Big Data-Based Methods for Functional Safety Case Preparation

Efim Rozenberg, Alexey Olshansky, Alexey Ozerov

Research and Design Institute for Information Technology, Signalling and Telecommunications on Railway Transport (NIIAS), Moscow, Russia, a.ozerov@vniias.ru

Contribution to the State of the Art

https://doi.org/10.7251/JIT2302091R UDC: 004.42.032.26:007.52

Abstract: The paper aims to overview the opportunities, approaches and techniques of studying and ensuring functional safety of transportation systems, including those driverless, with the use of Big Data. Examples are provided of machine learning/Big Data application in analysing the functional safety of complex control/management systems in railway transportation. The paper proposes the concept of application of supervised artificial neural networks combined with model checking. The following methods were used in the preparation of the paper: system analysis, logical and comparative analysis and historical principle. Updated requirements are defined for transportation systems using artificial intelligence as part of adaptive train schedule management and autonomous train control. That will ultimately allow developing an entire line of research from AI-based system functional safety estimation and machine learning to safety case preparation of intelligent supervised control/management systems based on formal verification.

Keywords: functional safety, safety case, code verification, supervised artificial neural networks (SANN), machine learning, Big Data, driverless control systems, train schedule, Markov chains.

Introduction

The modern technology that underpins next-generation transportation systems that operate in ever-evolving conditions, with significant numbers of passengers, requires modified control systems design. With the growth of agglomerations, many suburban and urban systems merge, and the headways approach those of the subway. In this context, man-machine systems are transforming into automatic ones with varied degrees of automation (from GoA1 to GoA4). Failures and delays that occur within such transportation systems cause significant disruptions that affect thousands of passengers and require the mobilization of infrastructure and technical assets. In the above scenarios of transportation system operations, it becomes impossible to use the conventional approaches to traffic schedule redesign and to transportation service planning using algorithms for static problem optimization (which, no doubt, can well be used in long-term planning).

Mitigating the above challenges requires solving the following primary tasks:

1. Improving the adaptive quality of the planning and management processes, in particular, in terms of schedule design/redesign;

2. Improving the resilience of the transportation system and its technical component to unsafe behaviours, failures and disruptive effects.

The significantly increasing costs of both an hour of time and infrastructure resource for companies, and trains and infrastructure assets favour the transition to computer simulation and formal analysis of all possible situations in the transportation system that is enabled by modern approaches and methods for developing adaptive and functionally safe management. In the context of the above problems, the conventional methods [1, 2, etc.] have quite limited capabilities.

Definition of the scope of research

Drawing on the existing standardization guidelines, we should consider the concept of "umbrella standard", i.e., a basic, high-level standard. In case of functional safety, that is IEC 61508 *Functional safety of electrical/electronic/programmable electronic safety-related systems*. IEC/GOST 61508 is a basic functional safety standard applicable to all industries.

According to IEC 61508, for the purpose of ensuring functional safety, first, the safety functions are to be defined that are required for reducing the risk associated with the controlled equipment, as well as for achieving and maintaining the safety of such equipment (e.g., emergency shut down function). It is very important that, according to the standard, a control system is to possess the property of so-called safety integrity, by which IEC 61508 means the probability that the system will correctly perform the safety functions under all specified conditions, within the specified period of time.

In practice, along with IEC 61508, industry-specific functional safety standards are used as well. For instance, in railway transportation, there is GOST 33433-2015 *Functional safety. Risk management on railway transport* that specifies the approach and general rules for managing risks in railway transportation associated with the functional safety of infrastructure and rolling stock. In addition, there is GOST 33432-2015 *Functional safety. Policy and program of safety provision. Safety proof of the railway objects* that defines the purpose of the "Safety Policy", "Safety Program" and "Safety Case", as well as specifies the primary requirements for the structure and content of those documents and the procedure for their development.

The following primary methods for safety case preparation are identified [3]:

- expert, based on expert evaluation of technical and design documentation;
- computational, based on analytical computations;
- simulation, based of experiments with computer models;
- experimental, based on experimental tests with a trial system (laboratory tests);
- full-scale testing that involves testing the system in actual operation conditions at the stage of commissioning and run-in, certification tests;
- information-based that involves the collection of statistical data on failures in the course of long-term operation of a single system or a number of same-type systems.

