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

Dependability,
DOI:10.21683
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 17.06.2020, Volume , Format 60x90/8, Paper gloss

The Journal is published quarterly since 2001.
The price of a single copy is 1045 Rubles, an annual subscription costs 4180 Rubles.
Phone: +7 (495) 967 77 05.
E-mail: dependability@bk.ru.

Papers are reviewed.
Papers are published in author's edition.

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

https://doi.org/10.21683/1729-2646-2020-20-2-3-11

Dependability in digital technology

Yuri P. Pokhabov, *Joint Stock Company NPO PM – Maloe Konstruktorskoye Buro (AO NPO PM MKB), Zheleznogorsk, Krasnoyarsk Krai, Russian Federation*
pokhabov_yury@mail.ru

Yuri P. Pokhabov

Abstract. Aim. *The migration towards the Industry 4.0 digital technology will soon enable "right first time" (virtually with no material expenditures for experimental testing and subsequent design improvement) creation of increasing numbers of entities with unique application properties. Calculating the dependability indicators of such entities based on reliable statistical data will be greatly challenging. However, the need for dependable entities will remain. Additionally, the approaches to digital technology based on physical models and engineering knowledge enable the creation of predictive dependability methods (based on the assumption of non-acceptability or, contrarily, intentional programming of failures). That inevitably causes a paradigm shift in the modern dependability theory associated with a forced deviation from the mathematical models as the basis of the dependability theory.* **Methods.** *According to the Russian tradition, dependability is normally defined by specifying the required functions through a set of parameters that characterize the ability to perform them and the allowable variation limits of the parameter values. If the criteria of some required functions cannot be specified through parameters, a technique can be used, whereas the operation of the item is substituted with an information model in the form of a black box, in which the performance of the required functions is characterized by probabilistic indicators of failures (statistical, logical, Bayesian, subjective). In order to account for the parameters and probabilities of performance of the required functions in a coordinated manner, finding the values of the parameters within the allowed range can be characterized by the probability as the degree of confidence in the occurrence of such event, for example accounting for design reserves. In this case the performance of all the required functions can be characterized by an additive dependability indicator that is identified using the method of dependability structure diagram. This indicator completely characterizes the predicted dependability level.* **Results.** *Predicted dependability is estimated using the method of design engineering analysis of dependability (DEAD). This method allows using a set of algorithm-based techniques to present the design (per GOST 2.102) and process control (per GOST 3.1102) documentation for a technical item in the form of a generalized parametric model of operation. Such model allows taking into consideration the individual specificity of the design of entities based on the unity of functionality, operability and dependability, and thereupon estimating the probability of failures. DEAD and digital design algorithms are completely compatible and driven by common problems related to the substantiation of design solutions for the purpose of elimination (reduction of probability) of errors able to cause failures based on analytical, computational and experimental verification.* **Conclusions.** *Digital technology provides a tangible opportunity of predicting, reducing the impact or eliminating possible failures. That can be achieved through the same means that often cause failures, i.e. design engineering. For that purpose, it is required to create new applications of the modern dependability theory based on engineering disciplines and design engineering methods developed for ensuring quality and dependability of entities.*

Keywords: *digital technology, dependability theory, dependability prediction, unique highly vital system, design engineering analysis of dependability (DEAD).*

For citation: *Pokhabov Yu.P. Dependability in digital technology. Dependability. 2020;2: 3-11. https://doi.org/10.21683/1729-2646-2020-20-2-3-11*

Received on: *24.01.2020* / **Upon revision:** *17.04.2020* / **For printing:** *17.06.2020.*

Introduction

During the closing of the MMR-2004 conference in Santa Fe (US), a discussion titled "Is Reliability Theory Still Alive?" was held that defined the theme of Igor A. Ushakov's article that concluded: "The need for pure theory may be not as pressing as it used to be, yet the need for applications of the dependability theory for solving practical tasks was, is and always will be!" [1]. The list of problems that can be solved through the development of new applications of the dependability theory was published four years ago at the plenary meeting of the MMR-2000 conference in Bordeaux (France) in the presentation titled "Dependability: past, present, future" [2]. Despite the efforts made, some of the mentioned problems of the dependability theory are still unsolved, including, for instance, dependability of unique highly vital systems (entities, items) [3].

Over the past years, the fourth industrial revolution added new unsolved problems to that list [4]. Todays' generation of engineers can hardly imagine the technological changes caused by the results of this revolution, but we must start preparing for it right now. Meanwhile, as a forerunner of the predicted future, new directions in engineering have emerged and have been developing: system engineering ("right first time" design) [5] and Industry 4.0 digital technologies ("right first time" entity development) [3]. A trend is becoming popular, whereas the dependability indicators in digital engineering are not considered as a target. It is thought that if target values of operational integrity and resource limitations (time, financial, technological, industrial, etc.) are achieved, the dependability is ensured by default [5, 6]. For example, the residual life of an entity can now be set and defined in an explicit form (parametric form) according to the results of numerical simulation of physical processes resulting in its loss. The term "dependability" becomes blurry: dependability seemingly still exist in digital technology (it still needs to be ensured), but it is not clear how to control the dependability indicators (most processes related to the development and optimization of entities are transferred into a virtual computing environment, production of material objects for – primarily – experimental testing is minimized, and the mathematics of the modern dependability theory are not adapted to this). However, the most important is that the basic principle of the modern dependability theory, that, according to Alexander S. Pronikov, consists in the *statement of a certain level of dependability for a machine with expired service life* [7] is not good for anyone even today.

If, as part of finding solutions to future problems, the approaches to dependability do not change, then for the next generation of engineers the value of the modern dependability theory may die down, as it does not contribute to the development of new engineering ideas. Nothing can be done about it, it has been and remains: "*We know that every age has its own problems, which the following age either solves or casts aside as profitless and replaces by new ones*" [8]. And it's time to ask a question much more dramatic than in the early 2000s: Do we really need a dependability theory in the digital age?[1] By asking such a provocative question, the author does not in any way reject the dependability theory, strength of materials or any other engineering (technical) disciplines, and therefore asks another question: What requirements should the dependability theory applications and other engineering disciplines meet in the digital age?

Problems of the dependability theory for digital technologies

Since any technical entities are developed by engineers working on computers, it would be fair to suppose that in the digital era the skills and knowledge related to calculating formulas are also necessary and important, just like before the advent of computer technology (at least for the purpose of evaluating the effectiveness of own activity). The more so since a computer is just a high-performance calculation device (whether it is used for drafting or finite-element calculations). When solving engineering problems, a computer does not independently search for areas where the end result should be, but only performs specified calculations using established algorithms. The human prerogative is to apply the available computing resources to the required area to obtain the most optimal result by setting the appropriate initial data [10]. Without the knowledge of the principles of natural science, engineering disciplines and the ability to do elementary engineering calculations, that problem can hardly be solved adequately. The more so since (according to one of the definitions in GOST 27.002–2015) dependability is intended to be somewhat of the pinnacle of engineering (in order to define "**the values of all parameters** *that characterize*… **required functions**"), and in the age of computers dependability acquires even greater significance, not the other way around. First, the items' ability to perform the required function with specified dependability will remain the main goal of any development. Second, failures in operation (depending on the purpose of the items) must become predictable (unacceptable or, conversely, intentionally programmed). Digital technologies are intended exactly for that, i.e. to simulate adverse events and thereby enable the selection of optimal results. In the author's opinion, new applications of modern dependability theory should be applied to those problems in order to prove useful in the implementation of digital technology.

[1] Though this question might seem unreasonable, today the idea is seriously discussed that a modern engineer who is involved with computing does need no knowledge on strength of materials as that will be replaced by software [9].

Barriers of modern dependability theory on the way of digital technology

Computer calculations are performed according to the established algorithms (exact requirements that define the sequence of elementary operations with initial data), be it simple arithmetic operations or numerical solutions of differential equations. It is impossible to directly calculate dependability using computer calculations, since it cannot be expressed, calculated or measured using physical values, primarily, due to the multifactorial and interdisciplinary nature of the causes of possible failures that cannot be algorithmized. For this reason, before the age of computers, special mathematics were developed for the purpose of calculating dependability indicators; they allow identifying dependability using a posteriori knowledge about possible failures, i.e., in fact, through the experience of undependability. The result is an endless vicious circle, whereas it is required to know a technical item's undependability to calculate its dependability. The availability of appropriate failure statistics makes it easy; difficulties start when there is no database to obtain failure statistics, for example, if there are no prototypes, the items are one-of-a-kind (unique)[1] or failures are unacceptable under the operating conditions. There are no regulations or guidelines on dependability that could provide guidance on what to do in this case.

The Reference Annex of GOST 27.002-89[2] explicitly states that the area of dependability indicators calculations (according to the rules of the statistical theory of dependability) is limited to large-series items only. For unique and small-series items, calculations using methods of the statistical dependability theory are limited to only the cases when the dependability indicators can be calculated according to known dependability indicators of components and elements. In [11] the feasibility is substantiated, but it requires data on the dependability of components and elements that can be obtained from statistical tests in the amount of parent universe, which, for example, is almost impossible in the case of unique highly vital systems due to financial limitations [12].

With the deployment of digital technology, the problem of dependability calculation based on statistical dependability theory is exacerbated, since the number of test (engineering) models used in the development and commencement of product manufacture will inevitably reduce due to the virtualization of real processes (actions, inspections and tests) related to material objects [6]. Thus, the basis of statistical methods of the dependability theory, i.e. accumulation and processing of statistical information on item failures, disappears due to the application of digital technology. There is no point in setting probabilistic dependability indicators as input data for digital computing not only from the standpoint of the fundamental principles of the statistical dependability theory (due to the lack of information on failures). The probability itself is not subject to direct computer calculations, as it is a numerical measure of the events that is independent from the algorithm of their occurrence.

Nevertheless, if it is possible to calculate the value of a certain parameter and correlate it with limit permissible parameters, then we can talk about the probability of finding the value of this parameter within a permissible region (as a degree of confidence in the occurrence of such event[3]). If a technical item can be represented with a set of parameters and permissible limit values, then this makes it possible to identify its dependability as an additive indicator that characterizes the performance of the required intended function when modeling possible scenarios of events in operation (essentially, as an indicator of predicted dependability). It is not difficult for modern computers to calculate the favorable (unfavorable) outcomes, provided an appropriate calculation algorithm is defined. However, modern applications of the dependability theory do not provide such algorithms.

Current quality of the solution of highly vital item dependability problems

The barriers of the modern dependability theory on the way of digital transformation equally impede the development of highly vital items without prototypes. The lack of the required statistical data and difficulty of calculation of dependability indicators lead to the realization of the fact that calculations themselves only serve an auxiliary function in the adoption of engineering solutions in the course of development, leaving the leading role to the methods of expert assessment and verification, which is reflected in foreign regulatory and literary sources:

• *NOTE The "probability of failure" and its corresponding reliability index are only notional values that do not necessarily represent the actual failure rates but are used as operational values for code calibration purposes and comparison of reliability levels of structures.* This is one of the explanations in the Eurocode EN 1990:2002 standard of the European design system;

---

[1] A one-of-a-kind (unique) product is a product that is one of a kind in terms of its design or unique in its extreme rarity/significance [OST 134-1032–2003, article 3.1].

[2] GOST 27.002–89 was cancelled in 2017 but the Reference Annex can be considered as a separate source, since it was written based on 12 publications, most of which constitute the very foundations of the modern dependability theory.

[3] Probability is a real number ranging from 0 to 1 related to a random event. Note: the number may indicate a relative frequency in a series of observations or the degree of confidence that some event will occur. The probability is close to 1 for a high degree of confidence [GOST P 50779.10–2000, article 1.1].

• *…it is more important to identify and, if possible, mitigate the consequences of failure modes by design measures than to know the probability of their occurrence.* That is an explanation of the definition of failure modes in IEC 60812:2006;

• "*…all methods of reliability assessment require expert evaluation. When we approach that, the probability values are much like a label that an engineer put on a structure to show what he thinks about its reliability*", said Charles Harlan, former Director, Safety, Reliability and Quality Assurance of the Space Shuttle program [13].

The above views were put into practice in the NASA and ESA standards, where calculations are part of the processes of analytical and experimental verification of rocket and space technology. However, in practice, the results of such verification still leave much to be desired. For example, after the crash of the STS-51L Challenger Shuttle, the application of one of the main tools of analytical verification (the FMEA method) was sharply criticized in the US engineering circles [13]. According to the results of preliminary analysis of possible failures and their consequences, only one out 10.000 flights was supposed to end in a crash. However, in practice, two shuttles crashed as a result of 135 flights (Challenger in 1986 and Columbia in 2003). That constitutes an unprecedented catastrophic error in the practice of FMEA application: the actual fail-safety was 0.985 instead of the predicted 0.999 9. A similar result follows from the 2009 – 2016 failure statistics of deployed structures on foreign and Russian spacecraft. The average fail-safety of deployment mechanisms did not exceed 0.996 instead of the permissible fail-safety of at least 0.999 5 (with reservations assuming that this assessment is overestimated due to incomplete failure statistics) [14]. It should be taken into account that, in practice, in each case the results of dependability calculation (verification) in accordance with the current regulations must confirm the above permissible fail-safety, otherwise the spacecraft would not have been launched due to design insufficiencies. The investigation of the real causes of failures was carried out to clarify why the results of dependability assessment do not correspond to reality. The investigation results showed that in most cases the causes are rare in their nature that, in turn, is defined by an unfavorable combination of manufacturing tolerances, unaccounted factors of technological heredity, as well as external effects that today's dependability verification methods do not consider [14]. It was also revealed that for highly vital products, any rare cause of failure can reduce the accuracy of the dependability assessment, while, in practice, the total calculation error can reach at least a magnitude order of the significant figure[1], which is confirmed by the above examples.

Approach to predicting the dependability of highly vital items

Let us assume that the operation of any item can be represented with a set of parameters, the values of which can vary within the given ranges (i.e. in strict accordance with one of the definitions of dependability). Each of these parameters is considered from the standpoint of resilience to possible failures under external effects that, in turn, determine the limits of value variation of the analyzed parameters [15]. In this case, combining the effect and resilience parameters it is possible to build a fail-safety operation model based on physical laws that takes into account the temporal variation of the limit values of the considered parameters. Such model, as opposed to mathematical models of the dependability theory is suitable for predicting dependability (an example of a similar model for spacecraft rotating rod is provided in [16]). In such model, the list of these parameters characterizes the item's functionality (properties determined by the presence and set of capabilities to perform the required functions), the specified range of parameter values variation characterizes its operational integrity (a state in which an item can perform the required functions), and the probability of the parameter values being within the given range during operation characterizes the dependability (the ability to maintain the performance of the required functions in specified modes and conditions of operation) [16].

Based on the fact that all item failures occur due to the physicality (causal connections) and physical necessity (consistency with the laws of nature) of the causes that generate them (whether we know these causes or not), then based on the knowledge of the laws of physics, it is possible to build a parametric model of the item operation that determines its functionality, operational integrity and dependability based on a single database of parameters and ranges of their permissible values. The construction of such model is based on the knowledge of the physical principles of nature at the levels of micro world (the world of elementary particles, atoms, molecules and molecular compounds), macro world (the world of persistent forms and values commensurate to human) and the mega world (the surrounding world commensurate to the universe). The values of the parameters of the parametric operation model are calculated by known methods of engineering disciplines, i.e. the theory of mechanisms and machinery, theory of theoretical mechanics, material resistance, machine components, etc.

If there are not enough knowledge and understanding at any level of the world structure to calculate the values of the parameters of the parametric model of item

[1] By analogy with engineering calculations, this corresponds to the accuracy of the sought result not by percentage points (usually, the error is 5÷10%), not even several times (for example, two or three times), but by orders of magnitude, i.e. not less than ten to a hundred times (!).

operation, it is possible to use the well-known technique, according to which the operation of any of the components of the item is replaced with the information model in the form of a black box where the performance of the required functions is characterized by probabilistic failure rates (statistical, logical, Bayesian, subjective). It is necessary to bring the values of parameters and probabilistic indicators to a consistent nondimensional form in order to take into account the probabilistic indicators of such information models and calculate the dependability using a generalized parametric model of item operation. For this purpose, the probability of the parameter values being within the acceptable range is identified (based on their physical understanding [16]), upon which all probabilities regardless of their origins (based on physical or information models) will be available to calculate the dependability using the method of structural dependability scheme [14, 16]. At the same time, this does not contradict the idea of calculating the dependability of unique and small-series items according to known dependability indicators of components and elements.

Two interchangeable methods can be used to determine the probabilities of parameter values being within the allowable range: deterministic (setting the design margins for each of the parameters in such a way as to guarantee with certain confidence that their values are within the allowable range) [14] and stochastic (for example, by assessing the individual structural dependability [17], i.e. calculating the probabilities of parameters being within the allowable areas based on individual characteristics of materials, loading/impact processes and product manufacturing processes). The interchangeability of these methods can be explained through the example of a strength calculating model of the "load parameter – strength parameter" type, whereas the probability of failure-free operation equals the probability that the value of the load parameter will never exceed the value of the strength parameter within a given period of time. Moreover, even if both parameters are random functions of time, it is possible to solve the dependability problem in a deterministic statement of the calculated values of the load and safety margins [16] by applying structural margins according to GOST R 56514–2015, i.e. by "expanding" the range of real values of the "load parameter" with safety factors and/or "narrowing" the allowable range of the "strength parameter" using safety margins. This method is widely used in the rocket and space industry.

Examples of structural margins used in practice in the form of redundancy, safety factors, safety margins and drive torque (forces), parametric redundancy, power and thermal decoupling, procedures for obtaining guaranteed results, for example, using minimax criteria or engineering psychology factors, are provided in [14, 16]. All structural margins are assigned based on the rules of the statistical dependability theory (for example, safety factors and safety margins [18]), proven application practices (for example, margins of drive torque (forces) [14, 19–20]), design methods aimed at removing limitations on output parameters variation (for example, by using power and thermal decoupling [21–22]), or other organizational and technical actions that reduce or eliminate the probability of failures.

In the general case, for example, for deploying structures of spacecraft, the dependability in terms of strength can be calculated using the deterministic method according to GOST R 56514–2015, and dependability in terms of operation can be calculated using the stochastic method [20], or in any other combinations [14, 17]. Furthermore, the use of structural margins for solving dependability problems in a deterministic formulation not only simplifies the selection and substantiation of parameters when designing items, it is also one of the important conditions for compiling the initial data for digital design in the form of a matrix of target indicators and their limitations [6].

Design engineering methods for solving the dependability problems of highly vital items

Various aspects of the above parametric approach to solving the dependability problems of highly vital items (philosophy, genesis, definitions, theoretical issues, models, calculations, practical applications, etc.) were considered in detail in [14]. They served as the basis for the design of the method of engineering analysis of dependability. That technique, relying both on engineering disciplines and the mathematical foundations of the dependability theory (if acceptable and justified), allows analyzing and taking into account individual design features of products, which makes it possible to predict dependability in the design and construction of technical objects without prototypes.

DEAD is based on a generalized parametric model of operation in the form of [16]

$$\{X_i\} = \left(X_1, X_2, ..., X_i\right)^T \ \forall i = \overline{1, n}; \tag{1}$$

$$D_x = \left\{X_i\left(t\right) | \alpha_i \le X_i\left(t\right) \le \beta_i\right\}; \tag{2}$$

$$R = P\left\{X_i\left(t\right) \in D_x, 0 < t < t_к\right\}, \tag{3}$$

where $\{X_i\}$ is a set of output parameters X_i, that determine the performance of the required functions in the form of a column-vector (functionality of the object); D_x is the acceptable region of output parameters $X_i(t)$ (operational integrity of the object in the permissible ranges of parameter values α_i and β_i); R is the dependability of the object as the probability P of values of the output parameters $X_i(t)$ being within the region of their permissible values of D_x within the time to failure $t_к$.

DEAD is a sequential set of algorithmic methods that allow presenting the design (in accordance with GOST 2.102) and process engineering (in accordance with GOST 3.1102) documentation of a technical item (i.e. its text-and-graphic or digital model depending on the development method) in the form of a generalized parametric model of operation (1) – (3). The procedures of the technique allow (in a generalized form):

• initialization of the item in the form of parameterization (turning it into a set of parameters and permissible ranges of their variation), which is carried out to establish conditions (1) – (2);

• calculations of theoretical dependability by design parameters carried out according to (3);

• providing evidence that the analysis (assessment) of dependability corresponds to the reality (the requirements of design and technological documentation, production conditions, quality control methods), for which the relevant risks assessment is carried out [23].

The application of the generalized parametric model of operation (1) – (3) and the DEAD [16] does not violate the basic principles of dependability theory. Along with the applied methods of the dependability theory (mathematical, statistical and physical), design engineering methods allow expanding the capabilities of the dependability theory for predicting the dependability of technical objects and making dependability problems understandable and accessible for engineers. DEAD was tested in the design of single-use mechanical space devices and hydraulic assemblies of oil well equipment [14], which allowed:

• detecting design and process engineering errors in the technical documentation;

• evaluating the effectiveness of the existing computational and experimental optimization of product design;

• assessing the adequacy of the established requirements in the design documentation;

• identifying unacceptable combinations of structural parameters based on the design constraints, actual manufacturing and control conditions;

• drawing conclusions regarding the propensity to failure of products;

• predicting the compliance to the specified dependability requirements;

• providing recommendations regarding design modifications to ensure specified dependability of products.

Comparability of DEAD with existing predictive approaches to dependability

The idea of dependability analysis (evaluation) with account of design and technological factors is not new. Its relevance was repeatedly noted and demonstrated, for example, in [24–26]. However, analysis and evaluation methods for design engineering factors that allow designers of highly vital systems making their decisions taking the dependability into account are yet to be developed (as far as the author knows).

Certain aspects of accounting for design factors that affect dependability are well known in the literature. For example, the basics of calculating the dependability by strength are set forth in [27], and approaches to calculating the dependability of the mechanical parts of an aircraft subject to the requirements of strength and undisturbed operation in case of deployment mechanisms actuation, are shown in [28, 29]. The parameters by which dependability is calculated in the indicated examples are part of the column-vector (1). Operational integrity and dependability are calculated using formulas (2) – (3), taking into account the physical foundations of ensuring the desired parameters. However, as practice shows [14], when calculating highly vital systems, it is required to take into account additional factors affecting dependability. Such factors may include, for example, sudden disappearance of gaps in kinematic pairs, insufficient vibration resistance of joints, presence of foreign objects in deployment mechanisms (components or adjacent parts of structures), instability of the mechanism settings, insufficient actuator stroke, critical operation execution modes being violated or not set etc. [14, 16–17, 23].

In order to establish the output parameters that affect dependability, a design engineering analysis of dependability is performed [14, 16] that produces a parametric description of the functionality (1), operational integrity (2) and dependability (3) of the structure. Moreover, the application of the method of mitigation [14, 16] that allows translating possible failures into the desired output parameters, actually allows considering model (1) – (3) as a condition for the failure-free operation of the structure. This greatly increases the effectiveness of analytical verification, for example, using FMECA [30], which is based on identifying undesirable failures by the severity of their consequences and conducting expert assessments of the risks of possible failures, but does not provide an answer on how to prevent the very possibility of failures. The use of DEAD allows managing failures by selecting the values of the design parameters under the conditions of given restrictions (modes and conditions of use) based on mathematical equations (1) – (3) that reflect the set of knowledge, ideas and hypotheses when implementing output effects based on the physical laws of nature.

When it comes to DEAD, it should be understood that it does not replace or undermine the existing foundations of dependability (generally accepted standards of dependability should be followed where possible). However, when there is no information on the dependability of components and statistical data on product failures is insufficient, this method allows avoiding a significant part of design errors, including those that cause unlikely failures. The use of DEAD puts the notion that dependability calculations are impossible and even meaningless

[12] for highly vital systems (0.997 and higher) into question. In the framework of DEAD, calculations of the dependability of highly vital systems are critical, but its procedure requires standardization [23].

Moreover, the use of DEAD in itself is a necessary, but insufficient condition for creating highly vital systems. Like any other tool, it requires skill. In this case, that is the knowledge of the physical principles of operation of technical items, the fundamentals of engineering disciplines and methods of design for ensuring quality and dependability. Furthermore, all the same is required when using digital design technologies. Fortunately, the need to follow the established DEAD algorithm together with the possibility of obtaining a posteriori knowledge (from the results of testing and operation) allows accumulating knowledge with each iterative cycle of analysis and, if necessary, creating check lists of design principles and design rules [14, 28], corresponding to a specific subject area of development (which only enhances the effectiveness of the method).

Compatibility and conditionality of DEAD and digital technologies

From the standpoint of being focused on dependability prediction, DEAD and digital design methods use common procedures, i.e. substantiation of design solutions in order to eliminate (reduce the probability of) errors that can cause failures based on analytical, computational and experimental verification.

DEAD is within the authority of the human dependability expert. It is therefore instrumental in compiling the initial data for computer calculations in the human – computer system, since the effectiveness of digital technology itself directly depends on their completeness and reliability.

Today, in the course of construction of a matrix of target indicators and limitations, as well as validation of the calculation results, each iteration would involve experts who rely only on their own knowledge and experience [6]. The use of DEAD enables algorithmized preparation and verification of input data for computer calculations using formulas (1) – (2) and validation of their results according to (3). Thus, two problems are solved:

- there is no need to search for unique and costly experts (who may just not be around at the right time);
- engineers in the human – computer system are able to use a system approach that increases the efficiency of their decisions and allows for effective actions when preparing and conducting computer calculations.

The benefits of the latter cannot be overestimated. The capabilities of computer hardware and software are constantly growing, while the human capabilities in terms of technology development have been deteriorating in recent years: the quality of thinking does not improve, analytical abilities do not increase and the educational level has noticeably degraded. If knowledge is not enhanced and human actions are not further algorithmized, an ever-widening gap in the human – computer system may lead to unpredictable consequences, the most harmless of which may be Robert Sheckley's prophecy in Ask a Foolish Question.

In theory, a generalized parametric model of operation (1) – (3) consisting solely of parameters can be obtained by simulating the operation of technical items at the micro-, macro- and megaworld levels (the principles of constructing digital models allow for that and are limited only by the available computational power). In this case, only an automated option is required, that would enable additive calculation of the predicted dependability resulting from the required measures aimed at preventing structural failures. Otherwise (if human knowledge or computing capabilities are insufficient), human participation is required for adjusting the calculation of predicted dependability by taking into account factors that require probabilistic assessment based on information models in the form of a black box.

The use of DEAD in digital technology may be essential for topological optimization of structures. In that case, it is important to distinguish between the goals of the tasks being solved. It is one thing when topological optimization is carried out to reduce production costs, while if the reduction of such costs may cause risks of excessively larger losses than the benefit of the savings is a totally different matter. For example, the mass of a mechanical spacecraft device can be reduced by 1 kg through topological optimization, which leads to savings of about 10^3 dollars based on the market price of blanks and the cost of manufacture. However, the failure of the mechanism in orbit as a result of the topological optimization can lead not only to losses of about 10^6 dollars, corresponding to the unit cost of payload deployment, but also to much more critical losses in the form of the cost of the lost spacecraft and the time of its creation, the cost of repeated satellite manufacture and financial losses due to potential reputational costs (for example, increasing cost of space risk insurance). In this case, predicting the dependability becomes a top priority that must be addressed using scientific methods.

Conclusions

Digital technology provides a tangible opportunity of predicting, reducing the impact or eliminating possible failures. That can be achieved through the same means that often cause failures, i.e. design engineering. For that purpose, it is required to create new applications of the modern dependability theory based on engineering disciplines and design engineering methods developed for ensuring quality and dependability of products.

Anyone interested in the problems described in the article are kindly asked to express their opinions, including personally at pokhabov_yury@mail.ru.

References

1. Ushakov I.A. Is Reliability Theory Still Alive? *Reliability: Theory & Applications.* 2017;3(46):45-68. [accessed 15.04.2020]. Available at: https://cyberleninka.ru/article/n/is-reliability-theory-still-alive-1/viewer.

2. Ushakov I.A. [Dependability: past, present, future: keynote speech of the opening of Mathematical Methods in Reliability (MMR–2000) conference, Bordeau, France, 2000]. *Reliability: Theory & Applications*. 2016;1(1):17-27. (in Russ.) (accessed 15.04.2020). Available at: http://www.gnedenko.net/Journal/2006/RTA_1_2006.pdf. (in Russ.)

3. Pokhabov Yu.P., Ushakov I.A. [On the fail-safety of unique highly vital systems]. *Metody menedzhmenta kachestva.* 2014;11:50-56. (in Russ.)

4. Schwab K. The fourth industrial revolution. Moscow: Eksmo; 2016.

5. Levenchuk A. [Systems engineering thinking in life cycle management]. (accessed 15.04.2020). Available at: https://ailev.livejournal.com/1121478.html. (in Russ.)

6. Borovkov A.I., Riabov Yu.A., Kukushkin K.V. [Digital tweens and the digital transformation of defense industry companies]. *Oboronnya teknika*. 2018;1:6-33. (in Russ.)

7. Pronikov A.S. [Dependability of machines]. Moscow: Mashinostroenie; 1978. (in Russ.)

8. Hilbert D. Mathematical Problems. *Bulletin of the American Mathematical Society: journal*. 1902;8(10):437-479.

9. Kuleshov A.P. [To overcome the resistance of materials: February 2, 2018 interview]. *Stimul: zhurnal ob innovatsiyakh v Rossii*. (accessed 15.04.2020). Available at: https://stimul.online/articles/interview/preodolet-soprotivlenie-materialov/?sphrase_id=1295. (in Russ.)

10. Doronin S.V., Pokhabov Yu.P. [Improving the reliability of structural strength estimates of technical items]. *Vestnik mashinostroyeniya*. 2013;6:85-88. (in Russ.)

11. Bolotin V.V. [Application of probability theory and dependability theory methods in structural analysis]. Moscow: Stroyizdat; 1971. (in Russ.)

12. Polovko A.M., Gurov S.V. [Introduction into the dependability theory]. Saint Petersburg: BHV-Peterburg; 2006. (in Russ.)

13. Lerner E. [Alternative to "starting at haphazard"]. *Aerokosmicheskaya tekhnika*. 1987;9:157-160. (in Russ.)

14. Pokhabov Yu.P. [Theory and practice of dependability of single-use mechanical devices]. Krasnoyarsk: SFU Publishing; 2018. (in Russ.)

15. Plahotnikova E.V., Safonov A.S., Ushakov M.V. The design of products with requirements of reliability parameters. *Izvestiya TulGU: Teknicheskie nauki*. 2015;7(1):134-139. (in Russ.)

16. Pokhabov Yu.P. Design for reliability highly vital systems on the example of a moving rod. *J. Sib. Fed. Univ. Eng. technol.* 2019;12(7):861-883. (in Russ.)

17. Timashev S.A., Pokhabov Yu.P. [Problems of comprehensive analysis and assessment of individual design dependability of spacecraft (with the example of rotating structures)]. Ekaterinburg: AMB; 2018. (in Russ.)

18. Gladky V.F. [Probabilistic methods of aircraft structural design]. Moscow: Nauka; 1982. (in Russ.)

19. Zolotov A.A., Pokhabov Yu.P., Ensuring the design reliability of unfolding structures spacecraft. *Polet*. 2018;7:36-45. (in Russ.)

20. Pokhabov Yu.P. [Method of selection of the servo for rotation of a structure in an articulated joint: pat. 2198387 Russian Federation. No. 2000129330/28; claim 23.11.2000; publ. 10.02.2003, bul. no. 4]. (in Russ.)

21. Pokhabov Yu.P. [Method of fastening items: pat. 2230945 Russian Federation. No. 2002113143/11; claim 18.05.2002; publ. 20.06.2004, bul. no. 17]. (in Russ.)

22. Pokhabov Yu.P., Nagovitsin V.N. [Method of fastening items with a statistically indeterminate system of connections: pat. 2125528 Russian Federation. No. 5067373/28; claim 29.09.1992; publ. 27.01.1999, bul. no. 3]. (in Russ.)

23. Pokhabov Yu.P. What should mean dependability calculation of unique highly vital systems with regards to single-use mechanisms of spacecraft. *Dependability*. 2018;18(4):28-35.

