$$

$$= \int_0^\infty \left(\frac{\lambda_1^2 e^{-\lambda_1 x} + \alpha_1 \lambda_1^{\alpha_1} x^{\alpha_1 - 1} e^{-(\lambda_1 x)^{\alpha_1}}}{1+\lambda_1} \right)\left(1 - \frac{\lambda_2 e^{-\lambda_2 x} + e^{-(\lambda_2 x)^{\alpha_2}}}{1+\lambda_2}\right)dx$$

$$= \frac{1}{(1+\lambda_1)(1+\lambda_2)} \int_0^\infty \left(\lambda_1^2 e^{-\lambda_1 x} + \alpha_1 \lambda_1^{\alpha_1} x^{\alpha_1 - 1} e^{-(\lambda_1 x)^{\alpha_1}}\right)\left(1+\lambda_2 - \lambda_2 e^{-\lambda_2 x} - e^{-(\lambda_2 x)^{\alpha_2}}\right)dx$$

$$= \frac{1}{(1+\lambda_1)(1+\lambda_2)} \left\{ \begin{array}{l} \lambda_1(1+\lambda_2) - \dfrac{\lambda_1^2 \lambda_2}{(\lambda_1+\lambda_2)} + \displaystyle\sum_{i=0}^\infty \frac{(-1)^{i+1}\lambda_1^2 \lambda_2^{\alpha_2 i}}{i!} \frac{\Gamma(\alpha_2 i + 1)}{(\lambda_1)^{\alpha_2 i+1}} + (1+\lambda_2) + \\ \displaystyle\sum_{i=0}^\infty \frac{(-1)^{i+1}\alpha_1 \lambda_1^{\alpha_1(i+1)}}{i!\,\lambda_2^{\alpha_1 i + \alpha_2 - 1}} \Gamma(\alpha_1 i + \alpha_2) + \sum_{i=0}^\infty \frac{(-1)^{i+1}\left(\dfrac{\lambda_2}{\lambda_1}\right)^{\alpha_2 i}}{i!} \Gamma\left(\frac{\alpha_2}{\alpha_1} i + 1\right) \end{array} \right\}.$$

12. Estimation of Parameters

If $X_1, X_2, \ldots X_n$ is a random sample from EW(α, λ), then its likelihood function is

$$L(\bar{x}; \alpha, \lambda) = \prod_{i=1}^n \frac{\lambda^2 e^{-\lambda x_i} + \alpha \lambda^\alpha x_i^{\alpha-1} e^{-(\lambda x)^\alpha}}{1+\lambda}.$$

It is familiar that maximum likelihood estimate of the parameters is the value of the parameter which maximizes the likelihood function. The partial derivatives of log L with respect to unknown parameters α and λ are,

$$\frac{\partial \log L(\bar{x}; \alpha, \lambda)}{\partial \alpha} = \sum_{i=1}^n \frac{\lambda^\alpha x_i^{\alpha-1} e^{-(\lambda x_i)^\alpha} \left\{\alpha \log \alpha(\alpha-1) - \alpha \lambda^\alpha x_i^\alpha \log(\lambda x_i) + 1\right\}}{\lambda^2 e^{-\lambda x_i} + \alpha \lambda^\alpha x_i^{\alpha-1} e^{-(\lambda x_i)^\alpha}}$$
, and

$$\frac{\partial \log L(\bar{x}.; \alpha, \lambda)}{\partial \lambda} = \sum_{i=1}^n \frac{\lambda e^{-\lambda x_i}\left[(1+\lambda)(2 - x_i\lambda) - \lambda\right] + \alpha(\lambda x_i)^{\alpha-1} e^{-(\lambda x_i)^\alpha}\left[\alpha(1+\lambda)\left(1 - (\lambda x_i)^\alpha\right) - \lambda\right]}{(1+\lambda)\left(\lambda^2 e^{-\lambda x_i} + \alpha \lambda^\alpha x_i^{\alpha-1} e^{-(\lambda x_i)^\alpha}\right)}$$

By equating the above equations to zero we get two non-linear equations.

$$\sum_{i=1}^n \frac{\lambda^\alpha x_i^{\alpha-1} e^{-(\lambda x_i)^\alpha} \left\{\alpha \log \alpha(\alpha-1) - \alpha \lambda^\alpha x_i^\alpha \log(\lambda x_i) + 1\right\}}{\lambda^2 e^{-\lambda x_i} + \alpha \lambda^\alpha x_i^{\alpha-1} e^{-(\lambda x_i)^\alpha}} = 0$$ and

$$\sum_{i=1}^n \frac{\lambda e^{-\lambda x_i}\left[(1+\lambda)(2 - x_i\lambda) - \lambda\right] + \alpha(\lambda x_i)^{\alpha-1} e^{-(\lambda x_i)^\alpha}\left[\alpha(1+\lambda)\left(1 - (\lambda x_i)^\alpha\right) - \lambda\right]}{(1+\lambda)\left(\lambda^2 e^{-\lambda x_i} + \alpha \lambda^\alpha x_i^{\alpha-1} e^{-(\lambda x_i)^\alpha}\right)} = 0.$$

A solution of the non-linear equation gives the maximum likelihood estimates of α and λ. The normal equations cannot be solved by analytically, so that Newton-Raphson's iteration method or any other numerical approximation methods are required. Since the maximum likelihood estimator (MLE)s cannot be expressed in explicit forms, we consider their asymptotic distribution and confidence interval for α > 0 and λ > 0. For large samples, the MLE $\left(\hat{\alpha}, \hat{\lambda}\right)$ of (α, λ) is asymptotically normal with mean zero and variance covariance matrix I^{-1}, where

$$I = \begin{bmatrix} E\left(-\frac{\partial^2 \log L}{\partial \alpha^2}\right) & E\left(-\frac{\partial^2 \log L}{\partial \alpha \partial \lambda}\right) \\ E\left(-\frac{\partial^2 \log L}{\partial \lambda \partial \alpha}\right) & E\left(-\frac{\partial^2 \log L}{\partial \lambda^2}\right) \end{bmatrix}.$$

We can derive the approximate (1- δ)100% confidence interval of the parameters α and λ. By using variance covariance matrix, the confidence intervals are $\hat{\alpha} \pm Z_{\frac{n}{2}} \sqrt{\operatorname{var}\left(\hat{\alpha}\right)}$ and $\hat{\lambda} \pm Z_{\frac{n}{2}} \sqrt{\operatorname{var}\left(\hat{\lambda}\right)}$ where $Z_{\frac{n}{2}}$ is the upper $100\left(\frac{\delta}{2}\right)^{th}$ percentile of the standard normal distribution.

13. Simulation Study

Here we performed a simulation study to validate the maximum likelihood estimation procedure for EW(α, λ) distribution using Newton-Raphson method. For this purpose, we generated samples of sizes 25, 50, 100, 500, 1000 for different combinations of α and λ. We computed the maximum likelihood estimates for each sample and repeated this process thousand times then computed the bias and mean square error(MSE)s of the parameter estimates.

The simulation is conducted for the selected values of α and λ. Here we considered $\alpha = 1$, $\lambda = 1$; $\alpha = 1$, $\lambda = 1.2$; $\alpha = 2$, $\lambda = 1.5$ and $\alpha = 3$, $\lambda = 1.5$ as initial parameter values. The Tables 1, 2, 3 and 4 gives the values of the estimates, bias and MSEs of the corresponding parameters. From the tables, it can be seen that, as sample size increases the bias and MSE of the estimates decreases.

Table 1: *Estimates, Bias and MSE for $\alpha = 1$ and $\lambda = 1$*

N	Estimates	Bias	MSE
25	α = 1.293439	0.2934385	0.81239
	λ = 1.038358	0.03835775	0.05436978
50	α = 1.087168	0.08716756	0.0942621
	λ = 1.019775	0.0197749	0.0244342
100	α = 1.043303	0.08716756	0.0942621
	λ = 1.009724	0.0197749	0.0244342
500	α = 1.011771	0.01177124	0.005638945
	λ = 1.000085	8.461249e-05	0.002088029
1000	α = 1.006076	0.006076378	0.002621145
	λ = 1.000816	0.0008162859	0.001182026

Table 2 *Estimates, Bias and MSE for α = 1 and λ = 1.2*

N	Estimates	Bias	MSE
25	$\alpha =$ 1.275423 $\lambda =$ 1.250335	0.2754229 0.05033506	0.7921809 0.07960997
50	$\alpha =$ 1.123425 $\lambda =$ 1.227627	0.1234254 0.02762709	0.1598412 0.03496351
100	$\alpha =$ 1.054752 $\lambda =$ 1.203634	0.1234254 0.02762709	0.1598412 0.03496351
500	$\alpha =$ 1.008541 $\lambda =$ 1.198176	0.008540759 -0.001823803	0.00660528 0.003205042
1000	$\alpha =$ 1.006074 $\lambda =$ 1.202171	0.006073959 0.002171475	0.003285917 0.001605966

Table 3: *Estimates, Bias and MSE for α =2 and λ = 1.5*

N	Estimation	Bias	MSE
25	$\alpha =$ 2.737266 $\lambda =$ 1.53967	0.7372658 0.03966983	6.287694 0.07501019
50	$\alpha =$ 2.304472 $\lambda =$ 1.512797	0.3044722 0.01279734	1.261294 0.03309165
100	$\alpha =$ 2.122918 $\lambda =$ 1.508089	0.122918 0.008089073	0.3660359 0.01584789
500	$\alpha =$ 2.029113 $\lambda =$ 1.50079	0.0291133 0.0007895803	0.04937387 0.003182324
1000	$\alpha =$ 2.012721 $\lambda =$ 1.500226	0.01272087 0.0002263662	0.02528183 0.001405722

Table 4: *Estimates, Bias and MSE for α = 3 and λ = 1.5*

N	Estimates	Bias	MSE
25	$\alpha =$ 4.016253 $\lambda =$ 1.514296	1.016253 0.01429607	10.68158 0.04879727
50	$\alpha =$ 3.388975 $\lambda =$ 1.508105	0.3889748 0.008105423	1.99424 0.02228412
100	$\alpha =$ 3.177544 $\lambda =$ 1.504892	0.1775441 0.004891745	0.7521783 0.01055644
500	$\alpha =$ 3.019977 $\lambda =$ 1.498705	0.01997658 -0.00129502	0.09545959 0.00199238
1000	$\alpha =$ 3.007664 $\lambda =$ 1.500396	0.007664266 0.0003962086	0.04595611 0.001004574

14. Data Analysis

In this section, we considered a real data set of survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli, observed and reported by Bjerkedal (1960). This data set is used to explain the supremacy of the EW(α, λ) distribution. The data are presented in Table 5:

Table 5: *Data set of survival times of 72 guinea pigs infected with virulent tubercle bacilli*

12 48 60 75 109 258 15 52 60 76 110 258 22 53 61 76 121 263 24 54 62 81 127 297 24 54 63 83 129 341 32 55 65 84 131 341 32 56 65 85 143 376 33 57 67 87 146 34 58 68 91 146 38 58 70 95 175 38 59 70 96 175 43 60 72 98 211 44 60 73 99 233

We fitted EW(α,λ) distribution to the given data set and compared the results with Exponential distribution and Weibull distribution. The comparison is carried out based on the values of Kolmogorov-Smirnov (K-S) statistic, p-value, log-likelihood value, Akaike information criterion (AIC) and Bayesian information criterion (BIC).

The Kolmogorov-Smirnov test is defined by:

$$D = \max_{1 \le i \le N}\left(F(Y_i) - \frac{i-1}{N}, \frac{i}{N} - F(Y_i) \right)$$

where F is the cdf of the distribution and N is the number of classes being ordered.

The AIC and BIC are defined by

$$AIC = 2k - 2\ln\left(\hat{L}\right)$$

and

$$BIC = \ln(n)k - 2\ln\left(\hat{L}\right)$$

where $\hat{L}$ denote the maximum value of the likelihood function and k is the number of parameters and n is the sample size. By comparing the values of K-S statistic, p-values, likelihood value, AIC and BIC, Table 6 shows that EW(α, λ) distribution has smallest K-S statistic value and largest p-value among them. The AIC and BIC values of EW(α, λ) distribution indicates that the amount of information lost by the model is less than that of Exponential and Weibull distributions. This points out that the proposed model provided a better fit as well as more precise estimates.

Table 6: *Parameter Estimates, log-likelihood, p-value, AIC and BIC values of model fitted*

MODEL	PARAMETER ESTIMATES	LOG LIKELIHOOD	K-S STATISTIC	P-VALUE	AIC	BIC
MEW	a=1.396574502 b=0.009055284	-397.1651	0.1459	0.09327	798.3302	802.8836
WEIBULL	a=1.393170797 b=0.009045585	-421.9635	0.14645	0.09114	847.927	852.4804
EXPONENTIAL	b= 0.01001842	-422.1097	0.2116	0.003168	846.2193	848.496

15. Conclusions

In this paper, the mixture of Exponential and Weibull distributions are considered. The failure rate or hazard rate function is given. Moments, generating functions, Conditional moments, Quantile function, Mean deviation, distributions of order statistics, and Bonferroni and Lorenz Curves are obtained. Reliability in stress-streangth model is computed. Method of estimation of parameters using maximum likelihood estimation method is described and a simulation study is conducted to validate the maximum likelihood estimation procedure. A real data analysis is given, which shows the mixture distribution, EW(α, λ) distribution is a better alternative in certain situations of reliability and survival analysis.

References

[1] Nadarajah, S., Bakouch, H.S., and R. Tahmasbi (2011). *A generalized Lindley distribution.* Technical Report, School of Mathematics, University of Manchester, UK

[2] Nadarajah, S., and A.K. Gupta (2007). The exponentiated gamma distribution with application to drought data. *Calcutta Statistical Association Bulletin* 59, 233-234.

[3] Abdelfattah Mustafa, Beih S. El-Desouky and Shamsan AL-Garash (2016). *Weibull Generalized Exponential Distribution*. Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.

[4] Bjerkedal T (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Am J Hyg* 72, 130-148.

[5] Gupta, R. D. and Kundu, D. (2001). Exponentiated Exponential family; an alternative to gamma and Weibull. *Biometrical Journal*, 43, 117-130

[6] Lawless, J. F. (2003). Statistical Models and Methods for Lifetime Data. *John Wiley and Sons, New York*.

[7] Chacko V.M., Deepthi K S, Beenu Thomas, Rajitha C (2018) Weibull-Lindly Distribution: A bathtub shaped failure rate model, *Reliability: Theory and Applications*, Vol.13, pp.9-20, No.4(51).

Building the Continuous Random Process Out of The Specified Sequence of Turning Points for Fatigue Testing

Irina Gadolina[1*], Rashid Zaynetdinov[2]

•

[1] IMASH RAS
gadolina@mail.ru
[2] Independent researcher

Abstract

The method of regeneration of continuous process is intended for the first place for the calculation of spectral density of regenerated process. The main feature of this method is preserving the values and sequence of turning points (extremes) given in "saw-kind" realization. While doing so, the methods which is based on cycle-counting methods will give the exactly the same fatigue durability estimation, because the initial condition MAX-MIN-MAX … is guaranteed. To investigate the random process standard deviation (RMS) by the spectral density, the extrapolation of the original sequence is provided by continuous cosine functions. Continuality of the process and its first derivative is ensured by the condition of compatibility at the turning points. To decide on frequencies, the information from some sample realizations obtained in exploitation was employed. As one of te applications, the method is intended to be used in the analysis of comparability of two competing approaches in the evaluation of loading in the tasks of evaluation of durability, namely, those that apply cycle-counting methods and which are based on the spectral density of processes methods. Some other speculations made on the modelled process are performed.

Keywords: material fatigue, durability estimation, cosine extrapolation, cycle counting, spectral densities

I. Introduction

The problem of transforming the digital random sequences is actual in many scientific pieces of research [1]. It is especially important when comparing two main approaches in dealing with random loading in fatigue – that is the spectral approach (frequency domain) and cycle-counting approach (time-domain). A large number of standard random sequences of random process turning points are known, which can be used in testing material samples in the of high cycle fatigue. Some of them (SAE [2], TWIST [3], FALSTAFF [4] etc.) are designed for random load testing of parts for some specific industries. A method for modeling three alternative types of random sequences of load peaks for fundamental research was also developed [5]. The problem of generating the random processes with special requirements (kurtosis, for example) is discussed in [6]. In [6], a time-domain procedure for a non-Gaussian random test is introduced. [7] addresses the question of linking fatigue damage with the prescribed input kurtosis. Direct applications of these results include improved fatigue life estimations.

While the estimation of loading in fatigue, most often a special transformation is being made. Usually the transition from continuous process to "the saw". The example of "the saw" is shown in Fig.1. That particular process (FALSTAFF, [4]) is being imitated in a special way for the testing of aircraft fighters. The data for imitation had been collected during the exploitation in varied conditions. The transformation of the initial continuous process, which express the physical nature of the loading process into the turning points sequence is recommended by most of the standards for fatigue testing under random processes. Most of the methods for cycle counting deal only with "the saw". The main reason for that is the non-dependence of fatigue property on frequency and the cycle shape. This initial processing stage provides the transformation, after which only the values of peaks and trough and their original order sequence still remains. Usually, the fatigue estimation is fulfilled with taking in mind some real exploitation unit, for example, 1 km run (for transport machines), one hour of work (for technological machines) or the number of flights for the airplane. In this situation the time and frequencies does not play an important role.