The choice of one or several specific methods depends on the developer's qualification and used regulatory framework.

The specificity of railway operations today

Currently, there is a number of distinctive features of the 1520 mm gauge railways that should be pointed out. In railway transportation, freight traffic is concentrated on certain lines. The primary load is on about 10 percent of its operational length. Historically, in the Russian Federation, about half of the total freight turnover is ensured by 1/6 of the railways. The situation is similar in passenger transportation.

Due to this uneven distribution of operations throughout the railway network, its certain parts become extremely busy, which ultimately affects the entire network. The primary cause of such "bottlenecks" is the insufficient capacity of the operational regions. Railway lines may also experience loads outside of the permitted capacity in case of insufficient traction power supply and length of receiving, marshalling, turnout and departure tracks at intermediate, line and marshalling stations. That reduces station capacity, causes train delays at entrance signals and generally reduces the service speed of passenger and freight trains.

In practice, improving the theoretical and practical capacity of lines subject to the existing limitations involves a comprehensive approach that requires significant investment. That includes the construction of main tracks, station tracks, delivery of modern locomotives, electrification, improvement of traction power supply, signalling upgrades (e.g., implementation of moving block sections). At the same time, improvements to the transportation management process allow reducing capital expenditures if local solutions are used that are adapted to specific facilities and sites.

In this context, point technological solutions should be used for the purpose of increasing the theoretical and practical capacity in the short term

until the completion of major infrastructure projects and to ensure the performance of the development program.

For instance, the last decade saw widespread deployment of digital telecommunications, process automation and remote data collection and management technology in railway transportation. Managing railway operations through the use of sensors and microcontrollers, as well as programmable and remotely controlled railway signals and switches, has resulted in increased system efficiency as well as operational flexibility. However, the use of network connectivity made railway data communications vulnerable to cyber-attacks. Today, an increasing number of railway data transmission networks are cyber-physical systems with interconnected physical, computer and communication components. Cyber-attacks against such systems can potentially cascade through those interconnections and cause significant damage. These systems are critical to safety due to the great financial implications and, more importantly, potential threats to human life. Therefore, safety and reliability requirements for such systems are to be taken into account at the very beginning of their design [4].

Thus, an accelerated adoption of new technology requires new approaches to safety case preparation and safety assurance. Most importantly, that involves the development of alternative and target schedules for large regions of operations and stations, especially, with possible large-scale application of unmanned vehicles in passenger transportation and shunting operations. Due to their complexity, all such new technical solutions require intelligent control elements. The solution may be in the artificial intelligence and deep neural networks along with big data processing as a more advanced computational method for functional safety case preparation.

Big Data-based safety case process for systems with artificial intelligence

One of the difficult aspects of a safety case is the identification of abnormal scenarios that may potentially cause a specific accident. At this difficult stage of safety assessment, experts are to be able to understand, among other things, the specificity of the artificial intelligence technology.

For example, [5] suggested an ACASYA-based model for calculating hazardous scenarios of automatic railway device operation. The software tools presented in this paper have two main functions. First, to record and store the experience associated with safety analysis. Second, to help those involved in system development and evaluation, as well as in solving the complex problem of evaluating safety trials. Currently, those tools are at the prototype stage, yet safety experts have noted the strong potential of the proposed approaches.

The approach chosen in this paper is based on several capabilities of artificial intelligence and, in particular, on the use of the following methods:

- accumulation of railway safety information, including potential accident scenarios;
- training through concept classification for the purpose of grouping accident scenarios into homogeneous classes associated with train collisions or derailments;
- rule-based machine learning (RBML) for automatic identification – based on a database of historical scenarios (experience feedback) – of appropriate safety rules that are often difficult to extract manually;
- knowledge-based system (KBS) that is filled with the process rules previously derived using machine learning for the purpose of creating a knowledge base for a functional safety analysis support tool.

Thus, the approach to railway transport safety assessment used in this paper is a hybrid method based on a classification algorithm, rule-based machine learning (RBML) and a knowledge-based system (KBS).

As shown in Fig. 1, the railway safety analysis and assessment methodology consists of 11 stages. The first eight steps are performed by the scenario classification module (CLASCA), while the last three are performed by the scenario evaluation module (EVALSCA).