24. Hecht H., Hecht M. Reliability prediction for spacecraft: Report prepared for Rome Air Development Center: no. RADC-TR-85-229, Dec. Rome Air Development Center; 1985.

25. Saleh J.H., Caster J.-F. Reliability and Multi-State Failures: A Statistical Approach, First Edition. NJ: John Wiley & Sons; 2011.

26. Testoedov N.A., Mikhnev M.M., Mikheev A.E. et al. [Spacecraft manufacturing process]. Krasnoyarsk: SibSAU; 2009. (in Russ.)

27. Dhillon B.S., Singh C. Engineering reliability. NJ: John Wiley & Sons; 1981.

28. Bowden M.L. Peter L. Conley, editor. Deployment devices. Space Vehicle Mechanisms: Elements of Successful Design. NJ: John Wiley & Sons; 1998.

29. Kuznetsov A.A. [Structural dependability of ballistic missiles]. Moscow: Mashinostroenie; 1978. (in Russ.)

30. ECSS Standard. Space product assurance. Failure modes, effects (and criticality) analysis (FMEA/FMECA). ECSS Secretariat, ESA ECSS-Q-ST-30-02C.

About the author

Yuri P. Pokhabov, Candidate of Engineering, Joint Stock Company NPO PM – Maloe konstruktorskoye buro (OAO NPO PM MKB), Head of Center for Research and Development, Russia, Krasnoyarsk Krai, Zheleznogorsk, e-mail: pokhabov_yury@mail.ru.

The author's contribution

The paper is the result of a long practice (since 1982) of design and assurance of dependability of space structure deployment mechanisms. Using the patented method of design engineering analysis of dependability (DEAD), between 2014 and 2019 expert assessment of the susceptibility to failure of structure-deploying mechanisms of spacecraft was conducted by leading Russian developers (with publication of scientific technical reports), that identified the insufficiency of todays' methods of analytical and experimental verification of dependability in the aerospace industry in terms of ensuring the required reliability above 0.999 and, paradoxically, reduced quality of the designs in terms of dependability subject to the application of digital design technology.

Selection of network structures of pipeline systems resilient to mixed damage

Igor A. Tararychkin, *V. Dahl Lugansk National University, Ukraine, Lugansk*
donbass_8888@mail.ru

Igor A. Tararychkin

Abstract. *Pipeline transportation systems are used for the purpose of delivering to consumers various substances, materials, including those required for continuous flow processes. The operation of such complex industrial facilities is associated with some risks and possible failures of individual units and assemblies due to various causes. The paper examines the specificity of pipeline transportation systems behaviour in emergency situations. The development of such processes may cause the disconnection from the source of some or all end product consumers. The process of damage may occur in accordance with the following mechanisms: progressive damage, when individual pipeline systems fail in a random order; progressive blocking, when individual transportation nodes fail in a random order. An accident scenario, in which progressive damage to linear elements and blocking of transportation nodes simultaneously occur within a system, represents mixed damage. The* **Aim** *of this paper is to develop the criteria for estimating a pipeline transportation systems' resilience to mixed damage, as well as the methods for solving routine problems of synthesis of network structures resilient to such process.* **Methods of research.** *The ability of a specific system to resist mixed damage depends on its network structure and is identified by means of simulation. The structural changes caused by mixed damage are described with a cyclogram, whose parameters indicate the number of damaged linear and blocked point elements within one cycle of system exposure. A comparison of the network structures' ability to resist mixed damage is only possible in case they are comparable. For that purpose, the analyzed systems must have identical numbers of nodes, linear elements, as well as end product consumers. Additionally, such systems must be exposed to mixed damage with identical cyclograms.* **Results.** *The simulation of the mixed damage process identified such characteristic as the average percentage of system components, whose failure causes disruption of the connection of all consumers to the source, as well as the average percentage of nodes, whose blocking causes a complete disconnection of the source from all consumers. The developed method of estimation of resilience to mixed damage allows solving the following structural synthesis problems: selection of the position of the source of the end product within the given network; selection of the position of new consumers within an existing system; definition of the locations of additional fragments' connection to the system; selection of coupling linear elements when additional fragments are connected to a transportation system.* **Conclusions.** *Mixed damage is a hazardous development scenario of an emergency situation and is associated with rapid degradation of the transportation capacity of pipeline systems. Various network structures vary in terms of their ability to resist mixed damage, while their resilience characteristics should be identified using computer simulation. A comparison of the mixed damage resilience characteristics is only possible for comparable network structures with equal numbers of nodes, linear elements and end product consumers. Additionally, the same cyclogram of mixed damage must be used.*

Keywords: *pipeline, system, structure, damage, network, accident, resilience.*

For citation: *Tararychkin I.A. Selection of network structures of pipeline systems resilient to mixed damage. Dependability. 2020;2: P. 12-17. https://doi.org/10.21683/1729-2646-2020-20-2-12-17*

Received on: *02.01.2020* / **Upon revision:** *21.04.2020* / **For printing:** *17.06.2020.*

Pipeline transportation systems are used for the purpose of delivering various substances, materials, including those required for continuous flow processes [1-7].

The operation of such complex industrial facilities is associated with some risks and possible failures of individual units and assemblies due to various causes [8-11]. In some cases, the failure of linear elements (pipelines) does not cause a noticeable limitation of the system's performance due to the presence of redundant connections and alternative ways of delivering the end product.

If, as the result of internal or external effects, a sequence of structural element failures occurs within a system, such development of the accident may cause the disconnection from the source of initially some, then all end product consumers. A failure in a random order of a set of a system' linear elements is called progressive damage [12, 13].

The failure of an individual transportation node makes it unable to handle transportation flows, and such point element of the system becomes blocked. Sequential blocking of point elements of a system in a random order is herein called progressive blocking [14, 15].

If, during accident development, both progressive damage to linear elements and blocking of transportation nodes occur simultaneously, such system exposure is considered as mixed damage.

The development of mixed damage is associated with a rapid degradation of the system's transportation capabilities, however technical literature does not provide any information regarding the patterns of this process, as well as methods of estimating the systems' ability to resist its development.

The aim of this paper is to develop the criteria for estimating pipeline transportation systems' resilience to mixed damage, as well as the methods for solving routine problems of synthesis of network structures resilient to such process.

Mixed damage is characterized by cyclogram T(α.β). Parameters α and β indicate the number of sequentially damaged linear elements and blocked transportation nodes within one system exposure cycle. Under the defined mixed damage cyclogram, for each moment of system time the exact number of operable structural elements of the analyzed network entity can be specified.

Using computer simulation [16, 17], the following resilience characteristics of the analyzed network entity were identified.

Average share of linear system elements φ_{EL}, whose damage causes the disruption of the connections between the source and all end product consumers.

Average share of transportation nodes φ_{UZ}, whose damage under conditions of mixed damage causes the disruption of the connections between the source and all end product consumers.

All calculations were performed using MathCAD [18, 19]. The above characteristics should be considered as the projections of vector $\vec{\Phi}^*$ on the coordinate axes that allows estimating the system's ability to resist mixed damage.

The module of this vector $\left|\vec{\Phi}^*\right| = \sqrt{\varphi_{EL}^2 + \varphi_{UZ}^2}$ generally characterizes the resilience of the analyzed network structure [20]. The higher is value $\left|\vec{\Phi}^*\right|$, the better is the examined item's resilience to mixed damage. Practically speaking, the value of the of the above characteristics consists in the fact that they enable comparative analysis of the resilience of various network entities.

However, a correct comparison of values φ_{EL}, φ_{UZ} and $\left|\vec{\Phi}^*\right|$ is possible only in the structures are comparable. For that purpose, they must have identical numbers of:

- linear elements;
- transportation nodes;
- end product consumers.

Additionally, the conditions of damage of the analyzed network structures must be similar, i.e. be described by the same damage cyclogram T(α.β).

A series of computing experiments has established that the correlation of the resilience indicators of the sets of comparable network structures does not depend on the adopted damage cyclogram, but is rather defined by the existing set of intrasystem communications. That means that for a random set of comparable network entities the correlation between their resilience does not depend on the specific conditions of mixed damage.

Then, the estimation of the correlation of the resilience of a number of comparable network structures only requires defining the corresponding values of $\left|\vec{\Phi}^*\right|$ in the conditions of test input with characteristics $\alpha = \beta = 1$. The above structures are regulated in terms of their resilience to mixed damage in such a way as to ensure correspondence between more resilient systems and higher values of $\left|\vec{\Phi}^*\right|$.

Let us note that in case of test input with cyclogram T(1.1) there is a consecutive alteration of random damage of one linear element and blocking of one transportation node of the system. This exposure pattern is further used for the purpose of comparative estimation of the ability of comparable network structures to resist mixed damage.

Routine problems of structural synthesis of pipeline systems resilient to mixed damage

The study of the specificity of network structures behaviour when affected by mixed damage is of practical interest. This specificity should be taken into consideration with regard to problems of structural analysis and synthesis of pipeline systems of various complexity and purpose. The properties of alternative network structures and design decision-making must be evaluated subject to the specified comparability requirements. Let us examine some typical design problems, as well as methods of solution.

Selection of the position of the source of the end product within the given network

Problem definition. Within the given network structure with known positions of consumers, it is required to identify the location of the source of the end product, whereas the system's resilience to mixed damage is the highest. The structural synthesis problem in this case is solved by comparing the values of $\left|\vec{\Phi}^*\right|$ of network entities with different locations of the source.

Let us examine alternative network structures designated SKA, … SKD shown in Fig. 1. They are characterized by varied location of source A and each have 13 nodes, 23 linear elements and 9 product consumers. In the course of the computation, each of the above structures was exposed to mixed damage with cyclogram T(1.1).

Under such conditions, all the above systems are comparable, while the comparison of corresponding values of $\left|\vec{\Phi}^*\right|$ proves to be correct. The calculation data obtained for samples of the size of 10^4 elements are shown in Fig. 1.

It can be seen that the worst location of the source corresponds to structure diagram SKA, while the highest value $\left|\vec{\Phi}^*\right|$ can be observed in case diagrams SKC and SKD are used.

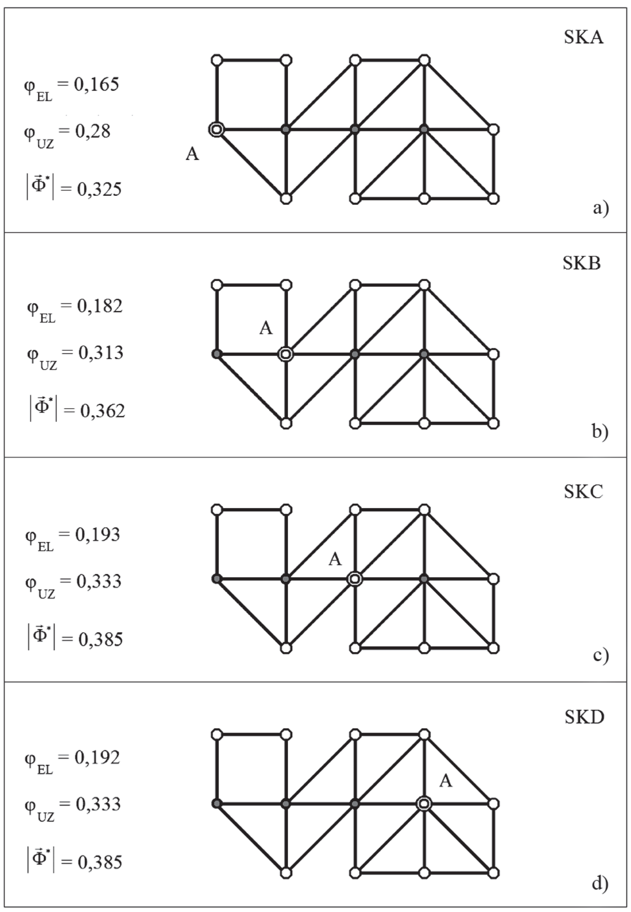


Fig. 1. Structure diagrams of the SKA (a), … SKD (d) pipeline systems with different locations of the source of the end product A

At the same time, the resilience of the pipeline systems with structure diagrams SKC and SKD proves to be identical. Thus, while solving the problem at hand, one of those structure diagrams should be selected. The ultimate solution in this case depends on the additional conditions or limitations that take into consideration, for instance, the possible cost of practical implementation of those two variants.

Selection of the position of new end product consumers within an existing system

Problem definition. In a system with a known location of the source and several end product consumers, it is required to identify the location of additional consumers, whereas the resilience to mixed damage is the highest.

Fig. 2 shows the layouts of network structures with source A and additional end product consumer B, C and D to be included in the system.

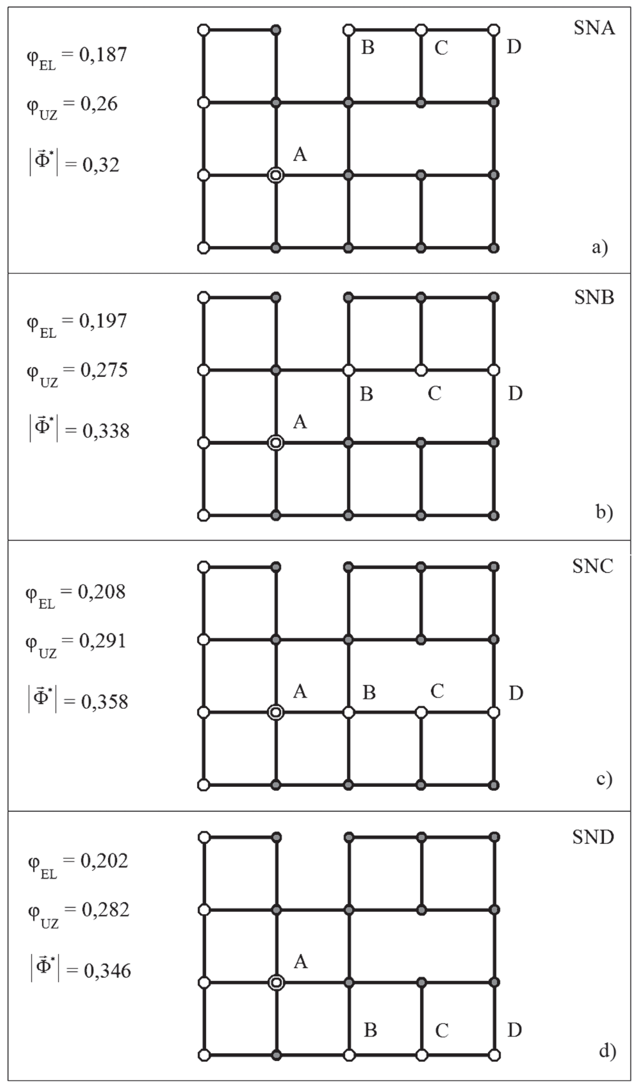


Fig. 2. Structure diagrams of the SNA (a), … SND (d) pipeline systems with different locations of the end product consumers B, C, D

The difference between the above options consists in the location of such consumers in the network. It is required to analyze and select the network structure with the highest resilience to mixed damage.

Let us assess the comparability of the structure diagrams shown in Fig. 2. They all include the same number of nodes, linear elements and product consumers. In case such structures are exposed to mixed damage with cyclogram T(1.1) the values of resilience indicators can be correctly compared to each other.

The values of resilience characteristics identified as the result of simulation for samples of the size 10^4 elements are shown in Fig. 2. It can be seen that the structure designated SNC is characterized by the highest resilience to mixed damage. The value of $\left|\vec{\Phi}^*\right|$ of such network entity is about 1.12 times higher than the corresponding value of structure SNA with the worst properties.

Thus, the network entity designated SNC should be considered as the solution of the structural synthesis problem.

Selection of the locations of an additional fragment connection to the system

Problem definition. The planned reconstruction of a pipeline system aims to extend its capacity and introduce an additional fragment with a number of consumers. There are several ways this extension can be implemented. The option must be selected that would enable the highest achievable resilience to mixed damage for the specific pipeline system.

An example structure diagram of a pipeline system and extension fragment are shown in Fig. 3. The extension includes consumers B, … G that are connected to each other and can be included into the original system with the creation of network structures SFA and SFB (Fig. 4).

In terms of the capabilities of the newly created system all the above extension options are equivalent. It is required to evaluate the resilience of SFA and SFB to mixed damage, as well as select the best extension option.

At the first stage of analysis it is required to identify if the above network entities are comparable. In this case the answer will be positive, as they have matching numbers of nodes, linear elements and end product consumers. Additionally, the analyzed entities are subsequently exposed to mixed damage with the same cyclogram T(1.1). The identified values of $\left|\vec{\Phi}^*\right|$ are shown in Figure 4.

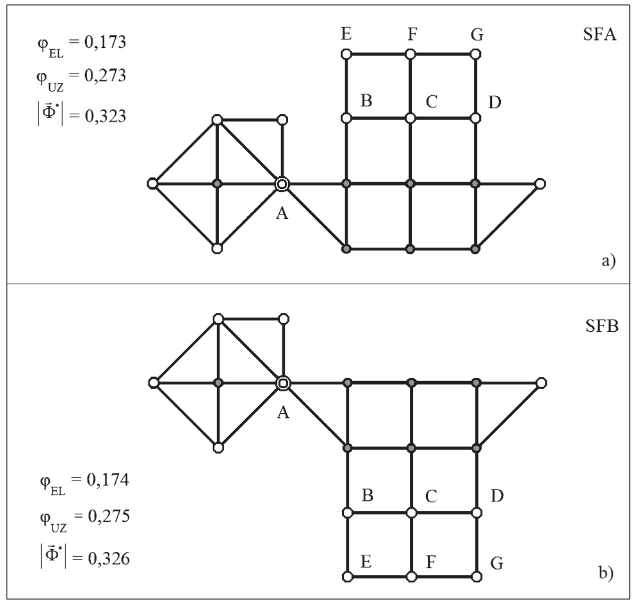


Fig. 4. Structure diagram of pipeline systems SFA (a) and SFB (b) corresponding to the different extension options

As simulation used samples of the size of 10^4 elements, the obtained expected values have 2 decimal significant figures. That means that the resilience to mixed damage of network structures SFA and SFB turns out to be identical. In this context, the ultimate choice is to be made subject to additional criteria, e.g., subject to the results of installation activities cost estimation.

Selection of coupling pipelines when an additional fragment is connected to a transportation system

Problem definition. The reconstruction of the pipeline transportation system is associated with the inclusion of a fragment that may contain several end product consumers. For the given number of additional pipelines, it is required to select the locations of their connection to the system and the fragment. The resulting network structure must have a high resilience to mixed damage.

Let us then examine the problem related to the connection of an extension fragment that includes 10 end product consumers (Fig. 5).

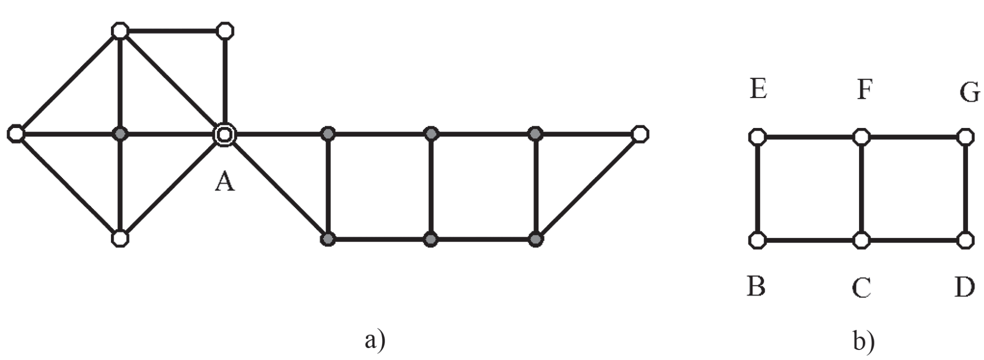


Fig. 3. Structure diagram of the pipeline system (a) and extension fragment with 6 end product consumers (b)

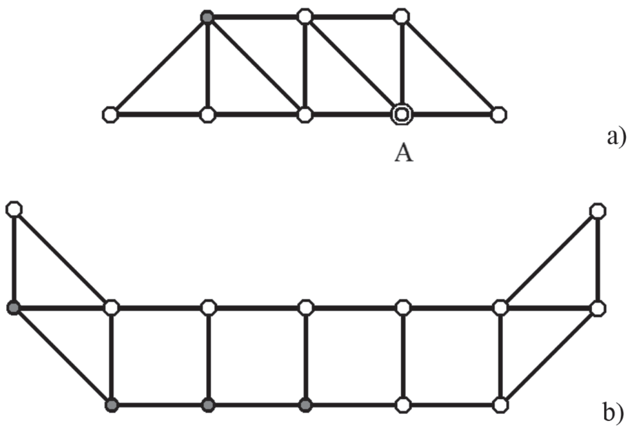


Fig. 5. Structure diagram of a pipeline transportation system with source A and 6 product consumers (a), as well as the diagram of the connected extension fragment with 10 consumers (b)

The connection is to use 4 additional pipelines. Fig. 6 shows the available connection options that allow achieving the goals of the reconstruction. At the same time, it is required to identify which of the examined options ensures the highest system resilience to mixed damage. Let us assess the comparability of network structures SOA, …. SOD shown in Fig. 6.

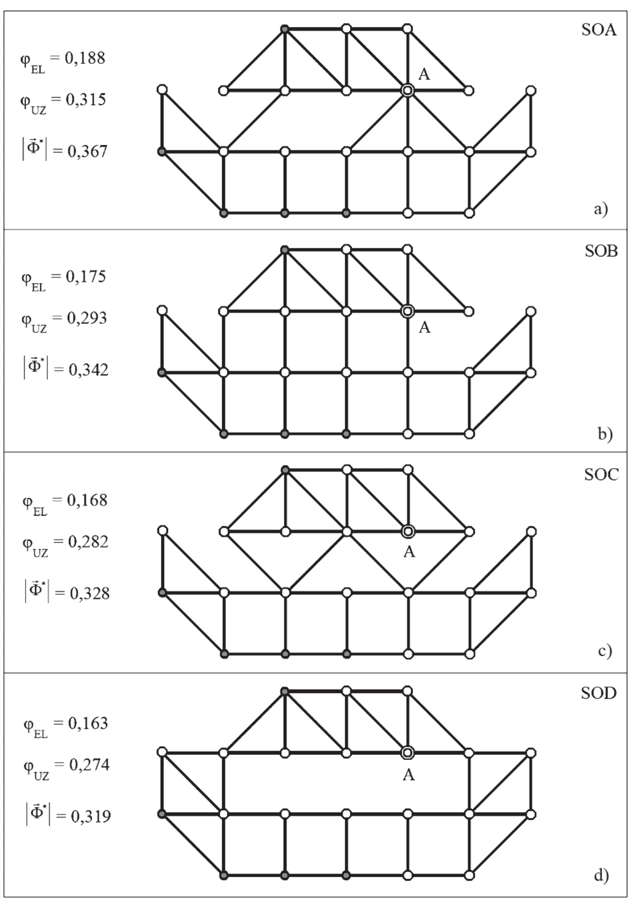


Fig. 6. Structure diagrams of pipeline systems SOA (a) … SOD (d) with various locations of the coupling pipelines

All of them have the same number of nodes, linear elements and product consumers, therefore in case of mixed damage with cyclogram T(1.1) the corresponding resilience characteristics can be correctly compared. The calculation data obtained for the above network structures using samples of the size of 10^4 elements are shown in Fig. 6.

It can be seen that the highest value of $\left|\vec{\Phi}^*\right|$ is observed in case of mixed damage to network structure SOA. For the examined extension options the highest value of $\left|\vec{\Phi}^*\right|$ exceeds the lowest one by 1.15 times. Thus, the structure diagram shown in Fig. 6 (a) should be considered as the solution of the synthesis problem.

Conclusions

1. Mixed damage is a hazardous development scenario of an emergency situation and is associated with rapid degradation of the transportation capacity of pipeline systems.

2. Various network structures vary in terms of their ability to resist mixed damage, while their resilience characteristics should be identified using computer simulation.

3. Comparison of the mixed damage resilience characteristics φ_{EL}, φ_{UZ}, $\left|\vec{\Phi}^*\right|$ is only possible for comparable network structures with equal numbers of nodes, linear elements and end product consumers. Additionally, the same cyclogram of mixed damage must be used.

References

[1] Barker G. The Engineer's Guide to Plant Layout and Piping Design for the Oil and Gas Industries. Elsevier Inc.; 2018.

[2] Baron H. The Oil & Gas Engineering Guide. Second Edition. Paris: Technip; 2015.

[3] Stewart M. Surface Production Operations Facility Piping and Pipeline Systems. Vol. III. Gulf Professional Publishing. Elsevier Inc.; 2016.

[4] Kleiber M. Process Engineering. Addressing the Gap between Studies and Chemical Industry. Walter de Gruyter GmbH.; 2016.

[5] Ellenberger J.P. Piping and Pipeline Calculations Manual Construction. Design Fabrication and Examination. Second Edition. Butterworth-Heinemann. Elsevier Inc.; 2014.

[6] Mokhatab S., Poe W.A., Speight J.G. Handbook of Natural Gas Transmission and Processing. Gulf Professional Publishing. Elsevier Inc.; 2006.

[7] Sambasivan M., Gopal S. Handbook of Oil and Gas Piping. A Practical and Comprehensive Guide. Taylor & Francis Group.; 2019.

[8] Nolan D.P. Handbook of Fire and Explosion Protection Engineering Principles for Oil, Gas, Chemical, and Related Facilities. Fourth Edition. Gulf Professional Publishing. Elsevier Inc.; 2019.

[9] Singh R. Pipeline Integrity Handbook. Risk Management and Evaluation. Elsevier Inc.; 2014.

[10] Cheng Y.F. Stress Corrosion Cracking of Pipelines. John Wiley & Sons Inc.; 2013.

[11] Jawad M.H. Stress in ASME pressure vessels, boilers, and nuclear components. The American Society of Mechanical Engineers. John Wiley & Sons, Inc.; 2018.

[12] Tararychkin I.A. Structural synthesis of pipeline transportation systems resilient to damage to linear elements. *Problemy sbora, podgotovki i transporta nefti i nefteproduktov*. 2017;1(107):96-106. (in Russ.)

[13] Tararychkin I.A., Blinov S.P. [Simulation of the process of damage to pipeline network structures]. *World of Transport and Transportation*. 2017;15(2):6-19. (in Russ.)

[14] Tararychkin I.A., Blinov S.P. [The distinctive features of damage to network structures and development of accidents in pipeline transportation facilities]. *Occupational Safety in Industry.* 2018;3:35-39. (in Russ.)

[15] Tararychkin I.A. Resistance of the Pipeline Transportation Systems to Damages of the Network Elements. *Occupational Safety in Industry*. 2018;6:63-68. (in Russ).

[16] Banks J., Carson J.S., Nelson B.L., Nicol D.M. Discrete-Event System Simulation. 5 edition. Prentice Hall; 2009.

[17] Bandyopadhyay S., Bhattacharya R. Discrete and Continuous Simulation. Theory and Practice. CRC Press; 2014.

[18] Porshnev S.V. [Computer simulation of physical processes using the MathCad suite]. Moscow: Goriachaya liniya – Telekom; 2004. (in Russ.)

[19] Okhorzin V.A. [Computer simulation in Mathcad: a study guide]. Moscow: Finansy i statistika; 2006. (in Russ.)

[20] Krasnov M.L., Kiseliov A.I., Makarenko G.I. et al. [All advanced mathematics. Vol. 1]. Moscow: Editorial URSS; 2003. (in Russ.)

About the author

Igor A. Tararychkin, Doctor of Engineering, Professor, V. Dahl Lugansk National University, Ukraine, Lugansk, e-mail: donbass_8888@mail.ru

The author’s contribution

The author analyzed the specificity of mixed damage in network structures of pipeline transportation systems. The paper proposes to describe the dynamics of a system’s stationary random exposure with a damage cyclogram. Resilience characteristics and conditions of comparability of network structures under mixed damage were identified.

A comparative analysis was conducted of the ability of various network structures to resist mixed damage as part of standard designs.

Discussion of dependability terminology

Editorial note

Our Journal has published a number of papers that deal with the subject of dependability-related terminology and its standardization. However, our readers continue to have a keen interest in the matter and send new papers dedicated to it. Unfortunately, many authors who express reasonable (as they believe) proposals often disregard a whole range of factors that must be taken into consideration with respect to the development of a general (multisectoral) dependability standard. Those include the requirements of general regulatory documents, Russian and international experience of standardization in this domain, provisions of other general standards, requirements of other industries, etc. In this context, we suggest interested experts should take part in discussion regarding the dependability-related terminology. This issue opens discussion by the article of V.A. Netes, member of the Editorial Board, in which he, drawing from his experience in dependability-related terminological standardization, lays down its general principles. The editors believe it to be reasonable – before even starting the discussion of specific terms and definitions – to agree upon such common principles and then examine their specific applications in certain areas of science and practices.

The principles of dependability terminology standardization

Victor A. Netes, *Moscow Technical University of Communication and Informatics, Russian Federation, Moscow*

Victor A. Netes

Abstract. *Aim. The paper continues a series of publications discussing the dependability terminology and its standardization. It aims not to review and discuss specific terms, but rather to formulate the main principles that should be used as the basis for the development of a general terminology standard for dependability in technics. Such consistent general principles will enable easier solutions regarding specific terms and definitions.* **Methods.** *The general principles and requirements set out in the regulatory documents on standardization are specified in the context of the dependability terminology standard. The provisions of a number of other general technical standards that have an impact on the standardization of dependability terminology are also taken into account. Current and former terminology standards are considered, both domestic (GOST 13377–67, GOST 13377–75, GOST 27.002–83, GOST 27.002–89, GOST R 27.002–2009 and GOST 27.002–2015) and international (IEC 60050-191:1990 and IEC 60050-192:2015). The author analyzed to what extent they comply with the general principles; the shortcomings of the reviewed standards are identified.* **Findings and conclusions.** *The main principles that a general dependability terminology standard should conform to are formulated: continuity in relation to previous similar domestic standards, alignment with the international IEC standard, consistency with other general technical standards, internal consistency and logical coherence, generality and universality to meet the needs of all industries.*

Keywords: *dependability, terminology, national and international standards, continuity, consistency and coherence, generality and universality.*

For citation: *Netes V.A. The Principles of dependability terminology standardization. Dependability. 2020;2: 19-23. https://doi.org/10.21683/1729-2646-2020-20-2-19-23*

Received on: *30.03.2020* / **Upon revision:** *18.04.2020* / **For printing:** *17.06.2020.*

Introduction

Over the last years the standardization of dependability, in particular the terminology it uses, has been dwelt upon by various publications (e.g. [1–6], some more papers to be referenced further on). On the one hand, such an interest in the topic is good news for the author who has been working in this area for many years and has participated in the development of the interstate and international terminology standards. On the other hand, however, this is indicative of problems in the area [5]. It is also upsetting that discussions sometimes go around in circles and it seems as if discussants do not read carefully what has been written on the subject before or even are not familiar at all with the key provisions in the area of standardization established by normative documents.

This situation initiated the writing of this paper whose major aim is not to review and discuss specific terms, but rather to formulate the overall principles that should be used as the basis for the development of a general terminology standard for dependability in technics. The author hopes that if the general principles are agreed upon first, then it will be easier to come to solutions for specific terms and definitions.

Naturally, any standard shall comply with the requirements specified in the Federal law "On standardization in the Russian Federation" (dated 29.06.2015 № 162-FZ), in the standards of complexes "Interstate standardization system" (GOST 1.) or "Standardization in the Russian Federation" (GOST R 1.). The procedure of the development of terms and definitions standards is prescribed by the guidelines [7]. These provisions will be considered in the paper in relation to a terminology standard for dependability. The description will be supported by examples from the experience related to the development of such standards.

Legacy

One of the principles of standardization specified by Article 4 of the Federal law "On standardization in the Russian Federation" is the provision of the continuity of activities in the area of standardization. The first general terminology standard for dependability in technics GOST 13377–67 was adopted in the USSR over half a century ago. It was consecutively replaced by GOST 13377–75, GOST 27.002–83, GOST 27.002–89, GOST R 27.002–2009 and GOST 27.002–2015. In this row GOST R 27.002–2009 stands out (its specifics will be considered later), yet all others maintained legacy to its predecessors. Of course, each new standard introduced new terms and elaborated some definitions, otherwise it would have not made sense to adopt it, nonetheless some basic provisions were kept intact.

