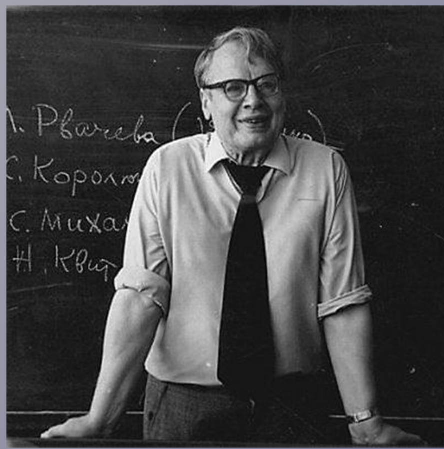
ELECTRONIC JOURNAL
OF INTERNATIONAL GROUP
ON RELIABILITY

Gnedenko Forum Publications



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

•

ISSN 1932-2321

•

VOL.9 NO.2 (33)
JUNE, 2014

RELIABILITY:
THEORY&APPLICATIONS

RELIABILITY
RISK ANALYSIS
MAINTENANCE
SAFETY

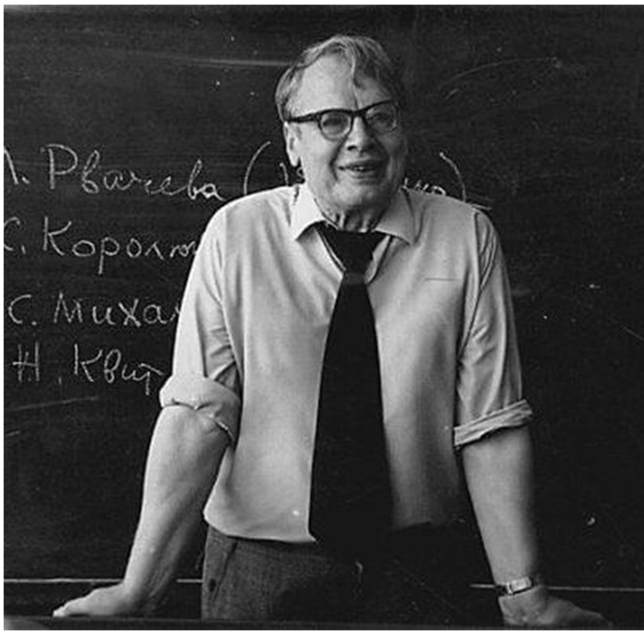
Vol.9
No.2 (33)
•
June
2014

San Diego

ISSN 1932-2321

http://www.gnedenko-forum.org/Journal/index.htm

Journal of International Group on Reliability



RELIABILITY:
THEORY & APPLICATIONS

Vol.9 No.2 (33),
June, 2014

San Diego
2014

Send your paper

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 MSWord format. (Times New Roman, 12 pt , 1.5 intervals).
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 to use professional linguistic proofs before sending a paper to the journal.

* * *

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

the Editor-in-Chief ,
Igor Ushakov
igusha22@gmail.com

or

the Deputy Editor,
Alexander Bochkov
a.bochkov@gmail.com

Table of Contents

Numerical experiments with biological networks are made and hubs loads are analysed. These experiments are based on algorithms of oriented graphs factorization. Numerical results allow to find narrow places in the networks.

The article describes a mathematical model of reliability of the information system security systems consisting of object information protection and security of the two systems. The functioning process of the studied information security system is described as a superposition of alternating renewal processes. Upper and lower bounds for the expectation of time before unauthorized access to information.

Estimation of parameters of individual reliability demands to estimate expediency of classification of all population of statistical data on the set versions of attributes. The method and algorithm optimum number of attributes classification of data is developed. Application allows calculate optimum on accuracy of an estimation of parameters of individual reliability, to reveal significant versions of attributes, to pass at the account of reliability of the concrete equipment with qualitative to a quantitative level.

The estimation of latent faults probability is a key indicator for quantitative assessment of critical systems reliability and safety. The article presents the method of latent faults probability estimating for critical software with a required (controlled) trustworthiness. Compositional set-theoretical model of residual and latent faults is proposed for framework assessment of latent faults. The procedure of the experimental calibration of faults sensitivity and diversity degree of inspection methods of source software faultlessness is presented. The "dotted" injections method using faults profile for the level of programming languages is proposed.

The AHP-method proved its efficiency in different situations of decision making under conditions when experts can perform pairwise comparisons of compared objects under study. But, our real life consists of uncertainties. How can we compare the objects if information about them is not available or extremely uncertain? Pairwise comparisons are impossible. In this case, we propose to use the approach of minimizing the functional «errors» in the evaluation of

ANALYSIS OF HUBS LOADS IN BIOLOGICAL NETWORKS

***G.Sh. Tsitsiashvili, **A.S. Losev, ***M.A. Osipova, ****Yu.N. Kharchenko**

IAM, FEB RAS, Vladivostok, Russia,
FAR Eastern Federal University
•
e-mails: *guram@iam.dvo.ru , **alexax@bk.ru,
***mao1975@list.ru, ****har@iam.dvo.ru.

ABSTRACT

Numerical experiments with biological networks are made and hubs loads are analysed. These experiments are based on algorithms of oriented graphs factorization. Numerical results allow to find narrow places in the networks.

INTRODUCTION

In articles [1], [2] we consider oriented graph with finite sets of vertices and edges. A sequential algorithm of graph factorization with square number of arithmetical operations by a number of graph vertices is used. A decrease of calculation complexity in a comparison with Boolean multiplication of contiguity matrixes is connected with an introduction of partial order matrix which is defined completely by assignment operations.

But a problem is to make numerical experiments, to analyze possibilities of suggested algorithm and to consider interesting protein networks with large numbers of nodes. For this aim we take protein network of Arabidopsis and analyze not only calculation complexity of the algorithm but find some new properties of considered network.

1. PRELIMINARIES

Consider oriented graph G with finite sets of vertices V and edges E. Say that two vertices $v_1, v_2 \in V$ of oriented graph belong to binary relation $v_1 \sim v_2$ if in this graph there is a cycle which contains both vertices. Call the oriented graph $[G]$ with the set of vertices $[V]$ and the set of edges $[E] = \{([v_1],[v_2]) : \exists v_1' \in [v_1], v_2' \in [v_2], (v_1', v_2') \in E\}$ the factor graph. It is obvious that $[G]$ is acyclic graph.

On the set $[V]$ of the graph $[G]$ vertices define the binary relation "$\succeq$":$[v_1] \succeq [v_2]$ if in the graph $[G]$ there is a way from the cluster $[v_1]$ to the cluster $[v_2]$. It is obvious that if $[v_1] \succeq [v_2]$ then in the graph $[G]$ there is a way from any vertex of the cluster $[v_1]$ to any vertex of the cluster $[v_2]$.

Describe now an algorithm of a factorization of the graph $[G]$ vertices and a construction of partial order "$\succeq$". Assume that the set $V = \{1,...,n\}$ consists of n vertices and denote $t = \min\left(f : 2^f \geq n\right) = [\log_2 n] + 1$ and put $A = \|a_{ij}\|_{i,j=1}^{n}$ the contiguity matrix of the graph G. Construct a sequence of zero-one matrixes $A(k) = \|a_{jr}(k)\|_{j,r=1}^{n}$ by the recurrent relation

$$a_{jr}(1) = \max_{1 \leq s \leq n} \min\left(a_{js}, a_{sr}\right),\ a_{jr}(k+1) = \max_{1 \leq s \leq n} \min\left(a_{js}(k), a_{sr}(k)\right),\ 1 \leq k < t.$$

It is obvious that if $a_{jr}(t)=1$ then in the graph G there is a way from the vertex j to the vertex r and $j' \in [j] \Leftrightarrow a_{jj'}(t)=a_{j'j}(t)=1$. The relation $[i] \succeq [j] \Leftrightarrow \exists i' \in [i], j' \in [j] : a_{i'j'}(t)=1$.

Cubic complexity of this algorithm is strong restriction for numerical experiments with oriented graphs occurred in an analysis of protein networks. So it is naturally to decrease calculation complexity.

Describe sequential algorithm of a construction of clusters and the matrix of partial order between clusters. On the step 1 there is the single vertex 1 which creates the cluster [1] and the set of clusters $K=\{[1]\}$. Introduce the matrix $a=\left\|a([p],[q])\right\|_{[p],[q]\in K}$ which characterizes the partial order "$\succeq$": $a([p],[q])=1$, if $[p]\succeq[q]$ and in opposite case $a([p],[q])=0$.

On the step 1 put $a([1],[1])=1$. Assume that on the step t there is the clusters set K and the matrix a. These clusters create a division of the set $V=\{1,...,n\}$ into non intersected subsets, each cluster $[k]\in K$ is indexed by maximal number $k\in V$ of its vertices.

On the step $t+1$ new vertex $t+1$ appears. It is connected with edges which come into vertices from the set $P\subseteq V$ and with vertices which come into the vertex $t+1$ from the vertices from the set $Q\subseteq V$. Each vertex from the set P (the set Q) comes into some cluster. Denote P the sets of clusters corresponding to the sets of vertices from P and define

$$K_{[p]}=\{[k]\in K: a([p],[k])=1\},\ [p]\in P,$$

$$K_{[p]}=\{[k]\in K: a([k],[q])=1\},\ [q]\in Q,\ A=\left(\bigcup_{[p]\in[P]} K_{[p]}\right)\bigcap\left(\bigcup_{[q]\in[Q]} K_{[q]}\right),$$

$$A_1=\left(\bigcup_{[p]\in[P]} K_{[p]}\right)\setminus A,\ A_2=\left(\bigcup_{[q]\in[Q]} K_{[q]}\right)\setminus A,\ B=K\setminus(A\cup A_1\cup A_2).$$

New vertex $t+1$ and clusters from the set A create new cluster

$$[t+1]:=\{t+1\}\cup A,\ \ K:=(K\setminus A)\cup\{[t+1]\}.$$

A recalculation of the matrix a on renewed set of clusters K is following [2]:

$a([t+1],[i]):=1$, $[i]\in A_1\cup\{[t+1]\}$, $a([i],[j]):=1$, $[i]\in A_2$, $[j]\in A_1\cup\{[t+1]\}$,

$a([i],[j]):=0$, $[i]\in A_1$, $[j]\in A_2\cup\{[t+1]\}\cup B$,

$a([i],[j]):=0$, $j\in A_2$, $[i]\in B\cup\{[t+1]\}$, $a([t+1],[i]):=0$, $a([i],[t+1]):=0$, $[i]\in B$.

All other meanings of the matrix a elements coincide with previous ones on the step t. This algorithm has square complexity (a number of arithmetical operations) by a number of graph vertices. All other operations are assignment operations.

2 NUMERICAL EXPERIMENT AND ITS INTERPRETATION

Comparison of algorithms efficiency. To compare suggested factorization algorithms we consider the protein network Arabidopsis with 2824 vertices and 7570 edges [3]. Using the server on a base of 2 processors Xeon with 6 kernels (each of them), the frequency 2300 Hz and 32 GB of working memory it is possible to realize the factorization procedure by the calculation of the matrixes $A(k),\ 1\le k<t,$ during 21 days. Using the sequential algorithm it is possible to fulfill the factorization procedure on the notebook with 2 kernels processor I3, the frequency 2300 Hz, and 4 GB of working memory during one and half hours.

Detection of network kernel. Results of the factorization of the network vertices are represented in Table 1 which allows to represent this protein network essentially more compact and to obtain very contrast distribution of clusters by numbers of their vertices.

numbers of cluster vertices	numbers of unisolated clusters	numbers of isolated clusters
1	1429	37
2	41	26
3	17	11
4	11	6
5	5	0
6	1	0
7	1	1
8	3	0
10	1	0
11	1	0
16	1	0
958	1	0

Table 1. Factorization of Arabidopsis protein network.

Table 1 shows that Arabidopsis protein network has a kernel with 958 vertices which composes 34 percents of all network vertices.

Analysis of kernel hubs. It is interesting to consider an interaction of kernel vertices between them and with vertices out of the kernel. Following [4], [5] we concentrate our attention on vertices which are incident with maximal numbers of edges - so called hubs. Using results of the factorization we find 10 largest numbers of edges which 1) log out from within to outside; 2) come from outside to inside; 3) log out from within to inside; 4) come from within to inside. Hubs from the set 1 may be called sources, hubs from the set 2 - sinks, hubs from the sets 3, 4 - dispatchers of the kernel. Table 2 shows that hubs from the set 1 have largest powers, hubs from the set 2 have smallest powers and hubs from the sets 3, 4 have average powers. So hubs - sources have maximal load and consequently restrict working intensity of the kernel work.

set 1	set 2	set 3	set 4
from within outside	from outside inside	from within inside	from within inside
102	23	47	50
70	19	47	39
70	16	43	38
69	14	43	37
63	12	39	33
55	8	38	27
54	7	36	27
43	6	35	26
41	5	35	24
36	4	32	24

Table 2. Numbers of edges correspond to different hubs.

Role of hubs in counteraction of kernel with its environment. Calculations show that sixty percent of edges emergent from the kernel are incident with vertices (hubs) which compose 7 percents of all appropriate kernel vertices. And sixty percents of all other types of edges are incident with kernel vertices which compose about 30 percents of all kernel vertices. This remark shows a role of hubs - sources in the considered network.

REFERENCES

1. Tsitsiashvili G.Sh. 2013. Sequential algorithms of graph nodes factorization. *Reliability: Theory and Applications.* Vol. 8(4). P. 30-33.
2. Tsitsiashvili G.Sh., Kharchenko Yu.N., Losev A.S. and Osipova M.A. 2014. Enumeration of Minimal Control Sets of Vertices in Oriented Graph. *Applied Mathematical Sciences.* Vol. 8, no. 39. P. 1941 - 1945.
3. Mingzhi Lin, Xi Zhou, Xueling Shen, Chuanzao Mao, Xin Chen. The predicted Arabidopsis Inteactome Resource and Network Topology-Based Systems Biology Analyses// Thwe Plant Cell. 2011. Vol. 23. P. 911-922.
4. Liu Y.Y. , Slotine J.J., Barabasi A.L. 2011. Controllability of complex networks. *Nature.* Vol. 473. P. 167-173.
5. Barabasi L.A., Albert R.1999. Emergence of scaling in random networks. *Science.* Vol. 286. P. 509-512.

MATHEMATICAL MODEL OF THE RELIABILITY OF INFORMATION PROTECTION WITH LAYERED SECURITY SYSTEM. INTRODUCTION

A. I. Pereguda

•

Obninsk Institute for Nuclear Power
Engineering, Obninsk, Russia

e-mail: pereguda@iate.obninsk.ru

ABSTRACT

The article describes a mathematical model of reliability of the information system security systems consisting of object information protection and security of the two systems. The functioning process of the studied information security system is described as a superposition of alternating renewal processes. Upper and lower bounds for the expectation of time before unauthorized access to information.

Keywords: reliability, security system, random variables, time to failure, random process, the expectation of the time, distribution function.

1. INTRODUCTION

When building a highly reliable security systems often used layered threat protection. To protect data from unauthorized access is required to provide a certain level of reliability of such systems, including the reliability of their hardware and software. When analyzing the reliability of such systems is not enough just to take into account the structure of the security system. Mathematical models of reliability of information protection systems considered in various works, such as [1, 2]. It is also necessary to take into account the reserve time available in security systems, which arises from the fact that to overcome the successive layers of protection attacker will take some time .Here we obtain relations to assess how this extra time is a reliability of information security.

2. DESCRIPTION OF THE MODEL

Let an information system which consists of a protection and safety of the two systems, each of which saves one of the varieties to attack protection. Security systems are not foolproof and may refuse. It is assumed that the failures of safety systems are independent. Failure detection security systems occurs only during periodic monitoring her condition. After failure detection security system performed its full restoration to its original state. If the first security system was not immediately able to parry, the unauthorized access to information can only be achieved after a while though, and only provided that during this time, none of the safety systems failed.

Denote χ_1- a random time before the first attack on the object of protection, detectable first security system. Random time reflection of the first attack on the object of protection, the detected first security system, denoted γ_1. Then a random time before the second attack on the object of protection after the first attack, the reflection of the first security system detected a protection object denoted χ_2, after reflection the second attack - χ_3 etc. We assume that all χ_i, $i = 1, 2, \ldots$, as usual, are independent and identically distributed with the distribution function $F_\chi(t) = P(\chi \le t)$.

Random time of reflection on the i attack to protect, detect the first security system will be denoted γ_i, $i = 1, 2, \ldots$,, as to which assume that - γ_i, $i = 1, 2, \ldots$ independent identically distributed random variables with distribution function $F_\gamma(t) = P(\gamma \le t)$.Denote δ_i a random time between the i attack on the object of protection, which was not detected the first security system, and unauthorized access. It is obvious that δ_i, $i = 1, 2, \ldots$, - independent identically distributed random variables with distribution function $F_\delta(t) = P\{\delta \le t\}$. Thus, the first security system can still prevent unauthorized access time interval $[\chi_i, \chi_i + \delta_i)$

Since the process of operation of the first security system consists of cycles "attack - a reflection attack", and, consequently, the moments of the completion of attack to protect, detect the first security system are regenerative information system.

If the first security system and has not worked in the specified time interval, the second security system can still prevent unauthorized access at the moment $\chi_i + \delta_i$. The duration of the second reflection attack security system after the first attack is denoted α_1, after the second attack - α_2, etc. Believe that α_i, $i = 1, 2, \ldots$- independent identically distributed random variables with distribution function $F_\alpha(t) = P\{\alpha \le t\}$.

Random time before the i attack on the object of protection, security system detects a second denote φ_i, $i = 1, 2, \ldots$.Let φ_i, $i = 1, 2, \ldots$- independent identically distributed random variables with distribution function $F_\varphi(t) = P\{\varphi \le t\}$. Duration reflection i-attack on the object of protection, which was discovered a second security system, denoted ψ_i, $i = 1, 2, \ldots$. Believe that ψ_i, $i = 1, 2, \ldots$ - independent identically distributed random variables with distribution function $F_\psi(t) = P\{\psi \le t\}$. Assume that the end points of reflection of such attacks are regenerative process of information system. Moreover, let the expectations χ_i, $i = 1, 2, \ldots$, γ_i, $i = 1, 2, \ldots$, δ_i, $i = 1, 2, \ldots$, α_i, $i = 1, 2, \ldots$, φ_i, $i = 1, 2, \ldots$and ψ_i, $i = 1, 2, \ldots$, exist and are finite .

Since the process of reflection attacks consists of cycles "attack - a reflection attack", and, consequently, the moments of the completion of the attack on the object of protection, first detected as a security system, and the second system security are regenerative information system.

We now consider the processes of safety systems. Let $\xi_i^{(1)}$, $i = 1, 2, \ldots$ for the i time between failure of the first security system. We assume that the time to failure of the first security system $\xi_i^{(1)}$, $i = 1, 2, \ldots$ - independent identically distributed random variables with distribution function $F_{\xi_i^{(1)}}(t) = P\{\xi_i^{(1)} \le t\}$. Duration of the first system restore security after the i-failure is denoted $\eta^{(1)}$, $i = 1, 2, \ldots$ Here $\eta^{(1)}$, $i = 1, 2, \ldots$– independent identically distributed random variables with distribution function $F_{\eta_i^{(1)}}(t) = P\{\eta_i^{(1)} \le t\}$. Fault detection security systems occurs only during periodic condition monitoring. So serviceability first security system update to the period $T^{(1)}$ and the duration of the periodic control requires time equal $\theta^{(1)}$.

If the first security system (SS) functioned properly random time $\xi^{(1)}$, for this time $\left[\frac{\xi^{(1)}}{T^{(1)} + \theta^{(1)}}\right]$ period was performed preventive control, and during these periods the first SS was operational $T^{(1)}\left[\frac{\xi^{(1)}}{T^{(1)} + \theta^{(1)}}\right]$ units of time. Between the last before giving up control prevention, which occurred at the time $(T^{(1)}+\theta^{(1)})\left[\dfrac{\xi^{(1)}}{T^{(1)}+\theta^{(1)}}\right]$, and the denial of time $\xi^{(1)}$ is still time $\xi^{(1)}-(T^{(1)}+\theta^{(1)})\left[\dfrac{\xi^{(1)}}{T^{(1)}+\theta^{(1)}}\right]$, during which the first SS is still OK.

On the last loop prevention control security system is idle time $(T^{(1)}+\theta^{(1)})-(T^{(1)}+\theta^{(1)})\left\{\dfrac{\xi^{(1)}}{T^{(1)}+\theta^{(1)}}\right\}$, because rejection has occurred in its time $\xi^{(1)}$, but was detected after the cycle test prophylaxis. Failure of the first security system will be detected in time $(T^{(1)}+\theta^{(1)})\left(\left[\dfrac{\xi^{(1)}}{T^{(1)}+\theta^{(1)}}\right]+1\right)$, and after the repair work lasting $\eta^{(1)}$ SS starts functioning correctly again.

Believe $\xi_i^{(2)}$, $i=1,2,\ldots$ – time between i-failure of the second-security, where $\xi_i^{(2)}$, $i=1,2,\ldots$ are independent and identically distributed random variables with distribution function $F_{\xi_i^{(2)}}(t)=P\{\xi_i^{(2)}\le t\}$. Duration of recovery the second security system after i-failure denote $\eta_i^{(2)}$, $i=1,2,\ldots$, where $\eta_i^{(2)}$, $i=1,2,\ldots$

then independent identically distributed random variables with distribution function $F_{\eta_i^{(2)}}(t)=P\{\eta_i^{(2)}\le t\}$.

Monitor the status of the second period the security system is denoted $T^{(2)}$, and the duration of its periodic monitoring – $\theta^{(2)}$. Failure of the second security system is detected in time $(T^{(2)}+\theta^{(2)})\left(\left[\dfrac{\xi^{(2)}}{T^{(2)}+\theta^{(2)}}\right]+1\right)$. After the repair work with duration equal to $\eta^{(2)}$ the second SS will again operate correctly. Thus, the processes of the functioning of the first and second security system to control prevention have alternating renewal process. Here we assume that during the control state security not perform their functions .

Moreover, let the expectations $\xi_i^{(1)}$, $i=1,2,\ldots$, $\eta_i^{(1)}$, $i=1,2,\ldots$, $\xi_i^{(2)}$, $i=1,2,\ldots$and $\eta_i^{(2)}$, $i=1,2,\ldots$ exist and are finite, and the distribution function of these random variables are not arithmetic.

Functioning processes of the first and second security systems appears as alternating intervals "work - restoration", which consist of independent, identically distributed random variables $\{\xi_i^{(1)}, i\ge 1\}$, $\{\eta_i^{(1)},\ i\ge 1\}$и$\{\xi_i^{(2)}, i\ge 1\}$, $\{\eta_i^{(2)},\ i\ge 1\}$ forming two alternating renewal process $\{(\xi_i^{(1)},\eta_i^{(1)}),\ i\ge 1\}$ and $\{(\xi_i^{(2)},\eta_i^{(2)}),\ i\ge 1\}$..

We introduce some notation and explanations of the process of preventive control, which will be further taken into account in the mathematical model. Through $[x]$ and $\{x\}$ denote the integral and fractional parts of a number x, $x^{+}=max(x,0), x\wedge a=min(x,a)$, $J_{x<a}$ — indicator of the event $x<a$.

Random time to unauthorized access denoted ω. Our task is to construct a mathematical model of reliability of information security systems with layered security system and obtain an estimate of the average time $M\omega$ to unauthorized access to information.

3. MAIN RESULTS

Since the process of functioning of the system of protection of information is described as a superposition of alternating renewal processes, the random time to unauthorized access can be written as the sum of the following:

$$\omega = \sum_{i=1}^{\nu-1} \sigma_i + \sigma'_\nu , \qquad (1)$$

where the duration of the regeneration process, the information system, which has not happened unauthorized access, equal

$$\sigma_i = \chi_i \wedge \varphi_i + \left((\beta_i + \gamma_i) J_{B_i} + (\delta_i + \alpha_i) J_{\overline{B_i}} \right) J_{\chi_i \le \varphi_i} + \psi_i J_{\varphi_i < \chi_i} ,$$

and duration of the regeneration process, the information system, which occurred unauthorized access, equal

$$\sigma'_i = \chi_i \wedge \varphi_i + \delta_i J_{\chi_i \le \varphi_i} .$$

In relation σ_i use the following notation: β_i - the length of time between the implementation of an attack on the object of protection, which should be parried the first security system, and actuation of this safety system, provided that it was able to parry; B_i -the event consists in the fact that the regeneration of the i-cycle process, the information system for the implementation of attack corresponding to the first security system, the security system retorted this wave in the time interval $\left[\sum_{j=1}^{i-1} \sigma_i + \chi_i , \sum_{j=1}^{i-1} \sigma_i + \chi_i + \delta_i \right)$; $\overline{B_i}$ - denotes the opposite event - first security system is not countered this attack on the object of protection in that time interval. For this model, it is obvious $0 \le \beta_i < \delta_i$.

For the distribution function of time to unauthorized access, obviously, we have the following relationship:

$$F_\omega(t) = P(\omega \le t) = P\left(\sum_{i=1}^{\nu-1} \sigma_i + \sigma'_\nu \le t \right).$$

For further transformations of $F_\omega(t)$ we use the method of conditional probability distributions and express complex events through the conditional probability of this event under the appropriate conditions. If the conditions are incompatible and form a complete group of events, the absolute transformation of the distribution function $F_\delta(t)$ for the total probability formula

$$\widetilde{F}_\omega(s) = Me^{-s\omega} = \sum_{n=1}^{\infty} M\left(e^{-s\omega} \mid \nu = n \right) P(\nu = n) ,$$

where $\widetilde{F}_\omega(s) = \int_0^\infty e^{-st} dF_\omega(t) = E\left[e^{-s\omega} \right]$.