Advanced information technology is increasingly employed as part of controlling and controlled equipment, whose correct operation might have an effect on the life and health of people. Information security is vital to ensuring comprehensive functional safety, therefore efficient methods for safety evaluation and safety case preparation are required [6].

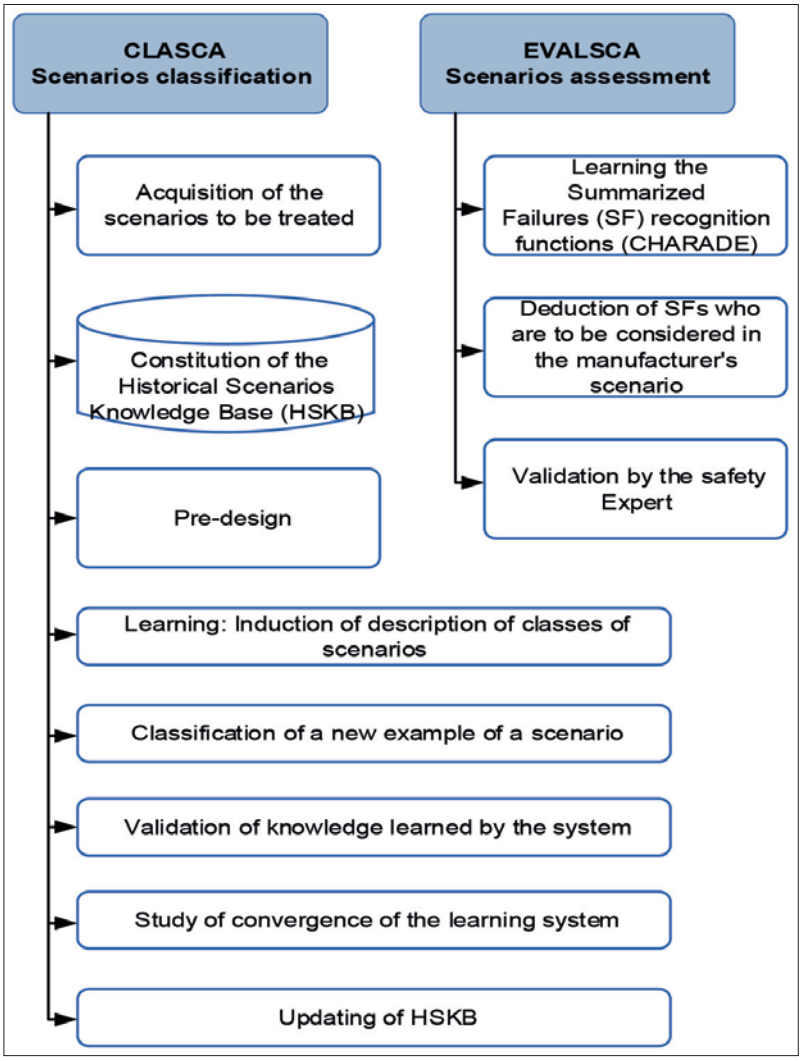

Fig. 1. *Safety analysis and assessment method*

Thus, [7] examined the process of safety case preparation of a complex information system using artificial intelligence.

Convergence indicators, speed and accuracy of the inverse error distribution (BP) neural network algorithm, particle swarm method (PSO), genetic algorithm, GA-PSO and PSO-BP algorithm were compared in the study of information system operation risks (Fig.2).

As a result of the simulation experiment, the error of the PSO-BP algorithm in predicting information system risks is practically 0, the error of the conventional BP algorithm is 3.87, while the maximum error of the PSO algorithm is 1.12 units.

However, it should be noted that a significant drawback of the proposed method consists in the requirement of prior knowledge of the possible risks for the examined system, whose insufficient availability is noted by the authors.

In general, the use of safety case assessment methods for systems with artificial intelligence is arguably a poorly studied domain that should be based on both the experience of successfully implemented standards, and new developments [8].

Today, the issue is particularly pressing for driverless cars. It is also common in the application of artificial intelligence for direct equipment control, where documents regulating the methods of safety assessment are lacking. There are many works dedicated to this topic, including comprehensive studies ([7], etc.) that suggest ways to solve the problem of shortage of safety assessment methods, yet they do not address the matter of safety integrity demonstration for artificial intelligence.