For information the table contains the number of terms used in each of the mentioned standards (as well as in the international standards considered below). It is evident that GOST R 27.002–2009 stands out for this parameter as well, with the number of terms in it exceeding the boundary of 200 units specified in [7].

Number of terms in the standards

Standard	Number of terms
GOST 13377–67	24
GOST 13377–75	86
GOST 27.002–83	89
GOST 27.002–89	116
GOST R 27.002–2009	212
GOST 27.002–2015	146
IEC 60050-191:1990	244 (only terms related to dependability considered)
IEC 60050-192:2015	260

In the author's opinion, the basic provisions maintained in the Russian standards are as follows:

1. The definition of dependability as the *property* of an item (product in GOST 13377–67, the relationship between these notions will be considered below).

2. The consideration of dependability in the conditions of *an item's application (use), maintenance and repair, storage and transportation.*

3. The definition of dependability as an integrated property that can incorporate several simpler properties: *reliability, maintainability, durability and storability*.

4. The distinction of two pairs of opposite states of an item: *up state – down state* and *perfect state – imperfect state*.

5. The presence of several *integrated dependability measures* apart from the availability factor: *total availability (utilization) factor*, *interval availability factor* (since 1975), *effectiveness retention (efficiency) ratio* (since 1983).

Alignment with international standards

One of the objectives of standardization as specified in the Federal law "On standardization in the Russian Federation" (Article 3) is the facilitation of the Russian Federation integration into the global economy and the international systems of standardization as an equal partner. In accordance with [7, Par. 3.12], one of the tasks of standardization in scientific and technical terminology is the harmonization (the provision of alignment) of scientific and technical terminology at the national and international levels. Therefore, it is recommended to use terminology standards and vocabularies of international organizations (ISO, IEC, etc.) to the maximum extent practicable in order to provide terminology support for the Russian national standardization system, and special section 8 is devoted to this topic in [7]. As stated therein, the application of international standards enables the achievement of several objectives. One of them is the use of the benefits of scientific and technological progress. Indeed, most of the advanced technologies, technical solutions, their hardware and software implementations come to us from abroad. Naturally, related terms come along with them as well. Another objective is to standardize terminology used within the framework of trade, economical, scientific and technological cooperation with other countries.

Also, a substantial part of our standards is currently harmonized with international standards. Therefore, even if the development of a terminology standard in dependability is not going to take an analogous international standard into account by any means, the terms and definitions from it will anyway find their way into our country along with other standards, this leading to undesirable collisions. Let us consider a typical example. In a widely used Russian standard [8], the term "dependability" has the following definition: "ability to perform as and when required" (its source will be provided below).

Unfortunately, the complete harmonization of the national terminology standard in dependability with an international standard is hardly possible at present, since it will be in plain contradiction to the legacy principle. The point is that the standardization of dependability terminology in our country began earlier than globally, though our representatives did not take any active part in the work of respective international organizations, so our experience has not practically been taken into account, and this has resulted in the fact that the Russian and international standards differ in some important positions (examples will be provided below). Thus, it is reasonable, on the one hand, to go step by step in the direction of alignment of the Russian terminology with the international one, and on the other hand, to make attempts to include those terms from our standard into the international one, which are missing therein.

A major role in dependability standardization belongs to IEC, namely to its Technical Committee (TC) 56, which is called "Dependability". By agreement with ISO, it is horizontal, i.e. it develops standards for all areas of technics, not only for electrotechnics. Those who are interested in its history and activities, can be referred to papers [9] and [10] (the author of the first one was the TC 56 chairman at the period of 2008–2017, and the author of the second one is the acting chairman); in Russian there is a publication [11] reflecting its authors' work experience in TC 56.

The current international terminology standard in dependability [12] represents the part 192 of the International Electrotechnical Vocabulary (IEV). IEV contains all the IEC standardized terms and their definitions in English and French. It has a publicly available online version "Electropedia" (http://www.electropedia.org/), where one can also find equivalents of terms in other languages. In particular, terms in dependability are provided in other 9 languages (apart from English and French), but unfortunately, the Russian language is not among them. The overview of the standard [12] in Russian, its comparison with the previous version of a similar standard [13] and the Russian terminology standard is presented in [14].

None of the mentioned basic provisions of the Russian standards is fully in line with the international standards. Therein dependability and its constituents (reliability, maintainability, etc.) are defined as an item's abilities, rather than as its properties; the storage and transportation of an item are not taken into account, thus storability not being incorporated into dependability; there are no analogs of perfect state and imperfect state, and no integrated dependability measures mentioned above. Nonetheless, as far as the relationship between a property and an ability is concerned, in the author's opinion, the difference is not so vital [2], and by the way, even in IEC TC 56, when defining dependability as the ability of an item, they say about it as a property [9].

It is [12] that was the source of the abovementioned definition of dependability from [8]. The thing is that standard [8] is identical to ISO 9000:2015, and during the development of this international standard they took the definition from [12] considering the major role of IEC in dependability standardization. By the way, it would be good for the developers of our standards to act in the same way, so that they not invent their own terms and definitions related to dependability, but rather take them from the terminology standard of the series "Dependability in technics".

The first attempt to find a compromise between legacy and alignment with an international standard was GOST R 27.002–2009, which was developed with the basic provisions of international standard [13] taken into account. Unfortunately, this attempt was a failure, as GOST R 27.002–2009 had a wide range of significant drawbacks.

A system of terms therein was poorly coherent and inconsistent. For instance, the definition of dependability taken from [13] used the term "maintenance support", whereas this term was not included into the standard. On the other hand, likewise in our previous standards, GOST R 27.002–2009 had the term "storability", though its relationship with dependability was absolutely unclear. Some of terms (e.g. "imperfect state") took on a meaning different from what was defined in previous national standards and became something common for specialists. For several terms taken from [13], bad Russian equivalents were chosen. Some definitions from [13] were translated with mistakes (omissions, wrong cases, etc.), resulting in distortion and ambiguity of the meaning. As was mentioned above, the number of terms therein is too large. Moreover, standard [13] taken as a basis had become obsolete by the time, and IEC TC 56 had been actively working over a new standard that was to replace it (unfortunately, the work took longer than it had been expected initially, and [12] was adopted only in 2015).

GOST R 27.002–2009 got severely criticized by the scientific and technical community, which resulted in the fact that Rosstandart made a decision to suspend GOST R 27.002–2009 and to renew the validity of the interstate standard GOST 27.002–89 (order dated 29.11.2012 № 1843-st). In parallel, they began developing a new terminology standard, which became the interstate standard GOST 27.002–2015. Unfortunately, while bringing this standard into force as a national standard of the Russian Federation (order dated 21.06.2016 № 654-st), Rosstandart did not cancel the contradicting GOST R 27.002–2009, as it should have been done according to Par. 6.2 of GOST R 1.8–2011 [14].

During the development of GOST 27.002–2015, a new attempt to find a compromise between legacy and alignment with IEC new standard [12] was made, however, compared to the previous time, the preference was given to legacy. At the same time, it included some terms from [12], which had been missing in our standards before.

Consistency with other general technical standards

One more principle of standardization specified by the Federal law "On standardization in the Russian Federation" is consistency of national standards. Indeed, contradictions between standards create difficulties for those who apply these standards ("What to believe?" [5]), shatter confidence and respect in the entire system of standardization. Therefore, a terminology standard in dependability ought to be consistent with basic general technical standards, in particular, with standards of "Unified system for design documentation" (Russian abbreviation ESKD, GOST 2.), "Unified system of technological documentation" (ESTD, GOST 3.), etc.

One may think that this requirement is obvious and should be clear for everyone. Unfortunately, in practice there have been cases when this principle was violated. For example, in GOST R 27.002–2009 the definition of the term "product" was fundamentally different from the definition of the same term in ESKD (in more detail it was covered in [16]).

The definition of dependability as a property is in line with this principle. Indeed, a general technical standard [17] defines the quality of products as the entirety of a product's properties underlying its capability to satisfy certain needs as to its purpose. One of these properties is dependability.

It is worth noting that consistency of standards should be provided by both ways. General terms used in dependability standards should have the same meaning as in basic general technical standards where these terms are in place. On the other hand, dependability terms in all standards should be used in line with the way they are defined in the terminology standard of the standards series "Dependability in technics".

Internal consistency and logical coherence

Even more obvious is the requirement for internal consistency and logical coherence of a standard itself. Unfortunately, sometimes this principle was violated as well. For example, in GOST R 27.002–2009 the terms "availability state" and "availability time" were by no means related to availability measures (availability factor, etc.); "perfect state" and "imperfect state" were not opposite to each other, whereas "imperfect state" was opposed by "on-call state".

There is also some logical inconsistency in [12], which is admitted even by its developers [9]. Dependability is therein defined as an ability (property) of an item, although it includes maintenance support performance, which is defined as the efficiency of an organization in relation to maintenance support, i.e. is not an item's property, but rather conditions under which it is used.

Generality and universality

The fact that a general dependability standard should be applicable to all branches of technics requires the maximum generality and universality of terms and definitions to be specified in it. They should be specified in such a way that they could be used in all industries. According to [7, Par. 6.3], the attributes introduced into a definition should be inherent in all objects comprising the scope of a term.

Of course, the application in various industries can necessitate further specification and elaboration. Therefore, at the beginning of all recent standards (since 1989) it is said that definitions specified therein can be changed if necessary, by introducing derived attributes, elaborating the meaning of terms used therein, specifying objects comprising the scope of a defined term. However, these changes should not modify the scope of terms defined in a standard.

Using this line of thinking, let us take a look at what the term "dependability" is related to, i.e. whose property or ability it is. For that purpose, GOST 13377–67 and GOST R 27.002–2009 used the term "product", while all others of our standards used the term "item". The scope of these terms and the relationship between them were analyzed in detail in [16], therefore, the issue will be considered briefly here.

Quite naturally that GOST 13377–67 and GOST R 27.002–2009 used rather general definitions of a product, the first one in its preamble, the second one as one of the main terms. However, they were different from the term in ESKD and, consequently, did not satisfy the principle of consistency. That is why there used to be cases when representatives of some industries dealing with buildings and constructions, power supply systems, telecommunications networks, etc. said that the standard did not apply to them, since their objects were not products. And they meant a product in a typical way, i.e. in line with the definition of ESKD.

Therefore, in all others of our standards the definition of dependability and other terms is given relation to an item (ob'ekt in Russian). By the way, among the languages in Electropedia, in which dependability terminology is provided, there are two Slavonic languages, namely Czech and Polish, and in those languages the terms *objekt* and *obiekt* are used respectively for this notion. At the same time, since a product is a specific case of an item, it is not prohibited to write about dependability or failures of products in industry documents, if the scope of consideration is limited to them.

The definition of the term "dependability" should also be general and universal to the maximum extent. Various approaches to the specification of such a definition were analyzed in detail in [2]. In particular, it compared two types of definitions: parametric and functional definitions. It was noted that a functional definition is more general, i.e. it is suitable for a wider range of situations. The possibility of cases when a parametric definition is not reasonable or possible was already mentioned in GOST 27.002–89 in the explanatory note to the term "dependability", though a parametric definition was used therein as a basic one, while a functional definition was provided just in the explanatory note. Therefore, a generality principle was violated in this case: the definition contained parameters pertaining not to all objects. The developers of GOST 27.002–2015 decided to follow this principle and, thus, chose a functional definition as a basic one, while providing a parametric definition in one of the notes to the term "dependability". Such a choice was also a step in the direction of alignment with [12] (the definition of dependability from it was given above).

When discussing this aspect, sometime one has to come across with the position that can be called "industrial egocentrism". Discussants request that a standard should incorporate the terms and definitions that are used in their industry and do not take the arguments of other industries' representatives, for whom such definitions are not suitable, refuse to reach a compromise by finding mutually acceptable universal solutions. In general, the importance of compromises in standardization (and not limited to it) was well written about in [18].

Conclusions

The paper has formulated the main principles that a general dependability terminology standard should conform to: continuity in relation to previous similar national standards, alignment with the international IEC standard, consistency with other general technical standards, internal consistency and logical coherence, generality and universality to meet the needs of all industries.

The author appeals to all concerned specialists to share their opinions and make constructive suggestions about these principles.

References

[1] Netes V.A., Tarasyev Yu.I., Shper V.L. Current issues of terminology standardization in dependability. Dependability. 2014:2:120-123.

[2] Netes V.A., Tarasyev Yu.I., Shper V.L. How we should define what "dependability" is. Dependability. 2014;4:15-26.

[3] Alpeiev A.S. Definitions of dependability. Dependability. 2016;16(4):43-45.

[4] Pokhabov Yu.P. On the definition of the term "dependability". Dependability. 2017;17(1):4-10.

[5] Yershov G.A., Semerikov V.N., Semerikov N.V. [What to believe? On the system of standards "Dependability in engineering"]. Standarty i Kachestvo. 2018;8:14-19. (in Russ.)

[6] Netes V.A. [How to regain trust? About the system of standards "Dependability in engineering"]. Standarty i Kachestvo 2019;2:19-24. (in Russ.)

[7] R 50.1.075–2011. [Development of standards on terms and definitions]. (in Russ.)

[8] GOST R ISO 9000-2015. Quality management systems. Fundamentals and vocabulary. (in Russ.)

[9] Loll V. Short History of IEC TC56 from 1965 to 2017. 6th Workshop on Open Systems Dependability. Tokyo; 2017.

[10] Van Hardeveld T. Developments in Dependability Standardization. 6th Workshop on Open Systems Dependability. Tokyo; 2017.

[11] Bogdanova G.A., Netes VA. [IEC/TC 56: standardization for dependability]. Metody Menedzhmenta Kachestva. 2009;5:44-47. (in Russ.)

[12] IEC 60050-192:2015. International electrotechnical vocabulary – Part 192: Dependability.

[13] IEC 60050-191:1990. International Electrotechnical Vocabulary (IEV) – Part 191: Dependability and quality of service.

[14] Netes V.A. New international standard for dependability. Dependability. 2016;3:54-58.

[15] GOST R 1.8-2011. Standardization in Russian Federation. Interstate standards. Rules of developing, application, renovation and cease of application in Russian Federation.

[16] Netes V.A. Item in dependability: definition and content of the concept. Dependability 2019;19(4):3-7.

[17] GOST 15467–79. Product quality control. Basic concepts. Terms and definitions

[18] Aronov I.Z. [The art of compromise]. Metodi Menedzhmenta Kachestva. 2012;6:1. (in Russ.)

About the author

Victor A. Netes, Doctor of Engineering, Professor of the Department of Telecommunications Networks and Switching Systems, Moscow Technical University of Communication and Informatics, Russian Federation, Moscow, e-mail: v.a.netes@mtuci.ru

The author’s contribution

Netes V.A. analyzed Russian and international dependability terminology standards and proposals for their improvement, identified their shortcomings and deviations from the requirements of regulatory documents for standardization, formulated the main principles a general dependability terminology standard is to comply with.

On the terminology of dependability

Viktor S. Mikhailov, *D.I. Mendeleev Central Research and Design Institute of Chemistry and Mechanics, Russian Federation, Moscow*
Mvs1956@list.ru

Viktor S. Mikhailov

Abstract. Aim. *Currently, there is a fully-fledged system of Russian dependability standards, the GOST R 27.xxx series. However, due to the suspension of the terminology standard (GOST R 27.002-2009) this system is now incomplete. In this situation, a compromise solution can be found with dual designation in the current dependability standard in Russia. The aim of the paper is to define the proposals for improved basic terminology in dependability.* **Methods.** *The paper uses methods of system analysis in respect to dependability terminology. The last decade was marked by active discussions regarding dependability terminology. Not only particular definitions, but the definition of the term "dependability" itself are addressed. The dependability terminology in the Russian Federation is currently represented in two standards: the Russian GOST R 27.002-2009 (suspended indefinitely) and the interstate GOST 27.002-2015. This paper continues the discussion regarding a limited set of concepts and terms that interest the author most. Such concepts as item, entity, failure, property, ability, calculation, estimation, prediction, requirements for dependability are examined. It is noted that the concept of technical entity is based on the product, the study object as a finished result of some technical activity, i.e. to make and at the same time provide the product with the ability to perform certain functions. It is shown that a product's properties characterize its abilities, therefore, while identifying, the focus should be on the ability of a product provided with properties (features) required for the performance of certain functions. The features (properties) themselves are primary only for the purpose of identification of the entity's required ability and are secondary for the purpose of dependability identification. It is demonstrated that there is no need to substitute the concepts of "calculation" and "estimation". The correctness of the definition of "prediction" in the Russian standard GOST R 27.002-2009, i.e. a computational process aimed at predicting the values of quantitative characteristics, is noted.* **Conclusions.** *Based on the terminological analysis performed in the paper, the following proposals were developed. Dependability terminology should be complemented with the definition of entity. An entity should be understood as a functional unit provided with abilities defined by the required properties. A failure should be understood as an event consisting in the disruption of the product's up state. The concept of item should be interpreted as in GOST R 27.001-2009: an item (entity, system) that is considered individually in terms of dependability, that consists of hardware and software or their combinations. The terms of dependability, reliability, durability, etc. should be defined as the identified ability of the product to perform the required function in the given circumstances. The term "requirements for dependability" should be specified in the dependability standards. The term "dependability estimation" should not be introduced in the interstate standard GOST 27.002-2015.*

Keyword*: item; entity; failure; property; ability; calculation; estimation; prediction; requirements for dependability.*

For citation: *Mikhailov V.S. On the terminology of dependability. Dependability. 2020;2: 24-27. https://doi.org/10.21683/1729-2646-2020-20-2-24-27*

Received on: *21.12.2020* / **Upon revision:** *16.04.2020* / **For printing:** *17.06.2020.*

Introduction

The last decade was marked by active discussions regarding dependability terminology [1 – 6]. Not only particular definitions, but the definition of the term "dependability" itself are addressed [1 – 6]. Due to the fact that the Russian standard GOST R 27.002-2009 [7] is suspended indefinitely dependability terminology in the Russian Federation is currently based on the interstate standard GOST 27.002-2015 [8]. GOST 27.002-2015 was adopted by the Interstate Council for Standardization, Metrology and Certification (proceedings of December 28, 2015 no. 83-P). It was voted for by Armenia, Kazakhstan, Kyrgyzstan, Moldova and the Russian Federation. The list of voters does not include Belarus who is the founder of the Customs Union.

This paper continues the discussion regarding a limited set of concepts and terms that interest the author most.

GOST 27.002-2015 or GOST R 27.002?

Currently, there is a fully-fledged system of Russian dependability standards, the GOST R 27.xxx series. However, due to the suspension of the terminology standard (GOST R 27.002-2009) this system is now incomplete. In this situation, a compromise solution can be found with dual designation in the current dependability standard in Russia. When applied in Russia, the standard is designated GOST R 27.002-2015, while for the purpose of interstate relations it is designated GOST 27.002-2015. While the national dependability standard GOST R 27.002-2009 is suspended, such solution would allow solving many problems.

Item, entity and failure

The concept of technical **entity** (hereinafter referred to as entity) is based on the product, the study object as a finished result of some technical activity that consists in making and at the same time providing the product with the ability to perform certain functions. For instance: a microscope has a high resolution, a diode is able to withstand high reverse voltage, an instrument is able to equalize voltage, a surface with a high reflective power. For that reason, an entity should be understood as a functional unit out of a produced set of products provided with required abilities defined by the necessary properties. That complies with the definition of entity per GOST R 27.002-2009 [7].

In accordance with Item 49 of GOST R 27.002-2009, a **failure** is the loss by an entity of the ability to perform the required function. The note to Item 49 states that "a failure is an event that causes a fault". On the other hand, in accordance with the interstate standard GOST 27.002-2015 a **failure** is "an event that consists in the disruption of operability of an **item**". Given that a fault does not always cause a failure (for example, minor chipping or dent on the surface of equipment, broken cap of a signal light, etc.), the definition according to the interstate standards GOST 27.002-2015 is preferable. GOST R P 27.002-2009 does not define item. Let us clarify what the concept of item consists in.

N.E. Yatsenko: "Item: 1. In philosophy, any phenomenon existing independently from human consciousness. 2. In a general sense, an object, phenomenon that people try to get to know and the human activity is directed at". "Object: 1. Any material phenomenon, a thing. 2. Something the thought, an action or a feeling is directed at" [9].

An item is a process or a phenomenon that causes a problem situation and that a researcher chose to examine. An object is something that is within an item. An item is the part of scientific knowledge a researcher is dealing with. The study object is the aspect of the problem, researching which we get to know a whole item by identifying its primary, most significant features. As scientific categories, item and study object are the general and the particular [9].

Dependability studies the quality of an item or, ultimately, quality as a property of an object, i.e. the feature that constitutes the identifying characteristic of the object of cognition. Therefore, the dependability terminology should include the concept of "**entity**" as study object, as the particular and the specific.

GOST 27.002-2015 introduces the definition of "technical item": "The subject matter covered by the terminology of dependability in engineering." Such definition of item is not universally accepted and raises a few questions:

- what should be the scope of the terminology of dependability in engineering;
- if the coverage is not to be complete, what should it encompass.

The concept of item is best defined in GOST R 27.001-2009 [10]: an item (entity, system) that is considered individually in terms of dependability, that consists of hardware and software or their combinations.

A developer must make a choice as to which term to use, item or entity, based on the need for a terminology.

On the term "dependability requirements"

Despite the fact that the term "requirements for dependability" [11] has established itself, the term "dependability requirements" can be frequently encountered [12]. According to [11], the dependability characteristics (requirements for dependability) can be specified (raised) by a supplier or a consumer. However, out of term "dependability requirements" follows that the requirements for the dependability of an entity are raised by the entity itself, which is nonsense. A supplier, by specifying requirements for dependability in the form of specified qualities (properties, indicators) as part of the entity's design, may expect from such entity the required ability to perform certain functions. Therefore, the concept and term "requirements for dependability" should be specified in the section dedicated to development.

Property or ability

A **property** is a feature [13, 14] that constitutes the identifying characteristic of the object of cognition [9]. An entity's properties characterize its abilities, therefore while identifying the dependability the focus should be on the **ability** of the entity provided with properties (features) required for the performance of certain functions. The features (properties) themselves are primary only for the purpose of identification of the entity's required ability and are secondary for the purpose of dependability identification. Therefore, the terms dependability, reliability, durability, etc. should be defined as the **ability** of the product to perform the required function in the given circumstances, which is in compliance with GOST R 27.002-2009. It should be noted that the interstate standard GOST 27.002-2015 defines the terms dependability, reliability, durability, etc. as the **property of an item**, which is incorrect due to the secondary status within the hierarchy of definitions. Let us give an example of how certain definitions of abilities are built.

Reliability is the identified ability of an entity to continuously perform the required function within the defined period of time (operation time) under the given conditions. This ability is defined by the entity's properties that are characterized by the following indicators: mean time between failures, mean time to failure, probability of no-failure, gamma-percentile time to failure, gamma-percentile time between failures, failure rate, assessed failure rate.

Availability is the identified ability of an entity to perform the required function under the given conditions assuming that the required external resources are provided. This ability is defined by the entity's properties that are characterized by composite indicators of availability.

Durability is the identified ability of an entity to perform the required function until the onset of the limit state under the given conditions of operation and maintenance. This ability is defined by the entity's properties that are characterized by the following indicators: average operating life, gamma-percentile life, mean lifetime, gamma-percentile lifetime.

Dependability is the ability of an entity to perform the required function in the given circumstances. An entity's ability is defined by particular abilities: availability, reliability, durability, maintainability and storability.

The list goes on. Thus, the following hierarchical structure is shown: dependability, ability, property, indicator.

Calculation, estimation and prediction

The concept of "dependability estimation" introduced in the interstate standard GOST 27.002-2015 as the identification of the numerical values of the indicators of an items' dependability, is broad and requires additional specifications in the standard. Dependability estimation implies that the identification of the numerical values of the dependability indicators is performed through either **calculation** based on reference data, or **estimation** based on the results of testing, where estimation means statistical estimation. According to GOST R 50779.10-2000, statistical estimation (the word "statistical" is always omitted) is understood as the statistics used for the purpose of estimating the population parameter. Statistics is the function of selective values [13]. The population parameter is some dependability indicator. Since in the dependability theory the word "estimation" is reserved to statistical estimation, the interstate standard GOST 27.002-2015 should not feature the term "dependability estimation". In the last resort, the term "Identification of numerical values of dependability indicators" should be introduced to imply the identification of the numerical values through calculation based on reference data or estimation based on test results. In the Russian standard GOST R 27.002-2009, there is no such term as "dependability estimation", since there is no need to substitute the concepts of "calculation" and "estimation".

The Russian standard GOST R 27.002-2009 sets forth a correct definition of "**prediction**", i.e. a computational process aimed at predicting the values of quantitative characteristics. Given the above, the concept of "**prediction** of dependability" introduced in the interstate standard GOST 27.002-2015 as the preliminary **estimation** of dependability based on prior experience or statistics should be modified in accordance with the Russian standard GOST R 27.002-2009.

Conclusions

1. Dependability terminology should be complemented with the definition of entity.

2. An entity should be understood as a functional unit provided with abilities defined by the required properties.

3. A failure should be understood as an event consisting in the disruption of an entity's up state.

4. The concept of item should be interpreted as in GOST R 27.001-2009: an item (entity, system) that is considered individually in terms of dependability, that consists of hardware and software or their combinations.

5. The terms of dependability, reliability, durability, etc. should be defined as the identified ability of the product to perform the required function in the given circumstances.

6. The term "requirements for dependability" should be specified in the dependability standards.

7. The term "dependability estimation" should not be introduced in the interstate standard GOST 27.002-2015.

References

[1] Netes V.A., Tarasyev Yu.I., Shper V.L. Current issues of terminology standardization in dependability. Dependability. 2014;2:120-123.

[2] Netes V.A., Tarasyev Yu.I., Shper V.L. How we should define what "dependability" is. Dependability. 2014;4:15-26.

[3] Netes V.A. New international standard for dependability. Dependability. 2016;3:54-58.

[4] Alpeiev A.S. Definitions of dependability. Dependability. 2016;16(4):43-45.

[5] Pokhabov Yu.P. On the definition of the term "dependability". Dependability. 2017;17(1):4-10.

[6] Netes V.A. Item in dependability: definition and content of the concept. Dependability 2019;19(4):3-7.

[7] GOST 27.002-2009. Dependability in technics. Terms and definitions. Moscow: Standartinform; 2011. (in Russ.)

[8] GOST 27.002-2015. Dependability in technics. Terms and definitions. Moscow: Standartinform; 2016. (in Russ.)

[9] Yatsenko N.E. [Explanatory dictionary of social science terminology]. Saint Petersburg; 1999. (in Russ.).

[10] GOST 27.001-2009. Dependability in technics. Dependability management system. Basic principles. Moscow: Standartinform; 2010. (in Russ.)

[11] GOST 27.003-2011. Dependability in technics. Dependability management. Guide to the assignment of dependability technical requirements. Moscow: Standartinform; 2013. (in Russ.)

[12] 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; 2016. (in Russ.)

[13] GOST R 50779.10-2000. Statistical methods. Probability and general statistical terms. Terms and definitions. Moscow: Standartinform; 2008. (in Russ.)

[14] GOST R 50779.11-2000. Statistical methods. Statistical quality control. Terms and definitions. Moscow: Standartinform; 2008. (in Russ.)

About the author

Viktor S. Mikhailov, Lead Engineer, D.I. Mendeleev Central Research and Design Institute of Chemistry and Mechanics (FGUP CNIIHM), 16a Nagatinskaya St., 115487, Moscow, Russian Federation, e-mail: Mvs1956@list.ru

The author's contribution

The author analyzed the current interstate and Russian dependability terminology standards. Based on the analysis, the author proposed the following: within the dependability terminology, the definitions of entity, failure, item (per GOST R 27.001-2009), as well as of the term "requirements for dependability" are to be made more specific; the terms dependability, reliability, durability, etc. are to be defined as the identified ability of an entity to perform the required function in the given circumstances; the term "dependability evaluation" is not to be introduced into the interstate standard.

Original article
https://doi.org/10.21683/1729-2646-2020-20-2-28-34

Accounting for the effect of correlations by modulo averaging as part of neural network integration of statistical tests for small samples

Alexander I. Ivanov[1*], Andrey G. Bannykh[2], Yulia I. Serikova[2]
[1]Penza Research and Design Electrical Engineering Institute, Russian Federation, Penza, [2]Penza State University, Russian Federation, Penza
**ivan@pniei.penza.ru*

Alexander I. Ivanov

Andrey G. Bannykh

Yulia I. Serikova

Abstract. *The* **Aim** *of the paper is to demonstrate the advantages of taking into consideration real correlations by means of their symmetrization, which is significantly better than completely ignoring real correlations in cases of statistical estimation using small samples.* **Methods.** *Instead of real correlation numbers different in sign and modulo, identical values of correlation numbers moduli are used. It is shown that the equivalence of transformation to symmetrization is subject to the condition of identical probabilities of errors of the first and second kind for asymmetrical and equivalent symmetrical correlation matrices. The authors examine the procedure of accurate calculation of equal data correlation coefficients by trial and error and procedure of approximate calculation of symmetrical coefficients by averaging the moduli of real correlation numbers of an asymmetrical matrix.* **Results.** *The paper notes a practically linear dependence of equal probabilities of errors of the first and second kind from the dimension of the symmetrized problem being solved under logarithmic scale of the variables taken into consideration. That ultimately allows performing the examined calculations in table form using low-bit, low-power, inexpensive microcontrollers. The examined transformations have a quadratic computational complexity and come down to using pre-constructed 8-bit binary tables that associate the expected probability of errors of the first and second kind with the parameter of equal correlation of data. All the table calculations are correct and do not accumulate input data round-off errors.* **Conclusions.** *The now widely practiced complete disregard of the correlations when performing statistical analysis is very detrimental. It would be more correct to replace the matrices of real correlation numbers with symmetrical equivalents. The approximation error caused by simple averaging of the moduli of coefficient of asymmetrical matrices decreases as the square of their dimension or the square of the number of neurons that generalize classical statistical tests. When 16 and more neurons are used, the approximation error becomes negligible and can be disregarded.*

Keywords: *replacement of statistical test with equivalent neurons; multicriteria statistical analysis of small samples; accounting for the effect of correlations; symmetrization of correlation matrices.*

For citation: *Ivanov A.I., Bannykh A.G., Serikova Yu.I. Accounting for the effect of correlations by modulo averaging as part of neural network integration of statistical tests for small samples. Dependability. 2020;2: 28-34. https://doi.org/10.21683/1729-2646-2020-20-2-28-34*

Received on: *14.06.2019* / **Revised on:** *07.05.2020* / **For printing:** *17.06.2020.*

The problem of application of classical statistical tests with small samples

Pearson's statistical hypothesis test was created in 1900 and proved to be very effective. Naturally, in 1900 computer technology did not exist, so only relatively computationally simple tests could be created, studied and used. Person's test set the trend in statistical study for decades. As the result, hundreds of mathematicians in the XX century created about 200 statistical tests applicable under various limiting conditions.

Unfortunately, all known statistical tests provide poor results with small samples. In such areas as biometrics, medicine, biology, economy, the samples of actual data are small. This circumstance impedes reliable statistical estimation. Thus, Pearson's chi square test over a 21-experiment sample yields poorly shared states for normal data and uniformly distributed data (Fig. 1).

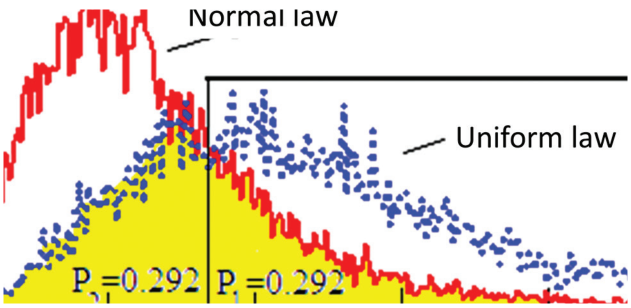


Fig. 1. The threshold of the chi square neuron $k = 7.5$ was defined based on the matching values of the probabilities of errors of the first and second kind $P_1 = P_2 = P_{EE} = 0.292$

The confidence probability of detection of normal data under the adopted conditions is not high: 0.708 (1 – 0.292 = 0.708), which makes practical application impossible. Practically acceptable confidence probabilities can be obtained only with large samples [1] of 200 and more examples.