Since all χ_i and γ_i are independent random variables, given that $\widetilde{F}_\sigma(s) = Me^{-s\sigma}$ and $\widetilde{F}_{\sigma'}(s) = Me^{-s\sigma'}$, and the process in the regeneration, the $P(\nu = n) = q(1-q)^{n-1}$, where q - likelihood of unauthorized access to the regeneration process, the information system. Easy to see that

$$M\left(e^{-s\omega}\mid\nu=n\right)=M\left(e^{-s\left(\sum_{i=1}^{\nu-1}\sigma_i+\sigma'_\nu\right)}\middle|\nu=n\right)=\left(\widetilde{F}_\sigma(s)\right)^{n-1}\widetilde{F}_{\sigma'}(s).$$

Consequently, the Laplace-Stieltjes $F_\omega(t)$ rewrite

$$\widetilde{F}_\omega(s)=\sum_{n=1}^{\infty}\left(\widetilde{F}_\sigma(s)\right)^{n-1}\widetilde{F}_{\sigma'}(s)q(1-q)^{n-1}=\frac{q\widetilde{F}_{\sigma'}(s)}{1-(1-q)\widetilde{F}_\sigma(s)}.$$

Expression for the expectation of time before unauthorized access can be directly calculated from the ratio of (1) or by using the well-known relation $M\omega=-\left.\frac{d\widetilde{F}_\omega(s)}{ds}\right|_{s=0}$. We obtain

$$M\omega=M\sigma'+\frac{1-q}{q}M\sigma. \tag{2}$$

To find the expectation of use of the complex to the first accident is necessary to calculate separately the expectations $M\sigma'$ and $M\sigma$. In calculating the expectation $M\sigma'$, we consider that these random variables are mutually independent. Then, the mathematical expectation $M\sigma'$, we can write as follows:

$$M\sigma'=M((\chi\wedge\varphi)+\delta J_{\chi\le\varphi})=M(\chi\wedge\varphi)+M\delta P(\chi\le\varphi)=$$

$$=\int_0^\infty(1-F_\chi(t))(1-F_\varphi(t)dt+\int_0^\infty(1-F_\delta(t))dt\int_0^\infty F_\chi(t)dF_\varphi(t).$$

By analogy with $M\sigma'$ compute the expectation σ, which is equal to

$$M\sigma=M((\chi\wedge\varphi)+((\beta+\gamma)J_B+(\delta+\alpha)J_{\overline{B}})J_{\chi\le\varphi}+\psi J_{\chi>\varphi})=M(\chi\wedge\varphi)+$$

$$+M((\beta+\gamma)J_B+(\delta+\alpha)J_{\overline{B}})J_{\chi\le\varphi})+M(\psi J_{\chi>\varphi})=\int_0^\infty(1-F_\chi(t))(1-F_\varphi(t)dt+$$

$$+P(\chi\le\varphi)((M\beta+M\gamma)P(B)+(M\delta+M\alpha)P(\overline{B}))+M\psi P(\varphi<\chi).$$

Substituting the calculated expectations $M\sigma'$ and $M\sigma$ in (2) we have

$$M\omega=M(\chi\wedge\varphi)+\delta J_{\chi\le\varphi})+\frac{1-q}{q}M((\chi\wedge\varphi)+((\beta+\gamma)J_B+(\delta+\alpha)J_{\overline{B}})J_{\chi\le\varphi}+\psi J_{\chi>\varphi}).$$

Note that the record of the distribution function for the random variables β_i, which depend on the magnitude of σ_i, it is not possible, but it is possible to obtain upper and lower bounds for the expectation of time before unauthorized access to information by using the order relation on the set of distribution functions [3]. This assessment is written as follows:

$$M\sigma'+\frac{1-q}{q}\left(M(\chi\wedge\varphi)+\left(M\gamma P(B)+(M\delta+M\alpha)P(\overline{B})\right)P(\chi\le\varphi)\right)+M\psi P(\varphi<\chi)\le M\omega\le$$
$$\le M\sigma'+\frac{1-q}{q}\left(M(\chi\wedge\varphi)+\left((M\delta+M\gamma)P(B)+(M\delta+M\alpha)P(\overline{B})\right)P(\chi\le\varphi)\right)+M\psi P(\varphi<\chi). \tag{3}$$

In order to calculate the upper and lower bounds, you must first assess the probability q unauthorized access to the regeneration process of the information system. To get an estimate of the probability of a closer look at the processes of the safety systems [3,6].

Consider a single cycle of the functioning of the information system and compute $P(\overline{B})$, by considering two auxiliary random variables U_n and V_n ,, defined by the relations

$$U_n = \sum_{i=1}^{n} \xi_i^{(1)} + \sum_{i=1}^{n-1} \left((T^{(1)} + \theta^{(1)}) - \left\{ \frac{\xi_i^{(1)}}{T^{(1)} + \theta^{(1)}} \right\} (T^{(1)} + \theta^{(1)}) \right) + \sum_{i=1}^{n-1} \eta_i^{(1)},$$

$$V_n = \sum_{i=1}^{n} \xi_i^{(1)} + \sum_{i=1}^{n} \left((T^{(1)} + \theta^{(1)}) - \left\{ \frac{\xi_i^{(1)}}{T^{(1)} + \theta^{(1)}} \right\} (T^{(1)} + \theta^{(1)}) \right) + \sum_{i=1}^{n} \eta_i^{(1)},$$

Where U_n - the time of n-failure of the first security, and V_n - the end of the first security system recovery after an-failure.

Because the process is functioning security system is alternating renewal process, we can write

$$U_n = \sum_{i=1}^{n} \xi_i^{(1)} + \sum_{i=1}^{n-1} \left((T^{(1)} + \theta^{(1)}) - \left\{ \frac{\xi_i^{(1)}}{T^{(1)} + \theta^{(1)}} \right\} (T^{(1)} + \theta^{(1)}) \right) + \sum_{i=1}^{n-1} \eta_i^{(1)},$$

$$V_n = \sum_{i=1}^{n} \xi_i^{(1)} + \sum_{i=1}^{n} \left((T^{(1)} + \theta^{(1)}) - \left\{ \frac{\xi_i^{(1)}}{T^{(1)} + \theta^{(1)}} \right\} (T^{(1)} + \theta^{(1)}) \right) + \sum_{i=1}^{n} \eta_i^{(1)}.$$

Then unauthorized access to the cycles of regeneration process, the information system provided if

$$U_n \le \chi < V_n - \delta, \; \delta \le V_n - U_n,$$

or if

$$V_{n-1} + T^{(1)} \le \chi < V_{n-1} + T^{(1)} + \theta^{(1)} - \delta;$$

$$V_{n-1} + (T^{(1)} + \theta^{(1)}) + T^{p} \le \chi < V_{n-1} + 2(T^{p} + \theta^{p}) - \delta;$$

$$\ldots$$

$$V_{n-1} + \left(\left[\frac{\xi_n^{(1)}}{T^{(1)} + \theta^{(1)}} \right] - 1 \right) (T^{(1)} + \theta^{(1)}) + T^{(1)} \le \chi < V_{n-1} + \left[\frac{\xi_n^{(1)}}{T^{(1)} + \theta^{(1)}} \right] (T^{(1)} + \theta^{(1)}) - \delta;$$

$$\delta < \theta^{(1)}.$$

For further exposition must enter non-negative random variables

$$\varepsilon_n = (T^{(1)} + \theta^{(1)}) - \left\{ \frac{\xi_n^{(1)}}{T^{(1)} + \theta^{(1)}} \right\} (T^{(1)} + \theta^{(1)})$$

$$\Delta_n = \eta_n^{(1)} + \varepsilon_n - \delta, \; \zeta = \theta^{(1)} - \delta,$$

Here ε_n - the length of time from the moment of first refusal to security failure detection and recovery is started.

Taking into account the circumstances of the unauthorized access to information, write the expression for the probability of $P(\overline{B})$:

$$P(\overline{B}) = \sum_{n=1}^{\infty} \int_{0}^{\infty} M \left(J_{U_n \le x < V_n - \delta} J_{\Delta_n > 0} + \sum_{i=1}^{\left[\frac{\xi_n^{(1)}}{T^{(1)} + \theta^{(1)}} \right]} J_{V_{n-1} + (i-1)(T^{(1)} + \theta^{(1)}) \le x < V_{n-1} + i(T^{(1)} + \theta^{(1)}) - \delta} J_{\zeta > 0} \right) dF_{\chi}(x).$$

It follows that

$$P(\overline{B})=\sum_{n=1}^{\infty}\int_{0}^{\infty}P\big(U_n\wedge(U_n+\Delta_n)\le x\big)dF_{\chi}(x)-\sum_{n=1}^{\infty}\int_{0}^{\infty}P\big(U_n+\Delta_n\le x\big)dF_{\chi}(x)+$$

$$+\sum_{n=1}^{\infty}\int_{0}^{\infty}M\left(\sum_{i=1}^{\left[\frac{\xi_n^{(1)}}{T^{(1)}+\theta^{(1)}}\right]}P\big((V_{n-1}+i(T^{(1)}+\theta^{(1)})-\theta^{(1)})\wedge(V_{n-1}+i(T^{(1)}+\theta^{(1)})-\theta^{(1)}+\zeta)\le x\big)\right)dF_{\chi}(x)-$$

$$-\sum_{n=1}^{\infty}\int_{0}^{\infty}M\left(\sum_{i=1}^{\left[\frac{\xi_n^{(1)}}{T^{(1)}+\theta^{(1)}}\right]}P\big(V_{n-1}+i(T^{(1)}+\theta^{(1)})-\theta^{(1)}+\zeta\le x\big)\right)dF_{\chi}(x)=q_1+q_2.$$

Note that

$$q_1=\sum_{n=1}^{\infty}\int_{0}^{\infty}\int_{0}^{\infty}\Big((F_{\xi^{(1)}}*(F_{\xi^{(1)}}*F_{\eta^{(1)}}*F_{\varepsilon})^{*(n-1)})(x)-(F_{\xi^{(1)}}*(F_{\xi^{(1)}}*F_{\eta^{(1)}}*F_{\varepsilon})^{*(n-1)})(x-y)\Big)dF_{\Delta}(y)dF_{\chi}(x),$$

where $F_{\xi^{(1)}}*F_{\eta^{(1)}}(t)=\int_{0}^{t}F_{\xi^{(1)}}(t-z)dF_{\eta^{(1)}}(z)$ – convolution of the distribution functions $F_{\xi^{(1)}}(t)$ and $F_{\eta^{(1)}}(t)$, $F^{*(n)}(t)=F*F^{*(n-1)}(t)$ – n-fold convolution of functions $F(t)$.

Probability q_1

$$q_1=\int_{0}^{\infty}\int_{0}^{\infty}(H_0(x)-H_0(x-y))dF_{\Delta}(y)dF_{\chi}(x),$$

where $H_0(x)=\sum_{n=1}^{\infty}F_{\xi^1}*(F_{\xi^1}*F_{\eta^1}*F_{\varepsilon})^{*(n-1)}(x)$ - 0-function recovery process operation of the first security system. The second term is converted analogously

$$q_2=\sum_{n=1}^{\infty}\int_{0}^{\infty}\int_{0}^{\infty}M\left\{\sum_{i=1}^{\left[\frac{\xi_n^{(1)}}{T^{(1)}+\theta^{(1)}}\right]}\big(P(V_{n=1}+i(T^{(1)}+\theta^{(1)})-\theta^{(1)}\le x)-P(V_{n=1}+i(T^{(1)}+\theta^{(1)})-\theta^{(1)}\le x-y)\big)\right\}dF_{\xi}(y)dF_{\chi}(x)$$

The resulting expression can be written using functions 0-recovery:

$$q_2=\int_{0}^{\infty}\int_{0}^{\infty}M\left\{\sum_{i=1}^{\left[\frac{\xi^{(1)}}{\mathrm{T}^{(1)}+\theta^{(1)}}\right]}\big(H_{0i}(x)-H_{0i}(x-y)\big)\right\}dF_{\xi}(y)dF_{\chi}(x),$$

where $H_{0i}(x)=\sum_{n=1}^{\infty}F_{2i,n}(x)$, and $F_{2i,n}(x)=P(V_{n=1}+i(T^{(1)}+\theta^{(1)})-\theta^{(1)}\le x)$ ..

Further simplify the obtained relations is not possible, but you can get the asymptotic estimates, using the limit theorems of renewal theory [4,5]:

$$q_1 \approx \frac{1}{M\eta^{(1)} + (T^{(1)} + \theta^{(1)})M\left(\left[\frac{\xi^{(1)}}{\mathrm{T}^{(1)} + \theta^{(1)}}\right] + 1\right)} \int_0^\infty y dF_\Delta(y) = \tag{4}$$

$$= \frac{\left(M\eta^{(1)} + (T^{(1)} + \theta^{(1)}) + M\left\{\frac{\xi^{(1)}}{\mathrm{T}^{(1)} + \theta^{(1)}}\right\}(T^{(1)} + \theta^{(1)} - \delta)\right)^+}{M\eta^{(1)} + (T^{(1)} + \theta^{(1)})M\left(\left[\frac{\xi^{(1)}}{\mathrm{T}^{(1)} + \theta^{(1)}}\right] + 1\right)}$$

and

$$q_2 \approx \frac{M\left[\frac{\xi^{(1)}}{\mathrm{T}^{(1)} + \theta^{(1)}}\right]}{M\eta^{(1)} + (T^{(1)} + \theta^{(1)})M\left[\frac{\xi^{(1)}}{\mathrm{T}^{(1)} + \theta^{(1)}}\right] + (T^{(1)} + \theta^{(1)})} \int_0^\infty y dF_\zeta(y) = \tag{5}$$

$$= \frac{M\left[\frac{\xi^{(1)}}{\mathrm{T}^{(1)} + \theta^{(1)}}\right] M\left(\theta^{(1)} + \delta)^+\right)}{M\eta^{(1)} + (T^{(1)} + \theta^{(1)})M\left[\frac{\xi^{(1)}}{\mathrm{T}^{(1)} + \theta^{(1)}}\right] + (T^{(1)} + \theta^{(1)})}.$$

Summing (5) and (4) we obtain the probability that the first security system is not to parry by protection time interval $[\chi, \chi + \delta)$

$$\mathrm{P}(\overline{\mathrm{B}}) \approx \frac{M\left(\eta^{(1)} + (T^{(1)} + \theta^{(1)}) - \left\{\frac{\xi^{(1)}}{T^{(1)} + \theta^{(1)}}\right\}(T^{(1)} + \theta^{(1)}) - \delta\right)^+}{M\eta^{(1)} + (T^{(1)} + \theta^{(1)}) + (T^{(1)} + \theta^{(1)})M\left[\frac{\xi^{(1)}}{T^{(1)} + \theta^{(1)}}\right]} +$$

$$+ \frac{\mathrm{M}\left[\frac{\xi^{(1)}}{T^{(1)} + \theta^{(1)}}\right]\mathrm{M}\left(\theta^{(1)} - \delta\right)^+}{\mathrm{M}\eta^{(1)} + (T^{(1)} + \theta^{(1)}) + (T^{(1)} + \theta^{(1)})\mathrm{M}\left[\frac{\xi^{(1)}}{T^{(1)} + \theta^{(1)}}\right]}.$$

Written explicitly $\int_0^\infty y dF_\Delta(y)$ and $\int_0^\infty y dF_\zeta(y)$ is quite difficult even for the relatively simple case of exponential distributions of random variables, but you can use the Monte - Carlo method to obtain the desired numerical estimates. Note also that

$$P(B) = 1 - P(\overline{B}).$$

Probability q unauthorized access to the regeneration process of the information system in accordance with the formula of total probability we can write

$$\mathrm{q} = \mathrm{P}\left(accident \text{ on the regenaration cycle} \Big|_{\chi \le \varphi} \right) \mathrm{P}(\chi \le \varphi) +$$
$$+ \mathrm{P}\left(accident \text{ on the regenaration cycle} \Big|_{\chi > \varphi} \right) \mathrm{P}(\chi > \varphi) =$$
$$= \mathrm{q}^{(1,2)} \mathrm{P}(\chi \le \varphi) + \mathrm{q}^{(2)} \mathrm{P}(\chi > \varphi) = \mathrm{q}^{(1,2)} \int_0^\infty \mathrm{F}_\chi(\mathrm{t}) d\mathrm{F}_\varphi(\mathrm{t}) + \mathrm{q}^{(2)} \left(1 - \int_0^\infty \mathrm{F}_\chi(\mathrm{t}) d\mathrm{F}_\varphi(\mathrm{t}) \right),$$

Where $q^{(1,2)}$ - it's likely that the first regeneration cycle, followed by a second security system will not be able to parry, and $q^{(2)}$ - the likelihood that the second regeneration cycle security system will not be able to fend off the corresponding type of attack. Taking into account that the safety systems are independent from each other, we can write

$$q^{(1,2)} = P(\overline{B}) q^{(2)}.$$

And finally, the estimate for the probability of $q^{(2)}$ can be obtained using the same approach that was used in calculating the above mentioned $P(\overline{B})$. Omit the intermediate calculations and present the final result immediately:

$$q^{(2)} \approx 1 - \frac{M\xi^{(2)} - \theta^{(2)} M\left[\dfrac{\xi^{(2)}}{T^{(2)} + \theta^{(2)}}\right]}{M\eta^{(2)} + (T^{(2)} + \theta^{(2)}) + (T^{(2)} + \theta^{(2)}) M\left[\dfrac{\xi^{(2)}}{T^{(2)} + \theta^{(2)}}\right]}$$

Note that $q^{(2)}$ - factor unavailability second security system that takes the minimum value during the control period with optimal prevention $T^{(2)}_{onm}$ [6].

Optimal prevention period determined by the formula

$$T^{(2)}_{onm} = \sqrt{2\theta^{(2)} (M\xi^{(2)} + M\eta^{(2)})}$$

Thus, we managed to get the upper and lower asymptotic bounds (3) for the mean time to unauthorized access.

The proposed mathematical model of reliability of information security systems with layered security system with recoverable elements allows to take into account the temporal redundancy, when one security system "insures" the other. The proposed two-sided estimate for the expectation to unauthorized access to information provides a fairly narrow range of values for this indicator system reliability, simply calculated and takes into account a large number of different parameters of functioning of the system of information protection. The relations obtained are valid without any assumptions regarding the distribution functions of random variables.

REFERENCES

1. Corneliussen, K. & Hokstad, P. 2003. *Reliability Prediction Method for Safety Instrumented Systems; PDS Method Handbook, 2003 Edition.* SINTEF report STF38 A02420, SINTEF, Trondheim, Norway.

2. Gnedenko, B.V., Ushakov, I.A. 1995. *Probabilistic Reliability Engineering*. John Wiley & Sons, Inc.

3. Pereguda, A.I. 2001. Calculation of the Reliability Indicators of the System Protected Object-Control and Protection System. *Atomic Energy* 90: 460-468.

4. Rausand, M., Høyland, A. 2004 *System Reliability Theory: Models, Statistical Methods and Applications.* John Wiley & Sons, Inc.

5. Stoyan, D. 1983 *Comparison Methods for Queues and Other Stochastic Models*. Wiley-Interscience.

6. Pereguda, A.I., Timashov D.A. Mathematical model of reliability of security information systems.// Information . — 2009. —№8. —c.10–17.

METHOD AND ALGORITHM OF THE CHOICE OPTIMUM NUMBER ATTRIBUTES DESCRIBING RELIABILITY OF THE EQUIPMENT OF ELECTRO INSTALLATIONS

Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliyev A.Z.
•
Azerbaijan Scientific-Research and Design-Prospecting
Institute of Energetic AZ1012, Ave. H.Zardabi-94
e-mail:fem1939@rambler.ru

ABSTRACT

Estimation of parameters of individual reliability demands to estimate expediency of classification of all population of statistical data on the set versions of attributes. The method and algorithm optimum number of attributes classification of data is developed. Application allows calculate optimum on accuracy of an estimation of parameters of individual reliability, to reveal significant versions of attributes, to pass at the account of reliability of the concrete equipment with qualitative to a quantitative level

I. INSTRUCTION

Objective estimation of reliability of the concrete equipment of electro installations demands attraction of data about the attributes describing their individuality and reliability of work. A source of these data are the passport of the equipment, data on conditions of operation, reports of preventive tests, data on restoration of deterioration at carrying out of emergency and scheduled repairs. Thus, the knowledge of features of functioning of the equipment, deterioration of its units, an operational experience and intuition of the researcher has great value. Today probably, it is not necessary to approve, that such calculations with such volume and a variety of data can be spent only on the basis the automated information systems.

Each attribute is characterized not less than two versions of attributes (VA). For set VA the concrete equipment to make sample of data does not represent any difficulties. Difficulties arise when we pass to an estimation of parameters of reliability. It is known, that at such classification data for calculation of parameters of reliability or absolutely absent or insufficient for calculation of their average and guaranteed values. For this reason in practice only average values of parameters of reliability are calculated, in other words, the attributes noted above simply are not considered. Thus, the opportunity of the objective decision of many operational problems connected with the organization of maintenance service and repair of the equipment excluded. It is obvious, that subjective decisions it is far not always and not for all problems correct. To overcome noted difficulties of an estimation of parameters of individual reliability it is possible by the account of the importance of the set versions of attributes. Classification to insignificant attributes is inexpedient. In practice of an estimation of parameters of reliability of the equipment calculated for one, less often – to two VA. As examples, estimations of parameters of reliability for the equipment of a various class of a voltage, type, and an installation site can serve, etc.

However, at such formal classification there are unresolved following questions:

- What reliability of the assumption of not casual divergence of the average estimations of parameters of reliability calculated on all population of statistical data (Π^*_Σ) and the estimations calculated on sample of data ($\Pi^*_{v,i,j}$) for everyone VA, where i–a serial number of attributes, and j– serial number VA;
- What reliability of the assumption of not casual divergence of estimations on versions i-го an attribute, for example $\Pi^*_{v,i,j}$ and $\Pi^*_{v,i,(j+1)}$.

Such the sort of a problem should be solved in conditions of high uncertainty when practically it is not known about multivariate data, including about a kind of the law of distribution of random variables, and any assumptions (for example, about normal the law of distribution and independence VA) put new questions, in particular, about objectivity and risks of the erroneous decision [1]. Dependence of statistical data on set VA demands transition from concept «general population » to concept «final population of multivariate statistical data», and, hence, to transition to corresponding methods of the statistical analysis.

Development of methods and algorithms of a choice of the optimum number VA, describing reliability of the equipment of electro installations, allows calculating authentic estimations of parameters of individual reliability and to lower expenses for carrying out of their maintenance service and repair.

Statement of a problem.

Methods and algorithms of a choice of optimum number of attributes depend, first, on parameters of reliability and type of a scale of measurement of an attribute. Formulas of an estimation of parameters of reliability on statistical data of operation of the equipment known and presented by two versions:

1. The parameters of reliability calculated as average arithmetic random variables. For example, average duration of idle time in emergency repair, average duration of a finding in a reserve, etc.
2. The parameters of reliability calculated as probability of occurrence of event. For example, probability of refusal, specific number of refusals, probability of overlapping of conditions, etc.

In practice of a version of many attributes are fixed. For example, an installation site, a kind of the equipment, area of storm activity, etc. Only at a small part from these attributes number of versions equally to two.

VA with a quantitative scale of measurement set heuristically in the form of intervals of data. Hit of a random variable, for example, in j-th an interval testifies to display j-oh VA.

In the present article, we shall consider a method and algorithm of a choice of optimum number VA for the parameters of reliability concerning the first kind, i.e. calculated as average arithmetic random variables and we shall estimate them for attributes with a nominal scale of measurement.

Method and algorithm for a nominal scale of measurement attribute. For simplification of a statement, specify a solved problem. Let it is known m realizations of duration of idle time in emergency repair (τ_a) the instruction of date of occurrence of refusal. It is required to estimate laws of change of average duration of idle time in emergency repair on months. Thus, realizations τ_a are characterized by twelve VA. For the accepted conditions of a problem, calculations spend in following sequence:

1. Average arithmetic value of all realizations is calculated τ_a under the formula $M^*_\Sigma(\tau_a) = m^{-1}\sum_{j=1}^{m}\tau_j$
2. Samples of random variables are formed τ_a, concerning each of set VA. For example, at refusals in January (the first VA) sample looks like $\{\tau_a\}_1$;
3. Calculate average arithmetic value of random variables of each sample;
4. Random variables of each sample are placed in ascending order, fixed their minimal $\tau_{min,i}$ and the maximal value $\tau_{max,i}$ with i=1, n statistical functions of distribution (s.f.d. also are formed.) $F^*_{v,i}(\tau_a)$ with i=1,n;
5. Importance VA it is coordinated with representative samples concerning final population of multivariate data, i.e. a file $\{\tau_a\}_\Sigma$. At the first analysis stage, the realizations $M^*_{v,i}(\tau_a)$ calculated according to sample, we divide into three groups. The first group includes n1 the estimations $M^*_v(\tau_a)$ casually differing from $M^*_\Sigma(\tau_a)$, the second group includes n2 estimations $M^*_v(\tau_a)$ which not casually exceed $M^*_\Sigma(\tau_a)$, and the third group includes n3 estimations $M^*_v(\tau_a)$ which

it is not casual less $M^*_\Sigma(\tau_a)$. The second and third groups consist from significant VA, and the first – from insignificant. To the second group of estimations $M^*_v(\tau_a)$ we carry VA, for which $M^*_\Sigma(\tau_a)<\tau_{min,i}$ also does not concern, if $M^*_\Sigma(\tau_a)>M^*_{v,i}(\tau_a)$. If $\tau_{min,i}<M^*_\Sigma(\tau_a)<M^*_{v,i}(\tau_a)$, $M^*_\Sigma(\tau_a)$ it is compared with quantile s.f.d. $F^*[M^{**}_{v,i}(\tau_a)]$, corresponding probability β=0,05. S.f.d. $F^*[M^{**}_{v,i}(\tau_a)]$ pays off as follows:

On s.f.d. $F^*_{v,i}(\tau_a)$ By the stated [2] method it is modeled m_i realizations of a random variable τ_a, where m_i number of realizations of sample for considered VA;
Average arithmetic value of these realizations $M^{**}_{v,i}(\tau_a)$ is calculated;
Calculations on i.5.1. and i.5.2. repeat N time;
Is under construction s.f.d. $F^*[M^{**}_{v,i}(\tau_a)]$ and on realization $M^{**}_{v,i}(\tau_a)$ with serial number Nβ It is defined квантиль the distribution $M^{**}_{v,i}(\tau_a)$ corresponding probability β.

Experience of calculations shows, that the number of realizations N=100 is quite enough for classification VA on the groups noted above. Performance of a condition $M^*_\Sigma(\tau_a)<M^{**}_{v,i}(\tau_a)$ with a significance value β testifies to an accessory considered VA to the second group. Notice, that conditions $M^*_\Sigma(\tau_a)<\tau_{min,i}$ unequivocally $M^*_\Sigma(\tau_a)>M^*_{v,i}(\tau_a)$ testify that considered VA, accordingly, belongs and does not belong to second group VA. However these conditions concern to a kind necessary, but insufficient for judgment about an accessory of all VA to the first group. They are rather useful to decrease in time of calculation;

6. If VA do not concern to the second group, except for a case, when $\tau_{min,i}<M^*_\Sigma(\tau_a)<\underline{M^{**}_{v,i}}(\tau_a)$, this VA is checked on conformity to the third group. For what the condition is checked: $M^*_\Sigma(\tau_a)>\tau_{max,i}$. If this condition carried out, VA belongs to the third group. Otherwise the condition $M^*_\Sigma(\tau_a)<M^*_{v,i}(\tau_a)$ is checked. If the condition carried out, VA belongs to the first group. If $M^*_{v,i}(\tau_a)<M^*_\Sigma(\tau_a)<\tau_{max,i}$ it is entered into consideration quantile the distributions $F^*[M^*_{v,i}(\tau_a)]$, corresponding probability (1-β). If it quantile $\overline{M^{**}_{v,i}}(\tau_a)$ appears less, than $M^*_\Sigma(\tau_a)$ considered VA concerns to the third group, and otherwise – to the first;
7. Further pass to consideration next VA with constant conditions of check representative samples;
8. The first stage of calculation comes to the end by consideration of all VA;
9. If it will appear, that n1≤1 and (or) n2≤1 classification of a final data population on VA comes to the end. From n VA no more than to two VA there correspond the estimations $M^*_v(\tau_a)$ differing from $M^*_\Sigma(\tau_a)$;
10. Otherwise it is necessary to continue calculations and to establish character of a divergence of parameters of reliability concerning the second group and (or) to the third group. Considering, that methodology for the analysis of importance VA for the second and third groups are similar, we shall consider only a case, when n1> 1;
11. From final population of multivariate data $\{\tau_a\}_\Sigma$ spend sample of the realizations corresponding one of n1 VA of the second group. We shall designate set of realizations of this sample as $\{\tau_a\}_{\Sigma 1}$;
12. All the subsequent calculations it is spent in strict conformity with i.(1÷8) with that difference, that instead of $\{\tau_a\}_\Sigma$ it is analyzed $\{\tau_a\}_{\Sigma 1}$.

This stage of calculations allows increase number of significant attributes with 3 up to 7. Experience of calculations for statistical data of operation of the equipment of electro installations shows, that on it calculations, as a rule, come to the end

The block the scheme of algorithm of a choice of optimum number VA is resulted in figure 1 and supplements practical realization of a method for a nominal scale of measurement VA.