Let us examine how the matters of safety management can be implemented in the train schedule

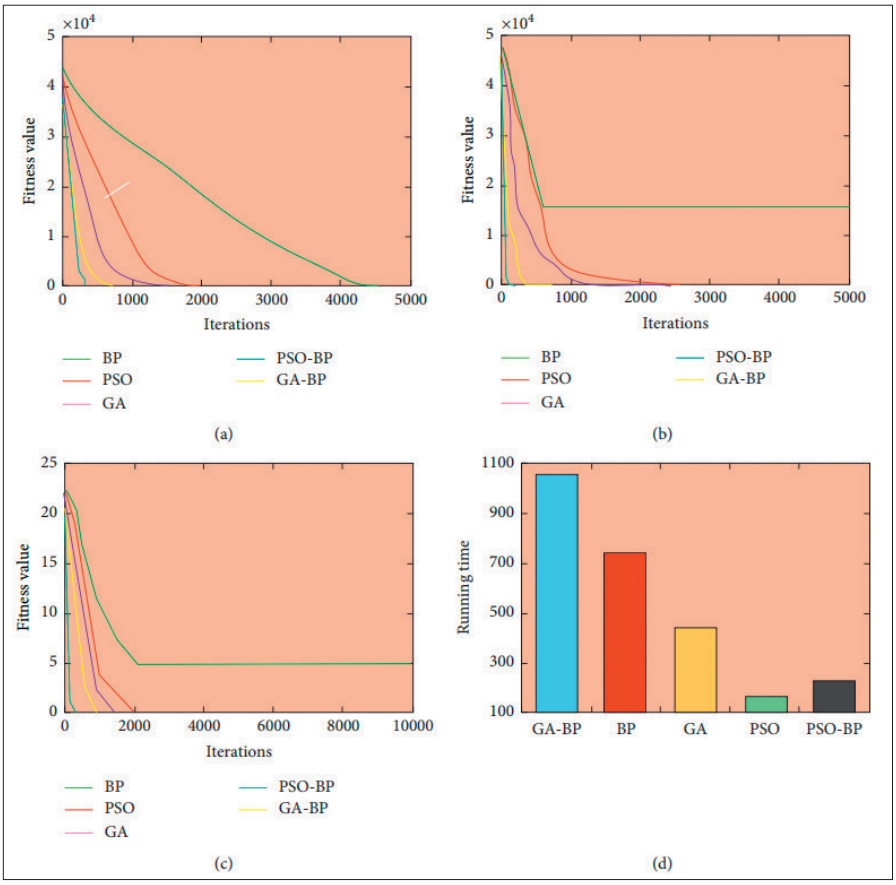

Fig. 2. *Convergence measures of five algorithms: (a) First training; (b) Second training; (c) Third training; (d) Time.*

process. It should be noted that the International Union of Railways (UIC) regards the transition to adaptive train schedule and its life cycle-based management as a most important component of the railways' digital transformation [9].

The notion of train schedule life cycle includes an entire set of phases from the concept, definition of operating conditions to the actual information implementation and logging of the schedule performance, each of which refers to the corresponding information fields involved in the associated information model.

These parameters may be defined by regulatory means (from the corresponding database) or by analysing the input information using an intelligent algorithm, including with the use of machine learning examples and historical context.

Up until now, literature has not considered the relationship between the phases of the train schedule life cycle and the matters of its functional dependability and safety. Although, essentially, each of the parameter values in the train schedule phases actually defines a certain level of safety on the line.

The set of automation devices, data communication and man-machine systems (operation of traffic controller, train driver, station duty officer) are elements of functional scenario trees, each of which results in an assessment of the risk of technology or process-related failure.

As a result, at each phase, a hazardous failure scenario can be defined and protection measures can be foreseen for a train schedule. Here, the term "failure" should be interpreted broadly, including a wide class of situations, in which the safety level is not below standard, yet the line capacity does not allow handling the specified train traffic [10].

In this situation, Markov chains may prove to be the most efficient method. For each scenario branch, Markov chains are thus formed with specified state transition rates that are compactly written in matrix form.

The URRAN methodology would be very applicable in this area [1]. That will enable a comprehensive approach to managing the functional safety and dependability of train schedules, in which key indicators, i.e., the transition rates, will be evaluated using Data Science.