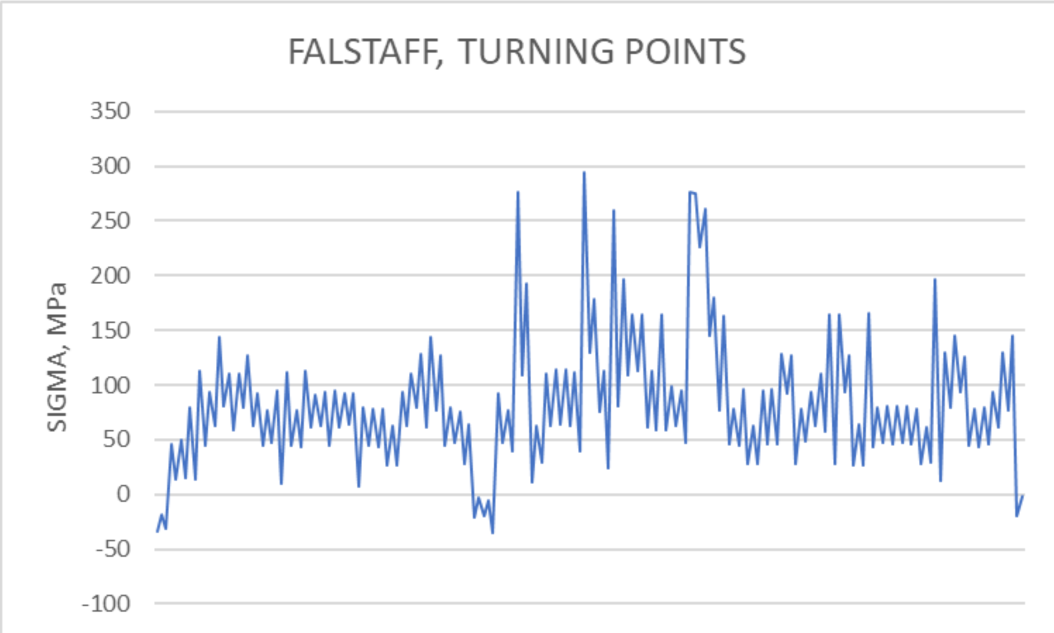


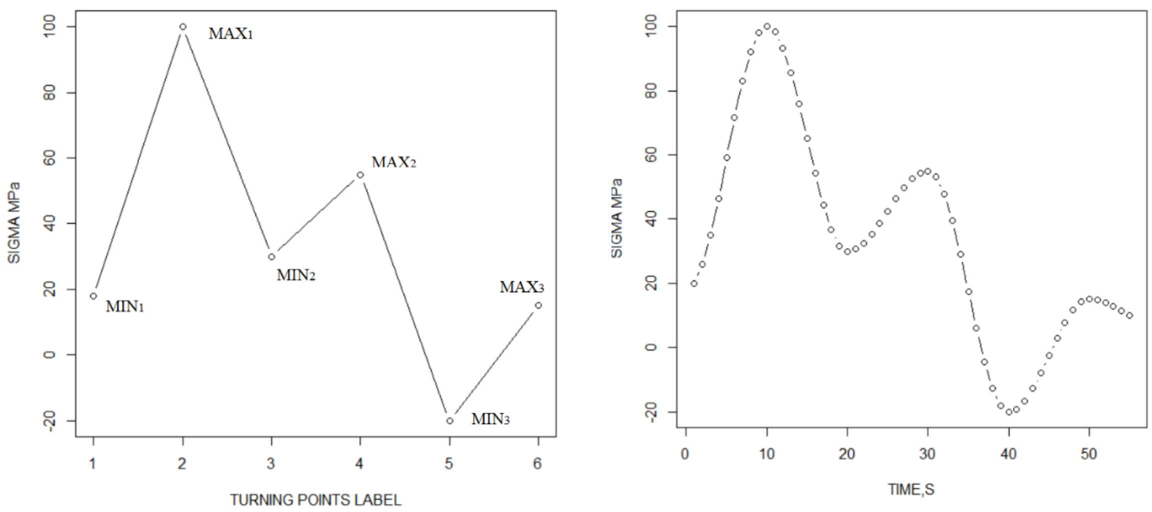
Figure 1: *Digital sequence MIN-MAX-MIN-MAX ("the saw") of representation of the part of random loading process FALSTAFF in the working parts of a fighter airplane [4 Falstaff]*

One the other hand, many applications require the necessity of the consideration of continuous process. Widespread nowadays-spectral methods for fatigue estimation [8] operate with the spectral density of the random process. Some other investigations, for example, the studying of the necessary discretization frequency, also require availability of continuous process. In problems of analyzing the strength of equipment, where there are miniature parts that do not allow registering stresses in operation using load cells or alternative measurement methods [9], it is necessary to resort to analyzing the stress distributions in the parts during vibrations by analyzing spectral densities. It is interesting to compare the results obtained using the rain method schematization with the results obtained based on the analysis of spectral densities.

While looking at the process in the Fig.1 one could easily see, that the process has the peculiarities at the peaks points and is far from being continuously smooth. It's first derivate is also non-continuous. Existing methods for functions extrapolation do not work here, because the main point here is to preserve the turning points as concerning their value as their sequence. The values of the turning points (extemes) of the process play the critical role while estimating machine parts durability under fatigue, and their sequence might as well be very important at the stage of estimation of crack propagation due to fatigue crack arrest effect after the overloading [10,11].

II. Method

To create the smooth extrapolating function with continuous derivate the cosine extrapolation has been used. Let's consider as an example a short part of "the saw" sequence, Fig.2a). Here only a few turning points are shown for the sake of simplicity.



a) *b)*

Figure 2: *Cosine extrapolation of the short peaks sequence. 1a) initial data; 2b) extrapolation*

Similarly to the shown in Fig.1, the data in Fig.2,a) is not, in reality, the process, but rather a kind of scheme, indicating the sequence of turning points MIN_1-MAX_1-MIN_2-MAX_2-MIN_3. The necessary requirement here if fulfilling the condition:

$$MIN\ i > MAX\ i > MIN\ i{+}1 \quad \&\ MAX\ i{-}1 < MIN\ i < MAX\ i \qquad \text{for } i\text{-}1,2 \ldots N \qquad (1),$$

where N is the total number of maximums (or minimums) in the sequence. It's worth mentioning, that to begin the sequence of extremes with the minimum, is a common practice among the researchers.

According to the idea [12], we consider each range in Fig.2a as a half-wave, described as a cosine function:

$$x(t) = A \cos (wt+\varphi) \qquad (2),$$

where x (t)– is the resulting continuous extrapolating function, which is described on the domain $t=0\ldots \pi/w$, because the period of cosine function is $T=2\pi/w$, s.

For each half-wave starting from the next extremum MAX or MIN, the parameters A, w, and φ are unique. The stress amplitude A, MPa is defined as the half of modulus of successive extremes:

$$A=mod(MAXi - MIN\ i{+}1)/2 \ \vee\ A=mod(MAXi - MIN\ i)/2 \qquad (3).$$

The half-waves are appended one by one. Due to the appropriate choice of extrapolating function, the conditions of coherence and continuity in the appending points at the turning points are fulfilled:

$$x^{-}(t_0) = x^{+}(t_0) \qquad (4)$$

$$\dot{x}^{-}(t_0) = \dot{x}^{+}(t_0) \qquad (5).$$

In expressions (3) and (4), the indices – and + mean that the appending points belong to different half-waves, to the left and right of the turning points MAX i and MIN i. Condition (4) is provided by fulfilling condition (3). The condition (5) for equality of derivatives is fulfilled due to modeling principle, according to which the cosine argument is equal to $\pi/2$ at the points of extremes corresponding to t = 0, hence the derivative at all points of alignment is zero.

The resulting continuous process is shown in Fig.2b). Turning points MIN - MAX- MIN - MAX -MIN are the same in both processes. In this example for each half-wave 10 readings were taken.

Choice of the half-cosine frequency w

Due to the fact, that the frequency component is not defined in known standards random sequences [2-5], there is a need to define it. An analysis was performed earlier to investigate the connection between the range and the period of the oscillation [13]. Intuitively one can imagine, that the greater the oscillation, the more time will be required to it performance. That fact is well-known to the personal, who perform the testing under random loading. For the real oscillation of the torsion stresses in a shaft of caterpillar machine the regression equation has shown the existence of linear dependence.

In Fig.3 the dependence of the range *Raz* of oscillation on the half- wave period *t* of the FALSTAFF oscillation is shown. Although the scatter is significant enough (correlation factor is only K=0.6664 < 1), one could still notice a tendency of increasing *Raz* with *t* increasing. That might be used later on while modelling the random process by its turning points.

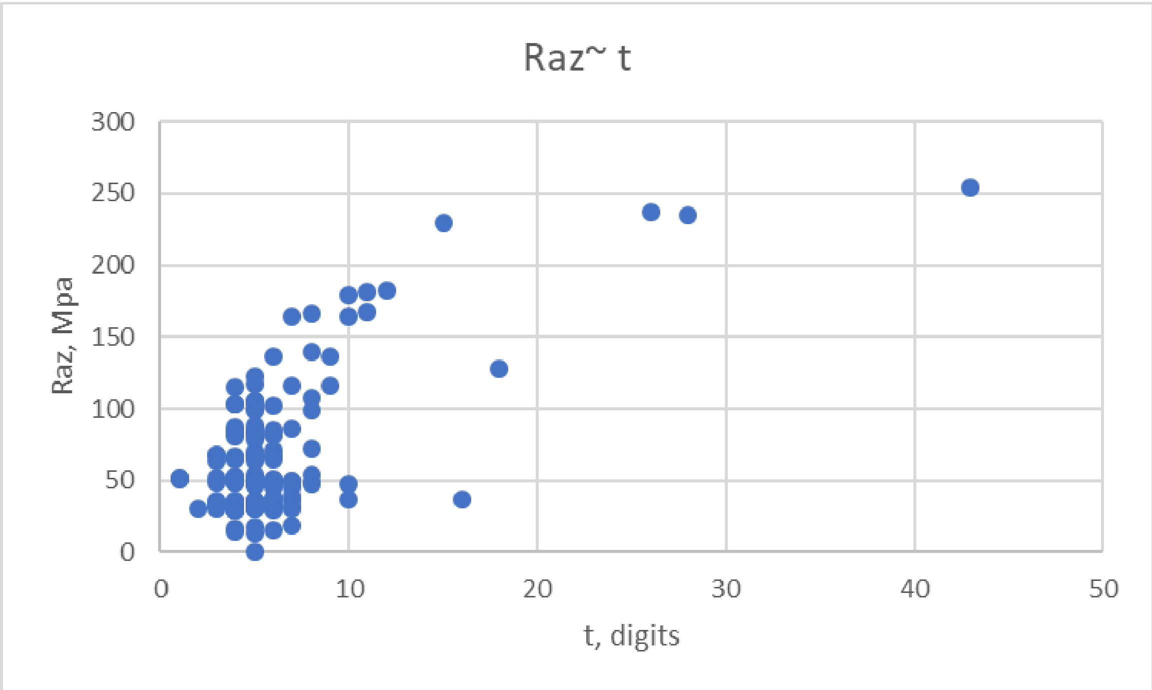


Figure 3: *Correlation of the period of oscillation t with its range Raz (FALSTAFF)*

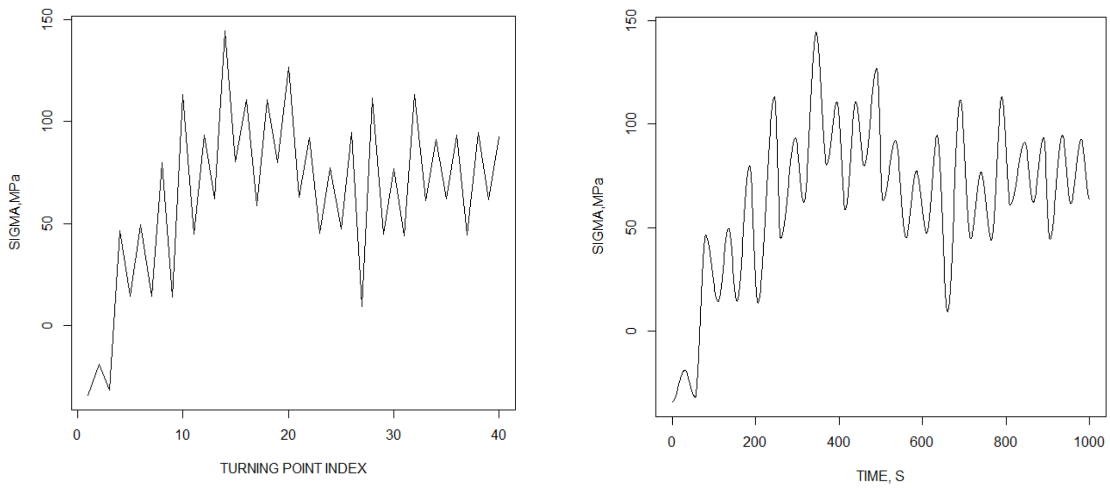
The proposed method of process imitation has been applied to the FALSTAFF sequence shown in Fig.1. [4]. Because one of the aims of this imitation is spectral density estimation for the spectral methods application [8], the number of reading in continuous process was chosen among the number:

$$2^n, \text{ where } n=9 \dots 12 \text{ (digit)} \tag{6}$$

to have the possibility to apply EXCELL for spectral density estimation.

III. Results

In Fig. 4 the part of the modelled by proposed method random process, corresponding to FALSTAFF sequence (Fig.1) is shown. Because the presence of random number generator in this algorithm this trial is not unique. Some duplications might exist. Because the aim of that modelling is the fatigue testing, the main factors for fatigue estimation were analyzed. They are: 1) irregularity factor *I;* 2) spectra fullness factor *V*. *V* depends on *m* – fatigue curve exponent and is always less than unity. Because of this particular modelling principle, *I* & *V* are the same for all duplications. That means, that when employing the rain-flow method for loading estimation, duplications will give the same result. The discrepancies will appear only while applying the spectral method [8] because the spectral densities function will differ due to random choice of *w* in (2).



a) *b)*

Figure 4: *FALSTAFF (the beginning part) a) digit sequence of the turning points; b) continuous process, modelled by proposed method*

Due to the fact, that for normal stationary processes a lot of tools were developed, the continuous modeled process (Fig.4, a) was investigated for its normality. To test the hypothesis that the generated process is normal, the Q-Q plot test was applied [14]. To check that the sample of random ordinates comes from a population that is Normally distributed? The Q-Q plot, shown in Fig. 5, indicates that it appears to be a fairly safe assumption. The points seem to fall about a straight line. Sample quantiles are on y-axis. The x-axis plots are the theoretical quantiles. Those are the quantiles from the standard Normal distribution with mean 0 and standard deviation 1.

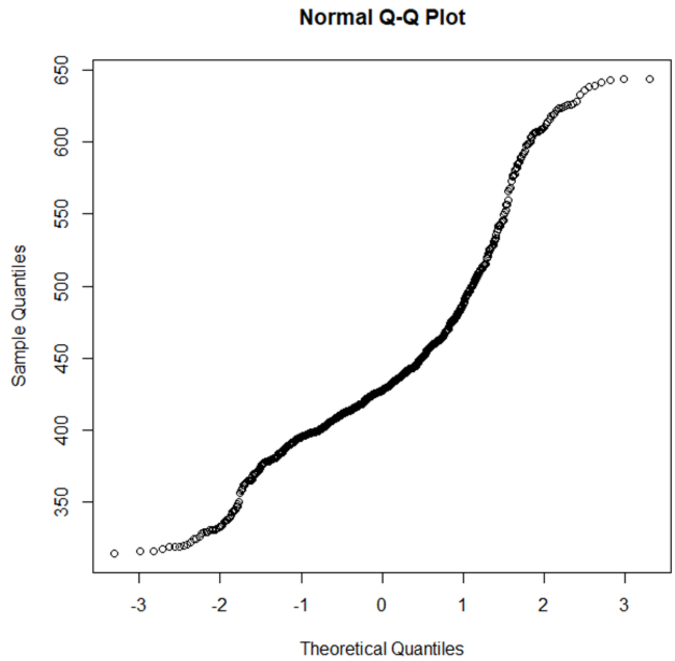


Figure 5: *Q-Q plot for random coordinates of the modeled process.*

In Fig. 6 the histogram of the selected by rain-flow cycle counting method [15] amplitudes for FALSTAFF sequence is shown. The fullness ratio *V,* which is similar to equivalent constant amplitude under the assumption of linear summation fatigue damages depends on the fatigue exponent *m.* This graph represents also the distribution of rain-flow amplitudes – the last one is exactly the same for continuous process! For the FALSTAFF example values *V* are shown in the Table 1 for some typical for construction materials *m*.

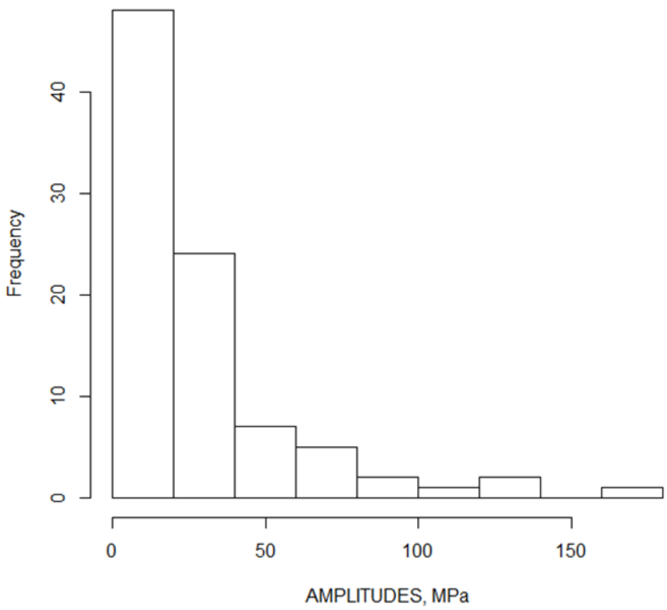


Figure 6: *Rain-flow amplitudes histogram for the selected realization*

Table 1: *Fulness ration V for FALSTAFF for some fatigue exponents m*

Fatigue exponent *m*	6	9	12
Fulness ratio *V*	0.512	0.622	0.6937

To investigate the loading characteristics of random processes sometimes the spectral densities apparatus is employed [8, 16]. For the normal stationary processes, the Rice theory [16] might be applied. It operates with the spectral densities. In particular, the irregularity factor I might be estimated through some moment, estimated on spectral density:

$$I = \left(\int_0^\infty S(w)w^2dw\right) / \sqrt{\int_0^\infty S(w)w^4dw \int_0^\infty S(w)w\,dw} \tag{7}$$

Here *S(w)* is the dependence of the spectral density on frequency *w*.

In Fig. 7 the estimated by EXCELL *S(w)* for continuous process is shown. The integrals (moments) in (7) were calculated numerically.

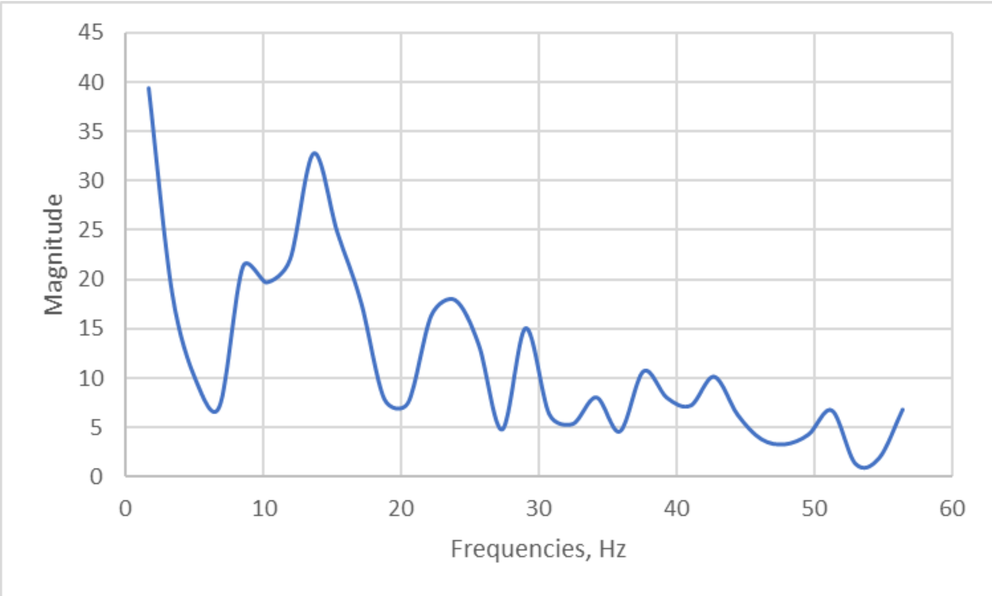


Figure 7: *Spectral density of imitated on FALSTAFF sequence process*

IV. Discussion and conclusions

The main result, which is important for fatigue estimation, is the coincidence of *V(m)* for both representations – the turning point scheme and for the continuous modeled process (Table 1 – true for both).