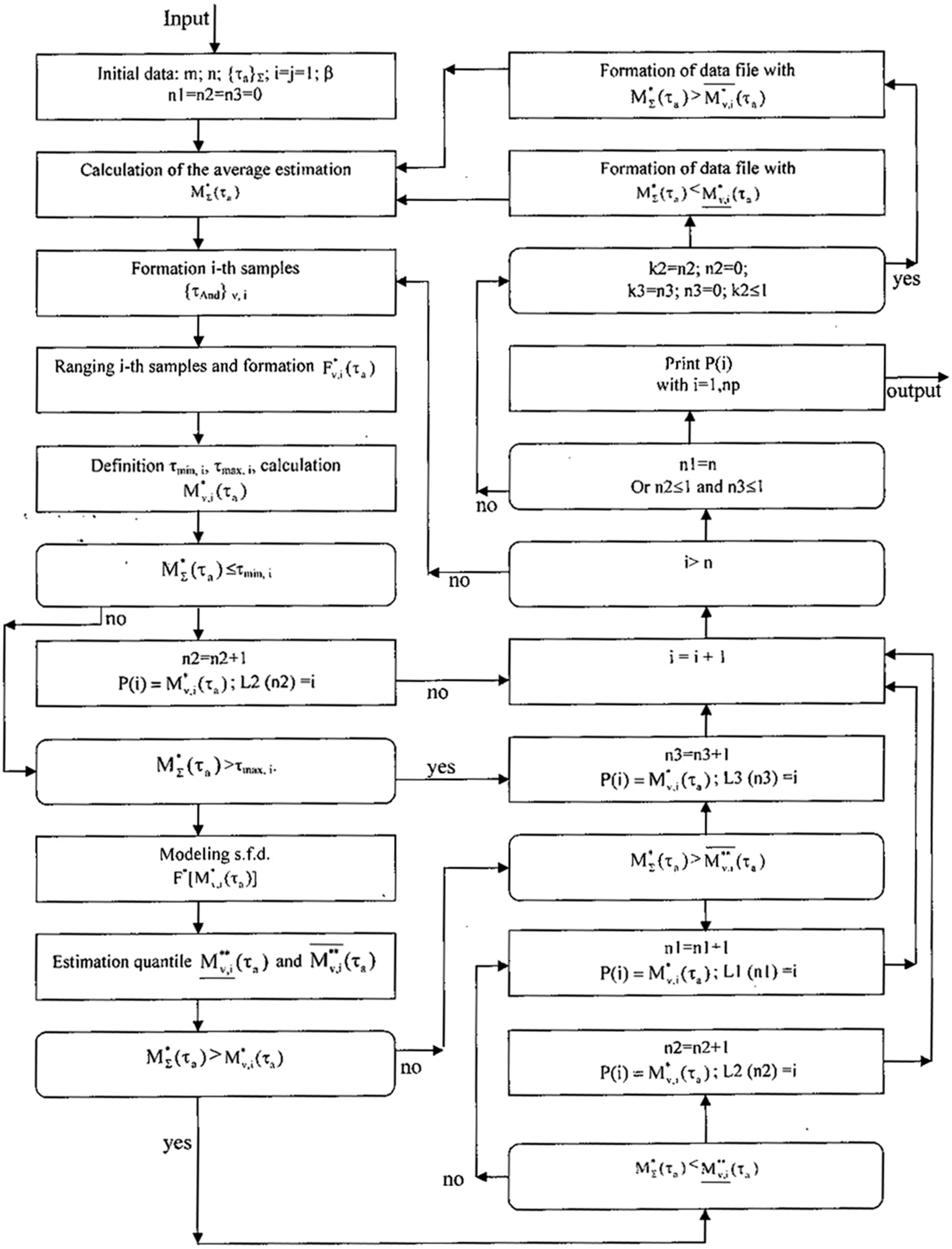


Fig.1. Algorithm of a choice of optimum number VA

Example of calculation. The illustration of calculations on the basis statistical data of operation is not always justified by the method stated above, as the true result and consequently, there is nothing to compare with results of calculations here is not always known, to estimate reliability. In these conditions, it is expedient to apply a method of the decision of "a return problem» according to which statistical data are modeled on in advance stipulated model, and results of calculation on these data should repeat structure of model.

In table 1 realizations of random variables are resulted τ For four VA with the set functions of distribution F(τ). Modeling of realizations was spent on the basis of random numbers ξ [2 tabl.9.1.] with are uniform distribution in an interval [0,1].

Table 1

Realization of random variables τ_a

N	Function of distribution of realizations for VA							
	τ=80ξ		τ=100 (ξ+0,5)		τ=50 (ξ+1,5)		τ=50 (ξ+3)	
	ξ	τ, c	ξ	τ, c	ξ	τ, c	ξ	τ, c
1	0,1009	8,1	0,1218	62,2	0,7942	114,7	0,9959	199,8
2	0,3754	30	0,6606	116,1	0,8868	119,3	0,6548	182,7
3	0,0842	6,7	0,3106	81,1			0,8012	190,1
4	0,9901	79,2	0,8526	135,3			0,7435	187,2
5			0,6357	113,6			0,6991	185,0
6			0,7379	123,8			0,0989	155,0
7			0,6234	112,3				
8			0,1180	61,8				

Given tables 1 allow to establish:

1. $M^*_\Sigma(\tau)=115.1c.$;
2. Boundary values of intervals of disorders of realizations of a random variable τ For considered four VA (models) ($\tau_{min} \div \tau_{max}$) are accordingly equal: (6,7÷79,2); (62,2÷135,3); (114,7÷119,3); and (155÷199,8);
3. Average values of each of four выборок $M^*_v(\tau)$ are accordingly equal: 31c., 100,8c., 117c., 183,1c.;
4. Comparison $M^*_\Sigma(\tau)$ to boundary values τ_{min} and τ_{max} shows:

4.1. The first VA for which $M^*_\Sigma(\tau)>\tau_{max,\,1}$, causes not casual divergence with $M^*_\Sigma(\tau)$;

4.2. For the second VA $\tau_{min}<M^*_\Sigma(\tau_a)<\tau_{max}$ and therefore $M^*_\Sigma(\tau_a)>M^*_{v,2}(\tau)$, it is required attraction of size $\overline{M^{**}_{v,2}}(\tau)$;

4.3. For the third VA also as well as at the second $\tau_{min,\,3}<M^*_\Sigma(\tau_a)<\tau_{max,\,3}$, but $M^*_\Sigma(\tau)<M^*_{v,3}(\tau)$ attraction of size $\underline{M^{**}_{v,3}}(\tau)$ also is required;

4.4. The fourth VA for which $M^*_\Sigma(\tau)<\tau_{min,\,4}$ causes not casual divergence $M^*_\Sigma(\tau)$ and $M^{**}_{v,4}(\tau)$;

5. Results of calculations $\overline{M^{**}_{v,i}}(\tau)$ and $\underline{M^{**}_{v,i}}(\tau)$ are resulted to table 2;

Table 2

Results of calculations of critical values of realizations $M^{*}_{v,i}(\tau_a)$

Parameter	Level the importance β	Conditional number VA (i)			
		1	2	3	4
$\underline{M^{*}_{v,4}}(\tau)$	0,05	11,5	84,4	116,2	171,8
$\overline{M^{**}_{v,4}}(\tau)$	0,95	61,8	118,2	119,1	193,6

6. As follows from table 2, $M^{*}_{\Sigma}(\tau) < \overline{M^{**}_{v,2}}(\tau)$, differently burning $M^{*}_{\Sigma}(\tau)$ and $M^{*}_{v,2}(\tau)$ differ casually, and $M^{*}_{\Sigma}(\tau) < \underline{M^{**}_{v,3}}(\tau)$ and therefore $M^{*}_{\Sigma}(\tau_a)$ and $M^{*}_{v,3}(\tau)$ differ not casually;
7. Results of calculations of an estimation of expediency of classification of statistical data on set VA it is resulted in table 3

Table 3

Results of classification of statistical data on VA

Number VA	Parameter	Numerical value	Significant VA
1	$M^{*}_{v,1}(\tau)$	31	Yes
2	$M^{*}_{v,2}(\tau)$	115,1	No
3	$M^{*}_{v,3}(\tau)$	117	Yes
4	$M^{*}_{v,4}(\tau)$	183,1	Yes
-	$M^{*}_{\Sigma}(\tau)$	115,1	-

CONCLUSIONS

1. Unreasonable classification of statistical data on set VA (for example, on a class of a voltage, type, an installation site, months of year, duration of operation and so forth) can lead to distortion of real values of parameters of reliability of the equipment of electro installations.
2. The method and algorithm of definition of optimum number VA on based on imitating modeling and the theory of check of statistical hypotheses is developed.
3. Application of this method allows to raise accuracy of estimations of parameters of reliability calculated as an average arithmetic random variables, to reveal significant VA, to raise efficiency of monitoring of reliability of the equipment of electro installations

REFERENCES

1. Lbov G.S. Method's foreheads of processing of polytypic experimental data. - Novosibirsk: the Science, 1981
2. Farhadzadeh E.M., Farzaliyev Y.Z., Muradaliyev A.Z. Comparison of methods of modeling of continuous random variables on empirical distributions. Baku, Problems of Power №1, 2013, with. 25-31

ESTIMATING OF CRITICAL SOFTWARE LATENT FAULTS PRESENCE WITH REQUIRED TRUSTWORTHINESS

Boris Konorev[1], Vladimir Sergiienko[1,2], Vyacheslav Kharchenko[1], Grygoriy Zholtkevych[3]

•

[1] N.E. Zhukovsky National Aerospace University "KhAI", Kharkiv, Ukraine
[2] I&C System Certification Center, Kharkiv, Ukraine
[3] V.N. Karazin Kharkiv National University, Kharkiv, Ukraine

e-mail: bkonorev@gmail.com, vov3d@mail.ru,
V.Kharchenko@khai.edu g.zholtkevych@gmail.com

ABSTRACT

The estimation of latent faults probability is a key indicator for quantitative assessment of critical systems reliability and safety. The article presents the method of latent faults probability estimating for critical software with a required (controlled) trustworthiness. Compositional set-theoretical model of residual and latent faults is proposed for framework assessment of latent faults. The procedure of the experimental calibration of faults sensitivity and diversity degree of inspection methods of source software faultlessness is presented. The "dotted" injections method using faults profile for the level of programming languages is proposed.

1. INTRODUCTION

Possible damage of the software fault propagation on the system level during operation of I&C systems critical to safety can be in the range "material losses - damage to the environment - a threat to the health and life of human being". Latent faults (which are not detected during testing) in critical software are considerable risk factors, as they may lead to failure of I&C systems during performing functions critical to safety. Therefore, the estimation of latent faults probability - one of the most important tasks for qualification testing and risk management of critical to safety systems maintenance.

At the moment, for the tests of reliability and functionality of critical software the model-checking approach is used more widely [Peled (2009), Baier (2008)]. Model checking can be implemented during static analysis [Moiseev (2013)]. Multi-model model-checking approach may be used to confirm the absence of latent faults [Konorev (2009), Konorev (2011), Konorev (2012)]. This approach consists in critical software source code validation by use of the invariants - oriented models. The invariant is an invariable attribute of software. The invariants - oriented models are generating at static software source code analysis to verify the correctness of each software invariant [Brukhankov (2010)]. Each such model of verifiable software project contains the necessary data to monitor a specific invariant. After that invariants measurement methods (using invariants - oriented models as input data) implement an algorithm of inspection of specific invariants values. Each invariants measurement method either confirms invariant consistency (unchangeability), or detects abnormality. Evidence of the approach is based on the usage of diversity principle - various (diverse) invariants measurement methods are used for latent faults detection. In general, the methods have different sensitivity to faults (diversity degree - the degree of differences between them). To evaluate the fault sensitivity and diversity degree of methods the procedure of artificial faults injection can be used [Cotroneo (2013)]. Faults can be injected both to the executable code, and at the level of the source codes [Lanzaro (2013)].

In the article the approach of latent faults probability estimation is proposed, which is based on an assessment of faults sensitivity and diversity degree of invariants measurement methods.

The basis for probability estimation of latent faults is a set-theoretical model of latent and residual software faults, the model allows getting a framework assessment of latent faults probability. The procedure of experimental calibration of sensitivity and diversity degree of invariants measurement methods is used for model parameterization, the procedure is based on the use of test faults "dotted" injection.

The result of the experimental calibration of invariants measurement methods is the definition of the partial sensitivity (as the probability of latent fault detection) of the j-th method to the i-th faults type from injectable test faults subset, expressed as a set of residual faults of the i-th type for the j-th method .

As a result of processing of data, collected during the calibration, diversity degree of invariants measurement methods can be determined, an assessment of source code test coverage completeness is performed, residual faults area and the border area of latent faults location are established.

2. MODEL OF RESIDUAL AND LATENT SOFTWARE FAULTS

To assess the effect of using different invariants measurement methods the set-theoretical model of the residual and latent software faults for the composition of invariants measurement methods is proposed (see Fig. 1). Latent fault is software code or documentation incorrectness which was not found during development and testing and could lead to one or more failures of a system component or an entire system. Efforts should be focused exactly on the removal or determining the location of this faults type during qualification testing, that includes software independent verification. Estimation of latent faults probability is encouraged to obtain indirectly from the residual faults data. Residual faults for invariants measurement methods are faults, which are not detected by these methods. Location area of residual faults can be set experimentally by procedure of invariants measurement methods calibration.

The model, presented by Venn diagram (Fig. 1), illustrates a possible positional relationship of residual and latent faults sets and allows estimating the advantages due to the use of invariants measurement methods for a variety of positional relationship (superposition) of residual faults sets.

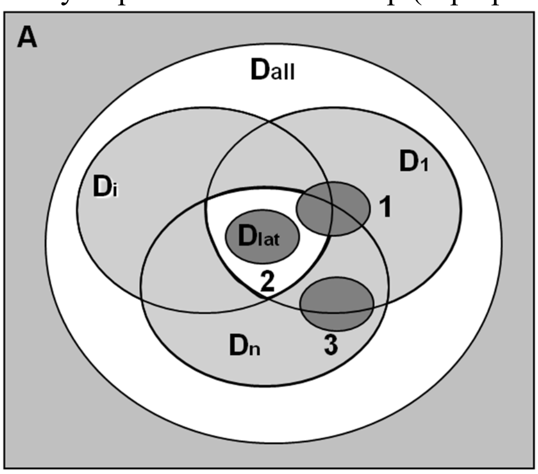


Figure 1. Model of latent faults for invariants measurement diverse methods, where D_{all} – the set of possible faults in the address space of the software A; D_i - the set of faults, which are not detected by measurement method of the i-type invariants; D_{lat} - the set of latent faults.

Assessment indicator of achievable effect for invariants measuring methods composition implementation at independent verification is the quantity (%), on which the boundary value of latent faults probability P_{lat} decreases.

$$I = \frac{\left| D_{lat} \setminus \bigcap_{i=1}^{n} D_i \right|}{\left| D_{lat} \right|} * P_{lat}, \tag{1}$$

where n – number of used methods.

The intersection of residual faults subsets is a frame assessment of the latent fault probability and can be used as estimation result.

The main benefit (advantage) of using invariants measurement methods is determined by the diversity degree of invariants measurement methods in terms of a specific project. It is possible to define 3 alternatives:

a) theoretically possible case of latent faults probability decrease to 100% (I=1, position 3 in Fig.1):

$$\left(\bigcap_{i=1}^{n} D_i \right) \cap D_{lat} = \varnothing, \tag{2}$$

b) ultimate unfavorable case (I=0, position 2 in Fig.1):

$$D_{lat} \subset \bigcap_{i=1}^{n} D_i, \tag{3}$$

c) common case(position 1 in Fig.1)

$$\left(\bigcap_{i=1}^{n} D_i \right) \cap D_{lat} \neq \varnothing, \tag{4}$$

If latent faults (elements of the set D_{lat}) are not detected during independent verification, though the indicator of achievable effect I=0, subset $\bigcap_{i=1}^{n} D_i$ can be used as boundary area of latent fault location (it is frame assessment of latent fault probability). The area (set) is established, where faults absence is guaranteed - $\overline{\bigcap_{i=1}^{n} D_i}$. The efforts should be focused on the area $\bigcap_{i=1}^{n} D_i$ analysis (searching the faults inside this area) to improve or refine the assessment. The relative benefit of using measurement invariants methods V_{rel} is defined as:

$$V_{rel} = 1 - \frac{\left| \bigcap_{i=1}^{n} D_i \right|}{\left| \bigcup_{i=1}^{n} D_i \right|}. \tag{5}$$

Thus, the proposed set-theoretical model allows defining the area of latent faults and to establish an absolute and relative gains during usage of diverse invariants measurement methods.

3. PROCEDURE OF EXPERIMENTAL CALIBRATION

In order to estimate the probability of latent faults presence for specific software project, experimental data on the fault sensitivity and the diversity degree of invariants measurement methods must be obtained in accordance with the proposed model of residual and latent faults (shown in Fig. 1). Parameterization of the model is getting through the procedure of experimental calibration of invariants measurement methods by the method of test faults "dotted" injection. The method is the "dotted" (single) injection of a specific type test fault in the selected location of software, sensitivity binary assessment of invariant measurement methods (detection/undetection of injected fault) and returning to its initial state (removal of the test fault). "Dotted" injection means a single fault injection, this guarantees the absence of mutual influence and the emergence of a secondary faults due to interference and mutation of injected faults.

To assess the completeness of test coverage and methods sensitivity during each step (layer) of calibration the following actions must be performed:

Generating of test faults profile with the use of results of syntactic and semantic parsing in the mode of software source code static analysis;

Consecutive choice of fault type and injection of a single fault of selected type into checked software;

Checking by all invariants measurement methods;

Fixation (registration) of detection/undetection of injected test faults for each method;

Backout to the software initial state (the state before fault injection) and repeat the procedure from step 2.

The procedure is performed cyclically for all possible fault types for a specific project.

Quantity of injected faults (number of steps) for each fault type is determined on the assumption of required accuracy (reliability, uncertainty) results.

Graphic representation of the calibration model is shown in Figure 2.

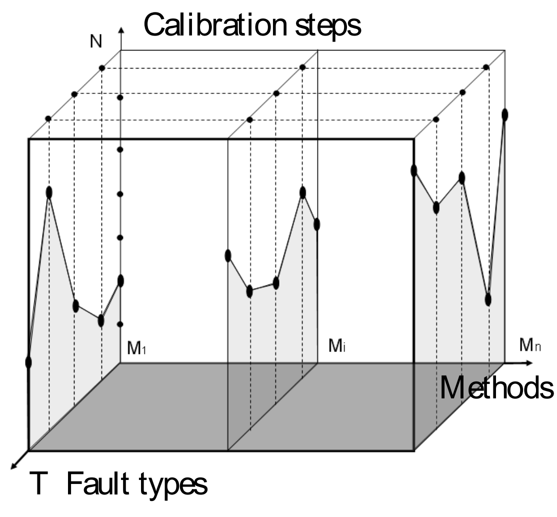


Figure 2. Calibration model.

Experimental calibration of fault sensitivity of each method and integral sensitivity of diverse methods composition is produced in the context of calibration tests space (Fig. 2), which represents Cartesian product:

$$P_{cal} \subseteq N \times M \times T , \qquad (6)$$

where N={n_i} – set of test faults injections that are implemented during the calibration;
M={m_j} – set of calibrated invariants measurement methods;
T={t_k} – set of injected test faults types or test faults profile (TFP).

TFP for a specific project is determined by the used language constructs composition. Each construction can be distorted by certain fault types, which injection will not lead to violation of the syntax and semantics of the program (i.e., the fault injection will not cause an error during compilation). These types of faults are forming test faults profile of a specific project.

P_{cal} defines the set (tuple) of the outcomes - the results of experiments during calibration. P_{cal} item value with the coordinate (n_i; m_j; t_k) possesses the value 0 if the test fault was not found, or 1 if the test fault was detected.

Injection of a tuple T={t_k} during method M = {m_j} calibration is a step or layer of calibration in the P_{cal} space.

Each calibration step (increment along the axis N) can be represented as a "secant plane": the injection of faults is performed from the test faults profile and verification is fulfilled by all methods.

At each step of "secant plane" the following elements may be defined:

Disjunction of undetected test faults for each M_i

$$D_{fault\_i} = \bigcup_{j=1}^{m} D_j , \tag{7}$$

Conjunction of undetected faults for each fault profile type for all M_i.

$$D_{t\_k} = \bigcap_{i=1}^{n} D_i , \tag{8}$$

Partial sensitivity $P_{part\ ji}$ of each (the j-th) method to the i-th fault type as a set of detected faults by the j-th method (sensitivity to the concrete faults type).

As a result of calibration the following characteristics are determined:

–Integral sensitivity of each method for chosen test faults profile TFP:

$$P_j = \bigcup_{i=1}^{k} P_{part\ ji} , \tag{9}$$

where k – total number of fault types;

–In pairs for all M_i (for all pairs) absolute and relative diversity degree :

$$m_{ijabs} = \left|P_i \Delta P_j\right| = \left|P_i \cup P_j\right| \setminus \left|P_i \cap P_j\right| , \tag{10}$$

$$m_{ij} = 1 - \frac{\left|P_i \cap P_j\right|}{\left|P_i \cup P_j\right|} , \tag{11}$$

where i and j take the values from 1 to n; n - number of used methods.

Using specified data the list of invariants measurement methods can be defined, this set ensures achievement of set point of test coverage completeness and validity (with a specified trustworthiness) of latent faults probability estimation.

4. FAULTS "DOTTED" INJECTION METHOD

The proposed method of faults "dotted" injection recommends the places (answers the questions) in which the test faults of TFP should be injected and the way how the injection procedure should be implemented.

Detailed algorithm of action sequences "fault injection - analysis of fault impact on invariants consistency - returning of software to its initial state" is shown in Figure 3.

1. Fault type is selected alternately from TFP.
2. The injection places (constructions) for each fault type are established, – only those places are chosen from full spectrum of software project operations, where injection of the fault does not lead to a violation of syntax and semantics program (i.e., the injection of the fault will not lead to an error during compilation). In other words, the faults in the software, which are detected during compilation, are not taken into account during the choice of faults for injection.
3. All possible injection places form a list of points for injection, - each established place gets its identification ordinal number.
4. If the number of possible injection places is very large (for a real project it can be tens of thousands), the number of injections n is determined by required reliability of the results.
5. The injection place is chosen randomly from the general list of points and injection is implemented (replacement of the correct construction with incorrect). The point, which has already been exposed to distortion, is excluded from the list (to prevent re-injection of the fault in the same place).
6. The modified code analysis (checking) is implemented by all invariants measurement methods which were determined for a specific project.
7. Check results are fixed in binary form (a fault is detected / not detected).
8. Verifiable code returns to a state, when the fault was not injected.
9. Steps 5-8 are executed number of times established on the 4th step.
10. If TFP is not covered, the procedure must be implemented from step 1 (the procedure should be performed for all fault types from TFP).

5. PROCESSING OF CALIBRATION RESULTS

Processing of collected during the calibration data allows calculating a number of metrics.

Numerical expression of partial sensitivity of method to each faults type is determined as:

$$S_{part\,ij} = \frac{n_{det\,ij}}{n_i} * 100\%\,, \tag{12}$$

where n_i - is general number of measurements of the i-th faults type; $n_{det\ ij}$ - number of measurements, in which the j-th method detected the i-th faults type.

Note: it is possible that two methods have equal partial sensitivity to some fault type, but at the same time they found faults in different injection places. In this case, to increase the degree of test coverage it is necessary to use both methods.

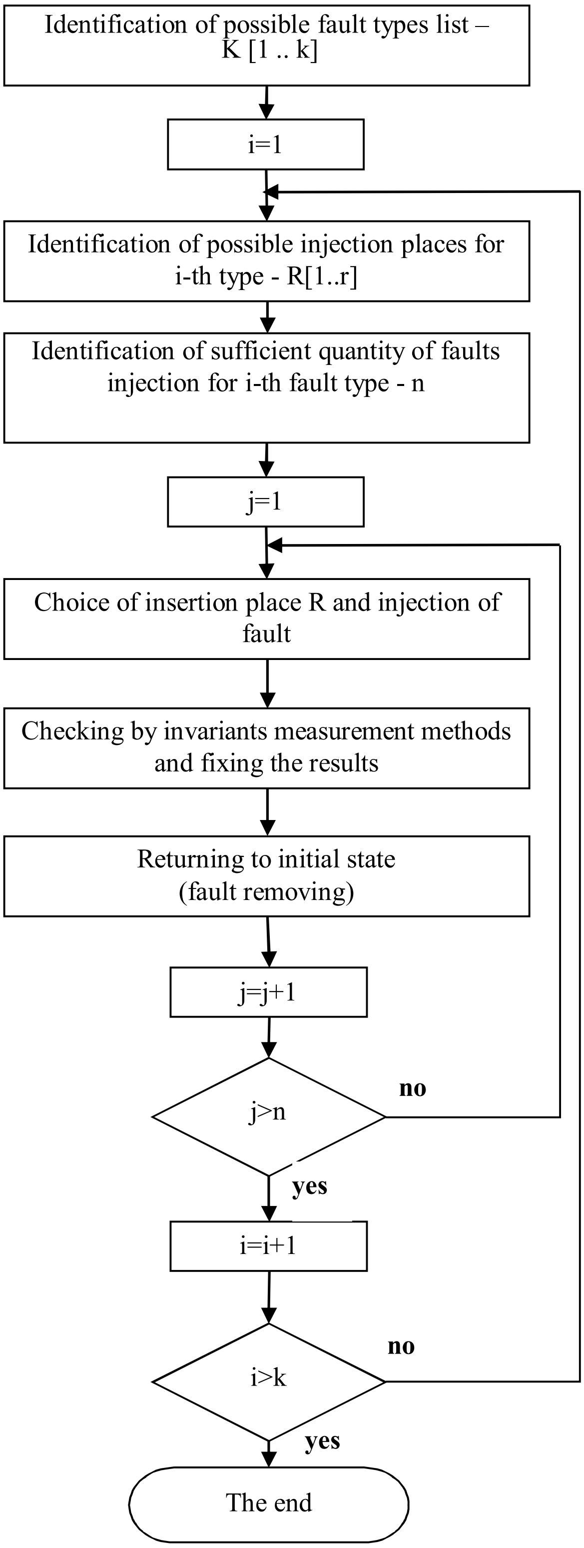


Figure 3. Test faults "drop" (dotted) injection algorithm.

The minimum required number of measurements n is calculated taking into account the established requirements of result reliability. At the same time, it is guaranteed that the discrepancy between true and calculated medium measured value does not exceed legitimate value.

Percentage of detected faults for each type $S_{det\ i}$:

$$S_{det\,i} = \frac{n_{det\,i}}{n_i} * 100\%\ , \qquad (13)$$

where $n_{det\ i}$ – number of measurements, in which at least one method detected i-th type of fault.

Percentage of detected faults for each method S_j:

$$S_j = \frac{\sum_{i=1}^{k} n_{det\,ij}}{\sum_{i=1}^{k} n_{det\,i}} * 100\%\ , \qquad (14)$$

where k – number of fault types.

As a result, the probability of latent faults existence P_{lat} is defined as:

$$P_{lat} = \frac{\sum_{i=1}^{k} (S_{det\,i} * r_i / 100)}{\sum_{i=1}^{k} r_i} * 100\%\ , \qquad (15)$$

where r_i – number of language structures in code, that can contain the i-th type of fault.

6. CASE STUDY

Testing of the proposed technique was performed in the framework of the Science and Technology Center in Ukraine project # 4726 «Safety-critical software independent verification and latent faults assessment based on diverse measurement of invariants». Model-checking verification has been implemented in static software source code analysis on IDE Eclipse platform. Designed Mobile instrumental complex enabled to check a number of invariants. These invariants were the following:

–Semantic invariants of software variables (physical dimension, variation interval, the accuracy of the representation);

–Control flow invariants: reducibility of control thread, the potential accessibility and "liveliness" (demand) of operators;

–RAM usage in a particular software project: memory leaks, repeated memory release;

–The structure of the threads of program control (the execution logic);

–Specific invariants for FPGA-based I&C systems ("list of sensitivity", "signal race", absence of "latches", etc) [Kharchenko (2012)].

