Kostogryzov Andrey 0000-0002-0254-5202
Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Moscow, Russia, akostogr (at) gmail.com

Nistratov Andrey 0000-0002-0688-4156
Scientific and Engineering Centre for Energy Safety, Moscow, Russia, andrey.nistratov (at) gmail.com

Probabilistic methods of risk predictions and their pragmatic applications in life cycle of complex systems

Keywords

analysis, model, operation, prediction, probability, rationale, risk, system, system engineering

Abstract

The probabilistic methods of risk predictions, focused on the prognostic modeling of complex systems, and their pragmatic applications are proposed. They are based on an implementation of the proposed interconnected ideas about system analysis in life cycle. The approach includes description of the probabilistic models, optimization methods for rationale actions and incremental algorithms for solving the problems of supporting decision-making and rationale preventive actions in uncertainty conditions. A suitability of the proposed models and methods is demonstrated by examples which cover wide reliability and safety applications for some intellectual systems, enterprises of coal company, a floating oil and gas platform, a set of manufacturing processes of gas preparation equipment, the systems of oil&gas transportation and distribution. The approach means practically proactive commitment to excellence in uncertainty conditions.

1. Introduction

Different views and methods of risk predictions and their pragmatic applications are recommended at level of the international standards of system engineering – for example, ISO/IEC/IEEE 15288 “System and software engineering – System life cycle processes”, ISO 17359 "Condition monitoring and diagnostics of machines – General guidelines", IEC 61508 “Functional safety of electrical/ electronic/ programmable electronic safety-related systems” etc. They are useful every time across life-cycle to meet reliability, safety and quality requirements on the base of estimating system behavior, feasibility, tracing critical quality characteristics, analyzing risks, sensitivity for changes of critical parameters values etc. The modern and perspective complex systems are notable for the fact that intelligence elements are used. These intellectual systems (IS) using such elements are operated by logic reasoning on the base of data processing. They also needs system analysis because of their complexities and uncertainty conditions.

Note. System is combination of interacting elements organized to achieve one or more stated purposes (according to ISO/IEC/IEEE 15288).

Note. According to ISO Guide 73 risk is defined as effect of uncertainty on objectives considering consequences. An effect is a deviation from the expected – positive and/or negative.

Considering intelligence elements specificity for complex IS there may be some scientific problems devoted to:

- system analysis of uncertainty factors, capabilities of operation in real time, information gathering and processing, protection from authorized access and dangerous influences;
- analysis of system requirements to acceptable conditions;
- system analysis and optimization in architectural design;
- comparative and prognostic estimations of quality, safety, interaction “user-system” and conditions, optimization of different processes, rationale of operation in uncertainty, etc.

DOI: 10.26408/srsp-2020-11

Now there isn’t enough universal effective approach to rationale of actions for complex IS operating in uncertainty conditions. In practice for each concrete case it is often used subjective expert estimations, a regression analysis of collected data, simulation and modeling processes – see [4], [6], [9]–[10], [12]–[13], [15]–[17], [23], [36]–[39] etc. It means, that scientific research of new methods for advanced rationale IS actions is today in high demand. For such systems the proposed approach is focused on probabilistic rationale of actions to operate in uncertainty conditions against existing approaches for which applied mathematical methods cover mainly information processing in the logician «if …, that …» and/or tracing situations by a man-operator. An application scope of this paper covers complex IS supporting decision-making in system engineering, the proposed methods are used to provide operation efficiency or/and increase reliability, safety and quality.

Note. The main efforts of this paper are not connected with illustrating the capabilities of modern and perspective complex systems, but they are focused on demonstrating the applicability of presented original probabilistic models and methods to improve some from existing capabilities for these systems.

For this goal by the use of the proposed probabilistic models the specific problems of supporting decision-making in uncertainty conditions are covered and explained by examples (see Section 6). These may be: the problem to rationale a rational variant for decision-making on the base of data monitored about events and conditions in real time; the problem to rationale preventive actions during long time period under limitations on admissible risks of system “failures” etc.

Note. Some relevant mechanical problems of robotics (for which different probabilistic methods are also applicable) are not covered by this paper.

The proposed approach is based on theoretical and practical research [1]–[39] and may be used either in combination or in addition to existing methods which are useful. There, where it is required often prognostic system analysis or where the used approaches are not effective, the proposed probabilistic approach can be used as rational basis or alternative. The ideas of this approach may be applied also by using another probabilistic models which supported by software tools and can predict successfulness or failures on a level of risks estimated by probability distribution functions (PDF) against consequences. Various fields of the examples applications have been chosen purposefully to demonstrate universality and analytical usefulness of the probabilistic methods.

The proposed models and methods have been presented at seminars, symposiums, conferences, ISO/IEC working groups and other forums since 80th in Russia, Australia, Canada, China, Finland, France, Germany, Italy, Kuwait, Luxembourg, Poland, Serbia, the USA, etc. The supporting software tools were awarded by the Golden Medal of the International Innovation and Investment Salon and the International Exhibition “Intellectual Robots”, acknowledged on the World's fair of information technologies CeBIT in Germany, noted by diplomas of the Hanover Industrial Exhibition and the Russian exhibitions of software.

2. Essence of the approach

The system efficiency corresponding to the rationale of actions for IS operation in uncertainty conditions means proactive commitment to excellence. In practice the achieved effects are often based on an implementation of the next proposed interconnected ideas 1–7.

Idea 1. It is concerning the usual concept and properties of PDF (see for example [4]–[6], [10], [12], [16], [23] etc.) for a continuous random variable of time. PDF for a time variable τ is nondecreasing function $P(t)$ whose value for a given point $t \geq 0$ can be interpreted as a probability that the value of the random variable τ is less or equal to the time value t, i.e. $P(t) = P(\tau < t)$. Additionally $P(t) = 0$ for $t < 0$, and $P(t) \to 1$ for $t \to \infty$. In general case the solutions for the problems in decision-making

are based on using concept of the probabilities of "success" and/or "unsuccess" (risk of "failure" considering consequences) during the given prognostic time period t_{req}. This probability is a value for a point t_{req} and is defined by created PDF in modeling.

Idea 2. The processes, connected with data processing, and used information should provide required system operation quality (because system performs functions by logic reasoning on the base of data processing). And corresponding probabilistic methods should be appropriate for prognostic estimations [6], [9], [11]–[17], [23], [31], [36]–[37].

Idea 3. The PDF should be presented as analytical dependence on input parameters. It needs to solve direct and inverse problems to rationale of system actions in real time. For example, for a simple element PDF $P(t)$ of time τ between losses of element integrity may be presented by analytical exponential approximation, i.e. $P(t) = 1 - \exp(-\lambda t)$, where λ is frequency of failures (losses of element integrity). At the same time frequency of failures may be represented as a sum of frequencies of failures because of specific reasons for each failure type – for example, failure from "human factor" λ_1, from hardware λ_2, from software λ_3 and so on. For this use case PDF may be presented as

$$P(t) = 1 - \exp[(-(\lambda_1 + \lambda_2 + \lambda_3 + \ldots)t)].$$

Then if the adequate function *P*(*t*) is built in dependence on different parameters and if admissible level for probability is given than inverse problem may be solved analytically [29].

Note. System integrity is defined as such system state when system purposes are achieved with the required quality 2. The rationale for exponential approximation choice in practice see for example in [14], [31].

Idea 4. The PDF should be adequate, it means a dependence on several essential parameters which define system operation and on which "success" or "failure" of system operation is mainly dependent. For example the way for risks prediction based on uses only one parameter – frequency of failures λ – is popular today. This implies the use of corresponding exponential PDF – see *Figure 1*. Only one connection of the frequency of failures λ with random time variable τ between losses of system integrity may be interpreted as the requirement: "to provide no failures during required time with probability no less than the given admissible probability P_{adm}. this required time should be no more than $t_{req} = 1/\lambda_{adm}$., here $\lambda_{adm} = -ln(1 - R_{adm})$". But for system element it is often rough and unpromising engineering estimations because capabilities of monitoring conditions and recovery of the lost element integrity are ignored. Such disregard deforms very essentially probabilistic estimations of probabilistic risk values and for complex system can't be useful for scientific search of effective counteraction measures against different threats. Deviations from more adequate PDF estimations are very high [2]–[3], [8], [11], [18]–[22], [25]–[29], [32]–[35]. In *Figure 2* the limitations to admissible risks, fragment of exponential and an adequate PDF of time between losses of system integrity with identical frequency of system integrity losses are illustrated (in conditional units). It means more adequate PDF allows more right understanding of probabilistic system vision of events prediction with scientific interpretation considering situations in time line.

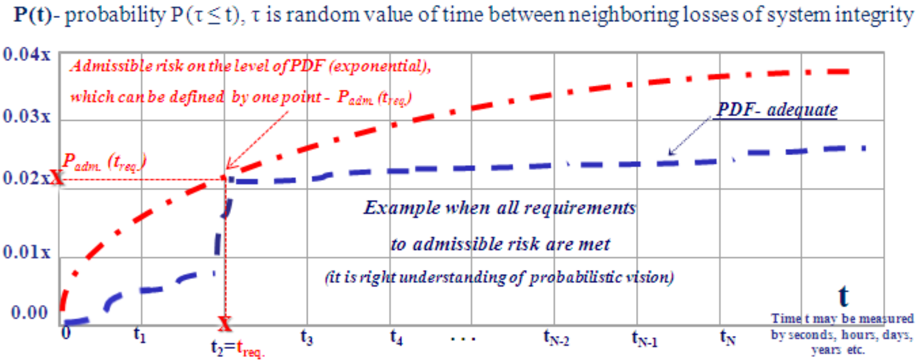


Figure 1. The possible variants of correlations for admissible risks, exponential and an adequate PDF of time between losses of system integrity

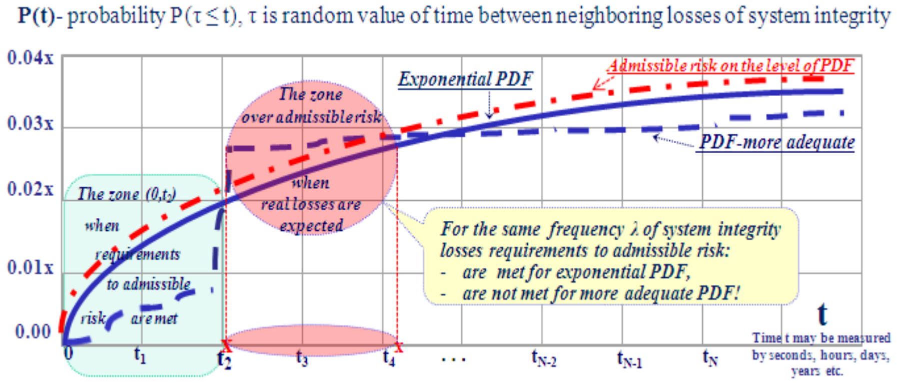


Figure 2. All requirements to admissible risk are met for an adequate PDF of time between losses of system integrity

Idea 5. Because the interested IS is a complex system and this system may be a subsystem or element of comprehensive complex system, the proposed approach should allow a generation of probabilistic models for prediction of “success” or “failure” of system actions in uncertainty conditions. In general case an input for generated models (including system used in real time) should consider system complexity, periodical diagnostics, monitoring between diagnostics, recovery of the lost integrity for every system element and also processes, connected with data processing, and used information. As an output of such generated models adequate PDF of time τ between losses of system (subsystem, element) integrity should be produced in analytical form.

Idea 6. Input for probabilistic modeling should be formed mainly from gathered data and established specific order of system operation and supporting actions.

Idea 7. To probabilistic rationale of actions for system operating in uncertainty conditions the problems of optimization should be solved. Optimization should be performed in real time by defined beforehand optimization problem statement. Every time the used optimization problem statement should be appropriated for solving specific analytical problems. For probabilistic rationale of actions the prognostic period should be defined so to be in time to do the given action or complex of actions on acceptable level according to optimization criterion or to perform preventive action (with which the initiation of performing an action or solving a problem is connected) or/and to recover operation capabilities (which can be lost and recovered on time line).

All ideas above may be applied also by using other probabilistic models which supported by software tools and can be used to predict successfulness or risks on a level of probability distribution functions.

For the approach implementation the next probabilistic models are proposed.

