THE JOURNAL IS INCLUDED IN THE LIST OF THE LEADING JOURNALS
AND EDITIONS OF THE HIGHER ATTESTATION COMMISSION (VAK)

ISSN 1729-2646
e-ISSN 2500-3909

THE JOURNAL PROMOTER:
“Journal “Reliability” Ltd

It is registered in the Russian Ministry of Press, Broadcasting and Mass Communications. Registration certificate ПИ 77-9782, September, 11, 2001.

Official organ of the Russian Academy of Reliability

Publisher of the journal
LLC Journal “Dependability”
Director
Dubrovskaya A.Z.
The address: 109029, Moscow,
Str. Nizhegorodskaya, 27,
Building 1, office 209
Ltd Journal “Dependability”
www.dependability.ru
Printed by JSC “Regional printing house,
Printing place” 432049, Ulyanovsk,
Pushkarev str., 27. Circulation: 500 copies.
Printing order
Papers are reviewed. Signed print ,
Volume , Format 60x90/8, Paper gloss

Papers are reviewed.
Papers are published in author’s edition. The opinion of members of the editorial board may not coincide with the point of view of authors’ publications. The reprint of materials is granted only with the written permission of the editorial board. Manuscripts are not returned.

THE JOURNAL IS PUBLISHED WITH THE PARTICIPATION AND SUPPORT OF THE JOINT-STOCK COMPANY «RESEARCH AND DESIGN INSTITUTE OF INFORMATISATION, AUTOMATION AND COMMUNICATION ON RAILWAY TRANSPORT» (JSC «NIIAS») AND LLC PUBLISHING HOUSE «TECHNOLOGY»

CONTENTS

Structural dependability. Theory and practice

Functional dependability. Theory and practice

Functional safety. Theory and practice

Dependability, vol. 17 no.3, 2017

Original article
DOI: 10.21683/1729-2646-2017-17-3-3-9

On the planning of the scope of new technology testing

Alexander V. Antonov, *Obninsk Institute for Nuclear Power Engineering (IATE MEPhI), Obninsk, Russia*
Valeri A. Chepurko, *Obninsk Institute for Nuclear Power Engineering (IATE MEPhI), Obninsk, Russia*
Vladimir E. Chekhovich, *State Scientific Center of the Russian Federation A.I. Leipunsky Institute of Physics and Power Engineering, (SSC RF-IPPE), Obninsk, Russia*
Vladimir F. Ukraintsev, *State Scientific Center of the Russian Federation A.I. Leipunsky Institute of Physics and Power Engineering, (SSC RF-IPPE), Obninsk, Russia*

Abstract. *This paper is a follow-up to [1]. It examines the matters of planning of the scope of highly dependable objects testing. The process of new technology development and manufacture involves determining its dependability indicators. The most objective method of identifying dependability characteristics of products is field testing. One of the widely used testing plans is the [N,U,T] plan. This plan that involves testing N nonreparable samples within the time interval between 0 and a certain T. It is assumed that during the tests k objects fail, while N-k objects successfully pass the tests. Thus, at the outcome of the experiment we have a mixed sample that includes k times to failure and N-k right censored observation. If the tested object is highly dependable it is quite possible that within the time period [0,T] failures will not happen, i.e. k will be equal to 0, therefore the probability of failure within this time period is extremely low and the number of tested objects is limited. Nevertheless even in this situation it would be desirable to be able to be in control of the accuracy of the estimation obtained during such experiments. It is clear that the accuracy of such estimation will depend not only on the number of tested objects N, but also on the experiment duration. For a fixed N, as the observation time T grows the estimation accuracy increases due to the increasing proportion of complete times, while the proportion of censored ones goes down. It should be noted that when we talk about identifying the dependability characteristics of complex and costly objects we cannot test large batches of finished products. Therefore the problem consists in defining testing duration and size of the product batch to be tested subject to specified requirements for the accuracy of estimation of dependability characteristics obtained as the result of the tests. The scope planning is based on the manufacturer's requirement to validate the lower bound of the probability of no failure $\underline{P}_0$ with a specified confidence level at a certain time point t_0.* **The aim** *of the paper is to identify the test scope of a batch of finished products N(T) under the condition of fulfilment of the manufacturer's requirement for compliance with the lower confidence bound of the probability of no failure with a specified confidence level 1 – . Three failure distributions are under examination: exponential distribution law, Weibull distribution and distribution with linear rate function. The considered types of distribution law enable the research of objects with decreasing, constant and increasing failure rate function.* **Methods.** *In this paper the authors deduce formulas for calculation of the scope of experiment for a number of experiment durations. The estimates are obtained using the maximum likelihood method and methods of researching asymptotic properties of estimates through the Fisher information quantity.* **Conclusions.** *The findings allow for a substantiated approach to planning the scope of highly dependable objects testing. It is shown that the longer is the experiment duration the fewer products must be supplied for testing. The dependence is non-linear, close to hyperbolic and is conditioned by both the input parameters and the parametrization of the failure rate function.*

Keywords: *test scope planning, experiment duration, probability of no-failure, failure rate, lower bound estimate of probability of no-failure, confidence level.*

For citation: *Antonov AV, Ukraintsev VF, Chekhovich VE, Chepurko VA. On the planning of the scope of new technology testing. Dependability 2017; 3: p. 3-9. DOI: 10.21683/1729-2646-2017-17-3-3-9.*

Introduction

The process of new technology development and manufacture involves determining its dependability indicators. The most objective method of identifying dependability characteristics of products is field testing. This paper examines the $[N,U,T]$ test plan. This plan that involves testing N nonreparable samples within the time interval between 0 and a certain T. It is assumed that during the tests k objects fail, while N-k objects successfully pass the tests. Thus, at the outcome of the experiment we have a mixed sample that includes k times to failure and N-k right censored observation. It is clear that the accuracy of the obtained estimate will depend not only on the number of tested objects N, but also on the experiment duration. For a fixed N, as the observation time T grows the estimate accuracy increases due to the increasing proportion of complete times, while the proportion of censored ones goes down. It should be noted that when we talk about identifying the dependability characteristics of complex and costly objects we cannot test large batches of finished products. Therefore the problem consists in defining testing duration and size of the product batch to be tested subject to specified requirements for the accuracy of estimation of dependability characteristics obtained as the result of the tests.

Problem definition

The experiment is conducted according to the plan $[N,U,T]$. This plan involves testing N samples within a given time interval between 0 and T with no replacement of failed products [2–4]. During the test k failures are observed. Let us designate the obtained operation times as $t_1, t_2, \ldots, t_k$. ν products successfully pass the tests $\nu = N–k$ (ν is the number of nonfailed samples with the operation time T). The nonfailed objects make up the sample of right censored operation times. Figure 1 shows the plan of the experiment.

Thus, sample no. 1 worked without failure up to the moment in time T. The second sample failed at the moment t_2, etc.

Let us assume that at a certain moment t_0 the lower confidence bound with a specified confidence level $1–\alpha_0$ for PNF $P(t_0)$ must not be lower than $\underline{P}_0$, i.e.

$$\Pr(P(t_0) \geq \underline{P}_0) \geq 1–\alpha_0. \quad (1)$$

It is obvious that it can be achieved by selecting the test scope $N(t_0)$ only in the case when the actual PNF in this point $P(t_0)$ is to the right of $\underline{P}_0$. Otherwise the problem is unsolvable.

It would be logical to assume that at a certain moment in time T ($T \geq t_0$) in order to ensure an equal accuracy of PNF estimation in point t_0 the required scope of finished products tests $N(T)$ is at least equal to that defined for point t_0. That is primarily due to the fact that the PNF is a non-increasing function. Let us designate the required test scopes as $N(t_0)$ and $N(T)$.

The aim of the paper is to identify the test scope of a batch of finished products $N(T)$ under the condition of fulfilment of the manufacturer's requirement for compliance with the lower confidence bound of the probability of no failure with a specified confidence level $1 – \alpha$. During the test we will identify the correlations between the test scopes of $N(T)$ and $N(t_0)$ provided that the requirements for the accuracy of the results for different test durations are equal. During the tests we will consider various parametrizations of the failure rate function $\lambda(t)$.

In the process of solving the problem we will be assuming that the failure rate function is defined by one of the formulas [1, 2]:

$$\lambda(t)=\lambda; \quad (2)$$

$$\lambda(t)=\lambda_1+\lambda_2 t; \quad (3)$$

$$\lambda(t)=\lambda_1 t^{\lambda_2}. \quad (4)$$

The formula (2) (the rate is constant) is typical to exponential distribution of time to failure, the formula (3) is typical to the distribution function with linear failure rate, while the function (4) is typical to the Weibull distribution law.

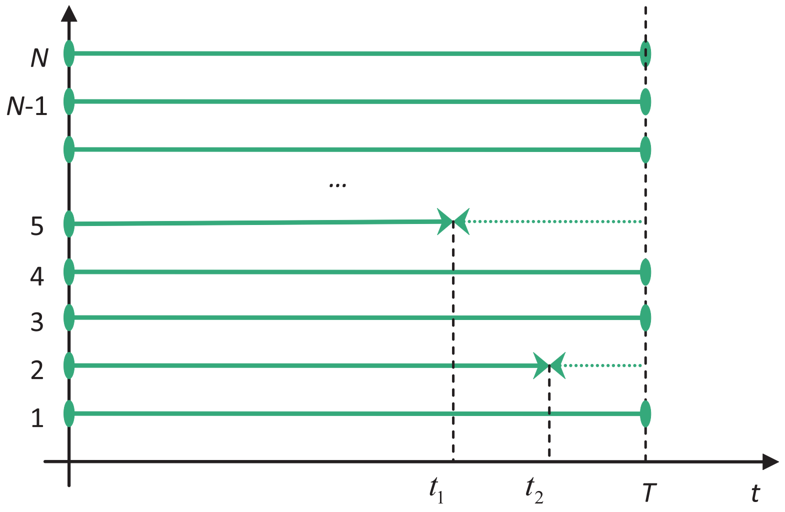


Figure 1. Plan of the experiment

In order to simplify the calculations, as in [1], let us transform the considered model as follows:

$$\lambda(t)=\lambda g(t) \qquad (5)$$

where g(t)=1 corresponds with the exponential distribution,

g(t)=a+bt corresponds with the distribution with a linear failure rate function, (6)

g(t)=t^a corresponds with the Weibull distribution. (7)

The function g(t) must meet two main conditions:

$$g(t)\geq 0,$$

while the integral function

$$G(t)=\int_0^t g(\tau)d\tau \to \infty \text{ if } t\to\infty.$$

We will assume that the coefficients a, b in (6), (7) are known, while the parameter λ is unknown and estimated by sample.

In the next section we will identify how the accuracy of this parameter's estimation depends on the duration T of the experiment and subsequently deduce the condition of its preservation under the chosen experiment plan.

Evaluation of parameter λ and identification of its accuracy

It is known [2] that the accuracy of an estimate obtained using the maximum likelihood method (MLM estimates) depends on the Fisher information quantity.

In the beginning, let us find the quantity of Fisher information contained in the initial statistics. The likelihood function that corresponds with the chosen plan of experiment [N,U,T] will be as follows:

$$L(\vec{t};T;\lambda)=\prod_{i=1}^{k} f_{t_i}(t_i;\lambda)\prod_{j=1}^{\nu} P_{t_j}(T;\lambda)=$$

$$=\lambda^k\cdot\prod_{i=1}^{k} g(t_i)\cdot\exp\left(-\lambda\left(\sum_{i=1}^{k} G(t_i)+\nu G(T)\right)\right),$$

where $f_t(t,\lambda)=\lambda g(t)\exp(-\lambda G(t))$ is the distribution density of time to failure.

Log-likelihood function:

$$l(\vec{t};T;\lambda)=\ln L=k\ln\lambda+\sum_{i=1}^{k} g(t_i)-\lambda\left(\sum_{i=1}^{k} G(t_i)+\nu G(T)\right).$$

We identify the partial derivative.

$$\frac{\partial}{\partial\lambda} l(\vec{t};T;\lambda)=\frac{k}{\lambda}-\left(\sum_{i=1}^{k} G(t_i)+\nu G(T)\right). \qquad (8)$$

Here $k=\sum_{i=1}^{N} \mathrm{I}\{t_i\leq T\}$ is a binomially distributed random value (r.v.):

$$Bin(N;F_t(T))=Bin\left(N;1-e^{-\lambda G(T)}\right). \qquad (9)$$

The information quantity (dispersion of the right part of the equation (8) will be defined by the sum of three summands. Let us find each individually.

$$\operatorname{var}\left[\frac{k}{\lambda}\right]=\frac{NP_t(T)(1-P_t(T))}{\lambda^2}=\frac{Ne^{-\lambda G(T)}\left(1-e^{-\lambda G(T)}\right)}{\lambda^2}.$$

$$\operatorname{var}\left[\sum_{i=1}^{k} G(t_i)+\nu G(T)\right]=\operatorname{var}\left[\sum_{i=1}^{N} G(t_i\wedge T)\right]=$$

$$=\frac{N\left(1-2\lambda G(T)e^{-\lambda G(T)}-e^{-2\lambda G(T)}\right)}{\lambda^2}.$$

$$\operatorname{cov}\left[\frac{k}{\lambda};\sum_{i=1}^{N} G(t_i\wedge T)\right]=$$

$$=\frac{1}{\lambda}\operatorname{cov}\left[\sum_{i=1}^{N}\mathrm{I}\{G(t_i)\leq G(T)\};\sum_{i=1}^{N} G(t_i)\wedge G(T)\right]=$$

$$=\frac{1}{\lambda}\left[\sum_{i=1}^{N}\mathrm{E}\left(G(t_i\wedge T)\cdot\mathrm{I}\{t_i\leq T\}\right)-\mathrm{E}\left(G(t_i\wedge T)\right)\mathrm{EI}\{t_i\leq T\}\right]=$$

$$=\frac{N}{\lambda^2}\left(e^{-\lambda G(T)}-e^{-2\lambda G(T)}-\lambda G(T)e^{-\lambda G(T)}\right).$$

By adding the dispersions to covariations we identify the Fisher quantity:

$$\mathrm{I}(\vec{t};T;\lambda)=\frac{N}{\lambda^2}\left(1-e^{-\lambda G(T)}\right)=\frac{N}{\lambda^2}(1-P(T))=\frac{NF_t(T)}{\lambda^2}, \qquad (10)$$

where $F_t(T)$ is the distribution function of time to failure.

Now, let us deduce the MLM estimate for λ:

$$\hat{\lambda}_T=\frac{k}{\sum_{i=1}^{N} G(t_i\wedge T)}. \qquad (11)$$

The estimate is acquired by equalling the right side of the equation (8) to zero. It is known that the MLM estimate is asymptotically unbiased, consistent, asymptotically efficient and asymptotically normal. Thus,

$$\hat{\lambda}_T\overset{as}{\sim} Norm\left(\lambda;\frac{1}{\mathrm{I}(\vec{t};T;\lambda)}\right)=Norm\left(\lambda;\frac{\lambda^2}{NF_t(T)}\right). \qquad (12)$$

In the next section we will research the Fisher information quantity.

Research of estimate $\hat{\lambda}_T$

Let us introduce the designation of the numerator of Fisher information quantity (10):

$$K(T)=NF_t(T). \qquad (13)$$

Obviously $K(T)$ will also depend on the parameter λ, because it affects the distribution function $F_t(T)$, yet we are primarily interested in the dependence of the introduced indicator on time. In terms of significance, $K(T)$ will be the expectation of r.v. k that is equal to the average number of failures by the moment of time T distributed according to the law (9).

Due to the well-known asymptotic property of the distribution function

$$K(T) \underset{T\to\infty}{\to} N. \tag{14}$$

Obviously sooner or later all N samples will fail. In this case the estimate (11) will tend to the estimate based on the complete sample:

$$\hat{\lambda}_T = \frac{k}{\sum_{i=1}^{N} G(t_i \wedge T)} \underset{T\to\infty}{\to} \hat{\lambda}_\infty = \frac{N}{\sum_{i=1}^{N} G(t_i)}. \tag{15}$$

If the test scope N is a time-independent constant, then the function $K(T)$ and function $F_t(T)$ will be non-decreasing, while $K(0)=0$.

Let is consider the condition that will ensure the achievement of the required lower bound of the PNF in point t_0. Under the chosen PNF estimation method its lower confidence bound will be $\underline{P}_0$. This value is identified by calculating the upper bound of the parameter λ: $\underline{P}_0 = e^{-\bar{\lambda}_{t_0} G(t_0)}$, where $\bar{\lambda}_{t_0}$ is the upper bound of the parameter λ calculated as if the testing of samples ended at the moment t_0. This indicator will be identified through the solution of the equation

$$1-\alpha_0 = P\left(\hat{\lambda}_{t_0} < \bar{\lambda}_{t_0}\right) \approx \Phi\left(\frac{\bar{\lambda}_{t_0}-\lambda}{\lambda}\sqrt{K(t_0)}\right), \tag{16}$$

where $\Phi(x)$ is the standard normal law distribution function Norm(0,1). By solving the equation (16) we obtain:

$$\bar{\lambda}_{t_0} = \lambda\left(1+\frac{u_{1-\alpha_0}}{\sqrt{K(t_0)}}\right).$$

If the test lasts to the moment in time T, the upper bound of the parameter will be identified through the solution of an equation similar to (16), where t_0 is replaced with T. The following formula is obtained:

$$\bar{\lambda}_T = \lambda\left(1+\frac{u_{1-\alpha_0}}{\sqrt{K(T)}}\right), \tag{17}$$

where $u_{1-\alpha_0}$ is the quantile of the normal law Norm(0,1) of the level $1–\alpha_0$.

Estimate (17) will be used in the calculation of the lower bound of the PNF in point t_0. According to the stated aim of the research it is required to ensure the fulfilment of the condition according to which the lower bound of the PNF calculated with the confidence level $1–\alpha_0$ was not lower than $\underline{P}_0$ regardless of the duration of experimental observations. As the failure rate is one-to-one expressible through PNF, a similar condition can be formulated for the rate. Therefore, for the random positive moment of time T the following can be written:

$$\bar{\lambda}_T = \bar{\lambda}_{t_0}. \tag{18}$$

Thus, it is required to choose the size of the batch of products to be tested in such a way as to ensure the upper bound of failure rate $\bar{\lambda}_T$ did not depend of the observation time. In other words, the accuracy of estimation of the parameter λ at the moment of time t_0 must be equal to that at the moment T. That can be achieved if the Fisher information quantity is assumed to be constant. This condition will be written as $K(T)$=const.

Condition of preservation of accuracy of estimate λ

The error is estimated at the moments t_0 and T will be identical if the information quantities in those points are equal. We will achieve the equality of the information quantities by selecting the required test scope $N(T)$ and $N(t_0)$.

The condition of equality of information quantities for two random moments in time t_0 and T will be as follows:

$$\mathrm{I}(\vec{t};t_0;\lambda) = \mathrm{I}(\vec{t};T;\lambda). \tag{19}$$

This condition (19) will ensure an estimation accuracy of the unknown parameter λ based on the findings of experiment $(N(T),U,T)$ as if we estimated λ based on the findings of experiment $(N(t_0),U,t_0)$.

Out of (10), (14) and (19) follow the properties:

$$\frac{N(T)}{N(t_0)} = \frac{F(t_0)}{F(T)} \text{ or}$$

$$K(T) = Const = K(t_0) = Ek(t_0) = K(\infty) = N(\infty). \tag{20}$$

Thus, for any T the constant $K(T)$ will be equal to both the expected number of failures at the moment t_0 and the number of samples $N(\infty)$. In the following section let us identify the constant $N(\infty)$.

Solution of the problem

Formula (15) defines the estimation of parameter λ if $T\to\infty$. Let us define the asymptotic number of samples $N(\infty)$ based on the accurate distribution of estimate $\hat{\lambda}_\infty$. It is known [8] that r.v. $2\lambda G(t_i)$ will have the ch-square distribution with two degrees of freedom: $2\lambda G(t_i) \sim \chi_2^2$. Due to the independence of summands:

$$2\lambda\sum_{i=1}^{N} G(t_i) \sim \chi_{2N}^2. \tag{21}$$

Proposition (21) allows constructing a right-hand confidence interval for parameter λ.

$$\Pr\left(\hat{\lambda}_\infty \le \bar{\lambda}\right) = \Pr\left(2\lambda\sum_{i=1}^{N} G(t_i) \ge \frac{2\lambda N}{\bar{\lambda}}\right) = 1-\alpha_0.$$

Table 1. Asymptotic values for the number of test (α_0=0.05)

q_0	0,051	0,178	0,273	0,342	0,394	0,436	0,469	0,498	0,522	0,543
$N(\infty)$	1	2	3	4	5	6	7	8	9	10

Table 2. Values of $N(\infty)$ depending on α and $\underline{P}_0$.

$\underline{P}_0$ / α	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,99	0,999
0,01	5,08	6,54	7,81	8,98	10,09	11,16	12,19	13,19	14,16	14,64	15,02	15,11
0,02	4,10	5,25	6,26	7,18	8,06	8,90	9,71	10,50	11,27	11,65	11,95	12,01
0,03	3,53	4,52	5,37	6,16	6,90	7,61	8,30	8,97	9,62	9,94	10,19	10,25
0,04	3,14	4,01	4,75	5,44	6,09	6,72	7,32	7,90	8,47	8,75	8,97	9,02
0,05	2,84	3,62	4,28	4,90	5,48	6,04	6,57	7,09	7,60	7,85	8,04	8,09
0,06	2,60	3,30	3,91	4,46	4,99	5,49	5,97	6,44	6,90	7,12	7,30	7,34
0,07	2,40	3,04	3,59	4,10	4,58	5,03	5,47	5,90	6,31	6,52	6,68	6,72
0,08	2,23	2,82	3,32	3,79	4,23	4,64	5,05	5,44	5,82	6,00	6,15	6,18
0,09	2,08	2,62	3,09	3,52	3,92	4,31	4,68	5,04	5,39	5,56	5,69	5,72
0,1	1,95	2,45	2,89	3,28	3,65	4,01	4,35	4,69	5,01	5,17	5,29	5,32

Out of which we obtain a transcendent equation to find $N(\infty)$:

$$\frac{\chi^2_{\alpha_0}(2N)}{2N} = \frac{\lambda}{\bar{\lambda}} = \frac{\ln P_0}{\ln \underline{P}_0} = q_0, \qquad (22)$$

where $\chi^2_{\alpha_0}(2N)$ is the quantile of ch-square distribution with $2N$ degrees of freedom of level α_0.

Table 1 shows an example of calculation of asymptotic values of $N(\infty)$ for various q_0.

We can use the asymptotic distribution of estimate (12) and by solving the equation (16) under N obtain the following results:

$$N(\infty) \approx \left(\frac{u_{1-\alpha_0} q_0}{1-q_0}\right)^2. \qquad (23)$$

The analysis of (23) has shown that approximate calculation is very optimistic (see Figure 2) and the estimate of the required test scope turns out to be significantly conservative.

If the PNF P_0 in point t_0 is unknown, we can, as in [1], assume that:

$$P_0 = \frac{1+\underline{P}_0}{2}. \qquad (24)$$

In fact, due to the significant asymmetry of the binomial distribution in case of highly dependable equipment the following inequation will be fulfilled: $P_0 > \frac{1+\underline{P}_0}{2}$. Therefore, $q_0 < \ln\frac{1+\underline{P}_0}{2} \Big/ \ln \underline{P}_0$ and the asymptotic value estimate $N(\infty)$ will be exaggerated, i.e. the estimate will be pessimistic.

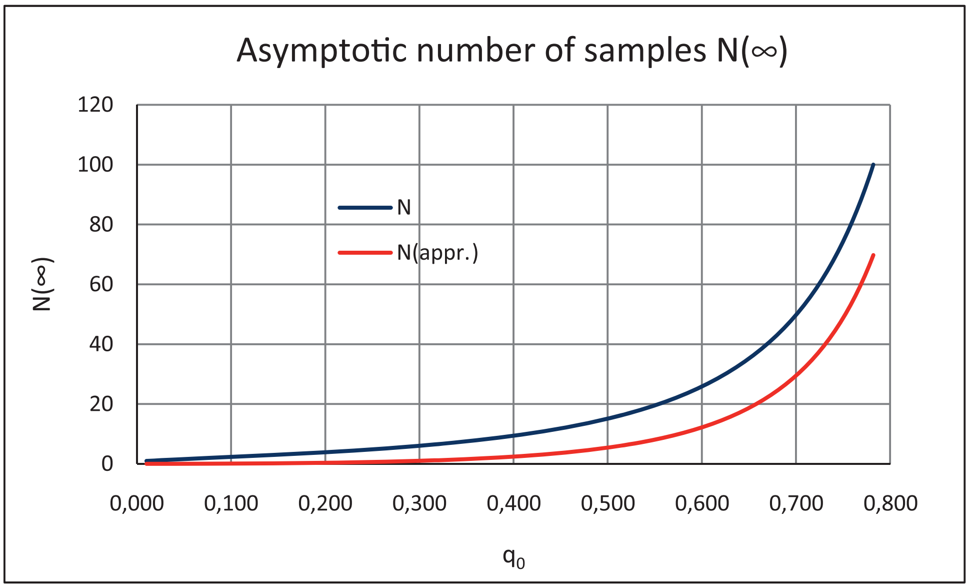


Figure 2. Exact and approximate values of $N(\infty)$

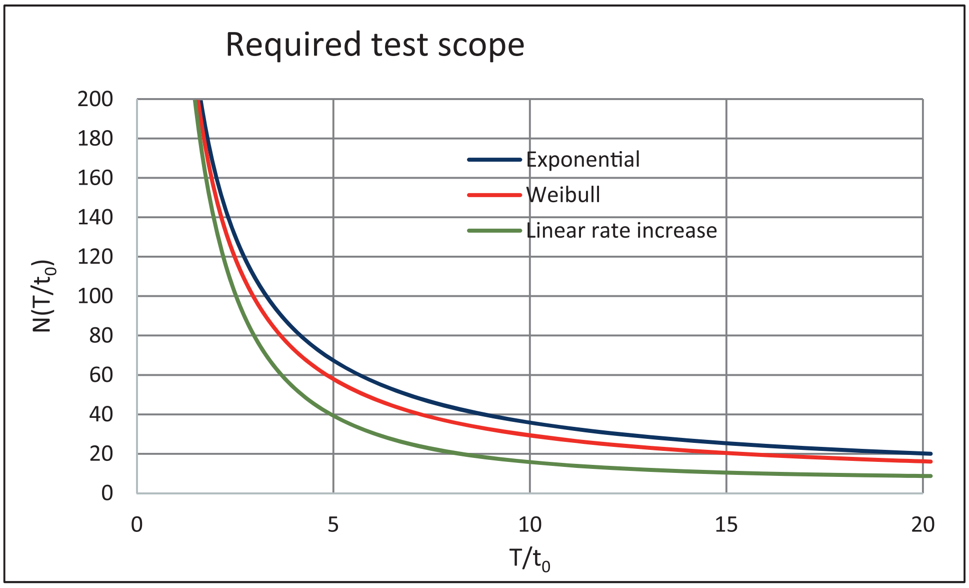


Figure 3. Dependence of the test scope on the time under different distribution models

Table 2 shows calculated values of $N(\infty)$ depending on the significance of α and lower bound of PNF $\underline{P}_0$.

In order to evaluate the test scope in random point t it remains to apply (20).

Out of (20) follows $N(T)=\frac{N(\infty)}{F(T)}$. As the true value of the parameter λ is unknown it can be evaluated based on the condition $\lambda=-\frac{\ln P_0}{G(t_0)}$. From which

$$N(T)=\frac{N(\infty)}{1-P_0^{G(T)/G(t_0)}}. \tag{25}$$

The time dependance in case of low $\lambda G(T)$ is almost hyperbolic under $G(T)$.

$$N(T)\approx\frac{N(\infty)}{\lambda G(T)}. \tag{26}$$

Fig. 3 shows the behavior of the required test scope (25) depending on T on the time scale T/t_0. The input parameters are as follows: $\underline{P}_0$=0.95; α_0=0.05, t_0=5, $N(\infty)$=7.85 (see Table 2). During the calculations the distribution parameters were chosen as follows: for the Weibull distribution the parameter a=1.1. For the model with linearly increasing rate function a=1; b=0.1.

It can be seen that models with an increasing rate function compared to those with a constant failure rate require a relatively smaller number of tests.

Conclusion

The obtained results allow for a substantiated approach to planning the scope of highly dependable objects testing. The input information is the manufacturer-supplied information on the requirement to confirm the lower bound of the product's PNF with given confidence level. The formulas obtained in the paper enabled the research of the dependence of the scope of testing from the experiment duration. It is shown that the longer is the experiment duration, the fewer products must be supplied for testing. The dependence is non-linear and conditioned by the parametrization of the failure rate function. New asymptotic results have been obtained that allow adequately assessing the number of tests for a given time period.

References

1. Antonov AV, Chepurko VA, Chekhovich VE, Ukraintsev VF. Regarding the planning of testing scope for new equipment samples. Dependability 2016;(3):3-7. DOI:10.21683/1729-2640-2016-16-3-3-7.

2. Antonov AV, Nikulin MS, Nikulin AM, Chepurko VA. Teoria nadiozhnosty. Statisticheskie modeli: Uchebnoie posobie [Dependability theory. Statistical models: A study guide]. Moscow: INFRA-M; 2015 [in Russian].

3. Gnedenko BV, Beliaev YuK, Soloviev AD. Matematicheskie metody v teorii nadiozhnosti [Mathematical methods in the dependability theory]. Moscow: Nauka; 1965 [in Russian].

4. Beliaev YuK, Bogatyrev VA, Bolotin VV et al. Ushakov IA, editor. Nadiozhnost tekhnicheskikh system: Spravochnik [Dependability of technical systems: Reference book]. Moscow: Radio i sviaz; 1985 [in Russian].

5. Antonov AV, Nikulin MS. Statisticheskie modeli v teorii nadiozhnosti; Ucheb. posobie [Statistical models in the dependability theory: A study guide]. Moscow: Abris; 2012 [in Russian].

6. Zarenin YuG, Stoyanova II. Opredelitelnye ispytania na nadiozhnost [Determinative dependability testing]. Moscow: Izdatelstvo standartov; 1978 [In Russian].

7. Cramer H. Matematicheskie metody statistiki [Mathematical methods of statistics]. Moscow: Mir; 1975 [In Russian].

8. Antonov AV. Sistemny analiz; Uchebnik dlia vuzov [System analysis. Textbook for higher educational institutions]. Moscow: Vysshaya Shkola; 2008 [in Russian].

9. David H. Poriadkovye statistiki [Order statistics]. Moscow: Nauka. Main office of physics and mathematics literature; 1979 [in Russian].

About the authors

Alexander V. Antonov, Doctor of Engineering, Professor. Obninsk Institute for Nuclear Power Engineering (IATE MEPhI), Professor of Automated Control Systems. Russia, Obninsk, e-mail: antonov@iate.obninsk.ru

Vladimir F. Ukraintsev, Candidate of Physics and Mathematics, Associate Professor. State Scientific Center of the Russian Federation Leipunsky Institute of Physics and Power Engineering, Lead Specialist. Russia, Obninsk, e-mail: ukraintsev@mail.ru

Vladimir E. Chekhovich, State Scientific Center of the Russian Federation Leipunsky Institute of Physics and Power Engineering, Head of Unit. Russia, Obninsk, e-mail: 89158916216@rambler.ru

Valeri A. Chepurko, Candidate of Engineering, Associate Professor. Obninsk Institute for Nuclear Power Engineering (IATE MEPhI), Professor of Automated Control Systems. Russia, Obninsk, e-mail: chepurko@iate.obninsk.ru

Received on 22.03.2017

Original article
DOI: 10.21683/1729-2646-2017-17-3-10-16

Particular characteristics of today's microelectronics and matters of highly dependable and secure control systems design

Aleksei P. Kirpichnikov, *V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russia*
Stanislav N. Vasiliev, *V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russia*

Abstract. Aim. *Drawing the readers' attention to the growing number of industrial disasters, associated damage, increasing human casualties and the connection of this phenomenon with computer-based automation systems. The authors produce arguments regarding the requirement for design technology with extended security features in view of the multifold growth of abnormal natural and industrial effects. The paper describes and analyzes distinctive features of control systems of critical application facilities and consequences of disregarding additional inspection of circuitry and software. Of special note is the growing risk caused by the introduction of unmanned technologies and their mass application in railway and automotive transportation. The paper examines the problems of control systems resilience to faults and external actions depending on the used components. Statistics of industrial disasters are provided, their connection with the indicators of control systems instability is examined. A special emphasis is put on the distinctive features of today's microelectronic components and the effects of technological progress on the systems' interference immunity and fault rate. Of note is the growing number of hazardous failures in systems based on 0.13-µm and lower microcontrollers. A significant attention is given to the research of the distinctive features of modern chips, their layout, particularly of the main element of a control system, i.e. the microcontroller and digital signal processor, the influence of the external effects on the chip. The matters related to CMOS layout in microprocessor-based units are considered, the dependance is shown between the rising noise influence and migration to new CMOS technology. Attention is drawn to the requirement to train an appropriate class of specialists able to work with such systems who have not only software engineering skills, but also profound knowledge of physics, fundamentals of control systems design and their stability.* **Results.** *A comparative evaluation of stability of 0.5 µm and 130 nm CMOS stability has been conducted. The resultant difference in threshold power of interference is over 4000 times. It is noted that most developers who design software for such systems are mislead by the non-availability of any public information on the fault rate of processing elements from the manufacturing companies. By taking the dependability figures as the main parameter they misjudge the safety integrity level, as instead of the fault rate parameters they erroneously use the microchip's dependability figures provided by the manufacturer. Additionally, standard methods of improving the safety level used by developers (e.g. redundancy) often prove to be inefficient.* **Conclusions.** *Designing highly dependable and safe control systems must take into consideration the distinctive features of today's computer components given the fact that new generations of modern microchips due to their fault rate characteristics are often unusable in highly dependable system design. It appears to be of relevance improving existing standards and developing new ways of increasing the stability and safety of systems. Also noted is the requirement of maintaining the level of education and awareness of a wide community of developers who work with control systems in transportation, energy, industrial automation, weapon systems, etc. as regards the importance of ensuring the required level of functional safety.*

Keywords: *computer-based control systems, microelectronic hardware components, industrial disasters, faults, SIL, safety unit.*

For citation*: Kirpichnikov AP, Vasiliev SN. Particular characteristics of today's microelectronics and matters of highly dependable and secure control systems design. Dependability 2017;3: p. 10-16. DOI: 10.21683/1729-2646-2017-17-3-10-16*

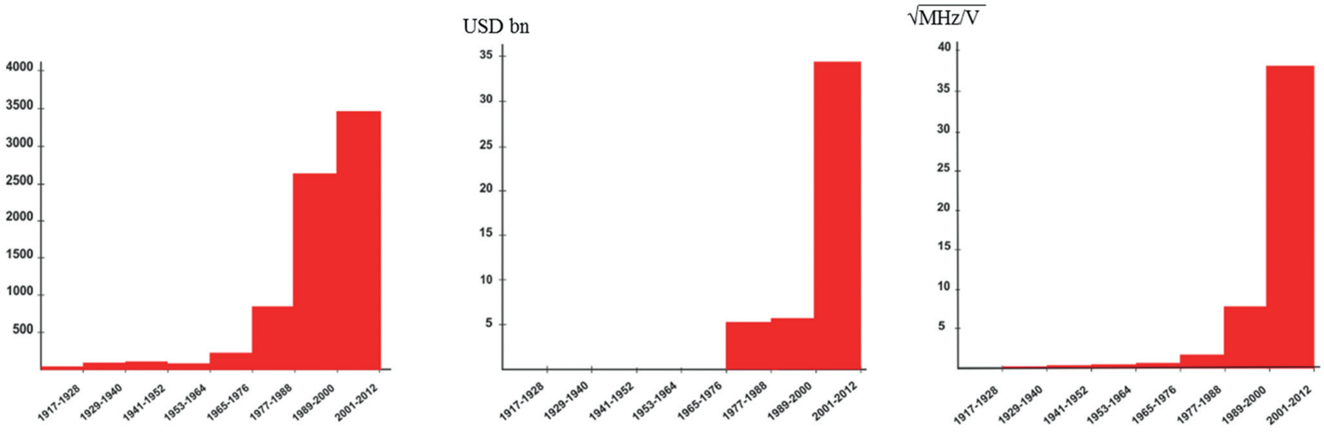
Introduction

The last few decades were marked by a menacing trend of constant growth of the number of industrial disasters that cause increasing damage and number of casualties. One of the primary reasons should be the general deployment of computer-control systems that replace old relay-based systems with no regard given to the specificity of the modern electronic components. The critical facilities control systems involve a very high level of safety that is practically unachievable with the new components without the use of very non-trivial methods. As the result, after modernization the facilities with the busiest operational schedules fall into the high risk group. Signal processing and safety systems design experts should pay special attention and be ready for very critical and intense work, which is not always typical for the modern times and unusual for the young generation. However, the result, i.e. saved human lives, is worth it. Therefore it is not recommended to leave such activities unattended, disregard additional inspection of circuitry and software (SW), on-receipt inspection of components and most importantly entrust the supply department with making decisions regarding alternative components. The consequences may be unpredictable and even tragic. The authors touched upon these aspects of safety in the context of railways and subways [1]. Moreover, for the teams of ICS RAS and OOO AVTEKS, the development of the BARS train control safety unit of 81-760 cars of the Moscow Metro was motivated by the ambition to counter the above mentioned trends.