Unlike in 1900, we possess the capability to multiply the complexity of calculations as part of statistical analysis. For instance, we can use several different statistical tests. We can associate artificial neurons with each statistical test [2, 3] and use them simultaneously. Fig. 1 shows one of the chi square neuron settings with 5 inputs. Each of such inputs of the neuron analyzes one of the bins of the histogram of the tested sample.

The output comparator of the artificial neuron is set in such a way as to allow the probability P_1 of errors of the first kind to be close to the probability P_2 of errors of the second kind. This technique allows reducing the dimension of the problem at hand by replacing two variables with one $P_1 = P_2 = P_{EE}$. Formally, the variables P_1 and P_2 may differ from each other, however if they are artificially made identical (symmetrical), we will be able – by means of symmetrization – simplify the calculations.

Table 1 shows the values of matching probabilities of errors of the first and second kind for 8 different statistical tests (neurons), where:

χ^2 is the chi square test [2, 3, 4, 5];

ad^2 is the Anderson-Darling test [4, 5];

adL is the logarithmical form of the Anderson-Darling test [4, 5];

sg is the geometric mean test [6, 7, 8];

sg_d is the differential-integral variant of the geometric mean test [5, 7];

ω^2 is the Cramér-von Mises test [5, 7];

ω^2_c is the Smirnov-Cramér-von Mises test [4, 5];

su^2 is the Shapiro-Wilk test [5, 9].

It is obvious that, using eight statistical tests instead of one is made easily possible through modern computer technology with low-bit microcontrollers (4-bit processors of RFID identification cards, 8-bit processors of modern controllers, low-power processors of SIM cards and microSD cards). The neural network implementation of such engineering solution will result in code condition 00000000, when all tests (all neurons) make a decision in favour of the normal distribution law of small sample values. If all neurons make a decision in favour of even distribution of values, the output code will be 11111111.

In practice, the bits of a neural network's output code do not always have identical states. In such cases the decision is made based on the majority of observed states. In other words, all codes with a majority of states 0 are taken as the decision of detection of normal value distribution in a 21-experiment input sample.

All transformations that can be performed using low-bit microcontrollers can also be performed on personal computers using appropriate software. Such approach is acceptable as part of scientific research; however, it cannot be used in large-scale biometrical calculations. In order to ensure compliance with cyber security requirements, biometric neural network calculations and cryptographic transformations must be performed only in a trusted computational environment, normally implemented on a low-bit, low-energy, low-cost microcontroller.

Rough statistical estimation under the hypothesis of complete absence of correlations between the responses of generalized statistical tests

Table 1 shows data of only 8 statistical tests (statistical neurons). For that reason, we can conduct a numerical experiment and identify the probabilities of each of the

Table 1. Values of error probability for criteria of statistical hypothesis testing for 21-test samples

No., i	1	2	3	4	5	6	7	8
Test	χ	ad^2	adL	sg	sg_d	ω	ω	su^2
P_{EEi}	0.292	0.349	0.320	0.320	0.278	0.351	0.311	0.322

256 code conditions. If we increase the number of neurons from 8 to 256, calculating the probabilities of all code conditions will be technically very complicated. As the number of simultaneously working neurons grows, the complexity of such computational task increases exponentially.

As we do not know how to precisely take into consideration the effect of correlations between the bits of the output code, we will opt for a simplification and accept the hypothesis of independence of the analyzed data. In this case the mutual reinforcement of the eight tests may be estimated as the product of equally possible errors from Table 1:

$$P_{EE(8)} = \prod_{i=1}^{8} P_{EEi} \approx 0,0001. \quad (1)$$

The geometric mean of the probabilities P_{EE} of eight tests is 0.316. Assuming that 256 simultaneously used statistical tests are independent, and their harmonic mean is 0.316, we obtain a very optimistic estimate of the probability of errors:

$$P_{EE(256)} \approx \prod_{i=1}^{256} P_{EEi} \approx \left\{ \sqrt[8]{\prod_{i=1}^{8} P_{EEi}} \right\}^{256} \approx 0,316^{256} \approx 10^{-128}. \quad (2)$$

The data of the actual numerical experiment for 8 statistical neurons from Table 1 are about 80 times worse than the optimistic estimate (1). That means that the hypothesis of independence of the conditions of statistical neurons is not applicable to our case. We cannot neglect the existing correlations while performing neural network integration of the many classical statistical tests.

Accounting for correlations through their symmetrization: estimation of the correctness of the hypothesis of equal correlation of the responses of the generalized statistical test

As the real correlations cannot be disregarded, since about 1999, neural network biometrics [10, 11, 12, 13] have been using the practical technique of symmetrization of correlations. The essence of the technique consists in the fact that the actual correlation number matrix is replaced with some equivalent with identical elements out of range:

$$\begin{bmatrix} 1 & r_1 & r_2 & ::: & r_n \\ r_1 & 1 & r_{n+1} & ::: & r_{2n-2} \\ r_2 & r_{n+1} & 1 & ::: & r_{3n-3} \\ ::: & ::: & ::: & ::: & ::: \\ r_n & r_{2n-2} & r_{3n-3} & ::: & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & \tilde{r} & \tilde{r} & ::: & \tilde{r} \\ \tilde{r} & 1 & \tilde{r} & ::: & \tilde{r} \\ \tilde{r} & \tilde{r} & 1 & ::: & \tilde{r} \\ ::: & ::: & ::: & ::: & ::: \\ \tilde{r} & \tilde{r} & \tilde{r} & ::: & 1 \end{bmatrix}. \quad (3)$$

The condition of correct symmetrization (3) comes down to matching probabilities of errors of the first and second kind for the initial asymmetrical model and the final symmetrical model:

$$P_{EE} \approx P_2\left\{\left[r_{i,j}\right]\right\} \approx P_1\left\{\left[r_{i,j}\right]\right\} \approx \tilde{P}_2\left\{\left[\tilde{r}\right]\right\} \approx \tilde{P}_1\left\{\left[\tilde{r}\right]\right\}. \quad (4)$$

For any actual correlation matrix, a symmetrical equivalent correlation matrix can be chosen that would have identical values data correlation coefficient. In order perform an accurate symmetrization, an iterative fitting of parameter $\tilde{r}$ is required. Such approach to the solution of the problem is similar to training artificial neurons through an iterative algorithm by criterion of systems movement towards the fulfillment of condition (4). The computational complexity of such iterative processes strongly depends on the dimension of the problems at hand. It is generally believed that iterative fitting as part of neural network has a polynomial computational complexity (for our case, the polynomial order is always lower than the dimension of the symmetrized matrix).

It is interesting to note that the first approximation of the equal correlation coefficients can be obtained through a simple averaging of the correlation coefficient modules of the initial asymmetrical matrix (this procedure has a quadratic computational complexity):

$$\tilde{r} \approx \frac{2}{n^2 - n} \cdot \sum_{i=1}^{\frac{n^2-n}{2}} |r_i|, \quad (5)$$

where i is the numbers of the correlation coefficients outside the diagonal of the initial asymmetrical correlation matrix.

Table 2. Correlation numbers between pairs of examined statistical tests

	χ	ad^2	adL	sg	sg_d	ω	ω	su^2
χ	**1**	0.423	0.672	0.037	-0.042	0.559	0.401	-0.726
ad^2	0.423	**1**	0.644	0.018	-0.145	0.226	0.393	-0.113
adL	0.672	0.644	**1**	0.056	0.209	0.827	0.832	-0.917
sg	0.037	0.018	0.056	**1**	0.132	0.414	0.402	-0.212
sg_d	-0.042	-0.145	0.209	0.132	**1**	-0.242	-0.142	-0.041
ω	0.559	0.226	0.827	0.414	-0.242	**1**	0.885	-0.667
ω	0.401	0.393	0.832	0.402	-0.142	0.885	**1**	-0.764
su^2	-0.726	-0.113	-0.917	-0.212	-0.041	-0.667	-0.764	**1**

It is obvious that formula (5) is an approximation, therefore it is required to evaluate the approximation error $\Delta\tilde{r}$ as the dimension function *n* of the matrix. In order to evaluate the rate of error reduction, let us use the correlations of the 8 neural network implementations of statistical tests, whose data is given in Table 2.

The correlation data from Table 2 can be used in estimating the degree of convergence of the examined computational operation. For that purpose, it will suffice to randomly select sets of three out of the eight statistical tests and apply approximate relationship (5) to their data. The histogram of the results of such calculations is shown in Fig. 2 (red line).

This procedure must also be done with sets of five randomly selected from the data of Table 2. As the result, we obtain a histogram of data also shown in Fig. 2 (blue line).

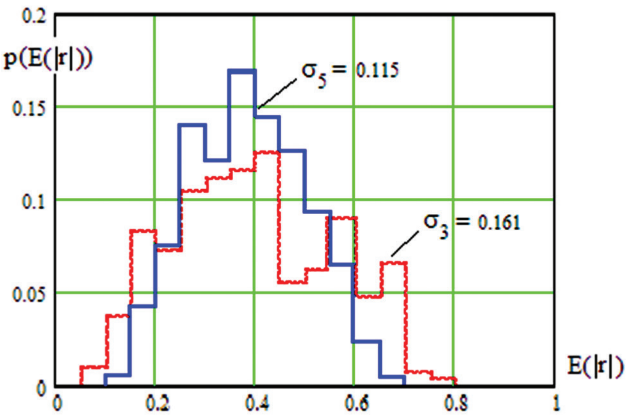


Fig. 2. Histogram of the value distribution of effective moduli of correlation numbers of non-repeating sets of three and five statistical tests from Table 2

Fig. 2 shows that as matrix dimension grows, the standard deviate of data decreases from value σ_3=0,161 to value σ_5=0,115. As the matrix dimension further grows, the distributions of possible values of effective moduli contract. Additionally, normalization of the distributions of possible values of calculation errors $\Delta\tilde{r}$ of symmetrization can be observed.

Numerical estimation of convergence of the symmetrization procedure of correlations of actual biometric data

It should be noted that activities aimed at neural network integration of several statistical tests started only recently [2, 4, 5] and, as consequence, actual statistical data is not yet sufficient. In neural network biometrics the situation is completely different [10, 11, 12, 13]. The biometric neural network authentication technology has been in active development in Russia and other countries since the beginning of the XXI century. As consequence, large anonymized biometric databases have been created using various technical methods, however, they cannot be used due to ethical limitations. Access to such reliable information if restricted both in Russia and abroad.

Ethical restrictions are removed if the problem of access to large volumes of reliable biometric information is solved using the BioNeiroAvtograf simulation environment [14, 15]. That is a free software product that is designed in such a way as to allow Russian-speaking universities to organize their educational process. The product analyzes the dynamics of handwritten reproduction of letters and/or words using a mouse or any graphic tablet. Using two-dimensional Fourier transform, BioNeiroAvtograf extracts 416 biometric parameters from handwriting dynamics and based on the GOST R 52633.5-2011 standard trains a single-layer 256-neuron network.

All data on the biometric parameters, weighting parameters and neuron connections are observable [15] (stored in viewable *.txt files). Using that data, let us generate a training database out of 30 examples of the handwritten

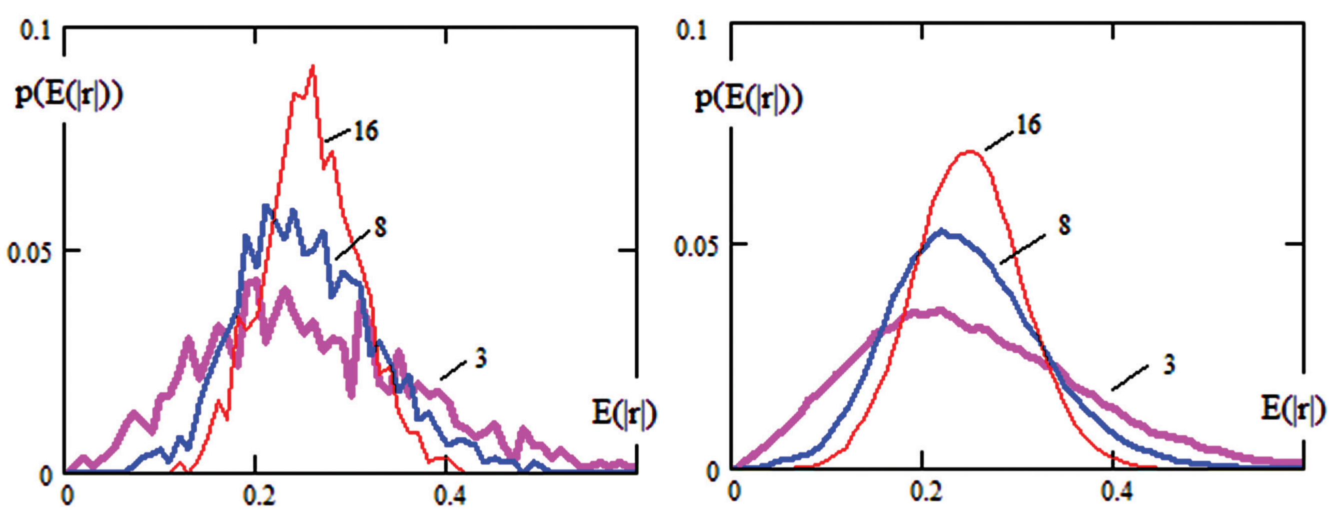


Fig. 3. Examples of distribution of symmetrization parameter values $\tilde{r}$ without smoothing (left part of the figure) and after smoothing (right part of the figure) for matrix dimensions from 3 to 16

word "Penza" in one person's handwriting. Loading data on 30 examples of 416 biometric parameters in MathCAD enables us building a 416 × 416 matrix of correlation numbers. Ultimately, we obtain an amount of data that is much larger than in an 8 × 8 matrix of Table 2.

That allows randomly generating 1024 samples of 3 biometric parameters each and modulo-averaging their coefficient correlation numbers. The resulting distribution of the values of symmetrization results is given in Fig. 3. Similar distributions are shown in the figure for random samples of 8 and 16 biometric parameters.

Fig. 3 shows that the constructed distributions normalize sufficiently quickly. In case of symmetrization of correlation coefficients of a 16 × 16 or larger matrix, the distribution can be considered to be normal. In other words, the distributions are normalized sooner than for the chi square test. It is allowable to replace asymmetrical chi square distributions with normal ones only when 32 and more parameters are taken into consideration. The matters of approximation of chi square distributions by other laws are examined in more detail in [16]. The effect of data normalization for the considered symmetrization procedures ensues sooner as compared with normalization of data of a well-researched chi-square test.

Another important feature of symmetrization is that the uncertainty introduced by this simplification monotonously declines $\sigma_3 > \sigma_4 > ... > \sigma_{256}$. For that exact reason, accounting for mutual correlations for vectors of the length of 256 binary states of a long password or cryptographic key enables sufficiently accurate predictions if a simple symmetrization procedure is used [12, 13]. In the first approximation it can be considered that the uncertainty decreases proportionally to $\sqrt{n^2 - n}/\sqrt{2}$. That means that the standard deviation σ_3=0,161 (see Fig. 2) in case of neural network integration of 100 statistical tests must decrease about 50 times to $\sigma_{100} \approx$ 0,0032.

Simple nomogram for predicting the operational quality of neural network integrations of statistical tests of various dimension

A sufficiently accurate prediction of the attainable probabilities of error under various conditions is possible if simulation tools are used to reproduce the operation of 1, 2, …, 8 neurons under various values of equal correlation coefficients $\tilde{r} = \{0,3,\ 0,4,\ ...,\ 0,7\}$. The results of simulation are well described with a linear approximation in logarithmical coordinates [17] as shown in Fig. 4.

The nomogram in Fig. 4 calculated for the probabilities of error in each of the neuron shows the geometric mean value of the geometrical probabilities of errors in each neuron 0.316. This nomogram easily transforms for other geometric mean values of the probabilities of errors in each of the neurons. For that purpose, it suffices to offset data upwards, if the probability of errors increases and downwards if the probability of errors decreases.

Fig. 3 shows that driving the strength of statistical tests up is less profitable than driving their correlation down. Thus, under correlation value $\tilde{r}$ = 0,4, in a group of 8 examined tests, the probability of errors of 0.001 would require 70 neurons (70 statistical tests). If the level of mutual correlation is reduced to $\tilde{r}$ = 0,3, the same level of probability of failures would require only 17 neurons (17 statistical tests).

Conclusion

In this paper we attempted to show that methods of symmetrization of multidimensional problems are sufficiently simple and efficient. Upon the symmetrization of the error probability estimation of several neurons accounting for their mutual correlations, a simple nomogram is constructed

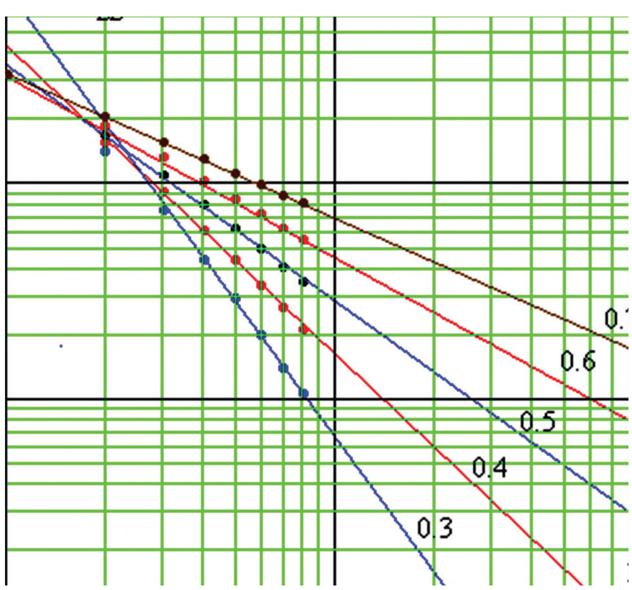


Fig. 4. Nomogram of association of identical probabilities P_{EE} of errors of the first and second kind of neural network integration for averaged values of correlation numbers 0.3, 0.4, 0.5, 0.6, 0.7

that predicts how many neurons will be required in order to achieve a certain probability of errors of the first and second kind.

Currently, the available computing facilities do not impose any restrictions on the number of statistical tests generalized by a neural network. It is only a matter of the degree of mutual correlation of hundreds of classical statistical tests. Unfortunately, most classical statistical tests provide strongly correlated results. The high level of their correlation is the next technical limitation. That indicates the growing relevance of the problem associated with the synthesis of new statistical tests, of which the data is weakly correlated in relation to the majority of known statistical tests.

Nevertheless, it is safe to say that in the years to come the confidence probability of statistical estimations based on small samples should significantly increase. Neural network integration of hundreds of already known statistical tests is not a complex scientific problem, but rather a sufficiently simple engineering task. Additionally, the approximations set forth in this paper allow taking into consideration the effect of correlations on the implementation of computations using low-bit, low-power microcontrollers of RFID cards, SIM cards and microSD cards, which should facilitate widespread application of the examined transformations as part of the solution of problems associated with biometric cryptographic authentication of persons.

References

1. R 50.1.037-2002. Recommendations for standardization. Applied statistics. Rules of check of experimental and theoretical distribution of the consent. Part I. Chi-square criteria. Moscow: Gosstandart Rossii; 2001. (in Russ.)

2. Ivanov A.I., Kupriyanov E.N., Tureev S.V. Neural network integration of classical statistical tests for processing small samples of biometrics data. Dependability 2019;2:22-27. DOI: 10.21683/1729-2646-2019-19-2-22-27.

3. Akhmetov B.B., Ivanov A.I. Estimation of quality of a small sampling biometric data using a more efficient form of the chi-square test. Dependability 2016;16(2):43-48.

4. Volchikhin V.I., Ivanov A.I., Bezyaev A.V., Kupriyanov E.N. The Neural Network Analysis of Normality of Small Samples of Biometric Data through Using the Chi-Square Test and Anderson–Darling Criteria. *Engineering Technologies and Systems*. 2019;29(2):205-217. DOI: 10.15507/2658-4123.029/2019.02.205-217.

5. Ivanov A.I., Bannych A.G., Kupriyanov E.N. et al. Collection of artificial neuron equivalent statistical criteria for their use when testing the hypothesis of normality of small samples of biometric data. Proceedings of the I All-Russian Science and Technology Conference Security of Information Technology. Penza. 2019. 156-164.

6. Perfilov K.A. [Criterion of geometric mean used for validity verification of the statistical hypotheses of biometric data distribution]. Proceedings of the Science and Technology Conference of the Penza Information Technology Security Cluster. Penza. 2014;9:92-93. [accessed 14.04.2020]. Available at: http://www.pniei.penza.ru/RV-conf/T9/S92. (in Russ.)

7. Ivanov A.I., Malygina E.A., Perfilov P.A. et al. The comparison of performance between the criterion mean geometric and the criterion of Cramér-von Mises on a small sample of biometric data. *Models, Systems, Networks in Economics, Engineering, Nature and Society*. 2016;2:155-158.

8. Ivanov A.I., Perfilov K.A., Malygina E.A. Multivariate statistical analysis of the quality of biometric data on extremely small samples using the criteria of the geometric mean tests calculated for the analyzed probability functions. *Measuring. Monitoring. Management. Control.* 2016;2(16):58-66. (in Russ.)

9. Ivanov A.I., Vjatchanin S.E., Malygina E.A. et al. Precision statistics: neuron networking of chi-square test and Shapiro-Wilk test in the analysis of small selections of biometric data. *Proceedings of the International Symposium Dependability and Quality*. 2019;2:131-134. (in Russ.)

10. Ivanov A.I. [Biometric identification of a person based on the dynamics of unconscious movement. Monograph]. Penza: PSU Publishing; 2000. (in Russ.)

11. Ivanov A.I. [Neural network technology of biometric identification of open system users. Author's summary of the Doctor of Engineering thesis per study program 05.13.01 System analysis, management and processing of information. Penza; 2002. (in Russ.)

12. Malygin A.Yu., Volchikhin V.I., Ivanov A.I. et al. [Fast algorithms of testing neural network mechanisms of biometric cryptographic protection of information]. Penza: Penza State University Publishing; 2006. (in Russ.)

13. Akhmetov B.S., Volchikhin V.I., Ivanov A.I., et al. [Algorithms for testing biometric neural network mechanisms of information protection]. Kazakhstan, Almaty: Satbayev University; 2013. [accessed 14.04.2020]. URL: http://portal.kazntu.kz/files/publicate/2014-01-04-11940.pdf. (in Russ.)

14. Ivanov A.I., Zakharov O.S. [The BioNeiroAvtograf simulation environment: a software product (created by the Laboratory of Biometric and Neural Network Technology, freely available since 2009 on the website of the Penza Electrical Engineering Research and Development Institute)]. [accessed 14.04.2020]. Available at: http://пниэи.рф/activity/science/noc/bioneuroautograph.zip. (in Russ.)

15. Ivanov A.I. [Automatic training of large artificial neural networks in biometric applications: a study guide to laboratory works performed in the BioNeiroAvtograf simulation environment]. Penza: Penza Electrical Engineering Research and Development Institute; 2013. [accessed 14.04.2020]. Available at: http://пниэи.рф/activity/science/noc.htm. (in Russ.)

16. Kobzar A.I. [Applied mathematical statistics. For engineers and researchers]. Moscow: FIZMATLIT; 2006. (in Russ.)

17. Ivanov A.I., Lozhnikov P.S., Bannykh A.G. A simple nomogram for fast computing the code entropy for 256-bit codes that artificial neural networks output. Journal of Physics: Conference Series. 2019;1260(2):022003.

About the authors

Alexander I. Ivanov, Doctor of Engineering, Associate Professor, Academic Advisor, Penza Research and Design Electrical Engineering Institute, Russian Federation, Penza, 9 Sovetskaya Str., phone: (841 2) 59 33 10, e-mail: ivan@pniei.penza.ru.

Andrey G. Bannykh, third year post-graduate student, Department of Information Security Technology, Penza State University, 440026, Russian Federation, Penza, 40 Krasnaya Str., 40, phone: (841 2) 36 82 23, e-mail: ibst@pnzgy.ru.

Yulia I. Serikova, third year post-graduate student, Department of Computer Technology, Penza State University, 440026, Russian Federation, Penza, 40 Krasnaya Str., e-mail: julia-ska@yandex.ru.

The authors' contribution

A.I. Ivanov proposed a method for estimating the correctness of procedures of approximate calculation of equal correlation coefficients through simple averaging of real coefficient modules of an asymmetrical correlation matrix.

A.G. Bannykh synthesized 8-bit tables that associate predictable probabilities of errors of the first and second kind with the equal correlation parameter for a predefined number of artificial neurons in a log grid.

Yu.I. Serikova developed the software to supervise the degree of convergence of the computational processes considered in the paper.

Dependability, vol. 20 no. 2, 2020 Original article
https://doi.org/10.21683/1729-2646-2020-20-2-35-42

Development of algorithms of self-organizing network for reliable data exchange between autonomous robots

Alexander V. Ermakov[1*], **Larisa I. Suchkova**[1]
[1]Polzunov Altai State Technical University, Russian Federation, Altai Krai, Barnaul
*tour0@ya.ru

Alexander V. Ermakov

Larisa I. Suchkova

Abstract. *Factors affecting the reliability of data transmission in networks with nodes with periodic availability were considered. The principles of data transfer between robots are described; the need for global connectivity of communications within an autonomous system is shown, since the non-availability of information on the intentions of other robots reduces the effectiveness of the robotics system as a whole and affects the fault tolerance of a team of independent actors performing distributed activities. It is shown that the existing solutions to the problem of data exchange based on general-purpose IP networks have drawbacks; therefore, as the basis for organizing autonomous robot networks, we used developments in the domain of topological models of communication systems allowing us to build self-organizing computer networks. The requirements for the designed network for reliable message transfer between autonomous robots are listed, the option of organizing reliable message delivery using overlay networks, which expand the functionality of underlying networks, is selected. An overview of existing popular controlled and non-controlled overlay networks is given; their applicability for communication within a team of autonomous robots is evaluated. The features and specifics of data transfer in a team of autonomous robots are listed. The algorithms and architecture of the overlay self-organizing network were described by means of generally accepted methods of constructing decentralized networks with zero configurations. As a result of the work, general principles of operation of the designed network were proposed, the message structure for the delivery algorithm was described; two independent data streams were created, i.e. service and payload; an algorithm for sending messages between network nodes and an algorithm for collecting and synchronizing the global network status were developed. In order to increase the dependability and fault tolerance of the network, it is proposed to store the global network status at each node. The principles of operation of a distributed storage are described. For the purpose of notification on changes in the global status of the network, it is proposed to use an additional data stream for intra-network service messages. A flood routing algorithm was developed to reduce delays and speed up the synchronization of the global status of a network and consistency maintenance. It is proposed to provide network connectivity using the HELLO protocol to establish and maintain adjacency relations between network nodes. The paper provides examples of adding and removing network nodes, examines possible scalability problems of the developed overlay network and methods for solving them. It confirms the criteria and indicators for achieving the effect of self-organization of nodes in the network. The designed network is compared with existing alternatives. For the developed algorithms, examples of latency estimates in message delivery are given. The theoretical limitations of the overlay network in the presence of intentional and unintentional defects are indicated; an example of restoring the network after a failure is set forth.*

Keywords: *dependability; message delivery; guaranteed data delivery; overlay network; autonomous robot; group interaction; multi-agent robotic system.*

For citation: *Ermakov A.V., Suchkova L.I. Development of algorithms of self-organizing network for reliable data exchange between autonomous robots. Dependability. 2020;2: 35-42. https://doi.org/10.21683/1729-2646-2020-20-2-35-42*

Received on: *21.11.2020* / **Upon revision:** *14.04.2020* / **For printing:** *17.06.2020.*

Introduction

Successful task execution by a team of robots depends on the reliability of communications between its members, namely guaranteed delivery of messages to the actor and receipt of feedback.

The problem associated with improving the reliability of data transmission in a network with flickering nodes was considered by Lavrov D.N. in [1–2]. A flickering node is understood as an intermediate device able to transmit messages and is characterized by unstable operation or varied presence in the network, e.g., as the result of spatial movement of the node.

The concept of alternate availability nodes is applicable to major mobile objects, such as ships, airplanes, trains, robotics systems [3]. Know-how in the area of topology models of communication systems exists and can be applied to self-organizing computer networks. The key specificity of such algorithms consists in the impossibility to guarantee information transfer via the specified route due to the dynamic nature of the network and varying topology.

The matters related to the interaction within a team of mobile robots emerged in the 1980s; before that, research was focused on individual robotics systems or distributed systems not associated with robotics [4].

Jun Ota's research [5] confirms the existence of a class of problems that are optimally resolved through the use of swarm robotics, one of the fundamental tenets of which consists in the ability of parallel and independent execution of subtasks that reduces the total time of task performance. Any system that uses a set of interchangeable agents allows improving fault tolerance by simply replacing a failed robot, however, the creation of multifunctional agents is associated with high costs as compared to the creation of single-purpose agents. The distributed approach allows designing specialized robots that perform tasks that other agents struggle with [6].

As we know from the paper by Michael Krieger, Jean-Bernard Billeter and Laurent Keller [7], when tasks are partitioned among robots, reduced system efficiency may be observed. For instance, even if the total cost of a multiagent system proves to be lower than that of an integral solution, managing such system may be difficult due to the decentralized nature or absence of a global data storage. The absence of information on the intentions of other agents can cause a situation when individual robots will interfere with each other in terms of task performance. In order to avoid that, global connectivity is required that would ensure reliable data exchange between autonomous robots for the purpose of local and global planning and subsequent performance of local tasks by each agent.

Data exchange in constantly changing external conditions is a factor that directly affects the operational stability and efficiency of a team of robots. In this context, the development and research of reliable communication algorithms are relevant and serve to improve the functional dependability of a robotics system as a whole. In [8], experimental research of fault tolerance is cited that shows the importance of assuring reliable communication in respect to a team of robot.

Of special interest is the research of the algorithms of communication between autonomous robots, as the dependability and stability of their operation affect the decision-making time and coordinated activities of the team as a whole.

Currently, short-range communication is based on mesh networks that are distributed self-organizing networks with meshed topology deployed using Wi-Fi networks [9].

Higher-level protocols, such as TCP, guarantee reliable delivery of messages over such networks. However, due to the growing scope of communications on the Internet and requirement of fault-free operation of the network, adding new basic protocols and modifying their structure in order to provide new services became difficult [10]. Overlay networks allow extending a network's functionality without interfering with lower-level basic protocols [11] and can provide the following services: establishment of disruption-tolerant networks [12], rendezvous points [13], search [14–15]. It is difficult to provide such services at the IP level.

The commonly used overlay networks ensure reliable delivery of messages to a network node in various ways. Some networks specialize on anonymity (tor [16], I2P [17]) guaranteeing safe delivery, others rely on fast Wi-Fi network deployment (MANET [18], netsukuku [19]).

Overlay networks separate themselves from lower-level protocols. Thus, for instance, a network may use a different communication media in different segments of a heterogeneous network. The only requirement for the networks an overlay network operates with is the availability of an inter-subnet route. Figure 1 shows an example of overlay network constructed on top of an IP network.

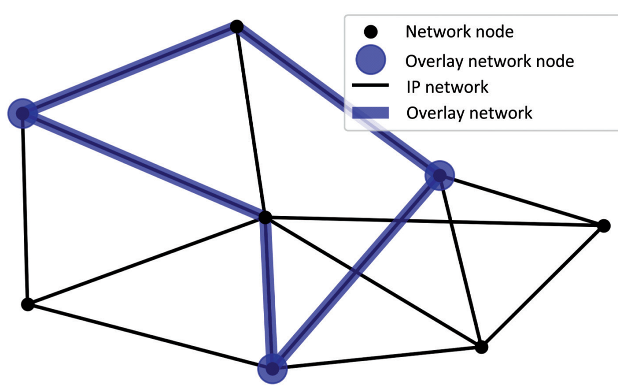


Fig. 1. An example of overlay network constructed on top of an IP network

The authors of [10] classify overlay networks into two major groups:

- controlled networks, where each node is aware of all nodes of the network and their capabilities;
- non-controlled networks, where none of the nodes is aware of the whole network topology.