3. Description of proposed models

In general case a probabilistic space (Ω, B, P) for probabilistic modeling is created (see for example [4]–[6], [12], [16]–[17], [23], [39] etc.), where:

- Ω – is a limited space of elementary events;
- B – a class of all subspace of Ω – space, satisfied to the properties of σ–algebra;
- P – is a probability measure on a space of elementary events Ω.

Because, $\Omega = \{\omega_k\}$ is limited, there is enough to establish a reflection $\omega_k \rightarrow p_k$ $P(\omega_k)$ like that $p_k \geq 0$ and $\sum_k p_k = 1$.

In order not to overload the reader with mathematical details, the final proposed formulas are presented in the *Appendixes A* and *B*.

3.1. System operation quality

The created models [12]–[18], [23], [14], [31] help to implement ideas 1, 2.

In general case system operation quality is connected with requirements for reliable and timely producing complete, valid and/or, if needed, confidential information. The gathered information is used for proper system specificity. The abstract view on a quality of used information is presented by *Figure 3*.

The proposed models for the estimation of information systems operation quality are described in *Table A.1* of *Appendix*.

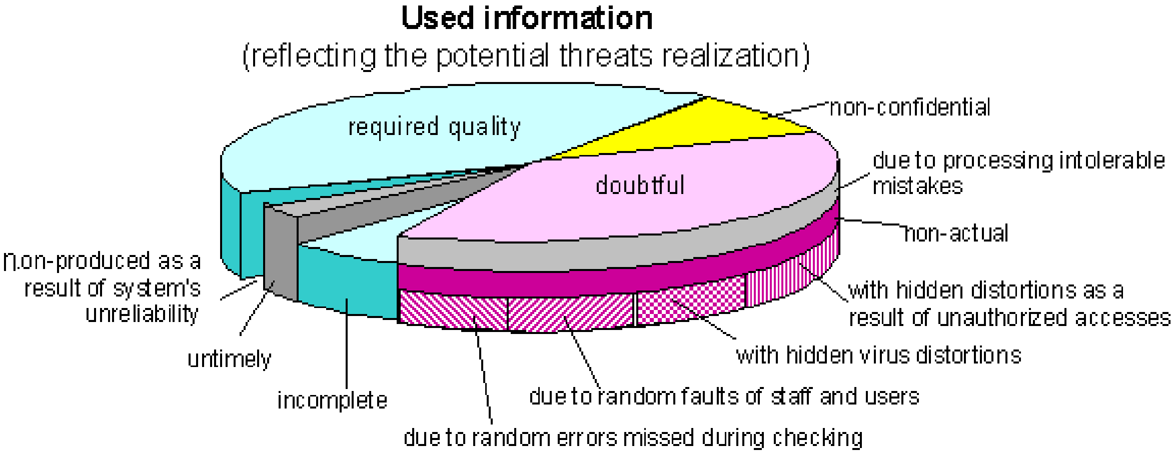


Figure 3. Abstract quality of used information against required one

The main analytical models are the next:

- “The model of functions performance by a complex system in conditions of unreliability of its components”;
- “The models complex of calls processing”;
- “The model of entering into IS current data concerning new objects of application domain”;
- “The model of information gathering”;
- “The model of information analysis”;
- “The models complex of dangerous influences on a protected system”.

“The models complex of an authorized access to system resources”. Risk to lose integrity (R) is an addition to 1 for probability of “success” (P), i.e. $R = 1 - P$ considering consequences.

These models, supported by different versions of software tools registered by Rospatent [18], may be applied for solving problems connected with decision-making for information gathering, processing and producing.

3.2. Black box formalization for prediction a risk of failure

The models below help to implement ideas 1, 3, 4 [1]–[3], [7]–[8], [11]–[12], [14], [17]–[22], [23]–[35]. In general case successful system operation is connected with counteraction against various dangerous influences on system integrity – these may be counteractions against failures, defects events, “human factors” events on time line, etc. There are proposed the formalization for two general technologies of providing counteraction against threats: periodical diagnostics of system integrity (technology 1, without monitoring between diagnostics) and additionally continuous monitoring between diagnostics (technology 2). As a rule one from these technologies is implemented by an interested IS.

Technology 1 is based on periodical diagnostics of system integrity, that is carried out to detect danger sources penetration into a system or consequences of negative influences (see *Figure 4*).

The lost system integrity can be detect only as a result of diagnostics, after which system recovery is started. Dangerous influence on system is acted step-by step: at first a danger source penetrates into a system and then after its activation begins to influence. System integrity can’t be lost before penetrated danger source is activated. A danger is considered to be realized only after a danger source has influenced on a system.Technology 2, unlike the previous one, implies that system integrity is traced between diagnostics by operator (operator functions may be performed by a man or special intelligence system element or their combination). In case of detecting a danger source an operator recovers system integrity. The ways of integrity recovering are analogous to the ways of technology 1 – see *Figure 5*.

According model assumption faultless operator’s actions provide a neutralization of danger source trying to penetrate into system. A penetration of a danger source is possible only if an operator makes an error but a dangerous influence occurs if the danger is activated before the next diagnostic. Otherwise the source will be detected and neutralized during the next diagnostic.

It is supposed for technologies 1 and 2 that the used diagnostic tools allow to provide necessary system integrity recovery after revealing danger sources penetration into system or the consequences of influences.

The probability of correct system operation within the given prognostic period, i.e. probability of “success” (P) may be estimated as a result of use the models presented in *Appendix B*. Risk to lose integrity (R) is an addition to 1 for probability of correct system operation (P), i.e. $R = 1 - P$ considering consequences.

3.3. Generation algorithm of probabilistic modeling for complex system

The proposed method for a generation of probabilistic models helps to implement ideas 1, 5.

The basic ideas of correct integration of probability measures are based on a combination and development of models. For a complex systems with parallel or serial structure described there are used the next method to generate adequate probabilistic models [1]–[4], [7]–[8], [10]–[12], [14], [18]–[22], [24]–[35], [39] etc. This method uses the usual way of probability theory for independent random variables. However, given the importance to analytical rationale the generation of new probabilistic models for complex system, the approach is described below. Let's consider the elementary structure from two independent parallel or series elements. Let’s PDF of time between losses of *i*-th element integrity is $B_i(t) = P(\tau_i < t)$, then:

1. time between losses of integrity for system combined from series connected independent elements is equal to a minimum from two times τ_1: failure of 1st or 2nd elements (i.e. the system goes into a state of lost integrity when either 1st, or 2nd element integrity is lost). For this case the PDF of

time between losses of system integrity is defined by expression

$$B(t) = P[min(\tau_1, \tau_2) \le t]$$

$$= 1 - P[min(\tau_1, \tau_2) \le t]$$
$$= 1 - P(\tau_1 > t)P(\tau_2 > t)$$
$$= 1 - [1 - B_1(t)][1 - B_2(t)]. \qquad (1)$$

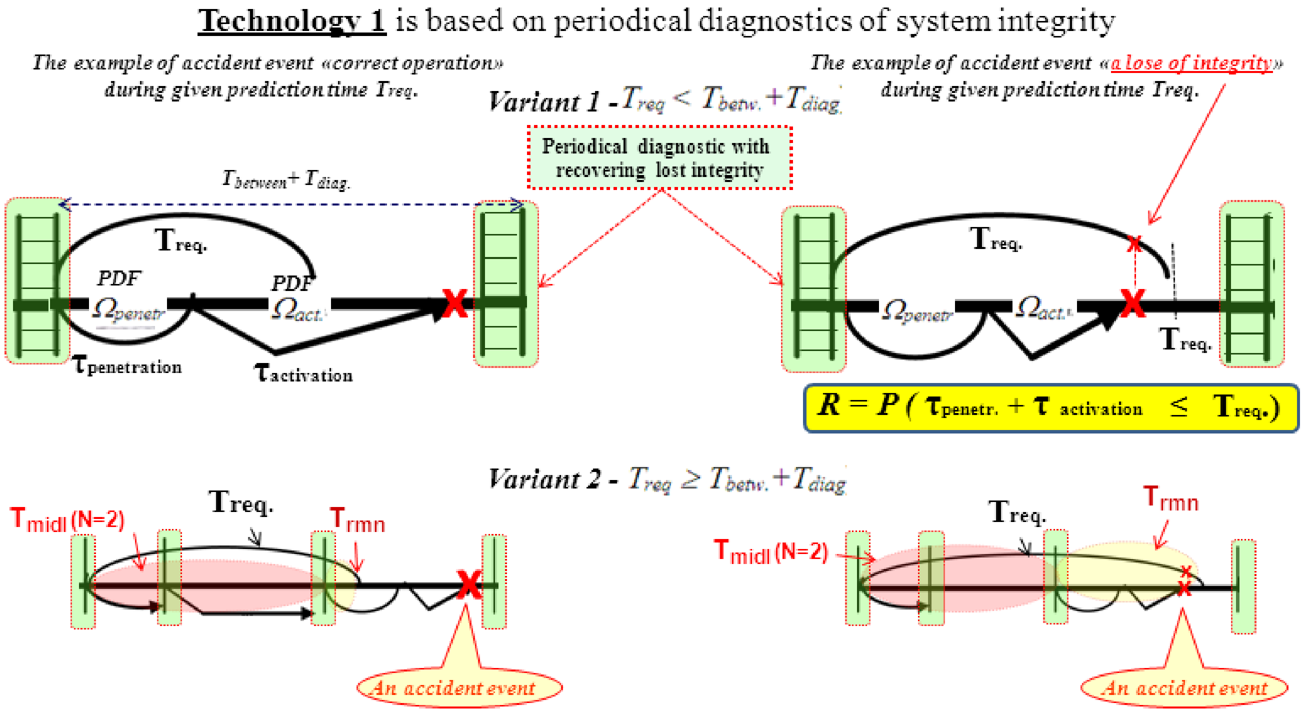

Technology 1 is based on periodical diagnostics of system integrity
The example of accident event «correct operation» during given prediction time Treq.
Variant 1 - $T_{req} < T_{betw.} + T_{diag}$
The example of accident event «a lose of integrity» during given prediction time Treq.
Periodical diagnostic with recovering lost integrity
$T_{between} + T_{diag}$
Treq.
PDF
Ω_{penetr}
PDF
$\Omega_{act.}$
$\tau_{penetration}$
$\tau_{activation}$
Ω_{penetr}
$\Omega_{act.}$
Treq.
$R = P(\tau_{penetr.} + \tau_{activation} \le T_{req.})$
Variant 2 - $T_{req} \ge T_{betw.} + T_{diag}$
Tmidl (N=2)
Treq.
Trmn
An accident event
Treq.
Trmn
Tmidl (N=2)
An accident event

Figure 4. Some accident events for technology 1 (left – correct operation, right – a lose of integrity during prognostic period T_{req})