It appears to be very important to achieve an adequate level of education and general awareness of a wide community of developers regarding functional safety. The requirement for extremely low probabilities of hazardous failures (e.g. 10^{-9} h^{-1} for SIL 4) makes the practical confirmation of the safety level by means of live testing of equipment practically impossible, which often pushes the matter into the realm of theoretical constructs and provides grounds for various legends and misunderstandings. Thus, a number of microcontroller and digital processor (DSP) manufacturers have announced new 90 nm and 65 nm SIL3 microchips with special error correction features. But all engineers know that microprocessors manufactured using old but still common 0.5 μm and 0.35 μm CMOS technology, provided there are no circuitry and layout design errors, de facto ensure fault rates as per SIL4 even under weak electromagnetic interference. Additional tests would improve those figures by two more orders of magnitude, which was rarely required in practice, when the matter was just in the failures due to dependability-related reasons. Modern 0.13 μm and lower microcontrollers and DSP, even given the embedded correction features in perfect laboratory conditions, are not always capable of complying even with SIL3 (10^{-7} h^{-1}) and in no case can be used in vital systems. This paper aims to at least partially clarify this matter.

Growth of the number of industrial disasters as the trend of the last decades

Let us start the examination with the industrial disasters. Their number (Figure 1) is obviously rising because it is proportional to the total number of technology units in use on the one hand and their growing contribution to the critical events on the other hand. The latter manifests itself both in the diversity of deployments (applications in facilities) and the variations of the step scale of consequences. As the result, we obtain a nonlinear (power) relationship. Despite the limited scale of the events, their effects are considerable and the specificity obvious: if supervision is disturbed (reliably operating safety systems are not in place) technology, especially energy-saturated, is elemental. The consequences of industrial disasters are locally always destructive, if only because technology (unlike volcanoes) is situated almost always near and among people. Of special significance in this



Number of registered industrial disasters (1917 – 2012)

Damage from industrial disasters (1917 – 2012)

Indicator of automation systems vulnerability to EMP (1917 – 2012)

Figure 1. Statistics of industrial disasters and changes in the automation systems stability

context is transportation and specifically railway and subway automation systems. Beside the regular risks the relentless progress (but in reality misunderstanding of the real situation and susceptibility to populism) dictates general deployment of automatics, unmanned technology, etc., which significantly increases the requirements for control systems and digital signal processing (DSP) units. Of special interest are the plans of mass deployment of unmanned automotive and air transportation (hopefully the GOST R IEC 61508 and ISO 26262 standards will not be disregarded in the process), which will have a significant impact on the safety of life.

Meanwhile, the statistics, even without those new and promising trends, look terrible. Let us take a look at the numbers, i.e. the statistics of industrial and natural disasters and their consequences over the last 100 years (based on data by the Centre for Research on the Epidemiology of Disasters, CRED) [2]. Thus, within 30 years between 1910 and 1940 (when the population of the Earth was about 2 bn people) the number of registered industrial disasters was just 162 and the number of people injured and killed was around 50 ths with the total damage amounting to about USD 102.5 mil. But within the same period between 1982 and 2012 (with the population of the Earth about 7 bn people) those numbers for the industrial disasters were about 6.7 ths with 4 mil people injured and killed and damage amounting to about USD 45.5 bn. In other words, the growth of the number of disasters was **35 times**, and **450 times** in terms of damage! The graphs shown in Figure 1 speak for themselves (for convenience, the data was processed with 12-year intervals, which is close to solar cycles on the one hand, and to the typical period of capital-intensive equipment modernization on the other hand).

Human thinking in population is characterized by delayed reaction (delay of one generation or more). On the one hand there is always «high-pass filtering» when small tactical events overshadow global trends, especially those where the time constant is more than several decades, and the generation develops a habit. That involves an adequate real-time reaction to current events from only a few professional communities at best, while the society at large ignores them.

How and by what means will the modern humanity try to react to the threat to its safety and the growing number of catastrophic events? Most probably with the same microprocessor-based protections, systems for information collection, processing and control. In this situation, we should be at least be concerned with purely professional matters of ensuring safety functionality, i.e. creation of technology that is capable to operate in adverse conditions when the probability of influencing factors abnormal in their magnitude grows multiple times. So, what is the typical resilience of automation systems to various effects (humidity, impacts, electromagnetic fields, etc.) and how was it changing over the past century? While in terms of mechanical strength and quality of manufacture the answer is obvious, the electromagnetic resilience must be evaluated. As the criterion let us take the conditional parameter of threshold effect that causes failure of such system [3, 4] measure it in units V / $\sqrt{MHz}$ in a similar way to noise spectral density. We omit intermediate calculations. The findings in the form of inverses are shown in the right-hand graph in Figure 1. A special attention should be given to the similar nature of the last two histograms: damage and instability of control systems.

Looking at the given graphs it becomes clear that standards must be overhauled and additional, non-market mechanisms of improving the stability of vital control systems must be discussed. Amidst the growing flow of catastrophe-inducing faults, «precedent» thinking with delayed reaction may fail. By taking that into consideration let us proceed to the examination of the specificity of modern microelectronic technology.

General problems of today's microelectronic components

The list of components used n the control and safety systems is quite long and includes a large number of active an passive elements: from resistors, capacitors and transistors to large scale integrated circuits and radio frequency modules. Per each section of the list we should consider the effect of modernized technologies (most importantly nanotechnologies) on the dependability of components, their resilience to environmental effects and change of the probability of error associated with the performance of functions in the circuit. As the result, in particular, a new type of defect of surface mounted components was identified that is associated with mechanical effects in the process of manufacture and operation of products and lack of protection of surfaces of the modules that house the components. For most electronic modules we have a situation when the assembly guidelines have hardly changed in decades, yet the density of the layout has grown, the components stripped of casings (if those remained they were reduced to thin coatings that do not provide protection from mechanical damage), i.e. according to old standards most modern modules would be considered microassemblies and would require respective additional protection that is disregarded nowadays. Another uncovered factor was the changed nature of breakdown of active components which is primarily due to the use of low-voltage technology and significantly thinner layers in the structure of semiconductors compared to previous generations of similar products. We should also note the general application of power products based on MOS transistors, which on the one hand reduced the loss of power during switching, but on the other hand significantly increased the probability of breakdown of such switches to closure (due to the above mentioned problems of nanotechnology introduction). The list could go on as we discussed the impact of each factor on the safety and fault tolerance of control systems, but we shall confine ourselves to what we have said.

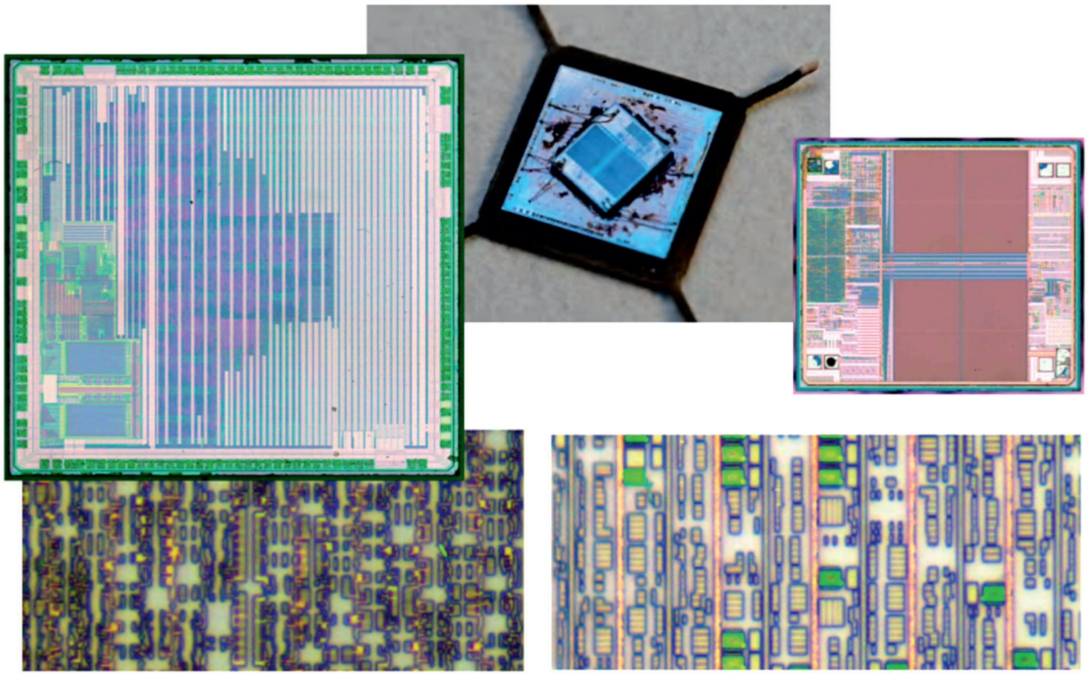


Figure 2. Cortex M3 microcontroller with the lid and pin matrix removed. On the left is a scaled image of a processor chip, on the right is FLASH memory, below, in more detail, are layout elements

Let us examine, even if briefly, the main element of a control system the processor unit (microcontroller, DSP, etc.) certainly is, the element on which depends the correctness of algorithm execution and safety of system's reactions.

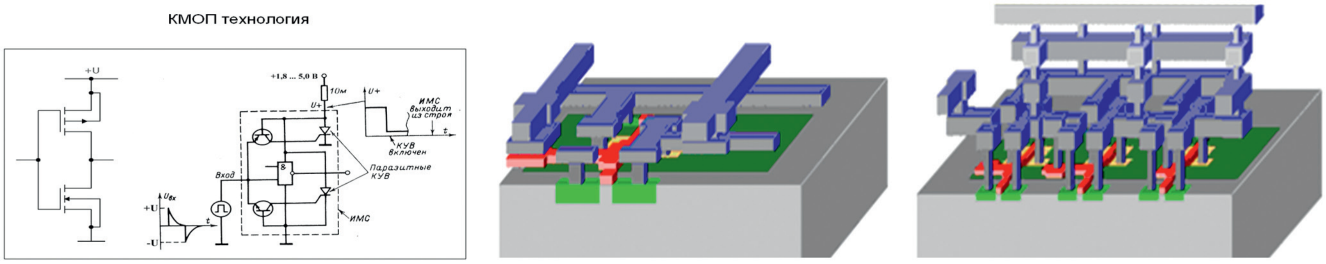
An example of such high technology component, that is well known but hardly recognizable in this form by IT experts, is shown in Figure 2.

Let us note that most developers who deal with system programming hardly understand what a microchip for which they design software is as an object in terms of physics and radio technology. As an example let us consider a popular Cortex M3-based processor, the successor to the well-proven ARM7. Figure 2 shows what such processor is without the case and the familiar pin matrix. Most importantly, we see not one, but two separate chips installed one on top of the other. The upper one is FLASH memory and is connected to the lower one (hardware core) with long links. The analysis of those connections immediately allows concluding on the difference in interference resistance of operations with RAM and FLASH due to those circuits alone. Other factors affecting the resilience will be the differences in chip technology, details of assembly, printed circuit boards (PCB), etc. However, our task does not consist in examining the features of a single component. Therefore let us, while briefly, touch upon the general problems of the CMOS technology one of the modifications of which is used in most modern microprocessor units.

Problems of the CMOS technology in modern microprocessor units

First of all, let us focus on a less popular fact: the CMOS logical gate, like any stage based on active elements at the moment of switching is an amplifier. For instance, in the case of an inverter (see diagram in the left-hand side of Figure 3) one transistor is the other one's load and the typical amplification ratio of such pair is usually around 80 – 100 in a broad band. In the right-hand part of the figure it is shown that this band changed from 90 MHz to 1.5 GHz with the transition from 0.5 μm to 0.13 μm. It should be taken into consideration that the threshold voltage and power supply of the transistors for the new technology is at least 4 times lower. That means that noise energy can be received by the circuit manufactured using the technology shown in the right-hand side part of Figure 3 from a band 16 times broader, given that its resilience to noise amplitude is at least 4 times lower.

While omitting intermediate calculations we note that the interference resistance of the currently used CMOS modifications with various design rules of external energy effects differs by orders of magnitude. The comparative evaluation of the **0.5 μm** and **0.13 μm** versions shown in the figure we obtain the value of the influencing interference that is different more than **4000** times **(!)**. The comparison of resilience for **0.5 μm** and **90 nm** shows the non-linear nature of this dependence, while the result of calculation is already over **15 000** times, all that is against the new chips that are so convenient for the users and software designers. That makes one wonder about the applicability of technology below 0.25 μm in critical application devices. It should be taken into consideration that the level of electromagnetic interference in modern civilization is very high in public places and energy-saturated infrastructure facilities. Thus, the operation of every safety device that is not resilient enough becomes a game of roulette.



Logic gate CMOS

0.5 μm, 2 layers of metal, up to 200K transistors, 90 MHz

0.13 μm , 7 layers of metal, up to 500 mil transistors, 1.5 GHz

Figure 3. CMOS technology: different density of metal layers and gates bandwidth for 0.5 and 0.13 μm

Having briefly considered the influence of environmental factors let us address the internal factors (i.e. those hidden within the microelectronic devices themselves). If in the context of previous generations of microprocessor devices the effect of radiation on the chip caused by ceramic elements of the case that with finite probability induced faulty switching of the chip's gates was discussed as a significant factor, now the definitively dominant effect is the poor magnetic compatibility of the densely situated elements and units of the chip itself, as we can see in Figures 2 (below) and 3. The factor of growing fault rate is the close location of the active elements and dense multilayer connections with the effect of crosstalk. It should be noted that for most modern processors the maximum length of a connection in the chip is 1 cm and more, which is also times more than the typical length of connection for the previous generation of circuitry.

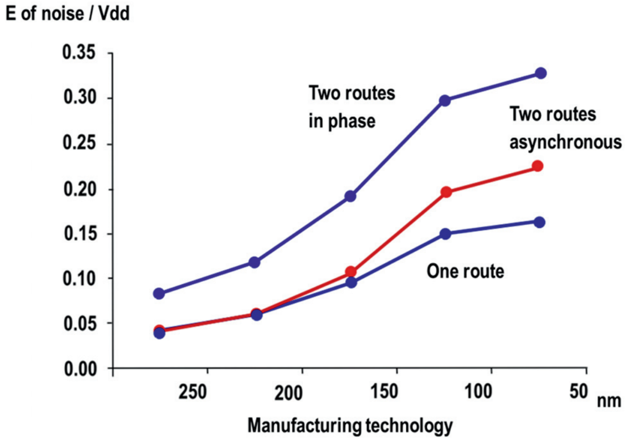


Figure 4. Dependence of the relative value of disturbance voltage on supply rail on the 1 mm-long routes

Figure 4 shows the dependence of the relative value of disturbing voltage in the power rail from 1-mm fragments of routes under topologically minimal distance between them as per design rules. The upper curve: signal in the source of interference changes in phase, middle curve: asynchronously, lowed curve: only one conductor with interference exists. Having taken into consideration the actual number of closely spaced routes on the chip and their lengths and having calculated the probability of peak values for the chosen noise model we obtain the approximate number of gate switching threshold overruns and an estimate of the fault rate of a specific microprocessor architecture (naturally, with no regard for the error compensation). Even this rough two-dimensional model yields results that are hardly compatible with the safety requirements. The results of 3D modelling of individual fragments of circuitry provide even more discouraging results. Additionally, large chips display a significant growth of loss current as an extra destabilizing factor that also indirectly influences the probability of faults.

That caused the interference resistance to fall and the fault rate to rise so much that it became noticeable in simple real-time systems (not safety-critical). The manufacturers were forced to use special fault compensation measures, i.e.: data surveillance and error correction units embedded in the chip, vital signals redundancy, etc. As the result, 65-nm chips with guaranteed SIL2 have appeared (NB, provided EMC in the equipment is perfect!). In other words, this has nothing to do with functional safety, as for products like that it counts as an achievement if some structures on one chip do not interfere too much with each other and can operate well under the fault rate of 10^{-6} h^{-1}. But that is absolutely not enough to build with their use even SIL2 safety system. Among the factors that radically affect safety will, for instance, be any electromagnetic effect with an appropriate probability of peak level within the considered time interval.

The usual practice of redundancy as a way to improve safety for such units also turns out to be not very effective due to fundamental reasons: under this technology, most faults [8] are common cause failure (CCF). Thus, even a low-power external electromagnetic effect that matches the frequency spectrum of a module's chips and connections "antennas" causes avalanching faults that cannot be mitigated with standard means of error correction that are designed to deal with non-numerous or isolated events. An effect threshold of sorts emerges, when the system's behavior becomes unpredictable. In case of safety systems it should be deemed absolutely intolerable and such components unusable. Their use must be confined to consumer and telecommunication devices, where the fault rate defines the availability of service and signal quality, but is in no way related with the safety of life.

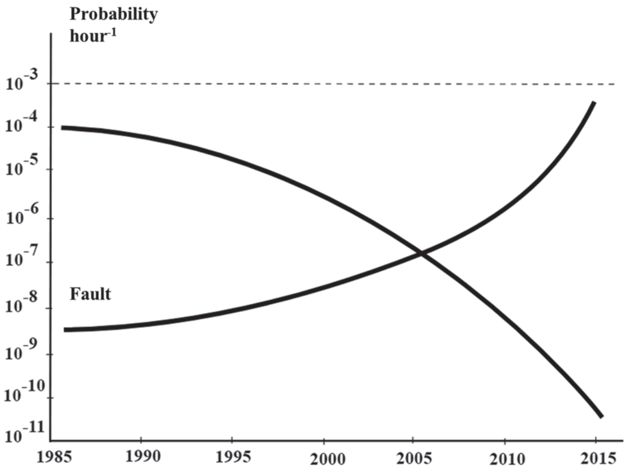


Figure 5. Evolution of the correlation between the failure probability and fault probability in microprocessor generations

As a result we have an unusual situation (Figure 5) when the dependability of microelectronic components is many times higher than their fault rate. So virtually the least dependable component of today's electronic module is the PCB and connections, while the least unsafe is the processor element and sensitive digital circuits. In terms of fault rate, the most advanced electronics have achieved the level typical to human operators (i.e. we have achieved perfection and reproduced ourselves!). Fault rate data for most processor elements is not available in standard documentation and can be obtained only after long and tiring negotiations with the manufacturer (if it even bothers with testing the products and obtain such data). As the result, the developers while evaluating the safety integrity level (SIL) erroneously use the microchip's dependability figures provided by the manufacturer instead, which is absolutely incorrect. As to the fault rate numbers, the information is confidential and is not discussed by the manufacturer.

This situation has a host of consequences. For instance, the replacement of a microcontroller with a newer version and fully SW-compatible can radically change the product's SIL. Additionally, beside the processors there are additional interface circuits, analog-to-digital converters, clock speed generator and other microchips that contain digital modules manufactured using the technology the developer is not aware of that have their own fault rate numbers. Some of those faults can be mitigated algorithmically by the processor itself. But what can be done if the source of the fault is a phase jump of the reference signal, that causes totally uncontrollable consequences for digital systems? That can happen during a simple replacement of one component (a generator) in the specification with another one that is identical in terms of appearance and specifications (e.g. upon the supply departments' advice).

Figure 6, as an example, shows five surface mounted board (SMB) generators of uniform application, i.e. intended for the same device, but by different manufacturers. Next are shown the results of "dissection" after the removal of the resonator. We can see digital circuits with totally different layout solutions, while the way the chip is assembled subsequently implies radically different interference resistance.

In conclusion, let us note that the development of microprocessors for SIL3 and let alone SIL4 control systems is hardly an easy task, especially for today's design teams. Those teams are mostly composed of IT experts, who tend to disregard the physical and even radio technical effects in their work. This may prove to be fatal in the case of critical application devices and safety-critical applications...

References

1. Vasiliev SN, Kirpichnikov AP, Botvinionok AA. Problemi obespechenia bezopasnosti v sovremennykh mikroprotsessornykh sistemakh oupravlenia podviznym sostavom, vyzvannye osobennostiami sovremennoy elementnoy bazy, i ikh reshenie na primere bloka bezopasnosty "BARS" vagonov 81-760 Moskovskogo metropilitena [Challenges of ensuring safety in today's computer-based train control systems caused by the specifics of modern computer components and their solution as in the case of the BARS safety unit of 81-760 cars of the Moscow Metro]. Bulletin of the JSC RZD Joint Academic Board 5:13–25 [in Russian].

2. Centre for Research on the Epidemiology of Disasters (CRED) <www.emdat.be>.

3. Kirpichnikov AP. Voprosy otkazoustoychivosty i bezopasnosty v oustroystvakh TsOS kriticheskikh prilozheniy [Matters of fault tolerance and safety in CSP devices of critical applications]. In: Proceedings of the Fourteenth

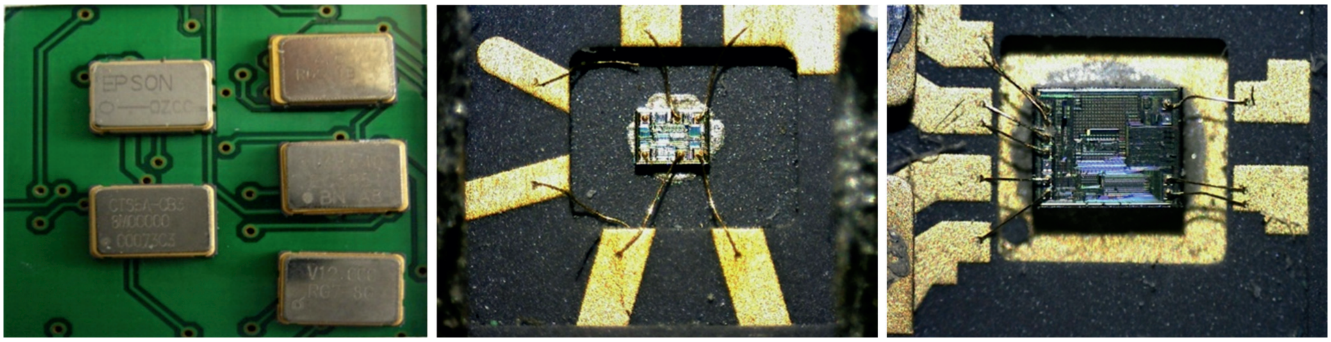


Figure 6. Interference resistance of MD generators Left to right: generators in cases SMD (5 types); low interference resistance version version acceptable in terms of process engineering and design

International Conference Digital Signal Processing and its Applications. Volume 1. Moscow (Russia); 2017. p. III–V [in Russian].

4. Kirpichnikov AP. Novaiya rol mikroprotsessornykh system: obespechenie bezopasnosti pered litsom katastrof [The new role of computer-based systems: ensuring safety in the face of catastrophes]. In: Proceedings of the Sixteenth International Conference Digital Signal Processing and its Applications, DSPA-2014. Volume 1. Moscow (Russia); 2014. p. 25-29 [in Russian].

5. Patent No. 2439666 RF. Kirpichnikov AP. Safety unit with validity checking of input information, 2010.

6. Patent No. 2449900 RF. Kirpichnikov AP. Safety unit, 2010.

7. Kirpichnikov AP, Botvinionok AA, Medunitsin NB. Mnogokanalnaya mikroprotsessornaya systema oupravleniya so sverkhvysokoy bezopasnostiu dlia poiezdov Moskovskogo metropolitena [Multichannel computer-based control system with ultrahigh dependability for the Moscow Metro trains]. Datchiki i sistemy 2014;9:38-45 [in Russian].

8. Shubinsky IB. Funksionalnaya nadiozhnost infromatsyonnykh sistem [Functional dependability of information systems]. Moscow: Nadiozhnost; 2012 [in Russian].

About the authors

Aleksei P. Kirpichnikov, head of unit, V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Russia, Moscow, phone: +7 (495) 334 89 10, e-mail: abramo@ipu.ru

Stanislav N. Vasiliev, member, RAS, Doctor of Physics and Mathematics, Chief Researcher, V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Russia, Moscow, phone: +7 (495) 334 89 10, e-mail: snv@ipu.ru

Received on 20.03.2017

Ensuring dependability of unique highly vital systems

Yuri P. Pokhabov, *Joint Stock Company NPO PM – Maloe konstruktorskoye buro, Russia, Krasnoyarsk Krai, Zheleznogorsk*

Abstract. Aim. *Dependability of products is usually researched with no regard to its genesis, while the causes of undependability are conventionally regarded as generalizing stochastic relationships that take into consideration "the result of interaction of a number of factors: the environment, system properties, process-specific, operational and other requirements." Consequently, the evaluation of dependability indicators is based on the assumption that by the beginning of operation the product is in working order. Respectively, the relations between the dependability and the time are considered only for the product operation period. The best known dependability-to-time relation is the empirical failure function, the so-called U-shaped dependability curve, which no one yet was able to describe with simple mathematical formulas usable in engineering calculations. The presence of the first "hump" in the U-shaped curve is associated with the manifestation of design errors, manufacturing defects or incorrect assembly of products, yet the specific causes of this "hump's" existence are not clarified in publications. The definition of the term "operability" does not rule out, and in practice there are often cases when design and development activities do not cover all the parameters that characterize the product's ability to perform the specified functions or when some of the documented requirements are not coordinated with the values of functional parameters, while during manufacture the values of such parameters may exceed the specified limits. As the result, a seemingly operable structure that passes experimental development may not be fit in terms of specified dependability indicators.* **Methods.** *The dependability properties of any product are specified long before the operation and can only fully manifest themselves after its beginning. The paper shows a graph that reflects the conditional probability of fault-free operation per lifecycle stages of products long before the beginning of operation. The dependability of unique highly vital systems (UHVS) may be ensured from the very early lifecycle stages based on consecutive execution of certain design, process engineering and manufacturing procedures, as well as application of engineering analysis of dependability.* **Results.** *The paper examines the role and significance of each lifecycle stage in ensuring UHVS dependability. The procedures of the engineering method of ensuring dependability are listed, the principles of UHVS design principles are set forth. Basic tools for increasing dependability and its evaluation principles are shown.* **Conclusions.** *The paper shows the possibility of ensuring the dependability of UHVSs using engineering procedures implemented at each lifecycle stage before the beginning of operation. Such procedures would enable an adequate level of design, development, preproduction, manufacture, as well as the development of a UHVS dependability evaluation method based on a single theoretical and methodological basis.*

Keywords: *unique highly vital system, ensuring dependability, dependability analysis, dependability evaluation, lifecycle stages, design, development.*

For citation*: Pokhabov YuP. Ensuring dependability of unique highly vital systems. Dependability 2017; 3: p. 17-23. DOI: 10.21683/1729-2646-2017-17-3-17-23.*

Introduction

Dependability of products is usually researched with no regard to its genesis, while the causes of undependability are conventionally regarded as generalizing stochastic relationships that take into consideration *"the result of interaction of a number of factors: the environment, system properties, process-specific, operational and other requirements."* [1] Consequently, the evaluation of dependability indicators is based on the assumption that by the beginning of operation (moment when operation time calculation begins) the product is in working order [2] and if t is the total operation time, while τ is the product's operation time to first failure, then the probability of no-failure (PNF) over time t is defined as follows:

$$P(t)=P(\tau>t). \qquad (1)$$

Respectively, the relations between the dependability and time are considered only for the product operation period. The best known dependability-to-time relation is the empirical failure function, the so-called *U*-shaped dependability curve [3], which no one yet was able to describe with simple mathematical formulas usable in engineering calculations. The presence of the first "hump" in the *U*-shaped curve is associated with the manifestation of design errors, manufacturing defects or incorrect assembly of products, yet the specific causes of this "hump's" existence are not clarified in publications [4].

In this context it is appropriate to recall a quizzical remark that I.A. Ushakov makes in his informal history of the dependability theory: "Dependability is calculated by people who cannot achieve it [5]." Indeed, in terms of theory the dependability of any complex technical system is a multidimensional problem, the definition and solution of which requires taking into consideration a multitude of interdependent parameters, stochastic in their nature, which is practically impossible to implement. At the same time, engineers have learned to practically achieve a more or less acceptable level of reliability of complex technology by using qualitative dependability criteria [6]. For instance, the designers of deployable structures of spacecraft know well that the product that is being designed, firstly, must be solid not to break before or during the loading, secondly, it must be operable so that the design allows deploying after flight loads, and thirdly, dependable in order to ensure stability of deployment time and again in given modes and conditions of operation.

As it is known, the state of operability *is a state of an object under which the values of all parameters that characterize the ability to perform the specified function comply with the requirements of regulatory technical and/or design documentation* [7]. Obviously, the definition of the term "operability" does not rule out cases, and in practice there are often cases when design and development activities do not cover all the parameters that characterize the product's ability to perform the specified functions or when some of the documented requirements are not coordinated with the values of functional parameters, while during manufacture the values of such parameters may exceed the specified limits. As the result, a seemingly operable structure that passes experimental development may not be fit in terms of specified dependability indicators. A prime example is the repetitive non-deployment of solar array panels of the Soyuz TMA-14M (in 2014) and Soyuz TMA-17M (in 2015) spacecraft due to jammed array mounting elements.

The role of the lifecycle stages preceding the operation in ensuring dependability of unique highly vital systems

In [8-10] the authors show the impossibility of developing UHVSs from the perspective of ensuring specified dependability with no account for the principles of its genesis. The dependability properties of any product are specified long before the operation and can only fully manifest themselves after its beginning.

As it is known from practice while in the state of expectation of operation any objects bear the risk γ of failures due to design, process engineering and manufacturing errors that may reach 80 % [10-11]. Up to 80 to 85 % of costs in the machine building industry is defined by the design solutions that are created in the process of technology design and development [12].

In [10] the impact of design and process preproduction on the dependability is examined and it is suggested to evaluate UHVS PNF (1) as follows:

$$P(t)=(1-\gamma)\cdot P(\tau>t). \qquad (2)$$

Formula (2) is to focus the developer's attention on the initial lifecycle (LC) stages, i.e. the design and development that are the only stages at which it is possible to take such design solutions that will ensure maximum dependability of the future product. At further LC stages of the product such opportunities do not present themselves, not to mention that *"it is impossible to improve technology dependability in the course of operation"*[6].

In one of the oldest Russian standards – GOST 2.103 – the stages of design documentation (DD) release are divided into the development of detailed design documentation (DDD) and working design documentation (WDD). The DDD stage in turn consists of three stages, i.e. technical proposal, draft design and engineering design. Each stage of design and development performs strictly defined tasks and has a quite specific significance that is implied by the definitions sets forth in the respective engineering regulations:

1) Design is *the process of description required for the creation in given conditions of a not yet existing object based on the initial description of such object and/or the algorithm of its operation or the algorithm of transformation (in some cases repeated) of the initial description, optimization of the specified characteristics of the object and the algorithm of its operation or the process algorithm, elimination of errors in the initial description and consecutive presentation (if required) of descriptions in various languages*. [13].

The design stage corresponds with the development of the technical proposal and/or draft design and its deliverable is the DDD of the technical proposal as per GOST 2.118 and/or draft design as per GOST 2.119.

2) Development is *the stage of design preproduction performed using a CAD system during which a detailed 3D model of the product is developed, along with 3D models of units, assemblies and primary (basic) parts, that are used in the preparation of 2D projections (drawings), improved design calculations and modelling. The deliverables are completed as information objects that are placed in the integrated information environment. As per GOST 2.120 this stage and its deliverable are called engineering design* [14].

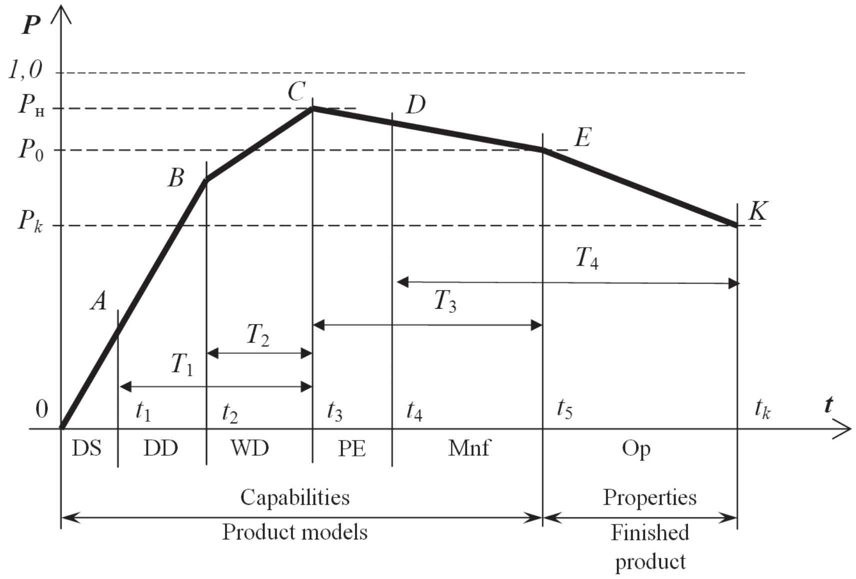
3) Development *is a stage of design preproduction performed using a CAD system, during which 3D models of all original parts and their 2d projections (drawings) are developed, specifications and bills for materials, components and standardized products are formalized, checking calculations an modelling are performed. The deliverables are completed as information objects that are placed in the integrated information environment. As per GOST 2.103 the deliverable of this stage is WDD* [14].