The proposed method was applied for the standard sequence, recommended to fighter airplane parts testing. Before now there was no tool to generate continuous process out of sequence of digits, preserving the values and sequencing the turning points, which are critical for fatigue estimation.

The irregularity factor [15]

$$I=No/Ne \tag{8}$$

is estimated numerically by a simple manner by counting the number of crossing mean load level *No* and the number of extremums *Ne* slightly differs for two representations due to varied process mean estimation (Table 2).

On the other hand, the irregularity factor I =0.28, estimated for imitated by the proposed model process, according the Rice theory (7) differs significantly from I =0.52… 0.55, estimated numerically by the realization by formula (8). That might lead to discrepancies in fatigue estimation using the spectral approach.

Table 2: *Irregularity factor I*

Estimated numerically by turning points sequence by (8)	0.52
Estimated numerically by continuous modelled process by (8)	0.55
Estimated theoretically by spectral densities by (7)	0.28

All calculations and process imitation were performed in R [14].

References

[1] Michał Böhm. Fatigue life assessment with the use of spectral method for materials subjected to standardized wind loading spectrums. October 2018 AIP Conference Proceedings 2029(1):020005. DOI: 10.1063/1.5066467. Conference: MECHATRONICS SYSTEMS AND MATERIALS 2018At: Zakopane, Poland. Volume: AIP Conference Proceedings 2029, 020005 (2018)

[2] Tucker L., Bussa S. The SAE Cumulative Fatigue Damage Test Program / in book Fatigue under Complex Loading. Ed. R.M. Wetzel. 1977. Society of Automotive Engineering. P.1-44.

[3] D. Schütz, H. Lowak, de Jonge, J.Schijve. LBF – Report FB-106, NLR – Report TR 73, 1973. (TWIST).

[4] J. B. de Jonge et al. FALSTAFF. DESCRIPTION OF A FIGHTER AIRCRAFT LOADING STANDARD FOR FATIGUE EVALUATION (FIugzeugwerke, Emmen (Switzerland). 1976. 71 p.

[5] R. Fisher, E. Haibach. Modeling functions loading in experiments on the evaluation of materials / In the book "Behavior of steel under cyclic loads". Edited by V. Dahl. 1983. P. 368-405.

[6] Song Cui, Enlai Zheng, Min Kang. A time-domain procedure for non-Gaussian stationary environmental testing using zero-memory nonlinear transformation// Journal of Vibration and Control · January 2020. DOI: 10.1177/1077546319894853

[7] Frédéric Kihm, Neil S. Ferguson, Jérôme Antoni. Fatigue life from kurtosis-controlled excitations//Procedia Engineering 133 (2015) 6th Fatigue Design conference, Fatigue Design 2015. P. 698 – 713.

[8] D.Benasciutti, R.Tovo. Frequency-based analysis of random fatigue loads: Models, hypotheses, reality// Mat.-Werkstofftech. 2018, 49, 345-361.

[9] Youssef M., Konstantin G. et al. Solder Joint Lifetime Modeling Under Random Vibrational Load Collectives// JOM: the journal of the Minerals, Metals & Materials Society. December 2019 DOI: 10.1007/s11837-019-03947-1

[10] S.G. Lebedinskii. Design Modeling of Propagation of the Fatigue Cracks in the Steel of Molded Parts of the Railway Structures // Journal of Machinery Manufacture and Reliability, 2018, Vol. 47, No. 1, P. 62–66. ISSN 1052-6188, © Allerton Press

[11] Irina V. Gadolina, Andrey A. Bautin, Evgenii V. Plotnikov. Studying the Crack Growth Rate Variability by Applying the Willenborg's Model to the Markov's Simulated Trials// In book: Advances in Intelligent Systems, Computer Science and Digital Economics. January 2020. DOI: 10.1007/978-3-030-39216-1_17

[12] Gryzlova T. P. Signal half-wave analysis (HW analysis) for recognizing the state of complex sources. // Proc of All-Russian conference "Modern signal processing technologies STOs-2018". - Moscow, RNTORES named after A. S. Popov, 2018. P. 240 – 244.

[13] Gryzlova T. P., Erpalov A.V., Gadolina I. V., Plotnikov E. V., Martynenko A. A. An analysis of the possible approaches for the consideration of loading the case of a narrow-band random loading processes. Half-wave method. //Proc. of the VI International conference " Fundamental research and innovative technologies in mechanical engineering", Moscow, IMASH RAS, November 26-27, 2019, p. 133-135.

[14] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2017. URL https://www.R-project.org/

[15] Interstate Standard GOST 25.101-83 Strength calculation and testing. Representation of random loading of machine elements and structures and statistical evaluation of results. - M: Standartinform, 2005. – 25p. [in Russian]

[16] Erpalov, A.V., Shefer, L.A. Fatigue-based Classification of Loading Processes Procedia Engineering 2nd International Conference on Industrial Engineering (ICIE-2016) Volume 150, (2016) 2016.–Vol. 150. pp.144-149.

Transient Solution of a Heterogeneous Queuing System with Balking and Catastrophes

Rakesh Kumar*[1], Sapana Sharma[1], Bhavneet Singh Soodan[1],
P. Vijaya Laxmi[2], Bhupender Kumar Som[3]

•

[1]School of Mathematics, Shri Mata Vaishno Devi University, Katra, J&K, India-182320
[2]Department of Applied Mathematics, Andhra University, Visakhapatnam, India-530003
[3]Lloyd Business School, Greater Noida, India-110019

Email: rakesh.kumar@smvdu.ac.in
*Corresponding author

Abstract

In this paper we consider a Markovian queuing system with heterogeneous servers, balking and catastrophes. The time-dependent behavior of the system is analyzed by using generating function technique. The expressions for mean and variance of the system are obtained in transient state. At last, some special cases of the model are derived and discussed.

Keywords: Transient solution, Catastrophes, Generating function, Heterogeneous servers, Balking

I. Introduction

Queuing models are playing an important role in modeling the queuing situations in computer-communication networks, hospitals, supply chain management, and in production processes. Customers' impatience is one of the most important aspects in modeling of queuing systems. Queuing systems with customers' impatience are comparatively less profitable than the ones without impatience. In real life many queuing situations arise in which there may be a tendency for customers to be discouraged by a long queue. As a result, the customers either decide not to join the queue (i.e. balk) or depart after joining the queue without getting service due to impatience (i.e. renege). The study of customers' impatience in queuing theory is started in the early 1950's. Haight (1957), Ancker and Gafarian ((1963a), (1963b)) are the pioneer researchers in the area of queueing with customers' impatience. Barrer (1957) analyzes an M/M/c queue with customers' impatience of constant duration. El-Paoumy and Nabwey (2011) study a Poisson queue with balking function, reneging and two heterogeneous servers. Kumar and Sharma (2012a) study a single server Markovian queuing system with balking and retention of reneging customers. They obtain the steady-state probabilities of the model. Kumar and Sharma (2012b) obtain the stationary system size probabilities of a finite capacity Markovian multi-server queuing system with balking and retention of reneging customers. Kumar and Sharma (2018) analyze the transient state probabilities of a multi-server queuing system with balking and retention of reneging customers.

The queuing systems with heterogeneous servers are more applicable as compare to their homogeneous counterparts, because in real-life situations the servers work at different rates. Morse (1958) introduces the concept of heterogeneous service. The heterogeneous service mechanisms are scheduling methods that allow customers to receive different quality of service. Most of the operations in manufacturing systems have heterogeneous service mechanism. That is why, the queuing systems with heterogeneous servers have gained significant attention in the literature. Saaty (1961) further discusses Morse's problem and derives the steady-state probabilities and the mean number in the system. Sharma and Dass (1989) analyze the initial busy period of multichannel Markovian queueing system and obtain the expression of its density function in closed form. Dharmaraja (2000) obtains the transient solution of a two-processor heterogeneous system with Poisson arrival of jobs having exponentially distributed processing times. Kumar and Sharma (2019) obtain the transient solution of a two-heterogeneous servers Markovian queuing model with retention of reneging customers.

Queuing models with catastrophes have been used in modeling a variety of real life systems, such as computer-communication networks under virus attack, manufacturing systems with sudden disasters, and call centers with sudden power breakdowns and corruption of hard disk of computer systems. Recently, due attention has been paid to the study of queuing systems with catastrophes. The occurrence of catastrophes leads to the annihilation of all the customers in the queuing system and momentarily inactivates the service facility until a new arrival occurs. In order to study the impact of noise bursts and virus on queues in computer networks Chao (1995) develops the queuing network model with catastrophes. He obtains the product-form solution of a queuing network model with catastrophes. Kumar and Arivudainambi (2000) incorporate the effect of catastrophes in a single server Markovian queuing system. They derive its transient solution using generating function technique explicitly. Di Crescenzo et al. (2003) discuss the application of M/M/1 queuing model with catastrophes in the phenomenon of muscle contraction. Jain and Kumar (2007) derive the transient solution of a queuing system with correlated arrivals, variable service capacity and catastrophes. Sudesh (2010) studies a single server queuing system with catastrophes and customers' impatience. He derives the transient solution of the model explicitly using generating function technique. Sudesh et al. (2016) derive the transient solution of a two-heterogeneous servers queuing system with catastrophes, server repair and customers' impatience. Jain and Kanethia (2006) study a single server queuing model with change in environment and catastrophes. They obtain both the transient and the steady-state solutions to the model. Tarabia (2011) performs the transient and steady-state analysis of a single server Markovian queuing system with balking, catastrophes, server failures and repairs. Yechiali (2007) studies single and multiple-server queuing models with catastrophes and impatient customers. Ammar (2014) derives the transient solution of a two-processor heterogeneous system with catastrophes, server failures and repairs. Dharmaraja and Kumar (2015) obtain the transient solution of a queuing model with multiple heterogeneous servers in presence of catastrophes. Kumar et al. (2001) obtain the transient solution of an M/M/2 heterogeneous servers queuing system in presence of catastrophes.

Yaseen and Tarabia (2017) analyze the transient and steady-state behavior of Markovian queuing system with balking and reneging subject to catastrophes and server failures. Suranga Sampath and Liu (2018) study an M/M/1 queuing system with reneging, catastrophes, server failures and repairs. They obtained the transient as well as steady-state solution of the model. The applicability of our queuing model can be seen in hospital emergency departments and computer communication system.
The remainder of the paper is structured as follows. In section 2, the queuing model is described. A mathematical model is formulated in section 3. In section 4 transient solution of the model is

studied. The time-dependent mean and variance of the model are obtained in section 5. Section 6 deals with the special cases of the model. Finally, the paper is concluded in section 7.

.

II. Queuing Model Description

In this section, we describe the queueing model. The model is based on following assumptions:

1. In accordance with a Poisson process, the arrivals occur one by one with intensity λ.
2. The system has multi-servers (say, c) having distinct service rates and the service times at each server are exponential distributed. This means that the customers are always served by the fastest servers. That is, when such a server becomes available, a customer may switch to a fastest server.
3. On arrival customer either decides to enter the queue with probability p or balk with probability $1-p$.
4. Apart from this, the catastrophes may also occur at the service facility as a Poisson process with rate ψ, when the system is not empty. At the moment when catastrophe occurs at the system, all the customers are destroyed, all the servers get inactivated momentarily and after the catastrophe, the servers become ready for service immediately.
5. The queue discipline is FCFS and the capacity of the system is infinite.
6. Initial condition: $P_0(0) = 1$.

III. Mathematical Formulation of the Model

Define, $\mathrm{P_n(t) = P\{X(t) = n\}, n = 0, 1, ...}$ The queuing model under investigation is governed by the following differential-difference equations:

$$\frac{dP_0(t)}{dt} = -(\lambda+\psi)P_0(t) + \mu_1 P_1(t) + \psi \tag{1}$$

$$\frac{dP_n(t)}{dt} = -(\lambda+\psi+\textstyle\sum_{i=1}^{n}\mu_i)P_n(t) + \sum_{i=1}^{n+1}\mu_i\, P_{n+1}(t) + \lambda P_{n-1}(t), \quad 1 \le n < c \tag{2}$$

$$\frac{dP_c(t)}{dt} = -(\lambda p+\psi+\textstyle\sum_{i=1}^{c}\mu_i)P_c(t) + \sum_{i=1}^{c}\mu_i\, P_{c+1}(t) + \lambda P_{c-1}(t), \; n = c \tag{3}$$

$$\frac{dP_n(t)}{dt} = -(\lambda p+\psi+\textstyle\sum_{i=1}^{c}\mu_i)P_n(t) + \sum_{i=1}^{c}\mu_i\, P_{n+1}(t) + \lambda p P_{n-1}(t), \; n > c \tag{4}$$

IV. Transient solution of the model

Theorem 1. The transient state probabilities of a Markovian queuing system with multi heterogeneous servers, balking and catastrophes which is governed by the differential-difference equations (1) − (4) are given by:

$$P_k(t) = b_{k,0}(t) + \psi\int_0^t b_{k,0}(u)du + \gamma\int_0^t b_{k,c-1}(u)P_c(t-u)du, \quad k = 0, 1, \dots, c-1$$

$$P_c(t) = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{(-1)^m}{\gamma}\left(\frac{\alpha}{2\lambda p}\right)^{n+1}(n+1)\binom{n}{m}\Big[\int_0^t A(t-u)\int_0^u B^{C(m)}(u-v)\exp\{-(\lambda p+\gamma+\psi)v\}\frac{I_{n+1}(\alpha v)}{v}dudv + \psi\int_0^t H(t-u)\int_0^u B^{C(m)}(u-v)\exp\{-(\lambda p+\gamma+\psi)v\}\frac{I_{n+1}(\alpha v)}{v}dudv\Big]$$

and, for n = 1, 2, ...

$$P_{n+c}(t) = n\beta^n\int_0^t \exp\{-(\lambda p+\psi+\gamma)(t-u)\}\frac{I_n(\alpha(t-u))}{(t-u)}P_c(u)du$$

where $H(t) = \int_0^t A(u)du$ and $B^{C(m)(t)}$ is m – fold convolution of B(t) with itself with $B^{C(0)} = \delta(t)$, the Dirac - delta function.

Proof. Define the pgf P(z, t) for the transient state probabilities $P_n(t)$ by

$$P(z,t) = q_c(t) + \sum_{n=1}^{\infty} P_{n+c}(t) z^{n+1}; \qquad P(z,0) = 1 \tag{5}$$

with

$$\sum_{n=0}^{c} P_n(t) = q_c(t) \tag{6}$$

Adding the equations (1) - (3), we get

$$\frac{d}{dt}\big(q_c(t)\big) = -\lambda p P_c(t) + \sum_{i=1}^{c}\mu_i\, P_{c+1}(t) - \psi q_c(t) + \psi \tag{7}$$

On multiplying equation (4) by z^n and summing, we get

$$\frac{d}{dt}\left[\sum_{n=1}^{\infty} P_{n+c}(t)z^{n+1}\right] = \left[(\lambda p+\psi+\sum_{i=1}^{c}\mu_i)+\left(\lambda pz+\frac{\sum_{i=1}^{c}\mu_i}{z}\right)\right]\sum_{n=1}^{\infty} P_{n+c}(t)z^n+\lambda pzP_c(t)-\sum_{i=1}^{c}\mu_i P_{c+1}(t) \tag{8}$$

By adding (7) and (8) the following differential equation is obtained:

$$\frac{\partial P(z,t)}{\partial t} = \left[\left(\lambda pz+\frac{\sum_{i=1}^{c}\mu_i}{z}\right)-(\lambda p+\psi+\sum_{i=1}^{c}\mu_i)\right]P(z,t)-\left[\left(\lambda pz+\frac{\sum_{i=1}^{c}\mu_i}{z}\right)-(\lambda p+\sum_{i=1}^{c}\mu_i)\right]q_c(t)+\lambda p(z-1)P_c(t)+\psi \tag{9}$$

On solving (9), we get

$$P(z,t) = \exp\left\{\left[\left(\lambda pz+\frac{\sum_{i=1}^{c}\mu_i}{z}\right)-(\lambda p+\psi+\sum_{i=1}^{c}\mu_i)\right]t\right\}+\int_0^t[\lambda p(z-1)P_c(u)-\left(\left(\lambda pz+\frac{\sum_{i=1}^{c}\mu_i}{z}\right)-(\lambda p+\sum_{i=1}^{c}\mu_i)\right)q_c(u)]\times\exp\left\{\left[\left(\lambda pz+\frac{\sum_{i=1}^{c}\mu_i}{z}\right)-(\lambda p+\psi+\sum_{i=1}^{c}\mu_i)\right](t-u)\right\}du+\psi\int_0^t\exp\left\{\left[\left(\lambda pz+\frac{\sum_{i=1}^{c}\mu_i}{z}\right)-(\lambda p+\psi+\sum_{i=1}^{c}\mu_i)\right](t-u)\right\}du \tag{10}$$

If $\gamma=\sum_{i=1}^{c}\mu_i$, $\alpha=2\sqrt{\lambda p\gamma}$ and $\beta=\sqrt{\frac{\lambda p}{\gamma}}$, then using the modified Bessel function of first kind $I_n(.)$ and the Bessel function properties, we get

$$\exp\left\{\left(\lambda pz+\frac{\gamma}{z}\right)t\right\}=\sum_{n=-\infty}^{\infty}(\beta z)^n I_n(\alpha t) \tag{11}$$

Using (11) in (10), we get

$$P(z,t)=\exp\{-(\lambda p+\psi+\gamma)t\}\sum_{n=-\infty}^{\infty}(\beta z)^n I_n(\alpha t)+\lambda p\int_0^t P_c(u)\exp\{-(\lambda p+\psi+\gamma)(t-u)\}\sum_{n=-\infty}^{\infty}(\beta z)^n\left[\beta^{-1}I_{n-1}(\alpha(t-u))-I_n(\alpha(t-u))\right]du+\int_0^t q_c(u)\exp\{-(\lambda p+\psi+\gamma)(t-u)\}\sum_{n=-\infty}^{\infty}(\beta z)^n\left[-\lambda p\beta^{-1}I_{n-1}(\alpha(t-u))+(\lambda p+\gamma)I_n(\alpha(t-u))-\beta\gamma I_{n+1}(\alpha(t-u))\right]du+\psi\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}\sum_{n=-\infty}^{\infty}(\beta z)^n I_n(\alpha(t-u))du \tag{12}$$

Now, comparing the coefficients of z^n on either side of (12), we obtain for $n=1,2,\ldots$

$$P_{n+c}(t)=\exp\{-(\lambda p+\psi+\gamma)t\}(\beta)^n I_n(\alpha t)+\lambda p\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}\left[I_{n-1}(\alpha(t-u))\beta^{n-1}-I_n(\alpha(t-u))\beta^n\right]P_c(u)du-\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}q_c(u)\left[\lambda pI_{n-1}(\alpha(t-u))\beta^{n-1}-(\lambda p+\gamma)I_n(\alpha(t-u))\beta^n+\gamma I_{n+1}(\alpha(t-u))\beta^{n+1}\right]du+\psi\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}\beta^n I_n(\alpha(t-u))du \tag{13}$$