Non-controlled networks are normally constructed using local area networks. Controlled overlay networks, on the contrary, are centralized or have one of the mechanisms of distributed storage of global network status, e.g. distributed hash tables (DHT) [20–22].

The tor network uses TCP flows for communication between network nodes and onion routing for passing messages within the network. It is not fully decentralized, as catalog servers exist that store information on the network status [16]. The tor network requires the availability of the Internet.

Other peer-to-peer or BPE-based networks do not have the function of passing messages to other nodes, and they require Internet in order to operate.

Thus, currently there are no network solutions for reliable communication within a team of autonomous robots.

In the context of the development of algorithms of reliable communication between autonomous robots, a self-organizing network must meet the following requirements:

- no need for manual node setting;
- a network client must be simple and easy to install (among other things, not require kernel patches or a specific version of the operating system);
- the network must operate at user level with no specific privileges;
- the network must operate on top of standard TCP and/or UDP protocols.

The existing overlay networks that were considered above do not meet the requirements, which urged the decision to develop an algorithm for reliable data exchange using an overlay network.

Structure of a self-organizing network

It is proposed to use an overlay network for the purpose of data exchange between autonomous robots. In such network, information is exchanged at the application layer based on the OSI model on top of the standard TCP and UDP protocols.

In order to connect to an overlay network, a client computing unit, e.g. the onboard computer of an autonomous robot, runs software that established connection with other network nodes and performs data transfer between intermediate nodes. Each node of a network at each moment in time maintains several connections while ensuring channel redundancy.

Structure of the message for delivery logic

A message is the minimal data unit transmitted within an overlay network. UDP packets are used for announcing changes within a network and low-priority notifications; they are received and parsed completely, which reduces the processing time. TCP requires a more complex processing algorithm. However, the fixed-size message header and

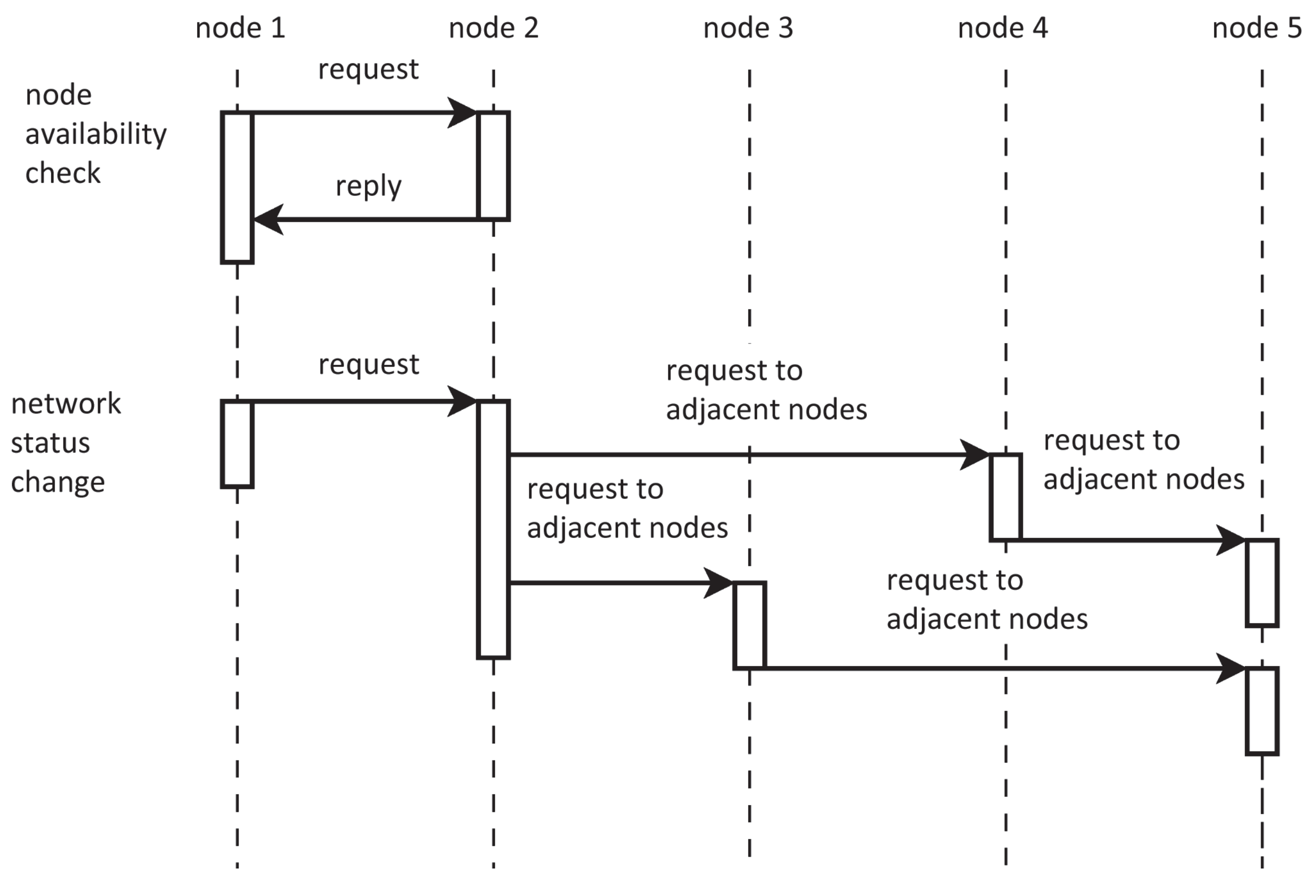


Fig. 2. Sequence diagram of received message processing by adjacent nodes with the example of availability check of an adjacent node and reporting of network status changes

presence of the data length field enable parsing the incoming TCP flow into individual messages.

A message is described with a formal structure: {IDsrc; IDdest; CMD; LEN; P}, where:

IDsrc is the source node identifier (8 bytes);

IDdest is the destination node identifier (8 bytes);

CMD is the type of the message (1 byte);

LEN is the length of the data field (unsigned integer, 2 byteы);

P is the protobuf2-coded data field of the length of LEN bytes [23].

Thus, a message consists of a header of the length of 19 bytes and variable-length data field.

A cell header {IDsrc; IDdest; CMD; LEN} contains a sender identifier and a recipient identifier. Node identifiers are 64-bit numbers comprised of a pair of IP addresses: {IPiface; IPext}, where:

IPiface is the network interface IP;

IPextis is the external IP address.

We believe such identification to be sufficient for our purposes. It allows nodes to generate unique identifiers for themselves, thus reducing the probability of collision to zero.

Depending on the type of CMD message, the message handler is selected. Cells are classified into two groups, i.e. the control and the transmit cells. Control cells are processed by destination nodes. That may include, for instance, cell availability check commands, network status change requests and responses (see Fig. 2). Transmit cells contain data that need to be processed if the recipient identifier matches the current node or transmit farther along the network.

The message processing procedure proposed by the authors has been implemented as software [24].

Algorithm of collection and synchronization of network status

A network intended for exchanging data between autonomous robots is controlled, i.e. it has a globally updateable status that contains information on all network nodes. The entire information on the network status is stored in each node. Thus, data redundancy increases the dependability and overall reliability of the network.

When a new node is added to the network, other LAN nodes are detected by means of a broadcast. In case of successful detection, communication is established with adjacent nodes and network status is synchronized. At this stage, message exchange occurs in all channels simultaneously (flooding) [1].

Flood routing is used for notification of changes in network status, for instance, when a new node is added. A network node sends the received packets to all of its direct neighbours, except the one, from which it was received. That approach improves the reliability of service information delivery and increases the probability of message reception by all the nodes of the network. The problem of message duplication is resolved through cashing of the received messages and inhibition of repeated message sending.

Figure 3 shows the block diagram of the system operation. A node receives a message from the network; depending on the value of CMD message type, flood routing algorithm is launched. The received packet is checked in the packet cache buffer in the random-access memory. If the packet is found in the cache, i.e. it was received earlier, the algorithm stops, discarding the packet and not processing it. Otherwise, the packet is added to the cache, replacing the oldest entries in the cache. Then, data field P from the packet is decoded and applied to the acquired global network status. Upon committing changes, the network status is rehashed. With that, local changes are complete; after that, adjacent nodes are notified by bulk messaging of the received message. An updated list of adjacent nodes is made, while the packet is immediately sent over UDP to nodes, from which a HELLO packet was recently received; for other network nodes, the generated packet is queued for subsequent asynchronous sending.

Besides flood routing, network status data consistency is maintained in all nodes through scheduled sending to adjacent nodes the hash of the list of known network node identifiers. In case of matching hash, adjacent nodes are synchronized. The capability to receive network status data from an adjacent node allows reducing the time of new node inclusion and not overwhelm the network with many complex messages [25].

Each node sends HELLO packets to the adjacent nodes, notifying them of its availability. Before network client shutdown it sends the respective notification over the network. In case of disconnection of or upon HELLO packet timeout a node makes several attempts to reconnect to the lost node. If the node proves to be unavailable, a message on the removal of the identifier from the network status is generated. Thus, network malfunctions are detected and prompt reaction is enabled [26].

Dependability is a complex physical property, therefore there is no single generalized criterion and indicator that would characterize dependability of technology in a sufficiently complete manner. Only a family of criteria allows evaluating the dependability of a complex technical system. The choice of criteria depends on the type of the technical item, its function and required completeness of dependability estimation [27].

One of the criteria of dependability of an overlay network is the latencies. It is assumed that an overlay network is to ensure reliable message delivery in cases of temporary unavailability of communication between adjacent nodes, including due to intentional or unintentional defects. Additionally, the specificity of application of the proposed network with autonomous robotics systems must be taken into consideration. In this context, priority is given to immediate message delivery, and a brief fault of delivery is preferable to a long delay (500 ms or more for some tasks). Another characteristic feature is the mutual independence of messages. In the proposed network the order of delivery is not important, which allows us optimizing the delivery algorithms and protocol subject to this criterion.

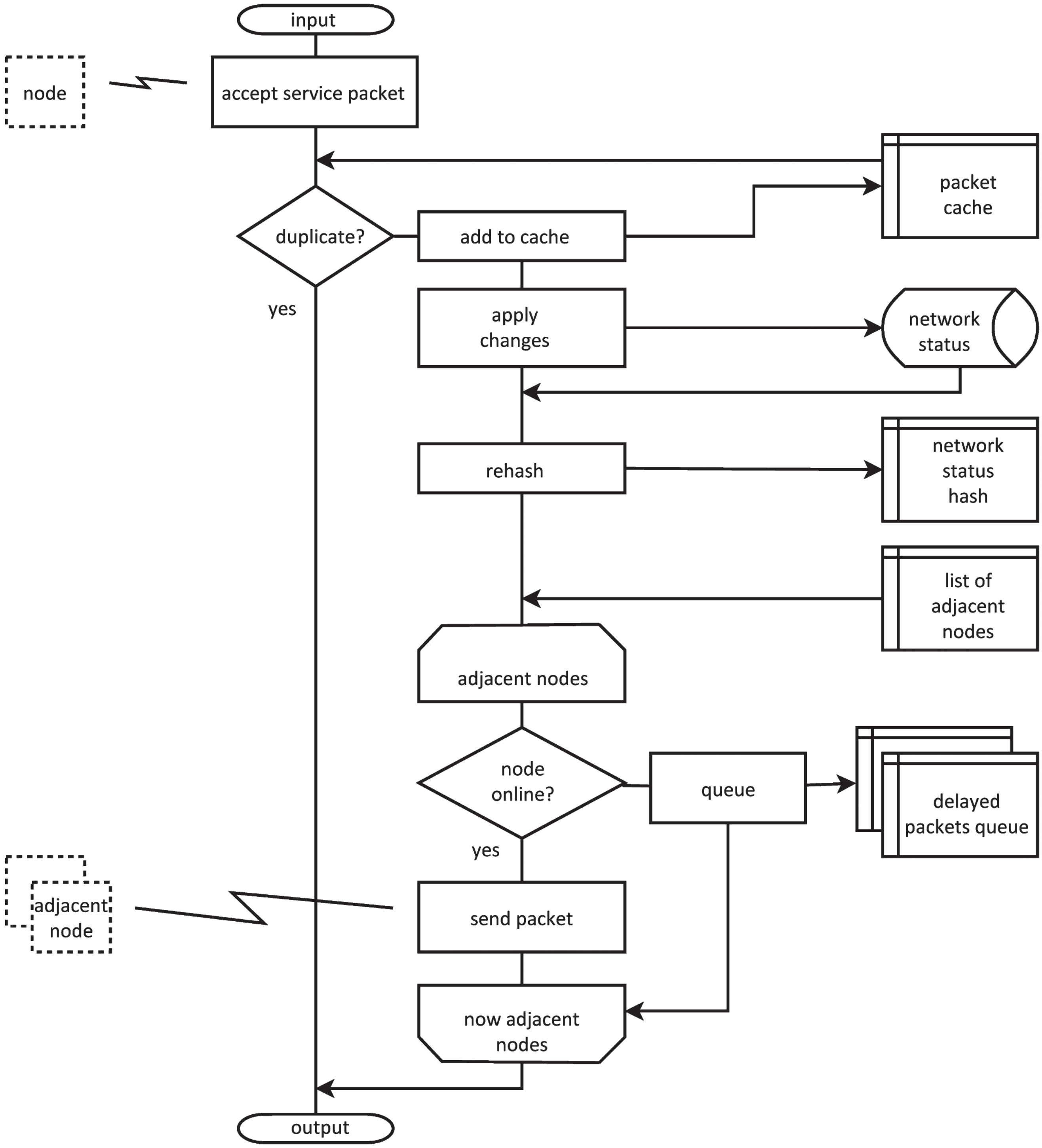


Fig. 3. System operation diagram for the flooding routing algorithm

Experimental testing of the developed data exchange algorithms

For the purpose of the experiment, a test network was created that consisted of a Cisco Catalyst 2960 router and six computers running under Ubuntu 18.04. In order to emulate several networks, five VLAN were configured on the PCs, with two computers in VLAN 0 and one in the others. The routing rules prohibited direct exchange of IP packets between all subnetworks except VLAN 0. As the result of the experiment network self-organization was confirmed; the operability of the developed algorithms was studied.

The existing TCP and UDP network protocols were experimentally tested using an active test network. For that purpose, data was sent between two routed nodes. Packet losses were emulated using the network interface of a node with an iptables rule and the statistic module that allows rule-based selection of a part of packets. For TCP, one connection was opened, within which overlay network cells were sent. UDP lacks a mechanism of delivery confirmation,

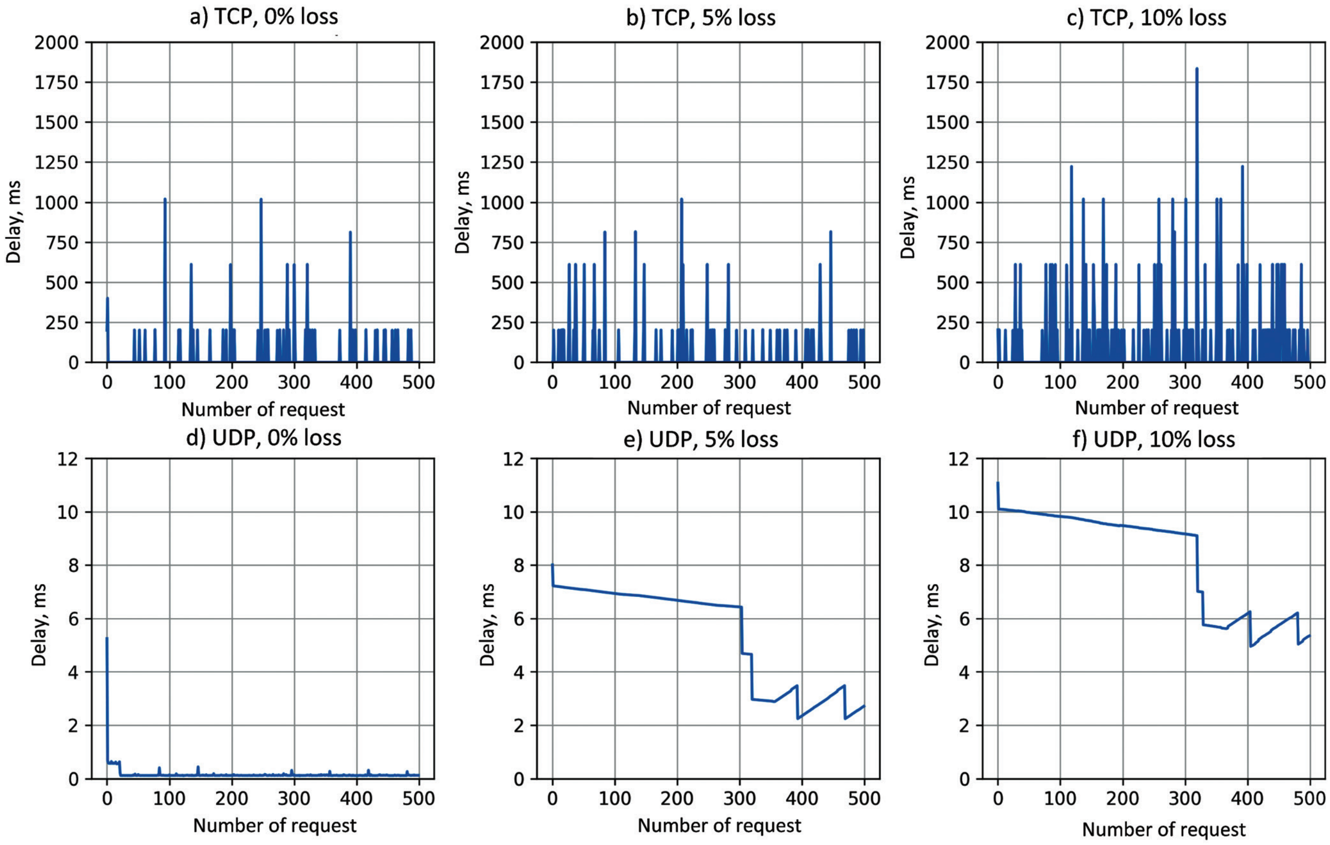


Fig. 4. Comparison of the delays of packet delivery by standard protocols subject to packet loss in the network: a) TCP, 0% loss; b) TCP, 5% loss; c) TCP, 10% loss; d) UDP, 0% loss; e) UDP, 5% loss; f) UDP, 10% loss

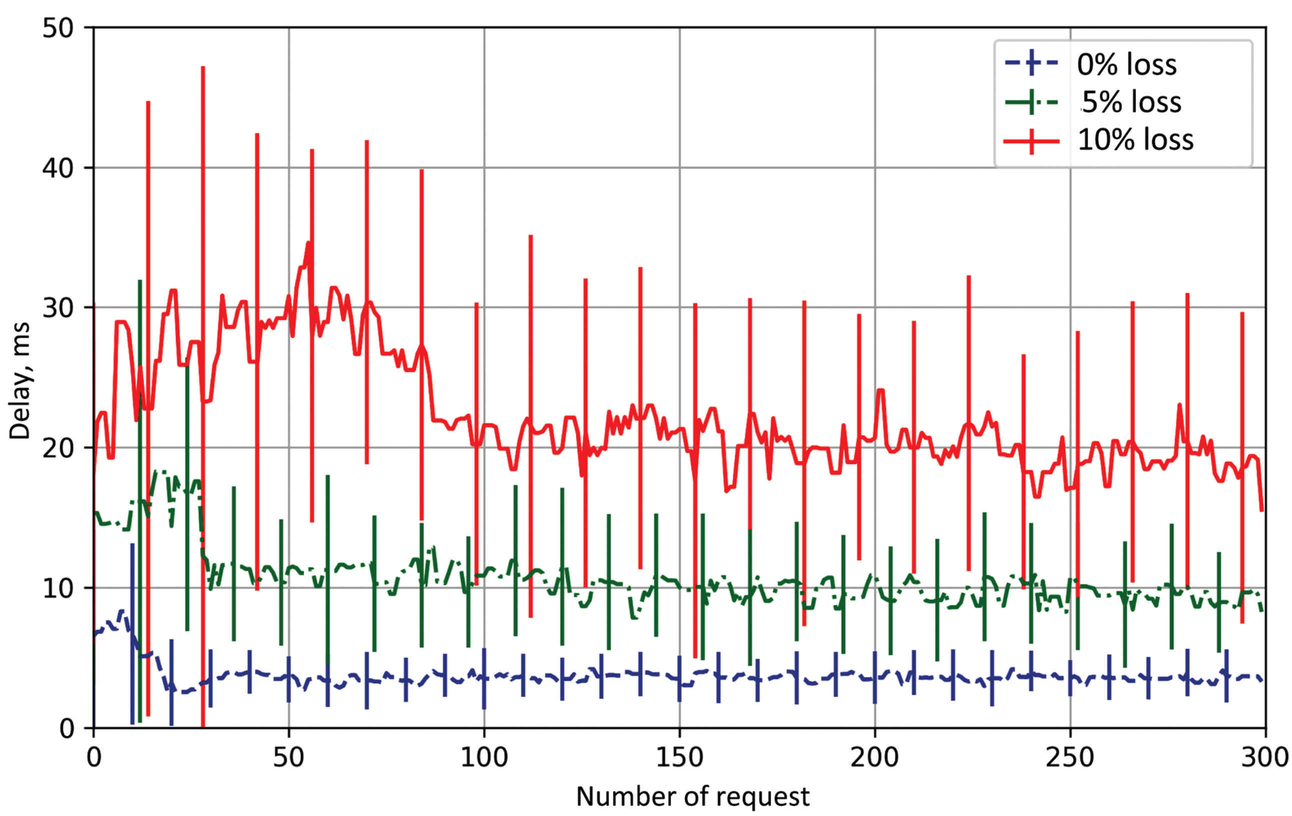


Fig. 5. Comparison of the delays of packet delivery within an overlay network subject to network defects

therefore the reception of each packet was confirmed by the receiving entity. If, upon time-out, no confirmation arrived, the packet was sent again.

Figure 4 shows the delays of delivery of each packet using standard TCP and UDP protocols under planned losses of 0%, 5% and 10% packets obtained during the experimental research.

Under no losses, UDP demonstrated minimal delays in data transfer, however, even under minimal losses the delay and amount of repeatedly sent data increase. Upon 300 to 400 sent packets, the delay settles (Figure 4, e and 4, f).

When TCP was used, establishing connection took long (up to a second in some cases) and reconnection was observed after disconnections. Such long single delays are unacceptable in networks used with autonomous robotics systems.

Knowing the results of research of delays associated with the use of existing data exchange protocols, we can estimate the dependability of the developed algorithms. Overlay networks were tested under the same conditions.

The proposed data exchange algorithm is characterized by shorter delays in normal operation and demonstrates higher dependability by enabling immediate delivery and minimization of failures that would occur because of untimely delivery of messages. The use of 0-RTT handshake (zero delay connection establishments) ensured the required performance of the overlay network.

The solutions' stability was verified over a month of operation with daily use of the network. No performance degradation or increasing delays were observed. The final results of the experiment are shown in Figure 5.

Conclusion

The authors developed operation algorithms of an overlay network subject to the particular conditions of its use by autonomous robots. The proposed approach will enable reliable data exchange within an autonomous system, thus ensuring the effect of collective mission performance with distribution of roles and subtasks, which would be impossible with no inter-agent communication and continuous data exchange.

Such algorithms are the foundation of a test software system intended for the research of the data exchange process within a team of autonomous robots.

References

1. Lavrov D.N. Principles of building a protocol for guaranteed message delivery. *Mathematical structures and modeling*. 2018;4(48):139-146. DOI: 10.25513/2222-8772.2018.4.139-146. (in Russ.)

2. Guss S.V., Lavrov D.N. Approaches to implementing a secure delivery network protocol for multipath data transfer. *Mathematical structures and modeling*. 2018;2(46):95-101. DOI: 10.25513/2222-8772.2018.2.95-101. (in Russ.)

3. Sorokin A.A., Dmitriev V.N. Description of communication systems with dynamic network topology by means of model "flickering" graph. *Vestnik of Astrakhan State Technical University. Series: management, computer science and informatics*. 2009;2:134-139. (in Russ.)

4. Parker L.E. Distributed Intelligence: Overview of the Field and its Application in Multi-Robot Systems. *AAAI Fall Symposium: Technical Report, FS-07-06*. 2008:5–14. DOI: 10.14198/JoPha.2008.2.1.02.

5. Ota J. Multi-agent robot systems as distributed autonomous systems. *Advanced engineering informatics*. 2006;20(1):59-70. DOI: 10.1016/j.aei.2005.06.002.

6. Arai T. et al. Advances in multi-robot systems. *IEEE Transactions on robotics and automation*. 2002;18(5):655–661.

7. Krieger M.J.B., Billeter J.B., Keller L. Ant-like task allocation and recruitment in cooperative robots. *Nature*. 2000;406(6799):992–995. DOI: 10.1038/35023164.

8. Winfield A.F.T., Nembrini J. Safety in numbers: Fault tolerance in robot swarms. *International Journal on Modelling Identification and Control*. 2006;1:30–37. DOI: 10.1504/IJMIC.2006.008645.

9. Bicket J. et al. Architecture and evaluation of an unplanned 802.11 b mesh network. Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM. ACM; 2005:31–42. DOI: 10.1145/1080829.1080833.

10. Srinivasan S. Design and use of managed overlay networks. Georgia Institute of Technology; 2007.

11. Clark D. et al. Overlay Networks and the Future of the Internet. *Communications and Strategies*. 2006;63:109.

12. Benson K.E. et al. Resilient overlays for IoT-based community infrastructure communications. 2016 IEEE First International Conference on Internet-of-Things Design and Implementation (IoTDI). 2016:152–163. DOI: 10.1109/IoTDI.2015.40.

13. Stoica I. et al. Internet indirection infrastructure. ACM SIGCOMM Computer Communication Review. ACM. 2002;3(4):73–86.

14. Ripeanu M. Peer-to-peer architecture case study: Gnutella network. Proceedings first international conference on peer-to-peer computing. IEEE. 2001:99–100. DOI: 10.1109/P2P.2001.990433.

15. Leibowitz N., Ripeanu M., Wierzbicki A. Deconstructing the kazaa network. Proceedings the Third IEEE Workshop on Internet Applications. WIAPP. 2003:112–120. DOI: 10.1109/WIAPP.2003.1210295.

16. Dingledine R., Mathewson N., Syverson P. Tor: The second-generation onion router. Naval Research Lab Washington DC; 2004.

17. Herrmann M., Grothoff C. Privacy-implications of performance-based peer selection by onion-routers: a real-world case study using I2P. International Symposium on Privacy Enhancing Technologies Symposium. Berlin, Heidelberg: Springer. 2011:155–174. DOI: 10.1007/978-3-642-22263-4_9.

18. Tandon N., Patel N. K. An Efficient Implementation of Multichannel Transceiver for Manet Multinet Environment. 10th International Conference on Computing, Communication and Networking Technolo-

gies (ICCCNT). IEEE. 2019:1–6. DOI: 10.1109/ICCC-NT45670.2019.8944505.

19. Kudriashova E.E., Vovchenko A.V., Oleynikov R.O. [A study of the NETSUKUKU fractal set-based network]. *Vestnik Mezhdunarodnoy akademii sistemnykh issledovaniy. Informatika, ekologiya, ekonomika*. 2008. 11(1):55-57. (in Russ.)

20. Kumar A. et al. Ulysses: a robust, low-diameter, low-latency peer-to-peer network. *European transactions on telecommunications*. 2004. 15(6):571–587. DOI: 10.1002/ett.1013.

21. Rowstron A., Druschel P. Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems. IFIP/ACM International Conference on Distributed Systems Platforms and Open Distributed Processing. Berlin, Heidelberg: Springer; 2001.

22. Stoica I. et al. Chord: A scalable peer-to-peer lookup service for internet applications. *ACM SIGCOMM Computer Communication Review*. 2001. 31(4):149–160.

23. Feng J., Li J. Google protocol buffers research and application in online game. IEEE conference anthology. 2013:1–4. DOI: 10.1109/ANTHOLOGY.2013.6784954.

24. Ermakova A.V., Suchkova L.I. [Implementation of the data communication protocol between smart autonomous robots]. Certificate of official registration of computer software no. 2019666759 of December 13, 2019.

25. Ermakov A.V., Suchkova L.I. Designing a network communication environment for the implementation of the management system in the autonomous robots team. *South-Siberian Scientific Bulletin*. 2019;2(4):28–31. DOI: 10.25699/SSSB.2019.28.48969.

26. Ermakov A., Suchkova L. Development of Data Exchange Technology for Autonomous Robots Using a Self-Organizing Overlay Network. 2019 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon). IEEE. 2019:1–5. DOI: 10.1109/FarEastCon.2019.8934727.

27. Polovko A.M., Gurov S.V. [Introduction into the dependability theory: study guide: second edition, updated and revised]. Saint Petersburg: BHV-Peterburg; 2006. ISBN 5-94157-541-6. (in Russ.)

About the authors

Alexander V. Ermakov, post-graduate student, IT and Information Security Department, Altai State Technical University, 656038, Altai Krai, Barnaul, 46 Lenina Ave., Administration of Information and Telecommunication Support, Information Systems Development Unit, k.a. of Ermakov A.V., e-mail: tour0@ya.ru.

Larisa I. Suchkova, Doctor of Engineering, Prorector for Learning and Teaching, Altai State Technical University, 656038, Altai Krai, Barnaul, 46 Lenina Ave., k.a. of Suchkova L.I., Prorector for Learning and Teaching, e-mail: li.suchkova@yandex.ru

Authors' contribution

Ermakov A.V., overview of literature, development of algorithms.

Suchkova L.I., formalization of requirements and problem definition.

Dependability, vol. 20 no. 2, 2020
Original article
https://doi.org/10.21683/1729-2646-2020-20-2-43-53

Application of machine learning methods for predicting hazardous failures of railway track assets

Igor B. Shubinsky[1], **Alexey M. Zamyshliaev**[1], **Olga B. Pronevich**[1], **Alexey N. Ignatov**[2], **Evgeny N. Platonov**[2]*
[1]*JSC NIIAS, Russian Federation, Moscow,* [2]*Moscow Aviation Institute, Russian Federation, Moscow*
**en.platonov@gmail.com*


Igor B. Shubinsky


Alexey M. Zamyshliaev


Olga B. Pronevich


Alexey N. Ignatov


Evgeny N. Platonov

Abstract. *The Aim of the paper is to reduce the number of hazardous events on railway tracks by developing a method of prediction of rare hazardous failures based on processing of large amounts of data on each kilometre of track obtained in real time from diagnostics systems. Hazardous failures are rare events; the set of variate values of the number of such events for an individual kilometre of track per year is: [0, 1]. However, for a railway network as a whole the yearly number of such events is in the dozens and efficient management requires the transition from the estimation of the probability of hazardous failure occurrence to the identification of the most probable location of failure.* **Methods.** *The problem of identification of rare, but hazardous possible events out of hundreds of thousands of records of non-critical railway track parameter divergences cannot be solved by conventional means of statistical processing. Hazardous events are predicted using the above statistics and artificial intelligence. Big Data and Data Science technology is used. Such technology includes methods of machine learning that enable item classification based on characteristics (features, predicates) and known cases of undesired event occurrence. The application of various algorithms of machine learning is demonstrated using the example of prediction of track superstructure failures using records collected between 2014 and 2019 on the Kuybyshevskaya Railway.* **Findings and conclusions.** *The result of facility ranking is the conclusion regarding the location of the most probable hazardous failure of railway track. That conclusion is based on the correspondence analysis between the actual characteristics of an item and conditions of its operation and the cases of adverse events and cases of their non-occurrence. The practical value of this paper consists in the fact that the proposed set of methods and means can be considered as an integral part of the track maintenance decision-making system. It can be easily adapted for online operation and integrated into the automated measurement system installed on a vehicle.*

Keywords: *machine learning, railway track facility failure, decision trees.*

For citation: *Shubinsky I.B., Zamyshliaev A.M., Pronevich O.B., Platonov E.N., Ignatov A.N. Application of machine learning methods for predicting hazardous failures of railway track assets. Dependability. 2020;2: 43-53. https://doi.org/10.21683/1729-2646-2020-20-2-43-53*

Received on: *29.02.2020* / **Upon revision:** *18.04.2020* / **For printing:** *17.06.2020.*

1. Introduction

The role of digital technology in the process management is on a steady rise. Automated management systems (AMS) enable much higher rate of business operations performance; autonomous control systems are deployed in trains and airplanes ensuring traffic safety at speeds beyond human reaction time. Today's diagnostics tools detect things the human eye is unable to capture and are used in healthcare, engineering, space exploration and other areas of science and industry. But the digital world is not limited to the automation of processes humans cannot perform, especially in case of major manufacturing facilities. Since 2016, JSC RZD has constructed an electronic document management system that connects over a thousand companies involved in freight transportation [1]. In the Lastochka EMUs diagnostic information is collected using 342 sensors and instruments. Together with the locomotive diagnostics systems, JSC RZD employs dozens of AMSs that provide the company with information on the condition of track [2, 3], power supply equipment [4], traffic safety systems [5], train graph [6] and a large number of other items and processes. Each of JSC RZD's AMSs is designed to solve individual problems, but in order to manage railway transportation in a holistic manner corporate-level systems were developed: EK ASU I (Single Corporate Automated Infrastructure Management System), EKP URRAN (Single Corporate Platform for Managing Resources, Risks and Dependability at Lifecycle Stages), EK ASU TR (Single Corporate Automated Workforce Management System), EK ASU FR (Single Corporate Automated Financial and Assets Management System). The existing data collection and storage systems, as well as the corporate systems that aggregate information from various sources, enable JSC RZD to successfully apply the Data Science technology (see. Fig. 1).