The modern "managerial" approach to solving industrial engineering problems is based on declarative reduction of DD development time primarily through the reduction or even omission of the DDD stage. As the results, the stages of design and development are often lumped together and are presented as *a process of design and development as a set of processes that ensure translation of requirements into specified characteristics or product, process or system specifications* [15]. The design and development process itself is divided into stages: design as *the process that translates the requirements into product characteristics set forth in the detailed design documentation* and development as *the process of development of design technical documentation for the product for subsequent preproduction and product manufacture* [16]. In some cases due to "time constraints" or sometimes due to thoughtlessness the design documentation developed in such abridged manner is handed over directly to the manufacturing facility for product manufacture while omitting the stage of process preproduction.

The potential results of such "managerial" approach in terms of achieving specified UHVS dependability can be seen in the figure that shows the conditional PNF (CPNF) per LC stages [10].

The figure reflects the general (qualitative) nature of UHVS development across LC stages subject the provisions of engineering regulations, generally accepted rules, the Common System of Design Documentation (CSDD), Common System of Process Documentation (CSPD) and quality management system (QMS), e.g. ISO 9001. The angles and shapes of the UHVS temporal variation curve in each specific case of design and development of products may somewhat differ from the shown graph, while retaining the general trend. The location of points *A*, *B*, *C*, *D* and *E* on the y axis depends on the adequate performance of the dependability procedures, which may not only largely reduce, but, in case of improper performance, significantly increase the risks γ of failures due to design, process engineering and manufacturing errors.

The figure reflects an important detail, i.e. it clearly shows the distinction between the LC stages, at which the future product exists in the form of a model and is characterized by the capability to manifest the property of dependability, and stages, at which the model materialized as finished product does manifest the property of dependability. This division allows the following:



0-t_1, release of the operational requirements and/or design specifications (DS), t_1-t_2, design and development of the product (DD), t_2-t_3, release of production drawings (WDD), t_3-t_4, manufacturing preparation (PED), t_4-t_5, product manufacture (PM), t_5-t_k, operation of finished product (OFP)

Figure 1. Graph of UHVS CPNF development across LC stages

– visualizing of hidden causes of the first "hump" in the well-known *U*-shaped curve of dependability;

– gaining the capability to compare the initial value of CPNF P_0 at the beginning of operation with the current values of CPNF in the process of UHVS development, which enables standardization of dependability per LC stages based on the specified PNF value at the end of operation P_k.

The significance of dependability procedures at lifecycle stages

The position of points *A* and *B* on the graph reflects the presence (absence) of "gross" errors due to the progress of fundamental research in the properties of structural materials, acquisition of reliable information on the external factors and loads (for point *A*), rationality of the chosen structural design solutions, observation of the design principles and rules (for point *B*) [10].

In case of absence of "gross" errors in the design the position of point *B* on the y axis may be close to 1, but not reach it due to two groups of causes. One of the groups of causes is associated with the project activities integration process that may last throughout the design stage and usually causes insufficient elaboration of the scope and content of requirements for product manufacture that are supposed to ensure specified dependability [9]. The second group of causes is set forth in [12]. It consists in various inevitable "small" errors due to the imperfection of the design methods, non-observance of regulatory technical documentation, insufficient qualification of the designers, their psycho-physiological properties, i.e. lack of attention, working speed, overall tiredness, etc.

Underestimation or disregard for the design stage significantly increase the risk of failing to achieve specified dependability indicators. If the above "minor details" are not properly dealt with during the design stage, the position of point *C* on the y axis may remain unchanged or be even lower that the position of point *B*. The aim of the WDD stage is to improve UHVS dependability by correcting the design errors and establishing required and sufficient requirements for manufacture. The position of point *C* on the graph reflects the maximum possible level of dependability P_d for this design that at subsequent LC stages can only decrease.

The position of points *D* and *E* is defined by the probability of errors of process preproduction and product manufacture. QMS is very important in this context as it is supposed to improve the production practices while reducing to an acceptable level the probability of production error. It is important to realize that the QMS in place at a manufacturing facility does not directly reflect on the quality and dependability of the finished product itself, as it is in practice a declaration of the fact that the enterprise is capable of releasing products of sufficient quality. Without proper design and process engineering support the QMS itself is incapable of solving the dependability problem, yet without proper QMS it is impossible to ensure dependability.

The position of point *K* is defined by the probability of errors of product operation. As in accordance with GOST 2.102 operational documentation is part of the DD the above errors are defined, on the one hand, by the establishment of clear requirements for observance of operating procedures, and, on the other hand, their proper observance.

In theory, for UHVS the position of points *C*, *D*, *E* and *K* on the y axis may reach its maximum possible values close to 1 upon condition of sufficient development and implementation of design and process engineering methods of dependability analysis and assurance [9]. The significance of activity per LC stages according to the figure is as follows:

– graph section 0-*A*-*B* is the elimination of "gross" design errors;

– graph section *B*-*C* is the correction of "small" design errors;

– graph section *C*-*D* is the elimination of errors of process preproduction;

– graph section *D*-*E* is the prevention of manufacturing defects;

– graph section *E*-*K* is the elimination of errors in operation;

Basic dependability method

The idea of design and process engineering support of dependability consists, on the one hand, in the implementation of procedures aimed at establishing required and sufficient DD requirement and ensuring the compliance with those requirements in manufacture. On the other hand, it consists in providing formalized confirmation of compliance with all design, process engineering and manufacturing procedures by means of associated analyses. The dependability analyses are considered the most important and integral part of the UHVS dependability methodology.

The design and process engineering analyses and dependability procedures are based on the common foundation, i.e. status and quality of the prepared DD and process engineering documentation (PED) and common principles of procedure performance based on the logical formula: done→to be confirmed, what's done→to be documented. The common foundation of design and process engineering analysis and dependability methods allows developing a methodology that can be used for three purposes, i.e. as a roadmap of design and development, a tool for verification of design and development and as a peer review tool. Thus, by using the design and process engineering methods of analysis and dependability the efficiency of the development process can be improved through division of powers and elimination of conflicts of interest between the developer who constantly thinks how the system will work and the dependability expert, a critical inspector of sorts, who must think how the system will fail to work [9].

It is important to note that design and process engineering dependability analysis (DPEDA) aims to research human decisions and errors (by designers, process engineers, production engineers) throughout consecutive LC stages, the FMEA method widely used in the West and its Russian counterpart, AVPKO, are designed to research product properties and processes. Unlike FMEA the association of the results of human activities with the end of each LC stage allows integrating DPEDA methods and the *Stage-Gate*-based design for reliability (DFR) methods [17].

From the point of view of dependability procedures the design and process engineering methods must be used alongside the QMS requirements fulfilment. In the verification of dependability requirements DPEDA must be used with other analysis methods in a strict order: functional analysis (FA), worst case analysis (WCA), DPEDA itself and dependability analysis (estimation) (DA) [9] as the results of each previous analyses serve as the source of data for the subsequent analysis.

Procedures of the engineering method of ensuring dependability

The design and process engineering support of dependability includes 4 procedures shown in the figure:

1) Procedure T_1 is to substantiate the finding within the set limits the values of parameters and indicators that characterize the ability to perform the required functions in specified modes and conditions of operation. The procedure is based on engineering calculations performed according to the most appropriate methods (strength and stiffness analysis, thermal analysis, dimension chain analysis, etc.) by any appropriate means: deterministic, semi-probabilistic or probabilistic method [18]. The duration of procedure T_1 includes the time of DDD and WDD development. The calculations are performed iteratively with the elaboration and detailing of the design, e.g. from analytical estimation of strength using beam idealization to numerical evaluation of full-size *3D* models with the finite elements method. This procedure as part of strength analysis of fixed (non-reconfigurable) structures is considered to be dependability calculation if it is performed using semi-probabilistic or probabilistic methods;

2) Procedure T_2 is used to establish DD requirements of which the fulfilment during manufacture ensures unconditional identification of the values of indicators and parameters with the specified tolerances. As the result of procedure T_2 performance each parameter (indicator) in DD must correspond with specified requirements in graphic or text form which eventually will ensure unconditional performance by the product of its functions;

3) Procedure T_3 serves to ensure guaranteed fulfilment of DD requirements at the stage of process preproduction and product manufacture. The function of this procedure is to eliminate any distortions and interpretations by process engineers and production engineers of dependability requirements stipulated in DD and to confirm the fact that the design, process engineering and metrological methods of manufacture, assembly, and supervision are based on single principles;

4) Procedure T_4 serves to ensure supervision of DD requirements fulfilment by the supervisory services at the manufacturing facility.

All 4 procedures are considered as a single and indivisible set of processes that ensure the fulfilment of specified dependability requirements. If for some reason PED is not prepared, the process engineering component of DPEDA may not be performed up to the moment of PED development. Then all conclusions regarding the product dependability are based on the assumptions that the manufacturing environment allows manufacturing the product in strict compliance with DD, i.e. errors of process engineering and manufacture are impossible. In this case formula (2) transforms as follows:

$$P(t)=P(A|B)\cdot P(\tau>t), \quad (3)$$

where A is an event that characterizes the readiness of the product to operate without failure allowing for the risk of malfunction due to design errors; B is the event that characterizes the readiness of the product to operate without failure allowing for the risk of malfunction due to process engineering and manufacturing errors.

Formulas (2) and (3) are connected with this formula:

$$P(A|B)>1-\gamma,$$

i.e. the dependability estimation based on the DD analysis alone will always be exaggerated.

Unique highly vital systems design principles

Design is considered as a sum of two equally significant processes that are implemented from the perspective of single principles of implementation of the procedures T_1 and T_2: visualization of the future product in the form of drawings (2*D* projections) or 3*D* models and parametric modelling (digitization) of the structure.

Structure digitization consists in the generation of column vectors of parameters (indicators) X and tolerances ΔX:

$$X=(X_1 \ldots X_n)^{\mathrm{T}}, \quad (4)$$

$$\Delta X=(\Delta X_1 \ldots \Delta X_n)^{\mathrm{T}}, \quad (5)$$

that quantify the properties of the future products that ensure their operability. The procedure of substantiation of parameters (indicators) T_1 comes down to the confirmation of the fact that all parameters and indicators X (4) (area of conditions E) are within the specified tolerances ΔX (5) (area of operability G):

$$\mathrm{E}\subset\mathrm{G},$$

here $G=\{X_i(t)|X_{\min(i)}\leq X_i(t)\leq X_{\max(i)}\}$.

Thus, the parametric modelling is a key component of design. On the one hand, it allows optimizing the designs and avoiding fundamental design errors, and on the other hand the generation of column vectors of parameters (indicators) (4) can support the generation of the check list for criterial supervision of the requirement and sufficiency of the requirements stipulated in DD (implementation of procedure T_2). In this case all calculations as part of substantiation of parameters (procedure T_1) that are associated with design and development process are implemented in order to confirm the DD requirements, with the total number of such calculations being defined by the list of the requirements.

Basic dependability improvement tools

The graph in the figure allows visualizing the tools used in design and development activities.

At the LC stage 0-t_1 (graph section 0-*A*) the primary operational requirements of the future product are specified. The achievement of the characteristics is defined by the most general strategic principles as fundamental truths that allow generating design solutions for the development of future products. Rational operating principles of the future product are the guarantee of its dependability. The number of the principles is not large. They aim to solve target tasks and reflect the general rules for achieving that.

At the LC stage t_1-t_2 (graph section *A-B*) with the integration of project activities the chosen principles of product creation must be implemented in the form of design and development solutions by means of the design rules [10]. The rules are based on the principles and defined by them. The number of rules may be considerable, they are aimed at solving specific tasks and reflect certain trends in the causal relationships.

The principles reflect the nature of a phenomenon, while the rules pertain to its individual aspects. Design principles and rules are universal for a certain type of products, therefore they can be reduced to a standardized set of design rules.

Dependability is ensured by fulfilling the requirements as a realized need to comply with the conditions that must be observed. Such requirements must be mandatorily specified at the LC stage t_2-t_3 (graph section *B-C*) and explicitly set forth in DD. The number of requirements is always larger than that of the used principles and rules as they are individual for each product in development and are used for elaboration of the adopted design solutions.

The principles and rules, if duly formalized, can be used for making check lists used in product design. While specifying DD requirements it is required to use digitization of the structure that will later be the basis for the preparation of the check list that in turn will support criterial evaluation of the completeness of the specified requirements during the development stage.

Principles of dependability evaluation

As dependability is a property, its measure is a qualitative characteristic. A requirement in DD is the expectation that a product, after its manufacture, will achieve such properties that unconditionally ensure the performance of the required functions under given conditions and modes of operation. The desired properties can always be identified by means of system analysis. If during the FA all modes and conditions of operation are identified, and the WCA identifies the worst combination of relative positions, mutual actions and interactions of critical elements, then during the design and production engineering dependability analysis there is always the chance of identifying such properties of critical elements that are required for achieving the set goals. That is possible due the antithesis method [10] when in given modes and conditions of operation under the worst possible combination of factors the causes of failures are identified, while the desired property is identified as the result of construction of logical formulas of type *“in order to eliminate the cause of failure in the form of .../ it is required that (a) critical element has a property of...”*. Next, each property is expressed quantitatively in the form of parameters (indicators) and their allowed values. Each of such properties is characterized with the probability of events that consist in finding the associated parameter (indicator) within the specified margins. In case of sequential occurrence of *i* events the overall evaluation of the product’s PNF *P*(*t*) as the result of the procedure T_1 equals to:

$$P(t)=\prod_{i=1}^{n} P_i(t), \tag{6}$$

here

$$P_i(t)=P\left[X_{\min(i)} \le X_i(\tau) \le X_{\max(i)};\ 0\le \tau \le t\right].$$

Formula (6) without regard to procedures T_2-T_4 yields an exaggerated result, because non-fulfilment or improper fulfilment of any of the design or process engineering procedures reduces the dependability. Therefore the DA is performed based on the results of the procedures T_1-T_4 and the current status of DD and MP. In case of non-fulfilment or improper fulfilment of any of the dependability procedures per any parameter or indicator, the formula (6) must include decreasing adjusting coefficients k_i that are defined, for instance, using the method of failure criticality rating as per GOST 27.310. As the result:

$$P(t) = \prod_{i=1}^{n} k_i P_i(t).$$

Conclusion

The paper shows the possibility of ensuring the dependability of UHVSs using engineering procedures implemented at each lifecycle stage before the beginning of operation. Such procedures would enable an adequate level of design, development, preproduction, manufacture, as well as the development of a UHVS dependability evaluation method based on a single theoretical and methodological basis.

References

1. Bolotin VV. Teoria nadiozhnosti mekhanicheskikh sistem s konechnym chislom stepeney svobody [Dependability theory of mechanical systems with a finite number of degrees of freedom]. Izvestia AN SSSR. Mekhanika tviordogo tela 1969;5:31-35 [in Russian].

2. GOST R 56526-2015. Reliability and safety requirements for space systems, complexes and unmanned spacecrafts of unique (small series) production with long life of active operation. Moscow: Standartinform; 2013.

3. Barlow R, Proschan F. Statistical theory of reliability and life testing. Moscow: Nauka; 1984.

4. Timoshenkov SP, Simonov BM, Goroshko VN. Osnovy teorii nadiozhnosti [Foundations of the dependability theory]. Moscow: Yurait; 2015.

5. Ushakov IA. Nadiozhnost – moi kompas zemnoi, a udacha nagrada za smelost. Human factors in reliability ili Neformalnaya istoriya teorii nadiozhnosti [Dependability is my compass on Earth, while fortune is the reward for bravery. Human factors in relaibility or the Informal history of the dependability theory], <http://gnedenko-forum.org/history.htm>; 2003 [accessed 31.08.2016] [in Russian].

6. Polovko AM, Gurov SV. Osnovy teoruii nadiozhnosti [Introduction into the dependability theory]. Saint-Petersburg: BHV-Petersburg; 2006.

7. GOST 27.002-89. Industrial product dependability. Basic concepts. Terms and definitions. Moscow: Izdatelstvo standartov; 1990.

8. Pokhabov YuP. About the philosophical aspect of reliability exemplified by unique mission critical systems. Dependability 2015;3:16-27.

9. Pokhabov YuP. Approach to ensuring of dependability of unique safety critical systems exemplified by large flexible structures. Dependability 2016;1:24-36.

10. Pokhabov YuP, Valishevsky OK. Genesis of dependability of unique safety critical systems. Dependability 2016;3:47-53.

11. Bart TV. Upravlenie kachestvom [Quality management]. Moscow: Izdatelstvo MIEMP; 2010 [in Russian].

12. Bushuev VV. Praktika konstruirovania mashin: spravochnik [Practice of machine design: Reference book]. Moscow: Mashinostroenie; 2006 [in Russian].

13. GOST 22487-77. Automated designing. Terms and definitions. Moscow: Izdatelstvo standartov; 1978 [in Russian].

14. R 50.1.031-2001 Continuous acquisition and life-cycle support. Glossary. Part 1. Product life-cycle stages. Moscow: Izdatelstvo standartov; 2001 [in Russian].

15. GOST ISO 9000-2011. Quality management systems. Fundamentals and vocabulary. Moscow: Standartinform; 2012 [in Russian].

16. STO 154-238-2014. Spacecraft design and development management using the requirements of foreign standards. Zheleznogorsk: AO ISS; 2014 [in Russian].

17. Design for Reliability. Crowe D, Feinberg A, editors. New York: CRC Press; 2001.

18. Rucker W, Hille F, Rohrmann R. SAMCO Final Technical Report: F08a Guideline for the assessment of existing structures. Berlin: Federal Institute of Materials Research and Testing (BAM); 2006.

About the author

Yuri P. Pokhabov, Candidate of Engineering, Joint Stock Company NPO PM – Maloe konstruktorskoye buro, Head of Research and Development Center, phone: +7 (913) 593 43 89, Russia, Krasnoyarsk Krai, Zheleznogorsk, e-mail: pokhabov_yury@mail.ru

Received on 31.08.2016

Original article
DOI: 10.21683/1729-2646-2017-17-3-24-31

Use of deduced Esary-Proschan assessments for evaluation of system dependability

Alexander G. Labutin, *Moscow Technical University of Communication and Informatics, Russia, Moscow*
Boris P. Filin, *Russia, Moscow*

> *…When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind…*
>
> *William Thomson (Lord Kelvin)*

Abstract. *In [1-2] it is shown that the widely known Esary-Proschan assessments [3-6] (EPA) are NP-complete [7]. In the process of their calculation a mutual cross-over of those assessments occurs despite the fact that the procedure of enumeration of complete sets of simple chains (SChs) and simple cuts (SCus) is performed all the way. This is confirmed by special research of these paradoxical phenomena in EPA conducted in [8] that concludes that EPAs are not assessments, as assessments cannot be NP-complete. In [7] it is clearly stated that in general an enumeration of a complete set of SCh (or SCu) alone already is an NP-complete problem. It implies directly that any NP-complete method cannot be an assessment one. In [9-10] a number of problems are classified depending on the associated computational complexity. As we can see out of those presented the most favourable is the intellectual intensity, as it allows controlling the computational process in the most desirable way, i.e. allows implementing the forced interruption principle (FIP) in regards to the computational procedure that is assessed by a certain parameter. For example, the parameter of achieved relative computational error. It should be noted that the devices, mechanisms and other systems we deal with in real life are called automated because such man-machine systems implement the FIP at the discretion of the human operator. We deal much less with automatic systems.* **The aim** *of this paper is to set forth the formal rules that allows quite easily the conventional NP-complete Esary-Proschan assessments to be transformed to the class of intelligent (IN-class) assessment methods that implement the FIP. Complete sets of SCh and SCu do not need to be enumerated here. Expanding the class of existing [1-6, 8, 11-29] methods that in one way or another implement the FIP is without a doubt a relevant problem for experts involved in structural dependability analysis of complex systems. It is an axiom that any of the tools of such system analysis, of which the exhaustive events (EE) are the "delivery nurse", contributes to the design of structurally dependent systems, while developing at the same time the analysis tool system itself. Essentially, the problem consists in casting the classic EPAs in the form of logic symbol multiplication (LSM) of logical operands the method uses.* **The result** *consists in the fact that we remove the "hardships" of NP-completeness from the classic EPAs and obtain a sufficiently efficient analysis tool.*

Keywords: *dependability, probability, two-pole network, simple chain, simple cut, working condition, faulty condition.*

Abbreviations: *EPA – Esary-Proschan assessment; SCh – simple chain; SCu – simple cut; FIP – forced interruption principle; EE – exhaustive events; LSM – logic symbol multiplication; RG – random graph; BCN – boundary couple nodes; TPN – two-pole network; WC – working condition; FC – faulty condition; PC – probability of connectedness; LAM – logic algebraic multiplication; cEEc – completion EE conjunction.*

For citation: *Labutin AG, Filin BP. Use of deduced Esary-Proschan assessments for evaluation of system dependability. Dependability 2017;3: p. 24-31. DOI: 10.21683/1729-2646-2017-17-3-24-31.*

1. Introduction

In [9–10] a number of problems are classified depending on the associated computational complexity: a) linear; b) polynomial; c) exponential; d) N-factorial; e) intellectual intensity (*IN*). Among experts involved in structural dependability analysis of complex systems the last one has the following definition: intellectual intensity is the principle of forced interruption of the computation process by a human being, when he is satisfied with the achieved error of function estimation by the given period time. As EPA contains logic multiplication of logical operands, describing the state of SCh and SCu excluding their structural interdependency, this article includes LSM developed rules of such operands taking into account this dependency. As a result, classic EPAs in computational complexity transform into N-class, as they become forcibly interruptible.

The article proposes formal rules that allow the conventional NP-complete Esary-Proschan assessments to be transformed to the convenient tool, excluding enumeration of complete sets of SCh and SCu. Let us name such assessments as reduced EPA.

2. Problem definition

2.1. Generalities. Let it be a simple edge random graph (RG) (Fig. 1) [11–14, 30–31] G: simple excludes loops, parallels and isolated nodes; random means that its elements are either present in the graph with probability of p or absent with probability of q = 1 – p (as $p + q = 1$, then presence and absence of the edge in the graph compose EE denoted as symbol I [31]); edge means that only edges are unreliable in the graph (this assumption is not essential, but simplifies the paradigm of the statement). Graph G contains a node set $V=\{v_i\}$ with the power $m_v=|V|$ and edges $L=\{l_{i,j}\}$ with the power $m_L|L|$, where the function of incidence and type adjacency $\Phi(l_{i,j})=v_i\&v_j$ reflects their interdependence: an edge is incident to its boundary couple nodes (BCN) where BCN are adjacent to each other by the edge $l_{i,j}$.

Let us assume two-pole network (TPN) be set on the graph G, whose pole nodes are denoted as s (source) and t (drain) with the parameters: m_V=4 and m_L=5. For example, as it is shown in Figure 1 (experts often define such structure as *«bridge»* [16]). It should be noted that currently TPN has transit nodes in the following form: $V_{s,t}^T = V \setminus \{s,t\}$. In this situation: $V_{s,t}^T = \{2,3\}$ (Fig. 1).

Figure 1b shows a graph with numbers renumbered by edges (using the method of «arc crossing» [16]), ranging from (m_V+1) with step 1 to (m_V+m_L), and Figure 1c presents the same renumbered RG, but here symbol «*l*» on edges is omitted for the purity of the picture. We will often resort to Figure 1. Thus, «continuous» numbering of RG elements is presented in Figure 1c. It should be noted that symbols «*v*» and «*l*» will be used rarely when creating any structures on RG (for example, SCh and SCu).

Let us form in TPN (Fig. 1c) a set of its SCh according to the principle: **«Every *parent*-node chooses a *child*-node from nodes adjacent with it and not occupied, that has the lowest number** k [17]»:

$$F_{s,t} = \left\{ \begin{matrix} f_1 = \{5,7,9\}, f_2 = \{5,8\}, \\ f_3 = \{6,7,8\}, f_4 = \{6,9\} \end{matrix} \right\} \Rightarrow \left(m_F = \left| F_{s,t} \right| = 4 \right). \quad (1)$$

own in Figure 1, every transit node was both «*parent*» and «*child*», but s (source) was only «*parent*» and t (drain) – only «*child*».

With the same principle let us form SCu set (method of «continuous start» [17]):

$$R_{s,t} = \left\{ \begin{matrix} r_1 = \{5,6\}, r_2 = \{6,7,8\}, \\ r_3 = \{5,7,9\}, r_4 = \{8,9\} \end{matrix} \right\} \Rightarrow \left(m_R = \left| R_{s,t} \right| = 4 \right). \quad (2)$$

SCh is in good order when all its edges are in working condition (WC): $\overline{\overline{f}}_n = \forall l_k \in f_n \left[\overline{\overline{l}}_k \right]$, and is in fault order when even one of its edges is in faulty condition (FC): $\overline{f}_n = \exists l_k \in f_n \left[\overline{l}_k \right]$ (here $\forall$ is an *«generality»* quantifier, $\exists$ is an *«existential»* quantifier, and $\in$ means *«belongs to»*

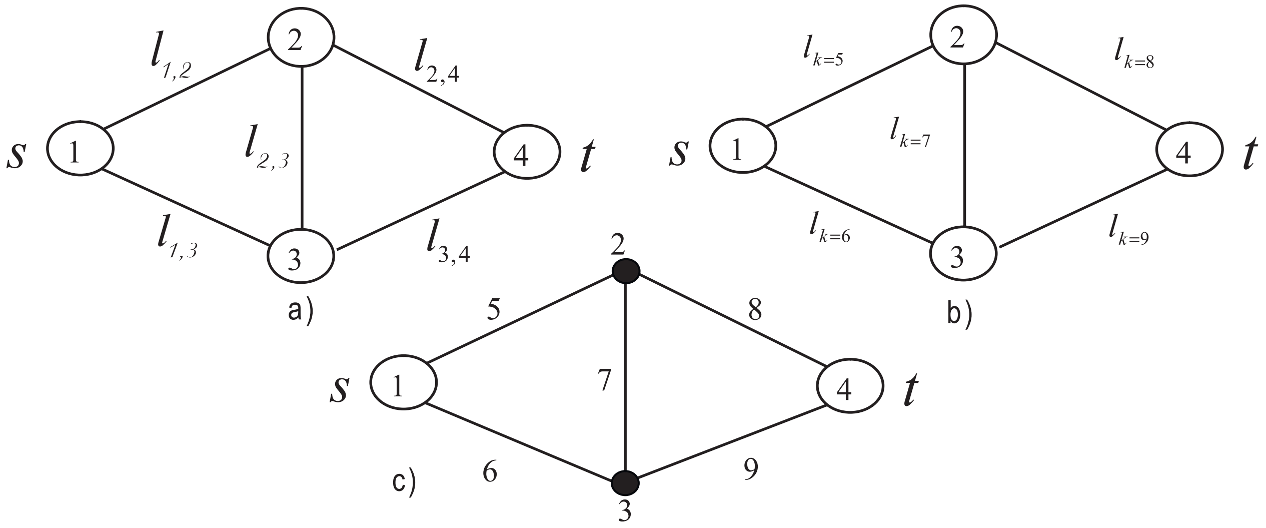


Figure 1. Renumbered simple random edge graph

[30–33]). SCu is in good order when all its edges are in FC: $\bar{\bar{r}}_n = \forall l_k \in r_n \left[\bar{l}_k \right]$, and is in fault order when even one of its edges is in WC: $\bar{r}_n = \exists l_k \in r_n \left[\bar{\bar{l}}_k \right]$.

2.2 Explaining example. Let us take initial data on AF of RG TPN edges:

$$(L \neq \varnothing) \Rightarrow (p_5 = 0,5,\ p_6 = 0,6,\ p_7 = 0,7,\ p_8 = 0,8,\ p_9 = 0,9), \quad (3)$$

where ⇒ is the «*sequence*» symbol [31–33].

Let us calculate an exact value of probability of connectedness (PC) of our TPN. Firstly, based on (1), let us describe complete event of connectedness (CoC) of TPN (when TPN nodes-poles are connected with even one working SCh), denoted as symbol $E_{s,t}$:

$$\begin{aligned} E_{s,t} &= 5 \bullet 7 \bullet 9 + 5 \bullet 8 \bullet \overline{7 \bullet 9} + \\ &+ 6 \bullet 7 \bullet 8 \bullet \bar{5} + 6 \bullet 9 \bullet \left(\bar{5} \bullet \overline{7 \bullet 8} + 5 \bullet \bar{7} \bullet \bar{8} \right), \end{aligned} \quad (4)$$

where • is the symbol of logic algebraic multiplication (LAM) (* is the symbol of LSM, in which in contrast to LAM, structural dependence of the multiplied logical operands is taken into account).

It should be noted that here (and hereafter) we don't «load» the edge WC with double bar over elements of operands as it shows that the edge is in faulty condition.

Double negation is the edge WC (bar cancels bar), i.e. «faulty» and «working» are synonyms. Let us write calculation formula for $P_{s,t}$, based on the result (4):

$$\begin{aligned} P_{s,t} &= p_5 \cdot p_7 \cdot p_9 + p_5 \cdot p_8 \cdot \overline{p_7 \cdot p_9} + \\ &+ p_6 \cdot p_7 \cdot p_8 \cdot q_5 + p_6 \cdot p_9 \cdot \left(q_5 \cdot \overline{p_7 \cdot p_8} + p_5 \cdot q_7 \cdot q_8 \right). \end{aligned} \quad (5)$$

Let us expand in (5) our initial data (3) and deduce that the exact (within the accuracy of initial data (3)) PC value of our TPN will be equal to:

$$P_{s,t} = 0,315 + 0,148 + 0,168 + 0,136 = \underline{\underline{0,766}}. \quad (6)$$

Now we have some «criterion» in the form of $P_{s,t}$=0,766, that allows us to make a comparison between «old» and «new», i.e. offered.

2.3. Task definition. Firstly, let us write according to [3] the formal representation of EPA for general cases:

$$\left. \begin{aligned} E^B_{s,t} &= I - \prod_{n=1}^{m_F} \bullet \bar{f}_n = I - \prod_{n=1}^{m_F} \bullet \left(I - \prod_{e=1}^{m_n} \bullet \bar{\bar{l}}_{k_e} \right) = \\ &\quad I - \prod_{n=1}^{m_F} \bullet \left(\coprod_{e=1}^{m_n} \bullet \bar{\bar{l}}_{k_e} \right) = \coprod_{n=1}^{m_F} \bullet \left(\coprod_{e=1}^{m_n} \bullet \bar{\bar{l}}_{k_e} \right) \\ E^H_{s,t} &= \prod_{n=1}^{m_R} \bullet \bar{r}_n = \prod_{n=1}^{m_R} \bullet \left(I - \prod_{e=1}^{m_n} \bullet \bar{l}_{k_e} \right) = \prod_{n=1}^{m_R} \bullet \left(\coprod_{e=1}^{m_n} \bullet \bar{l}_{k_e} \right) \end{aligned} \right\}. \quad (7)$$

where $\prod$ is the conjunction (logical (or arithmetical) products); $\coprod$ is the complement of EE (or *I*) conjunction (logical (or arithmetical)) (cEEc).

If we exclude the intermediate transformation, then EPA should be written as follows:

$$\left. \begin{aligned} E^B_{s,t} &= I - \prod_{n=1}^{m_F} \bullet \bar{f}_n = \coprod_{n=1}^{m_F} \bullet \left(\coprod_{e=1}^{m_n} \bullet \bar{\bar{l}}_{k_e} \right) \\ E^H_{s,t} &= \prod_{n=1}^{m_R} \bullet \bar{r}_n = \prod_{n=1}^{m_R} \bullet \left(\coprod_{e=1}^{m_n} \bullet \bar{l}_{k_e} \right) \end{aligned} \right\}. \quad (8)$$

In relation to our RG example (Fig. 1), these events (according to (1), (2) and (8)) can be graphically represented as shown in Fig. 2.

The redundant TPN CoC on SCh is read as: «The complement of EE conjunction of the complement of EE conjunction of WC edges included in *n*-th SCh». At the same time, insufficient description of TPN CoC on SCu reads differently: «The conjunction of complement of EE conjunctions of FC edges included in *n*-th SCu».

If TSE PC TPN have been calculated, then the calculation of approximate estimate of $P^{\approx}_{s,t}$, relative error of $\Delta_{s,t}$ (a priori) and absolute error of $W_{s,t}$ (a posteriori) is simple:

$$P^{\approx}_{s,t} = \frac{P^B_{s,t} + P^H_{s,t}}{2},\ \Delta_{s,t} = \frac{\left| P^B_{s,t} - P^H_{s,t} \right|}{2},$$
$$W_{s,t} = \left| P_{s,t} - P^{\approx}_{s,t} \right| \text{ and } W_{s,t} < \Delta_{s,t}. \quad (9)$$

1: It is required to prove (the first in (8)), that eliminating of structural interdependency of logical operands of conjunction leads to the redundancy in description of TPN CoC.

2. It is required to prove (the second in (8)), that eliminating of structural interdependency of logical operands of cEEc leads to insufficiency in description of TPN CoC.