Comparing the terms free of z on either side of equation (12), that is, for $n=0$, we get

$$q_c(t)=\exp\{-(\lambda p+\psi+\gamma)tI_0(\alpha t)\}+\lambda p\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}\left[I_1(\alpha(t-u))\beta^{-1}-I_0(\alpha(t-u))\right]P_c(u)du-\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}q_c(u)\left[\alpha I_1(\alpha(t-u))-(\lambda p+\gamma)I_0(\alpha(t-u))\right]du+\psi\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}I_0(\alpha(t-u))du \tag{14}$$

After simplifying (13), we obtain

$$\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}q_c(u)\left[\lambda pI_{n+1}(\alpha(t-u))\beta^{n-1}-(\lambda p+\gamma)I_n(\alpha(t-u))\beta^n+\gamma I_{n-1}(\alpha(t-u))\beta^{n+1}\right]du$$
$$=\exp\{-(\lambda p+\psi+\gamma)t\}(\beta)^n I_n(\alpha t)+\lambda p\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}\left[I_{n+1}(\alpha(t-u))\beta^{n-1}-I_n(\alpha(t-u))\beta^n\right]P_c(u)du+\psi\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}\beta^n I_n(\alpha(t-u))du \tag{15}$$

Substituting (15) in (13), we get

$$P_{n+c}(t)=n\beta^n\int_0^t\exp\{-(\lambda p+\psi+\gamma)(t-u)\}\frac{I_n(\alpha(t-u))}{(t-u)}P_c(u)du,\quad n=1,2,\ldots \tag{16}$$

On solving (1) and (2), we obtain the remaining probabilities $P_n(t), n=0,1,2,\ldots,c$. Equations (1) and (2) can be written in matrix form as:

$$\frac{d\boldsymbol{P}(t)}{dt}=A\boldsymbol{P}(\boldsymbol{t})+\gamma P_c(t)\boldsymbol{e_1}+\psi\boldsymbol{e_2} \tag{17}$$

where the matrix $A=(a_{i,j})_{c\times c}$ is given as:

$$A=\begin{bmatrix} -(\lambda+\psi) & \mu_1 & . & . & . & 0 \\ \lambda & -(\lambda+\psi+\mu_1) & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & \sum_{i=1}^{c-1}\mu_i \\ 0 & 0 & . & . & . & -(\lambda+\psi+\sum_{i=1}^{c-1}\mu_i) \end{bmatrix}$$

$\boldsymbol{P(t)}=(P_0(t)\ P_1(t)\ \dots\ P_{c-1}(t))^T$, $\boldsymbol{e_1}=(0\ 0\ \dots 1)^T$ and $\boldsymbol{e_2}=(1\ 0\ \dots 0)^T$ are vectors of order c. Let the Laplace Transform (LT) of $\mathbf{P(t)}$ is $\boldsymbol{P}^*(\boldsymbol{s})=(P_0{}^*(s)\ P_1{}^*(s)\ \dots\ P_{c-1}{}^*(s))^T$. Taking the Laplace Transform of equation (17) we get

$$\boldsymbol{P}^*(\boldsymbol{s})=(sI-A)^{-1}\left\{\gamma P_c^*(s)\boldsymbol{e_1}+\boldsymbol{P(0)}+\frac{\psi}{s}\boldsymbol{e_2}\right\} \tag{18}$$

with $\mathbf{P(0)}=(1\ 0\ \dots 0)^{\mathrm{T}}$. If $\mathrm{e}=(1\ 1\ \dots 1)^{\mathrm{T}}{}_{c\times 1}$, then

$$e^T\boldsymbol{P}^*(\boldsymbol{s})+P_c^*(s)=q_c^*(s) \tag{19}$$

Define

$$f(s)=[(s+\lambda p+\gamma+\psi)-\sqrt{(s+\lambda p+\gamma+\psi)^2-\alpha^2}]$$

Taking LT of (14), we obtain

$$s(s+\psi)q_c^*(s)=(s+\psi)+sP_c^*(s)\frac{1}{2}[f(s)-\alpha\beta] \tag{20}$$

Using (20) in (19), we get

$$P_c^*(s)=\left(\frac{s+\psi}{s}\right)\frac{1-se^T(sI-A)^{-1}\left(\boldsymbol{P(0)}+\frac{\psi}{s}\boldsymbol{e_2}\right)}{\left\{(s+\lambda p+\psi)-\frac{1}{2}\{f(s)\}+(s+\psi)\gamma\boldsymbol{e}^T(sI-A)^{-1}\boldsymbol{e_1}\right\}} \tag{21}$$

Let us assume that $(sI-A)^{-1}=\left(b_{ij}^*(s)\right)_{c\times c}$

We observe that $(sI-A)^{-1}$ is almost lower triangular. Following Raju and Bhat (1982), we obtain, $i=0,1,\dots,c-1$

$$b_{ij}^*(s)=\begin{cases}\frac{1}{\sum_{k=1}^{j+1}\mu_k}\frac{\mu_{c,j+1}(s)\mu_{i,0}(s)-\mu_{i,j+1}(s)\mu_{c,0}(s)}{\mu_{c,0}(s)}, & j=0,1,\dots,c-2\\ \frac{\mu_{i,0}(s)}{\mu_{c,0}(s)}, & j=c-1\end{cases} \tag{22}$$

where $\mu_{i,j}(s)$ are recursively given as

$$\mu_{i,i}(s)=1, \qquad i=0,1,\dots,c-1$$

$$\mu_{i+1,i}(s)=\frac{s+\lambda+\psi+\sum_{k=1}^{i}\mu_k}{\sum_{k=1}^{i+1}\mu_k} \qquad i=0,1,\dots,c-2$$

$$\mu_{i+1,i-j}(s)=\frac{\left(s+\lambda+\psi+\sum_{k=1}^{i}\mu_k\right)\mu_{i,i-j}-\lambda\mu_{i-1,i-j}}{\sum_{k=1}^{i+1}\mu_k}, \quad j\le i, \qquad i=1,2,\dots,c-2$$

$$\mu_{c,j}(s)=\begin{cases}[s+\lambda+\psi+\sum_{k=1}^{c-1}\mu_k]\mu_{c-1,j}-\lambda\mu_{c-2,j}, & j=0,1,\dots,c-2\\ s+\lambda+\psi+\sum_{k=1}^{c-1}\mu_k, & j=c-1\end{cases} \tag{23}$$

and $\mu_{i,j}(s)=0$, for other i and j. We have suppressed the argument s to facilitate computation. The advantage in using these relations is that we do not evaluate any determinant. Using these in equation (21), we get

$$P_c^*(s)=\left(\frac{s+\psi}{s}\right)\frac{1-(s+\psi)\sum_{i=0}^{c-1}b_{i,o}^*(s)}{\left\{(s+\lambda p+\psi)-\frac{1}{2}\{f(s)\}+(s+\psi)\gamma\sum_{j=0}^{c-1}b_{j,c-1}^*(s)\right\}} \tag{24}$$

and for $k=0,1,\dots,c-1$ from equation (18), we get

$$P_k^*(s)=\left(1+\frac{\psi}{s}\right)b_{k,0}^*(s)+\gamma b_{k,c-1}^*(s)P_c^*(s) \tag{25}$$

We observe that $b_{i,j}^*(s)$ are all rational algebraic functions in s. So, by partial fraction decomposition the inverse transform $b_{i,j}(t)$ of $b_{i,j}^*(s)$ can be obtained. Let $s_i, i=0,1,\dots,c-1$, be the characteristic roots of the matrix A. Then after simplification, $P_c^*(s)$ equals to

$$\frac{\left(1+\frac{\psi}{s}\right)A^*(s)}{\frac{1}{2}[(s+\lambda p+\gamma+\psi)+\sqrt{(s+\lambda p+\gamma+\psi)^2-\alpha^2}]\left[1-\frac{2\gamma(1-B^*(s))}{[(s+\lambda p+\gamma+\psi)+\sqrt{(s+\lambda p+\gamma+\psi)^2-\alpha^2}]}\right]} \quad (26)$$

where

$$A^*(s)=\sum_{i=0}^{c-1}\frac{A_i}{s-s_i} \quad (27)$$

$$B^*(s)=\sum_{i=0}^{c-1}\frac{B_i}{s-s_i} \quad (28)$$

with constants A_i and B_i given by

$$A_i=\lim_{s\to s_i}(s-s_i)\left[1-\sum_{l=0}^{c-1}(s+\psi)b_{l,0}^*(s)\right] \quad (29)$$

$$B_i=\lim_{s\to s_i}(s-s_i)\left[\sum_{l=0}^{c-1}(s+\psi)b_{l,c-1}^*(s)\right] \quad (30)$$

Hence, (26) simplifies into

$$P_c^*(s)=\sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{(-1)^m}{\gamma}\left(\frac{\alpha}{2\lambda p}\right)^{n+1}(n+1)\binom{n}{m}\left(1+\frac{\psi}{s}\right)A^*(s)\left(B^*(s)\right)^m\frac{[(s+\lambda p+\gamma+\psi)+\sqrt{(s+\lambda p+\gamma+\psi)^2-\alpha^2}]^{n+1}}{(n+1)\alpha^{n+1}} \quad (31)$$

Taking Laplace inverse of (31), we obtain

$$P_c(t)=\sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{(-1)^m}{\gamma}\left(\frac{\alpha}{2\lambda p}\right)^{n+1}\binom{n}{m}\left[\int_0^t A(t-u)\int_0^u B^{C(m)}(u-v)\exp\{-(\lambda p+\gamma+\psi)v\}\frac{I_{n+1}(\alpha v))}{v}dudv+\psi\int_0^t H(t-u)\int_0^u B^{C(m)}(u-v)\exp\{-(\lambda p+\gamma+\psi)v\}\frac{I_{n+1}(\alpha v))}{v}dudv\right] \quad (32)$$

where $H(t)=\int_0^t A(u)du$ and $B^{C(m)(t)}$ is m – fold convolution of B(t) with itself with $B^{C(0)}=\delta(t)$, the Dirac - delta function. Now, the Laplace inverse of equation (25) yields,

$$P_k(t)=b_{k,0}(t)+\psi\int_0^t b_{k,0}(u)du+\gamma\int_0^t b_{k,c=1}(u)P_c(t-u)du,\ k=0,1,\dots,c-1 \quad (33)$$

where $P_c(u)$ is given in (32). Thus, the equations (16), (32) and (33) determine all the transient state probabilities. Hence, the time-dependent probabilities of the model are obtained explicitly.

V. Mean and Variance

In this section we derive the expressions for time-dependent mean and variance of the queuing system.

Mean, M(t): The mean number of customers in the system at time t is given by:

$$E[X(t)\}=M(t)=m(t)+r(t)=\sum_{n=1}^{c-1}nP_n(t)+\sum_{n=c}^{\infty}nP_n(t) \quad (34)$$

$$M(0)=m(0)+r(0)=\sum_{n=1}^{c-1}nP_n(0)+\sum_{n=c}^{\infty}nP_n(0)$$

$$M'(t)=m'(t)+r'(t)=\sum_{n=1}^{c-1}nP_n'(t)+\sum_{n=c}^{\infty}nP_n'(t)$$

Multiplying (1)-(3) by n and summing over the range of n, we get

$$M'(t)=\lambda\sum_{n=0}^{c-1}P_n(t)-\psi[m(t)+r(t)]-\lambda p\sum_{n=c}^{\infty}nP_n(t)+\lambda p\sum_{n=c+1}^{\infty}nP_{n-1}(t)+\sum_{i=1}^{n+1}\mu_i\sum_{n=1}^{c-1}nP_{n+1}(t)-\sum_{i=1}^{n}\mu_i\sum_{n=1}^{c-1}nP_n(t)-\sum_{i=1}^{c}\mu_i\sum_{n=c}^{\infty}nP_n(t)+\sum_{i=1}^{c}\mu_i\sum_{n=c}^{\infty}nP_{n+1}(t)$$

On solving above equation we get

$$M'(t)=-\psi M(t)+\sum_{n=1}^{c-1}P_n(t)(\lambda-\sum_{i=1}^{n}\mu_i)+\lambda P_0(t)+\sum_{n=c}^{\infty}P_n(t)\,(\lambda p-\sum_{i=1}^{c}\mu_i)$$

The above equation is of the form $y'+Py=Q$ whose solution is

$$E[X(t)\}=M(t)=(\lambda-\sum_{i=1}^{n}\mu_i)\sum_{n=1}^{c-1}\int_0^t P_n(u)\exp\left(-\psi(t-u)\right)du+\lambda\int_0^t P_0(u)\exp\left(-\psi(t-u)\right)du+$$

$(\lambda p - \sum_{i=1}^{c} \mu_i) \sum_{n=c}^{\infty} \int_0^t P_n(u) \exp(-\psi(t-u))\, du$ (35)

Variance, V(t): The variance of number of customers in the system at time t is given by:

$$V(t) = K(t) - [M(t)]^2 \quad (36)$$

$K(t) = E[X^2(t)] = k(t) + l(t) = \sum_{n=1}^{c-1} n^2 P_n(t) + \sum_{n=c}^{\infty} n^2 P_n(t)$

$K'(t) = k'(t) + l'(t) = \sum_{n=1}^{c-1} n^2 P'_n(t) + \sum_{n=c}^{\infty} n^2 P'_n(t)$

Multiplying (1)-(3) by n^2 and summing over the range of n, we get

$K'(t) = \lambda \sum_{n=0}^{c-1} (2n+1) P_n(t) - \psi K(t) + \lambda p [\sum_{n=c}^{\infty} P_n(t)] + 2 \sum_{n=c}^{\infty} n P_n(t) + \sum_{i=1}^{c} \mu_i (c-1)^2 P_c(t) + 2 \sum_{i=1}^{c} \mu_i P_{c-1}(t) - \sum_{n=1}^{c-1} \sum_{i=1}^{n} \mu_i (2n+1) P_n(t) + \sum_{i=1}^{c} \mu_i [\sum_{n=c+1}^{\infty} P_n(t) - c^2 P_c(t) + 2c P_c(t) - 2 \sum_{n=c}^{\infty} n P_n(t)]$

On solving above equation we get

$K'(t) = -\psi K(t) + [\lambda p + \sum_{i=1}^{c} \mu_i] \sum_{n=c}^{\infty} P_n(t) + 2 \sum_{i=1}^{c} \mu_i P_{c-1}(t) + 2[\lambda p - \sum_{i=1}^{c} \mu_i] r(t) + [\lambda + \sum_{i=1}^{n} \mu_i] \sum_{n=1}^{c-1} P_n(t) + 3\lambda P_0(t) + 2[\lambda - \sum_{i=1}^{n} \mu_i] m(t)$

The above equation is of the form $y' + \mathrm{P}y = \mathrm{Q}$ whose solution is

$K(t) = [\lambda p + \sum_{i=1}^{c} \mu_i] \sum_{n=c}^{\infty} \int_0^t P_n(u) \exp(-\psi(t-u))\, du + 2 \sum_{i=1}^{c} \mu_i \int_0^t P_{c-1}(u) \exp(-\psi(t-u))\, du + 2[\lambda p - \sum_{i=1}^{c} \mu_i] \int_0^t (M(u) - m(u)) \exp(-\psi(t-u))\, du + 2[\lambda - \sum_{i=1}^{n} \mu_i] \int_0^t m(u) \exp(-\psi(t-u))\, du + [\lambda + \sum_{i=1}^{n} \mu_i] \sum_{n=1}^{c-1} \int_0^t P_n(u) \exp(-\psi(t-u))\, du + 3\lambda \int_0^t P_0(u) \exp(-\psi(t-u))\, du$

Therefore,

$$V(t) = [\lambda p + \sum_{i=1}^{c} \mu_i] \sum_{n=c}^{\infty} \int_0^t P_n(u) \exp(-\psi(t-u))\, du + 2 \sum_{i=1}^{c} \mu_i \int_0^t P_{c-1}(u) \exp(-\psi(t-u))\, du + 2[\lambda p - \sum_{i=1}^{c} \mu_i] \int_0^t (M(u) - m(u)) \exp(-\psi(t-u))\, du + 2[\lambda - \sum_{i=1}^{n} \mu_i] \int_0^t m(u) \exp(-\psi(t-u))\, du + [\lambda + \sum_{i=1}^{n} \mu_i] \sum_{n=1}^{c-1} \int_0^t P_n(u) \exp(-\psi(t-u))\, du + 3\lambda \int_0^t P_0(u) \exp(-\psi(t-u))\, du - [M(t)]^2 \quad (37)$$

where $M(t)$ is given in equation (35).

VI. Special Cases

Case 1 When there is no balking (i.e. $p = 0$), then the transient state probabilities are same as that of the model studied by Dharmaraja and Kumar (2015).

Case 2 If we remove the catastrophe from the model (i.e. $\psi = 0$), then the results of our model resemble with the model studied by Kumar and Arivudainambi (2001).

V. Conclusions

In this paper the transient analysis of a Markovian queuing system with heterogeneous servers, balking and catastrophes is performed. The time-dependent mean and variance of the number of customers in the system are also obtained. Some important queuing models are derived as the special cases.

References

[1] Ammar, S.I. (2014). Transient behavior of a two-processor heterogeneous system with catastrophes, server failures and repairs. *Appl. Math. Model.*, 38:2224-2234.

[2] Ancker Jr. C. J. and Gafarian, A. V. (1963a). Some queueing problems with balking and reneging I. *Oper. Res,*. 11:88-100.

[3] Ancker Jr. C. J. and Gafarian, A. V. (1963b). Some queueing problems with balking and reneging II. *Oper. Res.*, 11:928-937.

[4] Barrer D.Y. (1957). Queuing with impatient customers and ordered service. *Oper. Res.*, 5:650-656.

[5] Chao, X. (1995). A queuing network model with catastrophes and product form solution. *Oper. Res. Lett.*, 18:75-79.

[6] Dharmaraja, S. (2000). Transient solution of a two-processor heterogeneous system. *Math. Comput. Model.*, 32:1117-1123.

[7] Dharmaraja, S. and Kumar, Rakesh (2015). Transient solution of a Markovian queuing model with heterogeneous servers and catastrophes. *Opsearch*, 52:810-826.

[8] DiCrescenzo, A., Giorno, V. and Nobile, A.G. (2003). On the M/M/1 queue with catastrophes and its continuous approximation. *Queueing Syst.*, 43:329-347.

[9] El-Paoumy, M.S. and Nabwey, H.A. (2011). The Poissonian queue with balking function, reneging and two heterogeneous servers. *Int. J. Basic Appl. Sci.*, 11:149-152.

[10] Haight, F. A. (1957). Queuing with balking I. *Biometrika,* 44:360-369.

[11] Jain, N. K. and Kumar, Rakesh (2007). Transient solution of a catastrophic-cum-restorative queuing problem with correlated arrivals and variable service capacity. *Int. J. Infor. and Manag. Sci.*, 18:461-465.

[12] Jain, N. K. and Kanethia, D. K. (2006). Transient analysis of a queue with environmental and catastrophic e_ects. *Int. J. Infor. and Manag. Sci.,* 17:35-45.

[13] Kumar R. and Sharma S.K. (2012a). An M/M/1/N queuing system with retention of reneged customers. *Pak. J. Stat. Oper. Res.,* 8:859-866.

[14] Kumar, R. and Sharma, S.K. (2012b). An M/M/1/N queuing model with retention of reneged customers and balking. *Amer. J. Oper. Res.*, 2:1-5.