Thus, the transition to a trinity of models is possible, i.e., a historical, a dynamic and a predictive data landscapes, built on common principles that allows handling information data models in relation to different time horizons and in accordance with local goals. Evidently, obtaining reliable results requires that the data generated using Big Data methods are cleared of noise, validated and submitted to other standard procedures.

Today, train schedules are implemented in the form of specific management of assets: train routing, section and switch location occupancy supervision, etc., that are associated with specific time parameters. The fact that changes in the asset status have no effect on the target schedules that are modified no more than once a day represents a particular difficulty. The migration towards adaptive train schedules with continuous replenishment with data generated using Big Data-based methods will enable a shorter transition process caused by the normalization of the operational situation in case of certain disturbances. Currently, under the existing process that does not involve Big Data-based methods, the schedule is largely unable to keep up with the evolution of asset statuses.

In [12], the matter of supervised artificial neural networks was examined given their application in process management for the purpose of ensuring the required functional safety of systems.

Control and management systems are conventionally assessed for Lyapunov's stability. In this case, the behaviour of a stable system can with a 100% probability be predicted in the neighbourhood of the ε-tube [13].

For the examined supervised systems, in which stability is ensured through the introduction of a supervisor algorithm, speaking of a strict Lyapunov's stability would not be correct.

Let us consider the operation of the above diagram (see Fig. 3). At the first step, an external signal is fed to the input of the supervised ANN. The latter operates as an output-controlled system. The output of the network (with the addition of feedback control) is formally verified. If the solution belongs to the set of acceptable processes D, such signal is fed to the actuator and further to the controlled item (which may include the network). If the output signal is outside the set of acceptable processes, the limiter is triggered. The limiter is the history of processes and reactions to a specific implementation (managerial decision-making algorithm with no clear indication of the nature of such algorithm). The algorithm suggests a decision (in terms of the output vector) with confidence P%. If the decision is deemed belonging to the set of acceptable processes, it is fed to the actuator.

In order to clarify the specificity of the proposed diagram, let us note that, firstly, SANN has latency (if a correct decision is not produced at the second step, then between the duration of 2 steps and indefinitely, until it is abruptly interrupted by the DM using the time control unit), and secondly, the algorithms

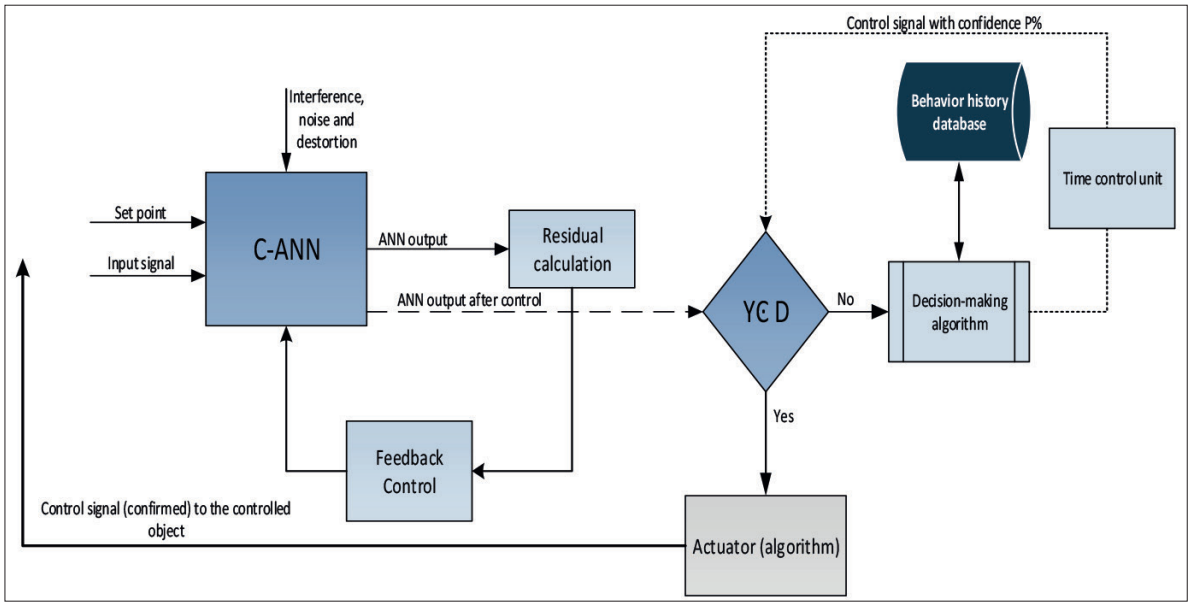

Fig. 3. *Block diagram of the supervised ANN*

for acceptability evaluation and development of decisions in the supervisor are to be sufficiently fast, so that the total latency was reasonable. In addition, the confidence level P will always be less than 100%.