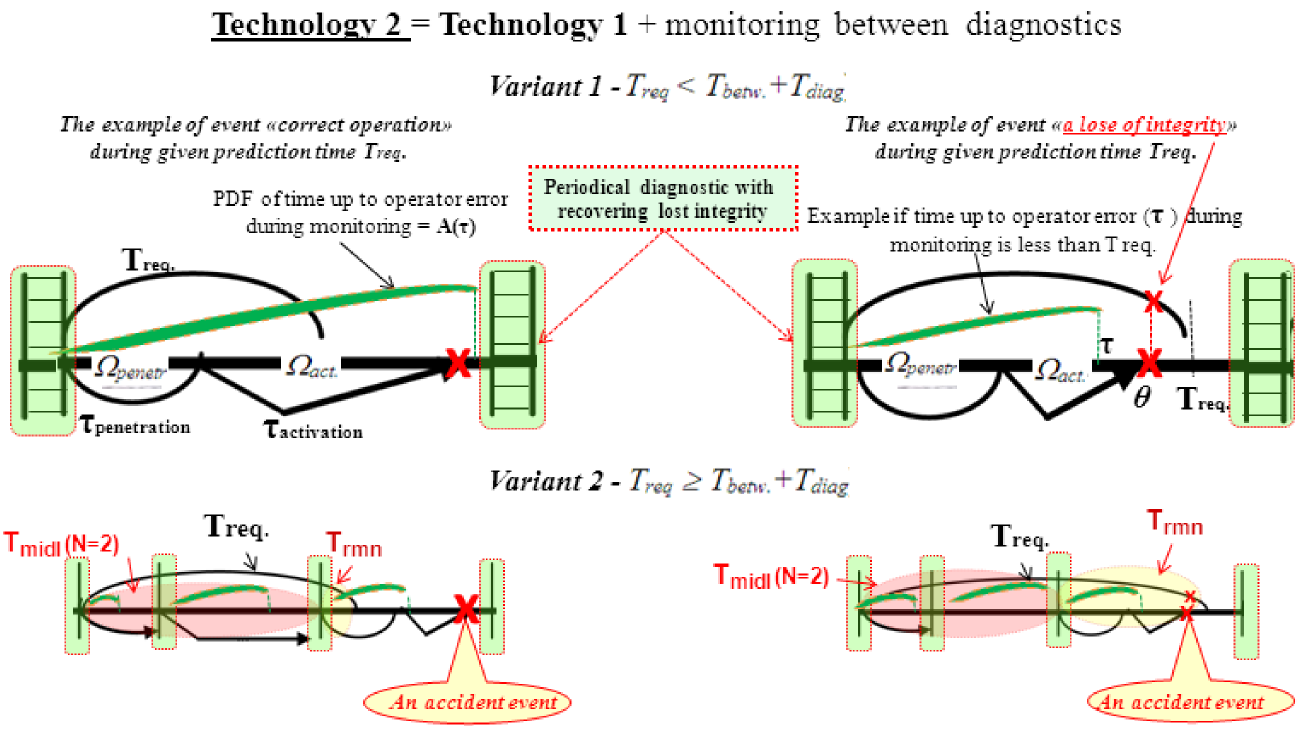

Technology 2 = Technology 1 + monitoring between diagnostics
Variant 1 - $T_{req} < T_{betw.} + T_{diag}$
The example of event «correct operation» during given prediction time Treq.
The example of event «a lose of integrity» during given prediction time Treq.
PDF of time up to operator error during monitoring = A(τ)
Periodical diagnostic with recovering lost integrity
Example if time up to operator error (τ) during monitoring is less than T req.
Treq.
Ω_{penetr}
$\Omega_{act.}$
$\tau_{penetration}$
$\tau_{activation}$
Ω_{penetr}
$\Omega_{act.}$
τ
θ
Treq.
Variant 2 - $T_{req} \ge T_{betw.} + T_{diag}$
Tmidl (N=2)
Treq.
Trmn
An accident event
Treq.
Trmn
Tmidl (N=2)
An accident event

Figure 5. Some accident events for technology 2 (left – correct operation, right – a lose of integrity during prognostic period T_{req})

2. time between losses of integrity for system combined from parallel connected independent elements (hot reservation) is equal to a maximum from two times τ_1: failure of 1st and 2nd elements (i.e. the system goes into a state of lost integrity when both 1st and 2nd elements have lost integrity). For this case the PDF of time between losses of system integrity is defined by expression

$$B(t) = P[max(\tau_1, \tau_2) \leq t] = P(\tau_1 \leq t)P(\tau_2 \leq t) = B_1(t)B_2(t). \quad (2)$$

Applying recurrently expressions (1)–(2), it is possible to build PDF of time between losses of integrity for any complex system with parallel and/or series structure and theirs combinations.

An example of complex system integrating two serial complex subsystems (abstraction) is presented by *Figure 6*.

For this integration the next interpretation of elementary events is used: complex system integrating compound components "Intellectual tructure 1 and 2" is in condition "correct operation" ("success") during given period $T_{req.}$ if during this period "AND" component "Intellectual tructure 1" "AND" component "Intellectual tructure 2" (both are special complex subsystems including IS subsystems and elements) are in condition "correct operation" ("success").

All ideas for analytical modeling complex systems are supported by the software tools, registered by Rospatent [18]–[22].

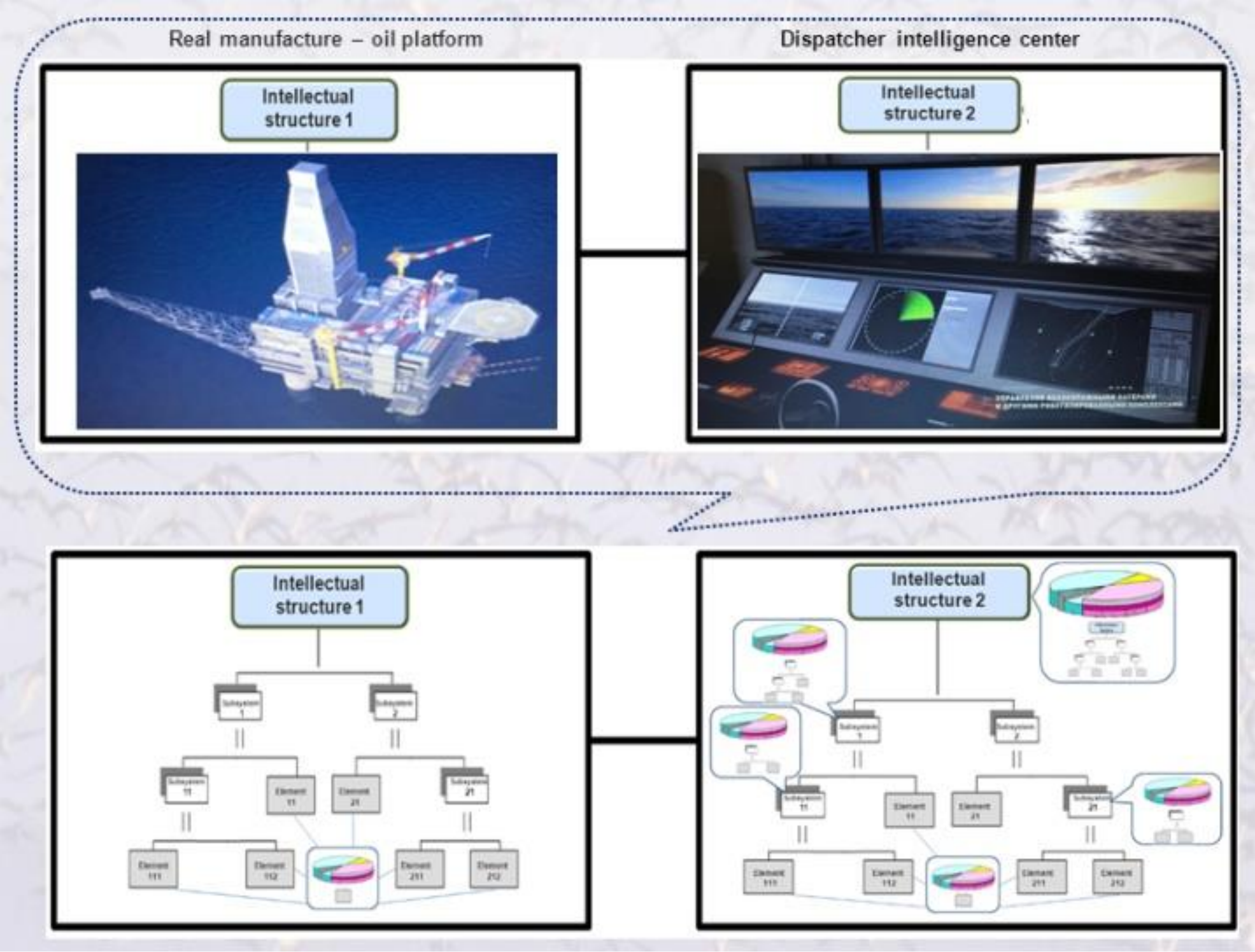


Figure 6. An example of complex system integrating two serial complex intellectual structures which also are complex subsystems (abstraction)

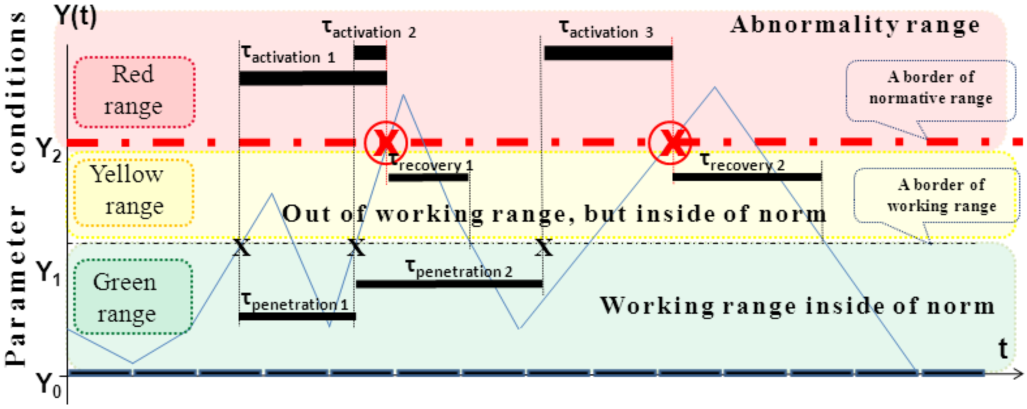


Figure 7. An example of universal elementary ranges for monitoring data about events and conditions

3.4. Input for probabilistic modeling

The proposed practical way to input forming helps to implement idea 6 for any monitored system (including real time systems).

For each critical parameter (for which prognostic estimations are needed to do actions) the ranges of acceptable conditions can be established. The traced conditions of monitored parameters are data about a condition before and on the current moment of time. For example, the ranges of possible values of conditions may be established: “Working range inside of norm”, “Out of working range, but inside of norm”, “Abnormality” for each traced separate critical parameter. If the parameter ranges of acceptable conditions are not established in explicit form than for modeling purpose the may be implead and can be expressed in the form of average time value. These time values are used as input for probabilistic modeling. For example, for coal mine some of many dozens heterogeneous parameters are: for ventilation equipment – temperature of rotor and engine bearings, a current on phases and voltage of stator; for modular decontamination equipment – vacuum in the pipeline, the expense and temperature of a metano-air mix in the pipeline before equipment, pressure in system of compressed air, etc. It may be interpreted similarly by light signals – "green", "yellow", "red" [11] – see *Figure 7* and example in Section 6.

Note. In general case the ranges may be established by subjective mode if objective one is impossible.

4. Optimization problem statements for rationale preventive actions

The proposed optimization problem statements for rationale actions helps to implement idea 7.

For example the proposed ideas 2–6 may be supported by the next typical optimization problem statements for system [11]–[12], [14], [18], [23]:

1) on the stages of system concept, development, production and support: system parameters, software, technical and control measures (Q) are the most rational for the given prognostic period if on them the minimum of expenses (Z_{dev}) for creation is reached

$$Z_{dev}(Q_{rational}) = \min_{Q} Z_{dev}, \qquad (3)$$

a) at limitations on probability of an admissible level of quality

$$P_{quality}(Q) \geq P_{adm}$$

and expenses for operation

$$C_{oper}(Q) \leq C_{adm}$$

and under other development, operation or maintenance conditions; or

b) at limitations on admissible risk to lose system integrity $R \leq R_{adm}$. and expenses for operation $C_{oper}(Q) \leq C_{adm}$ and under other development, operation or maintenance conditions; or

c) at limitations presented as combination 1a) and 1b);

2) on utilization stage:

2.1) system parameters, software, technical and control measures (Q) are the most rational for the given period of system operation if on them the maximum of probability of correct system operation is reached

$$P_{quality}(Q_{rational}) = \max_{Q} P_{quality}, \qquad (4)$$

a) at limitations on probability of an admissible level of quality

$$P_{quality}(Q) \geq P_{adm}.$$

and expenses for operation

$$C_{oper}(Q) \leq C_{adm}$$

and under other operation or maintenance conditions; or

b) at limitations on admissible risk to lose system integrity $R \leq R_{adm}$. and expenses for operation $C_{oper}(Q) \leq C_{adm}$ and under other operation or maintenance conditions; or

c) at limitations presented as combination 2.1a) and 2.1b);

2.2) system parameters, software, technical and control measures (Q) are the most rational for the given period of system operation if on them the minimum of risk to lose system integrity is reached

$$R(Q_{rational}) = \min_{Q} R(Q), \qquad (5)$$

a) at limitations on probability of an admissible level of quality

$$P_{quality}(Q) \geq P_{adm}$$

and expenses for operation

$$C_{oper.}(Q) \leq C_{adm.}$$

and under other operation or maintenance conditions; or

b) at limitations on admissible risk to lose system integrity $R \leq R_{adm}$. and expenses for operation $C_{oper}(Q) \leq C_{adm}$ and under other operation or maintenance conditions; or

c) at limitations presented as combination 2.2a) and 2.2b).