[15] Kumar R. and Sharma, S. (2019). Transient Solution of a Two-Heterogeneous Servers' Queuing System with Retention of Reneging Customers. *Bull. Malays. Math. Sci. Soc.,* 42:223-240.

[16] Kumar R. and Sharma S. (2018). Transient analysis of M/M/c queuing system with balking and retention of reneging customers. *Comm. Statist. Theory Methods*, 60:1318-1327.

[17] Kumar, B. K., Madheshwari, S. P. and Venkatakrishanan, K. S. (2007). Transient solution of an M/M/2 queue with heterogeneous servers subject to catastrophes. *Int. J. of Inf. Manag. Sci.*, 18:63-80.

[18] Kumar, B.K. and Arivudainambi, D. (2000). Transient solution of an M/M/1 queue with catastrophes. *Comput. Math. Appl.*, 10:1233-1240.

[19] Kumar, B. Krishna and Arivudainambi, D. (2001). Transient solution of an M/M/c queue with heterogeneous servers and balking, *Int. J. Inf. Manag. Sci.*, 12:15-27.

[20] Morse, P. M. Queues, Inventories and Management, Wiley, New York, 1958.

[21] Raju, S. N. and Bhat, U. N. (1982). A computationally oriented analysis of the G/M/1 queue. *Opsearch*, 19:67-83.

[22] Saaty, T. L. Elements of Queuing Theory with Applications. Mc-Graw Hill, New York, 1961.

[23] Sharma, O. P. and Dass, J. (1989). Initial busy period analysis for a multichannel Markovian queue. *Optimization*, 20:317-323.

[24] Sudhesh, R. (2010). Transient analysis of a queue with system disasters and customer impatience. *Queuing Syst.*, 66:95-105.

[25] Sudhesh, R, Savitha P., Dharmaraja, S. (2017). Transient analysis of a two-heterogeneous servers queue with system disaster, server repair and customers impatience. *TOP,* 25:179–205.

[26] Suranga Sampath, M. I. G. and Jicheng, L. (2018). Transient analysis of an M/M/1 queue with reneging, catastrophes, server failures and repairs. *Bull. Iran. Math. Soc.*, https://doi.org/10.1007/s41980-018-0037-6.

[27] Tarabia, A. M. K. (2011). Transient and steady-state analysis of an M/M/1 queue with balking, catastrophes, server failures and repairs. *J. Ind. Manag. Optim*., 7:811-823.

[28] Yassen, M. F. and Tarabia, A. M. K. (2017). Transient Analysis of Markovian Queueing System with Balking and Reneging Subject to Catastrophes and Server Failures. *Appl. Math. Inf. Sci.,* 11:1041-1047.

[29] Yechiali, Uri (2007). Queues with system disasters and impatient customers when system is down. *Queuing Syst.*, 56:195-202.

RT&A, No 1 (56)
Volume 15, March 2020

A Hybrid Technique of Combining AES Algorithm with Block Permutation for Image Encryption

Ritu Shaktawat[1], Rajdeep Singh Shaktawat[2], N. Lakshmi[3], Avinash Panwar[4], Arun Vaishnav[5]

•

[1,4,5]Department of Computer Science, Mohanlal Sukhadia University, Udaipur, India
[2]Department of Computer Science, College of Technology and Engineering, MPUAT, Udaipur, India
[3]Department of Physics, Mohanlal Sukhadia University, Udaipur, India

ritushaktawat85@gmail.com, shaktawat.rd@gmail.com, nlakshmi@mlsu.ac.in, avinash@mlsu.ac.in, arun.vaishnav@gmail.com

Abstract

In this paper a hybrid approach for image encryption is proposed by combining AES, a standard cryptography algorithm, along with splitting and block permutation. A standard image (Lena) is taken as an input and to enhance the security, the image is divided into 4*4 matrix followed by block permutation before encryption of the image with AES. Comparison of various parameters with and without block permutation confirms the superiority of the proposed method in terms of better results after using splitting and permutation functions.

Keywords - encryption, decryption, cryptography, AES, symmetric algorithm.

I. Introduction

Since the growth of digital communication using internet is increasing at a rapid rate, security of personal/sensitive or commercially valuable information against unauthorized access, misuse and disclosure poses an ever increasing challenge. Images are used in various fields like e- commerce, medical imaging, multimedia, telemedicine, military etc. and necessarily have very confidential information in most of these applications. Currently various standard algorithms exist which were initially designed for text data but are not very effective when it comes to image data [1, 2]. This is because since images have properties which are different from those of texts, these algorithms cannot be implemented directly over the image data and so a different encryption process needs to be adopted. Encryption is a part of cryptography which converts data into unintelligible form so that the original content cannot be accessed or utilized by an unauthorized person and is a common technique to protect data in the form of text or as images [3]. In particular, there has been a huge increase in the use of images in various applications such as in multimedia systems, internet communication, medical imaging, telemedicine, e-commerce, military communication etc., which, in turn, has generated intense research in the field of image encryption. The present work makes use of a symmetric block cipher algorithm, Advance Encryption Standard (AES) since it is a widely used and standard algorithm which is also readily available. The popularity of AES for encryption of texts is because it requires relatively less memory for implementation, is fairly robust against attacks and is also fast. However, the direct application of AES to images does not replicate the results as in the case of texts. In order to extend the application of AES to images with enhanced security, in the present work, functions like splitting, rotating and reshuffling of blocks have been performed over the image before implementing AES. To test the efficiency of the algorithm, a standard coloured Lena image of size 256*256 in jpeg format has been used as input image in this work.The encryption and decryption processes have been carried out in JAVA followed by testing and analysis of various parameters using MATLAB. Following the introduction which forms Section 1, the rest of the paper is organized as follows: Section 2 presents the literature review, Section 3 gives details of the AES algorithm, Section 4 presents the proposed approach, Section 5 discusses and analyzes the experimental results and section 6 consists of conclusions.

II. Background

Zeghid et al. [4] have implemented a new, modified version of AES over images. They have added a key stream generator to AES to remove textured zones in existing algorithms. The authors have also compared the scheme with other existing symmetric cryptography algorithms.
Using an AES algorithm, Deshmukh [5] has reconstructed the image without any distortion and concluded that the algorithm is strong enough against most common attacks such as the plaintext attacks, brute force attack and cipher attacks because of its extremely large security key space.
Karwande and Mirza [6] have first split an image into a 3*3 matrix and then applied the AES algorithm over the split image to provide more security. In the decryption process, the reverse order was followed; i.e., the split image was obtained first followed by the original image.
Brindha et al. [7] have carried out image encryption using symmetric block cipher algorithm and have made use of the DES algorithm for this purpose. They have first converted the image into byte array which was then converted into string which was used as an input. Authors have also compared the implemented algorithm with AES algorithm.
In an earlier work, Shaktawat et al. [8] had implemented three symmetric block cipher algorithms AES, DES and Blowfish over a real image and then made a comparative study to show the efficiency of these algorithms over the image data. The results of this study indicated that DES has very good performance followed by Blowfish while AES showed the lowest performance.

III. Advance Encryption Standard (AES) algorithm

Advance Encryption Standard is also known as Rijndael algorithm [1]. AES is a block cipher which was developed by two Belgium cryptographers, Joan Daemen and Vincent Rijmen, in 2000 [2]. Daemen and Vincent submitted their algorithm to NIST for the selection process of AES and it was among five finalist algorithms. The AES algorithm developed by Daemen and Rijmen became effective as a US Federal Government standard on 26th May 2002, after approval by the secretary of commerce [1-3]. It is being used worldwide for text encryption and is a well-known symmetric block cipher algorithm. It supports key sizes of 128, 192 and 256 bit. There are 10, 12and 14 rounds for 128, 192 and 256 bit keys respectively [9]. Figure 1 shows the block diagram of AES.

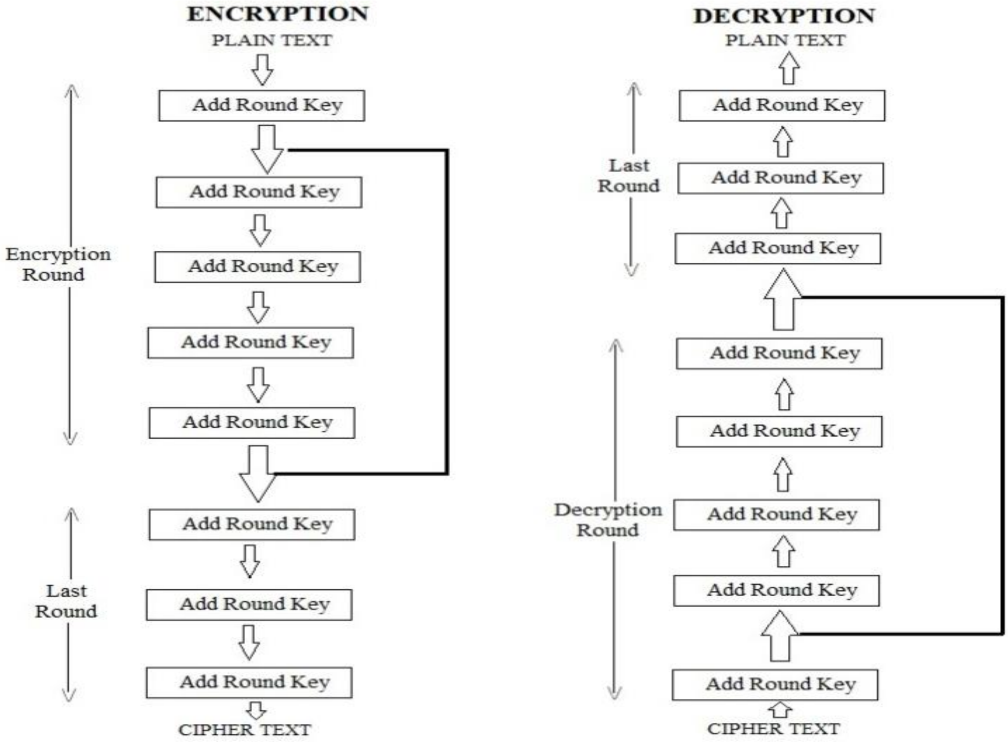


Figure 1: *Encryption and decryption process of AES*

IV. Proposed work

In this section, the proposed approach of implementing AES algorithm using block permutation on an image is presented. Implementation of AES algorithm is done in JAVA and experimental analysis is done using MATLAB. Initially a standard, coloured image (Lena) with size of 256*256 has been taken as an input from GOOGLE. The image is divided into 4*4 matrix which results in 16 blocks. These blocks are then permuted or rotated to change the position of blocks and pixels in the image. A 128 bit key, generated using a random number generator, is used to encrypt the reshuffled image through the AES algorithm. Figures 2 and 3 show the block diagram of proposed encryption and decryption methods respectively.

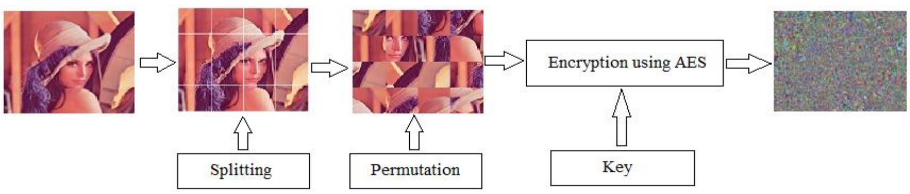


Figure 2: *Block diagram of proposed encryption method*

Figure 3: *Block diagram of proposed decryption method*

In the decryption process, the cipher image is taken as an input and decrypted with AES algorithm using the same key generated for encryption process. The permuted image is obtained after decryption and then reverse permutation operation is applied to the image to obtain the split image which is then combined with all sixteen blocks to get the original image. As shown in figure 3, after decryption the same image with same quality as that of the original is obtained as the original image.

V. Statistical Analysis and Result

i. **Histogram:** A histogram of an image is a graphical representation of the tonal distribution in an image. The entire tonal value can be judged by looking at the histogram since it shows the frequency of similar pixels in an image [10]. The histograms of images encrypted using the proposed method are uniform and are significantly different from the histogram of the original image (Figures 4, 5 and 6). On comparing the histograms of the image which has been encrypted without using permutation (Figure 5) with the one which has been encrypted after permutation (Figure 6), it is evident that the histogram displayed in Figure 6, i.e, of the image which was first split and then permuted followed by encryption is much more uniform. This means that using the proposed hybrid method of combining AES with block permutation operation results in a very low frequency of similar pixels. On comparison of the histograms in Figures 4 and 5 (original and encrypted images respectively) it is also evident that the decrypted image is similar to the original one.

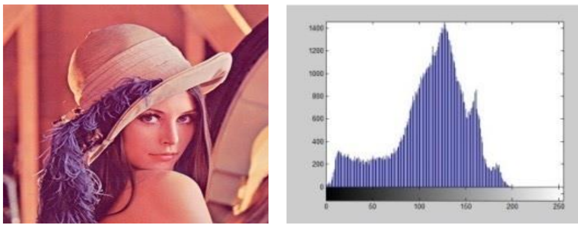


Figure 4: *Original image and Histogram*

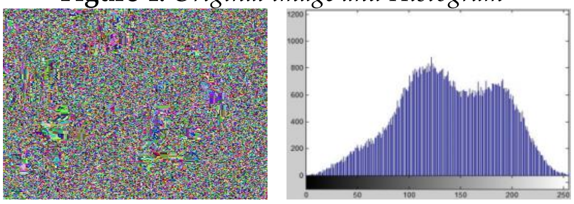


Figure 5: *Encrypted image and Histogram with AES (without block permutation)*

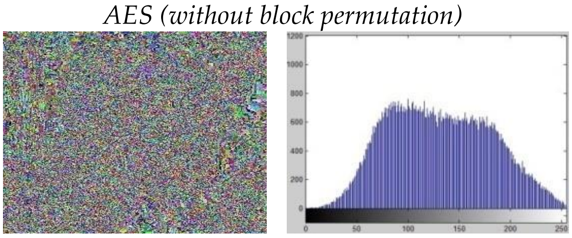


Figure 6: *Encrypted image and Histogram with AES using block permutation*

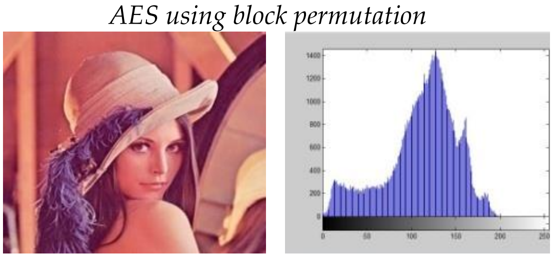


Figure 7: *Decrypted image and Histogram*

ii. **Entropy:** Entropy simply indicates the disorder of the content in an image and is calculated to test the randomness of pixels in an image. The highest value of entropy is 8 [11, 12]. It can be observed from table 1 that a higher level of entropy, with a value close to 8, is obtained with encryption process after using permutation function with AES algorithm.

iii. **PSNR:** PSNR computes the peak signal to noise ratio of an image and is used to measure the quality an encrypted image [13]. A low value of PSNR indicates that the original image and encrypted image are significantly different from each other pointing to the higher security of the encrypted image [14]. From table 1 it is evident that implementing the technique proposed in the present work, which consists of AES with splitting and permutation, yields a lower PSNR ratio than on application of directly implemented AES.

iv. **Correlation:** Correlation indicates any statistical relationship between the pixels of an image. It indicates the dependency of pixels or how the pixels in an image are correlated [15]. The adjacent pixels in an image are always highly correlated and encryption process spreads the pixels in the image. The highest value of correlation is 1. A low value of correlation implies that the adjacent pixels have less dependency and are not easily predictable [16, 17]. From table 1, it is observed that in the approach adopted in the present work, the correlation parameter has very low values, particularly for the horizontal case. This implies that using the proposed method results in a very low correlation among the adjacent pixels, which, in turn indicates stronger security than on direct implementation of AES.

Table 1: *Comparison between AES and AES using block permutation*

Algorithm	Entropy	PSNR	Correlation		
			Vertical	Horizontal	Diagonal
AES	7.5487	11.37	0.0324	0.0010	0.0778
AES using Block Permutation (present work)	7.7926	10.04	0.0205	0.0003	0.0410

VI. Conclusion

In this paper, a standard, coloured image of Lena is taken as an input and the outcome of the implementation of block based symmetric encryption algorithms AES is presented over the image based on some statistical parameters. The superior performance of this standard cryptography algorithm over the image data after the implementation of substitution and transposition approaches by using splitting and permutation functions is amply demonstrated. This is because potential attackers cannot predict that image has first been divided and re-arranged through a random permutation before being encrypted with AES. The study and results of statistical analysis in the case of the hybrid approach adopted in this work shows that the performance is very good with a positive outcome in terms of randomness of the pixels of the encrypted image as it has the higher entropy and PSNR values than the original or direct approach since the histogram and correlation test establishes the reduced dependency and predictability of pixels. This work thus shows how very good results can be achieved by a rather simple modification to a standard AES algorithm which already exists. We are currently in the process of extending this work by exploring the possibility of including other functions for substitution and/or transposition of the image to this standard cryptographic algorithm to improve its performance and hence to make it more versatile in the field of image oriented applications.

References

[1] Stallings W., 2011, *Network Security Essential: Applications and Standards*, 4th edn. Prentice Hall, United States of America.

[2] Kahate A., 2013, *Cryptography and Network Security*, 3rd edn. Tata McGraw Hill, NY, U.S.A.

[3] Gupta B., Agrawal D. P., Yamaguchi S., 2016, *Handbook of research on modern cryptographic solutions for computer and cyber security*, 1st edn. IGI Global, United States of America.

[4] Zeghid, M., Machhout, M., Khriji, L., Baganne, A., & Tourki, R., (2007), "A modified AES based algorithm for image encryption", *International Journal of Computer Science and Engineering*, 1(1), 70-75.

[5] Deshmukh, P., (2016), "An image encryption and decryption using AES algorithm", *International Journal of Scientific & Engineering Research*, 7(2), 210-213.

[6] Karwande, V. S. & Mirza, N., (2014), "Image Encryption using AES Encryption Algorithm", *International Journal & Magazine of Engineering, Technology, Management and Research*. 1(11), 214-218.

[7] Brindha, K., Sharma, R., & Saini, S., (2014), "Use of symmetric algorithm for image encryption", *International Journal of Innovative Research in Computer and Communication Engineering*, 2(5), 4401-4407.

[8] Shaktawat, R., Shaktawat, R. S., Suwalka, I., & Lakshmi, N., (2020), "Implementation of Block-Based Symmetric Algorithms for Real Image Encryption", In *Microservices in Big Data Analytics*, Springer, Singapore, pp. 127-140.

[9] Mulani, A. O., & Mane, P. B., (2019), "High-Speed Area-Efficient Implementation of AES Algorithm on Reconfigurable Platform", In *Computer and Network Security*. IntechOpen.

[10] Mathworks support Homepage, https://in.mathworks.com/help/images/ref/imhist.html, last accessed 2019/09/11.

[11] Mathworks support Homepage,https://in.mathworks.com/help/images/ref/entropy.html, last accessed 2020/01/10.

[12] Sarkar, A., Karforma, S., (2018), "Image Steganography using Password Based Encryption Technique to secure e-Banking Data", *International Journal of Applied Engineering Research*, 13(22), 15477-15483.