2. Relevance of track superstructure hazardous failure prediction

High train traffic and speed, environmental conditions, ageing cause tear and wear of railway infrastructure, primarily the track. Rail defects may cause derailments, accidents or crashes. Such hazardous events are associated with damage to the track, power supply systems, as well as cars and locomotive units with potential exclusion from the inventory rolling stock [7]. Derailed units of rolling stock may also intrude into the operational space of the adjacent track, which may cause a collision with an opposing train and, as the consequence, make damage catastrophic [8, 9]. A significant share of hazardous events attributed to the condition of track is typical not only to Russia's railways. Over the last decade, about one third of all railway incidents in the US were caused by track-related defects [10].

The analysis of derailments, accidents and crashes involving units of freight trains identified that such events caused by track malfunctions could occur on a kilometre of track rated, for instance, as "good". In this context, the aggregated estimate of a kilometre of track is not sufficient for predicting its condition, and it is required to take into consideration other parameters: number of widenings, realignments, etc. However, the collection of additional parameters alone will not suffice. According to [11], only a part of data on a controlled item is useful in terms of decision-making when managing specific events (see. Fig. 2).

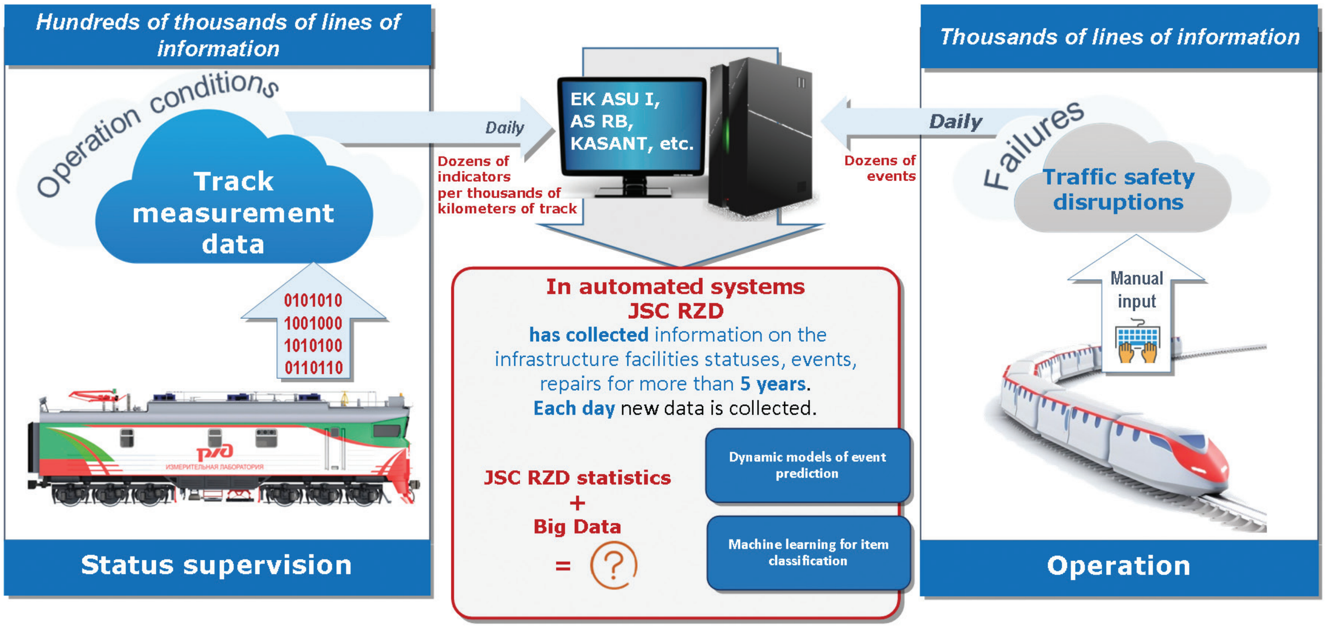


Fig. 1. JSC RZD AMSs as the foundation for Big Data application

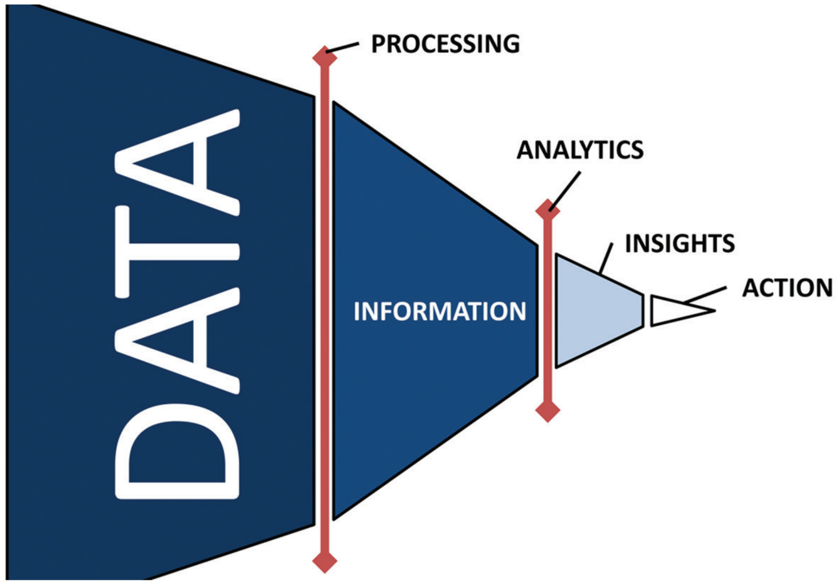


Fig. 2. Transformation of large volumes of raw data into actionable information

Modern methods of multiple factor data analysis and machine learning technology that allow including over 50 factors into models enable – based on existing knowledge of measured featured that characterize the condition of track – making conclusions regarding the need for urgent repairs in order to avoid track failures and derailments, accidents and crashes caused by an unsatisfactory condition of track. Conclusions regarding the efficiency of Big Data and Data Science can be made based on existing international practical experience, the analysis of which is set forth below.

3. Overview of the methods of machine learning and their application for the purpose of railway track defects analysis

With the growth of the amount of data collected by monitoring devices, such as wireless sensor networks or high-definition video cameras that are widely used for monitoring of critical railway infrastructure, machine learning also becomes increasingly popular in respect to improving the operational performance and dependability of railway systems [13].

Currently, due to the rapid technological advancements and widespread deployment of inexpensive sensors and wireless communications, the role of the Internet technology is increasing in the context of efficient implementation of maintenance strategies in a whole range of industries. In railway transportation, Data Science is also in active use [12]. Machine learning is increasingly popular as means of improving the dependability of railway systems. It also allows minimizing the daily cost of the maintenance [13].

Methods of machine learning can be subdivided into classical algorithms [14] and deep learning methods [15]. The main difference is the presentation level. The classical learning methods include the principal components method, support vectors method [16], solution trees [17], random forest [18, 19, 20], logistic regression [21] and nearest neighbours method [22].

In [23], the methodology of data classification for rail condition monitoring is presented. The authors put the emphasis on identifying the patterns of failure occurrence in sharp turns (horseshoe curves) using the principal components method and data obtained as the result of in situ inspections of the Swedish railway network.

In [24], the support vectors method is used for predicting a situation, when minor track defects deteriorate into major ones.

In [25], based on decision trees, a system is developed for preliminary automatic ranking of incidents that evaluates the probability of a pre-failure state based on the existing features.

Jiang and co-authors [26] proposed a hybrid approach to identifying contact fatigue based on ultrasound laser data.

In [10], the principal components method along with the support vectors method were applied to a set of data on 31 items collected on a US class I network for the purpose of detecting four types of surface defects.

As of late, the academic community has been making use of the advantages of the deep learning methods for studying rail defects. Researchers believe that deep learning may become an element of completely automatic railway monitoring systems [27].

Deep learning algorithms based on neural networks are employed as the primary tool for detecting structural defects in rails. The convolutional neural networks (CNN) are most widely used. That is due to the widespread use of video cameras that supply the research community with vast quantities of data and enable the application of more complex learning methods. However, CNN is a "black box" and practically cannot be interpreted. In other words, a researcher of machine learning cannot explain how a CNN model made its predictions or prove their reliability for the end user.

In [28], the CNN technology is used in examining the approaches to solving the problems of automated processing of images of track superstructure for the purpose of identifying the locations of potential defects. Images were used that had been collected by one of the trains of the Centre for Diagnostics and Monitoring of Infrastructure Facilities of the West Siberian Railway.

Lee and co-authors [29] used artificial neural networks and support vectors method for predicting the tear and wear of the ballast section based on such factors as the curvature, tonnage handled, etc. The authors however note that in order to obtain stable predictions, measurements must be taken over at least two years.

A more detailed overview of the application of various methods of machine learning in detecting track defects can be found in [30].

The diversity of the used models is evidence of the fact that the application of the machine learning technology currently represents a research process that includes the following stages:

- analysis of the sources of information on the track condition;
- data condition for machine learning;
- definition of machine learning objectives;
- training of models;
- selection of the best model;
- application of the model.

4. Algorithm of conditioning of railway track condition data as part of the JSC RZD machine learning application

Data received from JSC RZD AMSs are conditioned using an algorithm that includes 5 stages shown in Table 1.

Sample is one of the key concepts of machine learning. A sample is a finite set of cases (items, instances, events, test articles) and corresponding data (item characteristics) that form the description of the case. A sample that includes a full set of available data must include the target variable, i.e. an indicator, the prediction of whose value is the goal of machine learning. Additionally, a sample is subdivided into two parts: the learning sample and the test sample. The algorithm of conditioning of the data obtained from JSC RZD's AMSs for sampling as part of machine learning is shown in Fig. 3.

5. Algorithm of machine learning application for predicting hazardous failures of railway track

The problems of machine learning are normally described in terms of the ways a machine learning system is to process the learning sample. As the case of TSS learning sample, a kilometre of TSS was chosen, whose condition is characterized by 77 parameters, including the diagnostic results, operational conditions, qualitative estimates. The values of such parameters are represented in the form of vector $x \in R^n$, each element of which is the value of a feature.

Table 1. Stages of data conditioning

Name of stage	Aim	Conditions of stage performance	Relevance criterion of the stage
Data cleansing	Improvement of simulation through higher quality of data	Performed always	Performed always
Data conversion	Improvement of simulation through the capability to compare sequences with different physical units and/or value ranges	Performed if required for discrete sequences	1. Value variation ranges of various features differ more than 5 times. 2. Different physical units of features?
Data sampling	Extension of the scope of applicable models	Performed if required for continuous sequences	1. Target feature is a continuous value, but it is required to evaluate the probability of being within the range. 2. It is planned to employ a method that does not allow using continuous data.
Text cleansing	Improvement of simulation through higher quality of data	Performed if required for continuous sequences	It is planned to use information from the text in the simulation
Sampling	Quality verification of the developed models	Performed always	Performed always

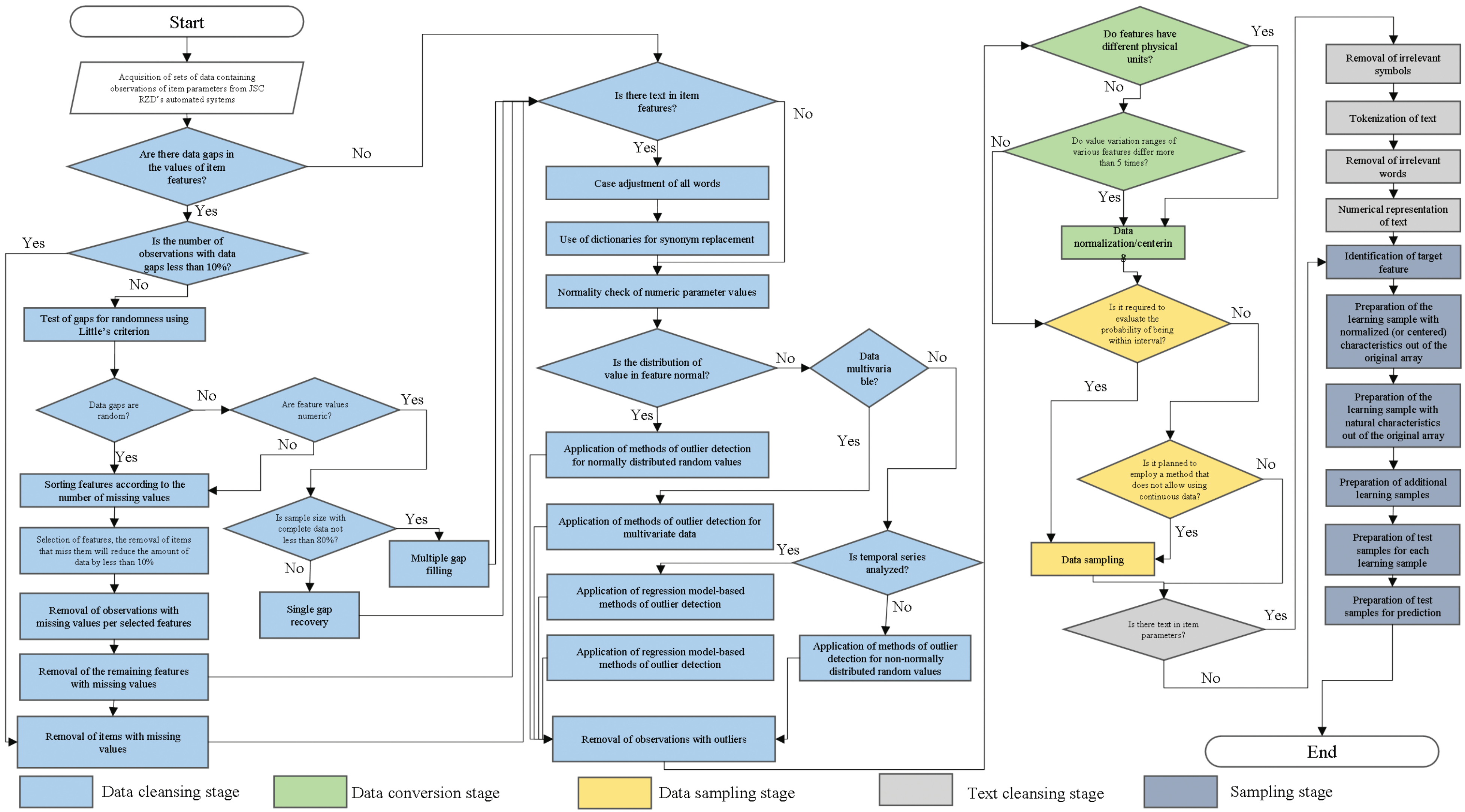


Fig. 3. Algorithm of AMS-generated data conditioning in sampling for machine learning

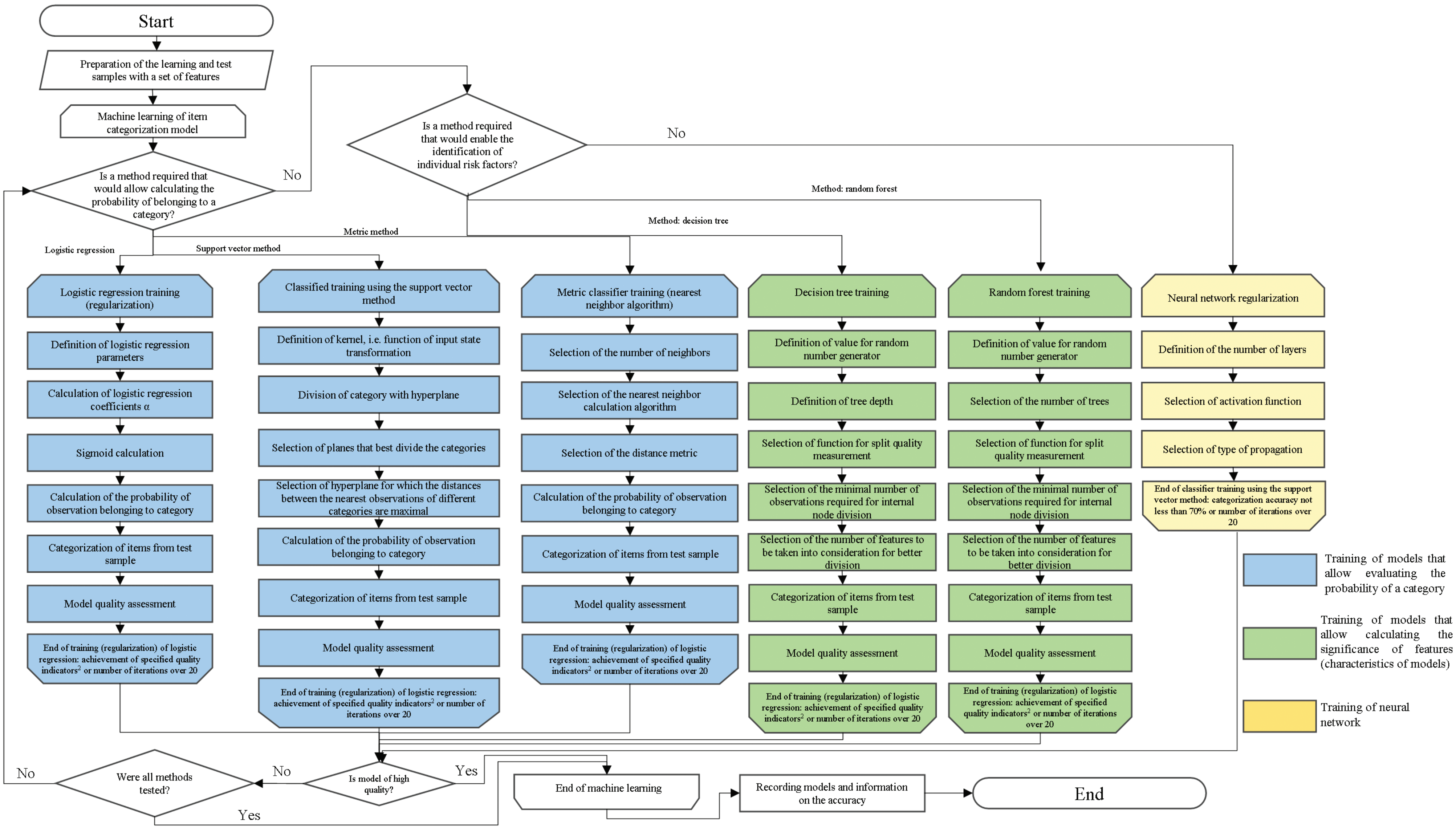


Fig. 4. Algorithm of machine learning methods application for classification of TSS condition

Classification, as the most common machine learning problem, consists in building models that serve to assign the examined item to one of the several known classes. With respect to that type of problems the classification algorithm is to answer the question as to which category the item belongs to. In terms of traffic safety (prevention of derailments, accidents and crashes) each item (kilometre of TSS) is divided into two classes: 0, a kilometre with no hazardous TSS failure; 1, a kilometre with a hazardous TSS failure.

From the learning sample we select the best parameters for the classification algorithm. On the test sample we calculate the classification error and in order to select the best algorithm.

Let X be an object space that is described by the set of features $X=\{X^1, ..., X^n\}^T$; $Y=\{0,1\}$ be the set of allowable responses; $y^*:X\rightarrow Y$ be the target dependence only known for the items of learning sample $Z^N=\left(x_i, y_i\right)_{i=1}^N$, where x_i is the vector of feature values, while $y_i=y^*(x_i)$ is the responses of the target variable, $i=1, ..., N$.

Let us denote $x=\{x_1, ..., x_n\}^T$, $y=\{y_1, ..., y_n\}^T$.

The learning problem consists in the requirement to re-establish the functional relationship between items and responses, i.e. to construct algorithm $a:X\rightarrow Y$ that approximates the target relationship y^* in the whole set X, not only the items of the learning sample Z^N.

Figure 4 shows the algorithm of application of six primary machine learning methods for kilometre of TSS classification.

6. Criteria of best model selection

A number of methods have been devised for the purpose of analysing the accuracy of the machine learning algorithm and comparing the accuracy of different algorithms.

For the purpose of problem binary classification, let us introduce the following designations:

TP, the number of correctly predicted category «1» items;

FN, the number of category «1» items with «0» prediction;

FP, the number of category «0» items with «1» prediction;

TN, the number of correctly predicted category «0» items.

Below are the primary measures of the quality of binary classification models.

1) General accuracy of the algorithm $AC=\frac{TP+TN}{TP+FP+FN+TN}$ that defines the overall efficiency of the classifier in terms of correct answers.

2) False alarm $FPR=\frac{FP}{FP+TN}$ that shows the efficiency of the classifier in terms of anomaly prediction.

3) Accuracy of the algorithm $PR=\frac{TP}{TP+FP}$ that shows the share of accurately predicted items identified as category «1».

4) Completeness of the algorithm $RE=\frac{TP}{TP+FN}$ that shows the share of items that are effectively category «1» and were predicted correctly.

5) *F*-measure of the algorithm, $F=\frac{2\cdot PR\cdot RE}{PR+RE}$ the harmonic average of accuracy and completeness.

6) Area under the curve of AUC errors, the global quality characteristic whose values are between 0 and 1. The value 0.5 corresponds to random guessing, while the value 1 implies correct recognition. AUC is the area under the ROC curve. The ROC curve shows the correlation between the share of false positive rate (FPR) and share of correct positive classifications (RE). The ROC curve is a sufficiently complex measure of algorithm accuracy; it is examined in more detail in [31].

7. Numerical experiment of line categorization based on failure prediction

Let us examine the problem of TSS failure classification. In order to prevent derailments, accidents and crashes, throughout the railway network, the condition

Table 2. Model quality indicators

Quality indicator	Logistic regression (sample 2)	Decision tree (sample 2)	Random forest (sample 2)	Support vectors method (sample 2)	Nearest neighbours method (sample 1)
1. AC	0.74	0.76	0.75	0.73	0.72
2. FPR	0.41	**0.28**	**0.28**	0.41	0.46
3. PR	0.78	0.94	0.94	0.94	0.89
4. RE	0.78	0.94	0.94	0.94	0.89
5. F-measure	0.78	0.86	0.86	0.94	0.88
6. AUC	0.68	0.83	0.83	0.76	0.71

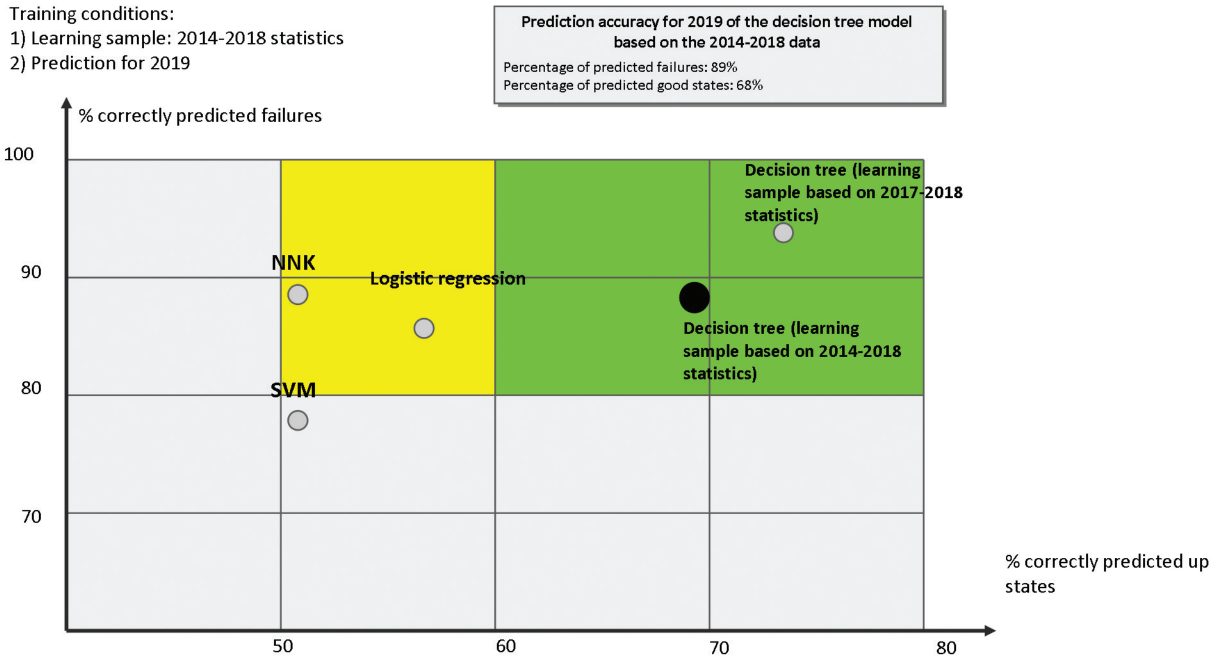


Fig. 5. Comparison of models in terms of quality

of track is checked for deviations from standard values using a geometry car. Based on the obtained data each kilometre of track is assigned a rating: "unsatisfactory", "satisfactory", "good" or "excellent" that is supposed to indicate the hazard of transportation incident caused by the condition of the track.

Between 2014 and 2019, TSS condition statistics were collected on the Kuybyshevskaya Railway. The following failures of railway infrastructure elements were registered: isolated joint, concrete tie, rail line as a whole, rail joint, geometrical parameters of the track, etc. Over a number of years, for each kilometre of track the following parameters were measured monthly: number of widenings, number of deviations, number of realignments, number of sags, traffic speed within the specific kilometre, etc.

If, within a kilometre of track, a failure is detected, the response is assigned the value of «1», otherwise, the value is «0», i.e. a set of category labels is of the form $Y=\{0,1\}$. It is required to solve the problem of binary classification based on the observations made in prior moments of time and verify the efficiency of the algorithm using the 2019 observations. Based on the performed classification, a hazardous failure is predicted.

194328 observations of various items (kilometres of track) were obtained. 267 items out of them were affected by hazardous failures. The data were subdivided

Table 3. List of test sample items within the zone of unacceptable risk

Date of check	Track maintenance department	Operational line	Track number	Kilometre	Probability of hazardous failure
29-JAN-19	9	2	1	979	0.55
29-JAN-19	9	1	1	969	0.51
14-JAN-19	9	2	1	979	0.48
14-JAN-19	9	1	1	969	0.48
29-JAN-19	9	2	1	1018	0.37
29-JAN-19	9	2	1	1003	0.28
14-JAN-19	9	2	1	1018	0.21
14-JAN-19	9	2	1	1003	0.17
23-JAN-19	20	1	1	36	0.003
25-JAN-19	20	2	1	36	0.0014

into the learning sample (192375 items, including 257 with hazardous failure, 2014 – 2018 data) and the test sample (1953 items, including 10 with hazardous failures, January 2019 data).

The classification problem was solved using several machine learning algorithms: logistic regression, solution tree-based algorithm, random forest method, methods of support vectors and nearest neighbours.

Learning samples were generated:

learning sample 1: 2014 – 2018 observations using standardized data;

learning sample 2: 2017 – 2018 observations using standardized data.

Additionally, data reduction was performed. The aim was to improve the quality of simulation through balanced learning samples, in which the number of observations with category «1» was at least 40% of the total number of observations.

Feature selection was done by means of recursive selection of the feature for each machine learning method.

Fig. 5 shows a comparison of the quality of models, Table 2 contains the indicators of model quality. The table shows models trained using the samples that demonstrated the best quality indicators for its type of model.

The results of model ranking: rank 1 is decision trees (trained using sample 6), rank 2 is random forest (trained using sample 6).

Table 3 shows a list of TSS elements with the highest probability of hazardous failure (corresponding to the highest levels of risk) in January 2019.

Upon an analysis of the data from Table 2 it can be concluded that the best possible results of item classification are ensured by using methods based on decision trees.

Shown in the last column of Table 3 are the values of frequency of trees classifying item category as “1”, i.e. the number of trees that identified an item as “kilometre with hazardous TSS failure”, in relation to the total number of constructed trees. Based on the results of the action of training sample classification algorithm, the threshold value of probability of failure is to be chosen depending on which classification error is the priority for us. The higher the threshold, the rarer the items will be classified as “kilometre with hazardous failure” (TP decreases, but TN grows). The lower the threshold, the lower will be the number of “kilometre with hazardous failure” items will be missed, but the higher the number of item with no hazardous failure (“0”) will be identified as having a hazardous failures (“1”) (TP and FP increase). In the context of TSS item classification, it is important not to miss an item with possible hazardous failure. Albeit at the cost of a larger number of items with no hazardous failure (“0”) that will be falsely identified as items with a hazardous failure (“1”).

Subject to the results of classification for the learning sample the threshold was chosen as $\overline{p}$=0,15. On the test sample that resulted in a situation, when out of 10 items with hazardous failures 8 were classified correctly and 5 items with no hazardous failure (marked “0”) were also classified as items with a hazardous failure. If the threshold was set at $\overline{p}$=0,10, the number of correctly identified items with a hazardous failure (“1”) would remain unchanged, while the number of incorrectly classified items with no hazardous failure (“0”) would have risen to 14. Under $\overline{p}$=0,001, all ten items with a hazardous failure (“1”) would have been classified correctly, but at the same time, the number of incorrectly identified items with no hazardous failure (“0”) would have risen to 251.

8. Conclusion

The paper presents the methodological foundations of prediction of rare hazardous events (failures) that can be used in the design of an automated system that performs real-time prediction of adverse events in railway transportation within a certain period of time by using and processing large amounts of information. The components of such system – mathematical models and methods, various metrics for model quality verification – should be defined subject to and based on the problem of prediction of railway track failures depending on various sets of factors. This problem was used in the process of optimization of the sequence of actions for taking the decision regarding the need for additional maintenance operations at any given railway line. For that purpose, models were compared using the proposed metrics. The ranking of facilities produced a conclusion regarding the presence of key indicators and their values of early warning of risk factors. That conclusion is based on the correspondence analysis between the actual characteristics of an item and conditions of its operation and the cases of adverse events and cases of their non-occurrence. The proposed set of methods and means can be easily integrated into an automated measurement system installed on a vehicle.

References

1. Sukonnikov G.V. [Application of the Internet of Things by JSC RZD]. www.rzd-expo.ru; 2017. Available at: URL: http://www.rzd-expo.ru/innovation/novosti/1.pdf. (in Russ.)

2. Bondarenko Yu.V., Kukso A.A., Markevich I.G. [Information technology in railway track diagnostics system management]. Proceedings of the IX international research and practice conference of students, post-graduate students and young scientists in 4 volumes. 2018;214-215. (in Russ.)

3. Nazarov D.G., Guda D.A. [On systems for automated track measurement]. Krasnodar: Kuban State Technological University: *Scientific works of the Kuban State Technological University*; 2019:135-146. (in Russ.)