3. The proved statements should be illustrated by graphical and numerical examples.

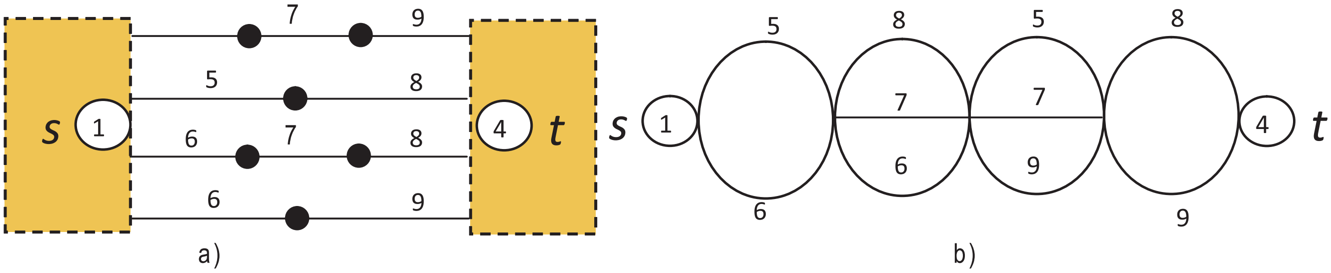


Figure 2. “Bridge” in terms of EPA elemental events

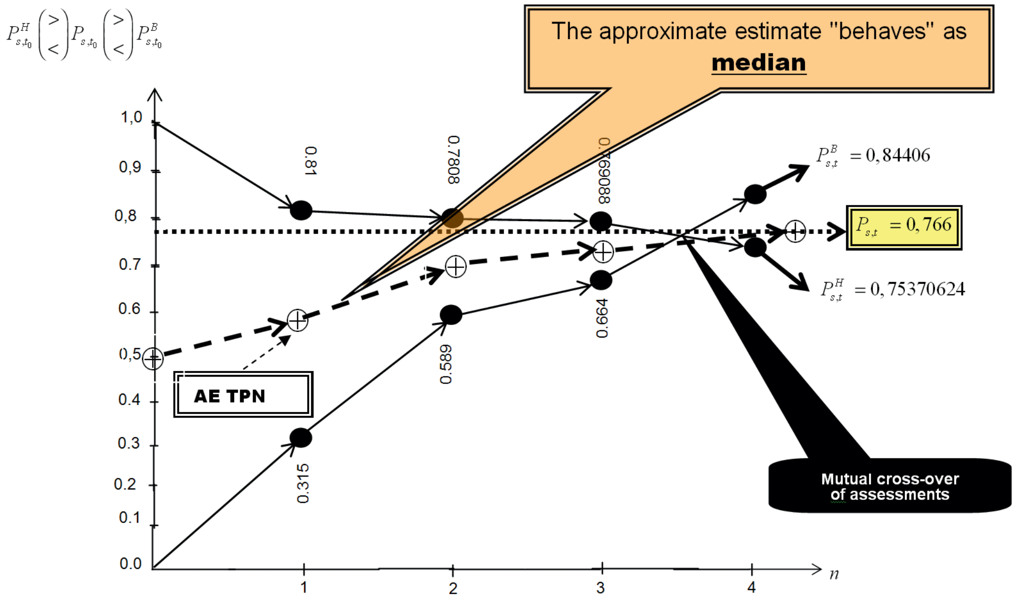


Figure 3. Graph of the dual image of the behavior of the PC TSE for the "bridge" under the adopted input data $\forall l_{(5\le\nu\le9)}\in L\left[p_{(5\le\nu\le9)}=0,(5\le\nu\le9)\right]$

3. Task solution

In accordance with the rules (7-8) and sets of SCh (1) and SCu (2) sets EPA in relation to the «bridge» (Fig. 1 and Fig. 2) looks the following:

$$\left.\begin{aligned}E^{U}_{s,t}&=\overline{\overline{5\bullet7\bullet9}\bullet\overline{5\bullet8}\bullet\overline{6\bullet7\bullet8}\bullet\overline{6\bullet9}}\\E^{L}_{s,t}&=\overline{\bar5\bullet\bar6}\bullet\overline{\bar6\bullet\bar7\bullet\bar8}\bullet\overline{\bar5\bullet\bar7\bullet\bar9}\bullet\overline{\bar8\bullet\bar9}\end{aligned}\right\}. \tag{10}$$

Based on (10), we deduce the following calculations formulas for TSE PC TPN (for UB and LB PC TPN):

$$\left.\begin{aligned}P^{U}_{s,t}&=\overline{\overline{p_5\cdot p_7\cdot p_9}\cdot\overline{p_5\cdot p_8}\cdot\overline{p_6\cdot p_7\cdot p_8}\cdot\overline{p_6\cdot p_9}}\\P^{L}_{s,t}&=\overline{q_5\cdot q_6}\cdot\overline{q_6\cdot q_7\cdot q_8}\cdot\overline{q_5\cdot q_7\cdot q_9}\cdot\overline{q_8\cdot q_9}\end{aligned}\right\}. \tag{11}$$

As in this case we have step-by-step «accumulation» of some numerical value as a result of multiplication (not additivity) then this *«step-by-step principle»* will be represented by arrows. Let us use our initial data (3) and calculate TSE PC TPN on EPA: UB $P^{U}_{s,t}$, and LB $P^{L}_{s,t}$ bear in mind (6), that $P_{s,t}$=0,766:

$$\left.\begin{aligned}P^{U}_{s,t}&=0,315\to0,589\to0,664\to0,84406=P^{U}_{s,t}\\P^{L}_{s,t}&=0,81\to0,7808\to769088\to0,75370624=P^{L}_{s,t}\end{aligned}\right\}. \tag{12}$$

Let us use the results and build the graph of dual image [17, 20] of the «behavior» of the TSE PC of our TPN (Fig. 1c, 2 and 3).

Analyzing the results we can see the paradox of classic EPA: a) UB PC TPN begins to accumulate from 0, and LB – from 1; b) TSE PC TPN are mutual intercrossing; c) after intercrossing TSE PC TPN diverge.

The axis of probabilities and its domains and intervals presented in Figure 4 should help us in the following questions.

Logically arguing, it is possible to understand that conjunctions and their additions (7) and (8) are nonequilibrium in different cases. For example, the arithmetical product of $\prod$ in initial state are set to entity (as the sum is set to «zero» in initial state):

$$\left.\begin{aligned}n&=\overline{1,N}\\n&:=0\end{aligned}\right\}\Rightarrow\prod:=1. \tag{13}$$

The conjunction in initial state also should be reduced to the form of «full possible event group»:

$$\left.\begin{aligned}n&=\overline{1,N}\\n&:=0\end{aligned}\right\}\Rightarrow\prod:=I. \tag{14}$$

Then cEEc in initial state should be reduced to the form of «impossible event»:

$$\left.\begin{aligned}n&=\overline{1,N}\\n&:=0\end{aligned}\right\}\Rightarrow\coprod=I-\prod=I-I=\otimes, \tag{15}$$

where ⊗ is the symbol of impossible event.

Let us set the complement of an arithmetic product to 1, as the numerical range varies from «zero» to «one»:

$$\left.\begin{aligned}n&=\overline{1,N}\\n&:=0\end{aligned}\right\}\Rightarrow\coprod=1-\prod=1-1=0. \tag{16}$$

We proceed from the fact that $I*I=I$, $I+I=I$, $\otimes*\otimes=\otimes$, $\otimes+\otimes=\otimes$, but $I+\otimes=\otimes$, $I+\otimes=I$.

Let us clarify only two abbreviations in Figure 4: ZPE – zero-probability events; OPE – one-probability events [9–10]. Considering that in theory of combinational dependability all our calculations are based on numerical values in the interval from 0 to 1 (including these boundaries), let us formulate some statements and theorems.

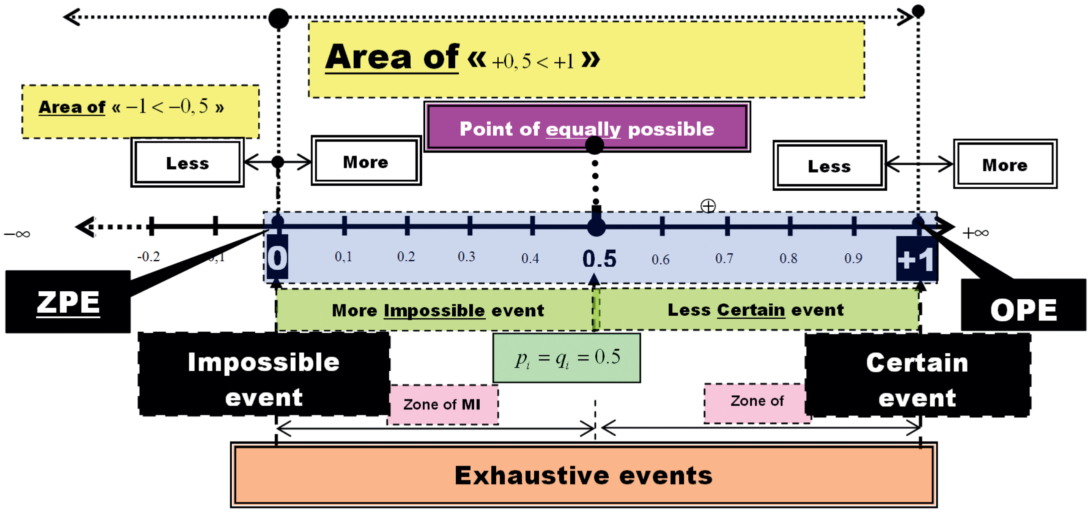


Figure 4. Event probability axis

Statement 1 (Fig. 4). With the increasing number of cofactors in the form of probabilistic numerical values the value of their product falls dramatically and converges to ZPE domain.

Statement 2 (Fig. 3). With the increasing number of logical cofactors, the veracity of their logical chain decreases.

In [16] it is written: «**When you remove the brackets remember the rule:** *p·p=p*». Let us name this rule by the author's name of this article: «Bogatyrev's rule».

Theorem 1. The connectedness event of TPN described by cEEc is redundant; every of TPN is cEEc where WC of RG edges are its cofactors, i.e.:

$$\left.\begin{matrix}f_1=\{a,b\}\\ f_2=\{a,c\}\end{matrix}\right\}\Rightarrow\left[\left(\overline{\overline{a\bullet b}\bullet\overline{a\bullet c}}\right)>\left(\overline{\overline{a\bullet b}*\overline{a\bullet c}}=a\bullet\overline{\overline{b}\bullet\overline{c}}\right)\right]. \quad (17)$$

The proof:

$$\left(\left(\overline{\overline{a\cdot b}\cdot\overline{a\cdot c}}\right)\Leftrightarrow\left(a\cdot\overline{\overline{b}\cdot\overline{c}}\cdot\right)\right)\Rightarrow$$

$$\Rightarrow\left(\left(I-(I-a\cdot b)\cdot(I-a\cdot c)\right)\Leftrightarrow\left(a\cdot(I-(I-b)\cdot(I-c))\right)\right)\Rightarrow$$

$$\Rightarrow\begin{pmatrix}I-\left(I-a\cdot b-a\cdot c+a^2\cdot b\cdot c\right))\Leftrightarrow\\ \left(a\cdot\left(I-\left(I-b-c+b\cdot c\right)\right)\right)\end{pmatrix}\Rightarrow$$

$$\Rightarrow\begin{pmatrix}\left(I-I+a\cdot b+a\cdot c-a^2\cdot b\cdot c\right)\Leftrightarrow\\ \left(a\cdot\left(I-I+b+c-b\cdot c\right)\right)\end{pmatrix}\Rightarrow$$

$$\Rightarrow\left(\left(a\cdot b+a\cdot c-a^2\cdot b\cdot c\right)\Leftrightarrow\left(a\cdot b+a\cdot c-a\cdot b\cdot c\right)\right)\Rightarrow$$

$$\left(\underline{\underline{(-a)\Leftrightarrow(-I)}}\right).$$

Turning to the stochastic side of «endspiel» and remembering that ($0<p_{(a,b,c)}<1$) it is easy to see that the comparison result is performed as follows ($(-p_a)>-1$), q.e.d. (⇔ is the symbol «compare»).

Corollary 1.1. As in EPA only LAM is used then according to (17) in EPA (the first in (8)) only the redundant TPN CoC is always described that leads to the crossing of UB from the *bottom* of the exact value of PC TPN thereby this UB is the false estimate.

Corollary 1.2. To eliminate this lack, it is necessary to use LSM rules instead of LAM when describing UB PC TPN in EPA [17, 20]. Then the redundant description of TPN CoC is impossible and the false UB PC TPN on EPA becomes true, i.e. in lower LB PC TPN. The LSM rules for SCh in FC (the first in (8) and [17, 20]) are the following:

$$\left.\begin{matrix}\overline{a}*\overline{a\bullet b}=\overline{a}\\ \overline{a\bullet b}*\overline{a\bullet c}=\overline{a\bullet\overline{\overline{b}\bullet\overline{c}}}\\ \overline{a\bullet\overline{\overline{b}\bullet\overline{c}}}*\overline{b\bullet d}=\overline{a}\bullet\overline{b\bullet d}+a\bullet\overline{b}\bullet\overline{c}\\ \overline{a}*\overline{b}=\overline{a}\bullet\overline{b}\end{matrix}\right\}. \quad (18)$$

Theorem 2. The connectedness of TPN described by conjunction of cEEc where WC of RG edges is its cofactors is redundant, i.e.:

$$\left.\begin{matrix}r_1=\{a,b\}\\ r_2=\{a,c\}\end{matrix}\right\}\Rightarrow\left[\left(\overline{\overline{a}\bullet\overline{b}}\bullet\overline{\overline{a}\bullet\overline{c}}\right)>\left(\overline{\overline{a}\bullet\overline{b}}*\overline{\overline{a}\bullet\overline{c}}=\overline{\overline{a}\bullet\overline{b\bullet c}}\right)\right]. \quad (19)$$

The proof: Let us also transfer the comparison to «endspiel»:

$$\left(\left(\overline{\overline{a}\bullet\overline{b}}\bullet\overline{\overline{a}\bullet\overline{c}}\right)\Leftrightarrow\left(\overline{\overline{a}\bullet\overline{b\bullet c}}\right)\right)\Rightarrow\left(\left(\left(I-\overline{a}\bullet\overline{b}\right)\bullet\left(I-\overline{a}\bullet\overline{c}\right)\right)\Leftrightarrow\right.$$

$$\Leftrightarrow\left[(I-\overline{a}\bullet(I-b\bullet c)))=(I-\overline{a}\bullet(I-(I-\overline{b})\bullet(I-\overline{c})))=\right.$$

$$=\left(I-\overline{a}\bullet\left(I-I+\overline{b}+\overline{c}-\overline{b}\bullet\overline{c}\right)\right)=\left(I-\overline{a}\bullet\left(\overline{b}+\overline{c}-\overline{b}\bullet\overline{c}\right)\right)=$$

$$\left.=\left(I-a\bullet b-a\bullet c+\overline{a}\bullet\overline{b}\bullet\overline{c}\right)\right]\Rightarrow$$

$$\Rightarrow\begin{bmatrix}\left(I-\overline{a}\bullet\overline{b}-\overline{a}\bullet\overline{c}+\overline{a}^{2}\bullet\overline{b}\bullet\overline{c}\right)\Leftrightarrow\\ \left(I-\overline{a}\bullet\overline{b}-\overline{a}\bullet\overline{c}+\overline{a}\bullet\overline{b}\bullet\overline{c}\right)\end{bmatrix}\Rightarrow\left(\overline{a}\Leftrightarrow\otimes\right).$$

Turning to the stochastic side of «endspiel» and remembering that ($0<p_{(a,b,c)}<1$), the inequality $q_a>0$ is always true. Therefore, the initial (19) is true, q.e.d.

Corollary 2.1. As in EPA only LAM is used then according to (19) in EPA (the second in (8)) only the redundant TPN CoC is always described that leads to the crossing of the LB from *above* of the exact value of PC TPN thereby this LB is the false estimate (note: we shouldn't forget that here we use GDI [17, 20]).

Corollary 2.2. To eliminate this lack, it is necessary to use LSM rules instead of LAM when describing LB PC TPN in EPA [17, 20]. Then the redundant description of CoC TPN is impossible and the *false* LB PC TPN on EPA becomes true and transforms into upper UB PC TPN. The LSM rules for SCu in FC (the second in (8) and [17, 20]) are the following:

$$\left.\begin{array}{c} a*\overline{\bar{a}\bullet\bar{b}}=a \\ \overline{\bar{a}\bullet\bar{b}}*\overline{\bar{a}\bullet\bar{c}}=\overline{\bar{a}\bullet\overline{b\bullet c}} \\ \overline{\bar{a}\bullet\overline{b\bullet c}}*\overline{\bar{b}\bullet\bar{d}}=a\bullet\overline{\bar{b}\bullet\bar{d}}+\bar{a}\bullet b\bullet c \\ a*b=a\bullet b \end{array}\right\}. \tag{20}$$

Graphic interpretation of the theorems 1 and 2 is presented in Figure 5.

Thus, the LSM rules and OsLF will be as follows:

Lower bound (based on SCh):

$$\left.\begin{array}{c} \bar{a}*\overline{a\bullet b}=\bar{a} \\ \overline{a\bullet b}*\overline{a\bullet c}=\overline{a\bullet\overline{\bar{b}\bullet\bar{c}}} \\ \overline{a\bullet\overline{\bar{b}\bullet\bar{c}}}*\overline{b\bullet d}=\bar{a}\bullet\overline{b\bullet d}+a\bullet\bar{b}\bullet\bar{c} \\ \bar{a}*\bar{b}=\bar{a}\bullet\bar{b} \end{array}\right\}\Rightarrow$$

$$\Rightarrow\left(E^{L}_{s,t}=\coprod_{e=1}^{m_F}*\bar{f}_e\,\middle|\,(e=0)\Rightarrow\coprod_{e=0}^{m_F}*I\right)\le E_{s,t}, \tag{21}$$

where $|-$ «under conditions», and

Upper bound (based on SCu):

$$\left.\begin{array}{c} a*\overline{\bar{a}\bullet\bar{b}}=a \\ \overline{\bar{a}\bullet\bar{b}}*\overline{\bar{a}\bullet\bar{c}}=\overline{\bar{a}\bullet\overline{b\bullet c}} \\ \overline{\bar{a}\bullet\overline{b\bullet c}}*\overline{\bar{b}\bullet\bar{d}}=a\bullet\overline{\bar{b}\bullet\bar{d}}+\bar{a}\bullet b\bullet c \\ a*b=a\bullet b \end{array}\right\}\Rightarrow$$

$$\Rightarrow\left(E^{U}_{s,t}=\prod_{e=1}^{m_R}*\bar{r}_e\,\middle|\,(e=0)\Rightarrow\prod_{e=0}^{m_R}*I\right)\ge E_{s,t}. \tag{22}$$

This completes the description of the main results.

4. Example

Let us take as an example «the bridge» (Fig. 1c), whose SCh and SCu sets are represented accordingly in (1) and (2) and initial data on AF edges are in (3). Let us calculate LB PC TPN «behavior» based on OsLF formal rules (19):

$$E^{L}_{s,t}=\overline{I*\overline{\overline{5\bullet7\bullet9}}}=\underline{\underline{5\bullet7\bullet9}}_1\to\overline{\overline{5\bullet7\bullet9}*\overline{5\bullet8}}=$$

$$=\overline{\overline{5\bullet\overline{\overline{7\bullet9}\bullet\bar{8}}}}=\underline{\underline{5\bullet\overline{\overline{7\bullet9}\bullet\bar{8}}}}_2\to\overline{\overline{5\bullet\overline{\overline{7\bullet9}\bullet\bar{8}}}*\overline{6\bullet7\bullet8}}=\dots$$

$$\dots=\underline{\underline{\left(\overline{\bar{5}\bullet\overline{6\bullet7\bullet8}+5\bullet\overline{7\bullet9}\bullet\bar{8}}\right)}}_3\to$$

$$\to\overline{\left(\overline{\overline{\bar{5}\bullet\overline{6\bullet7\bullet8}+5\bullet\overline{7\bullet9}\bullet\bar{8}}}\right)*\overline{6\bullet9}}=\dots$$

$$\dots=\left(\underline{\underline{\overline{\bar{5}\bullet\overline{6\bullet\overline{\overline{7\bullet8}\bullet\bar{9}}}+5\bullet\overline{\overline{\bar{7}\bullet\bar{6}}\bullet9}\bullet\bar{8}}}}_4\right). \tag{23}$$

The results at the corresponding step after LSM operands is marked by double underline. If we exclude from (21) all intermediate logical operations then we obtain the following:

$$E^{L}_{s,t}=5\bullet7\bullet9_1\to5\bullet\overline{\overline{7\bullet9}\bullet\bar{8}}_2\to$$

$$\to\left(\overline{\bar{5}\bullet\overline{6\bullet7\bullet8}+5\bullet\overline{7\bullet9}\bullet\bar{8}}\right)_3\to$$

$$\to\left(\overline{\bar{5}\bullet\overline{6\bullet\overline{\overline{7\bullet8}\bullet\bar{9}}}+5\bullet\overline{\overline{\bar{7}\bullet\bar{6}}\bullet9}\bullet\bar{8}}\right)_4. \tag{24}$$

Let us use our initial data (3) and show the dynamic of «normal» behavior of the LB PC TPN obtained according to the stated in the article OsLF rules (19), which allowed estimates to stop their mutual crossing and transform from LB to UB PC TPN which was the distinctive feature of «classic» EPA. The increasing LB PC TPN are the following:

$$P^{L}_{s,t}=0{,}315_1\to0{,}463_2\to0{,}631_3\to\underline{\underline{0{,}766}}_4=P_{s,t}. \tag{25}$$

It could be seen that at the last 4th step LB PC TPN is equal to the exact value of PC TPN, estimated in (6).

Let us describe the dynamic of UB PC TPN «normal» behavior using OsLF method, according to the rules (22):

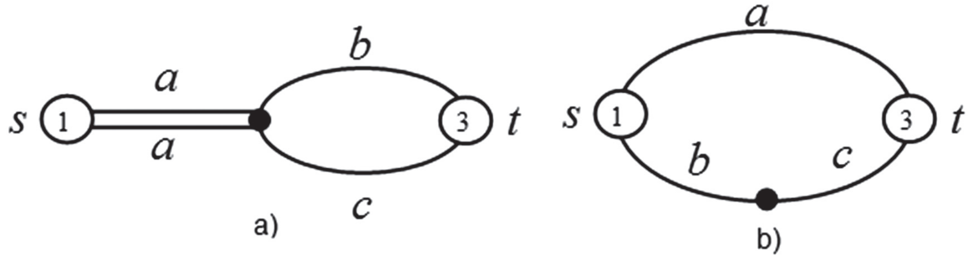


Figure 5. Graphic interpretation

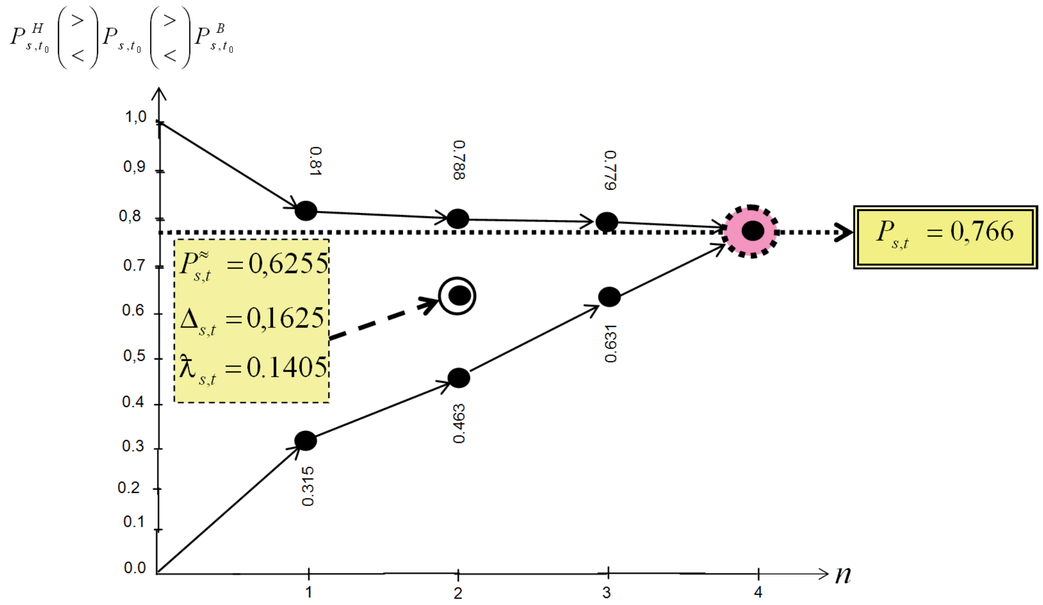


Figure 6. Graph of the dual image of the behavior of the PC TSE for the "bridge" under the adopted input data

$$\forall l_{(5\le \nu \le 9)} \in L\left[\, p_{(5\le \nu \le 9)} = 0, (5 \le \nu \le 9)\right]$$

$$E^{U}_{s,t} = I \bullet \overline{\overline{5}\bullet\overline{6}} = \underline{\underline{\overline{\overline{5}\bullet\overline{6}}}}_{1} \rightarrow \left(\overline{\overline{5}\bullet\overline{6}} * \overline{\overline{6}\bullet\overline{7}\bullet\overline{8}} = \underline{\underline{\overline{\overline{6}\bullet\overline{5\bullet\overline{\overline{7}\bullet\overline{8}}}}}}_{2}\right) \rightarrow$$

$$\rightarrow \left(\overline{\overline{6}\bullet\overline{5\bullet\overline{\overline{7}\bullet\overline{8}}}} * \overline{\overline{5}\bullet\overline{7}\bullet\overline{9}}\right) = \left(\underline{\underline{6\bullet\overline{\overline{5}\bullet\overline{7}\bullet\overline{9}}+}}\ldots\underline{\underline{+\overline{6}\bullet 5\bullet\overline{\overline{7}\bullet\overline{8}}}}\right)_{3} \rightarrow$$

$$\rightarrow \left(\begin{array}{c} \left(6\bullet\overline{\overline{5}\bullet\overline{7}\bullet\overline{9}} + +\overline{6}\bullet 5\bullet\overline{\overline{7}\bullet\overline{8}}\right) * \overline{\overline{8}\bullet\overline{9}} = \\ = \underline{\underline{6\bullet\overline{\overline{9}\bullet\overline{\overline{\overline{5}\bullet\overline{7}}}\bullet 8} + \overline{6}\bullet 5\bullet\overline{\overline{8}\bullet\overline{\overline{7}\bullet 9}}}} \end{array}\right)_{4}. \qquad (26)$$

Using our initial data (3) we can obtain numerical values of <u>falling</u> UB PC TPN equal to the PC TPN exact value:

$$P^{U}_{s,t} = 0{,}8 \rightarrow 0{,}788 \rightarrow 0{,}779 \rightarrow \underline{\underline{0{,}766}} = P_{s,t}. \qquad (27)$$

As a result, UB PC TPN became also equal to the exact PC TPN value calculated in (6).

Based on the results (25) and (27) it is possible to construct the graph of the dual image [17, 20] of LB PC TPN <u>increasing</u> and UB PC TPN <u>falling</u> dynamic for the exact PC TPN value. This graph is presented in Figure 6. Here it is also shown that FIP estimation procedure was realized at the 2^{d} step, the estimative parameters of the calculations were obtained under which the operator should decide whether he continuous calculations or not.

5. Conclusion

In summary, we think that we (co-authors) completely solve the task. Based on the results we can see that deduced EPA doesn't belong to *NP*-class. These new assessments based on human intelligence belong to *IN*-class. This approach is very popular among experts involved in structural dependability analysis of complex systems.

References

1. Krivulets VG. Ob otsenke otsenok Esary-Proschana v zadachakh analiza strukturnoy nadiozhnosti setey sviazi [On the assessment of Esary-Proschan assessments as part of analysis of structural dependability of communication networks]. Proceedings of the 55th scientific session dedicated to the Radio Day. RNTORES im. AS Popova; 2000 [in Russian].

2. Filin BP. O predelnom razviazyvanii klatterov v otsenkakh Polesskogo granits kombinatornoy nadiozhnosti sluchainykh binarnykh sistem [On the ultimate clutter disentanglement in Polessky assessment of combinatorial dependability of random binary systems]. Avtomatika i telemekhanika 2005; 9: 149-189 [in Russian].

3. Esary J, Proschan F. Coherent Structures of Non-Identifical Components. Technometrics 1963; 5 (2): 191-209.

4. Esary J, Proschan F. The reliability of coherent systems in Redundancy techniques for computing systems . Moscow: Radio i sviaz; 1966.

5.Barlow R, Proschan F. Mathematical theory of reliability. Gnedenko BV, editor. Moscow: Sov. radio; 1978.

6. Барлоу Р., Прошан Ф. Статистическая теория надёжности и испытания на безотказность: . . Gnedenko BV, editor. Moscow: Nauka, 1969.

7. Garey M, Johnson D. Computers and intractability. Moscow: Mir; 1982.

8. Labutin AG, Filin BP. Ob uklonenii ot *NP*-polnoty v otsenkakh Esary-Proschana [On avoiding NP-completeness in Esary-Proschan assessments]. Avtomatika i lelemekhanika 2017 [in print] [in Russian].

9. Gadasin VA. Triada substantsiy v mikromire "Korpuskula – Sluchainost – Volna" [Triade od substances in the microworld Corpuscule – Eventuality – Wave]. Proceedings of VNII PVTI; 2015 [in Russian].

10. Gadasin VA. Aksiomatika kontepsii triad – triokhmernaya gruppa [Axiomatics of the concept of triad – three-dimensional group]. Proceedings of the XV international conference The problem of safety of complex systems. Moscow: RAS ICS; 2007 [in Russian]. .

11. Polovko AM, Gurovich BI. Tekhnicheskaya kibernetika 1971;4:78 [in Russian].

12. Panteley VG, Shubinsy IB. Raschotnie metody otsenki nadiozhnosti priborov [Computational methods of devices dependability assessment]. Moscow: Mashinostroenie; 1974 [in Russian].

13. Shubinsky IB. Topologichskiy metod i algoritm opredelenia statsionarnykh pokazateley nadiozhnosti tekhnicheskikh sistem [Topological method and algorithm of identification of steady-state dependability indicators of technical systems]. Dependability and quality control 1984; 5:3-10.

14. Shubinsky IB. Strukturnaya nadiozhnost informatsionnykh sistem [Structural dependability of information systems]. Metody analiza [Analysis methods]. Moscow: Dependability; 2016 [in Russian].

15. Polessky VP. Razviazyvania klatterov, korreliatsionnie neravenstva i granitsy kombunatornoy nadiozhnosti [Clutter untyings, correlation inequalities, and bounds for combinatorial reliability]. Problems of information transmission 1997; 33 (3): 50-70 [in Russian].

16. Nosov MV. Metod polnogo razlozhenia mostikovykh soedineniy v zadachakh analiza sviaznosti strukturno-slozhnykh dvukhpoliusnykh setey [Method of complete decomposition of bridge connections in connectivity analysis problems of structurally complex bipolar networks]. Dependability 2015; 4: 68-74.

17. Filin BP. Metody analiza strukturnoy nadiozhnosti setey sviazi [Methods of structural dependability analysis of communication networks]. Moscow: Radio i sviaz, 1988 [in Russian].

18. Filin BP. Metod posledovatelnogo starta v opredelenii prostykh secheniy [Method of successive start in identifying simple crossections]. TsIVTI; 1977 [in Russian].

19. Bogatyrev VA. K raschiotu nadiozhnosti setey sviazi po sovokepnosti putey [On dependability calculation of communication networks per all route]. Elektrosviaz 1981; (2): 42-44 [in Russian].

20. Filin BP. O printsipe dualnosti v zadachakh analiza strukturnoy nadiozhnosti slozhnykh sistem [On the duality principle in the structural dependability analysis of complex systems]. Avtomatika i telemekhanika 1989; 6: 158-172 [in Russian].

21. Ivanitskaya LG. O funktsiakh nadiozhnosti ustroistv relsovogo deistvia [On the dependability functions of rail devices]. In: Proceedings of science and technology conference of VZEIS teaching staff chaired by Varakin L.E., Doctor of Engineering, Professor. Moscow: VZEIS; 1967 . 1: 111-132 [in Russian].

22. Hansler E. A fast recursive to calculate the reliability of a communication network. IEEE Trans. Commun. 1972 ; 3: 637-642.

23. Bogatyrev VA. K raschiotu nadiozhnosti seti po sovokepnosti putey [On dependability calculation of network per all route]. Elektrosviaz 1981; 5: 42-44 [in Russian].

24. Riabinin IA, Cherkesov GN. Logiko-veroyatnostnye metody issledovania nadiozhnosti strukturno slozhnykh sistem [Logical and probabilistic methods of dependability study of structurally complex systems]. Moscow: Radio i sviaz; 1981.

25. Ushakov IA, Litvak EI. Verkhiaia i nizhniaia otsenki parametrov dvukhpoliusnoy seti [Lower and upper bound estimates of a two-pole net]. . Izv. AN SSSR, Tekhnicheskaya kibernetika; 1977.

26. Ushakov IA, Harrison R., editors. Handbook of Reliability Engineering. New York: John Wiley and Sons inc.; 1994.

27. Filin BP. O metode ekspress-otsenki i kefficiente potentsialnoy strukturnoy neuyazvimosti sviazey v slozhnykh sistemakh [On the method of express assessment of and coefficient of structural invulnerability of connections in complex systems]. Avtomatika i telemekhanika. 1994; 5: 158-182 [in Russian].

28. Netes VA, Filin BP. Consideration of Node Failures in Network-Reliability Calculation. IEEE Transactions on Reliability 1996; . 49 (1): 67-68.

29. Filin BP, Shaparev AV. Ob odnom podkhode k raschiotu veroyatnosti sokhranenia maksimalnogo potoka [On an approach to the calculation of the probability of maximum flow maintenance]. Avtomatika i telemekahinika 2001; 1: 102-117 [in Russian].

30. Busacker R, Saaty T. Finite graphs and networks. Moscow: Nauka; 1974.

31. Wentzel ES. Probability theory. Moscow: Nauka; 1964.

32. Burachenko VA, Kolesnikov AN, Korzhik VI, Fink LM. Obshchya teoria sviazy [General theory of communication]. Leningrad: SM Budyonny Military Academy of the Signal Corps; 1970 [in Russian].

33. Davydenko VP, Loskutov NG, Ivanov LT. Osnovy voennoy kibernetiki [Introduction to military cybernetics]. Leningrad: SM Budyonny Military Academy of the Signal Corps; 1971 [in Russian].

About the authors

Alexander G. Labutin, Moscow Technical University of Communication and Informatics, master's student (1st year); OOO Progress, 2nd category engineer, e-mail: aglabutin@gmail.com, Russia, Moscow

Boris P. Filin, Doctor of Engineering, retired, e-mail: filinbp41@mail.ru, Russia, Moscow

Received on 24.04.2017

Original article
DOI: 10.21683/1729-2646-2017-17-3-32-40

Method of increasing fault tolerance of satellite communication networks under information technology interference

Sergey M. Klimov, *4-th Central Research and Design Institute of the Ministry of Defence of Russia, Russia, Korolyov*
Sergey V. Polikarpov, *4-th Central Research and Design Institute of the Ministry of Defence of Russia, Russia, Korolyov*
Andrey V. Fedchenko, *4-th Central Research and Design Institute of the Ministry of Defence of Russia, Russia, Korolyov*

Abstract. Aim. *The aim of the paper is to develop a method that would allow for integrated experimental, computational, analytical and expert assessment of the vulnerabilities of satellite communication networks, feasibility of information technology interference by intruders against such vulnerabilities and probability of fault tolerance under the chosen information protection solutions, trusted information technologies and fault tolerance sensors. The paper shows the relevance and importance of the method of increasing fault tolerance of satellite communication networks under information technology interference in service control channels and satellite equipment data. The authors examine targeted information technology interference that causes malfunction of satellite modems, control stations and connected user computer networks. The paper shows the unique nature of satellite communication networks operation due to the global operating range, availability of broadband radio signals from communication and retransmission spacecraft for technical analysis and processing within the operating range, potential possibility of unauthorized connection to communication services. The primary direction of development of procedural and process engineering guidelines on protection and fault tolerance of satellite communication networks are defined.* **Methods.** *A method has been developed that is based on three components: model of experimental identification of satellite communication network vulnerabilities; simulation, computational and analytical model of detection and identification of threats of information technology interference; decision-making algorithm for improvement of the fault tolerance of a satellite communication network under information technology interference. The model of experimental detection of satellite communication network vulnerability allows, as part of bench tests, establishing connections between existing vulnerabilities of the hardware and software of satellite modems, control stations, user networks and the potential information technology interferences by intruders. As part of the vulnerability model the authors describe the certificates of vulnerable radio technical and information technology parameters of the satellite communication network signals, as well as suggest an analytic expression for calculating the probability of detection of such network vulnerabilities. The paper presents a computational and analytical model of detection and identification of information technology interference threats as a structure of advanced means of detection, prevention and elimination of the consequences of information technology interference in satellite communication networks and the mathematical expression for identification of conditional probability of materialization of the threat of information technology interference in satellite communication networks. An algorithm is considered for improvement of fault tolerance of satellite communication networks under information technology interference, including preparation of parameters and evaluation of the fault tolerance of a satellite communication network, adjustment of the parameters of satellite communication network, information security facilities and fault tolerance sensors, situational adjustment of satellite communication network fault tolerance solutions.* **Conclusions.** *It is noted that the developed method enables improved fault tolerance of satellite communication networks under information technology interference based on a set of interconnected procedures of the model of experimental detection of satellite communication network vulnerabilities on testbed; simulation, computational and analytical models of detection and identification of information technology interference threats; application of the decision-making algorithm of improvement of satellite communication network fault tolerance.*

Keywords: *satellite communication networks, vulnerabilities, information technology interference, fault tolerance improvement, information protection facilities, trusted information technologies and sensors of fault tolerance.*

For citation: Klimov SM, Polikarpov SV, Fedchenko AV. Method of increasing fault tolerance of satellite communication networks under information technology interference. Dependability 2017;3: p. 32-40. DOI: 10.21683/1729-2646-2017-17-3-32-40.