7. CONCLUSIONS

The proposed technique allows evaluating latent faults probability with controlled (required) trustworthiness during independent verification of critical software. Trustworthiness is the controlled value during application of experimental calibration. It is determined by the test faults profile completeness and the accepted amount of injected faults.

The composite set-theoretical model of residual and latent faults is provided. This model allows evaluating of residual faults, to assess the area of latent faults location and test coverage completeness.

"Dotted" injection method of test faults is presented with the use of faults profile, which takes into account the specifics of a project at the level of the programming language.

The proposed procedure of experimental calibration of faults sensitivity and diversity degree of invariants measurement methods is based on the faults "dotted" injection method, it allows eliminating the possibility of interference (superposition) and mutation of faults in address field of software.

The proposed method is developed in framework of the invariant-oriented model-checking approach for critical software verification. Its application can be extended to other domains.

REFERENCES

1. Peled D., Pelliccione P., Spoletini P. (2009). Model Checking. *Wiley Encyclopedia of Computer Science and Engineering*. pp. 1904–1920.
2. Baier C., Katoen J.-P. (2008). *Principles of Model Checking*. The MIT Press. 984 p.
3. Moiseev M. (2013). Static Analysis Approach for Defect Detection in Multithreaded C/C++ Programs. In: *Software Engineering for Resilient Systems (SERENE 2013)* Springer LNCS, Vol. 8166. pp. 169-183.
4. Konorev B., Kharchenko V. (Eds) (2009). *Invariant-oriented assessment of space systems software quality*. Kharkiv: National Aerospace University KhAI. - 221 p. [in Russian]
5. Konorev B., Sergiyenko V., Chertkov G. (2011). The Evidential Independent Verification of Software of Information and Control Systems, Critical to Safety: Functional Model of Scenario. In. *IEEE East-West Design & Test Symposium (EWDTS'2011)* Kharkiv: KNURE, Ukraine. pp. 263-266.
6. Konorev B., Sergienko V., Zholtkevych G. , Chertkov G., Alexeev Y. (2012). Concept of Critical Software Independent Verification Based on Invariant-oriented Model-checking Approach *Radio-electronic and Computer Systems*. – Vol. 57, № 5. – pp. 184 – 190.
7. Brukhankov S., Konorev B., L'vov M., Zholtkevych G. (2010). About Static Analysis of Variables Physical Dimensions for Critical-mission Software. *Radio-electronic and Computer Systems*. – Vol. 47, # 6. – pp. 186 – 191.
8. Cotroneo D. & Madeira H. (2013). Introduction to Software Fault Injection. In: *Innovative Technologies for Dependable OTS-Based Critical Systems. Challenges and Achievements of the CRITICAL STEP Project*. Milan: Springer. pp. 1-15.
9. Lanzaro A., Natella R., Barbosa R. (2013). Tools for Injecting Software Faults at the Binary and Source-Code Level. In: *Innovative Technologies for Dependable OTS-Based Critical Systems. Challenges and Achievements of the CRITICAL STEP Project*. Milan: Springer. pp. 85-100.
10. Kharchenko V. (Ed.) (2012). *Case-assessment of critical software systems*, Volumes 1-3 / Volume 3. Safety. Kharkiv: National Aerospace University KhAI. – 302 p. [in Russian]

THE ANALYTIC HIERARCHY PROCESS MODIFICATION FOR DECISION MAKING UNDER UNCERTAINTY

Alexander V. Bochkov, Nikolay N. Zhigirev

•

LLC NIIgazeconomika – Science Research Institute
of Economics and Management in Gas Indystry, Moscow, Russia

e-mail: a.bochkov@gmail.com
e-mail: nnzhigirev@mail.ru

ABSTRACT

The AHP-method proved its efficiency in different situations of decision making under conditions when experts can perform pairwise comparisons of compared objects under study. But, our real life consists of uncertainties. How can we compare the objects if information about them is not available or extremely uncertain? Pairwise comparisons are impossible. In this case, we propose to use the approach of minimizing the functional «errors» in the evaluation of comparable objects (F-ratio test). Through successive iterations initially given equal weight values compared objects are close to the true value (within the stated error). This method does not require the consistency of the pairwise comparisons matrix and, moreover, can be applied to AHP incomplete matrix. The iterations are repeated as long as the new changes do not result in improved objects estimates. The practical example of the method use is considered.

Keywords: AHP, objects estimates, uncertainty, F-ratio test, incomplete matrix.

1. INTRODUCTION

Huge number of tasks in different fields of knowledge (such as: definition of relative weight purposes during carrying out multicriteria optimization; definition of expected result of activity, for example determination of investment project efficiency; definition of probabilities of implementation of accident development various scenarios in potentially dangerous object in the probabilistic analysis of its safety; a task of probabilities of outcomes in matrix of consequences at a choice of rational decision in the conditions of partial uncertainty, etc.) is solved in conditions when the behavior of studied process branches and it is necessary to define weights (relative probabilities) of possible scenarios (compared objects). In other tasks some parameters of mathematical models can accept various values depending on some not formalized (not described mathematically) factors. Definition of probabilistic distributions of these parameters is often impossible due to the lack of rather representative statistics. For value assignment of these parameters qualitative or qualitative and quantitative scales are used.

In all considered cases (both a scenario choice, and class definition or object state in some scale) mathematically the task is reduced to compared objects choice from a set possible by means of expert estimation. Whereby the interest is not only in choosing of the most probable compared objects, but also in defining compared objects weight (relative probabilities). This information is used further in algorithms of simulation modeling, and in simpler models in averaging operations taking into account compared objects weight.

Simplicity of relative compared objects weight calculation at their determination accuracy preservation often plays considerable part. The analysis shows that the most reliable and widely used method of problem solution of choice and definition of relative compared objects weight (relative probabilities) is the pairwise comparisons method on qualitative attribute with quantitative assessment of preferences [1].

2. LITERATURE REVIEW

The pairwise comparisons method [1] had broad application when determining relative indicators of compared objects importance against chosen qualitative criterion in decision-making support systems. Concerning some general property for compared objects [2-5] this method allows to carry out their ranging, to define a priority of one compared object (criterion, purpose, alternative) before another. One of the most developed and widely practiced is pairwise comparisons method modification, received the name of T. Saati's Analitic Hierarchy Process (AHP) [6].

One of essential shortcomings of pairwise comparisons method various realization is the essential increase in labor input of necessary calculations, with growth of number of estimated objects. In some cases it is too difficult for an expert to perform a large number of pairwise comparisons or he can't simply compare two objects on the offered preferences scale.

There is a problem of processing not completely certain pairwise comparisons matrixes. The solution of this task is connected to the following problems [7]. The first problem is in definition of necessary and sufficient conditions of calculation of all components of compared object importance indicators vector, and the second one is in definition of the algorithm of these components calculation itself.

In [8] these problems are solved concerning of pairwise comparisons results processing in fundamental scale by Eigenvector Following method. The method offered in [8], in point of fact, is reduced to restoration of full matrix of comparisons on available incomplete matrix and application standard algorithm of relative compared objects weight to the restored matrix. For method justification it is offered to use graph interpretation of pairwise comparisons method. Tops of graph G designate alternatives $A_i \in A$, $i = \overline{1,k}$ from which arcs d_{ij} release, connected them with each other. It is proved that d_{ij} is equal to geometrical average intensity of all possible ways connecting tops A_i and A_j. Necessary and sufficient condition for this method realization: between any pair of tops in graph G there has to be at least one way.

3. PROBLEM STATEMENT

We consider the problem of compared objects weight definition importance according to incomplete pairwise comparisons matrixes. In practice most precisely pairwise comparisons are given by experts between "the most significant objects and other objects" and between "the least significant objects and others", that is when differences between pairwise compared objects are contrast. When compared objects are close "naturally" to comparisons matrix element the expert, according to the scale [6], assigns value 1 that leads, in our opinion, to unjustified alignment of compared objects weight estimates.

We offer to consider, in this case, pairwise comparisons uncertain. Whereby instead of procedure of recovery of such uncertain data we offer reverse procedure lying in the fact that to find such subset of pairwise comparisons which is most coordinated and this subset, finally, assign values of compared objects weight. Thereby two aims are achieved. First, the required decision in the form of distribution of compared objects weight importance will be with some set accuracy is coordinated with pairwise comparisons matrix constructed on AHP method. Secondly, the remained pairwise comparisons will indicate communication circuits between compared objects, explaining distinctions of defined weights

.

So, the task is set as follows: not completely filled pairwise comparisons matrix is set and some initial (for example, uniform) distribution of weights importance objects W_i is set . The integrated majorant criterion of these weights mismatch and available at every moment submatrix of comparisons matrix is formulated – P.

It is necessary to construct iterative algorithm of ejection of the smallest number of least coordinated pairwise comparisons with simultaneous correction of objects weights W_i so that mismatch criterion P decreases to some acceptable level (probably equal to zero). Final distribution of compared objects weights, received as a result of such adaptive iterative procedure has to be considered as the required solution.

4. METHODOLOGY DESCRIPTION

Before describing offered algorithm, we will give some important definitions.

Definition 1. N – dimension of problem, amount of compared objects $O_i = \{O_1,...,O_N\}$.

Definition 2. $V = \{\upsilon_1,...,\upsilon_N\}$ – set of tops of estimated compared objects communication graph.

Definition 3. $W_i = W(V_i) > 0$ – normalized positive estimate of importance i object (compared objects).

Normalization condition:

$$\sum_i W_i = 1. \quad (1)$$

Definition 4. $H = H(V,V)$ – $N \times N$ dimencionality communication graph.

Whereby $H_{ij} = 1$, if V_i and V_j are connected, $H_{ij} = 0$, if such connection lacks due to lack of pairwise comparisons of O_i and O_j. It is estimated that communication graph H is related graph.

Definition 5. $G = G(V,V)$ – subgraph, is named spanning tree of graph H, if there is only one way from any top V_k to any top V_l.

Naturally quantity of spanning tree edges (quantity of pairwise comparisons) in spanning tree is equal to $N-1$. In general spanning tree is given by “snowflake” (Fig. 1).

Definition 6. To every nonzero communication graph edge G is given estimate S_{ij}, showing pairwise comparisons

$$W_i = S_{ij} \cdot W_j. \quad (2)$$

Please note if spanning tree is chosen then $N-1$ equality (2) and normalization requirement (1) uniquely allow to calculate all W_i. According to found W_i estimations $\widetilde{S}_{ij}$ are defined for other communication graph edges not including into spanning tree. The problem is that these estimations are not the same as pairwise comparisons estimations S_{ij}, given by experts.

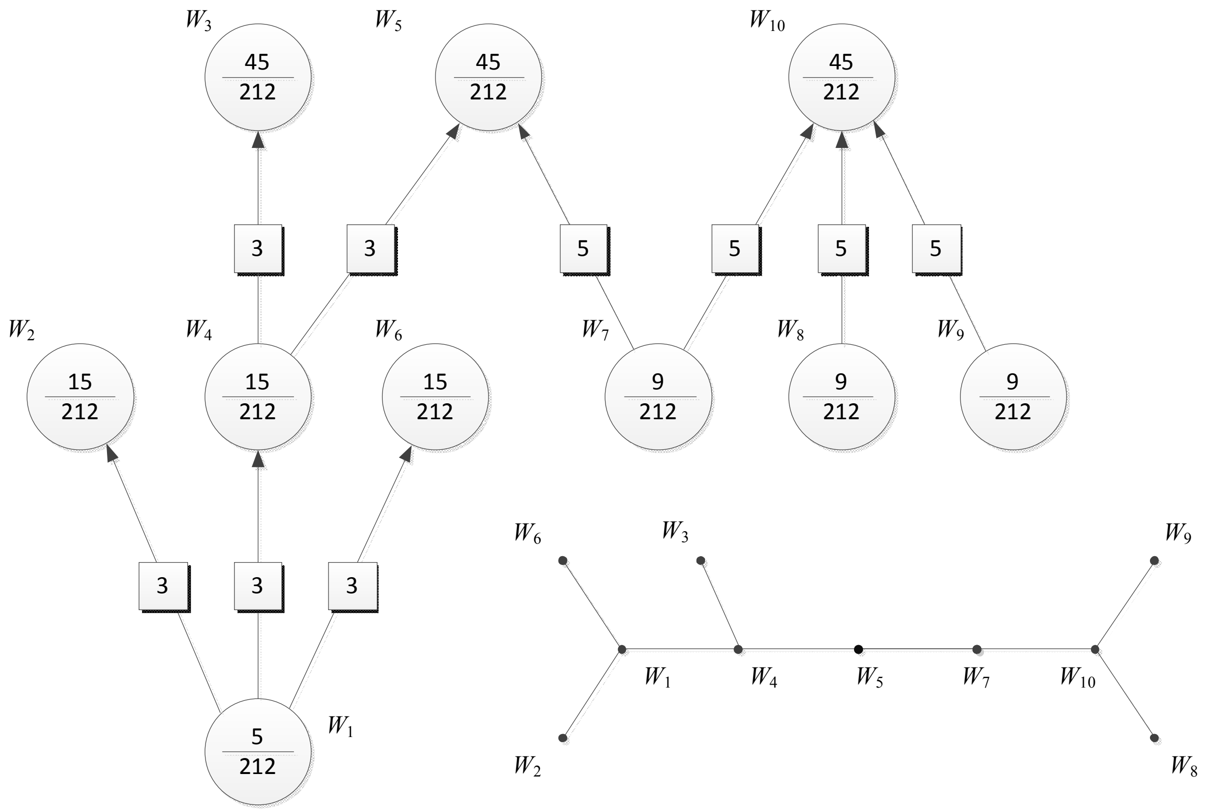


Fig. 1. Spanning tree (in circles – calculated weights of objects importance (tops), in squares – coefficients of pairwise comparisons of conditional example)

If connectivity graph is fully connected graph with N tops, relevant to Saati's method then quantity of spanning tree is estimated by Cayley formula with value N^{N-2} [9]. Thereby search for optimal spanning tree in some multitude similarity metrics $\{\widetilde{S}_{ij}\}$ and $\{S_{ij}\}$ by enumeration presents certain technical problems.

Simultaneously it must be admitted that pairwise comparisons estimations given by experts are rather crude and for high dimensions N «mechanistic». Thereby it is incorrectly to require proximate satisfaction of equalities (2) for all H graph edges, and in this way for edges entered into spanning tree G.

In this context, we take compelled step – we weaken conditions (2), allowing for tops V_i and V_j through S_{ij} mensurate opponent with some default level. Herewith of course it needs to consider that similarly measuring is asymmetric, i.e. the fact is not so critical that the object with big importance gets through another less important object importance estimation which is higher as it is. Also it is not critical when the object with little importance is underestimated because of more important objects.

Thus, for every communication graph edge we shall estimate value

$$f_{ij} = \frac{W_i}{S_{ij} \cdot W_j} - 1 . \tag{3}$$

"Ideally" all f_{ij} must be equal to zero, as for us we require them to be in some window with width $2P$. Here P is important parameter of suggested method which while solution construct shall go to zero (note that zero is realizable if we take decision on solution determination specifically in the shape of spanning tree). In addition, not to lose intergrating characteristics of pairwise comparisons, it is sufficient to concentrate on some level P_f, different from zero a little. In this situation connected subgragh $H(p)$ of graph H will be formed, exhibiting the characteristic when all estimations f_{ij} are set in window with width $2P_j$.

Take into account skewness of estimating f_{ij} for different values we consider two intervals $R(W_i)$ and $L(W_i)$ add up to size P (Fig. 2):

$$L(W_i) + R(W_i) = 2P . \tag{4}$$

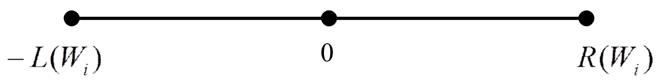


Fig.2.

We will consider that for big W_i values the centre of window displaces to the right of zero and for little values W_i, respectively – to the left of zero. Herewith in "middle" position ($W_i = \frac{1}{N}$) displacement is not occurred. Value of right border $R(W_i)$ under $W_i = 1$ we'll consider equal to $1{,}9P$, left one, respectively $-0{,}1P$. To the contrary for little W_i values right border is "forced against" level $0{,}1P$, and left border stands off zero at a distance of $1{,}9P$.

We suggest plotting correspondence in the shape of hyperbola:

$$R(W_i) = 0{,}1P + 1{,}8P \cdot \frac{(N-1) \cdot W_i}{1 + (N-2) \cdot W_i} . \tag{5}$$

At first for big values P inequations

$$-L(W_i) \le f_{ij} \le R(W_i) \tag{6}$$

are performed for all communication graph edgs.

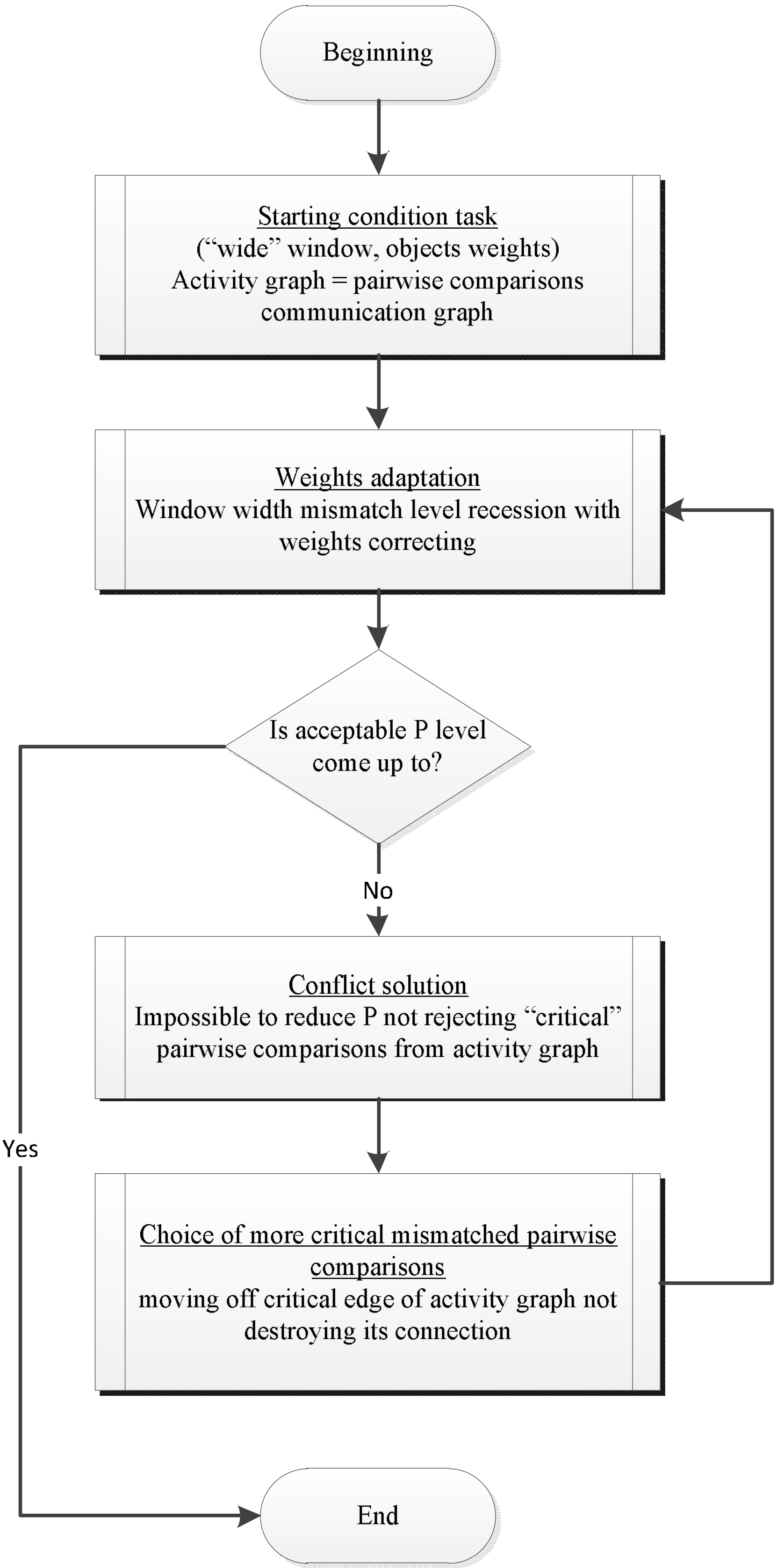


Fig. 3. Bookkeeping scheme algorithm

While reduction P exit vertex of one or several estimations f_{ij} will occur to the borders of windows. While f_{ij} exit to the right border of the window $R(W_i)$ value W_i must be extended, herewith the centre of this window will move to the right as simultaneous reduction of all other weights. While f_{ij} exit to the left border, W_i must be reduced with simultaneous increase of weights of all other tops.

Adaptation of tops weights in G coupling matrix happens. At the cost of this adaptation further reduction of P value contributes, which as the final result ends in situation when all tops W_i have even one edge situated according to its estimation f_{ij} at the border of relevant window. Correlation conflict happens. It is impossible to move even one edge G_{ij} inside window not "pushing out" any other edge G_{kl} in the amount of reached default level P. In this situation one of critical edges must be moved off coupling matrix. It is clear this edge mustn't be spanning, i.e. G graph connectivity mustn't be lost. By experiment we move off such edge G_{ij}, for which another edge G_{ik} exists very close to the border of the window. As a result all f_{ij} adjoined to top V_i constitute content to relevant border of the window, and make it possible to continue adaptation process, changing both value W_i and value W_j, by virtue of the fact that V_j after critical coupling release also became "less bounded".

Omitting detailed descriptions in this article we give bookkeeping scheme algorithm general view of method under discussion (Fig. 3).

5. APPLICATION EXAMPLE

We will consider control-flow chart operation by the example of pairwise comparisons matrix processing of ten objects. Basic data are given in Tab.1.

From 45 pairwise comparisons 18 comparisons are passed (NA). Other 27 estimations indicate preference of objects O_3 and O_5 over other objects, but mismatch of these estimations demanded clarification of the following circumstance: whether these estimations are enough to exclude demonstratively from leaders, let us say, object O_{10}.

We will note that the example is taken from authors' real task decision practice, instead of specially prepared for illustration of offered method opportunities.
So, initially choosing weights of all objects equal 0,1 through an exception of edges with simultaneous reduction of the size of the window P, the decision $W(O)$ provided in Fig. 1 and in two last columns of Tab. 2 (in the form of standard fraction and in a decimal form) is constructed.

In Tab. 2 you can find the summary data illustrating that estimations mismatch on different steps of algorithm differently influences compared objects importance weights ratios. So, in process of critical edges removal defining these edges estimations mismatch with all set of pairwise comparisons which have remained in working subgraph, these weights values execute a periodic motion relating to final solution given in the last column.

It should be noted that coordination of estimates of objects at P=0,742 level (the second column of Tab. 2) and at P=0,469 level (the third column of Tab. 2) demanded removal of mismatched edges, however begin with P=0,207 level all mismatches were found eliminated and further weights corrections were carried out only at the cost of "revaluation of objects importance weights" adaptation mechanism.

Table 1

	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}
O_1		$\frac{1}{3}$	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{1}{7}$	$\frac{1}{3}$	NA	NA	NA	$\frac{1}{4}$
O_2	$\frac{3}{1}$		$\frac{1}{5}$	NA	$\frac{1}{5}$	$\frac{1}{4}$	NA	NA	NA	$\frac{1}{3}$
O_3	$\frac{5}{1}$	$\frac{5}{1}$		$\frac{3}{1}$	NA	$\frac{3}{1}$	$\frac{3}{1}$	$\frac{5}{1}$	$\frac{5}{1}$	$\frac{3}{1}$
O_4	$\frac{3}{1}$	NA	$\frac{1}{3}$		$\frac{1}{3}$	NA	NA	NA	NA	$\frac{1}{3}$
O_5	$\frac{7}{1}$	$\frac{5}{1}$	NA	$\frac{3}{1}$		$\frac{3}{1}$	$\frac{5}{1}$	$\frac{5}{1}$	$\frac{5}{1}$	$\frac{3}{1}$
O_6	$\frac{3}{1}$	$\frac{4}{1}$	$\frac{1}{3}$	NA	$\frac{1}{3}$		NA	NA	NA	$\frac{1}{4}$
O_7	NA	NA	$\frac{1}{3}$	NA	0,2	NA		NA	NA	$\frac{1}{5}$
O_8	NA	NA	$\frac{1}{5}$	NA	$\frac{1}{5}$	NA	NA		NA	$\frac{1}{5}$
O_9	NA	NA	$\frac{1}{5}$	NA	$\frac{1}{5}$	NA	NA	NA		$\frac{1}{5}$
O_{10}	$\frac{4}{1}$	$\frac{3}{1}$	$\frac{1}{3}$	$\frac{3}{1}$	$\frac{1}{3}$	$\frac{4}{1}$	$\frac{5}{1}$	$\frac{5}{1}$	$\frac{5}{1}$	

Table 2

	Window size							
Beca	P=0,742	P=0,469	P=0,207	P=0,1	P=0,05	P=0,01	P=0,0001	P=0
W_1	0,047	0,0447	0,026	0,026	0,025	0,0237	0,0235	$\frac{5}{212}$ = 0,02359
W_2	0,084	0,0923	0,077	0,075	0,073	0,0711	0,0707	$\frac{15}{212}$ = 0,07075
W_3	0,209	0,1735	0,185	0,210	0,211	0,2110	0,2123	$\frac{45}{212}$ =0,21226
W_4	0,084	0,0923	0,077	0,075	0,073	0,0711	0,0707	$\frac{15}{212}$ =0,07075
W_5	0,209	0,1735	0,185	0,210	0,211	0,2110	0,2123	$\frac{45}{212}$ =0,21226
W_6	0,084	0,0923	0,077	0,075	0,073	0,0711	0,0707	$\frac{15}{212}$ =0,07075
W_7	0,047	0,0357	0,045	0,042	0,042	0,0425	0,0425	$\frac{9}{212}$ =0,04246
W_8	0,031	0,0357	0,045	0,042	0,042	0,0425	0,0425	$\frac{9}{212}$ =0,04246
W_9	0,031	0,0357	0,045	0,042	0,042	0,0425	0,0425	$\frac{9}{212}$ =0,04246
W_{10}	0,174	0,2243	0,238	0,203	0,208	0,2135	0,2123	$\frac{45}{212}$ =0,21226

It is also important to note that the first mismatches conflict occurred at P=1,15 level. For overcoming of this conflict communication estimation O_3-O_{10} was removed. Thus succeedent number of removals occurs as if the algorithm in advance knows the solution constructed on spanning tree edges.

To P=0,742, P=0,469 and P=0,207 levels given in Tab. 2 visual coloring of Tab. 3 corresponds, in which different shading show measures of disagreement of pairwise estimations without considering correcting displacements – f_{ij} value, mentioned in hypotheses that objects importance weights are the weights found at $P = P_f$.