[13] Zheng, Q., Wang, X., Khan, M. K., Zhang, W., Gupta, B. B., Guo, W., (2018), "A lightweight authenticated encryption scheme based on chaotic scml for railway cloud service", *IEEE Access*, 6, 711-722.

[14] Mathworks support Homepage, http://in.mathworks.com/help/vision/ref/psnr.html, last accessed 2020/01/05.

[15] Mathworks support Homepage, https://in.mathworks.com/help/images/ref/corr2.html, last accessed 2019/12/05.

[16] Sivakumar, T., Venkatesan, R., (2014), "A novel approach for image encryption using dynamic SCAN pattern", *IAENG International Journal of Computer Science*, 41(2), 91-101.

[17] Tiwari, H., Hamsapriye, N., (2018), "Logistic map based image encryption scheme", *International Journal of Applied Engineering Research*, 13(23), 16573-16577.

Variances of Linear Regression Coefficients for Safety Margin of Technical System

Gurami Tsitsiashvili, Marina Osipova
•
Russia, 690041, Vladivostok, Radio street 7
Institute for Applied Mathematics
Far Eastern Branch of Russian Academy Sciences
guram@iam.dvo.ru, mao1975@list.ru

Abstract

In this paper, we choose such a particular formulation of the problem of calculating linear regression coefficient, when the moments of observation form an arithmetic progression or depend in proportion to some degree of the observation number and variances of the estimation of the safety margin of the technical system decrease over time. It is proved that the variance of the trend estimation in this case decreases as a certain degree of the length of the series of observations. This makes it possible to evaluate the effectiveness of non-destructive testing for the safety margin of the technical system.

Keywords: linear regression coefficient, variance calculating, independent samples of observations.

1 Introduction

The problem of studying the variance of the linear trend estimation and its dependence on the length of the time series, that this estimation is based on, is of both theoretical and practical interest. This problem is closely related to the problem of small samples in mathematical statistics. In reliability theory, this problem occurs when using linear regression analysis to predict the safety margin of a technical system (see, for example, [1], [2]). In this regard, of particular interest is the task set by O. V. Abramov to study the variance of the estimation when the intervals between successive moments of observation and the variance of the estimation of the safety margin of the technical system decrease over time.

By analogy with [3] in this paper, we choose such an honest statement of this problem, when the observation moments form an arithmetic progression or grow as less than one degree of the observation number. Similarly, it is assumed that the variance of the observation estimate also decreases in proportion to some degree of the observation number. It is proved that the variance of the trend estimation in this case decreases as a certain degree of the length of the series of observations. This makes it possible to use short series of observations to evaluate the linear trend and also to evaluate the effectiveness of non-destructive testing for the safety margin of the technical system [4] – [6].

Consider the following linear regression model $x(t) = y(t) + \varepsilon(t),\ y(t) = at + b$. Assume that at times $t_1, \dots, t_n,\ 0 \le t_1 < t_2 < \cdots < t_N$, measured values are $y(t_1), \dots, y(t_N)$ with random errors $\varepsilon_1, \dots, \varepsilon_N$. The random variables $\varepsilon_1, \dots, \varepsilon_N$ are assumed to be independent, equally distributed with zero mean and variance σ^2.

To solve this problem, replace the variable $\tilde{t} = t - T_N$, $T_N = \frac{\sum_{k=1}^N t_k}{N}$, and define a linear function

$$\tilde{y}(t) = y(t + T_n) = at + b + aT_N = at + \tilde{b}, \quad \sum_{k=1}^N \tilde{t}_k = 0, \quad \tilde{b} = b + aT_N.$$

To do this, we compute $\tilde{t}_k$, $k = 1, \dots, N$, and construct the least squares [7], [8] estimates of the coefficients a, $\tilde{b}$ of the linear regression function $\tilde{y}(t) = at + \tilde{b}$ from observations

$$\tilde{x}_1 = \tilde{y}(\tilde{t}_1) + \sigma_1 \tilde{\varepsilon}_1, \dots, \tilde{x}_N = \tilde{y}(\tilde{t}_N) + \sigma_N \tilde{\varepsilon}_N.$$

Here $\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_1$ are independent random variables with zero means and unit variances.

The solution of this problem is a random vector consisting of estimates

$$\hat{a}_N = \frac{\sum_{k=1}^N \tilde{x}_k \tilde{t}_k}{\sum_{k=1}^N \tilde{t}_k^2}, \qquad \hat{b}_N = \frac{\sum_{k=1}^N \tilde{x}_k}{N}$$

of coefficients a, $\tilde{b}$ of linear function $\tilde{y}(t)$. The components of this vector have the following averages, variances, and covariance coefficient:

$$M\hat{a}_N = a, \qquad M\hat{b}_N = \tilde{b}, \qquad D\hat{a}_N = \frac{\sum_{k=1}^N \sigma_k^2 \tilde{t}_k^2}{(\sum_{k=1}^N \tilde{t}_k^2)^2}, \qquad D\hat{b}_N = \frac{\sum_{k=1}^N \sigma_k^2}{N^2}, \qquad cov(\hat{a}_N, \hat{b}_N) = \frac{\sum_{k=1}^N \tilde{t}_k \sigma_k^2}{N \sum_{k=1}^N \tilde{t}_k^2}. \tag{1}$$

2 Main Results

Statement 1. *The following equalities are true*

$$\sum_{k=1}^N \tilde{t}_k^2 = \sum_{k=1}^N t_k^2 - \frac{1}{N}\left(\sum_{k=1}^N t_k\right)^2, \tag{2}$$

$$\sum_{k=1}^N \sigma_k^2 \tilde{t}_k^2 = \sum_{k=1}^N \sigma_k^2 t_k^2 + \sum_{k=1}^N \frac{\sigma_k^2}{N^2} \left(\sum_{i=1}^N t_i\right)^2 - \frac{2}{N} \sum_{k=1}^N \sigma_k^2 t_k \sum_{i=1}^N t_i. \tag{3}$$

Proof. Statement 1 follows from the following equalities

$$\sum_{k=1}^N \tilde{t}_k^2 = \sum_{k=1}^N \left(t_k - \frac{1}{N}\sum_{i=1}^N t_i\right)^2 = \sum_{k=1}^N \left(t_k^2 - \frac{2t_k}{N}\sum_{i=1}^N t_i + \left(\frac{1}{N}\sum_{i=1}^N t_i\right)^2\right) =$$

$$= \sum_{k=1}^N t_k^2 - \frac{1}{N}\left(\sum_{k=1}^N t_k\right)^2,$$

$$\sum_{k=1}^N \sigma_k^2 \tilde{t}_k^2 = \sum_{k=1}^N \sigma_k^2 \left(t_k - \frac{1}{N}\sum_{i=1}^N t_i\right)^2 = \sum_{k=1}^N \sigma_k^2 \left[t_k^2 + \frac{1}{N^2}\left(\sum_{i=1}^N t_i\right)^2 - \frac{2}{N} t_k \sum_{i=1}^N t_i\right].$$

Equidistant series of observations. By equidistant series of observations, we will understand such a series in which the following equalities are fulfilled

$$t_k = k, \qquad \sigma_k^2 = \sigma^2, \qquad k = 1, \dots, N.$$

Statement 2. *The following relations are valid for an equidistant series of observations*

$$D\hat{a}_N = \frac{12\sigma^2}{N^3\left(1 - \frac{1}{N^2}\right)}, \qquad D\hat{b}_N = \frac{\sigma^2}{N}, \qquad cov(\hat{a}_N, \hat{b}_N) = 0, \tag{4}$$

Proof. By induction on N it is not difficult to obtain from Formula (2) the equality

$$\sum_{k=1}^N t_k = \frac{N(N+1)}{2}, \qquad \sum_{k=1}^N t_k^2 = \frac{N^3}{3} + \frac{N^2}{2} + \frac{N}{6}$$

and so it follows from the formula (3) that

$$\sum_{k=1}^{N} \tilde{t}_k^2 = \sum_{k=1}^{N} t_k^2 - \frac{1}{N}\left(\sum_{k=1}^{N} t_k\right)^2 = \frac{N^3}{12}\left(1 - \frac{1}{N^2}\right). \tag{5}$$

Therefore, from Formulas (1), (5), we get the equality (4).

Remark 1. *If independent random variables $\varepsilon_1, \dots, \varepsilon_N$ have normal distributions with zero averages and total variance σ^2, then estimates $\hat{a}_N$, $\hat{b}_N$ are independent.*

Here are the results of calculations performed using the formula (5).

N	2	3	4	5	6	7	8	9	10
$\sum_{k=1}^{N} \tilde{t}_k^2$	0.5	2	5	10	17.5	28	42	60	82.5
$12\sum_{k=1}^{N} \tilde{t}_k^2/N^3$	0.75	0.888	0.9375	0.96	0.9722	0.9795	0.9844	0.9876	0.99

N	11	12	13	14	15	16	17	18
$\sum_{k=1}^{N} \tilde{t}_k^2$	110	143	182	227.5	280	340	408	484.5
$12\sum_{k=1}^{N} \tilde{t}_k^2/N^3$	0.9917	0.9931	0.9948	0.9949	0.9955	0.9961	0.9965	0.9969

Tabl. 1. Meanings of terms $\sum_{k=1}^{N} \tilde{t}_k^2$, $12\sum_{k=1}^{N} \tilde{t}_k^2/N^3$.

From Tabl. 1 values of term $\sum_{k=1}^{N} \tilde{t}_k^2$ and its asymptotics at $N \to \infty$ it is seen that already at $N \geq 5$, the value of $\sum_{k=1}^{N} \tilde{t}_k^2$ with a five-percent and smaller relative error is approximated by the asymptotic $N^3/12$. Next, we will write this ratio as $\sum_{k=1}^{N} \tilde{t}_k^2 \approx N^3/12$.

Remark 2. *Let's now consider a series of values $t_k^{(m)}$, $k = 1, \dots, Nm$, represented as $1/m, 2/m, \dots, Nm/m$. Then approximate equalities are valid*

$$\sum_{k=1}^{Nm} \left(\tilde{t}_k^{(m)}\right)^2 \approx \frac{N^3 m}{12}, \qquad D\hat{a}_{N,m} \approx \frac{12\sigma^2}{m\, N^3}, \qquad D\hat{b}_{N,m} = \frac{\sigma^2}{Nm}.$$

Thus, replacing a series of observation moments t_k, $k = 1, \dots, N$ with a series $t_k^{(m)}$, $k = 1, \dots, Nm$, results in approximately a m reduction in the variances of the linear regression function coefficients estimates.

Series of observations with power dependencies.

Statement 3. *Now suppose that*

$$t_k = k^\alpha,\ \sigma_k = \sigma k^{-\beta},\ 0 < \beta, \alpha. \tag{6}$$

Then the following asymptotic relations are valid

$$D\hat{a}_N \sim \frac{\sigma^2 C_*(2\alpha+1)^2(\alpha+1)^4}{N^{2\alpha+2\beta+1}\alpha^4},\ D\hat{b}_N \sim \frac{\sigma^2}{N^{2\beta+1}(1-2\beta)},\ 0 < 1-2\beta, \tag{7}$$

$$D\hat{a}_N \sim \frac{\sigma^2 \ln N(2\alpha+1)^2(\alpha+1)^2}{N^{2\alpha+2}\alpha^4},\ D\hat{b}_N \sim \frac{\sigma^2 \ln N}{N^2},\ 0 = 1-2\beta, \tag{8}$$

$$D\hat{a}_N \sim \frac{\sigma^2 C^*(2\alpha+1)^2(\alpha+1)^2}{N^{2\alpha+2}\alpha^4},\ D\hat{b}_N \sim \frac{\sigma^2 C^*}{N^2},\ 1-2\beta < 0, \tag{9}$$

with

$$C_* = \frac{1}{2\alpha-2\beta+1} + \frac{1}{(1-2\beta)(\alpha+1)^2} - \frac{2}{(\alpha-2\beta+1)(\alpha+1)},\ 0 < 1-2\beta,$$

$$C^* = \sum_{k=1}^{\infty} k^{-2\beta}, \qquad 1-2\beta < 0.$$

Proof. The following inequalities are valid

$$\int_1^{N+1} \tau^\gamma d\tau \le \sum_{k=1}^N k^\gamma \le 1 + \int_1^N \tau^\gamma d\tau, \qquad \gamma < 0,$$

$$1 + \int_1^N \tau^\gamma d\tau \le \sum_{k=1}^N k^\gamma \le \int_1^{N+1} \tau^\gamma d\tau, \qquad 0 \le \gamma,$$

This is followed by asymptotic relations for $N \to \infty$

$$\sum_{k=1}^N k^\gamma \sim \frac{N^{\gamma+1}}{\gamma+1}, \qquad \gamma > -1; \ \sum_{k=1}^N k^{-1} \sim \ln N; \ \sum_{k=1}^N k^\gamma \to \sum_{k=1}^\infty k^\gamma < \infty, \qquad \gamma < -1. \tag{10}$$

and so

$$\sum_{k=1}^N \tilde{t}_k^2 \sim \frac{N^{2\alpha+1}\alpha^2}{(\alpha+1)^2(2\alpha+1)}, \qquad \alpha > 0. \tag{11}$$

Consider the case when $0 < 1 - 2\beta$, then from Formulas (10), (11) we get

$$\sum_{k=1}^N \sigma_k^2 t_k^2 \sim \frac{\sigma^2 N^{2\alpha-2\beta+1}}{2\alpha - 2\beta + 1}, \qquad \sum_{k=1}^N \sigma_k^2 t_k \sim \frac{\sigma^2 N^{\alpha-2\beta+1}}{\alpha - 2\beta + 1}, \qquad \sum_{k=1}^N \sigma_k^2 \sim \frac{\sigma^2 N^{1-2\beta}}{1-2\beta}$$

and consequently

$$\sum_{k=1}^N \sigma_k^2 \tilde{t}_k^2 \sim \sigma^2 C_* N^{2\alpha-2\beta+1},$$

$$D\hat{a}_N \sim \frac{\sigma^2 C_* (2\alpha+1)^2(\alpha+1)^4}{N^{2\alpha+2\beta+1}\alpha^4}, \ D\hat{b}_N \sim \frac{\sigma^2}{N^{2\beta+1}(1-2\beta)}. \tag{12}$$

If $1 - 2\beta = 0$, then

$$\sum_{k=1}^N \sigma_k^2 t_k^2 \sim \frac{\sigma^2 N^{2\alpha}}{2\alpha}, \qquad \sum_{k=1}^N \sigma_k^2 t_k \sim \frac{\sigma^2 N^{\alpha}}{\alpha}, \qquad \sum_{k=1}^N \sigma_k^2 \sim \sigma^2 \ln N$$

and so

$$\sum_{k=1}^N \sigma_k^2 \tilde{t}_k^2 \sim \frac{\sigma^2 N^{2\alpha} \ln N}{(\alpha+1)^2}, \qquad D\hat{a}_N \sim \frac{\sigma^2 \ln N (2\alpha+1)^2(\alpha+1)^2}{N^{2\alpha+2}\alpha^4}, \qquad D\hat{b}_N \sim \frac{\sigma^2 \ln N}{N^2}. \tag{13}$$

Now let's go to the case where $1 - 2\beta < 0, \ \alpha - 2\beta + 1 > 0$, then from Formulas (10), (11) we get

$$\sum_{k=1}^N \sigma_k^2 t_k^2 \sim \frac{\sigma^2 N^{2\alpha-2\beta+1}}{2\alpha - 2\beta + 1}, \qquad \sum_{k=1}^N \sigma_k^2 t_k \sim \frac{\sigma^2 N^{\alpha-2\beta+1}}{\alpha - 2\beta + 1}, \qquad \sum_{k=1}^N \sigma_k^2 \to \sigma^2 C^*,$$

and consequently

$$\sum_{k=1}^N \sigma_k^2 \tilde{t}_k^2 \sim \frac{\sigma^2 C^* N^{2\alpha}}{(\alpha+1)^2}, \qquad D\hat{a}_N \sim \frac{\sigma^2 C^* (2\alpha+1)^2(\alpha+1)^2}{N^{2\alpha+2}\alpha^4}, \qquad D\hat{b}_N \sim \frac{\sigma^2 C^*}{N^2}. \tag{14}$$

If $1 - 2\beta < 0, \ \alpha - 2\beta + 1 = 0$, then

$$\sum_{k=1}^N \sigma_k^2 t_k^2 \sim \frac{\sigma^2 N^{\alpha}}{\alpha}, \qquad \sum_{k=1}^N \sigma_k^2 t_k \sim \sigma^2 \ln N, \qquad \sum_{k=1}^N \sigma_k^2 \to \sigma^2 C^*$$

and so

$$\sum_{k=1}^N \sigma_k^2 \tilde{t}_k^2 \sim \frac{\sigma^2 C^* N^{2\alpha}}{(\alpha+1)^2}, \qquad D\hat{a}_N \sim \frac{\sigma^2 C^* (2\alpha+1)^2(\alpha+1)^2}{N^{2\alpha+2}\alpha^4}, \qquad D\hat{b}_N \sim \frac{\sigma^2 C^*}{N^2}. \tag{15}$$

Now consider the case where $\alpha - 2\beta + 1 < 0,\ 2\alpha - 2\beta + 1 > 0$, then from Formulas (10), (11) we get

$$\sum_{k=1}^{N} \sigma_k^2 t_k^2 \sim \frac{\sigma^2 N^{2\alpha-2\beta+1}}{2\alpha - 2\beta + 1}, \qquad \sum_{k=1}^{N} \sigma_k^2 t_k \to \sigma^2 \sum_{k=1}^{\infty} k^{\alpha-2\beta} < \infty, \qquad \sum_{k=1}^{N} \sigma_k^2 \to \sigma^2 C^*$$

and consequently

$$\sum_{k=1}^{N} \sigma_k^2 \tilde{t}_k^2 \sim \sigma^2 C^* N^{2\alpha} (\alpha+1)^2, \qquad D\hat{a}_N \sim \sigma^2 C^* (2\alpha+1)^2 (\alpha+1)^2 N^{2\alpha+2} \alpha^4, \qquad D\hat{b}_N \sim \sigma^2 C^* N^2. \tag{16}$$

If $\alpha - 2\beta + 1 < 0,\ 2\alpha - 2\beta + 1 = 0$, then

$$\sum_{k=1}^{N} \sigma_k^2 t_k^2 \sim \sigma^2 \ln N, \qquad \sum_{k=1}^{N} \sigma_k^2 t_k \to \sigma^2 \sum_{k=1}^{\infty} k^{\alpha-2\beta} < \infty, \sum_{k=1}^{N} \sigma_k^2 \to \sigma^2 C^*$$

and so

$$\sum_{k=1}^{N} \sigma_k^2 \tilde{t}_k^2 \sim \frac{\sigma^2 C^* N^{2\alpha}}{(\alpha+1)^2}, \qquad D\hat{a}_N \sim \frac{\sigma^2 C^* (2\alpha+1)^2 (\alpha+1)^2}{N^{2\alpha+2} \alpha^4}, \qquad D\hat{b}_N \sim \frac{\sigma^2 C^*}{N^2}. \tag{17}$$

Now let's go to the case where $2\alpha - 2\beta + 1 < 0$, then from Formulas (10), (11) we get

$$\sum_{k=1}^{N} \sigma_k^2 t_k^2 \to \sigma^2 \sum_{k=1}^{\infty} k^{2\alpha-2\beta} < \infty, \qquad \sum_{k=1}^{N} \sigma_k^2 t_k \to \sigma^2 \sum_{k=1}^{\infty} k^{\alpha-2\beta} < \infty, \qquad \sum_{k=1}^{N} \sigma_k^2 \to \sigma^2 C^*$$

and consequently

$$\sum_{k=1}^{N} \sigma_k^2 \tilde{t}_k^2 \sim \frac{\sigma^2 C^* N^{2\alpha}}{(\alpha+1)^2}, \qquad D\hat{a}_N \sim \frac{\sigma^2 C^* (2\alpha+1)^2 (\alpha+1)^2}{N^{2\alpha+2} \alpha^4}, \qquad D\hat{b}_N \sim \frac{\sigma^2 C^*}{N^2}. \tag{18}$$

From Formulas (12) – (18) we get the asymptotic relations for $N \to \infty$ (7) – (9).