4. Kurakina S.G., Shumakova E.G. [Automation of diagnostics and monitoring of railway catenary sections]. *Sovremennye innovatsii*. 2017;8(22):15-17. (in Russ.)

5. Zamyshliaev A.M. Premises of the creation of a digital traffic safety management system. *Dependability.* 2019;4(71):45-52.

6. Kuznetsova G.A., Krasheninnikov S.V., Kraysvitny V.P., etc. [Upgrading the GID Ural-VNIIZhT system]. *Avtomatika, sviaz, informatika*. 2016;11:15-19. (in Russ.)

7. Zamyshliaev A.M., Ignatov A.N., Kibzun A.I., Novozhilov E.O. Functional dependency between the number of wagons derailed due to wagon or track defects and the traffic factors. *Dependability.* 2018;18(1):53-60.

8. Liu X., Saat M., Barkan C. Analysis of causes of major train derailment and their effect on accident rates. *Transp. Res. Rec. J. Transp. Res. Board*. 2012;2289:154–163.

9. Zamyshliaev A.M., Ignatov A.N., Kibzun A.I., Novozhilov E.O. On traffic safety incidents caused by intrusion of derailed freight cars into the operational space of an adjacent track. *Dependability*. 2018;18(3):39–45.

10. Lasisi A., Attoh-Okine N. Principal components analysis and track quality index: a machine learning approach. *Transp. Res. Part C. Emerg. Technol;*91:230–248.

11. Zarembski A.M. Better railroading through Big Data. www.railwayage.com; 2018. Available at: https://www.railwayage.com/analytics/better-railroading-through-big-data.

12. Thaduri A., Galar D., Kumar U. Railway assets: a potential domain for big data analytics. *Proc. Comput. Sci*. 2015;53:457–467.

13. Li Q., Zhong Z., Liang Z. et al. Rail inspection meets big data: methods and trends. *18th International Conference on Network-Based Information Systems*. 2015:302–308.

14. Flach P. Machine Learning: the art and science of algorithms that make sense of data. Moscow: BMK Press; 2015.

15. Goodfellow I., Bengio Y., Courville A. Deep Learning. Moscow: DMK Press; 2018.

16. Widodo A., Yang B.S. Support vector machine in machine condition monitoring and fault diagnosis. *Mech. Syst. Signal Process*. 2007;21:2560–2574.

17. Sun W., Chen J., Li J. Decision tree and PCA-based fault diagnosis of rotating machinery. *Mech. Syst. Signal Process*. 2007;21:1300–1317.

18. Cerrada M., Zurita G., Cabrera D. et al. Fault diagnosis in spur gears based on genetic algorithm and random forest. *Mech. Syst. Signal Process*. 2016;70–71:87–103.

19. Santur Y., Karakose M., Akin E. Random forest based diagnosis approach for rail fault inspection in railways. *National Conference on Electrical, Electronics and Biomedical Engineering*. 2016:714–719.

20. Chistiakov S.P. [Random forests: an overview]. *Transactions of Karelian Research Centre of the Russian Academy of Sciences*. 2013;1:117-136.

21. Hosmer D., Lemeshov S., Sturdivant R.X. Applied Logistic Regression. New York: John Wiley & Sons; 2013.

22. Samworth R.J. Optimal weighted nearest neighbour classifiers. *Ann. Statist.* 2012;40(5):2733–2763.

23. Famurewa S.M., Zhang L., Asplund M. Maintenance analytics for railway infrastructure decision support. *Journal Qual. Maint. Eng*. 2017;23:310–325.

24. Hu C., Liu X. Modeling Track Geometry Degradation Using Support Vector Machine Technique. Joint Rail Conference; 2016.

25. Boyko P.Yu., Bikov E.M., Sokolov E.U., Yarotsky D.A. Application of Machine Learning to Incident Ranking at Moscow Railway. *Journal of Information Technologies and Computing Systems*. 2017;2:43-53. (in Russ.)

26. Jiang Y., Wang H., Tian G. et al. Fast classification for rail defect depths using a hybrid intelligent method. *Optik (Stuttg)*. 2019;180:455–468.

27. Gibert X., Patel V.M., Chellappa R. Deep multitask learning for railway track inspection. *IEEE Trans. Intell. Transp. Syst*. 2017;18:153–164.

28. Reznitskiy M.A., Arshinskiy L.V. Software implementation of the upper structure of the railway track defects detection automated system based on the technology of the convolutional neural networks. *Electronic Scientific Journal "Young Science of Siberia"*. 2018;1. (in Russ.)

29. Lee J.S., Hwang S.H., Choi I.Y. et al. Prediction of track deterioration using maintenance data and machine learning schemes. *J. Transp. Eng. Part A Syst*. 2018;144:04018045-1:9.

30. Nakhaee M.C., Hiemstra D., Stoelinga M. et al. The Recent Applications of Machine Learning in Rail Track Maintenance: A Survey. *Lecture Notes in Computer Science*. 2019;91–105.

31. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters. 2006;27:861–874.

About the authors

Igor B. Shubinsky, Doctor of Engineering, Professor, Deputy Director of Integrated Research and Development Unit, JSC NIIAS, Moscow, Russian Federation, phone: +7 (495) 786 68 57, e-mail: igor-shubinsky@yandex.ru

Alexey M. Zamyshliaev, Doctor of Engineering, Deputy Director General, JSC NIIAS, Moscow, Russian Federation, phone: +7 495 967 77 02, e-mail: A.Zamyshlaev@vniias.ru

Olga B. Pronevich, Head of Unit, JSC NIIAS, Moscow, Russian Federation, phone: +7 (985) 242 21 62, e-mail: oesune@rambler.ru

Alexey N. Ignatov, Candidate of Physics and Mathematics, Senior Lecturer, Moscow Aviation Institute, Moscow, Russian Federation, phone: +7 (906) 059 50 00, alexei.ignatov1@gmail.com

The authors' contribution

Shubinsky I.B. Definition of the requirements for the content of the algorithm of data conditioning for sampling, objectives of each stage. Definition of the requirements for classification of machine learning methods based on the capabilities of simulation data interpretation.

Zamyshliaev A.M. Aim definition, analysis of the problem and applicability of machine learning for prediction of hazardous failures of track superstructure, conclusions.

Pronevich O.B. Development of the superstructure condition data conditioning algorithm for the purpose of machine learning application, algorithm of machine learning application for predicting hazardous failures of track.

Ignatov A.N. Preprocessing and analysis of data for computation.

Platonov E.N. Overview of the methods of machine learning and their application for the purpose of railway track defects analysis. Classification problem definition.

Original article
Dependability, vol. 20 no. 2, 2020 https://doi.org/10.21683/1729-2646-2020-20-2-54-64

Towards safer rail control, command and signalling in the context of digitization

Alexey V. Ozerov, *JSC NIIAS, Russian Federation, Moscow*

Alexey V. Ozerov

Abstract. Aim. *The state of the art of railway computer-based control, command and signalling (CCS) systems is characterized by high requirements in terms of dependability, functional safety and cybersecurity under the conditions when digital transformation and challenges associated with the demand for increased competitiveness of railway transportation force the transition to new paradigms in engineering, testing, verification, validation and standardisation to facilitate and speed up the process of development and implementation. It is expected that while preserving the level of dependability and safety, at least, as it is, the industry has to enable the maximum possible introduction of innovative solutions and digital tools aimed at further automation of CCS systems to enhance the capacity and throughput of railways and the performance of systems, to minimize the impact of the human factor and reduce the number of failures and downtimes. In this context, the key factors are the interoperability (technical and operational compatibility) of systems and the technological independence of railway operators and infrastructure managers from the designer/supplier of railway automation systems, eliminating the vendor lock-in effect.* **Methods.** *The paper gives an overview of the state of the art of railway computer-based control, command and signalling using the example of the EU and provides an analysis of these systems in terms of dependability and safety in the context of migration to new grades of automation.* **Results.** *The author has considered the evolution of control, command and signalling systems in the EU using the example of the European Railway Traffic Management System (ERTMS). The analysis covered the general trends and approaches to engineering, testing, verification, validation and standardisation of railway CCS systems. The paper has overviewed the major EU research and design programmes of CCS development with the dependability and safety methodology taken into account. A special attention has been given to the methods of open engineering, remote lab testing and standardisation of ERTMS interfaces.* **Conclusions.** *In the context of digital transformation, the development of state-of-the-art railway computer-based CCS systems implies an accelerated introduction of a whole range of innovative solutions and a wide application of commercial off-the-shelf components (COTS), thus making systems more complex and being capable of affecting the dependability parameters. In order to maintain these parameters at a specified level and to minimize the impact of human factors, the railway community is increasingly using formal methods and automated means of engineering, diagnostics and monitoring at all stages of the system's lifecycle. A major factor of dependability is the standardisation of the system's architecture, interfaces, open source design and testing software, including the standardisation of approaches to remote lab testing of products by different manufacturers to prove the reliability of operation at the boundaries of systems of various manufacturers.*

Keywords: *CCS, railway signalling, train separation, dependability, safety, TSI, ERTMS/ETCS, GoA4, human factor, formal methods, verifiction, validation, certification, homologation, testing.*

For citation: *Ozerov A.V. Dependability of railway control, command and signalling in the context of digitization. Dependability. 2020;2: 54-64. https://doi.org/10.21683/1729-2646-2020-20-2-54-64*

Received on *18.02.2020* / **Revised on** *21.04.2020* / **For printing** *17.06.2020*

1. Introduction

The state of the art of railway computer-based control, command and signalling (CCS) systems is characterized by high requirements in terms of dependability, functional safety and cybersecurity under the conditions when digital transformation and challenges associated with the demand for increased competitiveness of railway transportation force the transition to new paradigms in engineering, testing, verification, validation and standardisation to facilitate and speed up the process of development and implementation. It is expected that while preserving the level of dependability and safety, at least, as it is, the industry has to enable the maximum possible introduction of innovative solutions and digital tools aimed at further automation of CCS systems to enhance the capacity and throughput of railways and the performance of systems, to minimize the impact of the human factor and reduce the number of failures and downtimes. In this context, the key factors are the interoperability (technical and operational compatibility) of systems and the technological independence of railway operators and infrastructure managers from the designer/supplier of railway automation systems, eliminating the vendor lock-in effect.

Strictly speaking, as regards railway CCS, digital transformation implies moving to a new paradigm of control and command of Industry 4.0. In terms of the basic principle of train separation, that means the evolution from simple separation of consecutive trains, first, in time, then in space (by fixed block sections) with further migration to radio-based control and command (such as in the European Railway Traffic Management System, ERTMS) and then to a dynamically changing headway between trains (including train convoys or virtual coupling, i.e. trains running closer than a safe breaking distance, like in road traffic). The transition implies a whole range of normative, regulatory, technological and technical changes [1].

One of the significant factors that underpin the need for a new methodology of engineering and maintenance of railway CCS systems is the increasing automation of train control with targets specified in the European programmes of research and innovation that aim to fully automate train operation, i.e. achieving driverless trains (so called GoA4, or Grade of Automation, according to IEC 62290) [2].

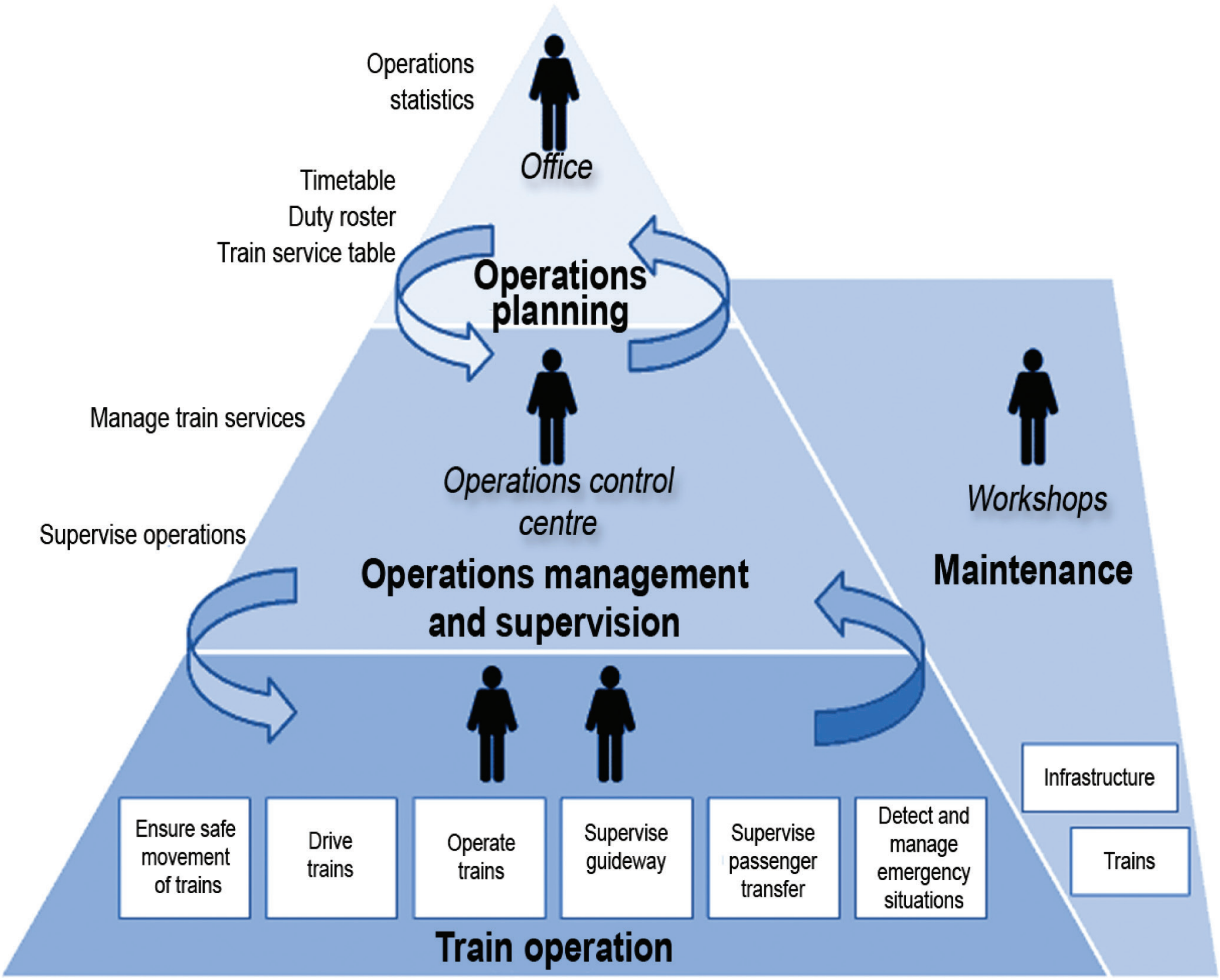


Fig. 1. Organization of railway operations

That emphasizes the importance of dependability and safety issues at all levels of railway operations management, where the human factor now still plays a significant role, especially at the level of safety-related (critical) systems (Fig. 1).

2. EU interoperability and dependability requirements and standards

Historically, practically each nation has its own railway normative requirements and operational rules, and often even a different railway gauge. For instance, before the EU was established, in Europe there were over twenty national CCS systems installed both trackside and onboard trains, as well as individual certification and homologation systems. After the establishment of the EU and opening of the Trans-European transport network corridors (TEN-T), the focus shifted to the issues related to interoperability (technical and operational compatibility) of railway systems and infrastructure and the provision of a common certification and homologation system (so called "cross acceptance system").

Later on, the EU approved the Interoperability Directives and Technical Specifications for Interoperability (TSI) for all components of the railway system including ERTMS that was developed by the European Railway Agency (ERA). In the directives, the interoperability is defined as the ability of a railway system to allow the safe and uninterrupted movement of trains which accomplish the required levels of performance [3].

The current version of TSI relating to Control, Command and Signalling (TSI CCS) is CCS 2016/919 [4]. It specifies the requirements for interoperability of ERTMS trackside and onboard assets, interfaces with external systems, as well as the parameters of reliability, availability, maintainability and safety (RAMS). The interoperability requirements are based on the body of functional requirements specifications for ERTMS subsystems and interfaces developed by the UNISIG group that combines the major European manufacturers of railway signalling equipment, under the aegis of ERA (so called "Subsets").

ERTMS has three core elements:

1. GSM-R (Global System for Mobiles – Railway) is the radio communication element based on the public GSM standard with specific railway frequencies and intended both for a voice communication between drivers and dispatchers and transmission of ETCS data (between the onboard train protection unit EVC – "European Vital Computer" – and the trackside control and command centre RBC – "Radio Block centre").

2. ETCS (European Train Control System) is the signalling system which is responsible for the control of speed, generation and execution of movement authorities, data exchange with interlockings of signals and points at stations.

3. ETML (European Traffic Management Layer) is the level of traffic management based on timetables and intended to optimize train speed profiles at routes using train running data in real time.

ERTMS/ETCS has three variants, or levels. Roughly speaking, Level 1 is the train protection using trackside signals and transponders (balises), with no GSM-R radio communication and, respectively, no RBC in place; Level 2 is the train control using GSM-R radio communication and, respectively, with RBC in place, as well as using balises as reference points along the route for the purpose of navigation (this being the system's variant most widely implemented both in Europe and elsewhere, with a rollout of over 100 ths. km. of railway lines); Level 3 foresees the additional application of onboard navigation and train integrity facilities and the implementation of moving block principle. So far, ERTMS/ETCS Level 3 is more of an experimental system being engineered and tested in the form of some hybrid solutions which integrate the application of satellite navigation, virtual balises and onboard digital route maps.

According to Subset-026 (System Requirements Specification), the ERTMS/ETCS reference architecture looks like as follows (Fig. 2) [5]:

The dash line in the diagram indicates the interfaces that are not yet standardised, and in this case the suppliers' proprietary (closed) protocols and solutions are used. This in particular applies to interfaces between RBC and interlocking installations (IXL) at stations and centralized traffic control (CTC), as well as the communication between radio block centres of different suppliers. This leads to both interoperability and RAMS-related issues.

Besides the list of mandatory functional specifications for subsystems and interfaces of ERTMS/ETCS, TSI CCS also contains a list of mandatory standards whose requirements shall be complied with for the certification of ERTMS/ETCS equipment, i.e.:

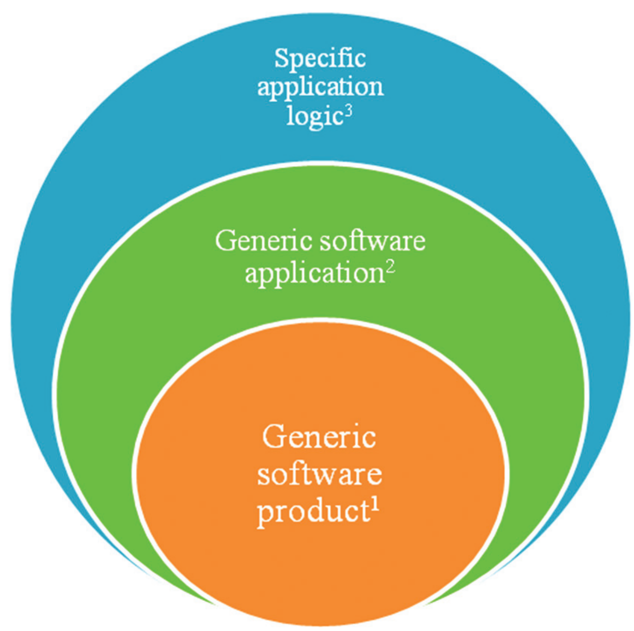


Fig. 3. Software layers

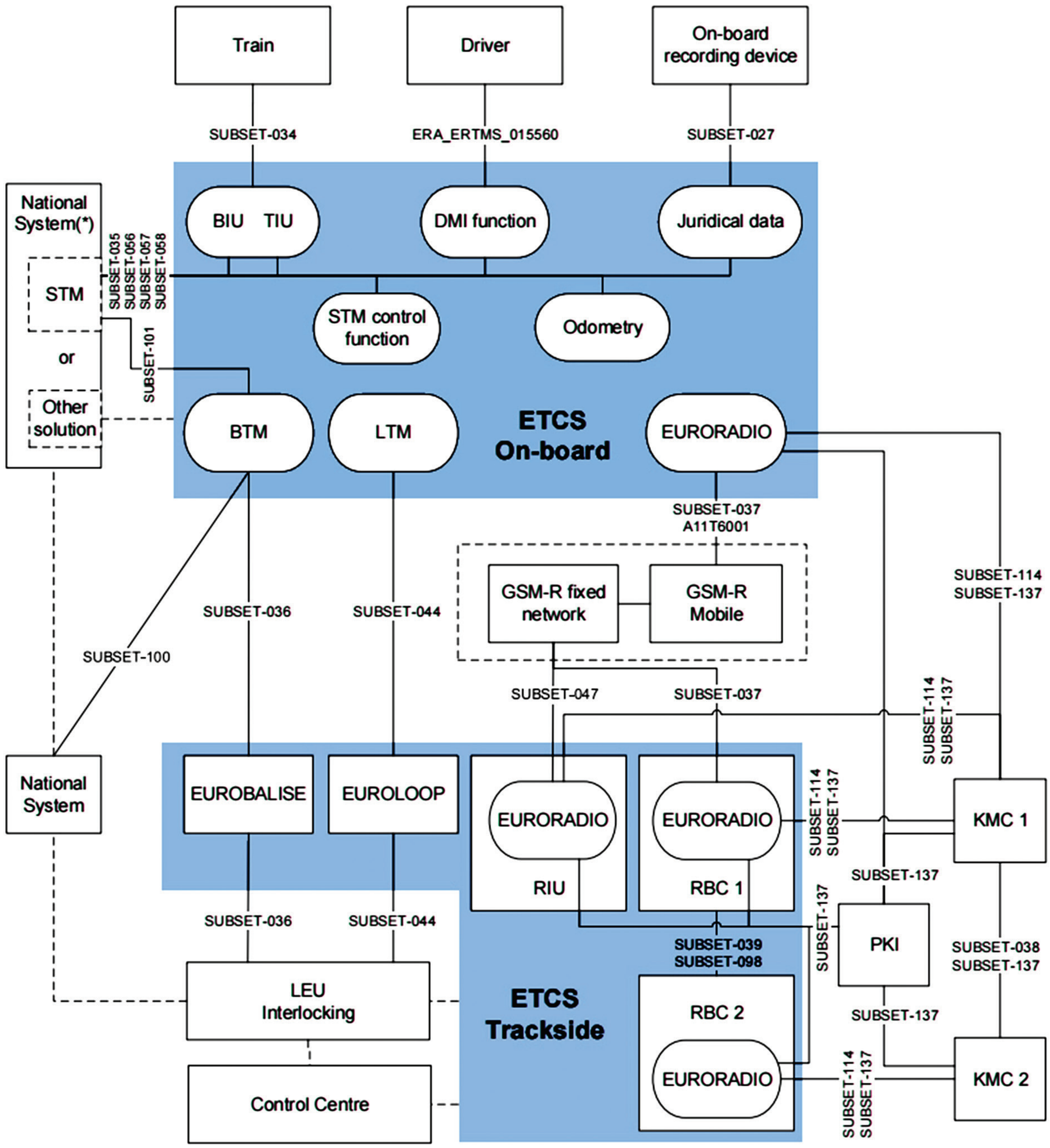


Fig. 2. ERTMS/ETCS reference architecture with functional interfaces specifications

1. EN 50126 Railway applications – The specification and demonstration of reliability, availability, maintainability and safety (RAMS).

2. EN 50128 Railway applications – Communication, signalling and processing systems – Software for railway control and protection systems.

3. EN 50129 Railway applications – Communication, signalling and processing systems – Safety related electronic systems for signalling.

4. EN 50159 Railway applications – Communication, signalling and processing systems.

As to CENELEC, in terms of software, ERTMS/ETCS engineering, verification & validation and certification are to be applied to three layers (Fig. 3):

If we take a look at the key element of ERTMS/ETCS Level 2, the RBC, then we can see that the first layer of RBC is its nucleus that contains a generic safety logic common for all railways where the product is implemented (the product

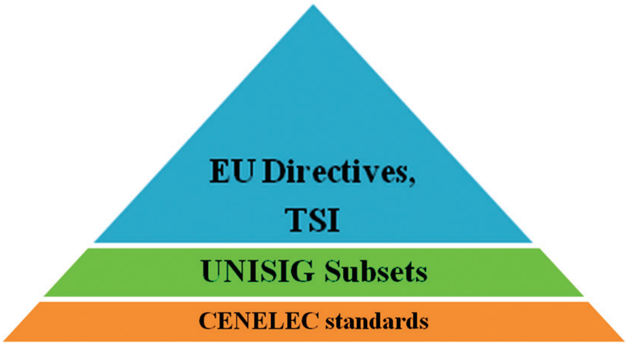


Fig. 4. ERTMS/ETCS regulatory pyramid

is certified once by a European notified body, provided that there are no further changes made to it); the second layer incorporates the signalling logic and rules of the country where the product is intended to be used, and is invariable for all applications of the product at the country's railway lines (requiring homologation for each country); the third layer is a project-specific signalling logic configured for a specific schematic plan and layout (requiring homologation for each project).

To summarize, the regulatory pyramid of ERTMS/ETCS can be presented in a schematic way as follows (Fig. 4):

3. ERTMS/ETCS dependability

The standards describing the RAMS methodology were developed as early as in the 1990s by the European Committee for Electrotechnical Standardisation (Comité Européen de Normalisation Électrotechnique, CENELEC). They apply an integrated approach to the management of RAM parameters directly related to the system dependability and safety (S) of a railway system based on risk assessment considering the lifecycle stages (V-model).

The standards are based on a probabilistic approach and provide quantitative parameters as well as recommendations for ensuring the specified RAMS by using well-proven methods (e.g. methods of programming, automated testing of software, detection and identification of errors and failures). Initially this approach was used in other manufacturing industries such as nuclear power engineering, aviation and space industry, from where it was adopted [6].

The certification of ERTMS/ETCS in compliance with CENELEC standards involves an extensive list of activities related to ensuring dependability and safety (RAMS), i.e. preparation and management of a large volume of documents at all stages of the system lifecycle as well as a strict observance of independence among the designer, the verifier/validator and the assessor of the system and the mandatory production quality management (manufacturing audit).

The RAM documentation includes a RAM programme and a RAM report (internal dependability calculation, check-lists of scheduled and unscheduled maintenance).

To preserve the dependability and operational parameters of the system during its lifetime, one shall define factors affecting RAMS, analyze and evaluate their consequences, use activities related to their control and prescribed by the standards.

According to EN 50126, the RAMS parameters of a railway system are influenced by three sources of failures:

– occurring within the system at any stage of the system lifecycle;

– adverse effects that affect the system in the course of operation;

– errors that affect the system during maintenance activities.

And all these three sources of failures can interact. The efficient management of these factors can keep RAMS as specified. In a schematic way, the relationship of the factors influencing dependability and safety is presented in Fig. 5 [7]:

The performance requirements of a railway CCS system are specific for each system and are thus specified in the agreement between the manufacturer and the infrastructure manager during the design phase. For a system as a whole, there are three defined types of failures:

– immobilizing failure (at least two trains have to be put in on-sight mode);

– service failure (one train at most has to be put in on-sight mode);

– minor failure (which requires unscheduled maintenance, though it doesn't fall under the previous categories).

For example, ERTMS/ETCS RAMS requirements specification (1998) provides the following specific parameters [8]:

– the probability of a train delay due to signalling failures shall not exceed 0.018, while the probability of a train delay due to ERTMS/ETCS failures shall not exceed 0.0027;

– the allowed average delay per train due to ERTMS/ETCS failures, at the end of an average trip of duration of 90 min., shall be not greater than 10 min.;

– the operational availability of ERTMS/ETCS due to all the causes of failure shall be not less than 0.99973;

– immobilizing failures shall not exceed the 10% of the total amount of failures which affect the system's operational availability; service failures shall not exceed 90% of the total amount of failures which affect the system's operational availability;

– the mean time to restore of trackside distributed equipment is 1.737 hours.

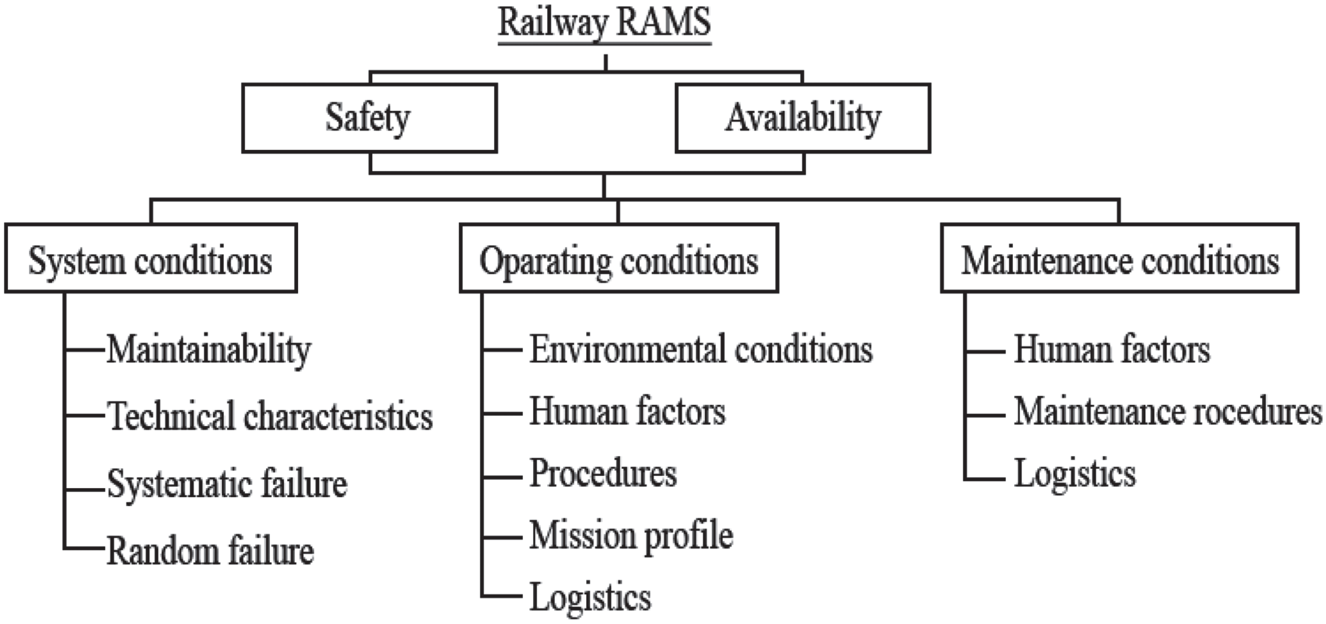


Fig. 5. Factors influencing RAMS (adapted from EN 50126)

However, it is worth noting that ERTMS/ETCS Level 2 is generally an overlay system, i.e. it is installed over a national signalling system and uses it as a kind of fallback in case of failure. There is an ongoing debate in literature about the need of redundancy in the form of external systems to increase the dependability of a primary signalling system [9].

The ERTMS/ETCS RAM programme shall include, as a minimum, the following activities:

– RAM programme planning;
– System conditions and mission profile;
– Periodical RAM programme reviews;
– Reliability modelling, prediction and apportionment;
– FMECA analysis;
– Software reliability analysis;
– Service dependability analysis and verification;
– Preventive maintenance analysis;
– Corrective maintenance analysis;
– Fault isolation and trouble-shooting plans;
– Reliability development/growth testing programme;
– Maintainability preliminary tests;
– Reliability demonstration tests;
– Maintainability demonstration tests;
– Failure data collection from the field (FRACAS).

Naturally, the human factor greatly affects RAMS as well – both at the design stage and in the course of operation. Since humans can considerably affect RAMS, the human factor should be taken into account to a greater extent than in other industries, when achieving the specified RAMS parameters of a railway system. This motivates all the efforts made by the railway community in terms of automation of operation and maintenance as well as of engineering, testing, verification and validation, particularly in the context of a global trend for digitization and the implementation of Industry 4.0 principles.