Introduction

According to the Doctrine of Information Security of the Russian Federation dated December 5, 2016, one of the primary tasks of information security in terms of state and public security is the improvement of the protection of critical information infrastructure and its operational stability, development of the mechanisms of detection and prevention of information threats and elimination of their consequences.

The satellite communication networks are one of the most complex and still-developing facilities of critical information infrastructure with stricter requirements for security and resilience to information technology interference (ITI). For instance, information and telecommunication systems based on DVB-RCS services of interactive satellite communication of large amounts of data are actively developing [1].

In the process of development of advanced models and methods of ensuring resilience (survivability) of information communication networks under destructive information technology interference [2] the characteristics of interrelated evaluation of actual security and fault tolerance of SCN can be taken into consideration by means of the proposed method of improving SCN fault tolerance under ITI.

Targeted and massive ITI against data communication protocols are a threat to the security and functional stability of SCN, which determines the requirement for the development of methods and facilities for detection, prevention and elimination of the consequences of such information technology threats based on improved fault tolerance of such networks [3, 5].

Currently, the most dangerous are those ITI that secretly penetrate, propagate, cause faults (failures) and damage to information technology resources of SCN.

The distinctive features of SCN, particularly those based on spacecraft (SC) in geostationary orbits, are:

- global operating range (e.g., the Yamal 401 SC serves the territory of the Russian Federation and cross-border regions)
- availability of broadband radio signals from communication and retransmission SC for technical analysis and processing within the operating range
- potential possibility of unauthorized connection to communication services within the distance up to 10 ths km between the users both from the territory of the Russian Federation and neighboring countries
- the hierarchical network infrastructure of SCN with geographically distributed users that interact through heterogeneous communication and retransmission SCs and landline networks using standardized data communication protocols.

This paper uses the following key terms:

- SCN vulnerability, software, architectural or logical deficiency of SCN that can be exploited to gain unauthorized access to protected information, compromise its integrity and availability, as well as cause SCN malfunction
- information technology interference (computer attack), targeted interference in automated and information and telecommunication systems by means of hardware and software facilities carried out for the purpose of causing malfunction and compromising information security in SCN
- SCN fault tolerance improvement method, a set of models and algorithms of analysis, detection of SCN vulnerabilities, detection of ITI threats, experimental, computational, analytical and expert evaluation of actual fault tolerance of SCN under ITI.

The primary future directions of development of procedural and process engineering guidelines of SCN protection and fault tolerance include the development of:

- simulators of targeted and massive ITI for SCN testing
- methods, algorithms and facilities for detection, prevention and elimination of consequences of ITI against SCN
- testbeds for testing SCN under ITI
- training facilities based on virtualization, cloud computing and multiplayer computer games involving ITI counteraction
- computer-assisted SCN operators training in preparation to actions under emergency threats of targeted and massive ITI.

Problem definition

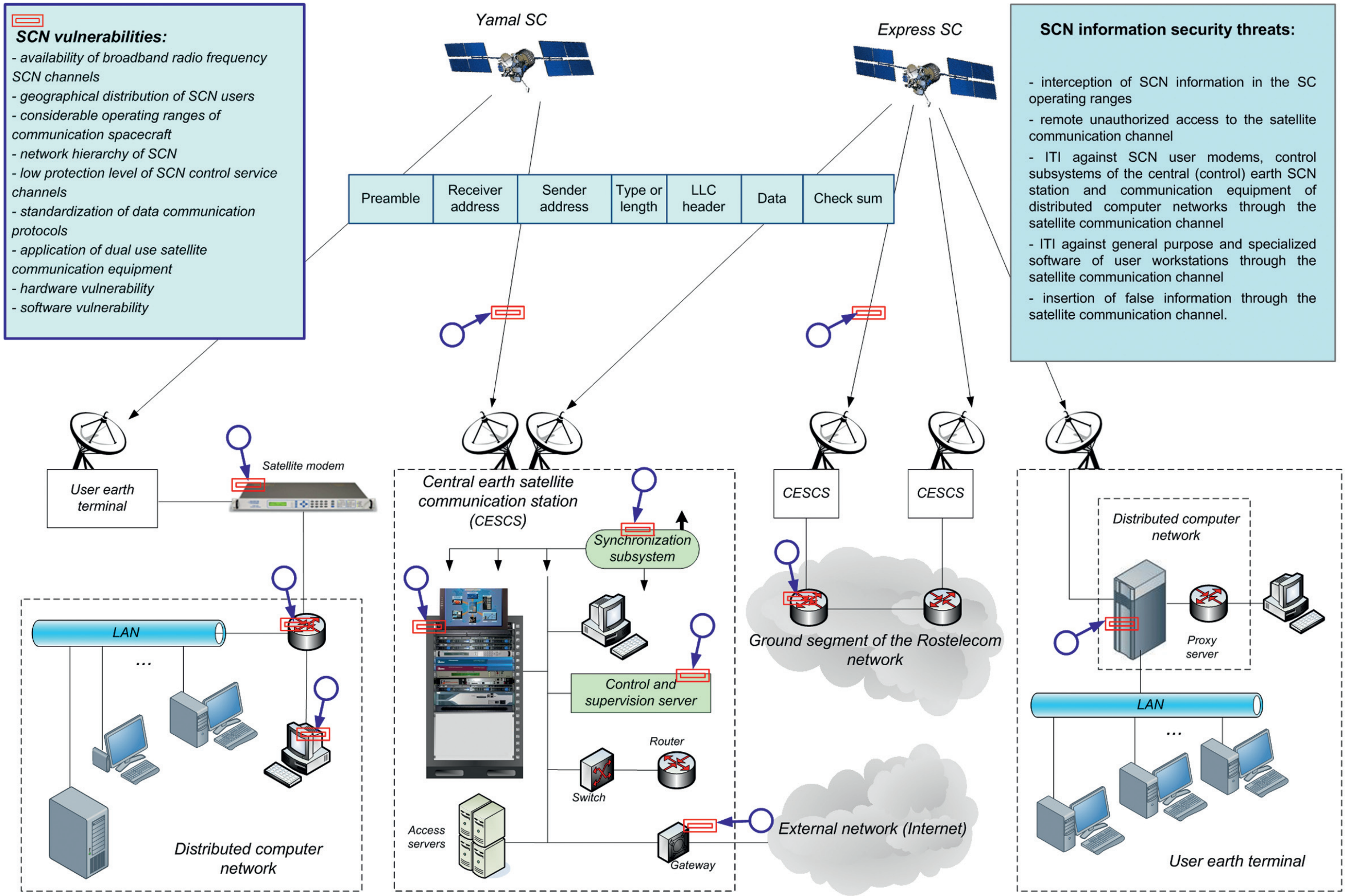
The method proposed in this paper is based on an interconnected set of experimental evaluations of actual SCN fault tolerance under simulated ITI, computational, analytical and expert evaluations of the achieved SCN fault tolerance under the chosen solutions, trusted information technologies (TIT), information security facilities (ISF) and fault tolerance sensors (detection, prevention and elimination of the consequences of ITI).

The method of improving SCN fault tolerance consists of the following models and algorithms:

1. Model of experimental identification of SCN vulnerabilities.
2. Simulation, analytical models of detection and identification of ITI threats.
3. Algorithm of decision making regarding the improvement of SCN fault tolerance under ITI.

Figure 1 shows the diagram of the model of experimental identification of SCN vulnerabilities. It is a generic SCN critical information technology infrastructure with typical vulnerabilities of geographically distributed facilities that interact through communication and retransmission SC and landline backbone networks. Vulnerabilities are associated with potential ITIs.

The primary vulnerability is the standard and open SCN data communication protocols that are not sufficiently secure in the service control and synchronization channels. High connectivity of heterogeneous segments of several SCNs enables unauthorized access to targeted objects that are not elements of individuals SCNs.



SCN vulnerabilities:
- availability of broadband radio frequency SCN channels
- geographical distribution of SCN users
- considerable operating ranges of communication spacecraft
- network hierarchy of SCN
- low protection level of SCN control service channels
- standardization of data communication protocols
- application of dual use satellite communication equipment
- hardware vulnerability
- software vulnerability

SCN information security threats:

- interception of SCN information in the SC operating ranges
- remote unauthorized access to the satellite communication channel
- ITI against SCN user modems, control subsystems of the central (control) earth SCN station and communication equipment of distributed computer networks through the satellite communication channel
- ITI against general purpose and specialized software of user workstations through the satellite communication channel
- insertion of false information through the satellite communication channel.

Yamal SC
Express SC
Preamble
Receiver address
Sender address
Type or length
LLC header
Data
Check sum
User earth terminal
Satellite modem
LAN
...
Distributed computer network
Central earth satellite communication station (CESCS)
Synchronization subsystem
Control and supervision server
Router
Switch
Access servers
Gateway
CESCS
CESCS
Ground segment of the Rostelecom network
External network (Internet)
Distributed computer network
Proxy server
LAN
...
User earth terminal

Figure 1. Model of experimental detection of SCN vulnerabilities

Experimental detection of SCN vulnerabilities is carried out by means of practical verification (testing) of the availability of control and information exchange protocols for technical analysis of radio technical parameters and revealing of information parameters of broadband signals of data communication channels through passive and active scanning.

The double red rectangle indicates the vulnerability, while the blue line with a circle indicates the ITI. The considered model of experimental detection of SCN vulnerabilities enables a preliminary evaluation of the exposure of an SCN with the employed information technologies and ISF to potential ITI. The experimental research is performed on actual SCN data communication channels with the use of communication and retransmission SC frequency resource, equipment for simulation of satellite communication channels between the control station and users taking into consideration the special features of the control cycle and with the use of a testbed system.

As part of the modelling process, the testbed hardware and software system enables technical analysis of the SCN radio technical and information technology parameters, simulation of ITI, interference environment according to the designed scenarios and input data. The testbed for experimental evaluation of SCN security under ITI must include antenna systems of various bandwidths, transceivers, demodulators, software that implement the actual operating process of SCN under ITI.

Table 1. An example of the certificate of vulnerable radio technical parameters of SCN signals

Description elements of vulnerable radio technical parameters	**Description of vulnerable SCN radio technical parameters**
Signal identifier	alphanumeric sequence that identifies a radio-frequency signal in the database
Name of communication and retransmission spacecraft	Yamal-401
Orbit location	90 e.l.
Transition frequency	3050 MHz
Signal level	-30 dBm
Carrier frequency	11245 MHz
Frequency band	20 MHz
Polarization type	horizontal (linear)
Modulation type	QPSK
Multiple access type	MF-TDMA
Type of interference immune coder	5/6
Type of turbo coder	Reed-Solomon
Scrambling	not used
Multiplexing	Yes

As the result, the model ensures the generation of the list of detected SCN vulnerabilities.

The certificate of vulnerable radio technical parameters of SCN signals available for technical review by an intruder is given in Table 1.

The analytic expression for calculation of the probability of detection of SCN vulnerabilities based on the experimental findings using [3-5] is as follows:

$$P^{\mathrm{SCN}}_{\mathrm{vuln}}\left(t_{\mathrm{sc}}\right) = P^{\mathrm{SCN}}_{\mathrm{NAA}}\left(t_{\mathrm{sc}}\right) + \left(1 + P_{\mathrm{NAA}}\left(t_{\mathrm{sc}}\right)\right) \cdot P_{\mathrm{PSC}}\left(t_{\mathrm{sc}}\right) + \\ + \left(1 - P_{\mathrm{NAA}}\left(t_{\mathrm{sc}}\right)\right) \cdot P_{\mathrm{ASC}}\left(t_{\mathrm{sc}}\right), \qquad (1)$$

where $P_{\mathrm{NAA}}\left(t_{\mathrm{sc}}\right)$ is the probability of points of unauthorized access in data communication channels between SCN users in the directions SC-to-Earth (back link with available broadband signals) and Earth-to-SC;

$P_{\mathrm{PSC}}\left(t_{\mathrm{sc}}\right)$ is the probability of passive scanning of data communication channels between SCN users within the time t_{sc};

$P_{\mathrm{ASC}}(t_{\mathrm{sc}})$ is the probability of active scanning of data communication channels between SCN users within the time t_{sc};

Let us represent the certificate of vulnerable information technology parameters of the SCN data communication channels using GOST R 56546-2015 (Table 2).

The simulation, computational, analytical models of detection and identification of threats of ITI in SCN define the list of threats to SCN information security that corresponds with the structure of the protocol of data communication between the control station and user segment, as well as SCN elements. The simulation model of ITI threats helps reveal the vulnerabilities of protocols, possible ways of implementing ITI threats against them and the consequences of SCN information security violations.

The presence of potential vulnerabilities and threats of ITI in SCN conditions the possibility of unauthorized connection to the network and interception of information.

The diagram of the model of detection and identification of SCN ITI threats is generated in the form of a structure of prompt response to ITI and improvement of fault tolerance is composed of three loops (Figure 2):

- first, SCN ITI
- second, control of information security and improvement of fault tolerance using the information from sensors of the ITI detection and identification facilities
- third, notification of information security violation using a generic form of computer incident description.

SCN TIT elements are secure hardware and software platforms that include operating systems (OS), database management systems (DBMS), translators from high-level programming languages and other general software. The hardware component of SCN TIT includes secure processors, memory modules, interface buses, satellite modems, control stations, trusted communication equipment of distributed landline networks that together must allow creating satellite and communication equipment immune to ITI.

Table 2. An example of the certificate of vulnerable information technology parameters of SCN

Elements of the description of vulnerable information technology parameters of SCN	Description of vulnerable information technology parameters of SCN
1. Name of vulnerability	Satellite modem control protocol vulnerability
2. Vulnerability identifier	USM-2017-00002
3. Brief description of vulnerability	Vulnerability allows interception of satellite modem software control channel
4. Vulnerability class	Satellite modem software vulnerability
5. Name of vulnerable element and its version	Satellite modem software ver. 7.34
6. Data communication protocol	Telnet control protocol, direct access to the satellite modem controls
7. Hardware and software design details	Hardware and software platform is based on the client/server and data communication protocol TCP/IP v.4.0 technology
8. Type of deficiency	Deficiencies related to operator authentication
9. Location of occurrence (manifestation) of vulnerability	Vulnerability exists due to the absence of legitimacy test of the source of satellite modem control
10. Deficiency type identifier	No data
11. Date of vulnerability detection	1.03.2017
12. Author of information on detected vulnerability	Information security unit
13. Means (rule) of vulnerability detection	Execution of step-by-step instructions
14. Vulnerability hazard criteria	Exceeding of specified risk probability value
15. Hazard level of vulnerability	High
16. Possible vulnerability elimination measures	Improvements to information protection facilities and satellite modem control protocols
17. Additional information	The network used satellite modems that allow remote software reboot via SCN

In the model of detection and identification of ITI the information security facilities (ISF) include on the one hand well-known ISFs such as automated trusted loading modules (ATLM), firewalls (FW), false network information entities (FNIE), and on the other had the set of sensors that implement the detection, prevention and elimination of consequences of ITI. In order to improve SCU fault tolerance under ITI the set of sensors records and identifies the facts of interference and generate the input data for SCN recovery.

SCN ISF elements should be implemented based on hardware and software facilities of the ITI monitoring and information security control station.

As part of SCN fault tolerance improvement the potential threats of SCN security and fault tolerance violations, potential target facilities, SCN TIT, ISF system and required testbed system for testing SCN under ITI are analyzed.

The distinction of improving the fault tolerance of SCN under ITI is that its required level must be ensured over a long period of operation in the context of ever-improving threats and control system paths with different data communication protocols.

The procedure of detection and identification of SCN ITI threats is as follows:

1. Installation of ITI detection server in the central control (Earth) station, installation of fault tolerance sensors at connected user control stations (terminals) and monitoring of SCN security.

2. Initial notification of the appearance of an unregistered user in case of unauthorized connection of an intruder to the SCN with active use of an unauthorized satellite modem.

3. Notification of a computer incident in case of ITI against the SCN users and control station by the fault tolerance sensors with the use of embedded signature-based and heuristic methods.

4. Analysis of targeted and massive ITIs, identification of their type and updating of ITI signatures database.

5. Decision-making regarding elimination of ITI consequences by means of blocking of the intruder's satellite modem based on the supervision of the parameters of the forward and reverse satellite channels, information from the fault tolerance sensors.

The monitoring of SCN security for ITI detection and identification is implemented by means of the following tests:

- supervision of radio technical parameters
- supervision of SCN control service channels
- supervision of SCN users connection and operation parameters
- supervision of SCN information traffic
- detection of insecure control and data communication channels

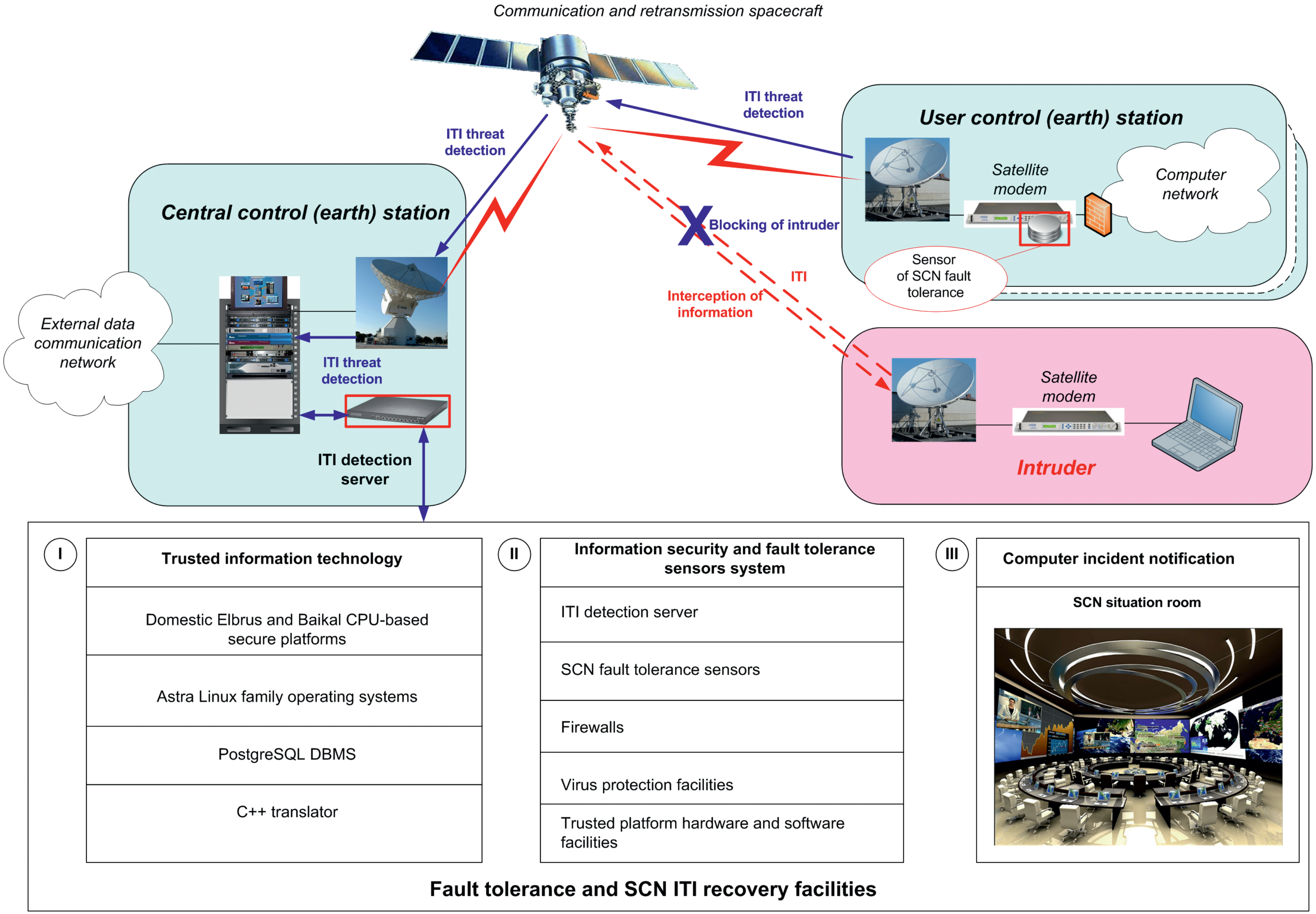

Communication and retransmission spacecraft
ITI threat detection
ITI threat detection
Central control (earth) station
External data communication network
ITI threat detection
ITI detection server
User control (earth) station
Satellite modem
Computer network
Sensor of SCN fault tolerance
Blocking of intruder
ITI
Interception of information
Satellite modem
Intruder
I
Trusted information technology
Domestic Elbrus and Baikal CPU-based secure platforms
Astra Linux family operating systems
PostgreSQL DBMS
C++ translator
II
Information security and fault tolerance sensors system
ITI detection server
SCN fault tolerance sensors
Firewalls
Virus protection facilities
Trusted platform hardware and software facilities
III
Computer incident notification
SCN situation room
Fault tolerance and SCN ITI recovery facilities

Figure 2. Diagram of the model of detection and identification of ITI threats

- detection of ITI in the forward and reverse SCN channels

- localization of the intruder (identification of ITI source and the supposed location)

- blocking of the intruder (disconnection from SCN)

- technical analysis and identification of SCN vulnerabilities

- supervision of the performance parameters of satellite equipment of the central and user stations, communication equipment, hardware and software, information security and fault tolerance improvement facilities.

The computational and analytical model is based on the mathematical expression for identification of the conditional probability of materialization of SCN ITI threats (with the use of [3-5]):

$$P\left(S_{\mathrm{O}}/Y_{\mathrm{ITI}}\right)=\frac{P\left(S_{\mathrm{O}i}\right)\cdot P\left(Y_{\mathrm{ITI}i}/S_{\mathrm{O}j}\right)}{P\left(S_{\mathrm{O}j}\right)\cdot P\left(Y_{\mathrm{ITI}i}/S_{\mathrm{O}j}\right)+P\left(S_{\mathrm{NO}k}\right)\cdot P\left(Y_{\mathrm{ITI}i}/S_{\mathrm{NO}k}\right)}, \quad (2)$$

where $P(S_{\mathrm{O}})$ is the probability of SCN being operable under ITI,

$P(Y_{\mathrm{ITI}i}/S_{\mathrm{O}j})$ is the conditional probability of materialization of the i-th ITI threat against the j-th operable SCN element,

$P(S_{\mathrm{NO}k})$ is the probability of the k-th SCN element being non-operable under ITI,

$P(Y_{\mathrm{ITI}i}/S_{\mathrm{NO}k})$ is the conditional probability of materialization of the i-th ITI threat against the k-th SCN element that causes violation of its fault tolerance.

Figure 3 shows the algorithm of improvement of SCN fault tolerance under ITI.

The algorithm of SCN fault tolerance is based on a set of experimental, computational and analytical and expert evaluations of SCN parameters under ITI for the purpose of selecting SCN fault tolerance solutions by means of their situational adjustment.

The essence of the algorithm of SCN fault tolerance under ITI consists in the implementation five stages:

1. Preparation of parameters (input data) for evaluation of SCN fault tolerance under ITI

2. SCN actual security evaluation.

3. SCN parameters adjustment.

4. ISF and fault tolerance facilities parameters adjustment.

5. Situational adjustment of SCN fault tolerance solutions.

The algorithm of SCN fault tolerance under ITI essentially defines the logic of stage-by-stage management of network fault tolerance, selection of the least vulnerable

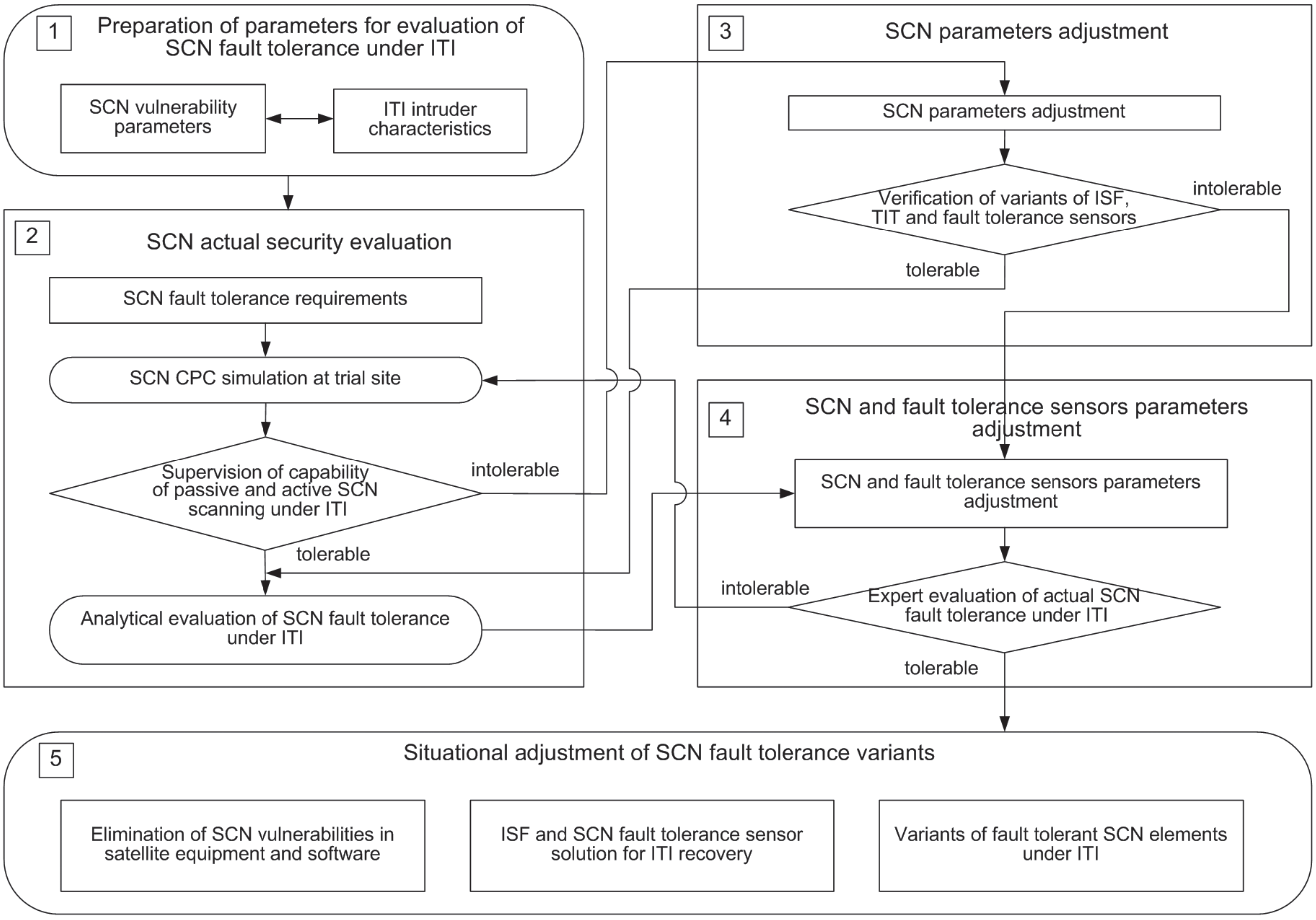

1
Preparation of parameters for evaluation of SCN fault tolerance under ITI
SCN vulnerability parameters
ITI intruder characteristics
2
SCN actual security evaluation
SCN fault tolerance requirements
SCN CPC simulation at trial site
Supervision of capability of passive and active SCN scanning under ITI
intolerable
tolerable
Analytical evaluation of SCN fault tolerance under ITI
3
SCN parameters adjustment
SCN parameters adjustment
Verification of variants of ISF, TIT and fault tolerance sensors
intolerable
tolerable
4
SCN and fault tolerance sensors parameters adjustment
SCN and fault tolerance sensors parameters adjustment
Expert evaluation of actual SCN fault tolerance under ITI
intolerable
tolerable
5
Situational adjustment of SCN fault tolerance variants
Elimination of SCN vulnerabilities in satellite equipment and software
ISF and SCN fault tolerance sensor solution for ITI recovery
Variants of fault tolerant SCN elements under ITI

Figure 3. Algorithm of SCN fault tolerance under ITI

and most secure SCN configuration with the capability to eliminate ITI consequences.

The SCN control cycle is characterized by the real-time operation, dynamic ITI environment, large numbers of TIT, which requires situational adjustment of the SCN fault tolerance parameters according to the current situation.

It is assumed that the initial stage of algorithm operation, the input ITI data, SCN vulnerabilities, TIT, ISF and fault tolerance sensors are undefined, which causes the requirement for experimental and analytical research to ensure ITI fault tolerance under ITI.

The algorithm is to be used both at the stage of secure SCN design and development and the stage of SCN parameters and fault tolerance facilities adjustments in operation.

The implementation of the stages of the algorithm of SCN fault tolerance adjustment under ITI ensures:

- preparation of input data and taking account of the factors for evaluation of SCN fault tolerance under ITI
- simulation of real SCN operation processes using the testbed system
- instrumental verification of the intruder's potential capabilities in terms of passive and active scanning of SCN vulnerabilities
- adjustment of SCN, ISF, TIT and fault tolerance sensors parameters based on the results of experimental, computational and analytical and expert evaluations
- situational adjustment of SCN fault tolerance solutions under dynamic ITI based on elimination of vulnerabilities of hardware and software, selection of SCN elements, ISF and fault tolerance sensors fault tolerance solutions
- selection of the SCN models, ITI, ISF, TIT, fault tolerance sensors for the purpose of identifying the measures of improving (insuring) SCN functional stability
- comprehensive experimental, computational and expert evaluation of SCN fault tolerance under ITI
- preparation of elimination of the consequences of ITI against SCN
- efficient management of SCN fault tolerance under uncertain ITI and various degradations in the control cycle in terms of satellite communication and data transmission services
- evaluation of fulfilment of the requirements for SCN fault tolerance under ITI and associated practical recommendations.

The many various parameters of SCN, ISF, TIT, fault tolerance sensors and ITI factors lead to a multitude of possible current SCN fault tolerance conditions. In practice, the number of adjustment (control) solutions for improving SCN fault tolerance is limited. In this context the considered algorithm based on situational adjustment of SCN fault tolerance solutions enables the fulfilment of the SCN fault tolerance requirements.

The situation is understood as the sum of vulnerabilities and states of SCN, ISF, TIT and fault tolerance sensors at a certain moment of operation for the purpose of identifying the requirement for an intervention in the SCN control process.

The procedures of supervision, verification and expert evaluation of the SCN fault tolerance parameters under ITI implies that based on the data regarding the testbed findings as part of decision-making at the stages of the algorithm implementation the facts are established of allowable and unallowable deviation of the current SCN state from the required values.

The interval of SCN fault tolerance state adjustment is chosen based on the significance of the users and units (control stations) and most probable ITI implementation.

If the state of SCN under ITI requires situational adjustment, its description is classified on the basis of the certificate of the SCN radio technical and information technology parameters, normal behavior profile and results of computer incident investigation. Each current state of the SCN under ITI can be assigned to a certain class that is associated with a certain set of adjustable parameters.

Situational adjustment of SCN fault tolerance is done in the following order: description of SCN state under ITI – preparation of adjustable parameters of SCN, ISF, TIT, fault tolerance sensors – control inputs for ITI counteractions and SCN recovery (initiation of backup segments).

The calculation formula for the probability of fault tolerance of SCN fault tolerance under ITI with the use of [3-5] is as follows:

$$P_{\text{fl}}(t)=\prod_{i=1}^{k}\left[1-\left(1-\prod_{j=1}^{r}P_{\text{PNF}i}\left(S_{\text{O}j}\right)\right)\right], \qquad (3)$$

where $P_{\text{PNF}i}\left(S_{\text{O}j}\right)$ is the probability of no-failure of the j-th SCN element under the i-th ITI threat;

k is the total number of ISF (sensors) within the SCN;

r is the total number of backup SCN elements.

Conclusion

The paper suggests a method of improving SCN fault-tolerance under possible ITI based on system analysis of potential vulnerabilities and unique operational characteristics of SCN. The method is based on a set of interconnected procedures of the model of experimental detection of SCN vulnerabilities on testbed, generation of the simulation, computational, analytical models of detection and identification of ITI threats and application of the decision-making algorithm of improvement of SCN fault tolerance under ITI.

References

1. Voronin AV, Ivanov VN, Sotov AM. Somov AM, Doctor of Engineering, editor. Tsyfrovoe televizionnoe veshchanie [Digital television broadcasting]. Moscow: Goriachaia linia-Telekom; 2017 [in Russian].

2. Velichko VV, Popkov GV, Popkov VK. Modeli i metody povyshenia zhivuchesty sovremennykh sistem sviazi

[Models and methods of improving the resilience of present day communication systems]. Moscow: Goriachaia linia-Telekom; 2016 [in Russian].

3. Klimov SM, Astrakhov AV, Sychiov MP. Eksperimentalnaia otsenka protivodeystvia kompiuternim atakam [Experimental evaluation of computer attack reaction]. Moscow: Bauman MSTU; 2013 [in Russian].

4. Kapur K, Lamberson L. Ushakov IA, editor. Reliability in engineering design. Moscow: Mir; 1980 [in Russian].

5. Shubinsky I.B. Nadiozhnie otkazoustoychivie informatsionnie sistemy. Metodi sinteza [Dependable failsafe information systems. Synthesis methods]. Ulianovsk: Oblastnaya tipografia Pechatny dvor; 2016 [in Russian].

About the authors

Sergey M. Klimov, Doctor of Engineering, Professor, Head of Division, 4-th Central Research and Design Institute of the Ministry of Defense of Russia. Russia, Korolyov, phone: + 7 (985) 928 13 55, e-mail: klimov.serg2012@yandex.ru.

Sergey V. Polikarpov, Deputy Head of Unit, 4-th Central Research and Design Institute of the Ministry of Defense of Russia. Russia, Korolyov, phone: +7 (916) 332 60 66, e-mail: polikarpov.s.v@yandex.ru.