Remark 3. *In the first quadrant of the coefficient values $\alpha > 0, \beta > 0$ in Fig. 1 shows the areas*

$$\Gamma_1 = \{0 < 1 - 2\beta\}, \qquad \Gamma_2 = \{0 = 1 - 2\beta\}, \qquad \Gamma_3 = \{1 - 2\beta < 0, \alpha - 2\beta + 1 > 0\},$$

$$\Gamma_4 = \{1 - 2\beta < 0, \alpha - 2\beta + 1 = 0\}, \qquad \Gamma_5 = \{\alpha - 2\beta + 1 < 0, 2\alpha - 2\beta + 1 > 0\},$$

$$\Gamma_6 = \{\alpha - 2\beta + 1 < 0, 2\alpha - 2\beta + 1 = 0\}, \qquad \Gamma_7 = \{2\alpha - 2\beta + 1 < 0\}$$

for which Formulas (12), (13), (14), (15), (16), (17), (18) are obtained for the variance of estimates of coefficients of the linear regression function, respectively.

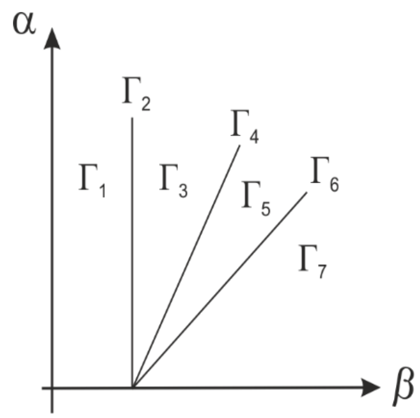


Fig 1. Domains $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5, \Gamma_6, \Gamma_7$ in the first quadrant of the coefficient values $\alpha > 0, \beta > 0$.

Remark 4. *If $\alpha < 1$ then distances between neighbour time moments t_k, t_{k+1}, $1 \le k < N$, decrease.*

3 Conclusion

The estimates of variances of linear trend coefficient estimates obtained in this work allow us to choose the moments of determining the technical system's strength reserves and their accuracy depending on the system's approximation to the rollback state.

4 Acknowledgements

The research of Gurami Tsitsiashvili is partially supported by Presidium of FEB RAS, program "Priority research for the integrated development of the far Eastern branch of RAS", project 18-5-044. The authors thank O.V. Abramov for help in setting the task in a meaningful way.

References

[1] Abramov, O. V., Tsitsiashvili, G. Sh. (2018). Prediction of failure of the controlled technical system. Informatics and control systems, 3: 42–49.
[2] Abramov, O. V., Tsitsiashvili, G. Sh. (2019). Interval estimation for the task of predicting failures of complex engineering systems. (In Russian). Proceedings of the International Symposium "Reliability and quality", 1, 113–114. (In Russian).
[3] Tsitsiashvili, G. Sh. (2019) Calculating the variance of the linear regression coefficient. (In Russian). Reliability: Theory and Applications, 14 (3): 65–68.
[4] Novokreshenov V. V. Nondestructive testing of welded joints in mechanical engineering: Textbook for academic baccalaureate /Novokreshenov V. V., Rodyakina R. V.; edited by Prokhorov N. N. 2nd ed. Moscow: Yurayt, 2018. (In Russian).
[5] Volchenko V. N., Gurvich A. K., Mayorov A. N., Kashuba L. A., Makarov E. L., Khasanov M. H. Quality control of welding /ed. - Textbook for engineering universities. Moscow: Mashinostroenie, 1975. (In Russian).
[6] GOST P 56542-2015. Non-destructive testing. Classification of types and methods. (In Russian).
[7] Rykov V. V. Mathematical statistics and experiment planning. MAKS Press, Moscow 2010. (In Russian).
[8] Borovkov A. A. Mathematical statistics and experiment planning. Additional chapters. Nauka, Moscow 1984. (In Russian).

RT&A, No 1 (56)
Volume 15, March 2020

Probabilistic Comparisons of Systems Operation Quality for Uncertainty Conditions

Andrey Kostogryzov[1], Petr Kanygin[2], Andrey Nistratov[2]

•

[1] Federal Research Center "Computer Science and Control"
of the Russian Academy of Sciences, Akostogr@gmail.com

[2] Scientific and Engineering Centre for Energy Safety,
pskanygin@gmail.com, andrey.nistratov@gmail.com

Abstract

For systems, including systems of intelligent manufacturing, information and intellectual systems, an achievement of operation quality at admissible expenses is very important and needs **prognostic** comparisons. Here the probabilistic **approach to** compare an **operation quality of functionally similar systems** for uncertainty conditions is proposed. To be compared t**here may be: different systems for an one operation time period or for different time periods with identical duration; or the same system for different time periods on time line.** The system operation outputs are considered in the forms of **material products,** information products **and products, combined from material and information products. For the given time of prediction the main results of the approach application are:** a relative **part of** functions executed with admissible quality, estimations of expenses considering inadmissible system operation quality, a relative part of system operation satisfaction connected with quality and cost. **The** approach is demonstrated by examples.

Keywords: analysis, model, estimation, operation, prediction, probability, quality, system

I. Introduction

Different complex systems needs to be compared in life cycle by covering many special aspects. For example, the output results of many system operation can be both material and information products. For compared systems the important problem is to estimate their **operation quality** for uncertainty conditions during given long time in future considering information quality and costs. **Many standards recommend to solve this problem by system analysis methods (**for example, see ISO/IEC/IEEE 15288 "Systems and software engineering — System life cycle processes", ISO/IEC/IEEE 15939 "Systems and software engineering—Measurement process", ISO 31000 "Risk management - Principles and guidelines", IEC 61508 "Functional safety of electrical / electronic / programmable electronic safety-related systems, etc.). But as a rule for the system with hypothetical or expected conditions in future these methods in details are the matter for creation. **Considering the practical needs here** the analytical **approach to** compare **complex systems operation quality** for uncertainty conditions is proposed. The systems operation with **material,** information and combined **outputs (for example, robotics) are researched.**

The approach develops the existing approaches [1-14]. It can be useful for analysis and comparisons of systems operation quality, system optimization, to rationale of quantitative system requirements and engineering solutions. Another probabilistic models which can predict probability of success or risks on a level of probability distribution functions (PDF) [1-7, 11-14] may be also applicable through this approach.

II. The assumptions and ideas for the methodological approach

The proposed methodological approach is developed by the assumptions that:

- for the compared systems an expected quality for system operation may be achieved;
- in general case all outputs of compared systems may be divided into **material products, information products and products, combined from material and information products;**
- for analyzed period of comparison in life cycle the possible expenses of satisfaction from **systems operation quality are** comprehended or approximately estimated.

The methods and models are proposed by the use of next 4 main ideas.

Idea 1. For the system and if needs, for each system elements, intended for producing **material,** information and combined **outputs** it is necessary to be able to estimate quantitatively the achieved quality level on time line:

a) concerning material **outputs** - depending on the frequency of significant changes in quality, the frequency of control measures for recovering the admissible system operation quality and the mean time of recovering;

b) concerning **information outputs** - depending on the system possibilities for reliable and timely producing complete, valid and, if needed, confidential information;

c) concerning system complexity - depending on the given set of functions executed.

Idea 2. From practice for different conditions of uncertainties there **may be compared different systems for an one operation time period or for different time periods with identical duration, or the same system for different time periods on time line. The functions may be executed with different material and/or information outputs.**

Idea 3. For the defined set of system functions it is essential to estimate systems operation on the level of relative **parts of** functions executing with admissible quality for hypothetical or expected conditions. Thus it is understood that taking into account conditions of uncertainties the level 100% is never achievable.

Idea 4. For different conditions of uncertainties the **systems operation quality may be compared** on the level of relative parts of satisfaction connected with quality and costs.

III. Proposed model for probabilistic estimations

According to the assumptions and the ideas every function execution can be described and researched by the next general models concerning **material, information and combined outputs**. The models allow to estimate probability of function execution with admissible quality.

A probabilistic space (Ω, B, P) is created [1-7], where: Ω - is a limited space of elementary events; B – a class of all subspace of Ω-space, satisfied to the properties of σ-algebra; P – a probability measure on a space of elementary events Ω. It is intended for probabilistic estimation of achieved quality level in function execution.

I. Model concerning **material outputs**

The next elementary events for function execution analysis are defined: "function execution with admissible system operation quality" and "function execution with inadmissible system operation quality". From the point of elementary event view at the moment "t" users need system ability to satisfy real requirements with the admissible quality. These requirements also are changed on time line. An elementary state of system operation can be changed on state «function execution with inadmissible system operation quality» instead of "function execution with admissible system operation quality" because of earlier significant changes in quality (at the moment t-τ). Also a time for recovering inadmissible system operation quality requires. It means system ability not to satisfy real requirements with admissible quality at the moment t.

The essence of "Black box" model is described by Figure 1.

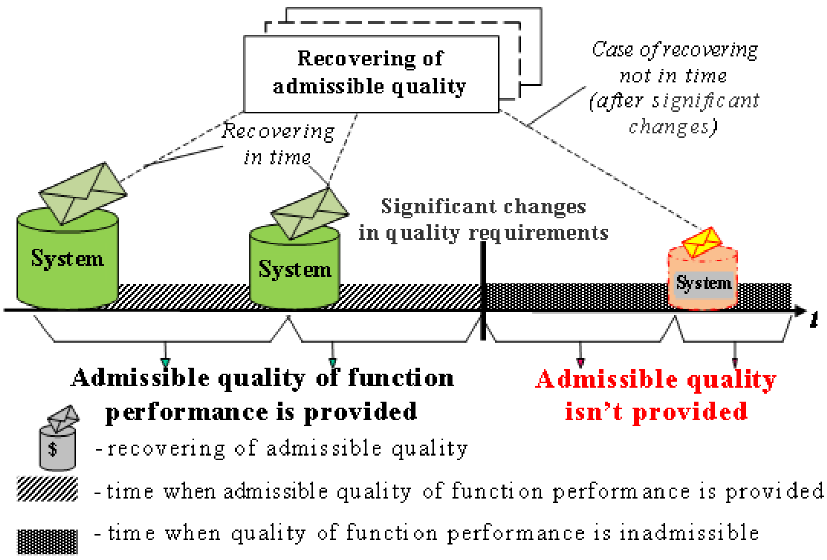


Figure 1: *Some events describing elementary events connected with function execution quality (abstraction)*

The next expression is proposed for estimation the probability of function execution with admissible system operation quality [1, 3, 5]:

$$P = \frac{\xi^2}{q(\xi+b)}\left[1 - \exp\left(-\frac{q}{\xi}\right)\right]. \tag{1}$$

Here

ξ is mean time between significant changes in items concerning admissible system operation quality, ξ^{-1} - is the frequency of significant changes in quality (considering changes of user needs);

q is mean time between control measures for recovering admissible system operation quality, q^{-1} - is the frequency of control measures for recovering admissible system operation quality;

b is mean time for recovering admissible system operation quality.

The proof see in [1, 3, 5].

For the practical use this means that the achieved quality level for material outputs can be estimated by the probability of function execution with admissible system operation quality (P) - it is calculated by (1).

II. Model concerning **information outputs**

The same space (Ω, B, P) is built [1, 3, 5] and proposed for using because system operation output quality may be considered as a quality of special information system, which reliable and timely produces complete, valid and, if needed, confidential information requested – see Figure 2.

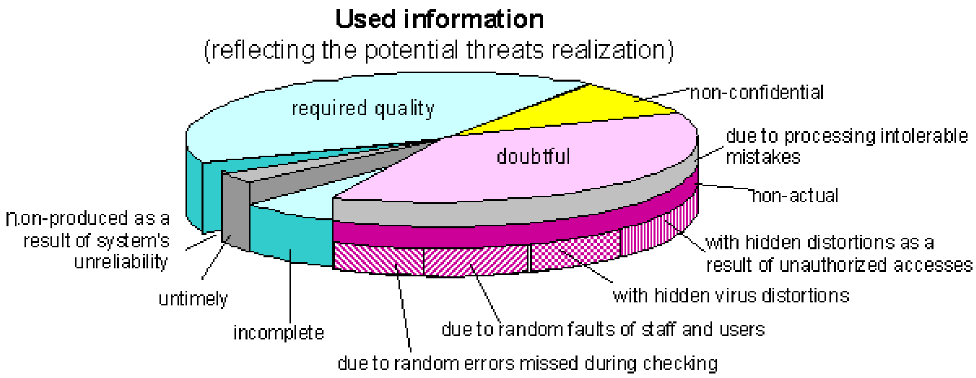


Figure 2: *An information output quality (abstraction)*

There are proposed for using the next "Black box" models and estimated measures [1, 3, 5]:

"The model of system operation by a complex system in conditions of unreliability of its components", the measures: T_{MTBF} - the mean time between failures; $P_{rel.}(T_{req.})$- the probability of reliable operation of system during the given prognostic period $T_{req.}$; $P_{man}(T_{req.})$ - the probability of providing faultless man's actions during the given prognostic period $T_{req.}$;

"The models complex of calls processing (for the different dispatcher technologies - for unpriority calls processing in a consecutive order for singletasking processing mode, in a time-sharing order for multitasking processing mode; for priority technologies of consecutive calls processing with relative and absolute priorities; for batch calls processing; for combination of technologies above), the measures: P_{tim} - the probability of well-timed processing during the given prognostic period; the relative portion of all well-timed processed calls; the relative portion of well-timed processed calls of those types for which the customer requirements are met C_{tim};

"The model of entering into system current data concerning new objects of application domain", the measure: P_{compl} - the probability that system contains complete current information about states of all objects and events;

"The model of information gathering", the measure: $P_{actual.}$ - the probability of information actuality on the moment of its use;

Note. This model is similar mathematically to the model III.I for material outputs.

"The model of information analysis", the measures: P_{check} is the probability of errors absence after checking; the fraction of errors in information after checking; $P_{process}$ - the probability of correct analysis results obtaining; the fraction of unaccounted essential information;

"The models complex of dangerous influences on a protected system", the measures: $P_{infl.}(T_{req.})$ - the probability of required counteraction to dangerous influences from threats during the given prognostic period $T_{req.}$;

"The models complex of an authorized access to system resources", the measures: P_{prot} - the probability of providing system protection from an unauthorized access by means of barriers; $P_{conf.}(T_{req.})$ - the probability of providing information confidentiality by means of all barriers during the given prognostic period $T_{req.}$.

III. Model concerning **outputs** which are combination of **material and information outputs**

For this case the models from subsections III.I and III.II are proposed to estimate every function execution quality – see Figure 3 and section IV. All the proposed models may be applied and improved for solving the problem to estimate and compare prognostic system **operation quality** for uncertainty conditions during given long time considering information quality and costs.

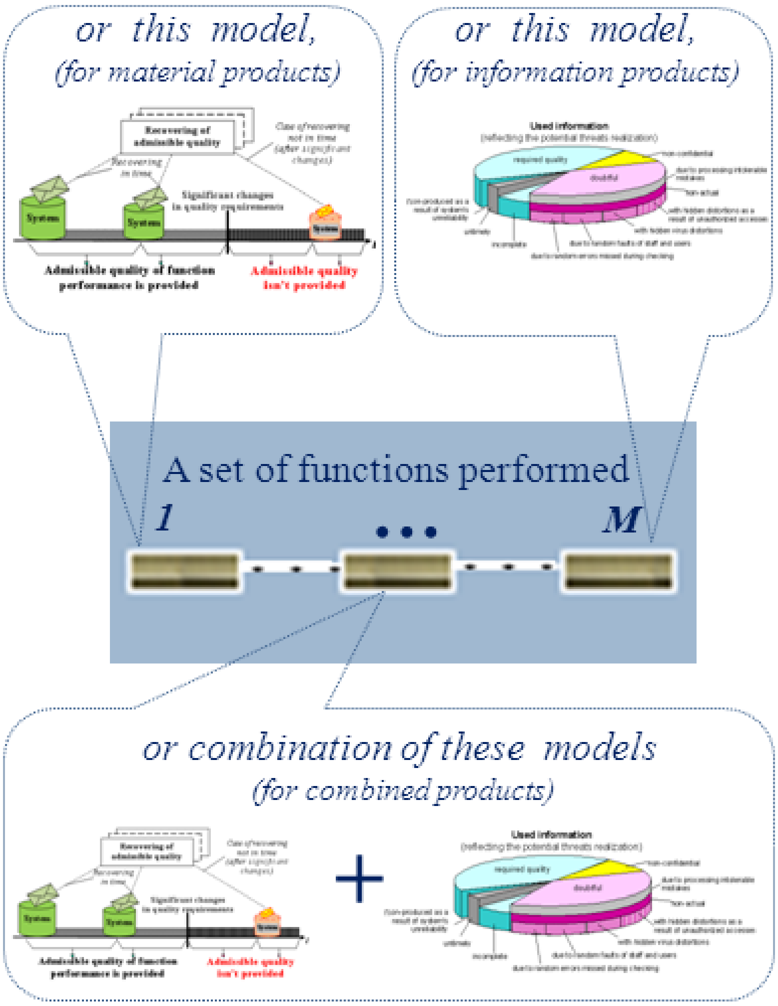


Figure 3: *Variants for a choice of the model for every function*

IV. Estimation of relative **parts of** functions executed with admissible quality

The next formula is proposed for calculation a relative **part of** functions executed with admissible system operation quality during the given prognostic period (here hypothetical conditions also may be considered):

$$S_{quality} = \sum_{m=1}^{M} P_m \cdot a_{hyp.\ m} \Big/ \sum_{m=1}^{M} a_{hyp.\ m} \tag{2}$$

where $a_{hyp.\ m}$ is frequency of m-th type function execution during the given prognostic period; P_m is the probability of m-th function execution with admissible system operation quality; M is full set of essential functions which are executed by the system and are considered in comparisons.