These statements may be transformed into the problems of expenses minimization in different limitations. There may be combination of these formal statements in system life cycle.

Note. Another variants of optimization problem statements are possible.

5. Incremental algorithms for solving problems in decision-making

The proposed algorithms for solving the problems in decision-making are based on using the proposed models and methods.

It is supposed that the terms "success" and accordingly "unsuccess" (“failure”) are defined in terms of admissible condition of interested system to operate for the purpose according to required quality – see the mode on *Figure 7*. For this definition a "failure" of equipment operation characterizes a threat to lose system norm integrity after danger influence (on the logic level this range “Abnormality” may be interpreted analytically as failure, fault, losses of quality or safety etc.).

The proposed steps for solving problems on the base of monitored data about events and conditions may be carried out by the next 4 steps – see *Figure 8*.

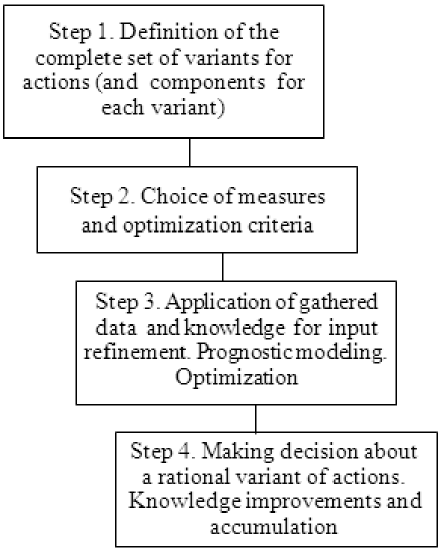

Step 1. Definition of the complete set of variants for actions (and components for each variant)
Step 2. Choice of measures and optimization criteria
Step 3. Application of gathered data and knowledge for input refinement. Prognostic modeling. Optimization
Step 4. Making decision about a rational variant of actions. Knowledge improvements and accumulation

Figure 8. Steps for solving problems

Step 1. The complete set of variants for actions should be defined, including for each variant – a definition of compound components is being. Each use case may be characterized by an expected benefit in comparable conventional units. If the objective value of a benefit can’t be defined, expert value of a level of "success" may be established, for example, on a dimensionless scale from 0 to 100 (0 – «no benefit», i.e. " failure", 100 – «the maximal benefit», i.e. complete "success").

Step 2. The measures and optimization criteria should be chosen (see Sections 3 and 4). As criteria there can be accepted:

- maximum of benefit as a result of system operation under the given conditions and limitations on the acceptable risk of “failure” and/or other limitations;
- maximum probability of "success" or minimum risk of "failure" under limitations.

Step 3. The knowledge should be used to refine the input for modeling. Using the probabilistic models and methods for each variant, the "success" measures are calculated for the given prognostic period. From a set of possible variants the rational one is chosen according to the step 2 criterion. Formal statements of optimization may be connected with maximization of benefit at limitations on admissible levels of quality and/or risks measures or with minimization of risks at limitations on admissible levels of benefit and/or quality and/or risks measures and/or under other operation or maintenance conditions.

Step 4. A decision for the optimal variant of actions (defined in step 3) should be made. In support of the efficiency of the functions, the achievable benefit calculated at step 3 is recorded. New knowledge is improved and systematized by comparing it with reality (including comparisons of probabilistic estimations and real events). A solution that meets all conditions may be not existing. In this case, there is no optimal variant of system operation on the base of monitored data about events and conditions. For this case other actions and/or criteria should be defined.

6. Examples

6.1. Period that guarantees successful intellectual system operation

The example is related to solving some problems concerning an estimation of successful system operation during given long time by IS capabilities in comparison against an usual system without or with usual sensors (i.e. without artificial intelligence capabilities to logic reasoning).

How long time may be period that guarantees successful IS operation? And what about conditions for this long period?

Those threats to IS operation which are known, traced at diagnostics and do not cause irreversible consequences at the first influence, are considered only. Besides, it is supposed, that an integrity can be operatively recovered after IS reaction at the earliest stages of detection of dangerous or guarding symptoms. Moreover, at modeling the time of full integrity recovering is artificially reduced till diagnostic time. Thus, the elementary condition “acceptable integrity” means such system state when system purposes are achieved with the required quality, it means the absence of danger source or neutralization of a penetrated source at the earliest stage prior to the its danger influence after activation. As supposed by the model it is enough for successful IS operation.

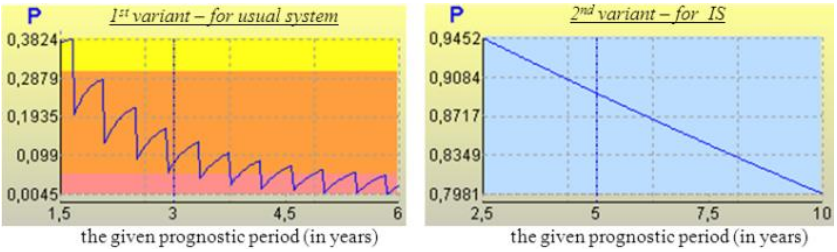


Figure 9. The probability of system integrity in dependence on the given prognostic period

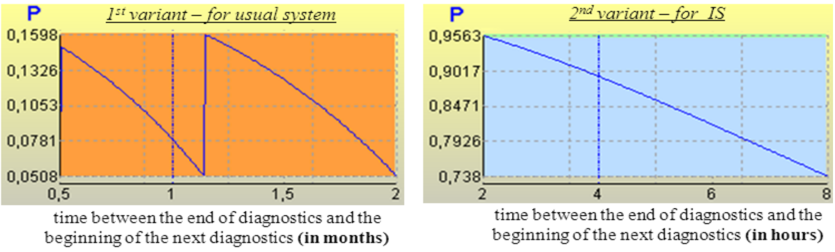


Figure 10. The probability of system integrity in dependence on the time between the end of diagnostics and the beginning of the next one

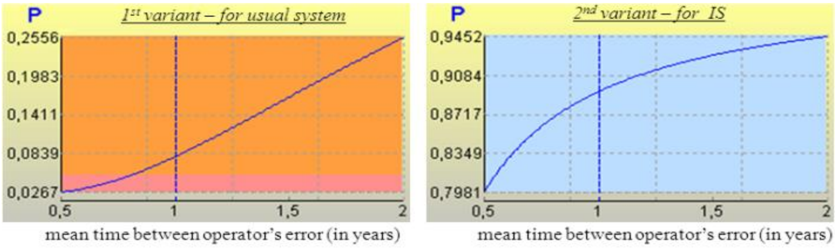


Figure 11. The probability of system integrity in dependence on the mean time between operator’s errors during continuous monitoring of IS integrity

Note. The above assumptions are supposed for modeling. In a reality it may be not always so. These conditions are considered for interpretation of modeling results.

To compare IS operation against an usual system (without artificial intelligence capabilities) for the same conditions we consider IS possibilities to provide “acceptable integrity” by continuous monitoring with artificial intelligence logic reasoning. Let's the threats to system integrity are being about 1 time a day because of natural or technogenic threats and “human factor”. Let's also after occurrence of a danger source an average activation time is equal to 6 hours, during which else it is possible to prevent or neutralize negative influence. Two variants of reaction caring of IS integrity are compared. 1st variant (for an usual system) considers the address to a recovering center about 1 time a month and reception of necessary recovering procedures within 4 hours after diagnostics. 2nd variant means IS performing functions of diagnostics every 4 hours and recovering acceptable integrity within one hour. For general technology 2 the mean time between operator’s error during continuous monitoring of system integrity is estimated not less than 1 year – see *Table 1*.

Table 1. Input for modeling

Input	Variants for comparisons	
	1st (usual system)	2nd (IS)
The given prognostic period (”in future”)	3 years	5 years
The frequency of influences for penetrating into system	1 day^{-1}	1 day^{-1}
The mean activation time	6 hours	6 hours
The time between the end of diagnostic and the beginning of the next diagnostic	1 month	4 hours
The diagnostic time	4 hours	1 hour
The mean time between operator’s error during continuous monitoring of system integrity	1 year	1 year

Some probabilities of providing system integrity in dependence on input, changing in diapason – 50% + 100% from *Table 1* data, are presented on *Figures 9–11*. They cover dependences on the given prognostic period, the time between the end of diagnostic and the beginning of the next diagnostic, the mean time between operator’s error during continuous monitoring of IS integrity. Deviations for other dependences are insignificant.

Results of modeling show that for 1st variant (for an usual system) the probability to provide “acceptable integrity” during 1 year is equal to 0.39, during 2 years – not less than 0.16, during 3 years – only 0.07. It means practically the inevitability of a failure during 2–3 years. 2nd variant (for IS) with operative recovering is more effective. Really, it is possible to provide “acceptable integrity” for system operation with IS capabilities within 3–5 years with probability about 0.90–0.93 – it may be interpreted as successful operation 9 times from 10 possible five-years periods. These results of modelling should serve a rationale for development counteractions against threats. Conditions for five-year period of successful system operation with IS capabilities are presented in *Table 1* for 2nd variant.

Note. Serrated and nonmonotonic character of dependence on *Figures 9–10* is explained by the periodic diagnostics, monitoring presence or absence and their quantitative values, and also because of parameter “N” is integer part – see *Appendix B*. Detais see in [31].

Of course the concepts “acceptable integrity” and “failure” of special system should be defined in details which produced input for modeling. However the expected modeling results against typical plausible input for this this simple example has also demonstrated a suitability of the proposed probabilistic “black box” models from Section 3.

6.2. Example of acceptable requirements to solve problems considering information quality

The example is connected with rationale a rational requirements to information quality for using IS. Information is input and output of IS operation. Also information may be used for following processing according to system purposes. This example summarizes the numerous results of researches performed for IS operating in government agencies, manufacturing structures (including power generation, coal enterprises, oil-and-gas systems), emergency services etc. [1]–[3], [7]–[8], [11]–[12], [14], [17], [23]–[35]. The results are based on the applications of proposed methods to provide quality of output information producing, quality of used information and security of IS operation (see *Appendix A*).

According to this generalization for the best practice of IS operation the acceptable requirements are the next – see the measures from *Table A.1*:

1) to provide quality of output information producing:
    - probability of providing reliable function performance during given time should be no less than 0.99;
    - system availability should be no less than 0.9995;
    - probability of well-timed calls processing during the required term should be no less than 0.95;
    - relative portion of well-timed processed calls of those types for which the customer requirements are met should be no less than 95%;
2) to provide quality of used information:
    - probability that system contains information about states of all real object and coincides should be no less than 0.9;
    - probability of information actuality on the moment of its use should be no less than 0.9;
    - probability of errors absence after checking should be no less than 0.97;
    - probability of correct analysis results obtaining should be no less than 0.95;
    - probability of providing information confidentiality during objective period should be no less than 0.999;
3) to provide security of IS operation:
    - probability of faultless (correct) operation under dangerous influence on IS during given time should be no less than 0.95;
    - probability of system protection against unauthorized access should be no less than 0.99.

Note. The special IS cases should be analyzed according to system purposes, the requirements may be specific.

These values characterizes some admissible limitations for probabilities of “success” (P) and risks of “unsuccess” ($R = 1 - P$) for information systems operation quality.

The fulfillment of these requirements is a certain scientifically proved guarantee of the quality of information used by IS.

Note. Important: the prognostic period that guarantees successful operation must match the given limitations.

6.3. Example of solving inverse problem to estimate the mean residual time before the next parameters abnormalities by periodic diagnostics

The example demonstrates IS possibility on the base of solving inverse problem by models described in Subsection 3.2 and *Appendix B*. The research results are applied to rationale actions in real time for the enterprises of coal company.

The conditions of parameters, traced by IS dispatcher intelligence center, are data about a condition before and on the current moment of time, but always the future is more important for all. With use of current data responsible staff (mechanics, technologists, engineers, etc.) should know about admissible time for work performance to maintain system operation. Otherwise because of ignorance of a residual time resource before abnormality the necessary works are not carried out. i.e. because of ignorance of this residual time it is not undertaken measures for prevention of negative events after parameters abnormalities (failures, accidents, consequences and-or the missed benefit because of equipment time out). And on the contrary, knowing residual time before abnormality these events may be avoided, or system may be maintained accordingly. For monitored critical system the probabilistic approach to estimate the mean residual time before the next parameters abnormalities for each element and whole system is proposed.