The diagram should also verify whether the output values belong to the set of acceptable processes D. The evaluation of the boundaries of domain D, continuous adjustment of the boundaries of the ε-tube of stability, and identification of the limiter's behaviour are to be done using Big Data-based algorithms and methods.

The extension of this practice (application of SANN and, more generally, supervised machine learning algorithms) to train schedule data processing enables the transition to functionally dependable and adaptive scheduling. This hypothesis is to be further studied, as it is required to:

1. Prove the controllability of algorithms and methods of information processing at least "at the output".
2. Synthesize a practically implementable method of control and evaluate its stability.

Such approach allowed examining the possible implementations of partially controlled intelligent systems in [10]. In literature, partially controlled (output) ANN with variable signal conductivity were first described in papers funded through the Russian Foundation for Fundamental Research grant no. 17-20-01065 *Development of the Theory of Neural Network-Based Control of Railway Transportation Systems* [14, 15]. The provisions on ANN control defined therein may, in principle, be extended to other machine learning algorithms. Only after those studies have been completed, specific software and hardware implementation should commence.

It should be noted that the idea of algorithms that are supervised in this manner may apply not only to ANN, but other intelligent algorithms as well. At the same time, the currently used rules and principles of strict verification of model checking algorithm branches should not be unconditionally rejected [16].

Scope of practical application of the findings

Unmanned train control systems are being rapidly deployed in the Moscow Central Circle and the Luzhskaya station. The most complicated issue is the safety case preparation, since the key component is the convolutional ANN-based machine vision. Using the methods proposed above, approaches have already been developed for specific engineering implementations that are undergoing trials both on specially designed laboratory benches, and in the field [17].

Additionally, future development of such intelligent systems will be ensured by in-depth research of signal recognition technology (sound analysis, machine vision, self-diagnosis systems, etc.). Matters of sensor interpretation, sensor elements and associated decision-making systems are in close relation with the proposed algorithms. In [18], a single safety case system is proposed for each intelligent autonomous transportation system. Within such system, a zone is defined that is not covered by SIL-4. That is the Intelligent Train Protection zone that is characterized by the uncertainty of system behaviour or partial observability. In these conditions, one of the ways to overcome these gaps may be to apply Big Data processing techniques to controlled algorithms (including the controlled ANNs described above). These techniques may prove to be especially useful for MAPE (Monitor – Analyse – Plan – Execute) signal processing. Identifying and eliminating abnormal signals will allow more clearly defining the boundaries of the set of acceptable processes D, thus, in some cases, increasing the speed of the decision algorithms by disabling an entire branch of unfavourable scenarios.

Given the developed and above-described approaches, as well as the papers on the examined matter, model checking is to be used in any case.

Conclusion

Thus, this paper analysed the current requirements for transportation systems, including those that use artificial intelligence, for the most promising areas, i.e., adaptive train schedule and unmanned systems, which allows defining an entire new line of research, i.e., assessment of the functional safety of systems using AI and machine learning.

A design is proposed that is promising in terms of further computer research, i.e., supervised ANN. The authors also substantiate further lines of research in the area of intelligent supervised systems, including the transition to the preparation of such system's safety cases based on formal verification of the developed control systems.

REFERENCES

[1] Gruntov P.S. [Centralized systems for automated railway management in market economy conditions]. Minsk: BelGUT; 2011. (in Russ.)

[2] Osminin A.T. Problems and ways of their scientific solving on the issues of railroading. *Bulletin of JSC RZD United Academic Council* 2015;4:41-54. (in Russ.)

[3] Fedukhin A.V. Mukha Ar.A., Sespedes Garsiya N.V. [Safety case of a computer system]. *MMS* 2016;3:93-101. (in Russ.)

[4] Neema H. Simulation testbed for railway infrastructure safety and resilience evaluation. In: Proceedings of the 7th Symposium on Hot Topics in the Science of Security. Association for Computing Machinery; 2020. P. 1-8.