Table 3

	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}
O_1		$\frac{0}{1260}$	$-\frac{560}{1260}$	$\frac{0}{1260}$	$-\frac{280}{1260}$	$\frac{0}{1260}$	NA	NA	NA	$-\frac{700}{1260}$
O_2	$\frac{0}{1260}$		$+\frac{840}{1260}$	NA	$+\frac{840}{1260}$	$+\frac{3780}{1260}$	NA	NA	NA	$\frac{0}{1260}$
O_3	$+\frac{1008}{1260}$	$-\frac{504}{1260}$		$\frac{0}{1260}$	NA	$\frac{0}{1260}$	$+\frac{840}{1260}$	$\frac{0}{1260}$	$\frac{0}{1260}$	$-\frac{840}{1260}$
O_4	$\frac{0}{1260}$	NA	$\frac{0}{1260}$		$\frac{0}{1260}$	NA	NA	NA	NA	$\frac{0}{1260}$
O_5	$+\frac{360}{1260}$	$-\frac{504}{1260}$	NA	$\frac{0}{1260}$		$\frac{0}{1260}$	$\frac{0}{1260}$	$\frac{0}{1260}$	$\frac{0}{1260}$	$-\frac{840}{1260}$
O_6	$\frac{0}{1260}$	$-\frac{945}{1260}$	$\frac{0}{1260}$	NA	$\frac{0}{1260}$		NA	NA	NA	$+\frac{420}{1260}$
O_7	NA	NA	$-\frac{504}{1260}$	NA	$\frac{0}{1260}$	NA		NA	NA	$\frac{0}{1260}$
O_8	NA	NA	$\frac{0}{1260}$	NA	$\frac{0}{1260}$	NA	NA		NA	$\frac{0}{1260}$
O_9	NA	NA	$\frac{0}{1260}$	NA	$\frac{0}{1260}$	NA	NA	NA		$\frac{0}{1260}$
O_{10}	$+\frac{1575}{1260}$	$\frac{0}{1260}$	$+\frac{2520}{1260}$	$\frac{0}{1260}$	$+\frac{2520}{1260}$	$-\frac{315}{1260}$	$\frac{0}{1260}$	$\frac{0}{1260}$	$\frac{0}{1260}$	

As is obvious in Tab. 2 – beginning with P=0,1 approximate solutions excursions differ from solution according to spanning tree less than for 1%. This circumstance allowed us to display visually f_{ij} sizes, reduced them to a common denominator.

Excluded at the first stage of algorithm work four pairwise comparisons (from P=1,15 to P=0,742) are painted in dark grey colour (Tab. 3). Four pairwise comparisons removed at the second stage (from P=0,742 to P=0,469) are painted in light grey colour. Two pairwise comparisons removed at the third stage (from P=0,469 to P=0,207), are marked with shading. The cells of Tab. 3 containing remained pairwise comparisons are marked with double shading. The cells containing pairwise comparisons, compatible with spanning tree edges, are painted in black colour.

Let us remark that estimations in table cells with double shading coincide with estimations recalculated through pairwise comparisons estimating, compatible with spanning tree edges. Actually existence of such "duplicating" communications in pairwise comparisons matrix as we understand, allows in the course of solution creation through mismatches conflicts elimination to support main preferences framework. Estimations in cells with double shading are rejected at the last stage at P going to zero – for spanning creation.

6. RESULTS DISCUSSION

Thus, suggested method, without denying AHP, allows to expand its opportunities significantly. At program realization of the algorithm described in the article it is possible to receive express estimations of compared objects importance without completing all pairwise comparisons matrix.

The method allows "to control" AHP matrix completing process as it reveals places of critical experts' estimates mismatch which can be more extensive than mismatch of estimates in threes (as usual among many).

In concept the method allows to work and with uncoordinated matrixes from which they refuse in AHP, in this case the number of iterations of edges exception from graph will naturally increase significantly.

Solution to mismatches conflicts is very similar to *Delphi*[1] method: "ponderability estimation dropping out of general series" – that is on the border – we "cancel" and we try to adapt other estimations to each other in "harder" circumstances.

7. CONCLUSIONS

Thus, we offer the method, allowing to receive compared objects estimations at incomplete comparisons and giving the chance to follow the structure of fundamental comparisons forming objects importance weights.

8. REFERENCES

1. Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs, I. the method of paired comparisons. Biometrika, 39, 324–345.
2. Markin, E.G. (1974). Problem of multiple select. – M.: Nauka (Science), 1974. – p. 256
3. Kendal, M. (1975). Grade correlation. – M.: Statistics, 1975. – p. 216.
4. David, G. (1978). Pairwise comparison method. – M.: Statistics, 1978. – p. 144.
5. Litvak, V. G. (1982). Expert information. Method for obtaining and analysis. – M.: Radio i Svyaz (Radio and communication), 1982. – p. 184.
6. Saati, T. L. (2008). Solution making at correspondences and backlinks: Analytic nets. – M.: Izdatelstvo (Press) LKI, 2008.– p. 360.
7. Totsenko, V. G. (2002). Methods and systems for supporting of solution making. Algorithmic aspect. – Riev: Naukova dumka, 202. – p. 381.
8. Harker, P. T. (1987) Incomplete pairwise comparisons in the analytic hierarchy process, Mathematical Modelling 9 (1987), 837–848.
9. A. Cayley (1889). «A theorem on trees». Quart. J. Math 23: 376-378

ALGORITHMS OF ISOMETRIC SURFACES CONSTRUCTIONS

G. Tsitsiashvili, M. Osipova

•

Russia, Vladivostok, Radio str. 7, IAM FEB RAS
e-mail: guram@iam.dvo.ru , mao1975@list.ru

In this paper a problem of a construction of isometric surfaces is considered. It is important in different technical applications. This problem usually is solved for surfaces with constant Gauss curvature using so called "gold" Gauss theorem. But in practical applications it is interesting to consider more wide class of surfaces. Here this problem is solved for surfaces of cylinder, conical and spherical types and their generalizations using a criterion of isometric based on equalities of coefficients of first quadratic forms.

1. INTRODUCTION

In the differential geometry a concept of the isometric is introduced as follows. Define a surface in three dimensional space E^3 as a set described by equalities

$$X = x(u_1,u_2),\ Y = y(u_1,u_2),\ Z = z(u_1,u_2),\ (u_1,u_2)\in G \subseteq E^2,$$

where the functions x, y, z on the domain G are smooth. Say that the surfaces

$$S_1 = \{X = x_1(u_1,u_2),\ Y = y_1(u_1,u_2),\ Z = z_1(u_1,u_2),\ (u_1,u_2)\in G\},$$

$$S_2 = \{X = x_2(u_1,u_2),\ Y = y_2(u_1,u_2),\ Z = z_2(u_1,u_2),\ (u_1,u_2)\in G\}$$

are isometric if for any pair of points on the surface S_1 (a length of a geodesic curve connected these points) coincides with a distance between appropriate points on the surface S_2. So the surface S_1 may be transformed into the surface S_2 without a dilation and a compression. Last condition has important significance in a solution of modern technical problems including reliability theory also [2].

From the Gauss theorem [1] the surfaces S_1, S_2 are isometric if and only if coefficients of their first quadratic forms coincide:

$$A^1_{i,j} = A^2_{i,j},\ i,j = 1,2, \quad (1)$$

where

$$A^k_{i,j} = \left(\frac{\partial x_k}{\partial u_i},\frac{\partial y_k}{\partial u_i},\frac{\partial z_k}{\partial u_i}\right)\cdot\left(\frac{\partial x_k}{\partial u_j},\frac{\partial y_k}{\partial u_j},\frac{\partial z_k}{\partial u_j}\right),\ i,j,k = 1,2.$$

Here "$\cdot$" is the scalar product operation. But in a literature a problem of the isometric is solved mainly for surfaces with constant Gauss curvature [3, p.126] using so called the "gold" Gauss theorem. This restriction is sufficiently inconvenient in an applied sense and it is worthy to use a direct application of the condition (1) and a search of more general relations for isometric surfaces. In this paper these relations are constructed for surfaces of cylinder, conical and spherical types and for some their generalizations.

2. SURFACE OF CYLINDER TYPE

Define the first quadrant S_1 in the space E^2 by the equalities $x_1 = u_1$, $y_1 = 0$, $z_1 = u_2$ and a surface of a cylinder type S_2 by equalities

$$x_2 = a(u_1),\ y_2 = b(u_1),\ z_2 = u_2.$$

Coefficients of the first quadratic form for the surface S_1 satisfy the relations $A^1_{i,j} = \delta_{i,j}$ where $\delta_{i,j}$, $i, j = 1,2$, is the Kronecker symbol. For the surface S_2 it is not difficult to obtain the equalities

$$A^2_{1,2} = A^2_{2,1} = 0,\ A^2_{2,2} = 1,\ A^2_{1,1} = \left(a(u_1)\right)^2 + \left(b(u_1)\right)^2 = 1.$$

So for the isometric of the surfaces S_1, S_2 it is sufficient to fulfill relations

$$\left(a'(u_1)\right)^2 \le 1,\ \left(b'(u_1)\right)^2 = 1 - \left(a'(u_1)\right)^2. \tag{2}$$

Consequently to solve considered problem for the cylinder surface S_2 it is sufficient using the function $a(t)$ to find the function $b(t)$ from the equality (2).

Example 1. A partial case of this problem solution is the pair of functions

$$a(u_1) = \cos u_1,\ b(u_1) = \sin u_1. \tag{3}$$

If in the first quadrant of E^2 we separate a stripe with sides not parallel to coordinate axes u_1, u_2, then from the equalities (3) it is possible to transform this stripe into spiral surface enveloped around the cylinder S_2.

3. SURFACE OF CONICAL TYPE

Assume that the angle S_1 in the space E^2 is described by the equalities $x_1 = u_2 \cos u_1$, $y_1 = u_2 \sin u_1$. Define the surface of a conical type S_2 by the formulas

$$x_2 = u_2 a(u_1),\ y_2 = u_2 b(u_1),\ z_2 = u_2 c(u_1).$$

Coefficients of the first quadratic form for the surface S_1 satisfy the formulas

$$A^1_{1,1} = u_2^2,\ A^1_{1,2} = A^1_{2,1} = 0,\ A^1_{2,2} = 1.$$

And coefficients of the first quadratic form for the surface S_2 satisfy the relations

$$A^2_{1,1} = u_2^2\left(\left(a'(u_1)\right)^2 + \left(b'(u_1)\right)^2 + \left(c'(u_1)\right)^2\right),\ A^2_{1,2} = A^2_{2,1} = \frac{u_2}{2}\left(a^2(u_1) + b^2(u_1) + c^2(u_1)\right)',$$

$$A^2_{2,2} = a^2(u_1) + b^2(u_1) + c^2(u_1).$$

So to fulfill the equalities in (1) it is sufficient to satisfy the formulas

$$a^2(u_1) + b^2(u_1) + c^2(u_1) = 1, \tag{4}$$

$$\left(a'(u_1)\right)^2 + \left(b'(u_1)\right)^2 + \left(c'(u_1)\right)^2 = 1, \tag{5}$$

Seek a solution of this problem as follows
$a(u_1) = \cos F \cos G$, $b(u_1) = \sin F \sin G$, $c(u_1) = \sin G$, $F = F(s)$, $G = G(s)$, $s = s(u_1)$

where $F(s)$, $G(s)$ are arbitrary smooth functions of an argument s and $s(u_1)$ is a smooth function of the argument u_1. Such a choice of the functions a,b,c leads to the equality (4) independently on a choice of the function s. Simple calculations show that to fulfill the formulas (5) it is sufficient to satisfy the condition

$$1=\left(s'(u_1)\right)^2\left[\left(F'(s)\right)^2+\left(G'(s)\right)^2\right],$$

that leads to the formula

$$u_1=\pm\int\sqrt{\left(F'(s)\right)^2\cos^2 G(s)+\left(G'(s)\right)^2}\,ds. \qquad (6)$$

Example 2. In a partial case it is possible to take real α and put

$$G(s)=\alpha,\ F(s)=s,\ s(u_1)=u_1$$

and so

$$x_2=u_2\cos u_1\cos\alpha,\ y_2=u_2\cos u_1\cos\alpha,\ z_2=\sin\alpha. \qquad (7)$$

The relations (7) define usual conical surface and the formulas (6) specify a wide class of surfaces isometric to an angle in the space E^2.

4. SURFACE OG SPHERICAL TYPE

Suppose that S_1 is defined as follows

$$x_1=a(u_2)\cos u_1,\ y_1=a(u_2)\sin u_1,\ z_1=b(u_2),\ \left(a'(u_2)\right)^2+\left(b'(u_2)\right)^2=1.$$
$$A^1_{1,1}=a^2(u_2),\ A^1_{1,2}=A^1_{2,1}=0,\ A^1_{2,2}=1.$$

Assume that for some real γ the condition $\left(a'(u_2)\right)^2\le 1$ is true. Define the surface S_2 by the equalities

$$x_2=\gamma a(u_2)\cos\frac{u_1}{\gamma},\ y_2=\gamma a(u_2)\sin\frac{u_1}{\gamma},\ z_2=B(u_2).$$

Then the functions

$$A^2_{1,1}=a^2(u_2),\ A^2_{1,2}=A^2_{2,1}=0,\ A^2_{2,2}=\left(B'(u_2)\right)^2+\gamma^2\left(a'(u_2)\right)^2.$$

So the surfaces S_1, S_2 are isometric if

$$\left(B'(u_2)\right)^2+\gamma^2\left(a'(u_2)\right)^2=1,$$

consequently

$$B(u_2)=\int\sqrt{1-\gamma^2\left(a'(u_2)\right)^2}\,du_2.$$

Example 3. Assume that the equality $a(u_2)=\cos u_2$ is true and so the surface S_1 is a sphere with unity radius and with a cut along a meridian. Then the surface S_2 may coincide with a bending of a sphere by a type [3,Figure 3]: spindled surface of a rotating with a self overlap and etc
Consider possible generalizations of isometric surface with a spherical type. Suppose that smooth functions $F(s)$, $G(s)$ satisfy the condition

$$\left(F'(s)\right)^2+\left(G'(s)\right)^2=1 \qquad (8)$$

and $f(u_1,u_2)$ is smooth function. Define the surface S_1 as follows

$$x_1 = F\left(f(u_1,u_2)\right),\ y_1 = G\left(f(u_1,u_2)\right),\ z_1 = R(u_2), \tag{9}$$

then the functions

$$A_{1,1}^1 = \left(\frac{\partial f(u_1,u_2)}{\partial u_1}\right)^2,\ A_{1,2}^1 = A_{2,1}^1 = \frac{\partial f(u_1,u_2)}{\partial u_1}\cdot\frac{\partial f(u_1,u_2)}{\partial u_2},\ A_{2,2}^1 = \left(\frac{\partial f(u_1,u_2)}{\partial u_2}\right)^2 + \left(R'(u_2)\right)^2.$$

Take a real number γ and define the surface S_2 by the equalities

$$x_2 = \gamma F\left(\frac{f(u_1,u_2)}{\gamma}\right),\ y_2 = \gamma G\left(\frac{f(u_1,u_2)}{\gamma}\right),\ z_2 = R(u_2). \tag{10}$$

Then the functions

$$A_{1,1}^2 = \left(\frac{\partial f(u_1,u_2)}{\partial u_1}\right)^2,\ A_{1,2}^2 = A_{2,1}^2 = \frac{\partial f(u_1,u_2)}{\partial u_1}\cdot\frac{\partial f(u_1,u_2)}{\partial u_2},\ A_{2,2}^2 = \left(\frac{\partial f(u_1,u_2)}{\partial u_2}\right)^2 + \left(R'(u_2)\right)^2$$

and so the relation (1) is true. Consequently the surfaces S_1, S_2 defined by the formulas (8) - (10) are isometric.

REFERENCES

1. Andreev G.N. 2007. Additional chapters of geometry. Differential geometry of curved surfaces. Moscow.: MGIU. (In Russian).
2. Ushakov I.A. 2009. History of science through prism of inspiration. Glamour of geometry. Book 3. Moscow: URSS. (In Russian).
3. Fomenko V.T. 1998. Bending of surfaces// Soros educational journal. Vol. 5. P. 122-127. (In Russian).

FEASIBILITY OF OPERATING CONDITIONS AND THE LOCATION OF SENSOR VARIABLES IN THE ELECTRIC POWER SYSTEM

Irina I. Golub, Oleg N. Voitov, Evgeny V. Boloev[1]

•

Energy Systems Institute of SB RAS,
Irkutsk, Russian Federation

e-mail: golub@isem.sei.irk.ru

ABSTRACT

In the paper the methods of probabilistic load flow, including linear method of generalized disturbance, are used to detect sensor variables in an electric power system, identify their probabilistic characteristics, detect critical variables, for which the probability that they lie in the feasible range is lower than the required one, and select the control actions to increase this probability. The controls are selected by the method similar to the method of deterministic equivalent by subsequently and iteratively solving deterministic and probabilistic problems. The method of contribution factors makes it possible to choose from a set of possible controls a vector of control that contains the minimum number of components. The presented numerical results on the example of test and real networks demonstrate the efficiency of the proposed approaches to the electric power system operation control.

1. INTRODUCTION

State variables of the electric power system should lie within certain feasible limits, which meet reliability and quality requirements. The probability that the variable will go beyond its feasible limits in case of disturbances depends on the sensitivity of the variable to disturbances, a feasible range of its variation, and proximity of its current value to the limit.

Such network elements whose state variables change largely due to random external disturbances are called sensors [1]. Inhomogeneity of the electric power system that leads to the emergence of sensors is determined by both the operating conditions of the system and parameters of the network elements that are called weak places in [2].

Two approaches can be used to detect sensors and weak places. The first approach is related to the analysis of disturbance scenarios and responses of variables to the disturbances that are expressed, for example, by deviations of variables.

In the first approach the sensitivity of a variable to external disturbances can be determined on the basis of singular or spectral decomposition of the Jacobian matrix. If the spread of singular values (eigenvalues) is large, it indicates severe inhomogeneity of the electric network. The greater the difference between the first singular value of the Jacobian matrix and the rest of the values, the more reasons to make a conclusion about the behavior of the variables on the basis of the first singular decomposition summand connected with the minimum singular value.

[1] The study has been supported by grant 1507.2012.8 Leading Scientific School.

The idea of using the maximum components of the eigenvector that correspond to the minimum eigenvalue to choose the most informative set of measurements [3] was applied to detection of sensor nodes in an electric network by means of the spectral analysis in [1], [4] and singular analysis in [5].

The variables most sensitive to external disturbances can be identified by using a scalar value of the first generalized disturbance. The greatest contribution to the first generalized disturbance is made by the disturbances at nodes corresponding to the components of the left singular vector, while the components of the first right singular vector distribute the generalized disturbance among the nodes of the electric network.

In the second approach used in this study the response of a random variable to disturbances can be determined by the methods of probabilistic load flow, including the linear analytical method which uses the scalar variance value of the generalized disturbance. The numerical characteristics of the variables obtained in the calculation allow us to find the probability that the variables are feasible.

A large response of variables to a disturbance is significant in the case if it changes some criterion of power system operation, for example the criterion of operation feasibility which is considered in this study.

To increase the probability that the variable lies in the feasible region it is necessary either to reinforce the network, which will improve the Jacobian matrix conditioning, reduce the response of the variables to the disturbance, and expand the feasible region, or to find the appropriate controls to make the variable mean value shift inside the feasible region.

The required probability is provided by iteratively and consecutively solving the problem of probabilistic load flow and the problem of determining feasible operating conditions by the deterministic equivalent method, which implies searching for a feasible solution with a shift of the critical variable mean value inside the feasible region. The critical variables are the variables for which the probability to fall within the specified interval is less than the required one. If the solution exists, a control vector with the minimum number of components is chosen from a set of possible controls, using the method of contribution factor [6], and the required increment in the control vector is determined, which increases the probability that the critical variable lies in the feasible region.

Analysis of contemporary methods for the calculation of probabilistic load flow and their use to solve various power engineering problems is presented in the overviews [7] and [8].

The methods of probabilistic load flow can be divided into the methods of linear and nonlinear approximation, which can be both iterative and noniterative, and numerical methods. The noniterative linear methods include the method of moments [9] that are formed on the basis of the Jacobian matrix, and the method of convolution [10]. The foundations of the linear iterative method that was called the method of statistical linearization were developed in [11] and successfully applied in [12]. In the noniterative [13] and iterative [8] nonlinear methods the moments are formed on the basis of the Jacobian and Hessian matrices.

The method of Monte Carlo and point methods [14], in which ordinary programs for deterministic load flow are used for the calculations of probabilistic load flow are referred to the numerical methods.

The study to be mentioned among the first to consider constraints in the calculation of probabilistic load flows is [15]. The need to solve the indicated problem by the procedure for calculating constrained optimal power flow has resulted in the development of the methods that combine deterministic and probabilistic approaches. The theoretical foundations of this approach called the method of deterministic equivalent are presented in [16].

2. DETECTION OF SENSOR VARIABLES BY THE METHODS OF PROBALISTIC LOAD FLOW

The mean square deviations of nodal voltage magnitudes and phases can be determined by the specified mean square deviations of nodal powers in the linear analytical method, using the expression relating the changes in phases $\Delta\delta$ and magnitudes ΔU of nodal voltages and the changes in active ΔP and reactive ΔQ powers in the system of linear equations

$$\begin{pmatrix}\Delta\delta\\ \Delta U\end{pmatrix} = J^{-1}\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix}, \qquad (1)$$

where J – Jacobian matrix.

The means $m_{\Delta\delta,\Delta U}$ and covariances $\mu_{2\Delta\delta,\Delta U}$ of changes in the magnitudes and phases of voltages are determined through the means $m_{\Delta P,\Delta Q}$ and variances of loads $\mu_{2\Delta P,\Delta Q}$ at the point of solution to the nonlinear system of steady state equations of electric power system as

$$m_{\Delta\delta,\Delta U} = J^{-1} m_{\Delta P,\Delta Q}, \qquad (2)$$

$$\mu_{2\Delta\delta,\Delta U} = J^{-1}\mu_{2\Delta P,\Delta Q}\left(J^{-1}\right)^{T}. \qquad (3)$$

The expressions of the means and covariances of changes in the active and reactive power flows as well as differences of voltage magnitudes and phases can be written in an analogous way. In particular for covariances they will have the form

$$\mu_{2\Delta P_{ij},\Delta Q_{ij}} = J_{ij}J^{-1}\mu_{2\Delta P,\Delta Q}\left(J_{ij}J^{-1}\right)^{T}, \qquad (4)$$

$$\mu_{2\Delta(\delta_i-\delta_j),\Delta(U_i-U_j)} = M^{T}J^{-1}\mu_{2\Delta P,\Delta Q}\left(J^{-1}\right)^{T}M, \qquad (5)$$

where J_{ij} – matrix of partial derivatives of active and reactive power flows in the tie ij with respect to magnitudes and phases of nodal voltages, M – the incidence matrix.

The numerical characteristics of loads under the assumption about their normal distribution can be obtained by using the Laplace function [9], also called the error function.

The linear method of generalized disturbance is based on the combination of linear analytical method (2), (3) with the method of singular analysis which implies singular decomposition of the asymmetrical Jacobian matrix

$$J = W\Sigma V^{T} = \sum_{j=1}^{n} w_j\,\sigma_j\,\mathrm{v}_j^{T}, \qquad (6)$$

where $W = (w_1, w_2, ..., w_n)$ and $V = (v_1, v_2, ..., v_n)$ – orthogonal matrices, whose columns represent left and right singular vectors, and Σ – diagonal matrix of singular values $\sigma_1 < \sigma_2 < \sigma_3 < ... < \sigma_n$, arranged in the ascending order.

Taking into consideration decomposition (6), expression (1) can be written in the form

$$\begin{pmatrix}\Delta\delta\\ \Delta U\end{pmatrix} = J^{-1}\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix} = \sum_{i=1}^{n} v_i\,\frac{w_i^{T}}{\sigma_i}\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix} = \sum_{i=1}^{n} v_i\,\Delta S^{(i)}, \qquad (7)$$

where $\Delta S^{(i)}$ – the i-th generalized disturbance.

If the first singular value $\sigma_1 = \sigma_{min}$ is considerably lower than the rest of the singular values the largest contribution to the changes in phases and magnitudes of nodal voltages is made by the first term of the sum (7)

$$\begin{pmatrix}\Delta\delta\\ \Delta U\end{pmatrix}^{(1)} = v_1\,\frac{w_1^{T}}{\sigma_1}\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix} = v_1\,\Delta S^{(1)}. \qquad (8)$$

The first generalized disturbance $\Delta S^{(1)}$ is the same for all nodes and is distributed among state variables in proportion to the components of the first right singular vector v_1. In this case the maximum contribution to the first generalized disturbance is determined by the disturbances at the nodes that correspond to the maximum components of the first left singular vector w_1.

The number of disturbance variants is infinite. They may differ both in composition and in value. The linear method of generalized disturbance does not require the scenario of change in the nodal powers to be specified and makes it possible to assess a set of disturbance scenarios by using the specified value of covariance $\mu_{2\Delta S^{(1)}}$ of the generalized disturbance. The most significant disturbances are the generalized disturbances that correspond to the first minimum

$$\mu^{(1)}_{2\Delta\delta,\Delta U} = \frac{\nu_1 w_1^T}{\sigma_1} \mu_{2\Delta P,\Delta Q} \left(\frac{\nu_1 w_1^T}{\sigma_1} \right)^T = \nu_1 \mu_{2\Delta S^{(1)}} \nu_1^T \tag{9}$$

or k minimum singular values that are close to one another

$$\mu^{(k)}_{2\Delta\delta,\Delta U} = \sum_{i=1}^{k} \nu_i \ \mu_{2\Delta S^{(i)}} \nu_i^T \ . \tag{10}$$

Similar expressions including the variance of generalized disturbance can be written for the numerical characteristics of changes in power flows, differences of nodal voltage magnitudes and phases.

3. PROBABILISTIC CONSTRAINED LOAD FLOW

A strong response of the sensor variable to the disturbance is not dangerous in itself if the variable remains within feasible limits after the disturbance. Therefore, the most important indicator is the required value of probability that the random value lies in the specified feasible range.

In the case that the calculation of probabilistic load flow results in critical variables for which the probability to fall within the specified interval is less than the required one, the probability can be increased either by decreasing the mean square deviation of the variable or by shifting its mean value inside the feasible region.

The mean square deviation of the critical variable being also a sensor variable, can be decreased for example by the reinforcement of weak ties [1], [4]. Another possibility is to choose the control actions that decrease the distance between the mean value and median of the distribution density curve $f(x)$ on the feasible interval.

Such a criterion is used in the case where variable x has the law of distribution other than the normal law, and the approximated probability density curve can be obtained by three or a greater number of moments, using the Gram- Charlie expansion [9]

$$f(x) = \sum_{j=0}^{\infty} c_j H_j(x) \phi(x), \tag{11}$$

where $\phi(x)$ – probability density for normal distribution, $H_j(x)$ – orthogonal Hermite polynomials, c_j – coefficients built on the basis of the second- and higher-order moments.

The probability density curve when approximated on the basis of two moments is symmetrical. This makes it possible to transform the criterion for the selection of controls into the minimization of distance between the mean value of the variable and the center of a feasible interval of its change towards the point of the required probability value. When the constraints on the variable are specified symmetrically with respect to its nominal value the center of the interval is the nominal value of the variable.

In order to choose the controls to provide the required probability for the critical variables to lie within the feasible limits, the method similar to the method of deterministic equivalent [15], [16] is used. In this method successively deterministic and probabilistic problems are solved. Solving the deterministic problem suggests shifting the mean value to the center of the feasible interval but not narrowing this interval for each critical variable as it is done in the method of deterministic equivalent.

The variable mean value to be obtained as a result of control can be determined by using the inverse error function. This function allows one to determine the interval $\Delta\varepsilon$ of change in the normally distributed random variable, which makes it possible at a specified value of mean square deviation to provide the required probability that this variable falls within this interval.

The interval $\Delta\varepsilon$ is compared to the known feasible interval $\Delta\varepsilon_{feasible}$ of the variable change. If $\theta = (\Delta\varepsilon_{feasible} - \Delta\varepsilon) > 0$, then the mean of the variable should equal θ, otherwise a conclusion about the impossibility of providing the required probability that the variable lies within the feasible interval $\Delta\varepsilon_{feasible}$ and the need to shift the mean value of the variable to the center of the feasible interval is made. In the first situation the required shift of the mean $\mu_{\Delta z}$ of variable Δz will equal $\Delta_z = \theta - \mu_{\Delta z}$, and in the second situation $\Delta_z = \mu_{\Delta z}$.

An algorithm for increasing the probability that the variables fall within the feasible limits is iterative, its each k -th iteration contains the following main steps.