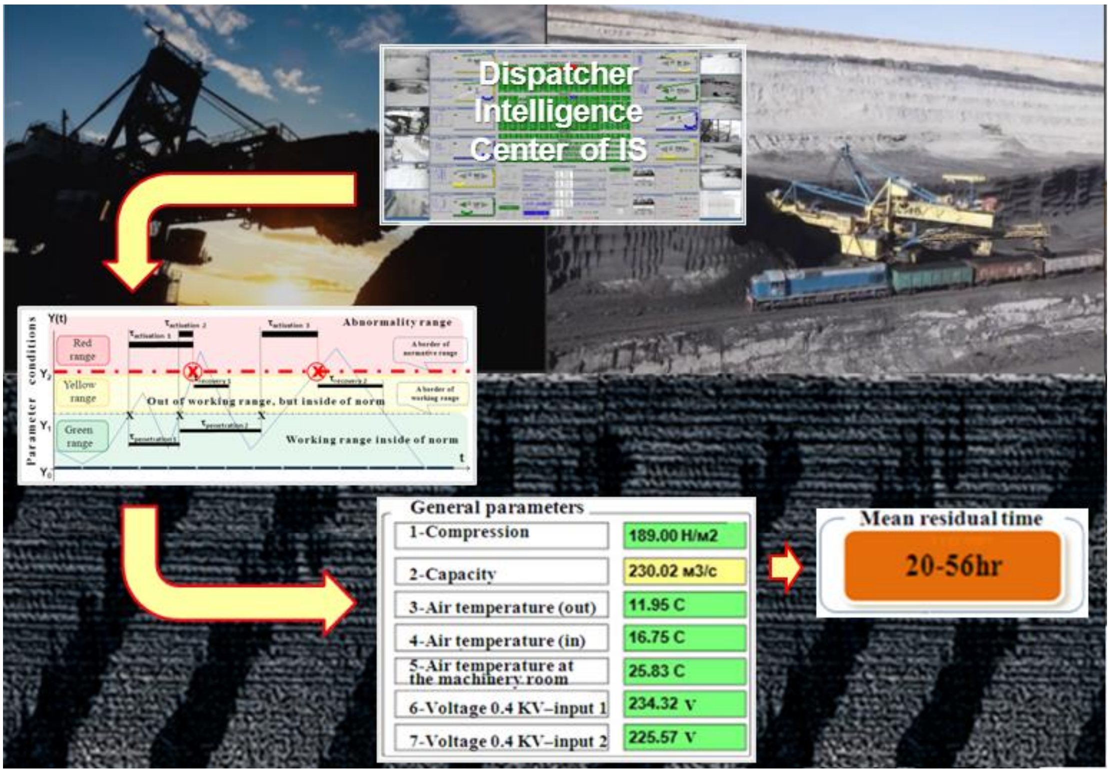
For every valuable subsystem (element) monitored parameters are chosen, and for each parameter the ranges of possible values of conditions are established: “In working limits”, “Out of working range, but inside of norm”), “Abnormality” (interpreted similarly light signals – "green", "yellow", "red") – see *Figure 7, 12*. The condition “Abnormality” characterizes a threat to lose system integrity.

For avoiding the possible crossing a border of “Abnormality” a prediction of residual time, which is available for preventive measures, according to gathered data about parameter condition fluctuations considering ranges is carried out. The approach allow to estimate residual time before the next parameter abnormality (i.e. time before first next coming into “red” range).

The estimated residual time T_{resid} is the solution t_0 of equation:

$$R(T_{penetr}, t, T_{betw}, T_{diag}, T_{req}) = R_{adm}(T_{req}) \quad (6)$$

concerning of unknown parameter t, i.e. $T_{resid} = t_0$. Here $R(T_{penetr}, t, T_{betw}, T_{diag}, T_{req})$ is risk to lose integrity, it is addition to 1 for probability $P(T_{req})$ of providing system integrity (“probability of success”), for calculations the formulas (B.1)–(B.3). T_{penetr} is the mathematical expectation of PDF $\Omega_{penetr}(\tau)$, it is defined by parameter statistics of transition from “green” into “yellow” range – see *Figure 7*.



General parameters	
1-Compression	189.00 H/м2
2-Capacity	230.02 м3/c
3-Air temperature (out)	11.95 C
4-Air temperature (in)	16.75 C
5-Air temperature at the machinery room	25.83 C
6-Voltage 0.4 KV–input 1	234.32 V
7-Voltage 0.4 KV–input 2	225.57 V

Figure 12. Example of a prognozed residual time before the next parameter abnormality

The other parameters T_{betw}, T_{diag} in (6) are known – see *Appendix B*. The main practical questions are: what about T_{req}. and what about a given admissible risk $R_{adm}(T_{req})$? For answering we can use the properties of function $R(T_{penetr}, t, T_{betw}, T_{diag}, T_{req})$:

- parameter t increases from 0 to ∞ for the same another parameters, the function $R(\dots, t, \dots)$ is monotonously decreasing from 1 to 0 (for *N* – real, i.e. no integer part), if the mean activation time of occurred danger (threat – from the 1st input at the "yellow" range to the 1st input in the "red" range) is bigger to lose integrity is less;
- if parameter T_{req} increases from 0 to ∞ for the same another parameters, the function $R(\dots, T_{req})$ is monotonously increasing from 0 to 1, i.e. for large T_{req} risk approaches to 1.

It means that the such maximal x exists when $t = x$ and $T_{req} = x$ and $0 < R(T_{penetr}, x, T_{betw}, T_{diag}, x) < 1$. I.e. the residual time before the next parameter abnormality (i.e. time before first next coming into "red" range) is equal to defined x with confidence level of admissible risk $R(T_{penetr}, x, T_{betw}, T_{diag}, x)$. So, if $T_{penetr} = 100\ days$, for $R_{adm} = 0.01$ residual time $x \approx 2.96\ weeks$ (considering decisions of recovery problems of integrity every 8 hours). Adequate reaction of responsible staff in real time is transparent for all interested parties. Details see [2], [11].

6.4. Example of solving problems for providing safety of a floating oil and gas platform

For estimation and rationale the possibilities of a floating oil and gas platform operation (considered as a system) the probabilistic modeling is being to answer the next special question: "What risks to lose system integrity may be for a year, 10 and 20 years if some subsystems are supported by special IS on the levels which are proper to skilled workers (optimistic view for future) and to medium-level workers (realistic view for now)?".

Let for studying efficiency a system is decomposed on 9 subsystems, for example – see *Figure 13.*

System components are: 1st – a construction of platform; 2nd – IS on platform for robotics monitoring and control; 3rd – an underwater communication modem; 4th – a remote controlled unmanned underwater robotic vehicle; 5th – a sonar beacon; 6th – an autonomous unmanned underwater robotic vehicle; 7th – non-boarding robotic boat – a spray of the sorbent; 8th – non-boarding robotic boat – a pollution collector; 9th – an unmanned aerial vehicle. Data is monitored from different sources and processed by the models described above in Section 3.

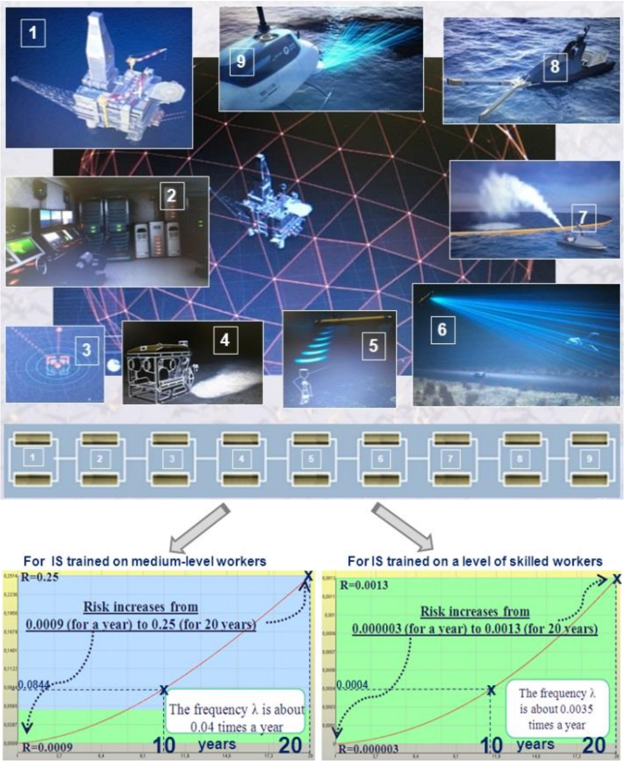


Figure 13. Subsystems operating for providing safety of a floating oil and gas platform

Note. Every subsystem also may be considered as a special complex system – see *Figure 6*. The studied structure should be focused on system purposes.

The information from monitored data and organizational data are used as input for models from *Table A.2* and performing *steps 1–4* (from *Figure 8*) in real time. Here risks to lose system integrity during given period T_{given} means risks to be at least once in state "Abnormality" within T_{req}. The functions of modeling may be performed on special servers (centralized or mapped). If virtual risks are computed for all points T_{req} from 0 to ∞, the calculated values form a trajectory of the PDF. The mathematical expectation of this PDF means the mean residual time to the next state "Abnormality". It defines mean time before failures ($MTBF$) from this PDF (see the similar calculations in 6.5). Requirements to IS operation quality should meet admissible levels recommended in Example 6.2.

To answer the question of this example let the next input are formed from data monitored and the time data of enterprises procedures. Let for every system component a frequency of occurrence of the latent or obvious threats is equal to once a month, mean activation time of threats is about 1 day. The system diagnostics are used once for work shift 8 hours, a mean duration of the system control is about 10 minutes, mean recovery time of the lost integrity of object equals to 1 day. The workers (they may be robotics, skilled mechanics, technologists, engineers etc.) are supported by capabilities of an intellectual system allowing estimations in real time the mean residual time before the next parameters abnormalities. Formally they operate as parallel elements with hot reservation. Workers are capable to revealing signs of a critical situation after their occurrence owing to the support of intellectual systems. If all subsystems are supported by intellectual systems on the level which is proper to skilled workers (optimistic view), workers can commit errors on the average not more often once a year. If all subsystems are supported by intellectual system on the level which is proper to medium-level workers (realistic view) only one difference is – medium-level workers can commit errors more often in comparison with skilled workers, for one element it is equal to 1 time a month instead of once a year.

Further the *steps 1–4* from *Figure 8* may be performed. Computed risks to lose system integrity on *Figure 13* means the risks of "failure" for every subsystem which can be detailed to the level of every separate critical parameter of equipment.

The fragments of built PDFs on *Figure 13* show:

- if all subsystems are supported by intellectual system on the level which is proper to skilled workers (optimistic view) the risk of "failure" increases from 0.000003 for a year to 0.0004 for 10 years and to 0.0013 for 20 years. The $MTBF$ equals to 283 years;
- if all subsystems are supported by intellectual system on the level which is proper to medium-level workers (realistic view) the risk of "failure" increases from 0.0009 for a year to 0.0844 for 10 years and 0.25 for 20 years. The $MTBF$ equals to 24 years. It is 11.4 times less against the results for optimistic view.

Such effects ($MTBF = 283\,years$ for optimistic view and $MTBF = 24\,years$ for realistic view) are owing to implemented technology of counteractions to threats. These are some estimations for example assumptions. Please, compare the effects against primary frequency of occurrence of the latent or obvious threats (it is equal to once a month, mean activation time of threats is about 1 day + workers errors).

6.5. Examples on quality prediction for manufacturing processes

A typical set of manufacturing processes of gas preparation equipment (GPE) on enterprise includes: processes connected with operation of entrance threads; processes of low temperature gas separations; process of economical measure of gas; processes of gas heating and reduction; processes of candle and torch separation; processes connected with storage and use methanol; processes connected with storage, supply and drainage dumps of the weathered condensation and diesel fuel; managing processes. Here are some results of modeling only for processes connected with operation of entrance threads and managing processes.