Andrey V. Fedchenko, Head of unit, 4-th Central Research and Design Institute of the Ministry of Defense of Russia. Russia, Korolyov, phone: +7 (916) 334 87 89, e-mail: fedchandr@yandex.ru

Received on 15.05.2017

Method of recovery of priority vector for alternatives under uncertainty or incomplete expert assessment

Alexander V. Bochkov, *Research and Design Institute of Economy and Business Administration in the Gas Industry, Moscow, Russia*
Nikolai N. Zhigirev, *Research and Design Institute of Economy and Business Administration in the Gas Industry, Moscow, Russia*
Alexandra N. Ridley, *Postgraduate Student, MAI (NRU), Moscow, Russia*

Abstract. Aim. *The so-called pair-wise comparison method is one of the most popular decision-making procedures owing to its efficiency, flexibility and simplicity. The primary disadvantage of this method in the context of expert evaluation of large numbers of alternatives or within a sufficiently wide field of knowledge is the impossibility to compare each element with each other, both due to the large number of such comparisons, random gaps and difficulties experienced by the expert while comparing some alternatives. The assessments are affected by gaps that complicate decision-making, as most statistical methods are not applicable to incomplete sets of data. The fairly popular algorithm for processing of pair-wise comparison matrices (the Saaty algorithm) cannot work with matrices that predominantly contain zero components. The purpose of the paper is to develop a method of processing comparison matrices in order to obtain weight coefficients (weights) of the considered alternatives that enable quantitative comparisons.* **Methods.** *In practice, there are several approaches to managing sets of data with gaps. The first, most easily implementable, approach involves the elimination of copies with gaps from the set with further handling of only complete data. This approach should be used in case gaps in data are isolated. Although even in this case there is a serious risk of "losing" important trends while deleting data. The second approach involves using special modifications of data processing methods that tolerate gaps in sets of data. And, finally, there are various methods of evaluation of missed element values. Those methods help to fill in the gaps in sets of data based on certain assumptions regarding the values of the missing data. The applicability and efficiency of individual approaches, in principle, depends on the number of gaps in data and reasons of their occurrence. In this paper, the pair-wise comparison matrix is considered in the form of a loaded graph, while the alternatives are the nodes and comparisons are the edges of the graph. Respectively, if a pair of alternatives occurs for which the expert could not specify a preference, the corresponding edge is absent. The paper considers a way of removing edges that correspond to the most controversial values, i.e. a cycle breakage algorithm that causes transformation of the initial graph to the spanning tree that allows for unambiguous comparison of any two alternatives. The algorithm of joint alignment of both the upper and lower boundaries of expert assessments is not considered in this paper.* **Results.** *The paper gives an example of practical application of the developed algorithm of processing incomplete matrices of pair-wise comparisons of ten objects obtained in a certain expert assessment. It also shows the efficiency of the suggested approach to priority recovery of compared alternatives, explores ways of automating computing and future lines of research.* **Conclusions.** *The proposed method can be used in a wide range of tasks of analysis and quantitative evaluation of risks, safety management of complex systems and objects, as well as tasks related to the verification of compliance with the requirements for such highly dependable elements as nuclear reactors, aviation and rocket technology, gas equipment components, etc., i.e. in cases when low (less than 0,01) probabilities of failure per given operation time are to be evaluated, while the failure statistics for such elements in operation is practically nonexistent. The proposed algorithm can be applied in expert assessment in order to identify the type and parameters of time to failure distribution of such highly dependable elements, which in turn will allow evaluating dependability characteristics with the required accuracy.*

Keywords: *missed data, expert assessment, loaded graph, spanning tree, pair-wise comparison graph, connectivity criterion.*

For citation*: Bochkov AV, Zhigirev NN, Ridley AN. Method of recovery of priority vector for alternatives under uncertainty or incomplete expert assessment. Dependability 2017;3: p. 41-48. DOI: 10.21683/1729-2646-2017-17-3-41-48*

Introduction

The decision-making procedures that involve experts in choosing the optimal variant(s) out of the allowed set are often used in a variety of areas for assessment, selection, definition of task priority, etc. Obviously, the comparison of various alternatives based on their preferability in terms of decision-making tasks in many cases is unfeasible using one criterion or one expert. Consequently, in most decision-making tasks there are procedures that allow combining the opinions of several experts regarding the alternatives presented to them [1, 2]. In most cases those procedures use the so-called pair-wise comparison method that assumes that an expert may prefer one alternative to another one while comparing them.

As each expert has a unique experience of solving specific problems, the opinions of various experts may significantly differ (indeed, there are many factors that affect an expert's preferences). This variety of expert assessments may cause a situation where some of them are unable to adequately express any degrees of preference by comparing two or more available alternatives. That may be caused by insufficient competence of the expert in an area of knowledge that pertains to the task or due to the fact that the expert is incapable of identifying the degree of preference of some of the presented variants over the others. In such situations such expert has to ensure fuzzy preference relation [3] or abandon the assessment of the presented pair of alternatives. A non-trivial task arises whereas missed data must be recovered in order to obtain acceptable results of expert assessment.

In practice, there are several approaches to managing sets of data with gaps. The first, most easily implementable, approach involves the elimination of copies with gaps from the set with further handling of only complete data [4]. This approach should be used in case gaps in data are isolated. Although even in this case there is a serious risk of "losing" important trends while deleting data. In the same case, when the number of gaps is too high, the removal of the respective copies may cause a data deficiency or even impossibility of further processing. The second approach involves using special modifications of data processing methods that allow gaps in sets of data. In [5], the authors describe a number of modifications of classification and clustering methods for managing data that contain missed values. And, finally, the third, most common, approach is the use of methods of evaluation of missed element values. Those methods help to fill in the gaps in sets of data based on certain assumptions regarding the values of the missing data. The applicability and efficiency of individual approaches, in principle, depends on the number of gaps in data and reasons of their occurrence. In terms of the nature of data origins, the categories of gaps that are usually identified are set forth in [6].

Quite frequently, empirical research has to reject the results of expert polls if some data is missing [7].

In [8], the effects of the above sets of pair-wise comparisons are researched. The paper compared the results for complete pair-wise comparison matrices and incomplete ones that were obtained by removing known elements from complete ones. The findings of [8] have shown that "random removal of up to 50 percent of comparisons provides good results with no loss of accuracy". Nevertheless, as this process is based on a priori knowledge of the complete pair-wise comparison matrix it is not applicable in practice. Thus, [8] suggests – for the cases of incomplete pair-wise comparison matrices – using methods that allow "completing" the matrix. A strong argument in support of this approach is set forth in [9]: "as a rule, a scenario with missed values disrupts the rating more significantly than the same scenario with a value". A system that helps build fuzzy preference relations in a solution was suggested in [10]. In the decision-making group, procedures that correct the absence of knowledge in a specific expert using information provided by the other experts, along with some aggregation procedures can be found in [11] and [12]. Those approaches have a number of disadvantages some of which are noted by the authors of [13].

In Russian literature, there are also works related to the potential solutions of the above problems, e.g. [14], however the approaches used in them do not provide a clear solution.

1. Problem definition

In the classic Saaty setting [15] there is a certain set of objects O_1, O_2, …, O_N (possible actions, parameters, alternative solutions, etc.) with a certain hierarchy. An expert's quantitative judgement regarding a pair of objects (O_i, O_j) is represented with a matrix of size n×n: $A=(a_{ij})$, (i, j=1, 2, …, n), where the numbers a_{ij}, of which the matrix consists correspond with the object's significance O_j compared to O_i and are non-negative. In order to identify the quantitative indicators of the relative significance of the compared objects the method suggests a scale of relative comparisons expressed in whole numbers from 1 to 9. Objects with equal significance are rated "1". The ratings along the main diagonal of the matrix are also "1" (objects are compared to themselves), i.e. $a_{ii}=1$. The Saaty matrix is antisymmetrical about the main diagonal, i.e. $a_{ij}=1/a_{ji}$.

Further, in [15], after the quantitative judgements regarding the pairs (O_i, O_j) have been formed in numeric expressions in terms of a_{ij} the task comes down to associating each of the compared objects numeric weights that would best match the stated expert judgements. In order to find the priority vector it is required to find the vector ω that fulfils the condition $A\omega=\lambda_{\max}\omega$. According to the theorem on the existence and uniqueness, while solving the eigenvalue problem for the non-negative matrix as per [15, 16] the resultant eigenvector is found, which after normalization becomes the priority vector of the compared objects.

We are seeking the solution for the case when in the matrix A some assessments are not defined, i.e. $\exists a_{ij}:a_{ji}=NA$ (NA stands for Not Available).

2. Method description

The incomplete pair-wise comparison matrix $\overline{A}$ is easily transformed into the skew-symmetric matrix $\overline{\overline{A}}$ by taking logarithms of the elements of the matrix of coefficients.

The weights W_i are transformed into V_i=ln(W_i), the pair-wise comparison coefficients matrix $\overline{S}_{ij} = \ln(S_{ij})$, while the residuals matrix $F_{ij} = \left(\frac{S_{ij} \times W_j}{W_i} - 1 \right)$ is transformed into the functionals matrix $\overline{F}_{ij} = \left(\overline{S}_{ij} + V_j - V_i \right)$ and becomes skew-symmetric.

$$W_i, S_{ij}, F_{ij} = \left(\frac{S_{ij} \times W_j}{W_i} - 1 \right) \rightarrow V_i = \ln\left(W_i\right),$$

$$\overline{S}_{ij} = \ln\left(S_{ij}\right), \overline{F}_{ij} = \left(\overline{S}_{ij} + V_j - V_i\right) \qquad (1)$$

Having identified the values V_i, we perform a backward transformation:

$$V_i \rightarrow W_i = \frac{\exp(V_i)}{\sum_{j=1}^{N} \exp(V_j)} \qquad (2)$$

in order to fulfil the Saaty weight normalization requirement, i.e.

$$\sum_{i=1}^{N} W_i = 1, W_i > 0 \left(i = 1, \ldots, N\right).$$

The matrix may be diagonalized in order to make it positive above the main diagonal. However, it should be noted that such diagonalization is not always possible even in case of a single expert and admittedly unnecessary as only the matrix graph connectivity matters. For a group of experts diagonalization is necessary, as depending on the vertex degree the upper and lower assessments swap places. It is assumed that for a group of experts total agreement is unachievable, while their preference coefficient $\overline{S}_{ij}$ is within a certain range:

$$B_{ij} \le \overline{S}_{ij} \le T_{ij} \qquad (3)$$

The lower bound B_{ij} (Bottom) corresponds to the minimal value, while the upper bound T_{ij} (Top) corresponds to the maximum value. In this paper we examine the implementation of the algorithm for the upper estimates $T_{i,j}$ or one expert $\overline{S}_{i,j}$.

Due to the particular properties of the matrix T_{ij} the initial values of weights can be random constant values in order not to be confused by the weight increment signs. The value of change depends on the type of line/column. The type equals to "-1" when in the j-th line of the matrix ($E_1 - \ldots - N_1; \ldots; E_n - \ldots - N_n$) above the main diagonal there are only indeterminate values (*NA*). The type equals to "1" when in the i-th column of the matrix ($E_1 - \ldots - N_1; \ldots; E_n - \ldots - N_n$) above the main diagonal there are only indeterminate values (*NA*). In those cases when actual data are present both in the j-th line and in the i-th column it is obvious that the type of the first object always equals to "1", while the type of the last object equals to "-1".

Initially the matrix $\left\| \overline{F}_{ij} \right\|$ equals to the matrix $\left\| \overline{S}_{ij} \right\|$. The arbitrary choice of the weights is due to the fact that the result does not depend on the choice of the initial approximation

$$\overline{F}_{ij} = \left(\overline{S}_{ij} + \left[V_j + const\right] - \left[V_i + const\right]\right) = \left(\overline{S}_{ij} + \left[V_j\right] - \left[V_i\right]\right). \quad (4)$$

The first important stage is finding the hardest contradiction. We seek the optimal weight displacement for all lines through the optimization of the weight V_i that involves both the j-th line and the i-th column. We chose lines in the descending order (i=N,…,1) until all displacements become zero[1].

Then, we take

$$\left|\Delta V_i\right| \le \varepsilon C = 1{,}0e^{-7}. \qquad (5)$$

Weight change happens step by step in accordance with

$$V_i = V_i + \Delta V_i. \qquad (6)$$

After the next step ΔV_i becomes zero.

The value ΔV_i is chosen using the algorithm based on the modulus optimization in the matrix $\left\| \overline{F}_{ij} \right\|$, but only as regards the i-th step (i-th line).

We minimize the maximum moduli of the following values

$$\max_{\square} \left\{ \left| R^i_{\max} - \Delta^i \right|, \left| R^i_{\min} - \Delta^i \right|, \left| C^i_{\max} + \Delta^i \right|, \left| C^i_{\min} + \Delta^i \right| \right\} \rightarrow \min_{\Delta} \square \quad (7)$$

for all i=1,…,N, where N is the dimension of the problem; $R^i_{\max}$ is the maximum value above the diagonal in the i-th line; $R^i_{\min}$ is the minimum value above the diagonal in the i-th line; $C^i_{\max}$ is the maximum value above the diagonal in the i-th column; $C^i_{\min}$ is the minimum value above the diagonal in the i-th column.

The modified weight is calculated using the following formulas:

$$W^i = W^i + \Delta^i. \qquad (8)$$

The respective solutions for any line are identified using the formulas given in Table 1.

The next step is the removal of cells. If it is allowable according to the connectivity criterion[2], we remove the first edge belonging to the right-hand side of the «floater» against which in the line there is an edge from the «left-hand» side of the «floater», when the edge (i,j) is the only edge that connects non-overlapping node (object) subset.

Upon removal of the edge the solution restarts from the initial conditions.

It should be noted that it is important to track possible changes in the type of the "line/column". Usually the change

[1] For instance, for a matrix with the dimension of 10 the number of required iterations is around 15.

[2] The connectivity criterion is important in cases when the spanning graph may be discontinued.

Table 1

Type	Condition	Calculation formula
0	There are elements in the line and elements in the adjacent column	
	if $R_{\max} > \max\left\{\left\|R_{\min}\right\|,\left\|C_{\min}\right\|,C_{\max}\right\}$, then	$\Delta^{(I)} = \frac{1}{2}\cdot\min\left\{R_{\max}+R_{\min};R_{\max}-C_{\max}\right\}$
	if $C_{\max} > \max\left\{\left\|R_{\min}\right\|,\left\|C_{\min}\right\|,R_{\max}\right\}$, then	$\Delta^{(II)} = \frac{1}{2}\cdot\min\left\{-C_{\max}-C_{\min};R_{\max}-C_{\max}\right\}$
	if $-R_{\max} > \max\left\{\left\|C_{\min}\right\|,C_{\max},R_{\max}\right\}$, then	$\Delta^{(III)} = \frac{1}{2}\cdot\min\left\{R_{\max}+R_{\min};R_{\min}-C_{\min}\right\}$
	if $-C_{\max} > \max\left\{\left\|R_{\min}\right\|,C_{\max},R_{\max}\right\}$, then	$\Delta^{(IV)} = \frac{1}{2}\cdot\min\left\{-C_{\max}-C_{\min};R_{\min}-C_{\min}\right\}$
	otherwise	$\Delta^{(V)} = 0$
1	All components of the column are equal to NA	$\Delta V_i = \frac{\left(R^i_{\max}+R^i_{\min}\right)}{2}$
-1	All components of the line are equal to NA	$\Delta V_i = \frac{-\left(C^i_{\max}+C^i_{\min}\right)}{2}$

is from “0” to “1”, i.e. when the last significant element in the respective column disappears.

The next stage is the procedure that consists of two embedded cycles. The external involves finding the next “boundary cycle” in the remaining matrix. The internal cycle involves finding the edges of which the “boundary cycle” consists.

The end of the external cycle is the condition:

$$\left|sup\left\|F_{ij}\right\|\right| < \varepsilon_F = 1{,}0\cdot e^{-5} \qquad (9)$$

The resulting solution will be the solution of the problem for the single expert case. For the group of exerts case, this solution is for reference only ($V_i^{\text{ref.}}$).

It is required to take into consideration the remoteness of the lower bounds defined by the matrix $\left\|B_{ij}\right\|$.

For that purpose, let us recover the initial configuration of the graph.

Let us calculate the resulting matrix $\left\|R_{ij}\right\|$:

$$R_{ij} = -(B_{ij}+V_j-V_i) \qquad (10)$$

and remove all negative elements.

Some edges can complement the reference configuration.

Then, let us perform the following procedure that also has two cycles.

In the external cycle we define the most critical edge that can be dropped. For that purpose, the scale of single displacement for each weight U_i is calculated, the size of single displacement step for each edge (i,j) D_{ij} and the size of the step h using the formula

$$h = \min_{i,j}\left(\frac{R_{ij}}{D_{ij}}\right). \qquad (11)$$

In this paper we omit the formulas for D_{ij} and U_i.

Next, in the internal cycle we identify the virtual optimal solution using the formulas:

$$V_i^{\text{opt.}} = V_i^{\text{ref.}} - h\times U_i,\ (i=1,\ldots,N) \qquad (12)$$

$$R_{ij}^{\text{opt.}} = R_{ij}^{\text{ref.}} - h\times D_{ij},\ (i=1,\ldots,N-1;\ j=i+1,\ldots,N) \qquad (13)$$

If the emerged zero elements $R_{ij}^{\text{opt.}}$ do not cause the graph to lose connectivity, the edge of the first one can be omitted. If, on the contrary, the output matrix loses connectivity, the

Table 2. Initial coefficient values

	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
O1	1	$\frac{1}{3}$	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{1}{7}$	$\frac{1}{3}$	NA	NA	NA	$\frac{1}{4}$
O2	$\frac{3}{1}$	1	$\frac{1}{5}$	NA	$\frac{1}{5}$	$\frac{1}{4}$	NA	NA	NA	$\frac{1}{3}$
O3	$\frac{5}{1}$	$\frac{5}{1}$	1	$\frac{3}{1}$	NA	$\frac{3}{1}$	$\frac{3}{1}$	$\frac{5}{1}$	$\frac{5}{1}$	$\frac{3}{1}$
O4	$\frac{3}{1}$	NA	$\frac{1}{3}$	1	$\frac{1}{3}$	NA	NA	NA	NA	$\frac{1}{3}$
O5	$\frac{7}{1}$	$\frac{5}{1}$	NA	$\frac{3}{1}$	1	$\frac{3}{1}$	$\frac{5}{1}$	$\frac{5}{1}$	$\frac{5}{1}$	$\frac{3}{1}$
O6	$\frac{3}{1}$	$\frac{4}{1}$	$\frac{1}{3}$	NA	$\frac{1}{3}$	1	NA	NA	NA	$\frac{1}{4}$
O7	NA	NA	$\frac{1}{3}$	NA	$\frac{1}{5}$	NA	1	NA	NA	$\frac{1}{5}$
O8	NA	NA	$\frac{1}{5}$	NA	$\frac{1}{5}$	NA	NA	1	NA	$\frac{1}{5}$
O9	NA	NA	$\frac{1}{5}$	NA	$\frac{1}{5}$	NA	NA	NA	1	$\frac{1}{5}$
O10	$\frac{4}{1}$	$\frac{3}{1}$	$\frac{1}{3}$	$\frac{3}{1}$	$\frac{1}{3}$	$\frac{4}{1}$	$\frac{5}{1}$	$\frac{5}{1}$	$\frac{5}{1}$	1

Table 3. Logarithms of coefficient values

	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
O1	0	-1,09861	-1,60944	-1,09861	-1,94591	-1,09861	NA	NA	NA	-1,38629
O2	1,098612	0	-1,60944	NA	-1,60944	-1,38629	NA	NA	NA	-1,09861
O3	1,609438	1,609438	0	1,098612	NA	1,098612	1,098612	1,609438	1,609438	1,098612
O4	1,098612	NA	-1,09861	0	-1,09861	NA	NA	NA	NA	-1,09861
O5	1,94591	1,609438	NA	1,098612	0	1,098612	1,609438	1,609438	1,609438	1,098612
O6	1,098612	1,386294	-1,09861	NA	-1,09861	0	NA	NA	NA	-1,38629
O7	NA	NA	-1,09861	NA	-1,60944	NA	0	NA	NA	-1,60944
O8	NA	NA	-1,60944	NA	-1,60944	NA	NA	0	NA	-1,60944
O9	NA	NA	-1,60944	NA	-1,60944	NA	NA	NA	0	10,65705
O10	1,386294	1,098612	-1,09861	1,098612	-1,09861	1,386294	1,609438	1,609438	1,609438	0

Table 4. Reordering of objects

	O5	O3	O10	O6	O2	O4	O1	O7	O8	O9
O5	0	NA	1,0986	1,0986	1,7041	1,0986	1,9459	1,7041	1,7041	1,7041
O3	NA	0	1,0986	1,0986	1,7041	1,0986	1,7041	1,0986	1,7041	1,7041
O10	-1,0986	-1,0986	0	1,3863	1,0986	1,0986	1,3863	1,7041	1,7041	1,7041
O6	-1,0986	-1,0986	-1,3863	0	1,3863	NA	1,0986	NA	NA	NA
O2	-1,7041	-1,7041	-1,0986	-1,3863	0	NA	1,0986	NA	NA	NA
O4	-1,0986	-1,0986	-1,0986	NA	NA	0	1,0986	NA	NA	NA
O1	-1,9459	-1,7041	-1,3863	-1,0986	-1,0986	-1,0986	0	NA	NA	NA
O7	-1,7041	-1,0986	-1,7041	NA	NA	NA	NA	0	NA	NA
O8	-1,7041	-1,7041	-1,7041	NA	NA	NA	NA	NA	0	NA
O9	-1,7041	-1,7041	-1,7041	NA	NA	NA	NA	NA	NA	0

last successful attempt is "memorized" as a real optimal value for the weights $R_{ij}^{\text{opt.}}$.

If the matrix $R_{ij}^{\text{opt.}}$ matches the connectivity matrix of the matrix $R_{ij}^{\text{ref.}}$, the single displacement U_i=1 (type *i*=1); U_i=0 (type *i*=–1).

After node run in $R_{ij}^{\text{opt.}}$ the optimal solution takes its final form. It connects to the lower bound and is the optimal solution deduced using the higher bound data. It is assumed that the conflict of interests is at the higher bound where each expert wants to define his/her own priorities.

3. Example of practical application of the method

Let us assume there is a pair-wise comparison matrix filled by experts using the Saaty method.

The matrix is not complete (missing assessments are marked *NA*), because the experts could not express their preferences while comparing some pairs of objects (e.g. O_1 and O_7, O_2 and O_4, etc.).

Let us transform the incomplete pair-wise comparison matrix into a skew-symmetric matrix by taking logarithms of the elements of the matrix of coefficients.

The matrix after diagonalization that we perform so that the matrix is positive above the main diagonal is shown in Table 4.

Let us examine the implementation of the algorithm[1] for upper estimates $T_{i,j}$ or single expert $\overline{S}_{i,j}$.

Due to the particular properties of the matrix T_{ij} the initial values of weights (column B, Table 5) can be random constant value in order not to be confused by the weight increment signs. The value of change (column C, Table 5), as we said above, depends on the type of line/column.

In order to find the hardest contradiction we seek the optimal weight displacement for all lines through the optimization of the weight V_i that involves both the *j*-th line and the *i*-th column. We select lines in the descending order (*i*=*N*,…,1) until **all displacements** in column C (Table 5) become zero.

Using the above algorithm, we find the boundary modulo cycle for the initial conditions. The result is given in table 6.

Which cell must be removed? On the right-hand side of the «floater» that are (2,3), (3,4), (4,5), (5,7), on the left-hand side that is only the edge (2,7). The edges on the right-hand side indicate that the 2-nd object is better that the 3-rd one, the 3-rd object is better than the 4-th one, the 4-th one is better that the 5-th one, the 5-th one is better than the 7-the one $e^{0{,}8029}$=2,3632 times.

[1] As the matrix is skew-symmetric let us omit the part below the main diagonal.

Table 5. Input data

A	B	C	D	E	F	G	H	I	J	K	L	M	N
i	V_j	V_j	Тип	1	2	3	4	5	6	7	8	9	10
1	10,00000	1,70280	1	0	NA	1,2528	1,3863	1,7918	1,2528	2,1528	1,7118	1,7918	1,6247
2	10,00000	1,48430	1		0	1,3863	1,3863	1,7047	1,0968	1,8718	1,2528	1,7047	1,8718
3	10,00000	0,24275	0			0	1,5041	1,2528	1,3863	1,5041	1,7918	1,8718	1,7047
4	10,00000	0,05265	0				0	1,6094	NA	1,2528	NA	NA	NA
5	10,00000	-0,20275	0					0	NA	1,3863	NA	NA	NA
6	10,00000	-0,06675	0						0	1,2528	NA	NA	NA
7	10,00000	-1,70280	-1							0	NA	NA	NA
8	10,00000	-1,52230	-1								0	NA	NA
9	10,00000	-1,78825	-1									0	NA
10	10,00000	-1,74825	-1										0

Table 6

A	B	C	D	E	F	G	H	I	J	K	L	M	N
i	V_j	V_j	Тип	1	2	3	4	5	6	7	8	9	10
1	10,81469	0,00	1	0	0	0,5956	0,0279	-0,373	-0,354	-0,5956	-0,176	-0,3619	-0,3686
2	10,74094	0,00	1		0	0,8029	0,1016	-0,387	-0,436	-0,8029	-0,5612	-0,3753	-0,0477
3	10,15750	0,00	0			0	0,8029	-0,255	0,4365	-0,5871	0,5612	0,3753	0,3686
4	9,45626	0,00	0				0	0,8029	0	-0,1372	0	0	0
5	8,64972	0,00	0					0	0	0,8029	0	0	0
6	9,20767	0,00	0						0	0,1114	0	0	0
7	8,06628	0,00	-1							0	0	0	0
8	8,92692	0,00	-1								0	0	0
9	8,66097	0,00	-1									0	0
10	8,82140	0,00	-1										0

Table 7

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	V_i	ΔV_j	W_i	Тип V_i	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}
v_1	2,688766	0,00000	0,162651	1	0	0	0	0	0	0	0	0	$-1e^{-16}$	$-2e^{-16}$
v_2	2,229765	0,00000	0,102782	1		0	0	0	0	0	0	$-1e^{-16}$	0	0
v_3	2,768765	0,00000	0,176197	1			0	0	0	0	$-2e^{-16}$	$-1e^{-16}$	$1e^{-6}$	$1e^{-6}$
v_4	2,517465	0,00000	0,137044	1				0	0	0	$-2e^{-16}$	0	0	0
v_5	2,650965	0,00000	0,156618	1					0	0	$2e^{-16}$	0	0	0
v_6	2,517465	0,00000	0,137044	1						0	$-2e^{-16}$	0	0	0
v_7	1,264665	0,00000	0,039154	-1							0	0	0	0
v_8	0,976965	0,00000	0,029365	-1								0	0	0
v_9	0,896966	0,00000	0,027107	-1									0	0
v_{10}	1,064066	0,00000	0,032038	-1										0

As the result, the 2-nd object is better that the 7-th object $(2{,}3632)^4=31{,}1890$ times. But cell (2,7) shows the opposite, i.e. the 2-nd object is worse than the 7-th one 2.3632 times. Thus, a contradiction arises. On the one hand the 2-nd object is 31.1890 times better that the 7-th one, on the other hand its is worse 2.3632 times. But most importantly the cycle can not be improved. Attempting to modify the weights of objects in the cycle automatically increases the modulus.

Therefore, the edge (2,3) must be removed. That is because from 0.8029 to 0.1016 (edge (2,4) the reduction is the greatest.

Following the algorithm we implement the embedded cycles (external and internal). The resulting solution (column B, Table 7) is the solution of the problem for the single expert case.

In order to account for the remoteness of the lower bounds defined by the matrix $\left\|B_{ij}\right\|$ let us recover the initial configuration of the graph (Table 8).

Next, let us calculate using (10) the resultant matrix R_{ij} and remove all negative elements (Table 9).

Some edges, e.g. (1,8) can complement the reference configuration. Next, in the external cycle we define the most

Table 8

	w$_{cp}$			w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}
w_1	2,688766			0	0	-0,773	-0,745	-1,3485	-0,745	-0,3677	0,2077	0,1824	0,1206
w_2	2,229765				0	-1,232	-0,981	-1,8075	-1,204	-0,539	0,3365	-0,1713	-0,2206
w_3	2,768765					0	-1,002	-0,5754	-0,665	0,4073	0,4055	0,3677	0,4519
w_4	2,517465						0	-1,3863	0	0,5596	0	0	0
w_5	2,650965							0	0	0,47	0	0	0
w_6	2,517465								0	0,5596	0	0	0
w_7	1,264665									0	0	0	0
w_8	0,976965										0	0	0
w_9	0,896966											0	0
w_{10}	1,064066												0

Table 9

	w$_{cp}$			w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}
w_1	2,688766			0	0	0	0	0	0	0	0,2077	0,1824	0,1206
w_2	2,229765				0	0	0	0	0	0	0,3365	0	0
w_3	2,768765					0	0	0	0	0,4073	0,4055	0,3677	0,4519
w_4	2,517465						0	0	0	0,5596	0	0	0
w_5	2,650965							0	0	0,47	0	0	0
w_6	2,517465								0	0,5596	0	0	0
w_7	1,264665									0	0	0	0
w_8	0,976965										0	0	0
w_9	0,896966											0	0
w_{10}	1,064066												0

Table 10

	V$_i$	W$_i$		w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}
w_1	2,48106	0,1611740		0	0	0	0	0	0	0	$1e^{-6}$	0	0
w_2	2,02867	0,1025226			0	0	0	0	0	0	0,1354	0	0
w_3	2,56767	0,1757537				0	0	0	0	0,2062	0,2044	0,1666	0,2508
w_4	2,31637	0,1366993					0	0	0	0,3585	0	0	0
w_5	2,44987	0,1562230						0	0	0,2689	0	0	0
w_6	2,11637	0,1119200							0	0,1585	0	0	0
w_7	1,26467	0,0477550								0	0	0	0
w_8	0,97697	0,0358156									0	0	0
w_9	0,89697	0,0330620										0	0
w_{10}	1,06407	0,0390751											0

critical edge that can be dropped. For that purpose, the scale of single displacement for each weight U_i is calculated, the size of single displacement step for each edge (i, j) D_{ij} and the size of the step using the formula (11), in the internal cycle we define the virtual optimal solution using the formulas (12) and (13).

In the considered example the matrix $R_{ij}^{\text{opt.}}$ matches the connectivity matrix of the matrix $R_{ij}^{\text{ref.}}$, the single displacement U_i=1 (type i=1); U_i=0 (type i=–1).

After (1,10) and (1,9) nodes run in $R_{ij}^{\text{opt.}}$ the optimal solution takes its final form (Table 10).

The solution connects with the edge (1,8) to the lower bound and is the optimal solution deduced using the higher bound data. It is assumed that the conflict of interests is at the higher bound where each expert wants to define his/her own priorities.

References

1. Evangelos T. Multi-criteria decision making methods: a comparative study. Dordrecht: Kluwer Academic Publishers; 2000.

2. Fodor J, Roubens M. Fuzzy preference modelling and multicriteria decision support. Dordrecht: Kluwer Academic Publishers; 1994.

3. Xu ZS. Goal programming models for obtaining the priority vector of incomplete fuzzy preference relation. International Journal of Approximate Reasoning 2004;36:3:261–270.

4. Little RJA, Rubin DB. Statistical analysis with missing data. Moscow: Financy i statistika; 1991.

5. Garcia-Laencina PJ, Sanco-Gomez J-L, Figueiras-Vidal AR. Pattern classification with missing data: a review. London: Springer-Verlag Limited; 2009.

6. Schafer JL, Graham JW. Missing data: Our view to the state of the art. Psychological methods 2002;7(2):147–177.

7. Millet I. The effectiveness of alternative preference elicitation methods in the analytic hierarchy process. J. Multi-Criteria Decis. Anal. 1997l;6(1):41–51.

8. Carmone FJ, Kara Jr A, Zanakis SH. A Monte Carlo investigation of incomplete pairwise comparison matrices in AHP. Eur. J. Oper. Res. 2001;102(3):533–553.

9. Ebenbach DH, Moore CF. Incomplete information, inferences, and individual differences: The case of environmental judgements. Org. Behav. Human Decis. Process. 2000;81(1):1–27.

10. Alonso S, Cabrerizo FJ, Chiclana F, Herrera F, Herrera-Viedma E. An interactive decision support system based on consistency criteria. J. Mult.-Valued Log. Soft Comput. 2008;14(3–5):371–386.

11. Kim JK, Choi SH. A utility range-based interactive group support system for multiattribute decision making. Comput. Oper. Res. 2001;28(5):485–503.

12. Kim JK, Choi SH, Han CH, Kim SH. An interactive procedure for multiple criteria group decision making with incomplete information. Comput. Ind. Eng. 1998;35(1/2):295–298.

13. Chiclana F, Herrera-Viedma E, Alonso S. A Note on Two Methods for Estimating Missing Pairwise Preference Values. IEEE Transactions On Systems, MAN, and Cybernetics – Part B: Cybernetics 2009;39(6):1628-1633.

14. Karlov IA. Vosstanovlenie propushchennukh dannykh pri chislennom modelirovanii slozhnykh dynamicheskikh system [Recovery of missing data in computational modeling of complex dynamic systems]. SPbSPU Journal. Computer Science. Telecommunication and Control Systems 2013;6(186):137–144 [in Russian].

15. Saaty T. Decision making. Analytic hierarchy method. Moscow: Radio i sviaz; 1993.

16. Ashmanov SA. Matematicheskie modeli i metody [Mathematical models and methods]. Moscow: Moscow State University publishing, 1980 [in Russian].