1. The deterministic problem is solved to obtain feasible operating conditions of electric power systems subject to

$$W(X,Y)=0, \tag{12}$$

$$X_{min} \le X \le X_{max}, \tag{13}$$

$$F_{min} \le F(X,Y) \le F_{max}, \tag{14}$$

$$Y_{min} \le Y \le Y_{max}, \tag{15}$$

where (12) – system of equations of nodal power balances, (13)–(15) – constraints on the dependent variables X, that include magnitudes and phases of nodal voltages and functional variables F, that contain active and reactive power flows; (15) – constraints on controls or independent variables Y, such as active and reactive power of generation and transformation ratios of tap-changing transformers.

Problem (12) to (15) is solved by combining the reduced gradient and quadratic programming methods [15], and if there exists a feasible solution for vectors X^k and Y^k then the algorithm goes to step 2. Otherwise, the algorithm stops.

2. The probabilistic load flow is calculated, the numerical characteristics of variables and probability of meeting the constraints (13)–(15) are determined. If for all the variables the required value of probability is provided, the algorithm stops. Otherwise, the number N_v of critical variables z_j^k is determined for which the required probability value is not provided and an estimate of the shift $\Delta_{z_j}^k$ of its mean which leads to an increase in the probability is calculated. If in the adjacent iterations for each critical variable z_j^k the condition $\left|\Delta_{z_j}^{k+1} - \Delta_{z_j}^k\right| \le \xi$ is met, where ξ – a set small number, the required probability values cannot be reached and the algorithm stops. Otherwise, the algorithm goes to step 3.

3. The deterministic optimization problem is solved to determine the vector of control $Y_* = Y^k + \Delta Y^k$, that provides the minimum of the criterion

$$min \sum_{j=1}^{N_v} \left(z_j(Y) - \left(z_j^k\left(Y^k\right) + \Delta_{z_j}^k\right)\right)^2, \tag{16}$$

when the constraints (12)–(15) are met. If a solution to the problem is found and criterion (16) equals zero, the algorithm goes to step 4, otherwise, $Y^{k+1} = Y_*$, $k = k + 1$, and it goes to step 2.

4. The minimum number of controls ΔY^k is chosen on the basis of the contribution factors method [6]. This method allows the identification of the ways of transmitting active and/or reactive power from generator nodes to load nodes and the contribution of generator power to the power of flows and loads. The information about tracing the flows is used to determine the so called significant controls that affect the critical variable to the greatest extent. The significant controls underlie the formation of variants with different number of controls ΔY^k, for each of which the solution to problem (16), (12)–(15) is searched for. When comparing the variants, the variants with the minimum number of controls are chosen. If there are several variants with equal number of controls, then the variant with the minimum active power losses is taken. Then the vector $Y^{k+1} = Y^k + \Delta Y^k$, $k = k+1$ is determined, and the algorithm goes to step 2.

3. CASE STUDY

The electric power system presented in Figure 1, which consists of 14 nodes and 15 ties, is used as a test scheme. The performance of the considered methods for this scheme is illustrated by an example of the nodal voltage magnitude control.

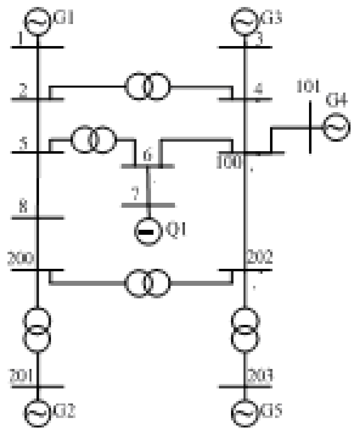


Figure 1. Scheme of a 14-node test network

The initial data on mean values and variances for the loads specified at all the nodes were obtained with the use of the Laplace function. Mean square deviations of nodal powers were assumed to be equal to 12 % of their mean values, which corresponds to 20 % of the load forecast error for a 0.9 probability of random value deviation from the mean.

Figure 2 presents the graphs of the mean square deviations of nodal voltage magnitudes obtained by the linear method, generalized disturbance method, and the Monte Carlo method. The graphs show that node 8 is the sensor node.

The conclusion that node 8 is the node with a sensor voltage magnitude can also be made on the basis of the singular analysis technology [5]. For this purpose the nodes can be projected on the plane in coordinates of the first and second singular vectors. The sensor nodes in such a graph will have maximum distance from the origin of coordinates. After interconnecting the nodes by the ties, the network graph projection in coordinates of the first and second right singular vectors is obtained, Figure 3.

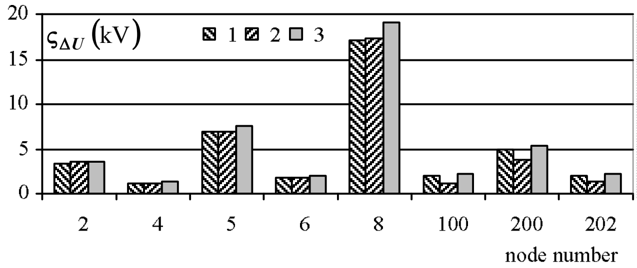


Figure 2. Mean square deviation of voltage magnitudes at the nodes of the test scheme obtained by the linear method – 1, generalized disturbance method – 2, the Monte Carlo method – 3

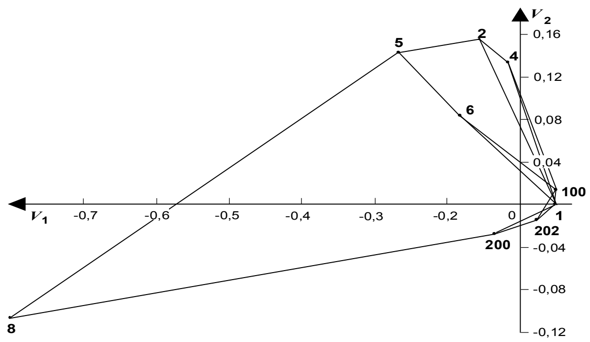


Figure 3. Projection of the network graph in coordinates of the first (***v*₁**) and second (***v*₂**) right singular vectors that correspond to voltage magnitudes

However, such a technology, unlike the probabilistic load flow, does not allow simultaneous identification of sensor variables and assessment of their possible variation ranges and probabilities that the variables lie within the feasible limits.

Table 1 contains the mean values and mean square deviations of voltage magnitudes obtained by the linear method of probabilistic load flow, differences between the mean values and nominal voltages, and the probabilities that voltage magnitudes fall within the feasible intervals. The feasible intervals for 500 kV voltages are taken equal to ±30 kV, and for 220 kV – ±25 kV.

Table 1. Probabilistic characteristics of voltage magnitudes at the test network nodes for the initial state

Nodes	m_U	ς_U (kV)	$m_U - U_{nom}$ (kV)	p
2	522.34	3.65	22.34	0.98
4	231.49	1.32	11.49	1.00
5	512.05	6.88	12.05	0.99
6	225.17	1.86	5.17	1.00
8	508.44	17.12	8.44	0.88
100	229.24	2.03	9.24	1.00
200	528.15	4.83	28.15	0.64
202	233.62	2.05	13.62	1.00

If the required value of probability that voltage magnitudes lie within the given intervals should be not less than 0.95, the voltage magnitudes at nodes 200 and 8 can be defined as critical. Another estimate of critical voltage magnitudes ΔU_i can be represented by the maximum ratio of mean square deviation $\varsigma_{\Delta U_i}$ to the feasible variable variation range determined by the proximity of the variable mean to its upper $\overline{\Delta U_i}$ or lower $\underline{\Delta U_i}$ limiting value $\phi_{\Delta U_i} = \varsigma_{\Delta U_i} / min\left(\overline{\Delta U_i} - \mu_{\Delta U_i}, \mu_{\Delta U_i} - \underline{\Delta U_i}\right)$.

Such a possibility is illustrated in Figure 4 which shows the values of components of vector $\phi_{\Delta U_i}$ and the values of probabilities that voltage magnitudes lie within the feasible limits. Under the initial operating conditions critical node 200 corresponds to the maximum value $\phi_{\Delta U_i}$ and minimum probability. At the nodes with the 0.98–1.0 probability that voltage magnitudes are within the feasible limits, the value of criterion $\phi_{\Delta U_i}$ does not exceed 0.5.

To provide the required probability that voltage magnitudes at nodes 8 and 200 fall within the feasible limits the two methods including the reinforcement of weak ties and selection of control actions to move the mean values of variables to the center of the feasible interval were compared. Weak ties are the ties, in which the reduction in resistances increases the minimum singular value of the Jacobian matrix, i.e. improves its conditionality and decreases the response of sensor variables to disturbances [5].

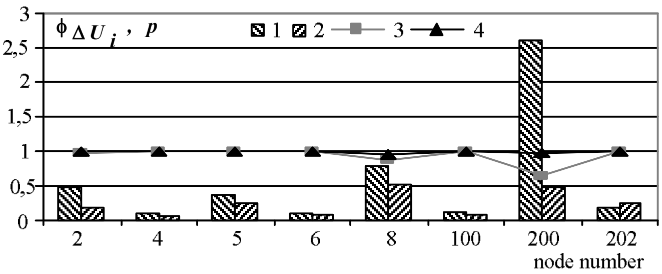


Figure 4. Values of index $\phi_{\Delta U_i}$ – 1, 2 and probability p that voltage magnitudes are within the feasible limits – 3, 4, for the initial operating conditions 1, 3 and the conditions obtained as a result of the network reinforcement and selection of control actions 2, 4

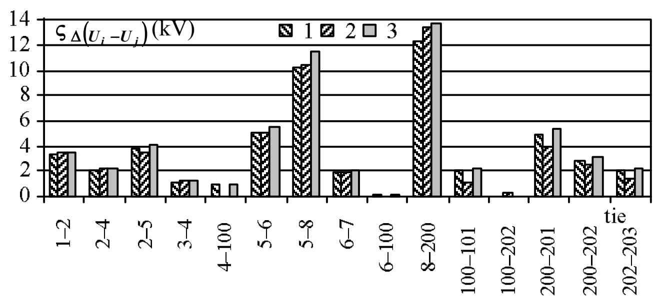


Figure 5. Mean square deviations of voltage magnitude differences in the ties of the scheme obtained by the linear method –1,the generalized disturbance method –2, the Monte Carlo method – 3

For the criterion used to detect weak ties the maximum values of the mean square deviations of changes in the voltage magnitude differences Figure 5 obtained by the linear method, the generalized disturbance method, and the Monte Carlo method are applied. According to these values, ties 5–8 and 8–200 in the test network Figure 1 are weak. The projection of the network graph Figure 3 shows that the weak ties are the longest.

At the initial point the mean square deviation of the voltage magnitude at node 8 amounts to $\varsigma_{\Delta U_8} = 17.12$ kV. This means that with the probability of 0.95 the voltage will be in the interval $\Delta\varepsilon = \pm 33.54$ kV. In the feasible interval of voltage changes equal to $\Delta\varepsilon_{feasible} = \pm 30$ kV, the mean of the variable will be $\theta = (30 - 33.54) = -3.54 < 0$, whence it follows that the mean should be moved to the center of the feasible interval. This will make it possible to provide the 0.9203 probability that the voltage falls in the feasible interval.

To provide the 0.99 probability that the voltage of node 200 with mean square deviation equal to $\varsigma_{\Delta U_{200}} = 4.83$ kV lies in the feasible interval, the centered value of the mean should equal $\theta = (30 - 11.77) = 18.23 > 0$ and the shift of the mean will be $\Delta_{\Delta U_{200}} = \theta - \mu_{\Delta U_{200}} = 18.23 - 28.15 = -9.92$ kV.

Since the critical variables are represented by voltages, their values are changed by the reactive power sources and transformers.

The test scheme has 11 controls with 6 reactive power sources at nodes 1, 3, 7, 101, 201, 203 and 5 tap-changing transformers in ties 2–4, 5–6, 200–201, 200–202, and 202–203.

The contribution factors method [6], [17] was used to determine the controls that make the required shifts of voltage values at nodes 8 and 200 in the test scheme by investigating the reactive power flows coming to the specified nodes, Figure 6.

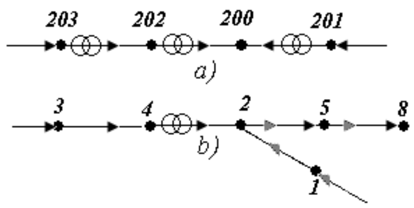


Figure 6. Directions of reactive power flows arriving at nodes 8 and 200 with critical voltage magnitudes

The analysis of the flow directions makes it possible to find the significant controls that include four sources of reactive power at nodes 1, 3, 201 and 203 and four tap-changing transformers 2–4, 200–201, 200–202 and 202–203.

The determined controls are the basis for the calculation of vector ΔY^k with the minimum number of controls that include the reactive power of node 203 and the transformation ratios of transformers 200–201, 200–202 and 2–4.

Figure 7 presents the distribution density curves of changes in voltage magnitudes of nodes 8 and 200 that were obtained for the following operating conditions: initial operating conditions (1, 5); operating conditions at reinforcement of weak ties (2, 6); operating conditions after the implementation of determined controls (3, 7); operating conditions under simultaneous implementation of the chosen controls and reinforcement of weak ties (4, 8).

An 11% decrease in the resistance of determined weak ties resulted in both a reduction in the mean square deviation of voltages at nodes 8 and 200 and a greater shift in the mean values with respect to the nominal voltage. Consequently, the probability that the voltage magnitudes at nodes 8 and 200 fall within the feasible limits decreased, as compared to the initial probability.

The implementation of control actions related to the change in the source reactive power at node 201 and to the regulation of transformation ratios of transformers, and the obtained shift in the mean values of voltage magnitudes at nodes 8 and 200 make it possible to increase the probability that they lie within the feasible limits.

Owing to the reduction in the mean square deviation and the shift in the mean, simultaneous implementation of control actions and reinforcement of weak ties increased the probability that the voltage magnitude at node 8 falls within the feasible limits. Graphs 2 and 4 in Figure 4, which correspond to the values of index $\phi_{\Delta U_i}$ and the probability that voltage magnitudes are within the feasible limits, show that in this case both indices attest to the absence of critical voltage magnitudes.

Let us illustrate the operation of the algorithms with an example of the real electric network consisting of 207 nodes and 224 ties. The projection of the nodes and ties of this network on the plane in coordinates of the first and second singular vectors corresponding to nodal voltage magnitudes is presented in Figure 8. Voltage magnitudes at nodes 77, 78, 12, and 17 (110 kV) are sensor variables. The feasible interval for 110 kV voltages is taken equal to ±10 kV.

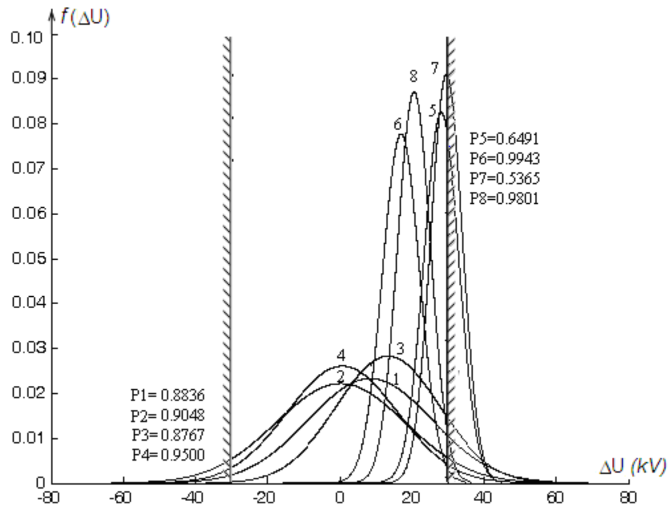


Figure 7. Probability density curves of changes in voltage magnitudes at nodes 8 and 200 for the initial operating conditions (1, 5), operating conditions after the reinforcement of weak ties (2, 6), operating conditions after the implementation of control actions (3, 7), and operating conditions under simultaneous implementation of control actions and reinforcement of weak ties (4, 8).

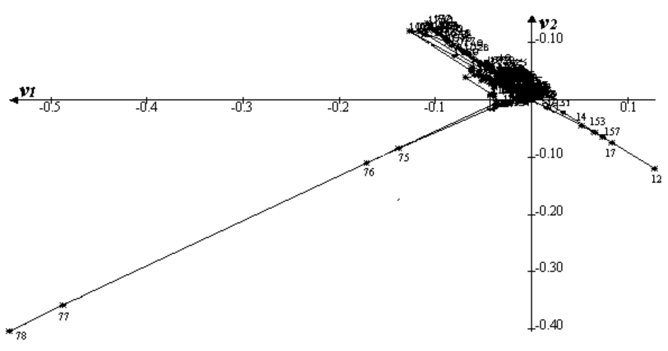


Figure 8. Projection of the real network graph on the plan in coordinates of the first ($\boldsymbol{v}_1$) and second ($\boldsymbol{v}_2$) right singular vectors that correspond to voltage magnitudes

Table 2 shows that the same order of the nodes with sensor voltage magnitudes is obtained by the linear method of probabilistic load flow for the initial and end states. In the initial state the critical variable is only the voltage magnitude at node 12, whose probability of lying within the feasible limits is close to zero. After the shift of the mean, the end state with the 0.95 probability for the critical value is determined. Such a result is obtained using only one control action found by the contribution factors method. The full control vector includes 93 components.

Table 2. Probabilistic characterictics of voltage magnitudes at the sensor nodes of real network for the initial (1) and end (2) states

Nodes	ς_U (kV)		$m_U - U_{nom}$ (kV)		p	
	1	2	1	2	1	2
78	1.267	1.264	–5,69	–5,56	0,999	0,999
77	1,189	1,186	–4,24	–4,12	1	1
12	0,923	0,877	–11,96	–8,55	0,017	0,951
17	0,621	0,594	–7,21	–4,71	1	1
157	0,601	0,573	–7,41	–4,25	1	1

3 CONCLUSIONS

1. The methods of probabilistic load flow allow the detection of those sensor variables in the electric power system which can be detected on the basis of singular analysis.
2. A combination of the analytic probabilistic method with the scalar value of the first generalized disturbance is suggested to obtain probabilistic indices of variables in an inhomogeneous network.
3. An approach is proposed to solve the problem of selection of the control actions which provide the required probability that the controlled sensor variables fall within the feasible limits.
4. A method is suggested to search for a solution to the control problem with minimum number of controls on the basis of the data on tracing the power flows.

REFERENCES

[1] A.Z. Gamm, I.I. Golub, “Detection of weak places in a power system”, Izv. RAS, Energetika, No.3, 1993, pp. 83-92, (*in Russian*).
[2] Gamm A.Z., Golub I.I., Observability of electric power systems, Novosibirsk: Nauka, 1990, p. 220, (*in Russian*).
[3] Fetzer E.E., Anderson P.M.,“Observability in the state estimation of power systems”, IEEE Trans. on Power Apparatus and Systems, Vol. 94, No.6, 1975, pp. 1981-1988.
[4] Gamm A.Z., Golub I.I., “The problem of weak places in electric power systems”, in Proc. 1995 International Conference on Electrical Power Engineering Power Tech, Stockholm, pp. 542-546.
[5] Gamm A.Z., Golub I.I., Bachry A., Styczynski Z.A.,”Solving several problems of power systems using spectral and singular analyses”, IEEE Trans. Power Syst., Vol. 20, No.1, 2005, pp. 138-148.
[6] Bialek J., “Tracing the flow of electricity”, IEE Generation Transmission &Distribution, Vol.143, No.4, 1996. pp. 313-320.
[7] Chen P., Chen Z., Bak-Jensen B., Probabilistic Load Flow: A Review http://vbn.aau.dk/ws/fbspretrieve/16272096/Probabilistic_load_flow_a_review.pdf
[8] Manusov V.Z., Mogirev V.V., Shepilov O.N., “Study of operating conditions of the USSR Unified Power System, considering the random character of initial data”, Elektrichestvo, No. 10, 1983, pp. 3-6, *(in Russian)*.
[9] Pugachev V.S., Theory of probabilities and mathematical statistics, M.: Nauka, 1979, p. 496, *(in Russian)*.

[10] Borkowska B., “Probabilistic Load Flow”, IEEE Trans. on Power Apparatus and Systems, Vol. 93, No. 3, 1974, pp. 752-755.
[11] . Kazakov I.E, Dostupov B.G., Statistical dynamics of nonlinear automatic systems, M.: Fizmatgiz, 1962, p. 332, *(in Russian)*.
[12] Manusov V. Z., Lykin A.V., “Probabilistic analysis of steady states of electric power systems”, Elektrichestvo, No.4, 1981, pp. 7-13, *(in Russian)*.
[13] Chen X. Li, X., Yin X., Xiang T., H. Liu, The Algorithm of Probabilistic Load Flow Retaining Nonlinearity, Proc. of 2002 Power Con, Int. Conf. on Power System Technology, Kunming, Vol. 4, 2002, pp. 2111-2115.
[14] Su C. L., “Probabilistic load-flow computation using point estimate method”, IEEE Trans. Power Systems, Vol. 20, No. 4, Nov. 2005, pp.1843-1851.
[15] Murashko N.A., Okhorzin Y.A., Krumm L.A. et al, Analysis and control of steady states of electric power systems, Novosibirsk: Nauka, 1987, p.240, *(in Russian)*.
[16] Yudin D.B., Mathematical methods of control under the conditions of incomplete information, M.: Sov. radio, 1974. p.400, *(in Russian)*.
[17] Gamm A.Z., Golub I.I., Application of Contribution factors at control of electric power system operating conditions, in Proc. 2005 International Conference on Electrical Power Engineering Power Tech, St. Petersburg, Rep. 649.

REACTIVE CAPACITY COMPENSATION AND VOLTAGE REGULATION MULTI-PURPOSE OPTIMIZATION METHOD IN POWER DISTRIBUTION NETWORKS

Guliyev H.B.

•

Azerbaijan Scientific-Research and Design-Prospecting
Power Institute AZ1012, Ave. H.Zardabi-94,
E-mail: huseyngulu@mail.ru

ABSTRACT

Using the most effective method of usefulness theory in multi-purpose optimization, algorithms and effective methods of reactive capacity compensations inside given technical restrictions of power network of electricity have been worked out.

Key words: power systems, electric network, reactive capacity compensation, voltage, optimization.

I. INSTRUCTION

With a big number of electric energy consumers and different characters electric energy quality depends on many factors in the modern power networks. It includes: power networks and working condition factors of consumers. One of them is the possibility of reactive power balances with an important reserve providing after emergency modes on the basic knots of the power system and voltage regulation on all networks.

As the length of networks of a power system increases in modern conditions, we can reduce the reactive power streams, as well as operational and capital expenses. Rational voltage mode brings to the front plan the technical – economic aspects of the power transmission EFFICIENCY. Analyses and economic calculations show that transferring the reactive power by short length lines means of a high voltage justifies. Therefore in most cases reduction of reactive power to the minimum is very effective for economically when the sources of reactive power settle down near the consumption centers.

The increase of consumer loading and its structure qualitative causes considerable increase of reactive power and constant reduction of a power factor in distributed power networks [1].

Thus, the tendency of modern power systems development is characterized by one side with the increase of reactive power consumption (in some systems to 1 kVAR/kVt), on the other side with decrease of power plant generators usage expediency and possibility for the reactive power compensation purpose [2-5]. In such conditions reactive power compensation attains a special urgency. Here the optimization's primary goal is optimum placing of reactive power sources and support of a necessary reserve of capacity Q_{rez} for voltage regulation on loading knot. For example, Polish power engineers consider that capacity of compensators should be 50% of the established capacity of generators in power plants. In France, Sweden and Germany the capacity of compensators is 35% of active peak loading, in the USA and Japan this volume is 70%. In different power systems of the USA the established capacity of compensators is 100% of generators capacities [6-11].

II. PROBLEM STATEMENT

Reactive power compensation problem is a multidimensional problem on the technical and economic aspects and consequently it is resulted with the finding of a global extremum of criterion function with the set of local extreme. In this article the voltage support within the technical restrictions and definition of optimal placing of the reactive power sources with a technique of multi-purpose optimization of reactive power in the power system is considered. By the problem consideration as one-target optimization within restrictions the criterion function is a linear combination from several factors. The problem decision is a unique optimum version and has lacks of alternative versions, and there is not dependency of an end result from the initial data.

Thus, the purpose of reactive power sources optimal placing in a power system consists of increase the quality of voltage in all central points of a network, control the stability of the system, reduce the power losses and capacities in networks. As a result these will increase the economic efficiency in the power system. From the economic efficiency point of view the new compensating units intended for installation should be proved and given corresponding optimum recommendations.

Considering the aforesaid as conditions of technical restriction of optimization it is possible to accept the followings:

1. Voltage support in the set limits of the technical restrictions on considered knots of loading of a power system:

$$U_{i\min} \le U_i \le U_{i\max}$$
$$PF_{\min} \le PF \le PF_{\max}$$
$$F(x,u) = 0$$

where, U_{imin}, U_{imax} – is respectively, admissible minimum and maximum values of voltage on i^th^ knot; *PF*- power factor; *F(...)* -

2. Voltage deviation should be minimum:

$$\delta U = \sqrt{\sum \left(\frac{U_i - U_{i\,nom}}{U_{i\,nom}} \right)^{2}} \rightarrow \mathbf{min}$$

3. Power loss reduction maximization in power system networks

$$\Delta\Delta W_{il} = (\Delta\Delta \mathrm{P} \cdot \tau) = (\Delta \mathrm{P}_{b,c} - \Delta \mathrm{P}_{a,c}) \cdot \tau \rightarrow \max$$

where, $\Delta\Delta \mathrm{P}$, $\Delta \mathrm{P}_{b,c}$, $\Delta \mathrm{P}_{a,c}$ - respectively, the change (reduction) by loss of active power in the maximum loading mode, losses of active power before and after compensation, τ - the loss time of maximum power.

4. Maximization of annual economic efficiency

$$\mathrm{E} = \left(\Delta\Delta W_{il} \cdot \beta - \frac{C_k}{n_k} \right) \rightarrow \mathbf{max}$$

where, β - the cost price of electric energy, for a power system; C_k -general expense of compensating devices; n_k - operating time of compensating devices.

5. Compensating devices expense outlay recovery time minimization:

$$T_0 = \frac{C_k - C_t - C_x}{\Delta\Delta W_{il} \cdot \beta} \to \min$$

where, C_t and C_X -incomes formed from increase of throughput of transformers and lines as a result of compensation:

$$C_{t,i} = K_{t,i} \frac{S_{i,\mathbf{1}} - S_{i,\mathbf{2}}}{S_{nom}}$$

$$C_{x,i} = K_{xi} \frac{I_{i,\mathbf{1}} - I_{i,\mathbf{2}}}{I_{b,i}}$$

$S_{i,1}$, $S_{i,2}$ – total power of transformers before and after compensation respectively; $\dot{I}_{i,1}$, $\dot{I}_{i,2}$ - currents proceeding through lines before and after compensation; S_{nom} - full rated power of the transformer; $\dot{I}_{a,i}$ - an admissible current of a line; k_t and k_X-costs of lines and transformers.

For voltage regulation on loading knot the necessary reactive power reserve should be provided:

$$Q_{rez} = (0{,}1 \div 0{,}15) Q_y$$

In this case capacity of the compensating device on consumer balance border belonging and power supplying organization should satisfy the below-mentioned condition:

$$Q_k \geq (1{,}1 \div 1{,}15) Q_y - Q_{ekv}$$

where, $Q_{ekv} = \sum Q_y - q$ is equivalent reactive power which is defined on the basis of the report; $\sum Q_y$ - the sum of reactive power demanded by the consumer; q - reactive power loss in a network.

The most effective method of efficiency for multi-purpose optimization is used. According to this method efficiency of each analyzed version is defined as below:

$$B_k = \sum l_{il} \cdot N_i \to \mathbf{max}$$

where N_i - is coefficients of purposes importance, their sum should be equal to one; l_{ik} - the standardized values of optimization criteria.