[5] Hadj-Mabrouk H. Contribution of Artificial Intelligence to Risk Assessment of Railway Accidents. *Urban Rail Transit* 2019;5:104-122.

[6] Shubinsky I.B., Schäbe H., Rozenberg E.N. On the functional safety of a complex technical control system with digital twins. *Dependability* 2021;21(1):38-44.

[7] Zhang J., Li J. Testing and verification of neural-network-based safety-critical control software: A systematic literature review. *Information and Software Technology* 2020;123.

[8] Shubinsky I.B., Shäbe H., Rozenberg E.N. On the safety assessment of an automatic train operation system. *Dependability* 2021;21(4):31-37.

[9] Ozerov A.V., Lysikov M.G., Olshansky A.M. Timetable as part of next-generation adaptive management system. *Nauka i tekhnologii zheleznykh dorog* 2021;5(1):50-64.

[10] Ozerov A.V., Olshansky A.M. [Approaches to the assessment of the functional safety of a driverless automatic train control system]. In: International Scientific Conference Proceedings "Advanced Information Technologies and Scientific Computing". Samara; 2021. P. 504-509. (in Russ.)

[11] Gapanovich V.A., Shubinsky I.B., Zamyshlyaev A.M. Mathematical and information support of the URRAN system. *Dependability* 2013;1:12-19.

[12] Ozerov A.V., Olshansky A.M. Safety model construction for a complex automatic transportation system. *Dependability* 2021;21(2):31-37.

[13] Dorf R., Bishop R. Modern control systems. Moscow: Lab. Bazovykh Znaniy; 2004.

[14] Rozenberg E.N. et al. [Hybrid neural network-based management of transportation systems]. *Automation, Communications, Informatics* 2017;12:2-5. (in Russ.)

[15] Ignatenkov A.V., Olshansky A.M. About neural network error control as an optimal control problem. *Izvestiya Samarskogo nauchnogo centra Rossiyskoy akademii nauk* 2016;18(4-4):733-738. (in Russ.)

[16] Clarke E.M., Henzinger T.A., Veith H., Bloem R., editors. Handbook of model checking. Vol. 10. Cham: Springer; 2018.

[17] Okhotnikov A.L., Popov P.A. Self-driving: yesterday, today and tomorrow. *Automation, Communications, Informatics* 2019;8:12-17. (in Russ.)

[18] Flammini F. et al. A Vision of Intelligent Train Control. In: Proceedings of the 4th International Conference on Reliability, Safety and Security of Railway Systems (RSSRail-22). Springer LNCS; 2022. (preprint 2022)

Received: September 10, 2023
Accepted: November 1, 2023

ABOUT THE AUTHORS

Efim Rozenberg is the First Deputy Director General of JSC NIIAS, Doctor of Engineering, Distinguished Professor. He is a recipient of the award of the Government of the Russian Federation in the field of science and technology. He was awarded the title of Honored Designer of the Russian Federation, Best Innovator of JSC Russian Railways. Professor Rozenberg leads research and development in train control and protection including signalling, train separation, automatic train operation, traffic safety, communication, cybersecurity. He is an author of about 300 research papers and about 400 patented inventions.

Alexey Olshansky is the Head of Centre for Mathematical and Computer Simulation of JSC NIIAS, PhD in Engineering. His areas of interest are system analysis, theory of automatic control of socio-economic systems, artificial neural networks and heuristic methods for designing control systems. He is the author of over 150 scientific papers and 20 patented inventions.

Alexey Ozerov is the Head of International Department – Head of Intellectual Property Management Centre of JSC NIIAS. He has been working with JSC NIIAS for over 16 years in various positions related to research, signalling business unit and international cooperation. He is finishing his PhD in traffic management and has authored over 50 papers and 8 patents in the field of advanced railway signalling and traffic management.

FOR CITATION

Efim Rozenberg, Alexey Olshansky, Alexey Ozerov, Big Data-Based Methods for Functional Safety Case Preparation, *JITA – Journal of Information Technology and Applications, Banja Luka*, Pan-Europien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, JITA 13(2023) 2:91-98, (UDC: 004.42.032.26:007.52), (DOI: 10.7251/JIT2302091R, Volume 13, Number 2, Banja Luka, December (57-120), ISSN 2232-9625 (print), ISSN 2233-0194 (online), UDC 004