It is required to predict quality of the production processes and reliability of equipment connected with operation of entrance threads. Designations are:

- 1 – subsystem connected with the processes of gas supply through entrance threads;
- 2 – subsystem connected with gas supply in a gathering collector;
- 3 – subsystem connected with gas refining;
- 4 – subsystem connected with gas exit into technological case;
- 5 – subsystem connected with refining liquid from mechanical impurity;
- 6 – subsystem connected with the decontamination and separation of liquid;
- 7 – subsystem connected with the methanol supply for catalyzing;
- 8 – subsystem connected with the methanol supply into entrance threads;
- 1…8 – for the system (i.e. for all subsystems).

Input data for modeling are formed as average statistical data and requirements to production processes of the enterprise. Separate quality of each group of processes is estimated, then quality of productions for GPE as a whole is predicted. Let an average time of recovery of each group of processes above is equal to duration of work of one shift, i.e. 8 hours. The prognostic period is 1 month, 1 year and 5 years at observance of set modes for processes.

Note. For a preemergency conditions input data can essentially differ, that will cause also change of modeling results.

For studying the models from Section 3 are used. The results of modeling processes connected with operation of entrance threads, are analyzed on *Figure 14*.

Analysis shows: owing to recovery in time technological and production processes as a result of periodic control the mean time between failures ($MTBF$), affecting quality, increases from 1361 hours to 20431 hours, i.e. in 15 times. It is reached at the expense of timely reaction during processes control. The integral probability of performing processes, connected with operation of entrance threads with the acceptable quality, is 0.97 for a month of GPE operation, 0.70 for GPE operation during a year and 0.32 for GPE operation during 5 years. The last probability (0.32) means, that it may be real one or more accidents or failures for 5 years of GPE operation, when counteremergency measures should be performed. Risk of this is about 0.68, i.e. twice more than probability of success.

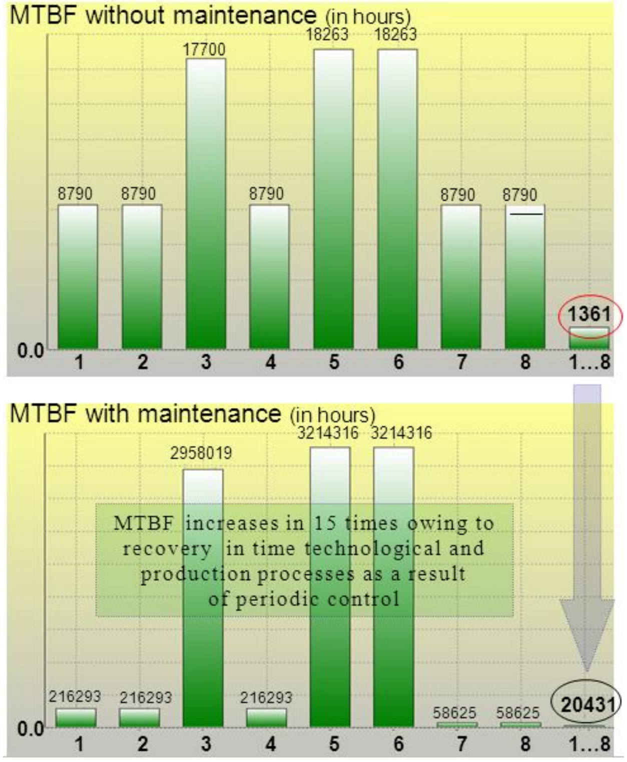


Figure 14. Prediction of quality of processes connected with operation of entrance threads

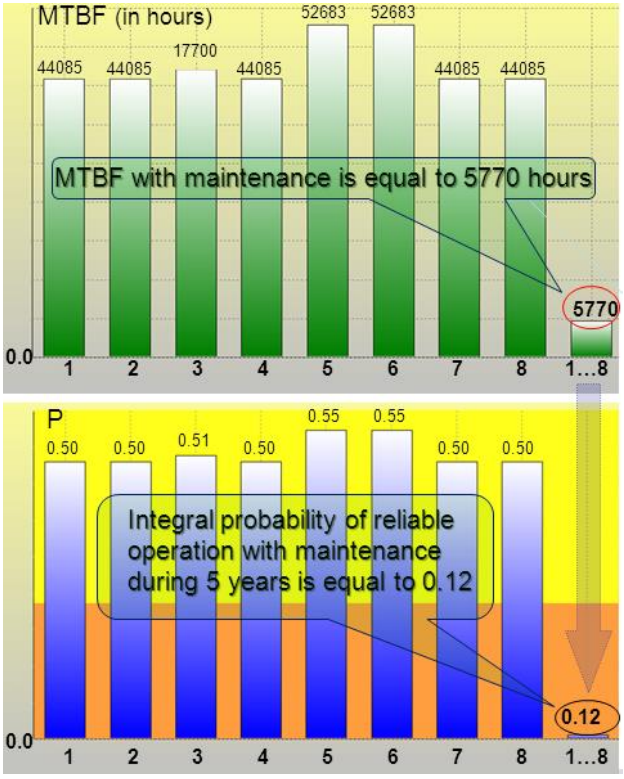


Figure 15. Predicted reliability of equipment connected with operation of entrance threads

And what about the reliability? The maintenance and diagnostic measures are performed every half a year according to recommendations of equipment suppliers. How much is it effectively for real operation conditions on the level of predicted reliability?

The results of predicting reliability of equipment connected with operation of entrance threads are demonstrated by *Figure 15*. Expected integral MTBF is equal to 5770 hours. It is 3.5 times less in comparison with 20431 hours owing to daily periodic control (see above). But the integral probability of reliable GPE operation during 5 years is only 0.12.

Summary: the account of daily results of the control and measurements is necessary. Otherwise, if to be guided by only guarantee recommendations of equipment suppliers occurrence at least one accident or failure demanding counteremergency measures of protection annually really is very possible and for 5 years it is inevitably.

6.6. Pragmatic effects for intellectual system

Author of this article took part in creation of the Complex of supporting technogenic safety on the systems of oil&gas transportation and distribution and have been awarded for it by the Award of the Government of the Russian Federation in the field of a science and technics.

The IS is a part of the created peripheral posts are equipped additionally by means of Complex to feel vibration, a fire, the flooding, unauthorized access, hurricane, and also intellectual means of the reaction, capable to recognize, identify and predict a development of extreme situations – see engineering decisions on *Figure 16*.

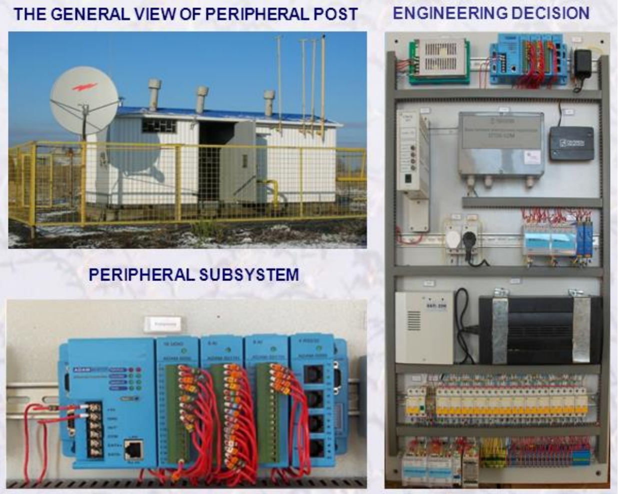


Figure 16. The IS as a hard-software part to support technogenic safety on the objects of oil&gas distribution

The applications of this Complex for 200 objects in several regions of Russia during the period 5 years have already provided economy about 8.5 Billions of Roubles. The economy is reached at the expense of effective implementation of the functions of risks prediction and processes optimization [1].

6.7. Example of implementation

The proposed ideas, probabilistic methods, models and justified normative requirements for IS are implemented in Russia at the level of national standards for system engineering. For example since 2020 the standard GOST R 58494–2019 “Mining equipment. Multifunctional safety systems of the coal mines. Remote monitoring system of dangerous industrial objects” is valid.

7. Conclusion

The proposed methods of risk predictions and their pragmatic applications in life cycle of complex systems include probabilistic models, optimization methods for rationale actions and incremental algorithms for solving the problems of supporting decision-making on the base of monitored data and rationale preventive actions in uncertainty conditions. Their suitability is demonstrated by examples which cover wide reliability and safety applications for some intellectual systems, enterprises of coal company, a floating oil and gas platform, a set of manufacturing processes of gas preparation equipment, the systems of oil&gas transportation and distribution. Effects are explained by solving the problems devoted to rationale variants for decision-making on the base of data monitored in real time and rationale preventive actions during long time period under limitations.

The proposed ideas, probabilistic methods, models and justified normative requirements for IS are implemented at the level of national standards for system engineering and are widely used in research practice and education.

Appendix A. Models to estimate intellectual system operation quality

The probabilistic models for the estimation of information systems operation quality are presented by the formulas (A.1)–(A.18) below. The proof and details – see [11], [14], [17], [23], [31].

A.1. The model of functions performance by a complex system in conditions of unreliability of its components.

Input:

$N(t)$ – is the probability distribution function (PDF) of time between neighboring failures (T_{MTBFnk} is the mean time); *W*(*t*) – is the PDF of repair time (T_{rep} is the mean time); *V*(*t*) – is the PDF of given time if this time is random value (T_{req} is the mean time).

Note. The next variants are used by the software tools [18], [21]–[22]: *N*(*t*), *W*(*t*) are exponentially distributed (i.e. enough mean times – T_{MTBFnk}, T_{rep}), $V(t)$ is determined (i.e. T_{req} is const).

Evaluated measures:

Probability P_{rel} of providing reliable function performance during given time.

$$P_{rel} = \frac{\int_0^\infty \left\{\int_t^\infty V(\tau - t)dN(\tau)\right\}dt}{\int_0^\infty td[N * W(t)]} \qquad \text{(A.1)}$$

* – is the convolution sign.

A.2. The models complex of calls processing for the different dispatcher technologies.

Input for *M*/*G*/1/∞:

λ_i – frequency of the *i*–th type calls for processing;

β_i – mean processing time of the *i*–th type calls (without queue).

Note. The software tools [18]–[22] allow to estimate and to compare effectiveness of the next dispatcher technologies for modeling by *M*/*G*/1/∞:

- technology 1 for apriority calls processing: in a consecutive order for single-tasking processing mode; in a time-sharing order for multitasking processing mode;
- priority technologies 2–5 of consecutive calls processing: technology 2 for calls processing with relative priorities in the order “first in – first out” (FIFO); technology 3 for calls processing with absolute priorities in the order FIFO; technology 4 for batch calls processing (with relative priorities and in the order FIFO inside a batch) [13], [15]; technology 5 is a combination of technologies 2, 3, 4 [16], [23].

Evaluated measures:

Probability $P_{tim.i}$ of well-timed processing of *i*–type calls during the required term $T_{req.i}$

$$P_{tim.i} = P(t_{full.i} \le T_{reg.i}) = \frac{\int_0^{\gamma_i{}^2 T_{reg.i}/T_{full.i}} t^{\gamma_i - 1} e^{-t} dt}{\int_0^\infty t^{\gamma_i - 1} e^{-t} dt} \qquad \text{(A.2)}$$

$$\gamma_i = \frac{T_{full.i}}{\sqrt{T^2_{full.i2} - T^2_{full.i}}} \; . \qquad \text{(A.3)}$$

Relative portion of all well-timed processed calls – *S* and relative portion of well-timed processed calls of those types for which the customer requirements are met – *C*:

$$S = \frac{\sum_{i=1}^{I} \lambda_i P_{tim.i}}{\sum_{i=1}^{I} \lambda_i} \qquad \text{(A.4)}$$

$$C = \frac{\sum_{i=1}^{I} \lambda_i P_{tim.i}[Ind(\alpha_1) + Ind(\alpha_2)]}{\sum_{i=1}^{I} \lambda_i}, \qquad \text{(A.5)}$$

$$Ind(\alpha) = \begin{cases} 0, & \alpha = true \\ 1, & \alpha = false' \end{cases} \qquad \text{(A.6)}$$

α_1 – (there is used criterion 1 and $T_{full.i} \le T_{req.i}$);

α_2 – (there is used criterion 2 and $P_{tim.i} \ge P_{req.i}$).