About the authors

Alexander V. Bochkov, Candidate of Engineering, Deputy Director of the Risk Analysis Center, Research and Design Institute of Economy and Business Administration in the Gas Industry, Moscow, Russia, phone: +7 (916) 234 40 32, e-mail: a.bochkov@gmail.com

Nikolai N. Zhigirev, Candidate of Engineering, Chief Researcher, Research and Design Institute of Economy and Business Administration in the Gas Industry, Moscow, Russia, phone: +7 (985) 782 47 16, e-mail: nnzhigirev@mail.ru

Alexandra N. Ridley, Postgraduate Student, MAI (NRU), Moscow, Russia, phone: +7 (929) 970 59 69, e-mail: alexandra.ridley@yandex.ru

Received on 18.01.2017

Original article
DOI: 10.21683/1729-2646-2017-17-3-49-57

Use of automatic signalling system for reduction of the risk of transportation incidents in railway stations[1]

Igor B. Shubinsky, *ZAO IBTrans, Moscow, Russia*
Aleksei M. Zamyshliaev, *JSC NIIAS, Moscow, Russia*
Aleksei N. Ignatov, *Moscow Aviation Institute, Moscow, Russia*
Andrei I. Kibzun, *Moscow Aviation Institute, Moscow, Russia*
Evgeni N. Platonov, *Moscow Aviation Institute, Moscow, Russia*

Abstract. Aim. *Evaluating the risk of collision between trains during shunting operations in railway stations. Risk is the combination of the probability and consequences of an event. The most complicated task related to risk assessment is the choice of the evaluation model for the probability of an undesired event. The model must ensure practical applicability of the results. In the context of railway facilities the construction of analytical models of probability evaluation is of principal interest due to the possibility to demonstrate the factors that are taken into consideration by the model. The main purpose of this paper is to examine the extent to which the Shunting Automatic Cab Signalling System (MALS) contributes to the probability of side collision of trains involving shunting engines in railway stations. The main function of the Shunting Automatic Cab Signalling System is to ensure that shunting engines do not pass signals at danger in stations.* **Methods.** *Methods of the probability theory and theory of random processes, addition, multiplication formulas, composite probability, properties of Poisson flows. In [2] a method is suggested for calculating the probability of collision as the result of shunting or train locomotive passing a signal at danger. The development of the method was based on the main assumption that the flow of shunting consists for each switch is a Poisson flow. This paper suggests a modification of this method that takes into consideration the possible use of the MALS system with shunting engines. The input data for the algorithm of calculation of the collision probability are the station topology, passenger train schedule and their possible routes through the station, average train lengths and speeds, as well as the frequency of shunting consists passing over switches.*

Results. *An algorithm has been developed for calculation of the probability of train-to-train collision involving shunting engines within a random time period. For different operating modes, e.g. pulling up, coupling, formulas are shown for calculation of the probability of collision with a passenger or freight train on a random switch. The algorithms consists in the following: 1) a time period is specified for which it is required to calculate the probability of collision; 2) passenger train timetable is designed using data from ASU “Express”; 3) overall number of passenger trains passing through the station within the specified time period is calculated; 4) passenger trains are renumbered according to the order of their arrival to the station; 5) probability of signal violation by shunting engine driver is calculated; 6) probability of violation of traffic safety by shunting engine driver in the “pull up” mode is calculated; 7) probability of violation of traffic safety by the shunting engine driver after coupling with the “coupling” mode off is calculated; 8) overall number of possible routes for each train is calculated; 9) for each train the frequency of one or another route is identified; 10) for each switch of each route a number is specified in the order of appearance; 11) probability that each passenger train on each route has at least one collision is calculated; 12) probability of at least one collision of each passenger train moving through the station is calculated; 13) probability of at least one collision in the station within the specified period of time is calculated. The paper considers the example of calculation of collision probability for an individual train route and the station as a whole within a month and a year. It shows that the use of MALS helps significantly reduce the probability of side collisions in railway stations.*

Keywords: *railway transportation, traffic safety, shunting operations, probability of collision.*

For citation*: Shubinsky IB, Zamyshliaev AM, Ignatov AN, Kibzun AI, Platonov EN. Use of automatic signalling system for reduction of the risk of transportation incidents in railway stations. Dependability 2017; 3: p. 49-57. DOI: 10.21683/1729-2646-2017-17-3-49-57.*

[1] Work performed with the support of the Russian Science Foundation (project no. 16-11-00062).

1. Introduction

Both passenger and freight traffic of certain density may be affected by various adverse events that may cause reputational and material costs to JSC RZD associated, for instance, with damage to infrastructure and/or rolling stock. Therefore, in order to identify quantitative characteristics of the hazard of one or another event the concept of risk is introduced that according to [1] is the functional of the probability and damage from the adverse event.

As noted in [2], during shunting operations in stations signal violations by one of the vehicles may cause not only collisions between two shunting consists, but particularly between a passenger train and a shunting consists, which as it is shown in [3] may cause loss of life and consequently significant damage. Despite the fact that in case of a single passenger train passing through a station the probability of its collision with a shunting consist is quite low, the probability of at least one collision per year per all the trains passing through a station is a significant value, especially if the stations feature intense shunting operations. The installation and use of the MALS system may help reduce the probability of collision between trains.

The installation of MALS on shunting engines enables several movement modes. The first one is the "typical" mode, in which MALS operates normally. The second one is the "pull up" mode, in which MALS is disabled and the shunting engine driver can drive the train closer than 20 meters to a restrictive signal and consequently may pass the restrictive signal, which may cause a collision between a shunting consist and a passenger or freight train. The third one is the "movement after coupling with the "coupling" mode off, in which when moving with the wagons in front of the engine MALS incorrectly determines the location of the head and tail of the consist, which may also cause signal violation. The second and third modes are abnormal. Obviously, most of the time a shunting engine is in the "typical" mode, yet in the other two modes the probability of signal violation is much higher, therefore each of those cases is characterized by its own probability of collision between a shunting consist with a passenger/freight train and is individually examined in this paper.

2. An algorithm for calculation of the probability of collision within a random time period if the MALS system is used

Let us introduce the following designations:

N is the total number of switches in the station;

L is the number of shunting engines;

N_l is the number of switches passed by shunting consist per hour, l=1,…L;

r_l is the number of half-runs, l=1,…L;

s_l is the number of couplings with the "coupling" mode off, l=1,…L;

T_l is the number of pull ups, l=1,…L;

l_{sh} is the average length of shunting consist;

l_{p} is the average length of passenger train;

v_{p} is the average speed of passenger train movement through the station;

v_{sh} is the average speed of shunting consist movement through the station;

v_{pu} is the average speed of shunting consist movement through the station in the "pull up" mode;

P_{p} is the probability of passenger train violating a restrictive signal;

$P_{\text{sh(one)}}$ is the probability of violation of restrictive signal by shunting engine driver while operating without an assistant driver;

$P_{\text{sh(two)}}$ is the probability of violation of restrictive signal by shunting engine driver while operating with an assistant driver;

P_{two} is the probability of shunting engine being manned with driver and assistant driver;

P_{Wsh} is the probability of coupling with subsequent movement of shunting engine with wagons;

$P_{\text{O}}(s)$ is the probability of failure by the station duty officer to prevent a violation of restrictive signal in the "pull up" mode;

P_{PUE} is the probability of signal violation by a shunting engine driver in the "pull up" mode when the shunting engine is at the head of the train;

P_{PUT} is the probability of signal violation by a shunting engine driver in the "pull up" mode when the shunting engine is at the tail of the train;

P_{ShM} is the probability of traffic safety violation by the shunting master;

λ_{S}^{i} is the frequency of shunting consists stopping on switches that did not violate traffic safety while crossing the switch, i=1,…,N;

τ_{S}^{i} is the average time a shunting consist dwells on a switch that did not violate traffic safety while crossing the switch after having stopped on the switch i=1,…,N;

τ_{PU} is the average time it takes a shunting consist to clear a switch after entering it in the "pull up" mode after having stopped of the switch.

1. A time period T is specified, for which it is required to calculate the probability of collision. Passenger train timetable is retrieved from ASU Express. According to the timetable all passenger trains that moving trough the station within the time T are identified.

2. I is identified, i.e. the total number of passenger trains that move through the station within the time T.

3. Passenger trains moving through the station within the considered time are numbered according to the time of their arrival to the station. Passenger trains are renumbered according to the order of their arrival to the station, i.e. the first arrived train is numbered 1, the second is numbered 2 and so on.

4. The probability of signal violation by shunting engine driver is calculated using the formula

$$P_{\text{Sh}} = P_{\text{TWO}} \cdot P_{\text{SH(TWO)}} + \left(1 - P_{\text{TWO}}\right) \cdot P_{\text{Sh(TWO)}},$$

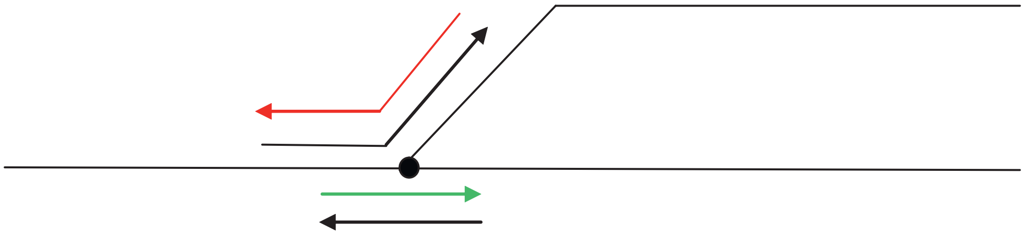


Figure 1. Movement direction (red arrow) of shunting consist across a switch (circle) that makes possible a collision with a passenger train moving in the direction shown with the green arrow

the probability of traffic safety violation by shunting engine driver in the "pull up" mode is calculated using the formula

$$P_{\text{PU}} = \frac{1}{2} P_{\text{O}}(s)(P_{\text{PUE}} + P_{\text{PUT}}),$$

the probability of traffic safety violation by shunting engine driver after coupling with the "coupling" mode off is calculated using the formula

$$P_{\text{Cu}} = \left(1 - \frac{1}{2} P_{\text{WSh}}\right) P_{\text{Sh}} + \frac{1}{2} P_{\text{Sh}} P_{\text{ShM}}.$$

The following is calculated for each switch

$\tilde{\lambda}_{\text{PU}}$, frequency of shunting consist in "pull up" mode being before a switch

$$\tilde{\lambda}_{\text{PU}} = \sum_{l=1}^{L} \frac{1}{N} \cdot \frac{T_l}{24};$$

$\tilde{\lambda}_{\text{Cu}}$, the frequency of a switch being crossed by a shunting consist after coupling with the "coupling" mode off

$$\tilde{\lambda}_{\text{Cu}} = \sum_{l=1}^{L} \frac{N_l}{N} \cdot \frac{s_l}{r_l};$$

$\tilde{\lambda}_{\text{Sh}}$, the frequency of a switch being crossed by a shunting consist in "normal" mode

$$\tilde{\lambda}_{\text{Sh}} = \sum_{l=1}^{L} \frac{N_l}{N} - \tilde{\lambda}_{\text{PU}} P_{\text{PU}} - \tilde{\lambda}_{\text{Cu}}.$$

The frequency[1] λ^1_{PU} of a switch being possibly crossed by a shunting consist in "pull up" mode in a certain direction 1 is calculated using formula $\lambda^1_{\text{PU}} = \tilde{\lambda}_{\text{PU}} / 4$.

The frequency λ^1_{Cu} of a switch being crossed by a shunting consist after coupling with the "coupling" mode off in a certain direction 1 is calculated using the formula $\lambda^1_{\text{Cu}} = \tilde{\lambda}_{\text{Cu}} / 4$.

The frequency λ^1_{Sh} of a switch being crossed by a shunting consist in the "normal" mode in a certain direction 1 is calculated using formula $\lambda^1_{\text{Sh}} = \tilde{\lambda}_{\text{Sh}} / 4$.

5. The value i is assumed to equal to one.

6. We calculate K, the total number of possible routes for the i-th train.

7. The value k is assumed to equal to one.

8. For the i-th train the frequency of using one or another routes is calculated[2]

$$P(R_k) = \frac{m_{R_k}}{n},$$

where m_{R_k} is the number of trains with the number i that took the route R_k, while n is the total number of passenger trains with the number i, that moved across the station over the observation period.

9. When the route R_k is processed, the included switches are numbered in the order the ith train passes them, i.e. the first switch passed by a passenger train is numbered 1, the second switch passed by the passenger train is numbered 2 and so on. Let m be the total number of switches that the i-th train crosses on the route R_k.

10. The value j is assumed to equal to one.

11. If the j-th switch of the route R_k of the i-th passenger train is uninsulated[3], the following are identified:

λ_{Sh} is the frequency of shunting consist crossing a switch in "normal" mode in a direction that makes a collision possible (selected out of λ^1_{Sh} obtained at step 4, see Figure 1);

λ_{Cu} is the frequency of passing by a shunting consist with a shunting engine that performed the coupling with the "coupling" mode off of a switch in a direction that makes a collision possible (selected from λ^1_{Cu} obtained at step 4, see Figure 1);

λ_{PU} is the frequency of possible passing by a shunting consist with a shunting engine that was before the switch in the "pull up" mode of a switch in a direction that makes a collision possible (selected from λ^1_{PU} obtained at step 4, see Figure 1);

λ_{S} is the frequency of shunting consists stopping on the j-th switch

τ_{S} is the average time a shunting consist dwells on the j-th switch if it stops on it;

P_{PS} is the probability of the i-th passenger train stopping on the j-th switch;

τ_{PS} is the average time the i-th passenger train dwells on the j-th switch;

[1] The frequencies can be specified during data collection for each switch in the considered station.

[2] If there is no data on the previous transits of trains across the station, all routes are assumed to be equally possible, i.e. for any k the probability $P(R_k)$ of using the route R_k is calculated using formula $P(R_k) = 1/K$.

[3] An insulated switch is a switch, on which a collision caused by signal violation is impossible, while an uninsulated switch is one on which a collision is possible.

the probability of collision through a fault of the shunting engine driver (in case of MALS failure) who violated a signal

$$P_{Sh}(A_{\mathrm{Sh}})=\lambda_{\mathrm{Sh}}P_{\mathrm{Sh}}\left(\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}+\frac{l_{\mathrm{Sh}}}{v_{\mathrm{Sh}}}+P_{\mathrm{PS}}\tau_{\mathrm{PS\,пс}}\right);$$

the probability of collision through a fault of the passenger train driver who violated a signal

$$P_{Sh}(A_{\mathrm{P}})=P_{\mathrm{P}}\lambda_{\mathrm{Sh}}\left(\frac{l_{\mathrm{Sh}}}{v_{\mathrm{Sh}}}+\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}\right)+P_{\mathrm{P}}\lambda_{\mathrm{Cu}}\left(\frac{l_{\mathrm{Sh}}}{v_{\mathrm{Sh}}}+\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}\right);$$

the probability of collision through a fault of the passenger train driver who violated a signal and allowed the train's collision with wagons dwelling on a switch

$$P_{Sh}(A_{\mathrm{S}})=P_{\mathrm{P}}\lambda_{\mathrm{S}}\tau_{\mathrm{S}};$$

the probability of collision through a fault of the shunting engine driver who violated a signal in the "pull up" mode

$$P_{Sh}(A_{\mathrm{PU}})=\lambda_{\mathrm{PU}}P_{\mathrm{PU}}\left(\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}+\frac{l_{\mathrm{PU}}}{v_{\mathrm{PU}}}+P_{\mathrm{PS}}\tau_{\mathrm{PS}}+\tau_{\mathrm{PU}}\right);$$

the probability of collision caused by signal violation by both the shunting engine and passenger train drivers

$$P_{Sh}(A_{\mathrm{PUP}})=\lambda_{\mathrm{PU}}P_{\mathrm{PU}}P_{\mathrm{PU}}\left(\frac{l_{\mathrm{PU}}}{v_{\mathrm{PU}}}+\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}\right)+$$
$$+\lambda_{\mathrm{Sh}}P_{\mathrm{Sh}}P_{\mathrm{P}}\left(\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}+\frac{l_{\mathrm{Sh}}}{v_{\mathrm{Sh}}}\right)+\lambda_{\mathrm{Cu}}P_{\mathrm{Cu}}P_{\mathrm{P}}\left(\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}+\frac{l_{\mathrm{Sh}}}{v_{\mathrm{Sh}}}\right);$$

the probability of collision through a fault of the shunting engine driver after coupling with the "coupling" mode off

$$P_{Sh}(A_{\mathrm{Cu}})=\lambda_{\mathrm{Cu}}P_{\mathrm{Cu}}\left(\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}+\frac{l_{\mathrm{Sh}}}{v_{\mathrm{Sh}}}+P_{\mathrm{PS}}\tau_{\mathrm{PS}}\right).$$

The resultant probability of collision on the *j*-th switch is calculated using the formula

$$P_{Sh}(A_{k:j})=\left(\lambda_{\mathrm{Sh}}\left(\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}+\frac{l_{\mathrm{Sh}}}{v_{\mathrm{Sh}}}\right)(P_{\mathrm{Sh}}(1+P_{\mathrm{P}})+P_{\mathrm{P}})+\right.$$
$$+\lambda_{\mathrm{Cu}}\left(\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}+\frac{l_{\mathrm{Sh}}}{v_{\mathrm{Sh}}}\right)(P_{\mathrm{Cu}}(1+P_{\mathrm{P}})+P_{\mathrm{P}})+$$
$$+\lambda_{\mathrm{PU}}\left(\frac{l_{\mathrm{PU}}}{v_{\mathrm{PU}}}+\frac{l_{\mathrm{P}}}{v_{\mathrm{P}}}+\tau_{\mathrm{PU}}\right)P_{\mathrm{PU}}(1+P_{\mathrm{P}})+\lambda_{\mathrm{S}}P_{\mathrm{P}}\tau_{\mathrm{S}}+$$
$$\left.+\left(\lambda_{\mathrm{Sh}}P_{\mathrm{Sh}}+\lambda_{\mathrm{PU}}P_{\mathrm{PU}}+\lambda_{\mathrm{Cu}}P_{\mathrm{Cu}}\right)P_{\mathrm{PS}}\tau_{\mathrm{PS}}\right)\cdot k_{\mathrm{S}},$$

where k_{S} takes on the value 1 if the switch is uninsulated, and 0 if it is insulated, while the other variables in the formula are defined at steps 1 and 4.

12. If *j*=*m*, proceed to step 14, otherwise proceed to step 13.

13. *j*:=*j*+1, proceed to step 11.

14. We calculate the probability that the *i*-th passenger train on route R_k has at least one collision that amounts to

$$P_{Sh}(A_i\mid R_k)=1-\prod_{j=1}^{m}(1-P_{Sh}(A_{k:j})),$$

Table 1. Data unit describing the station topology and shunting operations in it

Name	Notation	Number	Measurement units
Total number of switches in the station	N	102	pcs
Number of shunting engines	L	2	pcs
Number of switches passed by shunting consist per hour with engine no. 1	N_1	36	pcs
Number of switches passed by shunting consist per hour with engine no. 2	N_2	36	pcs
Average length of shunting consist	l_{Sh}	0,2	km
Average speed of shunting consist movement through the station	v_{Sh}	4,2	km/h
Probability of violation of restrictive signal by shunting engine driver while operating without an assistant driver	$P_{\mathrm{Sh(one)}}$	$2{,}1\cdot10^{-8}$	
Probability of violation of restrictive signal by shunting engine driver while operating with an assistant driver	$P_{\mathrm{Sh(two)}}$	$7\cdot10^{-9}$	
Probability of shunting engine being manned with driver and assistant driver	P_{two}	0,8	
Frequency of shunting consists stopping on switches that did not violate traffic safety while crossing the switch	λ_{S}^{1}, ..., $\lambda_{\mathrm{S}}^{102}$	0, ... 0	1/h, ..., 1/h
Average time a shunting consist dwells on a switch that did not violate traffic safety while crossing the switch after having stopped on the switch	τ_{S}^{1}, ..., τ_{S}^{102}	0, ... 0	h, ..., h

where the probabilities $P_{Sh}(A_{k,j})$ were calculated at step 11, the value m was calculated at step 9.

15. If $k=K$, proceed to step 17, otherwise proceed to step 16.

16. $k:=k+1$, proceed to step 8.

17. We calculate the probability of at least one collision involving passenger train no. i while moving through the station

$$P_{Sh}(A_i) = \sum_{k=1}^{K} P_{Sh}(A_i \mid R_k) P(R_k),$$

where the probabilities $P_{Sh}(A_i|R_k)$ were calculated at step 14, the probabilities $P(R_k)$ were calculated at step 8, the value K was calculated at step 6.

18. If $i=I$, proceed to step 20, otherwise proceed to step 19.

19. $i:=i+1$, proceed to step 6.

20. We calculate the probability of at least one collision within the time period T

$$P_{Sh}(A_{PU}) = P_{Sh}(A_1 + A_2 + \ldots + A_I) = 1 - \prod_{i=1}^{I}(1 - P_{Sh}(A_i)).$$

3. An example of calculation of the probability of collision between a passenger train and a shunting consist for some stations equipped with the MALS system

Let there be 4 possible movement directions from and to the station (station layout is shown in Figure 2): northeast, southeast, northwest, southwest. From the northeast trains arrive to the station via track F, from the southeast the trains arrive to the station via track D, from the northwest the trains arrive to the station via track B, from the southwest the trains arrive to the station via track B or track C. From the station, the northwest is accessed via track F, the southeast is accessed via track E, the northwest is accessed via track A, the southwest is accessed via track B or C. Let the commuter trains be able to stop only on tracks 1, 3, 7, 8, 11, 12, 13, while the long distance passenger and freight trains be able to stop or move through the station only via tracks 1, 3, 4, 5, 6, 7, 8. Let us assume that we have the train timetable for May and August. Let us fill in the table with data required for the application of probability calculation algorithm.

Let us consider passenger train no. 255N that passes through the station without stopping from northeast to northwest. It is known that this train will be directed to the first track. In this case it only has 6 routes through the station.

Let us calculate the probability of signal violation by a shunting consist, that depends on the probability P_{two} of the driver and assistant driver being onboard the engine, as well as the probabilities $P_{\text{Sh(two)}}$ and $P_{\text{Sh(one)}}$ that the shunting

Table 2. Additional data unit describing the station topology and shunting operations in it

Name	Notation	Number	Measurement units
Number of half-runs performed by engine no. 1	r_1	20	pcs
Number of half-runs performed by engine no. 2	r_2	22	pcs
Number of couplings with the «coupling» mode off performed by engine no. 1	s_1	2	pcs
Number of couplings with the «coupling» mode off performed by engine no. 2	s_2	0	pcs
Number of pull ups performed by engine no. 1	T_1	3	pcs
Number of pull ups performed by engine no. 2	T_2	3	pcs
Length of shunting engine (wagon)	l_{PU}	0,2	km
Average speed of shunting consist movement through the station in the «pull up» mode	v_{PU}	2	km/h
Average time it takes a shunting consist to clear a switch after entering it in the «pull up» mode after having stopped of the switch	τ_{PU}	0,01	h
Probability of coupling with subsequent movement of shunting engine with wagons	P_{WSh}	0,25	
Probability of failure by the station duty officer to prevent a violation of restrictive signal in the «pull up» mode	$P_{\text{O}}^{1}(s)$	10^{-2}	
Probability of signal violation by a shunting engine driver in the «pull up» mode when the shunting engine is at the head of the train	P_{PUE}	10^{-4}	
Probability of signal violation by a shunting engine driver in the «pull up» mode when the shunting engine is at the tail of the train	P_{PUT}	10^{-3}	
Probability of traffic safety violation by the shunting master	P_{ShM}	10^{-3}	

consist crosses a switch at danger, while the driver and assistant driver or only the driver is in the cab (see item 4 of the algorithm)

$$P_{\text{Sh}} = P_{\text{two}} \cdot P_{\text{Sh(two)}} + \left(1 - P_{\text{two}}\right)\cdot P_{\text{Sh(one)}} =$$
$$= 0{,}8 \cdot 7 \cdot 10^{-9} + \left(1 - 0{,}8\right)\cdot 2{,}1 \cdot 10^{-8} \approx 10^{-8}.$$

The probability of traffic safety violation by a shunting engine driver in the “pull up” mode equals to (see item 4 of the algorithm)

$$P_{\text{PU}} = \frac{1}{2} P_{\text{O}}(s)(P_{\text{PUE}} + P_{\text{PUT}}) =$$
$$= \frac{1}{2} \cdot 10^{-2} \cdot (10^{-4} + 10^{-3}) = 5{,}5 \cdot 10^{-6},$$

the probability of traffic safety violation by a shunting engine driver after coupling with the “coupling” mode off equals to (see item 4 of the algorithm)

$$P_{\text{CU}} = \left(1 - \frac{1}{2} P_{\text{WSh}}\right) P_{\text{Sh}} + \frac{1}{2} P_{\text{WSh}} P_{\text{ShM}} =$$
$$= \left(1 - \frac{1}{2} \cdot \frac{1}{4}\right) \cdot 10^{-8} + \frac{1}{2} \cdot \frac{1}{4} \cdot 10^{-3} \approx 1{,}25 \cdot 10^{-4}.$$

Frequency of shunting consist in “pull up” mode being before a random switch equals to (see item 4 of the algorithm)

$$\tilde{\lambda}_{\text{PU}} = \sum_{l=1}^{L} \frac{1}{N} \cdot \frac{T_l}{24} = \frac{1}{102} \cdot \frac{3}{24} + \frac{1}{102} \cdot \frac{3}{24} = 0{,}0025 \text{ (1/h)},$$

the frequency of random switch being crossed by a shunting consist after coupling with the “coupling” mode off equals to (see item 4 of the algorithm)

$$\tilde{\lambda}_{\text{CU}} = \sum_{l=1}^{L} \frac{N_l}{N} \cdot \frac{s_l}{r_l} = \frac{36}{102} \cdot \frac{2}{20} + \frac{36}{102} \cdot \frac{0}{22} = 0{,}035,$$

the frequency of a switch being crossed by a shunting consist in “normal” mode equals to (see item 4 of the algorithm)

$$\tilde{\lambda}_{\text{Sh}} == \sum_{l=1}^{L} \frac{N_l}{N} - \tilde{\lambda}_{\text{PU}} P_{\text{PU}} - \tilde{\lambda}_{\text{CU}} = \frac{36}{102} + \frac{36}{102} -$$
$$-0{,}0025 \cdot 5{,}5 \cdot 10^{-6} - 0{,}035 = 0{,}671.$$

Let us calculate the probability of at least one collision on route R_1. The probability $P(R_1)$ of the use of a route R_1 equals to (see item 8 if the algorithm)

$$P(R_1) = \frac{2}{2+1+0+0+0+0} = \frac{2}{3}.$$

Let us number the switches according to the order in which they are crossed by the train (see item 9 of the algorithm): 115→1, 121→2, 151-147→3, 149-161→4, 244→5, 238→6, 236→7, 174→8, 164→9, 154→10, 144→11, 138→12.

For switches nos. 144, 236, 149-161, 151-147 we have $\lambda_{\text{Sh}} = \tilde{\lambda}_{\text{Sh}} / 4 \approx 0{,}168$, $\lambda_{\text{PU}} = \tilde{\lambda}_{\text{PU}} / 4 \approx 0{,}0006$, $\lambda_{\text{CU}} = \tilde{\lambda}_{\text{CU}} / 4 \approx$ $\approx 0{,}009$, where division by 4 is performed as each shunting consist can cross a switch in 4 possible directions (see Figure 1).

The probability of collision on non-insulated switches nos. 144, 236, 149-161, 151-147 is calculated using the formulas (see item 11 of the algorithm)

$$P_{Sh}(A_{1:3}) = \lambda_{\text{Sh}} \left(\frac{l_{\text{P}}}{v_{\text{P}}} + \frac{l_{\text{Sh}}}{v_{\text{Sh}}}\right)(P_{\text{Sh}}(1+P_{\text{P}}) + P_{\text{P}}) +$$
$$+\lambda_{\text{CU}} \left(\frac{l_{\text{P}}}{v_{\text{P}}} + \frac{l_{\text{Sh}}}{v_{\text{Sh}}}\right)(P_{\text{CU}}(1+P_{\text{P}}) + P_{\text{P}}) +$$
$$+\lambda_{\text{PU}} \left(\frac{l_{\text{P}}}{v_{\text{P}}} + \frac{l_{\text{PU}}}{v_{\text{PU}}} + \tau_{\text{PU}}\right) P_{\text{PU}}(1+P_{\text{P}}) + \lambda_{\text{S}} P_{\text{P}} \tau_{\text{S}} +$$
$$+\left(\lambda_{\text{Sh}} P_{\text{Sh}} + \lambda_{\text{PU}} P_{\text{PU}} + \lambda_{\text{CU}} P_{\text{CU}}\right) P_{\text{PS}} \tau_{\text{PS}} =$$
$$= 0{,}168 \cdot \left(\frac{0{,}48}{42} + \frac{0{,}2}{4{,}2}\right) \cdot (10^{-8} \cdot (1+10^{-7}) + 10^{-7}) +$$
$$+0{,}009 \cdot \left(\frac{0{,}48}{42} + \frac{0{,}2}{4{,}2}\right) \cdot (1{,}25 \cdot 10^{-4} \cdot (1+10^{-7}) + 10^{-7}) +$$
$$+0{,}0006 \cdot \left(\frac{0{,}48}{42} + \frac{0{,}02}{2} + 0{,}01\right) \cdot 5{,}5 \cdot 10^{-6} \cdot (1+10^{-7}) = 6{,}8 \cdot 10^{-8}.$$

$$P_{Sh}(A_{1:4}) = P_{Sh}(A_{1:7}) = P_{Sh}(A_{1:11}) = 0{,}168 \cdot \left(\frac{0{,}48}{42} + \frac{0{,}2}{4{,}2}\right) \cdot$$
$$\cdot (10^{-8} \cdot (1+10^{-7}) + 10^{-7}) + 0{,}009 \cdot \left(\frac{0{,}48}{42} + \frac{0{,}2}{4{,}2}\right) \cdot$$
$$\cdot (1{,}25 \cdot 10^{-4} \cdot (1+10^{-7}) + 10^{-7}) + 0{,}0006 \cdot$$
$$\cdot \left(\frac{0{,}48}{42} + \frac{0{,}02}{2} + 0{,}01\right) \cdot 5{,}5 \cdot 10^{-6} \cdot (1+10^{-7}) = 6{,}8 \cdot 10^{-8}.$$

For the insulated switches nos. 138, 154, 164, 174, 238, 244, 121, 115 the probability of collision equals to zero, i.e.

$$P_{Sh}(A_{1:1}) = P_{Sh}(A_{1:2}) = P_{Sh}(A_{1:5}) = P_{Sh}(A_{1:6}) =$$
$$= P_{Sh}(A_{1:8}) = P_{Sh}(A_{1:9}) = P_{Sh}(A_{1:10}) = P_{Sh}(A_{1:12}) = 0.$$

The probability of at least one collision involving passenger train no. 255N on route R_1 equals to (see item 14 of the algorithm)

$$P_{Sh}(A_1 \mid R_1) = 1 - 1 \cdot 1 \cdot (1 - 6{,}8 \cdot 10^{-8}) \cdot (1 - 6{,}8 \cdot 10^{-8}) \cdot$$
$$\cdot 1 \cdot 1 \cdot (1 - 6{,}8 \cdot 10^{-8}) \cdot 1 \cdot 1 \cdot 1 \cdot (1 - 6{,}8 \cdot 10^{-8}) \cdot 1 = 2{,}7 \cdot 10^{-7}.$$

Let us calculate the probability of at least one collision on route R_2. The probability $P(R_2)$ of the use of a route R_2 equals to (see item 8 if the algorithm)

$$P(R_2) = \frac{1}{2+1+0+0+0+0} = \frac{1}{3},$$

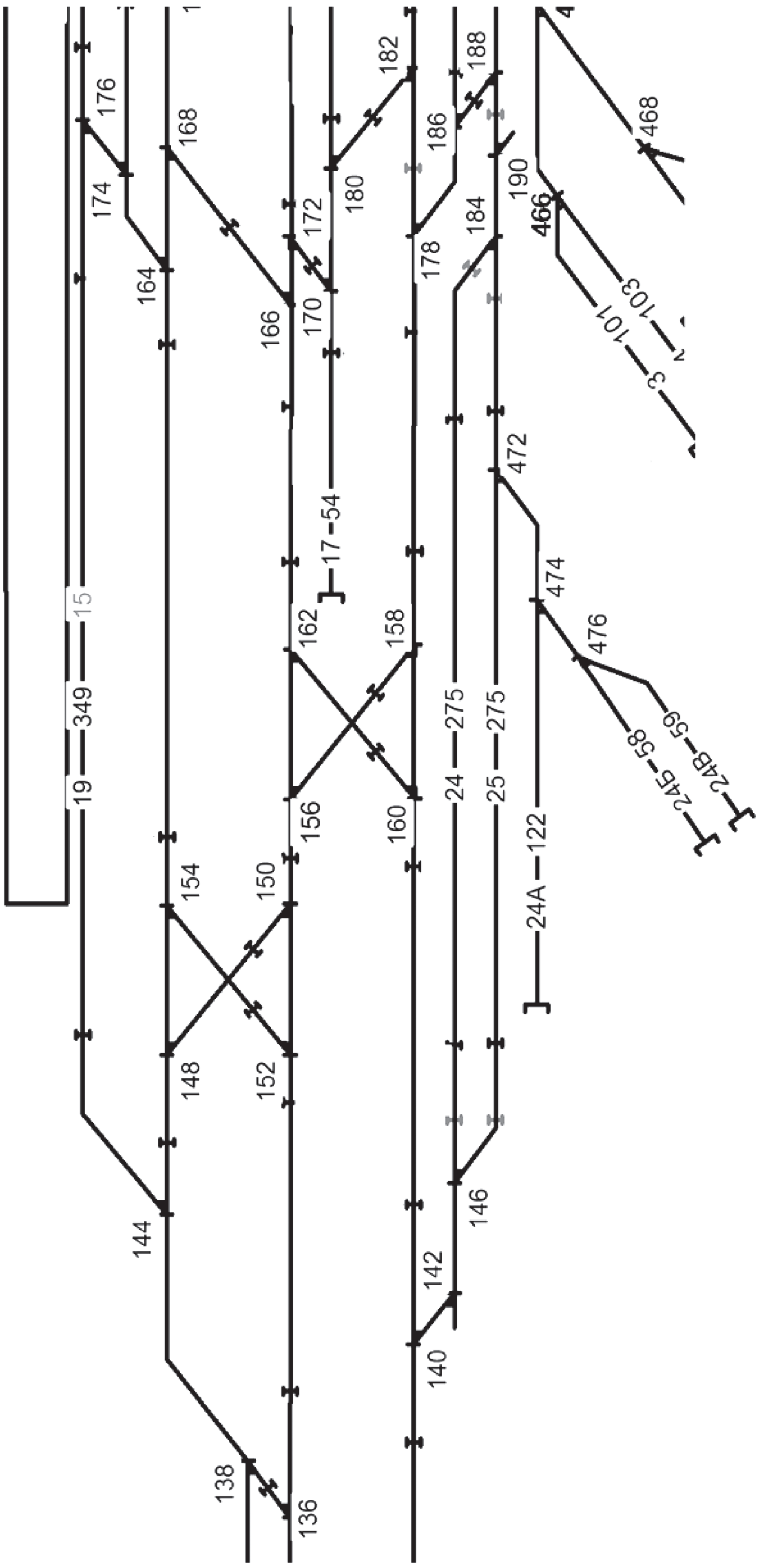


Figure 2. Layout of the station for which the collision probability is calculated

For switches nos. 144, 236, 149-161, 151-147 we have $\lambda_{\text{Sh}}=\tilde{\lambda}_{\text{Sh}}/4\approx 0{,}168$, $\lambda_{\text{PU}}=\tilde{\lambda}_{\text{PU}}/4\approx 0{,}0006$, $\lambda_{\text{CU}}=\tilde{\lambda}_{\text{CU}}/4\approx$ $\approx 0{,}009$.

Let us number the switches according to the order in which they are crossed by the train (see item 9 of the algorithm): 115→1, 121→2, 151-147→3, 149-161→4, 244→5, 238→6, 236→7, 176→8, 144→9, 138→10.