In general case for every m (1≤m≤M) the probability P_m is calculated by the next variants:

a) concerning **material outputs** P_m is calculated by the model (1);

b) concerning information outputs P_m is calculated by formula:

$$P_m(T_{req.}) = P_{rel.m}(T_{req.}) \cdot C_{tim\ m} \cdot P_{compl..m} \cdot P_{actual..m} \cdot P_{check\ m} \cdot P_{process.m} \cdot P_{infl.\ m}(T_{req.}) \cdot P_{man\ m}(T_{req.}) \cdot P_{prot.m} \cdot P_{conf.m}(T_{req.}), \tag{3}$$

where all measures are calculated by the models, proposed in subsection *III.II;*

c) for a combination of **material and information outputs**

$$P_m(T_{req.}) = P_{combined.m} \cdot P_{rel.m}(T_{req.}) \cdot C_{tim\ m} \cdot P_{compl..m} \cdot P_{actual..m} \cdot P_{check\ m} \cdot P_{process.m} \cdot P_{infl.\ m}(T_{req.}) \cdot$$
$$\cdot P_{man\ m}(T_{req.}) \cdot P_{prot.m} \cdot P_{conf.m}(T_{req.}), \tag{4}$$

where $P_{combined.m}$ is calculated by the model (1), and all the others measures are calculated by the models, proposed in subsection *III.II.*

For material outputs the result of calculation P_m by (1) means probability that m-th function is executed with admissible quality. For information outputs the results of calculation P_m by (3) means probability that m-th function is executed with admissible quality, i.e. requested information outputs are reliable and timely produced, are complete, valid and, if needed, confidential for the purpose use. For output**s, combined from material and information products,** the results of calculation P_m by (4) means probability that the m-th function is executed with admissible quality according to material outputs and the requested information outputs are reliable and timely produced, are complete, valid and, if needed, confidential for the purpose use.

For calculation a relative **part of** functions executed with admissible operation quality for the past compared conditions during the given prognostic period the next formula may be used:

$$S_{quality} = \sum_{m=1}^{M} U_{real\ m} \cdot a_{real\ m} \bigg/ \sum_{m=1}^{M} a_{real\ m} \tag{5}$$

where $a_{real\ m}$ is real frequency of m-th type function execution (or considered as real according to assumption) during the given prognostic period;

$U_{real\ m}$ is real **part of** functions executed with admissible operation quality (it is measured from 0 to 1).

If the definition of U_m in (5) is a problem, the formula (2) may be used for the conditions (different or identical to past conditions) during the same given time, for this case $U_m = P_m$.

V. Estimation of expenses in life cycle

If inadmissible system operation quality is not considered the next formula is proposed for the estimations of expected expenses [13]:

$$C_{exp.}(t) = C_{instal.} + C_{main.}(t) + t\sum_{m=1}^{M} a_{hyp.\ m} C_{hyp.\ m} \quad for\ t \le T_{life},$$

and $$C_{exp.}(t) = C_{instal.} + C_{main.}(t) + T_{life}\sum_{m=1}^{M} a_{hyp.\ m} C_{hyp.\ m} + C_{disposal} \quad for\ t > T_{life} \tag{6}$$

where $a_{hyp.m}$ is a frequency of m-th function execution;

the expected or real **costs are indicated:**

$C_{instal.}$ - for **system development and installation;**

$C_{main.}(t)$ - **for system maintenance during time t;**

$C_{hyp.\ m}$ - for **system operation in time unit for** m-th function execution **(for example, in a year if time t is expressed in years);**

$C_{disposal}$ - for **system disposal;**

***T_{life}* - system life time.**

Considering inadmissible system operation quality for the moment t (i.e. mathematical expectation of expenses) the next formula is proposed for the estimations of expenses:

$$C_{math.}(t) = C_{instal.} + C_{main.}(t) + \sum_{m=1}^{M}\left(a_{hyp\ m}C_{hyp\ m}\cdot t\cdot P_m + D_m\cdot N_m(t)\cdot(1-P_m)\right)\ for\ t\le T_{life},$$

$$C_{math.}(t) = C_{instal.} + C_{main.}(t) + \sum_{m=1}^{M}\left(a_{hyp\ m}C_{hyp\ m}\cdot T_{life}\cdot P_m + D_m\cdot N_m(T_{life})\cdot(1-P_m)\right) + $$

$$+ C_{disposal}\quad for\ t > T_{life}, \tag{7}$$

where the probability of m-th function execution with admissible system operation quality (P_m) is calculated by (1), (3)-(4) in dependence on the chosen model for m-th function;

D_m - a possible or real damage for inadmissible system operation quality of system for one loss of quality**;**

$N_m(t)$ - a prognostic number of damages **from installation to moment t.**

VI. Estimation of the relative parts of system operation satisfaction connected with quality and costs

Let two systems are compared.

The next formula is proposed for calculation a relative part of system operation satisfaction connected with quality:

$$S_{quality} = (S_{quality\ 1}/\ S_{quality\ 2})\cdot 100\%, \tag{8}$$

where the relative **part of** functions executed with admissible operation quality for 1-st system $S_{quality\ 1}$ and for the 2-nd compared system during the given prognostic period $S_{quality\ 2}$ are calculated by (2) or (5).

For calculation a relative part of system operation satisfaction connected with costs for one system the next formula is proposed:

$$S_{cost}(t) = [C_{math.}(t) / C_{exp.}(t)]\cdot 100\%, \tag{9}$$

where the expected expenses for satisfying quality requirements, if inadmissible system operation quality is not considered, is calculated by (6). The mathematical expectation of expenses for satisfying quality requirements, considering inadmissible system operation quality, for the moment t is calculated by (7).

For a preferability of the 1-st system in comparison with the 2-nd system the relative part of system operation satisfaction connected with quality $S_{quality}$ should be more 100%. And for system operation satisfaction connected with costs a relative part $S_{cost}(t)$ should be less than 100%.

Some parts from the described **methodological approach** for probabilistic estimations are supported by the different versions of software Complex for Evaluation of Information Systems Operation Quality (CEISOQ+, registered by Rospatent №2000610272) and the software tools "Mathematical modeling of system life cycle processes" – "know how" (registered by Rospatent №2004610858) [1, 3, 5].

VII. Examples

Example 1. This example summarizes the numerous calculation results in applications the approach to intellectual systems of government agencies, manufacturing structures (including power generation, oil-and-gas systems), emergency services etc. [1, 3, 5-6, 8-15]. The typical estimations of measures for information outputs are presented by Table 1.

Table 1: *Knowledge from the best practice*

The typical estimations of measures for information outputs quality	
Model tittle	Limits for measure value
The model of system operation by a complex system in conditions of unreliability of its components	$P_{rel.}$ - no less than 0.99 P_{man} - no less than 0.95
The models complex of requests processing for the different dispatcher technologies	C_{tim} - no less than 0.95
The model of entering into system current data concerning new objects of application domain	$P_{compl.}$ - no less than 0.9
The model of information gathering	$P_{actual.}$ - no less than 0.9
The model of information analysis	P_{check} - no less than 0.97 $P_{process.}$ - no less than 0.95
The models complex of an authorized access to system resources	$P_{prot.}$ - no less than 0.99 $P_{conf.}$ - no less than 0.999
The models complex of dangerous influences on a protected system	$P_{infl.}$ - no less than 0.95

These estimations are confirmed also by statistical measuring data of the operation quality of the real monitored objects of dangerous manufacturing [8, 13].

Example 2. Suppose a special intellectual control system (SICS) is planning to create for monitoring intelligent manufacturing. Considering results of example 1 there is estimating a SICS operation during its life cycle. For simplification two types of function are to be executed. An output of each function is combination of **material and information products. It means example is focused on using basic formula (4). A**ssumption is: to simplify this example we use only calculation measure for material output and identical "The model of information gathering", which is similar mathematically to the model *III.I* for information output. Another measures are constant on the level of Table 1 values.

So, whole set of functions is divided into 2 types - with more urgent (m=1) and less urgent (m=2, M=2) execution:

- for 1st type of functions (m=1) significant changes concerning information outputs produced for users occur once a month (ξ=1 month). The gathering, preparation and checking of data for entry into SICS b=2 hours, the system update of data after checking occurs once a day (q=1 day);

- for 2nd type of functions (m=2) – significant changes concerning material and information outputs relating to any of served users, also occur once a month (ξ=1month), the gathering, preparation and checking of outputs b=3 days, system update after checking occurs once a week (q=1 week).

At the stage of development this SICS was constructed on the assumptions that each year is about 20 million requests for execution of the 1st type function ($a_{hyp.\ 1}$) and 80 million type 2 requests ($a_{hyp.\ 2}$). An expense about 705 million of cost conditional units (c.u.) during development and 5 years of operation is considered for satisfying quality requirements.

However, according to the real results of the 1st year of operation, the number of requests of the 1st type ($a_{real\ 1}$) amounted to 10 million, while the number of requests of the 2nd type ($a_{real\ 2}$) is 190 million. At the same time, the frequency of significant changes relating to any of served users doubled to 2 times a month.

The comparisons in advance are needed.

At the development stage a prognostic degree of satisfaction of quality is due to be estimated. It can allow to develop rational technical solutions. At the beginning of SICS operation, as data are gathered, a prediction of prognostic satisfaction needs to take a reasonable improvement of the SICS operation. What about the results of the probabilistic estimations?

The next additional data are used for input definition: $C_{instal.}$=200 m c.u.; $C_{main.}$ (1 year) =1 m c.u.; $C_{\exp.m}$ = 1 c.u. for 1 request; N_1(1 year)=0, N_2(1 year)=0.01%, D_2=10.000 c.u.

The results of modeling by formulas (1), (3),(4) have presented by Figure 4 - the probabilities of function execution with admissible quality are:

- on development stage P_1=0.98, P_2=0.81;
- on operation stage P_1=0.96, P_2=0.67.

Evaluations of the relative hypothetical and the real **parts of** system operation connected with quality and costs satisfaction are the next - a relative prognostic **part of** functions executed with admissible quality:

on development stage is $S_{quality}$ = 0.844. It means 84.4% requests are satisfied;

after 1-st year of SICS operation S_{real} = 0.685, it means 68.5% requests are satisfied considering changes in requests flows.

On development stage, when inadmissible system operation quality is not considered, an prognostic expenses for satisfying quality $C_{exp.}$ (5 years) = 705 m c.u., on operation stage considering inadmissible system operation quality $C_{math.}$ (5 years) = 998 m c.u.

A relative part of system operation satisfaction connected with quality $S_{quality}$ =81.2%, with costs S_{cost} (t)=141.6%.

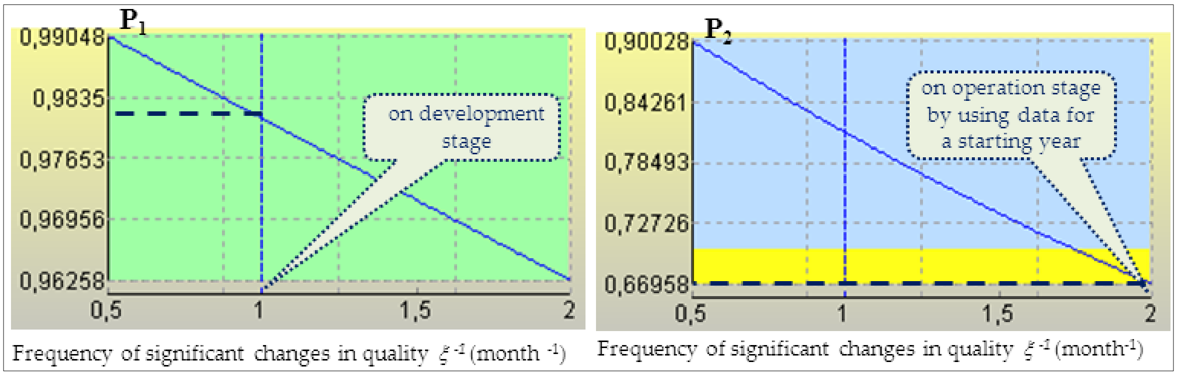


Figure 4: *Probability of 1-st and 2-nd type function execution with admissible system operation quality in dependence on frequency of significant changes in quality (ξ^{-1}, times in a month)*

If on development stage the level of 84.4% for a relative prognostic **part of** executed functions is acceptable, the level 68.5% (of requests executed with admissible quality) on operation stage means only 81.2% from accepted level. Moreover this result is achieved by the cost 41.6% over the admissible level. Such efficiency can't be estimated as satisfied for analyzed SICS.

Example 3. What about the comparable pragmatic effects? Authors of this article took part in creation of the Complex of supporting technogenic safety on the systems of oil&gas transportation and distribution. This Complex has been awarded by the Government of the Russian Federation in the field of a science and technics for 2014. Here peripheral posts are equipped additionally by means of monitoring operator actions to feel vibration, a fire, the flooding, unauthorized access, hurricane, and also intellectual means of the reaction in time, capable to recognize, identify and predict a development of extreme situations – see engineering decisions on Figure 5. Applications of Complex for 200 systems in several regions of Russia during the period 2009-2014 have already provided no accidents and economy about 8,5 Billions of Roubles. The economy is reached at the expense of effective implementation of adequate probabilistic modelling, risks prediction, justification of preventive measures against risks, processes optimization [10].

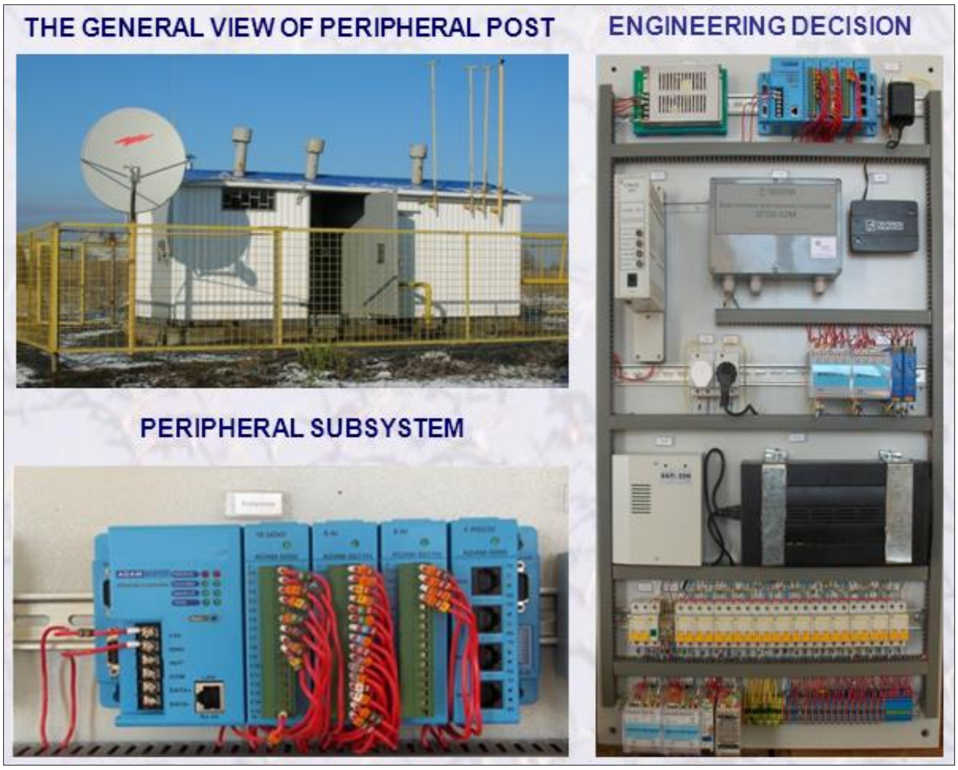


Figure 5: *Some elements of the Complex of technogenic safety on the systems of oil&gas transportation and distribution*

Conclusion

The proposed probabilistic **approach allows to** compare **different systems for an one operation time period or for different time periods with identical duration or the same system for different time periods on time line.** The approach can be useful for analysis and comparisons of systems operation quality, system optimization, to rationale of quantitative system requirements and engineering solutions for user satisfaction. The efficiency from implementation in life cycle of complex system is commensurable with expenses for system creation.

References

[1] Kostogryzov, A. and Nistratov, G. Standardization, mathematical modelling, rational management and certification in the field of system and software engineering. Armament.Policy.Conversion, Moscow, 2004.

[2] Zio, En. An Introduction to the Basics of **Reliability and Risk Analysis**, World Scientific Publishing Co.Pte.Ltd., 2006.

[3] Kostogryzov, A.I. and Stepanov, P.V. Innovative management of quality and risks in systems life cycle. Moscow, Armament.Policy.Conversion, Moscow, 2008.

[4] **Kolowrocki, K. and Soszynska-Budny, J. Reliability and Safety of Complex Technical Systems and Processes, Springer-Verlag London Ltd., 2011.**

[5] Kostogryzov, A., Nistratov, G. and Nistratov, A. Some Applicable Methods to Analyze and Optimize System Processes in Quality Management. Total Quality Management and Six Sigma, InTech, 2012: 127-196.

[6] Akimov, V., Kostogryzov, A., Mahutov N. at al. Security of Russia. Legal, Manufacturing&Economic and Scientific&Engineering Aspects. The Scientific Foundations of Technogenic Safety. Under the editorship of Mahutov N.A. Znanie, Moscow, 2015.

[7] Eid, M. and Rosato, V. Critical Infrastructure Disruption Scenarios Analyses via Simulation. Managing the Complexity of Critical Infrastructures. A Modelling and Simulation Approach, SpringerOpen, 2016: 43-62.

[8] Artemyev, V., Kostogryzov, A., Rudenko, Ju., Kurpatov, O., Nistratov, G., Nistratov, A. (2017). Probabilistic methods of estimating the mean residual time before the next parameters abnormalities for monitored critical systems. In: Proceedings of the 2nd International Conference on System Reliability and Safety (ICSRS), Milan, Italy, 2017: 368-373.

[9] Kostogryzov, A., Atakishchev, O., Stepanov, P., Nistratov, A., Nistratov, G., Grigoriev, L. (2017) Probabilistic modelling processes of mutual monitoring operators actions for transport systems. In: Proceedings of the 4th International Conference on Transportation Information and Safety (ICTIS), Canada, Banff, 2017: 865-871.

[10] Kostogryzov, A., Grigoriev, L., Golovin, S., Nistratov, A., Nistratov, G., Klimov, S. (2018) Probabilistic Modeling of Robotic and Automated Systems Operating in Cosmic Space. In: Proceedings of the International Conference on Communication, Network and Artificial Intelligence (CNAI), Beijing, China. DEStech Publications, Inc., 2018: 298-303.

[11] Artemyev, V., Rudenko, Ju., Nistratov, G. Probabilistic modeling in system engineering. Probabilistic methods and technologies of risks prediction and rationale of preventive measures by using "smart systems". Applications to coal branch for increasing Industrial safety of enterprises, IntechOpen, 2018: 23-51.

[12] Kershenbaum, V., Grigoriev, L., P. Kanygin, Nistratov, A. Probabilistic modeling in system engineering. Probabilistic modeling processes for oil and gas systems, IntechOpen, 2018: 55-79.

[13] Kostogryzov, A., Panov, V., Grigoriev, L., Zubarev, I., Nistratov, A., Nistratov, G. (2019) Estimation of stakeholders satisfaction in application to socially significant systems. In: Proceedings of the Proceedings of the 6th International Conference Actual Problems of System and Software Engineering, Moscow, Russia, 12-14 November, 2019: 10-16.

[14] Kostogryzov, A., Korolev, V. Probabilistic Methods for Cognitive Solving of Some Problems in Artificial Intelligence Systems, IntechOpen, 2019. Doi: http://dx.doi.org/10.5772/intechopen.89168

ISSN 1932-2321