Criterion 1 is if there is required $T_{full.i} \le T_{req.i}$ to be *i*–type calls processed in time, criterion 2 is if there is required $P_{tim.i} = P(t_{full.i} \le T_{req.i}) \ge P_{adm.i}$ to be *i*–type calls processed in time, $P_{adm.i}$ – is admissible level for well-timed processing of *i*–type calls during the required term $T_{req.i}$. $T_{full.i}$

The formulas for mean response time T_{full} of *i*–type calls and for 2nd moment $T_{full.i2}$ – see [11], [13]–[14], [16], [18]–[23], [31].

A.3. The model of entering into system current data concerning new objects of application domain.

Input:

q_m – the probability that m new objects appear in random moment, intervals between these moments are exponentially distributed with parameter λ.

$\Phi(z) = \sum_{m>0} q_m z^m$ – is productive (generating) function;

$B(t)$ – is the PDF of time for new information revealing and preparing, transfer and entering into data base.

Note. The next variants are used by the software tools [18]–[22]: $\Phi(z) = z$; $B(t)$ is exponentially distributed.

Evaluated measures:

Probability P_{comp} that system contains information about states of all real object and coincides

$$P_{comp} = exp\{-\lambda \int_0^\infty [1 - \Phi(B(t))]dt\}. \qquad \text{(A.7)}$$

A.4. The model of information gathering.

Input:

C(*t*) is the PDF of time between essential changes of object states, ξ_i – is the mean time;

B(*t*) is the PDF of time for information gathering and preparing, transfer and entering into system;

Q(*t*) is the PDF of time interval between information updating, q is the mean time (only for mode D_2);

the mode D_1 of gathering: information is gathered in order “immediately after an essential object state change; the mode D_2 of gathering: information is gathered without any dependencies on changes of objects current states (including regulated information gathering).

Note. The next variants are used by the software tools [18]–[22]:

B(*t*), *C*(*t*) are exponentially distributed, $Q(t)$ is *V*(*t*) is determined or exponentially distributed.

Evaluated measures:

Probability P_{act} of information actuality on the moment of its use:

1) for the mode D_1 when information is gathered in order “immediately after an essential object state change:

$$P_{act} = \frac{1}{\xi_i}\int_0^{\infty} B(t)\,[1 - C(t)]dt; \tag{A.8}$$

2) for the mode D_2 when information is gathered without any dependencies on changes of objects current states (including regulated information gathering)

$$P_{act} = \frac{1}{q_i}\int_0^{\infty}\{[1 - Q(t)]\,\left[1 - \int_0^{\infty} C(t+\tau)\,dB(\tau)\right]. \tag{A.9}$$

A.5. The model of information analysis.

Input:

$T_{req\cdot}$ – assigned term for analysis;

$N(t)$ is the PDF of time between type I analysis errors, η^{-1} is the mean time;

$M(t)$, is the PDF of time between the neighboring errors in checked information; $A(t)$ is the PDF of analyzed type II errors, T_{MTBF} is the mean time; μ is the relative fraction of errors in information content (destined for problems of checking) or the relative fraction of information essential for analysis (destined for problems of analysis);

$T_{rel} = V/\nu$ – is the real time for complete information analysis;

V – is a content of analyzed information;

ν – is an analyzed speed;

$T_{cont\cdot}$ – is time of continuous analyst’s work.

Note. The next variants are used by the software tools [18]–[22]:

$T_{req\cdot}$ – is an assigned term (deadline) for analysis; V, ν, $T_{cont\cdot}$ and $T_{req\cdot}$ are assigned as deterministic values;

$$N(t) = 1 - \exp(-t\cdot\eta);$$
$$M(t) = 1 - \exp(-t\cdot\mu\cdot\nu);$$
$$A(t) = 1 - \exp(-t/T_{MTBF}).$$

Evaluated measures:

Probability P_{after} of errors absence after checking (probability P_{after} of correct analysis results obtaining):

Variant 1. An assigned term for analysis is no less than the real analysis time ($T_{real} \le T_{req}$) and the content of analyzed information is such small that it is required only one continuous analyst’s work period ($T_{real} \le T_{req\cdot}$).

$$P_{after(1)}(V,\mu,\nu,n,T_{MTBF},T_{cont.},T_{req.})$$

$$= \left[1 - \hat{N}(V/\nu)\right]\cdot\left\{\int_0^{V/\nu} dA(\tau)[1 - M(V/\nu - \tau)] + \int_{V/\nu}^{\infty} dA(t)\right\}. \tag{A.10}$$

Variant 2. An assigned term for analysis is no less than the real analysis time (i.e. $T_{real} \le T_{req\cdot}$). But the content of analyzed information is comparatively large, i.e. $T_{real} > T_{req\cdot}$.

$$P_{after(2)} = \{P_{after(1)}\,(V_{part(2)}\,,\,\mu,\,\nu,\,\eta,\,T_{MTBF},\,T_{cont.},\,\tau_{part(2)})\}^N, \tag{A.11}$$

$N = V/(\nu\,T_{cont.})$, $V_{part(2)} = V/N$, $\tau_{part(2)} = T_{req.}/N$.

Variant 3. An assigned term for analysis is less than the real analysis time ($T_{real} > T_{req}$) and the content of analyzed information is such small that it is required only one continuous analyst’s work period ($T_{real} \le T_{req}$).

$$P_{after(3)} = (V_{part(3)}/V)\cdot P_{after(1)}\,(V_{part(3)}\,,\,\mu,\,\nu,\,\eta,\,T_{MTBF},\,T_{cont.},\,T_{req.}) + [(V - V_{part(3)})/V]\cdot P_{without}, \tag{A.12}$$

where $V_{part(3)} = \nu T_{req.\cdot}$, $P_{without} = e^{-\mu\,(V - V\,part(3))}$.

Variant 4. An assigned term for analysis is no less than the real analysis time (i.e. $T_{real} > T_{req\cdot}$), but the content of analyzed information is comparatively large, i.e. $T_{real} > T_{cont\cdot}$.

$$P_{after} = \begin{cases} [V_{part(4)}/V]\cdot{}_{after(1)}\ (V_{part(4)}, \mu, \nu, \eta, T_{MTBF}, T_{cont.}, T_{reg.}) \\ +[(V - V_{part(4)})/V]\cdot e^{-\mu(V - V\ part(4))}, if\ T_{reg.} \le T_{cont.}; \\ [V_{part(4)}/V]\cdot\{P_{after(1)}\ (V_{part(4.2)}, \mu, \nu, \eta, T_{MTBF}, T_{cont.}, \\ \tau_{part(4.2)})\}^N \\ +[(V - V_{part(4)})/V]\cdot e^{-\mu(V - V\ part(4))}, if\ T_{reg.} > T_{cont.}. \end{cases} \tag{A.13}$$

A.6. The models complex of an authorized access to system resources during objective period.

Input (for estimation of confidentiality):

M is the conditional number of a barriers against an unauthorized access;

$F_m(t)$ is the PDF of time between changes of the m-th barrier parameters;

$U_m(t)$ is the PDF of parameters decoding time of the m-th security system barrier, u_m – the mean time of a barrier overcoming;

$H(t)$ – is the PDF of objective period, when resources value is high.

Note. The next variants are used by the software tools [18]–[22]:

$U_m(t)$is exponentially distributed;

$F_m(t)$ and $H(t)$ are determined or exponentially distributed.

Evaluated measures:

Probability P_{value} of system protection against unauthorized access during objective period

$$P_{value} = 1 - \prod_{m=1}^{M} P_{over.m} \qquad \text{(A.14)}$$

where $P_{over\ m}$ – is the risk of overcoming the m-th barrier by violator during objective period when resources value is high,

$$P_{over} = \frac{1}{f_m} \int_0^{\infty} dt \int_t^{\infty} d\, F_m(\tau) \int_0^t d\, U_m(\theta) [1 - H(\theta)].$$

A.7. The models complex of dangerous influences on a protected system.

Input:

$\Omega_{penetr}(t)$ – is the PDF of time between neighboring influences for penetrating a danger source, for $\Omega_{penetr.}(t)$=1–$e^{-\sigma t}$, σ – is the frequency of influences for penetrating;

$\Omega_{activ}(t)$ – is the PDF of activation time of a penetrated danger source, for $\Omega_{activ}(t)$=1–$e^{-t/\beta}$, β – is the mean activation time;

T_{req} – is the required period of secure system operation;

$T_{betw.}$ – is the time between the end of diagnostic and the beginning of the next diagnostic, T_{diag} – is the diagnostic time.

Note. The next variants are used by the software tools [18]–[22]:

$\Omega_{penetr}(t)$ and $U_m(t)$ are exponentially distributed.

Evaluated measures:

Probability P_{infl} of faultless (correct) operation during given time:

variant 1 – the assigned period T_{req} is less than established period between neighboring diagnostics ($T_{req} < T_{betw.}+T_{diag}$)

$$P_{infl.(1)}(T_{req}) = 1 - \Omega_{penetr} * \Omega_{activ}(T_{req}), \qquad \text{(A.15)}$$

variant 2 – the assigned period T_{req} is more than or equals to established period between neighboring diagnostics ($T_{req} \geq T_{betw.}+T_{diag}$):

$$P_{\inf l.(2)} = \frac{N\left(T_{betw.} + T_{diag.}\right)}{T_{req.}} \cdot P_{\inf l.(1)}{}^{N}\left(T_{betw.} + T_{diag.}\right) + \\ + \frac{T_{req.} - N\left(T_{betw.} + T_{diag.}\right)}{T_{req.}} P_{\inf\ l.}(T_{betw.} + T_{diag.}), \qquad \text{(A.16)}$$

where N=**[** $T_{req.}/(T_{betw.}+ T_{diag.})$**]** – is the integer part.

The models complex of an authorized access to system resources.

Input (for estimation of confidentiality):

M is the conditional number of a barriers against an unauthorized access;

$F_m(t)$ is the PDF of time between changes of the m-th barrier parameters;

$U_m(t)$ is the PDF of parameters decoding time of the m-th security system barrier, u_m – the mean time of a barrier overcoming.

Note. The next variants are used by the software tools [18]–[22]:

$U_m(t)$ is exponentially distributed;

$F_m(t)$ is determined or exponentially distributed.

Evaluated measures:

Probability P_{prot} of system protection against unauthorized access:

$$P_{prot} = 1 - \prod_{m=1}^{M} P_{over\ m}, \qquad \text{(A.17)}$$

where $P_{over\ m}$ – is the probability of overcoming the m-th barrier by violator,

$$P_{over_m} = \frac{1}{f_m} \int_0^{\infty} [1 - F_m(t)] U_m(t) dt. \qquad \text{(A.18)}$$

Note. The final clear analytical formulas are received by Lebesque-integration of (A.1), (A.10)–(A.11), (A.14), (A.18).

Appendix B. Models to predict risks for black box

The proposed models allow to estimate preventive risks for being control in real time. The approach for modeling is based on algorithmic building probabilistic models**.** The proof and details – see [14], [23], [31].

B.1. The model for technology 1 ("black box").

Note. Technology 1 (without monitoring between diagnostics) is based on periodical diagnostics of system integrity, that are carried out to detect danger sources penetration into a system or consequences of negative influences. The lost system integrity can be detect only as a result of diagnostics, after which system recovery is started. Dangerous influence on system is acted step-by step: at first a danger source penetrates into a system and then after its activation begins to influence. System integrity can't be lost before a penetrated danger source is activated. A danger is considered to be realized only after a danger source has influenced on a system.

Input:

$\Omega_{penetr}(t)$ – is the PDF of time between neighboring influences for penetrating a danger source;

$\Omega_{activ}(t)$ – is the PDF of activation time of a penetrated danger source;

$T_{betw.}$ – is the time between the end of diagnostic and the beginning of the next diagnostic,

T_{diag} – is the diagnostic time.

Evaluated measures:

Risk to lose system integrity (R).

Probability of providing system integrity (P).

$R = 1 – P$ considering consequences.