The probability of collision on non-insulated switches nos. 144, 236, 149-161, 151-147 is calculated using the formulas (see item 11 of the algorithm)

$$P_{Sh}(A_{2:3})=\lambda_{\text{Sh}}\left(\frac{l_{\text{P}}}{v_{\text{P}}}+\frac{l_{\text{Sh}}}{v_{\text{Sh}}}\right)(P_{\text{Sh}}(1+P_{\text{P}})+P_{\text{P}})+$$
$$+\lambda_{\text{CU}}\left(\frac{l_{\text{P}}}{v_{\text{P}}}+\frac{l_{\text{Sh}}}{v_{\text{Sh}}}\right)(P_{\text{CU}}(1+P_{\text{P}})+P_{\text{P}})+$$
$$+\lambda_{\text{PU}}\left(\frac{l_{\text{P}}}{v_{\text{P}}}+\frac{l_{\text{PU}}}{v_{\text{PU}}}+\tau_{\text{PU}}\right)P_{\text{PU}}(1+P_{\text{P}})+\lambda_{\text{S}}P_{\text{P}}\tau_{\text{S}}+$$
$$+\left(\lambda_{\text{Sh}}P_{\text{Sh}}+\lambda_{\text{PU}}P_{\text{PU}}+\lambda_{\text{CU}}P_{\text{CU}}\right)P_{\text{PS}}\tau_{\text{PS}}=0{,}168\cdot$$
$$\cdot\left(\frac{0{,}48}{42}+\frac{0{,}2}{4{,}2}\right)\cdot(10^{-8}\cdot(1+10^{-7})+10^{-7})+0{,}009\cdot$$
$$\cdot\left(\frac{0{,}48}{42}+\frac{0{,}2}{4{,}2}\right)\cdot(1{,}25\cdot10^{-4}\cdot(1+10^{-7})+10^{-7})+0{,}0006\cdot$$
$$\cdot\left(\frac{0{,}48}{42}+\frac{0{,}02}{2}+0{,}01\right)\cdot5{,}5\cdot10^{-6}\cdot(1+10^{-7})=6{,}8\cdot10^{-8}.$$

$$P_{Sh}(A_{2:4})=P_{Sh}(A_{2:7})=P_{Sh}(A_{2:9})=0{,}168\cdot\left(\frac{0{,}48}{42}+\frac{0{,}2}{4{,}2}\right)\cdot$$
$$\cdot(10^{-8}\cdot(1+10^{-7})+10^{-7})+0{,}009\cdot\left(\frac{0{,}48}{42}+\frac{0{,}2}{4{,}2}\right)\cdot$$
$$\cdot(1{,}25\cdot10^{-4}\cdot(1+10^{-7})+10^{-7})+0{,}0006\cdot\left(\frac{0{,}48}{42}+\frac{0{,}02}{2}+0{,}01\right)\cdot$$
$$\cdot5{,}5\cdot10^{-6}\cdot(1+10^{-7})=6{,}8\cdot10^{-8}.$$

For the insulated switches nos. 138, 176, 238, 244, 121, 115 the probability of collision equals to zero, i.e.

$$P_{Sh}(A_{2:1})=P_{Sh}(A_{2:2})=P_{Sh}(A_{2:5})=$$
$$=P_{Sh}(A_{2:6})=P_{Sh}(A_{2:8})=P_{Sh}(A_{2:10})=0.$$

The probability of at least one collision involving passenger train no. 255N equals to (see item 14 of the algorithm)

$$P_{Sh}(A_1\,|\,R_2)=1-1\cdot1\cdot(1-6{,}8\cdot10^{-8})\cdot(1-6{,}8\cdot10^{-8})\cdot$$
$$\cdot1\cdot1\cdot(1-6{,}8\cdot10^{-8})\cdot1\cdot(1-6{,}8\cdot10^{-8})\cdot1=2{,}7\cdot10^{-7}.$$

The probability of the use of routes R_3, R_4, R_5, R_6 equals to (see item 8 if the algorithm) as

$$P(R_3)=P(R_4)=P(R_5)=P(R_6)=\frac{0}{2+1+0+0+0+0}=0.$$

Therefore the probabilities $P_{Sh}(A_1\,|\,R_3)$, $P_{Sh}(A_1\,|\,R_4)$, $P_{Sh}(A_1\,|\,R_5)$, $P_{Sh}(A_1\,|\,R_6)$ of at least one collision on those routes do non need to be calculated.

The probability of at least one collision involving passenger train no. 255N while crossing the station equals to (see item 17 of the algorithm)

$$P_{Sh}(A_1)=\sum_{k=1}^{K}P_{Sh}(A_1\,|\,R_k)P(R_k)=\sum_{k=1}^{6}P_{Sh}(A_1\,|\,R_k)P(R_k)=$$
$$=2{,}7\cdot10^{-7}\cdot\frac{2}{3}+2{,}7\cdot10^{-7}\cdot\frac{1}{3}=2{,}7\cdot10^{-7}.$$

Now, let us calculate the probability of at least one collision for a station equipped with MALS during a year.

Let the previous transits across the station for all trains be unknown. Then, the probability of at least one collision between passenger trains that pass the station without stopping and a shunting consist in May equals to

$$P_{Sh}(A_{\text{May}})=0{,}00002.$$

The probability of at least one collision between passenger trains that pass the station without stopping and a shunting consist in August equals to

$$P_{Sh}(A_{\text{August}})=0{,}00008.$$

The traffic density in August is peak, i.e. in other months the probability of at least one collision will be lower. This assumption is confirmed by the timetable analysis. From here it follows that the probability of at least one collision during a year is evaluated with the value

$$P_{Sh}(A_{\text{year}})=1-(1-0{,}00008)^{12}=0{,}00096.$$

If the MALS system is used in the station, the probability of at least one collision involving a passenger train moving through the station is most affected by the summand related to the coupling. That is due to the fact that the probability of signal violation by a shunting consist after coupling with the "coupling" mode off turns out to be higher than in the "pull up" mode or while performing the other shunting operations. If a freight train is considered, the probability of its collision with a shunting consist is most affected by the summand not associated with the coupling or pulling up. That is due to the high probability of signal violation by the freight train driver. The probability of at least one collision involving a passenger train is as expected lower than the probability of at least one collision involving a freight train. That is also due to the fact that the probability of signal violation by a freight train driver is two orders of magnitude higher than the probability of signal violation by a shunting consist driver, as well as the fact that freight trains are longer than passenger trains, therefore they occupy a switch longer and the probability of collision will last longer as well.

If we compare the resultant probability of at least one collision during a year with the respective probability in [2] for shunting engines not equipped with MALS, this probability will turn out to be an order of magnitude lower.

6. Conclusion

The paper has examined the matters related to the evaluation of the risk of collision in railway stations in which shunting engines may be equipped with the MALS system. The main focus is on the method of calculation of the probability of collision between passenger trains and shunting consists in a railway station based on the shunting consist traffic density and specific passenger train schedule. The use of MALS in a railway station helps significantly reduce the probability of collision within a year.

References

1. GOST R 33433-2015 Functional safety. Risk management in railway transportation.

2. Ignatov AN, Kibzun AI, Platonov EN. Otsenka veroiatnosti stolknoveniya zheleznodorozhnykh sostavov [Probability estimation of train-to-train collision]. Avtomatika i telemekhanika 2016;11:2016;43-59 [in Russian].

3. Shubinsky IB, Zamyshliaev AM, Ignatov AN, Kan YuS, Kibzun AI, Platonov EN. Dependability 2016;3:39-46.

4. Shubinsky IB. Funktsionalnaia nadiozhnost informatsionnykh system. Metody analiza [Functional reliability of information systems. Analysis methods]. Ulianovsk: Oblastnaya tipografia Pechatny dvor; 2012 [in Russian].

About the authors

Igor B. Shubinsky, Doctor of Engineering, Professor, Director, ZAO IBTrans, Moscow, Russia, phone: +7 (495) 786 68 57, e-mail: igor-shubinsky@yandex.ru

Aleksei M. Zamyshliaev, Doctor of Engineering, Deputy Director General, JSC NIIAS, Moscow, Russia, phone: +7 (495) 967 77 02, e-mail: a.zamyshlaev@gismps.ru

Aleksei N. Ignatov, Moscow Aviation Institute, postgraduate student, Moscow, Russia, phone: +7 (906) 059 50 00, alexei.ignatov1@gmail.com

Andrei I. Kibzun, Doctor of Physics and Mathematics, Professor, Moscow Aviation Institute, Head of Chair, Moscow, Russia, phone: +7 (499) 158 45 60, e-mail: kibzun@mail.ru

Evgeni N. Platonov, Candidate of Physics and Mathematics, Associate Professor, Moscow Aviation Institute, Faculty of Applied Mathematics and Physics, Moscow, Russia, phone: +7 (499) 158 45 60, e-mail: en.platonov@gmail.com

Received on 17.03.2017

State of the art and development prospects of functional safety norms and standards

Alexander F. Kolchin, *Korporativnie elektronnie sistemy, Russia, Moscow*
Oleg S. Yakimov, *KVF Interstandart, Russia, Moscow*

Abstract. Aim. *Familiarizing the readers with the state of the art and development prospects of functional safety norms and standards in the Russian Federation. As the safety of any product, service or process is its second most important characteristic after the function, safety-related systems (SRSs) are widely used in order to ensure the safety of industrial, transportation, energy, communication and critical facilities, buildings and structures, urban infrastructure, as well as machines, equipment and vehicles. Unfortunately, since 1980's the technologies used in the development of the SRSs have not gained full traction in Russia. As the result, a conservative approach is in use that often involves excessive requirements, which increases the cost of the developed safety systems but usually does not guarantee compliance with the requirements. Currently, functional safety (FS) is recognized globally as the primary SRSs characteristic, that indicates the probability of successful performance by the system of the safety function(s) under the given conditions within the given time period.* **Methods.** *Globally, the implementation, further development and practical application of the FS method is based on the development and application of a large number of regulatory documents at the international, regional and national levels, that help organize and perform activities related to the assessment and FS requirements compliance confirmation for a wide range of SRSs. In order to ensure methodological support and coordination of the activities aimed at the development of FS-related regulatory framework in the Russian Federation in accordance with the national standard GOST R 1.1-2013 Standardization in the Russian Federation. Technical committees for standardization. Rules of organization and function, the technical committee for standardization TK 058 Functional Safety has been established, is actively working and has so far developed around 50 FS-related standards. The TK 058 standardization activities are based on the provisions of the Federal Law dated June 29, 2015 no. 162-FZ On standardization in the Russian Federation.* **Conclusions.** *As in the Russian Federation a certain FS-related regulatory framework has already been established, while the market shows demand for services of FS requirements compliance evaluation, the main task for today is to develop, using national and international requirements, organizational support, regulatory and guidance documentation that would create a fully-fledged infrastructure that implements the national institution of FS requirements compliance verification. That will ensure not only a radical reduction of the risk of disasters and accidents, but also significantly increase the competitiveness of Russian products in the internal and foreign markets.*

Keywords: *functional safety, regulatory framework, standardization, safety integrity level, certification, compliance evaluation*

For citation: Kolchin AF, Yakimov OS. State of the art and development prospects of functional safety norms and standards. Dependability 2017;3: p. 58-62. DOI: 10.21683/1729-2646-2017-17-3- 58-62.

Introduction

Everything that people create to meet their needs is essentially DANGEROUS (for people and for the environment). Therefore, creating any object should include identification and analysis of the hazards associated with that object. Thus, along with ensuring the availability of required functions for the created object, one must ensure the correct safe functioning (behavior) of that object, taking into account the interrelationships of the object's various systems with each other and the environment at all life cycle stages of that object.

For that reason, the safety of any product, service or process is its second most important characteristic after the function.

Safety-related systems (SRSs) are widely used in order to ensure the safety of industrial, transportation, energy, communication and critical facilities, buildings and structures, urban infrastructure, as well as machines, equipment and vehicles.

Unfortunately, since 1980's the technologies used in the development of the SRSs have not gained full traction in Russia. Their main shortcomings are:

1. Domestic developers and manufacturers of the SRSs, design and construction organizations, with rare exceptions, prefer to use old and obsolete regulatory documents based on prescriptive approach, even though compliance with the requirements of these documents does not guarantee the systems' performance and therefore does not guarantee the safety of the facility in which these systems are installed.

2. The SBSs and their subsystems are seen as autonomous independent production units (thing in itself), and their hazard/safety is assessed without taking into account the interrelationships of their components with each other and the environment.

As a result, a conservative approach is in use that often involves excessive requirements, which increases the cost of the developed safety systems, but usually does not guarantee compliance with the requirements.

Currently, functional safety (FS) is recognized globally as the primary STS characteristic, that indicates the probability of successful performance by the system of the safety function(s) under the given conditions within the given time period.

The FS method presented in the set of standards GOST R IEC 61508:

1. Uses a single (industry-independent) system integrated process approach and is aimed at identifying, preventing and mitigating the consequences of all safe failures, detected hazards, as well as reasonably predictable hazards and rare hazards that can lead to catastrophic consequences in complex technical systems.

2. Introduces a single measure of safety assessment, i.e. safety integrity level (SIL), which is presented and evaluated in terms of unacceptable risk of harm to people, property and the environment. The FS method implies a regular iterative process of hazard and risk analysis, an overall risk assessment and risk reduction measures that are implemented at all stages of the lifecycle of the SBSs. It also prescribes the actions of all individuals who can influence safety at these stages.

3. Is actively used in industrialized countries all around the world. Its application is regulated by more than 200 international, regional and interstate standards in various industries.

Globally, further development and practical application of the FS method is based on the development and application of a large number of regulatory documents at the international, regional and national levels that help to organize and perform activities related to the assessment and FS requirements compliance confirmation for a wide range of SRSs.

Although a number of Russian researchers have been dealing with FS problems for more than 20 years, a wide range of engineering and technical specialists became acquainted with the practical application of the FS method after the well-known works by David J. Smith and Kenneth J.L. Simpson [1, 2] were published in the Russian Federation and the first edition of the basic FS standard GOST R IEC 61508-2007 Functional safety of electrical, electronic, programmable electronic safety-related systems. Parts 1-7 were released in 2007.

In order to ensure methodological support and coordination of the activities aimed at the development of FS-related regulatory framework in the Russian Federation in accordance with the national standard GOST R 1.1-2013 Standardization in the Russian Federation. Technical committees for standardization. Rules of organization and function, the Technical Committee (TC) for Standardization TK 058 Functional Safety has been established and later restructured. In addition, several other related national technical committees for standardization of leading industries are also presently participating in the establishment of a national regulatory framework in the field of FS.

The TK 058 standardization activities are based on the provisions of the Federal Law dated June 29, 2015 no. 162-FZ On standardization in the Russian Federation.

The participation in TK 058 is voluntary.

The TC was established to enable cooperation among concerned organizations and authorities when performing national, interstate and international standardization activities in the field of FS.

The main goals of the TC in the field of FS standardization are:

- Developing annual national standardization programs and overseeing the implementation of these programs;
- Considering application of international and regional standards at the national and interstate level proposals;
- Carrying out scientific and technical, legal and regulatory examination of national and interstate standard projects and existing standards change projects and submitting them for approval to federal executive authority in standardization.

The results of this activities achieved over the past 10 years are briefly reviewed below.

The Russian standards in the field of FS

The focus of the TK 058 was on preparing the second edition of the basic standard IEC 61508-2010, as well as baseline FS standards for various industries. As a result, a set of standards GOST R IEC 61508-2010 [3-9] was established in 2012.

At the moment, the IEC 61511 [10-12] set of standards is actively used at a number of petrochemical, gas and electrical power enterprises that use safety instrumented systems to ensure the safety of various industrial processes.

Several international standards were introduced directly for the construction industry and national FS standards [13–19] were developed on the basis of IEC 61508 [13-19]. A number of corporate standards for construction industry have already been developed based on these regulatory documents.

A lot of attention in the Russian Federation is paid to the implementation of the FS method in railway transportation. A fairly large number of corporate standards for maintaining the guaranteed safety and reliability of the transportation process in JSC RZD have been developed and are now in use in the company. There are a large number of national and interstate FS standards in place in the railway industry, some of which are presented in [20-23].

In the nuclear power plant engineering, several standards concerning monitoring and control systems for the industry's various products have been developed, that comply with FS requirements [24–32].

The problems of FS of machines and mechanisms FS problems are considered in regulatory documents [33-36].

The implementation of the FS method for road vehicles control systems is presented in [37-47].

The general principles for the implementation of the IEC 61508 series requirements for safety-related data communication, including possible data communication failures, recovery measures and considerations related to data integrity in industrial communication networks can be found in [48-53].

Immunity requirements for safety-related systems and for equipment intended to perform safety-related functions are presented in [54, 55].

FS requirements for programmable controllers are considered in [56].

At present, 4 more FS standards, including the 2-nd edition of IEC 61511-2016, have been developed and are awaiting approval.

From standardization to compliance assessment

Today it is practically impossible to export a complex technical system or its components without a confirmed safety integrity level compliance assessment.

Almost any complex systems delivered in the Russian Federation are marked with this characteristic that explicitly describes their safety.

Therefore, in most economically developed countries compliance confirmation (certification) agencies with corresponding institutions (testing laboratories/centers) providing measurements, tests, requirements compliance assessment calculations, are working to confirm FS requirements compliance of complex equipment, industrial facilities, systems and their components.

There is no national institution of FS requirements compliance verification in Russia yet, which not only radically increases the risk of disasters and accidents, but also significantly reduces the competitiveness of Russian products in the internal and foreign markets. Meanwhile, there are all the necessary conditions for the creation of such institution. A FS-related regulatory framework has already been established and is quite relevant, while the market shows demand for services of FS requirements compliance evaluation.

Therefore, the main task for today is to develop, using national and international requirements, organizational support, regulatory and guidance documentation that would create a fully-fledged infrastructure in the Russian Federation that implements the national institution of FS requirements compliance verification.

In March 2016 the Voluntary Certification System in the Field of Functional Safety (registration number ROSS RU.31461.04IDD0) was registered in the unified register of voluntary certification systems of the Federal Agency on Technical Regulation and Metrology. It certifies safety-related systems (SRSs), their components, related products, functional safety management systems of organizations and/or subdivisions that develop, produce and use SRSs: hazard and risk analysis, design, manufacture, installation, putting into operation, maintenance and repair, modernization, decommissioning of SRSs, safety-related systems development, manufacture and application tools. The FS certification authority is also to be accredited by the National Voluntary Certification System of the Russian Federation.

References

1. Smith DJ, Simpson KGL. Functional safety. A straightforward guide to applying IEC 61508 and related standards. Moscow: Tekhnologii; 2004.

2. Smith DJ. Reliability, maintainability and risk. Practical methods for engineers including reliability centered maintenance and safety-related systems. Moscow: Gruppa IDT; 2007.

3. GOST R IEC 61508-1-2012 Functional safety of electrical, electronic, programmable electronic safety-related systems. Part 1. General requirements

4. GOST R IEC 61508-2-2012 Functional safety of electrical, electronic, programmable electronic safety-related systems. Part 2. Requirements for systems

5. GOST R IEC 61508-3-2012 Functional safety of electrical, electronic, programmable electronic safety-related systems. Part 3. Software requirements

6. GOST R IEC 61508-4-2012 Functional safety of electrical, electronic, programmable electronic safety-related systems. Part 4. Terms and definitions

7. GOST R IEC 61508-5-2012 Functional safety of electrical, electronic, programmable electronic safety-related systems. Part 5. Examples of methods for the determination of safety integrity levels

8. GOST R IEC 61508-6-2012 Functional safety of electrical, electronic, programmable electronic safety-related systems. Part 6: Guidelines on the application of IEC 61508-2 and IEC 61508-3

9. GOST R IEC 61508-7-2012 Functional safety of electrical, electronic, programmable electronic safety-related systems. Part 7. Techniques and measures

10. GOST R IEC 61511-1-2011 Functional safety. Safety instrumented systems for the industrial processes. Part 1. Terms, definitions and technical requirements

11. GOST R IEC 61511-2-2011 Functional safety. Safety instrumented systems for the industrial processes. Part 2. Guidelines for the application of IEC 61511-1

12. GOST R IEC 61511-3-2011 Functional safety. Safety instrumented systems for the industrial processes. Part 3. Guidelines for the determination of the required safety integrity levels

13. GOST R ISO/IEC 14762-2013 Information technology – Functional safety requirements for home and building electronic systems (HBES)

14. GOST R 53195.1-2008 Functional safety of building/erection safety-related systems. Part 1. General

15. GOST R 53195.2-2008 Functional safety of building/erection safety-related systems. Part 2. General requirements

16. GOST R 53195.3-2015 Functional safety of building/erection safety-related systems. Part 3. Requirements for systems

17. GOST R 53195.4-2010 Functional safety of building/erection safety-related systems. Part 4. Software requirements

18. GOST R 53195.5-2010 Functional safety of building/erection safety-related systems. Part 5. Techniques and measures on risk reduction, estimation methods

19. GOST R EN 50491-4-1-2014 General requirements for Home and Building Electronic Systems (HBES) and Building Automation and Control Systems (BACS) – Part 4-1: General functional safety requirements for products intended to be integrated in Building Electronic Systems (HBES) and Building Automation and Control Systems (BACS)

20. GOST R 55980-2014 Risk management on railway transport. Hazardous events classification

21. GOST 33432-2015 Functional safety. Policy and programme of safety provision. Safety proof of the railway objects

22. GOST 33433-2015 Functional safety. Functional safety. Risk management on railway transport

23. GOST R IEC 62279-2016 Railway applications. Communication, signalling and processing systems. Software for railway control and protection systems

24. GOST R IEC 60880-2010 Nuclear power plants. Instrumentation and control systems important for safety. Software aspects for computer-based systems performing category A functions

25. GOST R IEC 62138-2010 Nuclear power plants. Instrumentation and control systems important for safety. Software aspects for computer-based systems performing category B or C functions

26. GOST R IEC 60987-2011 Nuclear power plants. Instrumentation and control systems important to safety. Hardware design requirements for computer-based systems

27. GOST R IEC 61513-2011 Nuclear power plants. Instrumentation and control important to safety. General requirements for systems

28. GOST R IEC 61225-2011 Nuclear power plants. Instrumentation and control systems important for safety. Requirements for electrical supplies

29. GOST R IEC 61226-2011 Nuclear power plants. Instrumentation and control systems important for safety. Classification of instrumentation and control functions

30. GOST R IEC 60709-2011 Nuclear power plants. Instrumentation and control systems important for safety. Separation

31. GOST R IEC 62340-2011 Nuclear power plants. Instrumentation and control systems important to safety. Requirements for coping with common cause failure

32. GOST R IEC 61500-2012 Nuclear power plants. Instrumentation and control important to safety. Data communication in systems performing category A functions

33. GOST R IEC 62061-2013 Safety of machinery. Functional safety of safety-related electrical, electronic and programmable electronic control systems

34. GOST R IEC 61800-5-2-2015 Adjustable speed electrical power drive systems. Part 5-2. Functional safety requirements

35. GOST R 60.1.2.1-2016 Robots and robotic devices. Safety requirements for industrial robots. Part 1. Robots

36. GOST R 60.1.2.2-2016 Robots and robotic devices. Safety requirements for industrial robots. Part 2. Robot systems and integration

37. GOST R ISO 26262-1-2014 Road vehicles. Functional safety. Part 1: Vocabulary

38. GOST R ISO 26262-2-2014 Road vehicles. Functional safety. Part 2: Management of functional safety

39. GOST R ISO 26262-3-2014 Road vehicles. Functional safety. Part 3. Concept phase

40. GOST R ISO 26262-4-2014 Road vehicles. Functional safety. Part 4. Product development at the system level

41. GOST R ISO 26262-5-2014 Road vehicles. Functional safety. Part 5. Product development at the hardware level

42. GOST R ISO 26262-6-2014 Road vehicles. Functional safety. Part 6: Product development at the software level

43. GOST R ISO 26262-7-2014 Road vehicles. Functional safety. Part 7: Production and operation

44. GOST R ISO 26262-8-2014 Road vehicles. Functional safety. Part 8: Supporting processes

45. GOST R ISO 26262-9-2014 Road vehicles. Functional safety. Part 9: Automotive Safety Integrity Level-oriented and safety-oriented analyses

46. GOST R ISO 26262-10-2014 Road vehicles. Functional safety. Part 10. Guideline on ISO 26262

47. GOST R 57300-2016/ISO/TS 15998-2:2012 Earth-moving machinery. Machine control systems (MCS) using electronic components. Part 2: Use and application of ISO 15998

48. GOST R IEC 61784-1-2016 Industrial communication networks. Profiles. Part 1. Fieldbus profiles

49. GOST R IEC 61784-3-2015 Industrial communications networks. Profiles. Part 3. Functional safety fieldbuses. General rules and profile definitions

50. GOST R IEC 61784-3-1-2016 Industrial communication networks. Profiles. Part 3-1. Functional safety fieldbuses. Additional specifications for CPF 1

51. GOST R IEC 61784-3-3-2016 Industrial communication networks. Profiles. Part 3-3. Functional safety fieldbuses. Additional specifications for CPF 3

52. GOST R IEC 61784-3-8-2016 Industrial communication networks. Profiles. Part 3-8. Functional safety fieldbuses. Additional specifications for CPF 8

53 GOST R IEC 61784-3-12-2016 Industrial communication networks. Profiles. Part 3-12. Functional safety fieldbuses. Additional specifications for CPF 12

54. GOST IEC 61326-3-1-2015 Electrical equipment for measurement, control and laboratory use. EMC requirements. Part 3-1. Immunity requirements for safety-related systems and for equipment intended to perform safety-related functions (functional safety). General industrial applications

55. GOST IEC 61326-3-2-2015 Electrical equipment for measurement, control and laboratory use. EMC requirements. Part 3-2. Immunity requirements for safety-related systems and for equipment intended to perform safety-related functions (functional safety). Industrial applications with specified electromagnetic environment

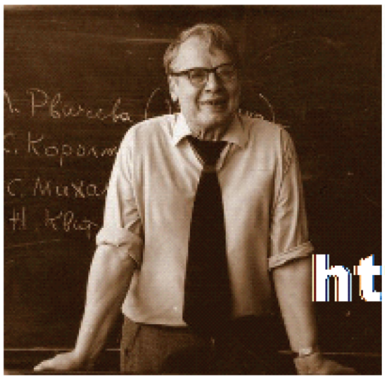
56. GOST R IEC 61131-6-2015 Programmable controllers. Part 6. Functional safety

About the authors

Alexander F. Kolchin, Deputy Director General, OOO Korporativnie elektronnie sistemy, Russia, Moscow, e-mail: kolchin@calscenter.ru

Oleg S. Yakimov, Director of Regulatory Support, KVF Interstandart, Chairman, Technical Committee for Standardization 058 Functional Safety, Russia, Moscow

Received on 30.06.2017



http://Gnedenko-Forum.org/

Dear colleagues!

In 2005 the informal Association of Experts in Reliability, Applied Probability and Statistics (I.G.O.R.) was established with its own Internet website GNEDENKO FORUM. The site has been named after the outstanding mathematician Boris Vladimirovich Gnedenko (1912-1995). The Forum's purpose is an improvement of personal and professional contacts between experts in the mathematical statistics, probability theory and their important branches, such as reliability theory and quality control, the theory of mass service, storekeeping theory, etc.

Since January 2006, the Forum has published a quarterly international electronic magazine

"Reliability: Theory and Applications".

The magazine is registered with the Library of Congress in the USA (ISSN 1932-2321). All rights reserved for authors so that articles can be freely published in any other publications or presented at conferences.

Join Gnedenko Forum!

Welcome!

More than 500 experts from 44 countries worldwide have already joined us!

To join the Forum, send a photo and a short CV to the following address:

Alexander Bochkov, PhD
a.bochkov@gmail.com

Membership is free.

REQUIREMENTS OF EDITION ON EXECUTION OF PAPERS IN JOURNALS OF PUBLISHING GROUP OF IDT PUBLISHERS

A letter from the organisation where the author (s) works or from the author (s) personally with the paper offered for publication should be sent to the de facto editorial office address: 107078, Moscow, 5 Orlikov lane, Office 755, LLC "JOURNAL DEPENDABILITY" or e-mail: E.Patrikeeva@gismps.ru (in scanned form).

The letter should be attached to a paper text containing the summary and keywords, information on authors, bibliographic list, and one complete set of figures. All listed items are to be presented in an electronic form (on CD or via the e-mail address provided above).

Attention! Titles of papers, names of authors, summary and keywords must be presented, in Russian and English languages, according to the requirements of the Higher Attestation Commission. The information on each author should contain the following standard data:

- Surname, name, patronymic;
- Scientific degree, academic status, honorary title;
- Membership of relevant public unions, etc.;
- Place of employment, position;
- The list and numbers of Journals of IDT Publishers in which papers of the author have been previously published;
- Contact information.

Texts should be presented in Word 97-2003 format in a 12-point typeface; the text should not be formatted. Paragraphs should be arranged by pressing the "return" key. The text of the paper should be double-spaced on pages of A4; on the left there should be a margin of 2 cm; pages should be numbered, the «first line indent» is obligatory.

All alphabetical designations represented in figures should be explained in the body text or in a legend. Inconsistencies between designations in figures and in the text are inadmissible. Numbering should only be applied to those formulas and equations that are referred to in the text.

Simple formulas appearing directly in the text (for example, m^2, n^2t, $c = 1 + DDF – A_2$), and the Greek letters and symbols, for example, β, © may be typed using the Symbol font. When it is not possible to type directly in the text editor, use the "Microsoft Equation" formula editor (available with the complete installation of Microsoft Office) or the "Mathtype" formula-editing program. Representation of formulae in the text in the form of images is not admissible. Photos and figures for papers should be provided in individual files with extension TIF, EPS or JPG with a resolution of not less than 300 dpi. The list of literature referred to in the paper (bibliography) is presented according to order of citation and provided at the end of paper. References to the literature in the text are marked by serial numerals in square brackets.

To authors that are published in journals of "IDT Publishers".

In addition to the journal, information on each author will be presented at the techizdat.ru site in the «Authors» section on the individual web page.

Authors of papers for publication have the opportunity to send an electronic photo and additional material to appear on this individualised Internet-business card. At their own discretion, authors can present more details about themselves, interesting examples and stories of solutions to technical problems, about contemporary problems according to subjects of corresponding journal, etc. This material should not exceed 1000 characters including spaces.

SUBSCRIPTION TO THE JOURNAL «DEPENDABILITY»

It is possible to subscribe to the journal:

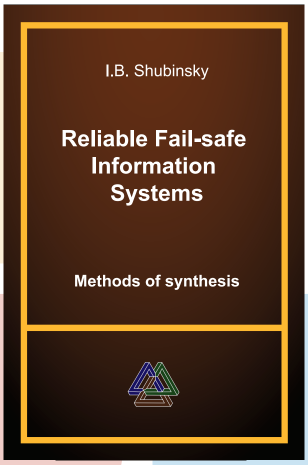
- Through the agency «Rospechat»
– for the first half of the year: an index 81733;

- Under the catalogue "Press of Russia" of the agency «Books-services»:
– for half a year: an index 11804;

- Through the editorial office:
– for any time-frame
tel.: 8-916-105-81-31; e-mail: E.Patrikeeva@gismps.ru

НАДЁЖНОСТЬ
DEPENDABILITY SCIENTIFIC-TECHNICAL JOURNAL
ISSN 1729-2646

REPRESENTS


I.B. Shubinsky

Reliable Fail-safe Information Systems

Methods of synthesis

Publication can be purchased through the editorial board of Journal Dependability Ltd.

+7 (495) 967-77-05, ext.186
+7-916-105-81-31
(Patrikeeva Evgenia)

E.Patrikeeva@gismps.ru,
www.dependability.pro

Igor B. Shubinsky

RELIABLE FAIL-SAFE
INFORMATION SYSTEMS
Methods of synthesis

Copy deadline 12.02.2016. format of the edition 70x100/16.
Offset printing. Offset paper. Conv. Sheet. P. 17,55.
Circulation of 700 copies. Order number 1452.

Journal Dependability Ltd,
109029, Moscow,
Nizhegorodskaya str.27, bldg.1, office 209
Tel./fax: +7 499 262 53 20
E-mail: E.Patrikeeva@gismps.ru

I.B.Shubinsky
Reliable Fail–safe Information Systems
2016

The book describes conceptual provisions to ensure structural and functional reliability of information systems at all stages of a life-cycle. It represents different types of redundancy taking into account limited efficiency of the failure detection system. Under these conditions a broad-based assessment of their efficiency is performed, with determination of capabilities of structural redundancy with an endless number of standby facilities. Ways to ensure functional reliability of software are represented, including the recommendations for the development of software programs requirement specification, with the description of the process of a reliable program architecture development and well proven rules and recommendations used for design and implementation of software, as well as for integration with system hardware.

The book also presents theoretical and practical provisions of adaptive fault tolerance (active protection) of information systems, including the methods and disciplines of active protection, as well as the ways of implementation. A method of synthesis of active protection and the results of research of information system reliability with various disciplines of active protection are offered. There are also certain assessments of the efficiency of active protection in relation to the traditional methods of structural redundancy.

You can find the description of the principles to ensure functional safety of information systems, with a substantiation of the possibility to restart independent channels in two-channel safe systems. The rules of determination of the allowed time for a guaranteed detection of single and double hazardous failures are developed, including the method of synthesis of a combined two-level information system developed with higher functional safety requirements.

To prove the conformance of reliability with functional safety the method of accelerated field testing of the information system has been developed. The book contains the description of this method, including the example of its practical implementation. You will also find the information about the procedures of certification tests based on the requirements of information safety and software certification conformance.

A checklist of the most complex and significant subjects is provided at the end of each chapter. The book is primarily intended for experts who are engaged in practical development, manufacture, operation and updating of information. It is intended for researchers in the field of structural reliability of different discrete systems, academic staff, post-graduate students and students specializing in the field of information systems and as well as those working in the field of automated control systems.

Publication can be purchased through the editorial board of Journal Dependability Ltd.
By phone +7 (495) 967-77-05, ext. 186; +7-916-105-81-31 (Patrikeeva Evgenia)
e-mail: E.Patrikeeva@gismps.ru, www.dependability.pro

SUBSCRIBER APPLICATION FOR DEPENDABILITY JOURNAL

Please subscribe us for 20___
from No. ________ to No._______ number of copies _________

Company name	
Name, job title of company head	
Phone/fax, e-mail of company head	
Mail address (address, postcode, country)	
Legal address (address, postcode, country)	
VAT	
Account	
Bank	
Account number	
S.W.I.F.T.	
Contact person: Name, job title	
Phone/fax, e-mail	

Publisher details: Dependability Journal Ltd.
Address of the editorial office: office 209, bldg 1, 27 Nizhegorodskaya Str., Moscow 109029,
Russia Phone/fax: 007 (495) 967-77-02, e-mail: E.Patrikeeva@gismps.ru
VAT 7709868505 Account 890-0055-006
Account No. 40702810100430000017
Account No. 30101810100000000787

Address of delivery:

To whom:____

Where:____

To subscribe for Dependability journal, please fill in the application form and send it by fax or email.
In case of any questions related to subscription, please contact us.
Cost of year subscription is 4180 rubles, including 18 per cent VAT.
The journal is published four times a year.

THE JOURNAL IS PUBLISHED WITH PARTICIPATION AND SUPPORT

OF JOINT-STOCK COMPANY RESEARCH & DESIGN INSTITUTE
FOR INFORMATION TECHNOLOGY, SIGNALLING AND TELECOMMUNICATIONS
ON RAILWAY TRANSPORT (JSC NIIAS)

JSC NIIAS is RZD's leading company in the field of development of train control and safety systems, traffic management systems, GIS support technology, railway fleet and infrastructure monitoring systems

Mission:

transportation

□ efficiency,

□ safety,

□ reliability

Key areas of activity

- Intellectual control and management systems
- Transportation management systems and transport service technology
- Signalling and remote control systems
- Automated transportation management centers
- Railway transport information systems
- Geoinformation systems and satellite technology
- Transport safety systems
- Infrastructure management systems
- Power consumption and energy management systems
- Testing, certification and expert assessment
- Information security
- Regulatory support

www.vniias.ru