For minimization purpose

$$l_{il} = \frac{Y_{k\max} - Y_{ik}}{Y_{k\max} - Y_{k\min}}$$

For maximization purpose

$$l_{il} = \frac{Y_{ik} - Y_{k\mathbf{min}}}{Y_{k\mathbf{max}} - Y_{k\mathbf{min}}}$$

In this problem decision the acceptance of N_i creates uncertainty and consequently its use creates difficulty. In the offered algorithm coefficients N_i are accepted in the below-mentioned form:

$$N_1 > N_2 > N_k$$

$$N_1 > (N_2 + N_3 + ... N_k) = m_i \sum_{i=2}^{k} N_i$$

$$N_1 > (N_{j+1} + N_{j+2} + ... N_k) = m_j \sum_{j+1}^{k} N_j$$

In the first case N_i is given by the researcher. And in the second case the change step is defined, and coefficients N_i are accepted as the least usage $N_i \ (N_k \approx 0{,}01 - 0{,}05)$:

$$a = \frac{1 + kN_k}{1 + 2 + ...(k-1)}$$

k - number of criterion functions.

For this case

$$N_{k-1} = N_k + a; \quad N_{k-2} = N_{k-1} + a = N_k + 2a; \quad N_1 = N_k + (k-1)a.$$

After definition N_i condition $\sum N_i = 1$ is checked and updating of coefficients to accuracy 0,001 is conducted.

In the third case for definition of N_i k number linear equations systems are constituted and from their decision N_i is found. For example, к=4 the following equation turns out

$$\begin{cases} N_1 + N_2 + N_4 = 1 \\ m_1 N_1 = (N_2 + N_3 + N_4) \\ m_2 N_2 = (N_3 + N_4) \\ m_3 N_3 = N_4 \end{cases} \quad \text{or} \quad \begin{cases} N_1 + N_2 + N_3 + N_4 = 1 \\ m_1 N_1 - N_2 - N_3 - N_4 = 0 \\ m_2 N_2 - N_3 - N_4 = 0 \\ m_3 N_3 - N_4 = 0 \end{cases}$$

The problem decision in the matrix form is as following:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ m_1 & -1 & -1 & -1 \\ 0 & m_2 & -1 & -1 \\ 0 & 0 & m_3 & -1 \end{vmatrix}^{-1} x \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \end{vmatrix} = \begin{vmatrix} N_1 \\ N_2 \\ N_3 \\ N_4 \end{vmatrix}$$

If coefficients m_1, m_2, m_3 are not given $m_1 = m_2 = \cdots = m_{k-1} = 1$ is accepted

III. PRACTICAL RESULTS

For defining of reactive power sources optimal voltages V_i values of coefficients efficiency for problem of four criterion optimizing have been accepted as V_i =0, 526; V_2 =0, 263; V_3=0, 158;

V_4=0, 053. Criteria report values are given at the table 1 and optimal variants have been given at the table 2.

As caan be seen from the tables 1 (0,789), 10 (0,7084) və 11-ci (0,7037) versions are much more optimal with the maximal efficiency point of view. For the optimal versions selfpeyment terms are 3,9; 3,4; 3,4 yerars.

As an example a distribution power network which has sixteen nodes has given. Before the installation of the capacitor banks in the electric network the power flow and power losses for the maximum load case has been calculated and given in the table 3 and table 4. As can be seen from the voltage graphic (Fig.1) in the most of the nodes voltage values are lower than the permittable minimum level (85 – 89%). The total losses of the network are 4, 37%.

Table 1

Standardizing and general efficiency

Variants	F1(0,526)	F2(0,263)	F3(0,158)	F4(0,053)	General effeciency
1	2	3	4	5	6
1	1	1	0	0	0,789
2	0,311	0,3162	0,6735	0,5714	0,3834
3	0,3484	0,3959	0,5102	0,1429	0,3756
4	0,7097	0,6427	0,4898	0,7143	0,6576
5	0,6503	0,5964	0,5102	0,5714	0,6098
6	0,4271	0,455	0,4898	0,1429	0,4293
7	0,3806	0,4242	0,4898	0,1429	0,3967
8	0,6968	0,6272	0,5102	0,7143	0,6499
9	0,4142	0,4396	0,5102	0,1429	0,4217
10	0,7794	0,6889	0,4898	0,7143	0,7064
11	0,7755	0,6864	0,4898	0,7143	0,7037
12	0,4103	0,437	0,5102	0,1429	0,4189
13	0,5226	0,491	0,5714	0,5714	0,5246
14	0,6748	0,6401	0,4286	0,4286	0,6137
15	0,5406	0,4961	0,5918	0,7143	0,5462
16	0,3226	0,3033	0,7347	0,7143	0,4034
17	0	0	0	1	0,211
18	0,6413	0,5938	0,5	0,5714	0,6028

Table 2

Optimal variants

№	Variants	General efficiency	F1 (man)	F2(MVt)	F3 (man)	F4 (year)
1	1	0,789	1758333	9,54	9400000	3,9
2	10	0,7084	1615833	8,33	7000000	3,4
3	11	0,7037	1613333	8,32	7000000	3,4

With the application of genetic algorithm eight nodes has been chosen for the optimal placement of the static capacitor banks.

The studied network with the allocation of capacitor banks (CB) in the nodes has been shown in fig.2. The voltage deviation from the nominal value has been given in the table 5. Also according to the power losses values of the static capacitor banks for which the electric network is considered as an optimal has been given in this table. After the allocation of the static capacitor banks the improvement of voltage quality and the decrease of power losses can be observed from the table 6. As can be seen the losses are decreased from 5,28% to 3,7%.

Table 3

The results of the flow distribution

Node	Voltage			Generation		Power		Node	Power Flow		Current	cos
Name	кV	%	Angle	MW	MVAr	MW	MVAr	Name	MW	MVAr	A	Power faktor. %
404	35	96,56	-1,2	0	0	6,178	2,682	**405**	-11,34	-5,818	217,8	89
								403	5,169	3,136	103,3	85,5
405	35	98,67	-0,6	0	0	2,434	1,46	**Dubendi**	-13,964	-7,53	265,2	88,0
								404	11,53	6,07	217,8	88,5
Buzovna	35	85,46	-5,2	0	0	0,454	0,272	**TS3**	-0,454	-0,272	10,2	85,8
								405	14,071	7,779	265,2	87,5
Dubendi	35	100,0	0,0	49,74	32,11	0	0	**N432**	8,511	5,385	166,1	84,5
								TS1	27,154	18,941	546,1	82,0
N406	35	89,79	-3,8	0	0	4,031	2,419	**TurkenII**	3,298	2,007	70,9	85,4
								N408	-16,894	-10,456	365	85
N407	35	85,69	-5,2	0	0	2,754	1,653	**TS2**	-3,665	-2,200	82,3	85,7
								TS3	0,911	0,547	20,4	85,7
N408	35	91,75	-3,0	0	0	2,104	1,263	**TS1**	-19,222	-12,178	409,1	84,5
								N406	17,118	10,915	365,0	84,3
N432	35	97,71	-0,5	0	0	8,365	5,184	**Dubendi**	-8,365	-5,184	166,1	85,0
N403	35	95,44	-1,5	0	0	5,124	3,074	**404**	-5,124	-3,074	103,3	85,8
TS1	35	95,59	-1,6	0	0	0	0	**Dubendi**	-26,439	-17,395	546,1	83,5
								Zire	6,744	4,197	137,1	84,9
								N408	19,694	13,197	409,1	83,1
TS2	35	87,13	-4,9	0	0	0	0	**N406**	-9,396	-5,667	207,7	85,6
								N407	3,717	2,252	82,3	85,5
								TurkenI	5,679	3,415	125,4	85,7
TS3	35	85,57	-5,2	0	0	0	0	**Buzovna**	0,455	0,273	10,2	85,7
								N407	-0,910	-0,546	20,4	85,7
								Seysm/st	0,455	0,273	10,2	85,7
Seysm/st	35	85,51	-5,2	0	0	0,455	0,273	**TS3**	-0,455	-0,273	10,2	85,8
TurkenI	35	86,94	-5,0	0	0	5,670	3,402	**TS2**	-5,670	-3,402	125,4	85,8
TurkenII	35	89,13	-3,9	0	0	3,278	1,987	**N406**	-3,278	-1,987	70,9	85,5
Zire	35	95,04	-1,7	0	0	6,716	4,162	**TS1**	-6,716	-4,162	137,1	85

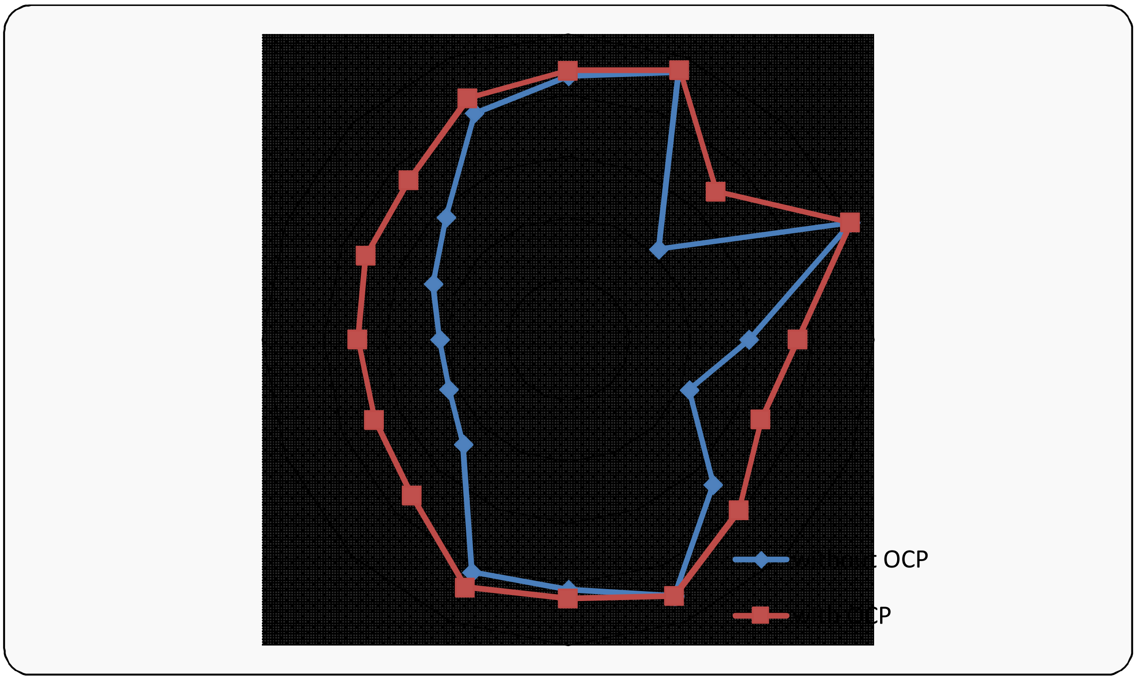


Fig.1. The voltage profile in the nodes

Table 4

The losses in the branches

Line	Flow -To		Flow- From		Losses		Node Voltage , %		Voltage drop %
Name	**MW**	**MVAr**	**MW**	**MVAr**	**кW**	**кVAr**	**To**	**From**	
Line 5	-11,347	-5,818	11,530	6,070	183,3	252,1	96,6	98,7	2,11
Line 6	5,169	3,136	-5,124	-3,074	44,8	61,6	96,6	95,4	1,12
Line 2	-13,964	-7,530	14,071	7,779	106,7	248,5	98,7	100,0	1,32
Line51	-0,454	-0,272	0,455	0,273	0,5	0,5	85,5	85,6	0,12
Line 8	8,511	5,358	-8,365	-5,184	146,1	200,8	100,0	97,7	2,29
Line18	27,154	18,941	-26,439	-17,395	715,8	1546,2	100,0	95,6	4,41
Line41	9,565	6,030	-9,396	-5,667	168,3	363,5	89,8	87,1	2,65
Line44	3,298	2,007	-3,278	-1,897	20,3	20,5	89,8	89,1	0,65
Line58	-16,894	-10,456	17,118	10,915	223,8	458,8	89,8	91,8	1,96
Line47	-3,665	-2,200	3,717	2,252	52,2	52,6	85,7	87,1	1,44
Line52	0,911	0,547	-0,910	-0,546	1,1	1,1	85,7	85,6	0,12
Line56	19,222	-12,178	19,694	13,197	472,0	1019,5	91,8	95,6	3,84
Line19	6,744	4,197	-6,716	-4,162	27,9	35,1	95,6	95,0	0,54
Line48	5,679	3,415	-5,670	-3,402	9,3	12,7	87,1	86,9	0,19
Line54	0,455	0,273	-0,455	-0,273	0,3	0,3	85,6	85,5	0,06
Total					2172,4	4273,9			

Table 5

The results of the optimal placement of capacitors

Node-Candidat	Voltage			Power Factor %	Information on the capacitors				Cost, ($)		
Name	kV	%	Angle		Nom. кВАр/ bank	Rate, кV	Number of banks	Total кVAr	Installation costs	Pursache, *thousand*	Service/year
404	35	96,998	-1,39	91,7							
405	35	98,850	-0,65	85,8							
Buzovna	35	92,124	-7,96	-70,4	1000	37	1	1000	1000	30	300
Dubendi	35	100,00	0,00	100,0	1000	37	1	1000	1000	30	300
N406	35	93,780	-5,35	85,8							
N407	35	92,057	-7,71	85,8							
N408	35	94,763	-4,21	85,8							
N432	35	97,712	-0,53	85,0							
N403	35	96,169	-1,80	97,1	1000	37	2	2000	1000	60	600
TS1	35	96,934	-2,11	100,0							
TS2	35	93,028	-7,11	100,0	1000	37	4	4000	1000	120	1200
TS3	35	92,123	-7,86	100,0							
Seysm/st	35	92,188	-7,95	-35,7	1000	37	2	2000	1000	60	900
TurkenI	35	92,885	-7,19	98,1	1000	37	3	3000	1000	90	900
TurkenII	35	93,406	-5,64	-99,3	1000	37	3	3000	1000	90	
Zire	35	96,385	-2,25	85,0							
Total							16	16000	7000	480	4200

Table 6

Summary results of losses in the nodes (at maximum load)

Node	Flow-To		Flow-From		Losses		Node Voltage,%		Voltage drop %
Name	MW	MVAr	MW	MVAr	кW	кVAr	From	To	
Line 5	-11,471	-4,026	11,636	4,253	165,2	227,1	97,0	98,8	1,85
Line 6	5,238	1,320	-5,202	-1,272	35,4	48,7	97,0	96,2	0,83
Line 2	-14,079	-5,719	14,176	5,946	97,5	227,2	98,8	100,0	1,15
Line51	-0,528	0,532	0,529	-0,531	0,9	0,9	92,1	92,1	0,00
Line 8	8,511	5,385	-8,365	-5,184	146,1	200,8	100,0	97,7	2,29
Line18	29,294	8,636	-28,685	-7,320	609,1	1315,7	100,0	96,9	3,07
Line41	10,912	-1,794	-10,764	2,113	147,6	318,7	93,8	93,0	0,75
Line 44	3,616	-0,419	-3,599	0,436	16,6	16,7	93,8	93,4	0,37
Line 58	-18,925	-0,425	19,111	0,807	186,3	381,8	93,8	94,8	0,98
Line 47	-4,241	0,001	4,285	0,044	44,5	44,9	92,1	93,0	0,97
Line 52	1,063	-1,908	-1,059	1,912	3,9	4,0	92,1	92,1	0,07
Line 56	-21,356	-2,154	21,750	3,004	393,7	850,4	94,8	96,9	2,17
Line 19	6,935	4,316	-6,907	-4,280	28,7	36,1	96,9	96,4	0,55
Line 48	6,479	1,305	-6,471	-1,294	8,1	11,1	93,0	92,9	0,14
Line 54	0,530	-1,381	-0,528	1,383	1,8	1,8	92,1	92,2	0,07
Total					1885,4	3686,1			

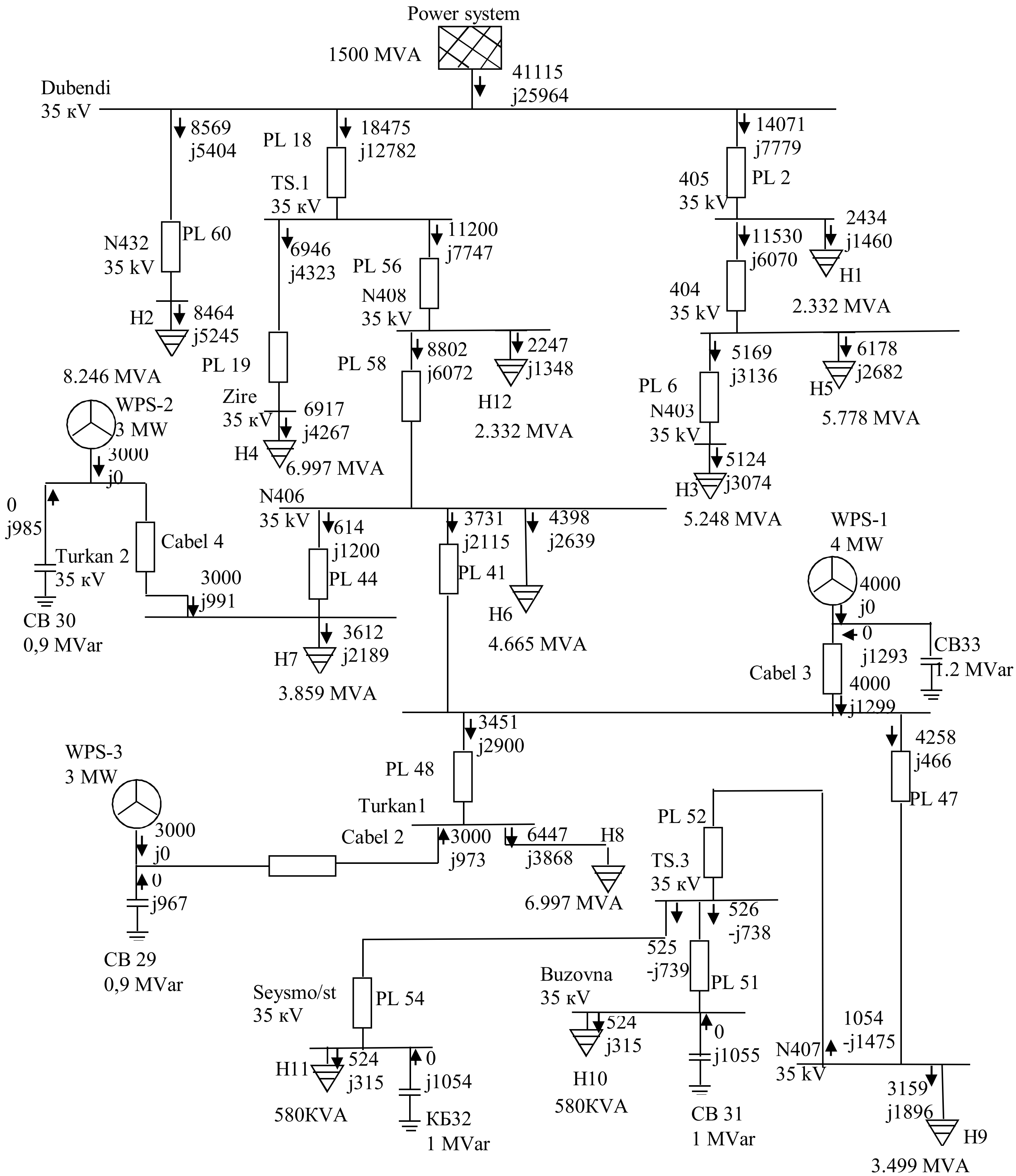


Fig. 2. The circuit of the network with the static capacitor banks

CONCLUSIONS

1. Methods and multi-purpose optimization compensations algorithms have been developed with support of a necessary reserve for preservation of normal level of voltage taking into account technical restrictions in knots of an electric network of a power system. Results of computerization to realization have shown speed and high efficiency the developed algorithm providing minimization of losses of active capacity in a net.

2. Based on genetic algorithm the power and installation locations of the static capacitor banks with the multicriteria optimization technique has given. In this case, as a criterion of optimality the minimum expenses for the installation and exploitation, the minimization of power losses during the required values of voltage and power factor and maximum saving and the minimum self-payment term are accepted.

3. The report of the real electricity network is given for two cases: operation without the CB; with optimal placement of CB. The application of the proposed method can reduce the average power losses approximately 13-14% in the electric network.

REFERENCES

1. Qotman V.İ., Markman Q.Z., Markman P.Q. İndemnification system jet capacity problems inspection. Industrial electricity. №8, 2006, c.50-55.

2. C. Lyra, C.Pissara, C. Cavellucci, A. Mendes, P.M. Franca. Capacitor placement in large-sized radial distribution networks, replacement and sizing of capacitor banks in distorted distribution networks by genetic algoritms. In *IEEE Proceedings Generation, Transmision & Distribution*, 2005, pp. 498-516.

3. M.H. Haque. Capacitor placement in radial distribution systems for loss reduction, In *IEEE Proceedings Generation, Transmision & Distribution*, 1999, pp. 501-505.

4. M.A.S.Masoum, M.Ladjevardi, A.Jafarian, E.F. Fuchs. Optimal Placement, Replacement and Sizing of Capacitor Banks in Distorted Distribution Networks by Genetic Algoritms, *IEEE Trans. Power Delivery*, 2004, vol.19, pp. 1794-1801.

5. T.M.Khalil, Hosam K.M.Youssef, M.M.Abdel Aziz. Optimal Capacitor Placement on Radial Distribution Feeders in Presence on nonlinear loads using binary particle swarm Optimization. 19th International Conference on Electricity Distribution, Vienna, 21-24 May 2007, paper No 180.

6. M.T.Hagh, M.Farsadi, S.Galvani. A Probabilistic Approach for Optimal Capacitor Placement in Unbalanced Distribution Systems Using NSGA_II. The 12th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS 2012), June 10-14, 2012, Istanbul, paper No 262.

7.Stebbins W. Power Distribution Systems and Power Factor Correction//Energy &Power Management. 2000.

8.Ovseichuk V.A.,Trofimov Q.Q. Technical and economic efficiency of jet capacity regulation and pressure in distributive electric networks. M, 2009.

9. Arzamaschev D.A. Lipes, A.V.Myzin. Optimization models of power system development. M, «Nauka», 1987.

10. H.B.Guliyev. Voltage regulation multi-purpoze optimization and reactive capacity indemnification on distributive netvorks of power systems. 3rd International Conference on Control and Optimization with Industrial Applications. Bilkent University, Ankara, 22-24 August 2011, p.113.

11. A.Rashtchizadeh, T.Sami, N.Rahmanov. Genetic Algoritm for Voltage Profile and Reducing Loss in the Power Distribution System Considering Distribution Generations and Capacitor Banks. The First Iranien Conference on Renewable Energies and Distributed Generation, ICREDG 2010, pp.1-5.

MULTI-CRITERIA DECISION TOOL APPLIED TO A SYSTEM RELIABILITY FOR THE PRIORIZATION OF SPARE PARTS.

V. González-Prida, General Dynamics – European Land Systems, Seville, Spain.
P. Viveros, Department of Industrial Engineering, Universidad Técnica Federico Santa María, Valparaiso, Chile.
A. Crespo, Department of Industrial Management, University of Seville, Spain.
C. Martin, ENIT- INPT, Toulouse University, France

ABSTRACT

This paper proposes a method for spare part priorization based on the system reliability behavior. The method considers the values taken by the reliability distribution parameters, as the result of a multi-criteria decision process. The range of values is divided into possible alternatives, which depend on the importance of different criteria. The presented exercise provides a quick view about how different spare part policies can be selected by the effect, not only of the design, installation quality or performed maintenance, but also due to factors that sometimes come from subjective assessments. Hence, the Analytic Hierarchy Process (AHP) includes both qualitative and quantitative criteria in the priorization scheme. The presented method is intended to be a starting point for the analysis of external factors that make an important influence on the decision–making of complex industrial assets, with high amounts of data, system configurations, and maintenance inputs, which will be analyzed in future researches with the support of a tailored software application.

1 INTRODUCTION

System reliability depends, mainly, on its design and installation quality. In addition to this, the reliability conservation will depend of course on the maintenance to be performed. Normally, these issues are analyzed from the point of view of a standard utilization of physical assets, under normal or controlled environments. However, external factors (as the usage profile) may affect the better or worse reliability conservation and, as a consequence, different maintenance plan can be tailored. Additionally, the current complexity of equipment provides difficulties for modeling the reliability behavior of a system, as far as components are generally a mix of taxonomies with different origins (electrical, mechanical, hydraulic, pneumatic etc.).

The objective of this brief study is to link some production criteria with those possible parameter values given for the system reliability. In order to illustrate this goal, an example is shown considering a Weibull distribution (as far as it is one of the most extended statistical distribution for modeling system reliability), and taking different values for its parameters. In order to simplify this sensitivity analysis, the present paper will consider failure rates intervals as well as different values for the shape parameter. Actually, the proposed methodology starts from a multi-criteria analysis with the target of selecting those parameter values that better match with the circumstances around the system.

The result will allow the analysis of the system reliability under specific parameter values. Finally, the maintenance or asset manager will be able to take a decision in aspects related to spare parts management, being able to be (for instance) more or less conservative or risky depending on the importance or weights considered during the multi-criteria decision process. With that purpose, this paper will start with a brief literature review on reliability linked to spare parts management. Afterwards, the proposed methodology for spare parts priorization is depicted. Then, with the support of a simple example, the study will approach the reliability uncertainty considering different alternatives for failure rate and shape parameter. The obtained results are shown and discussed in the following section, providing different points of view for the sensitivity analysis and how the selected parameter values may depend on diverse, external and sometimes subjective factors. Finally, there are some conclusions at the end of the paper, summarizing the main lines of this research.

2 BRIEF LITERATURE REVIEW

2.1 *On the Reliability Assessment*

Currently, different alternatives exist for the individual and systemic logical-functional representation of processes (Viveros et Al. 2011). The Reliability Block Diagram RBD methodology permits the representation of a system as a network of organized components, identify in terms of operation continuity, the contribution and effect of each on the system. This technique is better adjusted to non-repairable component systems and when the order of failure occurrence is not important. In addition to this, a complexity recognized in the productive processes of the mining industry is the large amount of equipment and systems, so that the industry prioritizes the use of functional tools with practical implementation and already proven in different plants and mining processes (Viveros et Al. 2011).

The RBD analysis methodology (Rausand & Hoyland, 2003; Guo & Yang, 2007), is a widely used technique in the mining sector, due to its adaptability to represent complex provisions and environments with large amounts of equipment, where they look for ways to simplify the reliability analysis through the use of block diagrams under systematic configurations in series, parallel or other more complex configurations (Guo & Yang, 2007). However, when working with reparable component systems and when the order in which the failures are represented is important, the Markov method is generally the most convenient (Rausand & Hoyland, 2003). For the calculation of Reliability of the subsystem in parallel or full redundancy, we use:

$$R_{sub}(t) = 1 - \prod (1 - R_i(t)) \tag{1}$$

Where $R_i(t)$ represents the individual reliability of the equipment. The k-out-of-n system structure is a type of structure in which system can operate, only if at least k of the n components are correctly working. For the reliability analysis, the formulation used is according to the probability of the Event Space Method (Lisnianski, 2007), where for a redundant system n over j, the reliability is represented by:

$$R_{sub}(t) = \sum_{j=m}^{n} \binom{n}{j} R^j * (1 - R_j)^{n-j} \tag{2}$$

However, this formulation assumes that the reliability of n equipments are similar, so none can be generalized to this case since the equipments have different adjustment parameters, and therefore its behavior in terms of reliability is different. For this reason, as an example: for a Subsystem, 3-out-of- 2, the formula will be:

$$\mathrm{R_{Sub}}(\mathrm{t}) = \prod_{i=1}^{3} R_i + (1 - R_3) * R_1 * R_2 + (1 - R_2) * R_1 * R_3 + (1 - R_3) * R_1 * R_2 \tag{3}$$

For a load sharing subsystem, example: for 60% and 40% distribution, the reliability formulation is:

$$\mathrm{R_{Sub}}(\mathrm{t}) = \sum_{i=1}^{n} R_i(t) * I_i + (1 - \sum_{i=1}^{n} I_i) * \prod_{i=1}^{n} R_i(t) \tag{4}$$

According to the last formulation, $R_i(t)$ represents the individual reliability of each equipment, and I_i is the impact on production system as a consequence of the failure of the equipment i. The second part of the equation (4) [$(1-\sum_{i=1}^{n} I_i) * \prod_{i=1}^{n} R_i(t)$], has importance when the capacity of the system, represented by the sum of the individual capacities of the equipment, is greater than 100% of demand, it means that exist a load sharing structure with overcapacity. Finally, the reliability of overall system, represented by a serial structure, would be equivalent to:

$$R_{Sub}(t) = \prod_{i=1}^{n} R_i(t) \tag{5}$$

Generally, if the reliability of a system needs to be improved, then efforts should first be concentrated on improving the reliability of the component that has the largest effect on reliability (Macchi et Al. 2012).