Variant 1 – the given prognostic period T_{req} is less than established period between neighboring diagnostics ($T_{req} < T_{betw.}+T_{diag}$):

$$P_{(1)}(T_{req}) = 1 – \Omega_{penetr} * \Omega_{activ}(T_{req}). \qquad (B.1)$$

Variant 2 – the assigned period T_{req} is more than or equals to established period between neighboring diagnostics ($T_{req} \geq T_{betw.}+T_{diag}$):

measure a)

$$P_{(2)}(T_{req}) = N((T_{betw}+T_{diag})/T_{req})\, P_{(1)}{}^{N}(T_{betw}+T_{diag}) + \\ + (T_{rmn}/T_{req})\, P_{(1)}(T_{rmn}), \qquad (B.2)$$

where

$$N=[T_{given}/(T_{betw.}+T_{diag.})]$$

is the integer part, $T_{rmn} = T_{given} – N(T_{betw}+T_{diag})$;

measure b)

$$P_{(2)}(T_{req}) = P_{(1)}{}^{N}(T_{betw}+T_{diag})\, P_{(1)}(T_{rmn}), \qquad (B.3)$$

where the probability of success within the given time $P_{(1)}(T_{req})$ is defined by (B.1).

B.2 The model for technology 2 (“black box”).

Note. Technology 2, unlike the previous one, implies that operators alternating each other trace system integrity between diagnostics (operator may be a man or special device or their combination). In case of detecting a danger source an operator recovers system integrity. The ways of integrity recovering are analogous to the ways of technology 1. Faultless operator’s actions provide a neutralization of a danger source trying to penetrate into a system. When operators alternate a complex diagnostic is held. A penetration of a danger source is possible only if an operator makes an error but a dangerous influence occurs if the danger is activated before the next diagnostic. Otherwise the source will be detected and neutralized during the next diagnostic.

Input:

Additionally to Input for technology 1:

$A(t)$ – is the PDF of time from the last finish of diagnostic time up to the first operator error.

Evaluated measures:

Risk to lose system integrity (R).

Probability of providing system integrity (P).

$R = 1 – P$ considering consequences.

Variant 1 – ($T_{req} < T_{betw.}+T_{diag}$):

$$P_{(1)}(T_{reg}) = 1 - \int_{0}^{T_{reg}} dA(\tau) \int_{\tau}^{T_{reg}} d\Omega_{penetr} * \Omega_{act}(\theta). \qquad (B.4)$$

Variant 2 – ($T_{req} \geq T_{betw.}+T_{diag}$):

measure a)

$$P_{(2)}(T_{req}) = N((T_{betw}+T_{diag})/T_{req})\, P_{(1)}{}^{N}(T_{betw}+T_{diag}) + \\ + (T_{rmn}/T_{req})\, P_{(1)}(T_{rmn}), \qquad (B.5)$$

measure b)

$$P_{(2)}(T_{req}) = P_{(1)}{}^{N}(T_{betw}+T_{diag})\, P_{(1)}(T_{rmn}), \qquad (B.6)$$

where N is the same and the probability of success within the given time $P_{(1)}(T_{req})$ is defined by (B.4).

References

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

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

[3] Artemyev, V., Rudenko, Ju. & Nistratov, G. 2018: Probabilistic methods and technologies of risks prediction and rationale of preventive measures by using “smart systems”. Applications to coal branch for increasing industrial safety of enterprises. *Probabilistic Modeling in System Engineering.* A. Kostogryzov (Ed.), IntechOpen, 23–51.

[4] Eid, M. & Rosato, V. 2016. Critical infrastructure disruption scenarios analyses via simulation. *Managing the Complexity of Critical*

Infrastructures. A Modelling and Simulation Approach, SpringerOpen, 43–62.

[5] Feller, W. 1971. *An Introduction to Probability Theory and Its Applications*. Vol. II, Willy.

[6] Gnedenko, B. V. et al. 1973. *Priority queueing systems*, MSU, Moscow.

[7] Grigoriev, L., Guseinov, Ch., Kershenbaum V. & Kostogryzov, A. 2014. The methodological approach, based on the risks analysis and optimization, to research variants for developing hydrocarbon deposits of Arctic regions. *Journal of Polish Safety and Reliability Association. Summer Safety and Reliability Seminars* 5(1–2), 71–78.

[8] Kershenbaum, V., Grigoriev, L., Kanygin, P. & Nistratov, A. 2018. Probabilistic modeling in system engineering. *Probabilistic Modeling Processes for Oil and Gas Systems*. A. Kostogryzov (Ed.), IntechOpen, 55–79.

[9] Kleinrock, L. 1976. *Queueing Systems, V.2: Computer Applications*, John Wiley & Sons, New York.

[10] Kołowrocki, K. & Soszyńska-Budny, J. 2011. *Reliability and Safety of Complex Technical Systems and Processes*, Springer-Verlag London Limited.

[11] Kostogryzov, A. & Korolev, V. 2020 Probability, combinatorics and control. *Probabilistic Methods for Cognitive Solving Problems of Artificial Intelligence Systems Operating in Specific Conditions of Uncertainties*. IntechOpen, 3–34.

[12] Kostogryzov, A. & Nistratov, G. 2004. *Standardization, Mathematical Modelling, Rational Management and Certification in the Field of System and Software Engineering (80 Standards, 100 Mathematical Models, 35 Software Tools, more than 50 Practical Examples).* Armament. Policy. Conversion, Moscow.

[13] Kostogryzov, A. I. & Nazarov, L. V. 1981. Batch processing of calls with relative priorities in queueing system, *News of Academy of Sciences of the USSR "Engineering Cybernetics"* 3, 194–198.

[14] Kostogryzov, A. I. & Stepanov, P. V. 2008. *Innovative Management of Quality and Risks in Systems Life Cycle (Modern Standards and Ideas of System Engineering, Mathematical Models, Methods, Techniques and Software Tools Complexes for System Analysis, Including Modelling through Internet, 100 Examples with an Explanation of Logic of the Reached Results, Useful Practical Recommendations).* Moscow, Armament. Policy. Conversion, Moscow.

[15] Kostogryzov, A. I. 1987. Conditions for efficient batch job processing of customers in priority-driven computing systems where the queueing time is construaned, *Avtomatika i Telemehanika*, 12, 158–164.

[16] Kostogryzov, A. I. 1992. Study of the efficiency of combinations of different disciplines of the priority service of calls in the computer systems, *Kibernetika i Sistemny Analiz*, 1, 128–137.

[17] Kostogryzov, A. I. 2000. Software tools complex for evaluation of iinformation systems operation quality (CEISOQ). *Proceedings of the 34th Annual Event of the Government Electronics and Information Association (GEIA)*, Engineering and Technical Management Symposium, USA, Dallas, 63–70.

[18] Kostogryzov, A. I., Bezkorovainy, M. M., Lvov, V. M., Nistratova, E. N., & Bezkorovainaya, I. V. 2000. *Complex for Evaluation of Information Systems Operation Quality – "know-how" (CEISOQ)*, registered by Rospatent №2000610272.

[19] Kostogryzov, A. I., Nistratov, G. A. & Nistratov, A. A. 2018. *Remote Analytical Support of Informing about the Probabilistic and Time Measures of Operating System and its Elements for Risk-Based Approach*, registered by Rospatent №2018617949.

[20] Kostogryzov, A. I., Nistratov, G. A. & Nistratov, A. A. 2018. *Remote Rationale of Requirements to Means and Conditions for Providing "Smart" Systems Operation Quality*, registered by Rospatent №2018618572.

[21] Kostogryzov, A. I., Nistratov, G. A., Nistratova, E. N. & Nistratov, A. A. 2004. *Mathematical Modeling of System Life Cycle Processes – "Know-How"*, registered by Rospatent №2004610858.

[22] Kostogryzov, A. I., Nistratov, G. A., Nistratova, E. N., Nistratov, A. A. 2018. *Complex for Evaluating Quality of Production Processes*, registered by Rospatent №2010614145.

[23] Kostogryzov, A. I., Petuhov, A. V. & Scherbina, A. M. 1994. *Foundations of Evaluation, Providing and Increasing Output Information Quality for Automatized System*. Moscow: Armament. Policy. Conversion.

[24] Kostogryzov, A. I., Stepanov, P. V., Nistratov, G. A., Nistratov, A. A., Grigoriev, L. I. & Atakishchev, O. I. 2015. *Innovative Management Based on Risks Prediction, Information Engineering and Education Science* – Zheng (Ed.). Taylor & Francis Group, London, 159–166.

[25] Kostogryzov, A., Atakishchev, O., Stepanov, P., Nistratov, A., Nistratov, G. & Grigoriev, L. 2017. Probabilistic modelling processes of mutual monitoring operators actions for transport systems. P*roceedings of the 4th International*

Conference on Transportation Information and Safety (ICTIS), Canada, Banff, 865–871.

[26] Kostogryzov, A., Grigoriev, L., Golovin, S., Nistratov, A., Nistratov, G. & Klimov, S. 2018. Probabilistic modeling of robotic and automated systems operating in cosmic space. *Proceedings of the International Conference on Communication, Network and Artificial Intelligence (CNAI)*, Beijing, China. DEStech Publications, Inc., 298–303.

[27] Kostogryzov, A., Grigoriev, L., Kanygin, P., Golovin, S., Nistratov, A. & Nistratov, G. 2018. The experience of probabilistic modeling and optimization of a centralized heat supply system which is an object for modernization. *Proceedings of International Conference on Physics, Computing and Mathematical Modeling (PCMM)*, Shanghai, DEStech Publications, Inc., 93–97.

[28] Kostogryzov, A., Nistratov, A., Nistratov, G., Atakishchev, O., Golovin, S. & Grigoriev, L. 2018. The probabilistic analysis of the possibilities to keep "organism integrity" by continuous monitoring. *Proceedings of the International Conference on Mathematics, Modelling, Simulation and Algorithms (MMSA)*, Chengdu, China. Atlantis Press, Advances in Intelligent Systems Research, volume 159, 432–435.

[29] Kostogryzov, A., Nistratov, A., Zubarev, I., Stepanov, P. & Grigoriev L. 2015. About accuracy of risks prediction and importance of increasing adequacy of used adequacy of used probabilistic models. *Journal of Polish Safety and Reliability Association. Summer Safety and Reliability Seminars* 6(2), 71–80.

[30] Kostogryzov, A., Nistratov, G. & Nistratov, A. 2013. The innovative probability models and software technologies of risks prediction for systems operating in various fields. *International Journal of Engineering and Innovative Technology (IJEIT)* 3(3), 146–155.

[31] Kostogryzov, A., Nistratov, G. & Nistratov, A. 2012. Some applicable methods to analyze and optimize system processes in quality management. Total quality management and Six Sigma, *InTech*, 127–196.

[32] Kostogryzov, A., Panov, V., Stepanov, P., Grigoriev, L., Nistratov, A. & Nistratov, G. 2017. Optimization of sequence of performing heterogeneous repair work for transport systems by criteria of timeliness. *Proceedings of the 4th International Conference on Transportation Information and Safety (ICTIS)*, Canada, Banff, 872–876.

[33] Kostogryzov, A., Stepanov, P., Grigoriev, L., Atakishchev, O., Nistratov, A. & Nistratov, G. 2017. Improvement of existing risks control concept for complex systems by the automatic combination and generation of probabilistic models and forming the storehouse of risks predictions knowledge. *Proceedings of the 2nd International Conference on Applied Mathematics, Simulation and Modelling (AMSM)*, Phuket, Thailand. DEStech Publications, Inc., 279–283.

[34] Kostogryzov, A., Stepanov, P., Nistratov, A. & Atakishchev, O. 2017. About probabilistic risks analysis during longtime grain storage. *Proceedings of the 2nd Internationale Conference on the Social Science and Teaching Research (ACSS–SSTR)* 18 (Advances in Social and Behavioral Science. Singapore Management and Sports Science Institute), PTE.Ltd., 3–8.

[35] Kostogryzov, A., Stepanov, P., Nistratov, A., Nistratov, G., Klimov, S. & Grigoriev, L. 2017. The method of rational dispatching a sequence of heterogeneous repair works. *Energetica* 63(4), 154–162.

[36] Martin, J. 1972. *System Analysis for Data Transmission*. V. II, IBM System Research Institute. Prentice Hall, Inc., Englewood Cliffs, New Jersey.

[37] Matweev, V. F. & Ushakov, V. G. 1984. *Queuing systems*. MSU, Moscow.

[38] Silhavy, R., Silhavy, P. & Prokopova, Z. 2017. Cybernetics approaches in intelligent systems. *Computational Methods in Systems and Software* Vol. 1, Springer.

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