4. New approaches and requirements

The analysis of the policy papers of the EU railway bodies and associations and those of the International Union of Railways (UIC) shows that one of the key drivers of the search for new approaches and solutions in the railway sector in the context of digital transformation is the low rate of innovations introduction due to a long period of certification and homologation, that is largely driven by the dominance of proprietary solutions in the absence of standardised protocols and interfaces as well as standardised methods of automated engineering. This leads to high costs of development and implementation, operations and maintenance, growing obsolescence of railway systems and vendor lock-in. Also, it potentially impacts their dependability and safety.

In order to find a way out, in 2014 the EU established a joint undertaking Shift2Rail with a total budget of about 900 million Euros [10]. This is an industry-scale innovation programme of railway transportation development that brings together railway manufacturers, operators and infrastructure managers. Its key objectives are the development, integration, demonstration and validation of innovative digital technologies for the railway transport intended to enhance its attractiveness for users.

Shift2Rail is expected to contribute to:

– reducing the lifecycle cost of railway transportation by as much as 50%;

– doubling the current railway capacity;

– increasing the reliability and punctuality of the railway transportation by as much as 50%.

Basically, the changes of approaches to the RAMS specification and demonstration and further on to certification are driven by the business requirements and considerations related to the need to reduce the costs for engineering, certi-

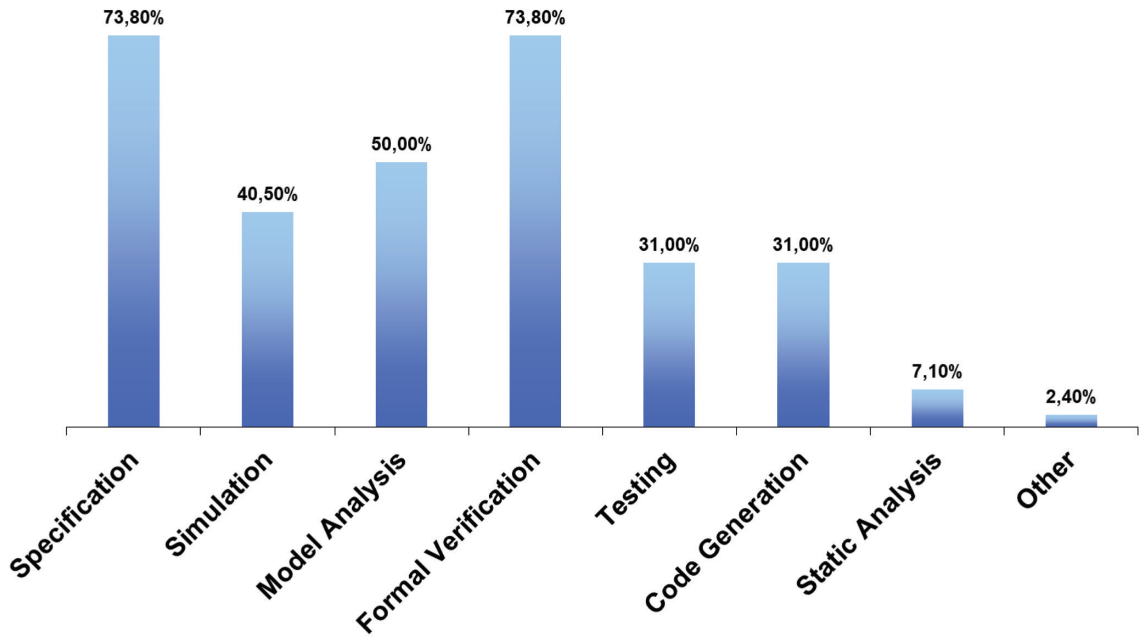


Fig. 6. The application of formal methods at system lifecycle stages

fication and homologation of products and the time to market and on-site installation. Not surprisingly that the Shift2Rail projects research various methods of automation of development, verification and validation, testing processes, including those that are used in other industries – first of all, in aviation and automotive engineering.

Based on selected and then standardised methods, the transition is supposed to be towards virtual certification. By virtual certification one means the maximum allowable use of evidence from virtual testing and simulation based on formal models to support the certification and homologation process [11]. For instance, this methodology is studied

Table 1. The list of the EU projects related to the use of formal methods in railway command, control and signalling

Project	ERTMS/ETCS/CBTC
CRYSTAL	http://www.crystal-artemis.eu/
Deploy	http://www.deploy-project.eu/
DITTO	http://cs.swansea.ac.uk/dittorailway/
EuRailCheck	https://es.fbk.eu/projects/eurailcheck-era-formalization-and-validation-etcs
MBAT	http://www.mbat-artemis.eu/home/69-abstract.html
OpenCOSS	http://www.opencoss-project.eu
OpenETCS	http://openetcs.org/
PERFECT	https://trimis.ec.europa.eu/project/performing-enhanced-rail-formal-engineering-constraints-traceability
	Distributed railway signalling
SafeCap	http://gow.epsrc.ac.uk/NGBOViewGrant.aspx?GrantRef=EP/I010807/1
	Interlocking
ADVANCE	http://www.advance-ict.eu/
EULYNX	https://eulynx.eu/
EuroInterlocking	http://test.swissrequirementsengineering.ch/en/projects/euro-interlocking-project
INESS	http://www.iness.eu
RobustRail	http://www.robustrails.man.dtu.dk

within the framework of the Shift2Rail project – PLASA 2. The objective is to substantially reduce the time required for provision of interface with the existing systems in place and the field testing by standardising interfaces and using formal methods for engineering, verification and remote lab testing.

In fact, it is worth noting that EN 50128 highly recommends the use of semiformal and formal methods for development and automated tools of testing, verification and validation, however there is still much to be done in terms of selection and standardisation of respective methods and tools [12].

According to [13], the approach to ensuring the "development quality" of software presented by the CENELEC standard alone cannot guarantee the correct operation of a computer-based system. It is to increase the "development quality" and to reduce the lifecycle cost of safety-critical computer-based systems, interlocking systems in the first place, why formal methods were introduced. The basic advantage of the methods is that they enable an exhaustive analysis of all possible scenarios of the programmed system behavior while ensuring the consistency between the formalized and proven behavior of the model and the behavior of the code embedded into the system.

5. History and further application of formal methods

The history of the use of formal methods in standardisation of railway signalling started in 1997 when the UIC published the European Railway Research Institute (ERRI) project report that presented a detailed analysis of functional conditions of interlocking systems and proposed the harmonization of functional requirements for signalling systems based on formal methods. Later on, a UIC working group developed a semiformal method called EURIS (European Railway Interlocking Specification), which defined building blocks (e.g. signal, track, point) and described the operations related to each building block using flowcharts. The UIC project EURO-INTERLOCKING (1998-2008) formalized the requirements for an interlocking system that were converted into a formal model visualized by a computer. It appeared that both the skills of a signal engineer and a modelling specialist were needed to do this work. Additionally, it became apparent that that is an iterative process requiring further quality improvements both in the verbal language representation and the requirements coverage [14].

This work was continued within the framework of the EULYNX project where using the SysML models the focus was on the formalized description of interfaces of trackside signalling systems of different supplies, including ERTMS/ETCS subsystems, to reduce the time of their development and software/hardware adaptation. As an extension of these approaches, the ERTMS Users Group and the EULYNX consortium then initiated the Reference CCS Architecture (RCA) project aimed at developing a new ETCS reference architecture integrating ATO functionality (and further migration to GoA4), harmonization of components and standardisation of interfaces and communication protocols based on the use of formal methods. In 2019, an alpha release of a future reference architecture was issued [15].

In parallel with RCA, the initiative of the railway infrastructure managers from the major European countries (Germany, France, Switzerland, the Netherlands, etc.) gave birth to the Open CCS Onboard Reference Architecture (OCORA) consortium with the objective to develop and standardise a next-generation open modular ETCS onboard architecture platform. The OCORA initiative plans to use the EULYNX and RCA approaches and is also focused on the requirements of an updated CCS TSI version to be released in 2022. OCORA strives to negate the vendor lock-in effect (by modularity, interoperability, replaceability, modifiability, security and usability) through the development of a new open CCS communications bus and standardisation of communications protocols of all onboard modules using accepted industry standards as much as possible. It is assumed that such approach will also allow achieving the tangible enhancement of the system performance, as a summary of reliability, availability, maintainability and safety, plus cyber security. According to the project master document, the OCORA deliverables are expected to be a comprehensive and coherent set of specifications as well as new supporting recommendations for integration, verification and validation of CCS onboard implementations with the maximum use of automated testing tools and formal methods [16].

Within the framework of the Shift2Rail-backed ASTRail project, the researchers from the Formal Methods and Tools (FMT) laboratory, which is part of the Institute of Information Science and Technologies (ISTI), one of the institutes of the Italian National Research Council, made an analysis and assessment of major languages and tools for formal simulation and verification used in the railway domain. For example, the research identified [17] that the following automated engineering tools most frequently appear in literature: Simulink, NuSMV, Atelier B, Prover, ProB, SCADE, IBM Rational Software Architect, Polyspace, S3.

The surveys made within the framework of the project revealed that developers use the above or other automated tools for the following purposes (Fig. 6):

The results of the survey showed that formal methods are typically used at the stages of the system specification and verification. The standardisation of approaches to the composition of functional and system requirements specifications (FRS, SRS) as well as to verification based on formal methods is covered by a number of the EU projects, let alone the Shift2Rail programme itself. Thus, starting from 1998 till now, 14 projects have addressed the use of formal

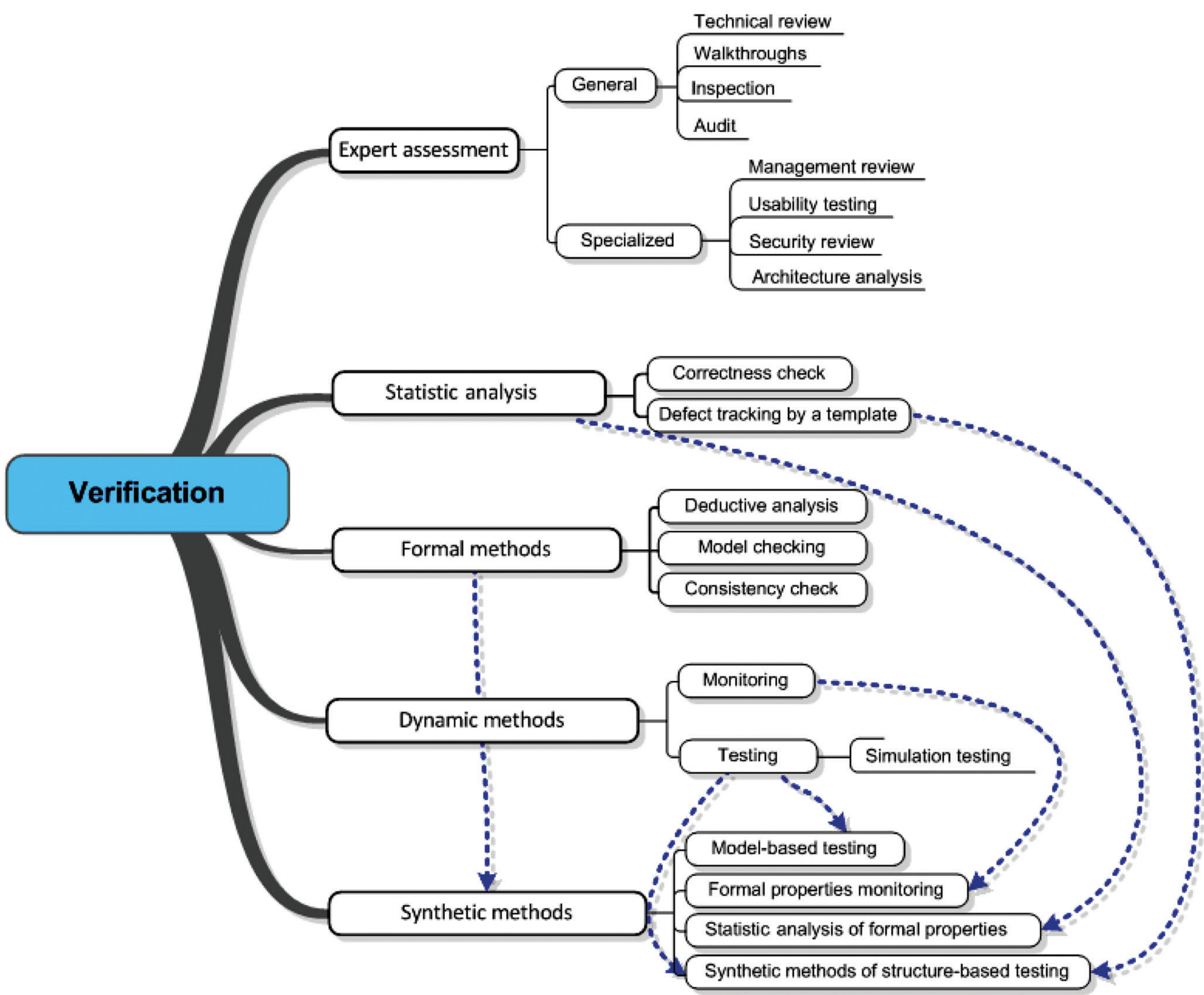


Fig. 7. Combined application of verification methods

methods in railway applications. A detailed description of them is obviously out of scope of the paper, so let us just list them (Table 1):

A detailed analysis of the conventional and formal methods of verification is given in [18]. Generally, the verification process of a safety-related system such as ERTMS/ETCS is made of a set of complementary methods and tools, that at present frequently takes into account not only RAMS parameters but also cyber security (one more area subject to further standardisation in the railway domain, Fig. 7):

A detailed analysis of the capabilities of discrete event simulation as applied to the lifecycle stages of ERTMS/ETCS, in particular to the verification phase, is given in [19]. The author notes that the ERTMS/ETCS system can be characterized by the fact that the system states are discrete, and the transition mechanism of states is driven by events. For safety-critical systems, current engineering methods cannot guarantee that the developed system will respect all its requirements and behave safely, and that shows an urgent demand to integrate verification processes into the system engineering as early as possible. This can be done by using formal languages and formal methods of engineering.

There is a long list of formal methods, but they share certain advantages:

– formal representations have precise semantics that is free from ambiguity;

– formal models can be mathematically verified and thus proven to be correct;

– formal models can be read by computers, and so enabling the automation of the engineering process.

Ideally, the application of formal methods allows avoiding the unsafe transitions of the system states as well as minimizing the number of errors introduced into the system by a designer, and therefore, the number of system failures, which directly affects its dependability.

One of the key sections of the European Shift2Rail railway initiative is its innovation programme IP2, whose objectives include the development of automated tools for simulation and lab testing (remote as well) to reduce the need of integration and validation tests on site (so called “Zero on-site testing”).

In the opinion of the Shift2Rail authors [20], today, as regards the testing of CCS system, the situation can be characterized as follows:

– In most cases, suppliers do product testing in the lab.

– System testing is still done with a large amount of on-site testing.

– On-site testing is often used as a fallback, if lab testing has not been finished in time.

– Lab testing is done mainly by a supplier-specific process and testing environment.

– Collaboration with different suppliers always causes the need for sophisticated adaptors with less chance to reuse them in subsequent projects while increasing costs.

– The test case derivation is not comparable since different approaches have been applied, which are proprietary.

Eventually, in terms of the targeted goals of the Shift2Rail programme and its research and innovation projects, the approaches at all the stages of the ERTMS/ETCS lifecycle are expected to be standardised taking into account the necessity of implementing innovative ideas such as moving block, virtual coupling, perception capabilities as part of GoA4, future railway mobile radio communication standard FRMCS that is under development by the UIC and will be based on IP to replace the obsolete GSM-R standard. The results of the projects are supposed to be the basis for new requirements of interoperability of the updated version of CCS TSI to be released in 2022, as well as, presumably, recommendations for changes to be made to the CENELEC standards.

5. Conclusions

In the context of digital transformation, the development of state-of-the-art railway computer-based CCS systems implies an accelerated introduction of a whole range of innovative solutions and a wide application of commercial off-the-shelf components (COTS), thus making systems more complex and being capable of affecting the dependability parameters. In order to maintain these parameters at a specified level and to minimize the impact of human factors, the railway community is increasingly using formal methods and automated means of engineering, diagnostics and monitoring at all stages of a system’s lifecycle.

A major factor of dependability is the standardisation of the system’s architecture, interfaces, open source design and testing software, including the standardisation of approaches to remote lab testing of products by different manufacturers to prove the reliability of operation at the boundaries of systems of various manufacturers. A potential future development of a common CCS ontology and standardisation of methods and tools for engineering, testing and maintenance based on the principles of interoperability and whitebox solutions to avoid vendor lock-in for railway companies can provide railway transportation with a competitive edge compared to other modes of transportation.

Evidently, there is yet another large area of research and practical activities which has by no means been covered in this paper, and that is the application of digital sensors and digital models, as well as integrated information systems intended for monitoring and prediction of the system dependability parameters, identification of pre-failure states based on the formal description and simulation of possible degradation scenarios using Data Science and Big Data. But this might be a topic for a separate study.

References

1. Doppelbauer J. Command and Control 4.0. IRSE News. 2018;246.

2. IEC 62290:2014. Railway applications – Urban guided transport management and command/control systems.

3. Interoperability Directive 2008/57/EC.

4. CCS (EU) No. 2016/919: Technical Specification of Interoperability relating to Control-Command and Signalling.

5. UNISIG Subset-026-2_v360.

6. Zamyshliaev A.M. [Applied information systems for management of dependability, safety, risks and resources in railway transportation]. Ulyanovsk: Pechatnyi dvor; 2013. (in Russ.)

7. BS EN 50126:1999. Railway applications – The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS).

8. EUG. ERTMS/ETCS RAMS Requirements Specification; 1998.

9. Rumsey A. Achieving high levels of signalling system availability – is there a role for secondary systems? IRSE News. 2018; 247.

10. https://Shift2Rail.org/.

11. Shift2Rail Plasa 2. Deliverable D 4.1: Virtual Certification: State of the art, gap analysis and barriers identification, benefits for the Rail Industry; 2019.

12. BS EN 50128:2011. Railway applications – Communication, signalling and processing systems – Software for railway control and protection systems.

13. Antoni M. Formal Validation Method and Tools for French Computerized Railway Interlocking Systems. IRJ. 2009;2(3):99-106.

14. Van der Werff M., Elsweiler B., Luttik B., Hendriks P. The use of formal methods in standardisation of interfaces of signalling systems. IRSE News. 2019;256.

15. EUG EULYNX. RCA Alpha – Architecture Overview; 2019.

16. OCORA Architecture – Alpha Release; 2019.

17. Shift2Rail ASTRail. D4.1 Report on Analysis and on Ranking of Formal Methods; 2019.

18. Estevan A.M. Dependability and safety evaluation of railway signalling systems based on field data. Doctoral thesis; Luleå 2015.
19. Xie Y. Formal Modeling and Verification of Train Control Systems. Thesis; 2019.
20. Shift2Rail Multi-Annual Plan; 2015.

About the Author

Alexey V. Ozerov, Head of International Department, JSC NIIAS, Moscow, Russian Federation, phone. +7 (495) 967-77-02, e-mail: A.Ozerov@vniias.ru

The author's contribution

Ozerov A.V. has analyzed the key parameters of railway control, command and signalling using the example of ERTMS/ETCS and considering RAMS requirements based on international standards, and outlined the major directions of evolution of the approaches to engineering, certification and homologation withing the framework of digital transformation and transition to new paradigms of innovations implementation.

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.

DEPENDABILITY JOURNAL ARTICLE SUBMISSION GUIDELINES

Article formatting requirements

Articles must be submitted to the editorial office in electronic form as a Microsoft Office Word file (*.doc or *.docx extension). The text must be in black, on a A4 sheet with the following margins: 2 cm for the left, top and bottom margins; 1.5 or 2 cm for the right margin. An article cannot be shorter than 5 pages and longer than 12 pages (can be extended upon agreement with the editorial office). The article is to include the structural elements described below.

Structure of the article

The following structural elements must be separated with an *empty line*. Examples of how they must look in the text are shown *in blue*.

1) Title of the article

The title of the article is given in the English language.

Presentation: The title must be in 12-point Times New Roman, with 1.5 line spacing, fully justified, with no indentation on the left. The font face must be bold. The title is not followed by a full stop.

An example:

Improving the dependability of electronic components

2) Author(s)' name.

This structural element for each author includes:

In English: second name and first name as "First name, Second name" (John Johnson).

Presentation: The authors' names must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with no indentation on the left. The font face must be bold. The authors' names are separated with a comma. The line is not followed by a full stop.

An example:

John Johnson[1], Karen Smith[2*]

3) The author(s)' place of employment

The authors' place of employment is given in English. Before the place of employment, the superscripted number of the respective reference to the author's name is written.

Presentation: The reference to the place of employment must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with no indentation on the left. The font face must be normal. Each place of employment is written in a new line. The lines are not followed by a full stop.

An example:

[1] Moscow State University, Russian Federation, Moscow

[2] Saint Petersburg Institute of Heat Power Engineering, Russian Federation, Saint Petersburg

4) The e-mail address of the author responsible for maintaining correspondence with the editorial office

Presentation: The address must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with no indentation on the left. The font face must be normal, all symbols must be lower-case. Before the address reference, symbol * is written. The title is not followed by a full stop.

An example:

*johnson_j@aaa.net

5) Abstract of the article

This structural element includes a structured summary of the article with the minimal size of 350 words and maximum size of 400 words. The abstract is given in the English language. The abstract must include (preferably explicitly) the following sections: Aim; Methods; Results/Findings; Conclusions. The abstract of the article should not include newly introduced terms, abbreviations (unless universally accepted), references to literature.

Presentation: The abstract must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with no indentation on the left. The font face must be normal, except "**Abstract**", "**Aim**", "**Methods**", "**Conclusions**", that (along with the full stop) must be in bold. The text of the abstract must not be paragraphed (written in a single paragraph).

An example:

Abstract. Aim.Proposing an approach … taking into consideration the current methods. **Methods.** The paper uses methods of mathematical analysis,…, probability theory. **Results.** The following findings were obtained using the proposed method … **Conclusion**. The approach proposed in the paper allows…

6) Keywords

5 to 7 words associated with the paper's subject matter must be listed. It is advisable that the keywords complimented the abstract and title of the article. The keywords are written in English.

Presentation: The text must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with no indentation on the left. The font face must be normal, except "**Keywords:**" that (along with the colon) must be in bold. The text must not be paragraphed (written in a single paragraph). The text must be followed by a full stop.

An example:

Keywords: dependability, functional safety, technical systems, risk management, operational efficiency.

7) Text of the article

It is recommended to structure the text of the article in the following sections: Introduction, Overview of the sources, Methods, Results, Discussion, Conclusions. Figures and tables are included in the text of the article (the figures must be "In line with text", not "behind text" or "in front of text"; not "With Text Wrapping").

Presentation:

The titles of the sections must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with no indentation on the left. The font face must be bold. The titles of the sections (except the Introduction and Conclusions) may be numbered in Arabic figures with a full stop after the number of a section. The number with a full stop must be separated from the title with a no-break space (Ctrl+Shift+Spacebar).

The text of the sections must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with a 1.25-cm indent. The font face must be normal. The text of the sections must be paragraphed. There must be no indent in the paragraph that follows a formula and contain notes to such formula, e.g.:

where n is the number of products.

An example:

1. State of the art of improving the dependability of electronic components

An analysis of Russian and foreign literature on the topic of this study has shown that …

Figures (photographs, screenshots) must be of good quality, suitable for printing. The resolution must be at least 300 dpi. If a figure is a diagram, drawing, etc. it should be inserted into the text in editable form (Microsoft Visio). All figures must be captioned. Figures are numbered in Arabic figures in the order of their appearance in the text. If a text has one figure, it is not numbered. References to figures must be written as follows: "Fig. 3. shows that …" or "It is shown that … (see. Fig. 3.)." The abbreviation "Fig." and number of the figure (if any) are always separated with a no-break space (Ctrl+Shift+Spacebar). The caption must include the counting number of the figure and its title. It must be placed a line below the figure and center justified:

Fig. 2. Description of vital process

Captions are not followed by a full stop. *With center justification there must be no indent!* All designations shown in figures must be explained in the main text or the captions. The designations in the text and the figure must be identical (including the differences between the upright and oblique fonts). *In case of difficulties with in-text figure formatting, the authors must – at the editorial office's request – provide such figures in a graphics format (files with the* *.tiff, *.png, *.gif, *.jpg, *.eps extensions).

The tables must be of good quality, suitable for printing. The tables must be editable (not scanned or in image format). All tables must be titled. Tables are numbered in Arabic figures in the order of their appearance in the text. If a text has one table, it is not numbered. References to tables must be written as follows: "Tab. 3. shows that …" or "It is shown that … (see. tab. 3.)." The abbreviation "tab." and number of the table (if any) must be always separated with a no-break space (Ctrl+Shift+Spacebar). The title of a table must include the counting number and its title. It is placed a line above the table with center justification:

Table 2. Description of vital process

The title of a table is not followed by a full stop. *With center justification there must be no indent!* All designations featured in tables must be explained in the main text. The designations in the text and tables must be identical (including the differences between the upright and oblique fonts).

Mathematical notations in the text must be written in capital and lower-case letters of the Latin and Greek alphabets. Latin symbols must always be oblique, except function designators, such as sin, cos, max, min, etc., that must be written in an upright font. Greek symbols must always be written in an upright font. The font size of the main text and mathematical notations (including formulas) must be identical; in Microsoft Word upper and lower indices are scaled automatically.

Formulas may de added directly into the text, for instance:

Let $y = a \cdot x + b$, then…,

or written in a separate line with center justification, e.g.:

$$y = a \cdot x + b.$$

In formulas both in the text, and in separate lines, the punctuation must be according to the normal rules, i.e. if a formula concludes a sentence, it is followed by a full stop; if the sentence continues after a formula, it is followed by a comma (or no punctuation mark). In order to separate formulas from the text, it is recommended to set the spacing for the formula line 6 points before and 6 points after). If a formula is referenced in the text of an article, such formula must be written in a separate line with the number of the formula written by the right edge in round brackets, for instance:

$$y = a \cdot x + b. \quad (1)$$

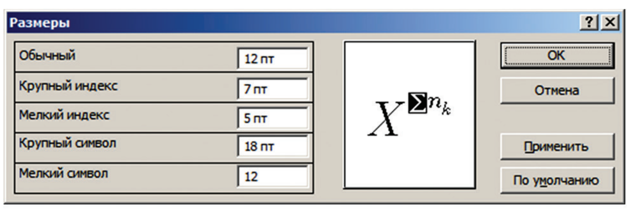
If a formula is written in a separate line and has a number, such line must be right justified, and the formula and its number must be tab-separated; tab position (in cm) is to be chosen in such a way as to place the formula roughly at the center. Formulas that are referenced in the text must be numbered in Arabic figures in the order of their appearance in the text.

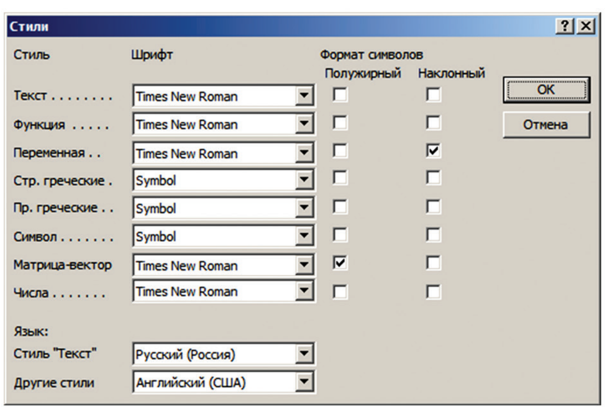
Simple formulas should be written without using formula editors (in MS Word, Latin should be used, as well as the “Insert” menu + “Special Characters”, if Greek letters and mathematical operators are required), while observing the required slope for Latin symbols, for example:

$$\Omega = a + b \cdot \theta.$$

If a formula is written without using a formula editor, letters and +, –, = signs must be separated with no-break spaces (Ctrl+Shift+Spacebar).

Complex formulas must be written using a formula editor. In order to avoid problems when editing and formatting formulas it is highly recommended to use Microsoft Equation 3.0 or MathType 6.x. In order to ensure correct formula input (symbol size, slope, etc.), below are given the recommended editor settings.





When writing formulas in an editor, if brackets are required, those from the formula editor should be used and not typed on the keyboard (to ensure correct bracket height depending on the formula contents), for example (Equation 3.0):

$$Z = \frac{a \cdot \left(\sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_i \right)}{n+m}. \quad (2)$$

Footnotes in the text are numbered with Arabic figures, placed page by page. Footnotes may include: references to anonymous sources on the Internet, textbooks, study guides, standards, information from websites, statistic reports, publications in newspapers, magazines, autoabstracts, dissertations (if the articles published as the result of thesis research cannot be quoted), the author’s comments.

References to bibliographic sources are written in the text in square brackets, and the sources are listed in the order of citation (end references). The page number is given within the brackets, separated with a comma and a space, after the source number: [6, p. 8].

8) Acknowledgements

This section contains the mentions of all sources of funds for the study, as well as acknowledgements to people who took part in the article preparation, but are not among the authors. Participation in the article preparation implies: recommendations regarding improvements to the study, provision of premises for research, institutional supervision, financial support, individual analytical operations, provision of reagents/patients/animals/other materials for the study.

Presentation:

The information must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with no indentation on the left. The font face must be normal.

9) References

The References must include only peer-reviewed sources (articles from academic journals and monographs) mentioned in the text of the article. It is not advised to references autoabstracts, dissertations, textbooks, study guides, standards, information from websites, statistic reports, publications in newspapers, websites and social media. If such information must be referred to, the source should be quoted in a footnote.

The description of a source should include its DOI, if it can be found (for foreign sources, that is possible in 95% of cases).

References to articles that have been accepted, but not yet published must be marked “in press”; the authors must obtain a written permission in order to reference such documents and confirmation that they have been accepted for publication. Information from unpublished sources must be marked “unpublished data/documents”; the authors also must obtain a written permission to use such materials.

References to journal articles must contain the year of publication, volume and issue, page numbers.

The description of each source must mention all of its authors.

The references, imprint must be verified according to the journals’ or publishers’ official websites.

Presentation:

References must be written in accordance with the Vancouver system.

The references must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with a 1.25-cm indent on the left. The font face must be normal. Each entry must be numbered in Arabic figures with a full stop after the number. The number with a full stop must be separated from the entry with a no-break space (Ctrl+Shift+Spacebar).

10) About the authors

Full second name, first name (in English); complete mailing address (including the postal code, city and country); complete name of the place of employment, position; academic degree, academic title, honorary degrees; membership in public associations, organizations, unions, etc.; official name of the organization in English; e-mail address; list and numbers of journals with the author's previous publications; the authors' photographs for publication in the journal.

Presentation:

The information must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with no indentation on the left. The font face must be normal.

11) The authors' contribution

Detailed information as to each author's contribution to the article. For example: Author A analyzed literature on the topic of the paper, author B has developed a model of real-life facility operation, performed example calculation, etc. Even if the article has only one author, his/her contribution must be specified.

Presentation:

The information must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with no indentation on the left. The font face must be normal.

12) Conflict of interests

A conflict of interests is a situation when people have conflicting and competing interests that may affect editorial decisions. Conflicts of interests may be potential or conscious, as well as actually existing. The objectivity may be affected by personal, political, financial, scientific or religious factors.

The author must notify the editorial office on an existing or a potential conflict of interests by including the corresponding information into the article.

If there is no conflict of interests, the author must also make it known. An example of wording: "The author declares the absence of a conflict of interests".

Presentation:

The text must be in 12-point Times New Roman, with a 1.5-line spacing, fully justified, with no indentation on the left. The font face must be normal.

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

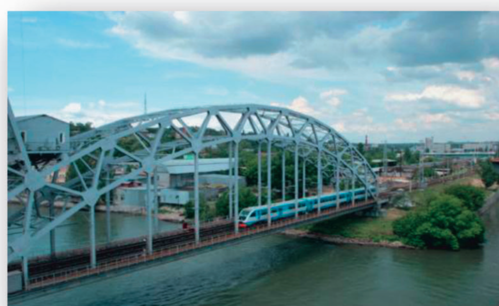
Mission:

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