2.2 *Link to the Spare Parts Management*

The literature regarding reliability analysis usually deals with system features like "Ratio of system failure", "Mean Time Between Failure” (MTBF), or "Mean Down Time” (Rausand and Høyland, 2003; American Institute of Chemical Engineers, 1989).On the other hand, the bibliography about Spare Parts Management covers a wide range of topics: stocking strategies (Molenaers et Al. 2012), inventory control (Kennedy et al. 2002), priorization based on demand (Syntetos et al. (2009), realistic classification approaches (Braglia et al. 2004), among many others. In order to link the reliability analysis with the system units (those units that constitute the industrial assets), the following chart (Table 1) shows the expressions in terms from the components themselves.

Formulas	Two subsystems in Series	Two subsystems in Parallel
System Failure Rate	$\lambda_{series} = \lambda_1 + \lambda_2$	$\lambda_{paralllel} = \lambda_1 \cdot \lambda_2 \cdot (MDT_1 + MDT_2)$
System MTBF	$MTBF_{series} = (MTBF_1 \cdot MTBF_2) / (MTBF_1 + MTBF_2)$	$MTBF_{parallel} = (MTBF_1 \cdot MTBF_2) / (MDT_1 + MDT_2)$
System Mean Down Time (MDT)	$MDT_{series} = (MTBF_1 \cdot MDT_2 + MTBF_2 \cdot MDT_1) / (MTBF_1 + MTBF_2)$	$MDT_{parallel} = (MDT_1 \cdot MDT_2) / (MDT_1 + MDT_2)$

Table 1. Summary chart of formulas

The conventional formulas mentioned above are sometimes not used in cases of hybrid configurations. Other methods such as a probabilistic formulation (Henley and Kumamoto, 1992) are usually applied for these hybrid cases. Nevertheless, using the a.m. conventional formulas, one can observe from the quantitative assessment of these parameters (González-Prida et Al. 2009), those subunits or components which can be critical in the functioning of the entire system (Crespo and Iung, 2007). The obtained result allows tailoring for example the possibility of a preliminary list of recommended spares (González-Prida et Al. 2010). This kind of analysis are easy to implement during the design phase, and useful once the system is launched to the market and starts working. Once the system performance is known, it is possible to apply a similar analysis with real data about the system behaviour, making easier and more realistic the decision taking for future batches of spare parts. Other interesting references in this field are (Barberá et Al. 2010), (Vintr 2007) or (González-Prida and Crespo, 2010).

2.3 *Risk analysis and probability of failure*

The common definition of risk (associated with hazard) is the probability that a hazard will occur and the (usually negative) consequences of that hazard. In essence, it comes down to the following expression (the most frequently used definition in risk analysis) ISO 31000, 2002, where R is the risk, P_{fi} is the probability of failure and C_{fi} represent the consequences of the unwanted event.

$$Risk = \sum_{i=1} P_{f_i} * C_{f_i} \tag{6}$$

According to Kaplan and Garrick, risk consists of three components; (1) the scenario, (2) the probability of the scenario and (3) the consequences of the scenario. Kaplan and Garrick, 1981. Also suggests that one has to take all hazards into account, which can be accomplished by summing up all possible hazards (scenarios) with their consequences for a certain activity. Particularly for the calculation of probability, we refer to the reliability of the equipment, which depend directly on the parameters of life of its distribution function. The changing and evolution of life parameters affect directly on the life expectancy (MTBF) and consequently in the number of changes (parts) in a finite time period.

For example, decision makers have to consider that the number of events equals the number of failures allowable for the system to continue running during the analysis period t, which should be less than or equal to the number of parts available. It is possible to obtain all the probabilities for the success or failures scenarios. Therefore, the results are probabilistic for discrete scenarios, so that the maximum allowable failures and the number of available spare parts are integer numbers. In the other hand, the consequence depend of the attributes or criteria considered, according to the significance of business itself.

3 PRIORIZATION OF SPARE PARTS

Despite the possible procedure of ranking the maintainable items according just to their reliability, a more complex replacement model for spare parts is developed by Molenaers et Al. 2012. Such a model considers other criticality criteria like:

- Frequency of a failure
- Possible consequences of a failure
- Probability of item failure
- Replenishment time
- Number of potential suppliers
- Availability of technical specifications
- Maintenance type

Different criteria can also be observed in case studies like Gonzalez-Prida et Al. 2011, where the AHP is implemented considering criteria like: availability of extra stock, repair/supply cost and terms, and (in that scenario), the number of assets under warranty and the number still to deliver. As introduced in the first section, the objective now is to link some of the above mentioned criteria (or choosing new ones) with those possible parameter values given for the reliability models. Gajpal et al. (1994) and Braglia et al. (2004) adopted the AHP approach for spare parts classification based on criticality. The proposed methodology considers the possible parameter values or intervals as alternatives in a multi-criteria decision tool (Table 2).

		Characteristics	
Alternatives	Parameters	Working Conditions	Items Type
A	Lambda Max. & Beta Min.	Severe	Mechanical
B	Lambda Max. & Beta Max.	Severe	Electrical
C	Lambda Min. & Beta Min.	Soft	Mechanical
D	Lambda Min. & Beta Max.	Soft	Electrical

Table 2. Example of alternatives for the sensitivity analysis

In other words, the novelty here is to consider for the reliability modeling, those factors that can make the parameters to take different values. These factors will change from the point of view of technical requirements, facilities and boundary conditions, etc. Spare parts are considered critical (Huiskonen 2001) for a production process (from the point of view of the consequences for an industrial plant) or functional control (from the point of view of industrial assets fleets). Therefore, many attributes can be taken into account as criteria for the hierarchy process. Based on their significance to the business itself, some examples for criteria are presented in Table 3.

Criticality criteria	Description
Production requirement	Expected quantity of products obtained in a defined term and under specific quality standard by the industrial process of transforming tangible (materials) and intangible (knowledge) inputs into goods or services.
Installation environment	Conditions that surrounds an assembly of systems. Under such conditions, the whole machinery should run and work properly.
System availability	Characteristic of an industrial resource, which is committable, operable, or usable on demand to perform its required function. This characteristic extends the definition to elements such as quantity and proximity of spares, tools and manpower to the resource itself.

Table 3. Example of criticality criteria

Once known the criteria and alternatives (Figure 1), maintenance and assets managers may proceed to calculate the relative importance or weight, according to their judgements, which are transformed into mathematical matrices.

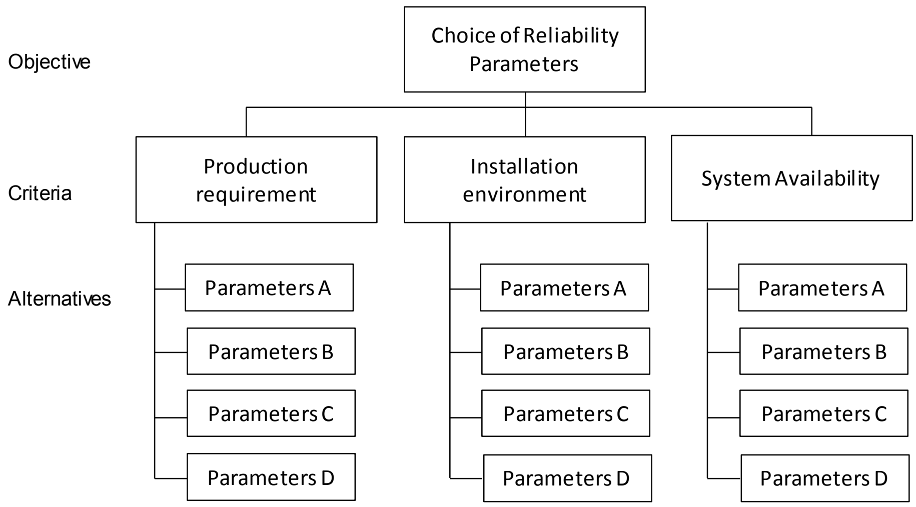


Figure 1. Hierarchy process

To sum up, the proposed method considers then the following steps:

i. Definition of statistical distribution in order to model the system reliability.
ii. Definition of plausible intervals and ranges for the reliability parameters.
iii. Definition of alternatives according to parameter values and intervals.
iv. Definition of criticality criteria for production process or functional control.
v. Calculation with the AHP procedure.
vi. Implementation of the selected alternative to the statistical distribution.
vii. Obtaining a ranking of maintainable items according to their reliability.
viii. Selection of spares (for example, by Pareto principle, or economic considerations)

4 APPROACHES TO THE RELIABILITY UNCERTAINTY

4.1 *Scenario and prior conditions*

In order to analyse the system reliability considering a specific failure rate interval as well as different values for the shape parameter, we will consider a very simple system constituted by four components or maintainable items. The configuration of the system functional blocks (ISO/DIS 14224), as we can see in Figure 2, will be two components (A and B) in series, plus another two components (C and D) in parallel. The intention here is to obtain the system reliability R(t), as a function of the component characteristics. In other words, applying the formulas from Table 1, we obtain for the whole system, the following expressions for system failure rate (7), and system MTBF (8).

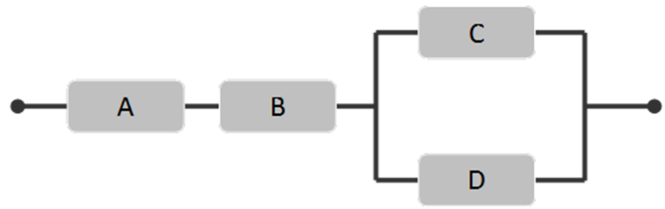


Figure 2. Block diagram example

$$\lambda_1 = \lambda_A + \lambda_B + \lambda_C \cdot \lambda_D \cdot (MDT_C + MDT_D) \tag{7}$$

$$MTBF_1 = \left[\frac{MTBF_A \cdot MTBF_B}{MTBF_A + MTBF_B} \cdot \frac{MTBF_C \cdot MTBF_D}{MDT_C + MDT_D}\right] \div \left[\frac{MTBF_A \cdot MTBF_B}{MTBF_A + MTBF_B} + \frac{MTBF_C \cdot MTBF_D}{MDT_C + MDT_D}\right] \tag{8}$$

Based on the above formulation and considering a constant failure rate for the system life cycle as well as a Weibull distribution as a model for the reliability behavior over time (9), it is possible to obtain reliability values R(t) (Meeker and Escobar, 1999) for the complete system.

$$R(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}} \tag{9}$$

In the expression for R(t), for each equipment:

- t: time
- β: shape parameter
- η: characteristic life

Particularizing to our example, we will analyze the system evolution during 1 year (twelve months). Therefore, t=[1, 12]; the shape parameter will assume a Weibull Distribution with β=0.5, till an Exponential Distribution (β=1); and finally, the characteristic life will be considered as $\eta = MTBF \cdot 10^{-6}$, which depends on the values taken for the failure rate. Basically, the assumed data for sensitivity analysis will be those ones included in the following chart (Table 4).

The assumed values for β consider the case when the elements are electrical items, then β=1 (the Weibull expression refers then to an Exponential distribution); or the case when the elements are mechanical items, then β=0.5 (Lawless J.F.). The physical explanation is that mechanical items are usually deteriorating overtime faster than the electrical ones (Parra et Al. 2006). In other words, the goal with the different values for shape parameter is to consider pure electrical components or pure mechanical components, as far as the curve trend in both cases are different depending if the components are just electrical or mechanical. Therefore, the shape parameter for mechanical components should be lower than the shape parameter for the electrical items.

Parameter	Min	Max
Beta	0.500	1.000
Lambda A	125.000	225.000
Lambda B	100.000	200.000
Lambda C	5.000	20.000
Lambda D	10.000	15.000

Table 4. Assumed data for the sensitivity example

Similarly, the values of λ are higher when the system is considered to be used under severe conditions (higher trend to the failure), or lower when the usage conditions are even softer than the standard conditions. In other words, the goal with a failure rate interval is to consider different usage profiles or environmental severities. Values for failure rates can be obtained from data bases as Oreda or Faradip (Sintef, 2002; Smith, 2001).

4.2 *Calculations*

Considering the assumed values for Lambda (failure rate) and Beta (shape parameter), it is possible to calculate the system reliability R(t) according to different combinations of the commented values. In our case, we will take into account just the four extreme cases as alternatives for our multi-criteria decision process:

A. Lambda Max. & Beta = 0,5
B. Lambda Max. & Beta = 1
C. Lambda Min. & Beta = 0,5
D. Lambda Min. & Beta = 1

Providing the calculations just for alternative A, the values for R(t) is shown in the following chart (Table 5):

t (month)	Item A	Item B	Item C	Item D	System
1	0.6642	0.6799	0.8852	0.8997	0.5699
2	0.5687	0.5873	0.8451	0.8644	0.4604
3	0.4980	0.5183	0.8123	0.8353	0.3836
4	0.4471	0.4682	0.7866	0.8123	0.3308
5	0.4054	0.4268	0.7640	0.7920	0.2891
6	0.3721	0.3937	0.7447	0.7747	0.2570
7	0.3430	0.3647	0.7269	0.7586	0.2298

t (month)	Item A	Item B	Item C	Item D	System
8	0.3181	0.3396	0.7107	0.7440	0.2071
9	0.2970	0.3183	0.6963	0.7309	0.1885
10	0.2777	0.2988	0.6825	0.7183	0.1719
11	0.2611	0.2819	0.6701	0.7070	0.1579
12	0.2456	0.2662	0.6580	0.6959	0.1452

Table 5. Reliability values considering λ max. and β=0,5

Graphically, the above mentioned chart can be represented with the curves shown in Figure 3. In the same way, it is possible to obtain the values for R(t) in alternatives B, C and D. Nevertheless, according to the proposed method, the alternative to be implemented to analyze the reliability evolution of the system should be the one obtained as a result of the AHP process. Such a process is not here included as far as the procedure is quite well known and does not provide a significant novelty to the current research.

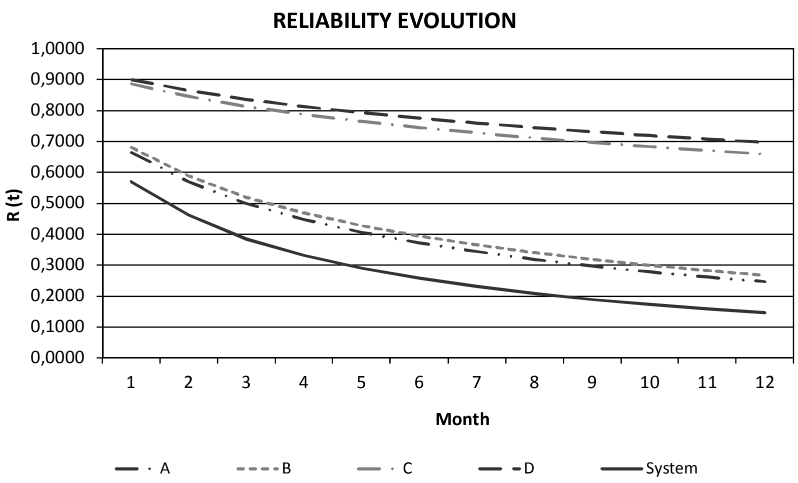


Figure 3. Graphical representation of alternative A

Another interesting quantitative proposal to analyze this phenomenon is according to (Ramirez-Marquez et Al. 2006) and (Barabady and Kumar, 2008), where the reliability importance, I, of component i in a system of n components is given by:

$$I(i) = \frac{dR_s(t)}{dR_i(t)} \tag{9}$$

Where $R_s(t)$ is the system reliability, and $R_i(t)$ is the component reliability.

4.3 ***Results***

From these values, it is possible to obtain for each alternative a ranking of “items reliability”. This ranking can be a helpful tool in order to decide which components should be prioritized in comparison to the rest, in order to draw up a list of recommended spares. Of course, this example has not consider the implementation of the AHP decision tool, or the monetary value of the spare parts. Nevertheless, it is possible to illustrate the curves for the system reliability behavior, according to each alternative. This representation is shown in Figure 4.

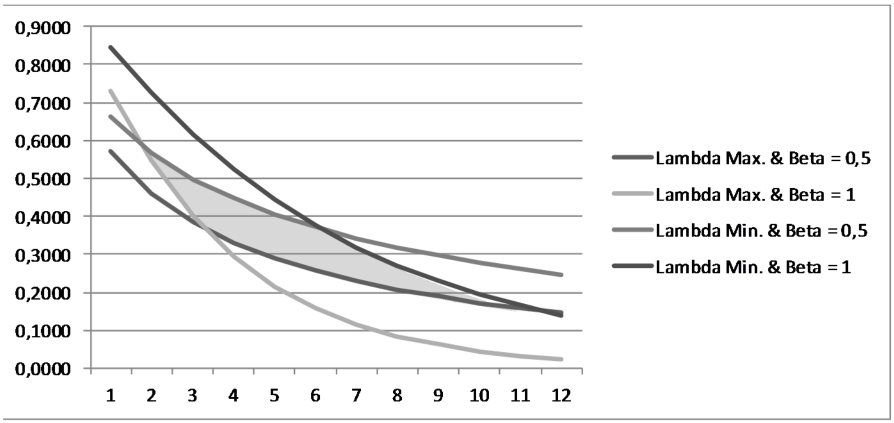


Figure 4. System reliability behavior according to each case

As already commented, if the four items are the possible spare parts they can be prioritized according to their unreliability. Therefore, considering each alternative, the unreliability ranking is shown in the following chart (Table 6).

RANKING	Item A	Item B	Item C	Item D
Lambda Max. & Beta = 0.5	4	3	2	1
Lambda Max. & Beta = 1	4	3	2	1
Lambda Min. & Beta = 0.5	4	3	1	2
Lambda Min. & Beta = 1	4	3	1	2

Table 6. Unreliability ranking of items acc. to each alternative

5 DISCUSSION AND FUTURE APPLICATIONS

With this easy example, we observe that:

a) Depending on the alternative considered, the need of a component as spare part varies.
b) The value of β also affect to the ranking of spare parts need (although it is not perceptible with the values of this example).
c) The exercise has assumed a pure mechanical case (all β=0.5) and a pure electrical case (all β=1). However, items in the reality are mixed, interacting electromechanical components (or hydraulic, pneumatic, etc.), which may present a different shape parameter.
d) The most conservative position would be the scenario that assumes the worst system performance (higher failure rate).
e) On the contrary, the most probable situation would be that one whose parameters take values adjusted to a more realistic behavior. That means, λ and β would take values within the range [min, max], but not necessarily in the extreme (as assumed by this example).

Therefore, the most probable situation does not have to coincide with the most conservative one. As a consequence, future applications that this study may provider are:

- If we foresee the system usage profile, it is possible provide to the end user a tailor-made list of recommended spare parts.

- Similarly to the spare parts, maintenance plan may also vary according to the severity of use or working environment. The possible changes can affect the task frequency, and obviously the order of application in the maintenance schedule.

The outlined example can be implemented in a more complex way (i.e., with more number of subsystems and levels for example to reach the maintainable items; or considering different statistical distributions with diverse parameters and rates). With that target, the next step will be the implementation of the proposed methodology, including the AHP calculation, in a software for the processing of a high amount of data, values and system configurations, which allow the performance of simulations, histograms etc. This future study will consider as decision variables not only the Reliability R(t), but also Maintainability, Cost, Production, Availability, Usage etc.

The effectiveness of production processes and the equipment that are part of them is generally measured according to the results of reliability and availability indicators, as well as through the economic analysis of its life cycle. In addition to this, the OEE indicator allows for the measurement of productive efficiency using the control parameters as a basis for calculating fundamentals in industrial production: availability, efficiency and quality. The productive processes in the mining industry (future development) have, as an additional complexity, a large amount of equipment and systems, which make the systematic analysis of the plant more difficult. Because of this, different analysis methodologies have been developed, like the last explained RBD methodology, widely used in the mining sector for its adaptability of representation in complex arrangements and environments with large amounts of equipment, where they look to simplify the reliability analysis through the use of diagram blocks under systemic configurations mainly in serial and parallel All decision parameters are actually the output of a series of qualitative or quantitative factors. These factors may be for instance the required level of service, the investor profile, the needed workload etc. obtaining as a result solutions packages which will be dependent on the system boundary conditions.

6 CONCLUSION

This paper suggests a methodology to select spare parts for an assumed system. The proposed method provides to the reader an easy view about the effect over the system maintenance, not only the design or installation quality, but also the consideration of external (and sometimes subjective) factors. The exercise considers different shape parameters and intervals for the failure rates. The failure rate has been here considered as a constant during the life cycle of the system. Nevertheless, all these parameters should be assessed based on different factors as a result of a multi-criteria decision tool. The presented study implies how usage profile and/or severity of the working conditions may affect to the decision of required spares, but also task frequency, or other aspects of the maintenance policies. Moreover, achieving a good level of reliability, especially in the critical assets, requires appropriate analysis and prioritization in the allocation of resources and adequate allocation of maintenance policies according to the criticality of maintainable each element. Therefore, a method for convenient and practical hierarchical becomes an important tool for the success of the maintenance function and, in some cases, its complement methodologies for auditing the resources allocation of critical maintenance activities. In conclusion we can say that is necessary to have this type of tool that dynamically, can re-establish procurement policies spares and adjust to changing business conditions the values of R (t) of our equipment.

7 REFERENCES

American Institute of Chemical Engineers Center for Chemical Process Safety. (1989). Guidelines for Process Equipment Reliability Data, with data tables. ISBN 0-8169-0422-7.

Barabady, J., Kumar, U. (2008). Reliability analysis of mining equipment: A case study of a crushing plant at Jajarm Bauxite Mine in Iran. Reliability Engineering and System Safety 93:647–653.

Barberá L, González-Prida V, Moreu P, Crespo A. (2010). Revisión de herramientas software para el análisis de la fiabilidad, disponibilidad, mantenibilidad y seguridad (RAMS) de equipos industriales. Revista Ingeniería y Gestión de Mantenimiento. Spain.

Braglia, M., Grassi, A., Montanari, R., 2004. Multi-attribute classification method for spare parts inventory management. Journal of Quality in Maintenance Engineering 10 (1):55–65.

Crespo A, Iung B. (2007). A Estructured Approach for the Assessment of System Availability and Reliability Using Montecarlo Simulation. Journal of Quality in Maintenance Engineering. 13(2):125-136.

Gajpal, P.P., et al., 1994. Criticality analysis of spare parts using the analytic hierarchy process. International Journal of Production Economics 35 (1-3):293–297.

González-Prida V, Barberá L, Crespo A. (2010). Practical application of a RAMS analysis for the improvement of the warranty management. 1st IFAC workshop on Advanced Maintenance Engineering Services and Technology, Lisbon, Portugal.

González-Prida V, Crespo A. (2010). Book Review: Reliability Engineering. Warranty Management and Product Manufacture (By D.N.P. Murthy and W.R. Blischke). Production Planning & Control: The Management of Operations. 21(7):720-721.

Gonzalez-Prida V. Parra C, Gómez J, Crespo A, Moreu P. (2009). Availability and Reliability Assessment of Industrial Complex Systems: a Practical View Applied on a Bioethanol Plant Simulation. Safety, Reliability and Risk Analysis: Theory, Methods and Applications. Taylor & Francis, pp 687-695. ISBN: 978-0-415-485.

Gonzalez-Prida V., Gómez J., Crespo A. (2011). Practical application of an Analytic Hierarchy Process for the improvement of the warranty management. Journal of Quality in Maintenance Engineering 17 (2), 163-182

Guo H, Yang X. (2007). A simple reliability block diagram method for safety integrity verification. Reliability Engineering System Safety, 92:1267–1273.

Henley E, Kumamoto H. (1992). Probabilistic Risk Assesment, Reliability Engineering, Design and Analysis. IEEE Press. ISBN 0-87942-290-4.

Huiskonen J. (2001). Maintenance spare parts logistics: special characteristics and strategic choices. International Journal of Production Economics 71 (1–3): 125–133.

ISO/DIS 14224 (2004). Petroleum, petrochemical and natural gas industries - Collection and exchange of reliability and maintenance data for equipment. International Standard.

ISO 31000, 2002. Current Draft Vocabulary for Risk Management – ISO/IEC Guide 73:2009.

Kaplan, S., Garrick, B.J., (1981). On the quantitative definition of risk. Risk Analysis. 1, 11–27.

Kennedy, W., Patterson, J., Fredendall, L., 2002. An overview of recent literature on spare parts inventories. International Journal of Production Economics 76 (2):201–215.

Lisnianski A. (2007). Extended block diagram method for a multi-state system reliability assessment. Reliability Engineering & System Safety; 92, 1601–1607.

Macchi M, Kristjanpoller F, Arata A, Garetti M, Fumagalli L. (2012). Introducing buffer inventories in the RBD analysis of production systems. Reliability Engineering & System Safety; 104, 84-95.

Meeker WQ, Escobar LA. (1999). Statistical Methods for Reliability Data. Journal of the Royal Statistical Society. Series D (The Statistician) Wiley. 48(3):456-458. ISBN: 0-471-14328-6.

Molenaers A. Baets H, Pintelon L, Waeyenbergh G. (2012).Criticality classification of spare parts: A case study. *Int. J. Production Economics* 140:570–578

Parra C, Crespo A, Cortés P, Fygueroa S. (2006). On the Consideration of Reliability in the Life Cycle Cost Analysis (LCCA). A Review of Basic Models. Safety and Reliability for Managing Risk. Taylor & Francis, pp. 2203-2214. ISBN: 0-415-41620-5

Ramirez-Marquez JE, Rocco CM, Gebre BA, Coit DW, Tortorella M. (2006). New insights on multi-state component criticality and importance. Reliability Engineering & System Safety; 91:894–904.

Rausand M, Hoyland A. System reliability theory: models, statistical methods, and applications. 3rd Ed. New York: Wiley, IEEE; 2003.

SINTEF Technology and Society. (2002). Offshore Reliability Data Handbook. 4th Edition, (OREDA). ISBN 8-21400-438-1.

Syntetos, A., Keyes, M., Babai, M., 2009. Demand categorisation in a European spare parts logistics network. International Journal of Operations & Production Management 29 (34):292–316.

Smith D. (2001). Reliability, maintainability, and risk: practical methods for engineers. Newnes. ISBN 0-75060-854-4.

Vintr M. (2007). Reliability Assessment for Components of Complex Mechanisms and Machines. 12th IFToMM World Congress. Besançon, France.

Viveros P, Zio E, Arata A, Kristjanpoller F. (2011). Integrated system reliability and productive capacity analysis of a production line. A Case Study for a Chilean Mining Process. Proceeding of the Institution of Mechanical Engineers, Part O: Journal of Risk and reliability.

ISSN 1932-2321