METHODS FOR ENSURING AND PROVING FUNCTIONAL SAFETY OF AUTOMATIC TRAIN OPERATION SYSTEMS

I.B. Shubinsky[1], E.N. Rozenberg[2], H. Schäbe[3]

•

[1]DsC, prof., NIIAS, Moscow, Russia, Igor-shubinsky@yandex.ru
[2]DsC, prof., NIIAS, Moscow, Russia, s,lavruhina@vniias.ru
[3]Doctor of Natural Sciences, TÜV Rheinland InterTraffic, Cologne, Germany, dr.hendrik.schaebe@gmail.com Abstract

Abstract

The paper examines the specificity of artificial intelligence-based automatic train operation systems. Justifying the functional safety (FS) of such systems is quite difficult. The paper proposes a process for proving the functional safety of intelligent systems. A hybrid control system for a shunting locomotive was developed and analysed. It combines machine vision (MV), train protection devices and manual control by a driver. A model is presented that allows examining the functional safety of a locomotive control system layer by layer, i.e., evaluating the time to safety degradation depending on the component failure and subsequent requirement of bringing the locomotive to a complete stop. This allows to improve the FS of the shunting locomotive control system with machine vision from SIL 2 to SIL 3 and maintaining it during sufficiently long periods of time (over a quarter of the mean time to system failure). The mean time of faultless operation of a locomotive control system until it has to be brought to a complete stop for safety reasons can be increased three times. A general approach is proposed to design the functional safety of automatic train operation systems. It is based on the division of the information processing process into two subprocesses, i.e., internal intelligent information processing onboard the locomotive for the purpose of decision-making regarding track vacancy and communication of initial visual information to the operating driver for decision-making. The division of this process must be combined with redundant machine vision facilities, regular comparison of the outputs of the onboard and fixed machine vision facilities, redundant comparison outputs, smoothing of the outputs in the process of locomotive movement.

Keywords: Functional safety, artificial intelligence, automatic train operation system, machine vision, dependability, safety justification, safety case, statistical and experimental methods, expert methods, simulation methods, heuristic semi-Markov graph methods, process methods of compliance confirmation, safety device, control system, Markov model, standards.

1. Introduction

The problem of ensuring functional safety of any technical system consists of two integral components. The first one consists in the development of proposals, techniques, procedures, methods for improving FS. The second component is intended for verifying the efficiency of the chosen method of improving safety. Essentially, the second component of the problem consists in

proving the acceptability of the achieved level of FS. Substantiating FS for railway control systems with the grade of automation GoA 2/3 (from manual driving with the function of automatic train operation to automatic train driving with no human driver) is quite difficult. These systems use artificial intelligence-based methods for the purpose of training information processing algorithms. One of the first works on artificial intelligence aimed at recognising patterns by training recognition algorithms was the monograph by Vapnik and Chervonenkis titled "Pattern recognition theory (statistical problems of training)" [1]. In [2 – 4], it was shown that an automatic train operation (ATO) system has a number of distinctive features. Those include the following:

1. Distributed system architecture.
2. Availability of machine vision and effect of weather conditions.
3. Close information interaction between the system and the environment via information communication channels.
4. Presence of a large and not always definite number of vulnerabilities within a system closely connected to the environment.
5. A high probability of evolving environmental effects and resulting changed system behaviour.
6. The changed control algorithm parameters as the result of neural network training using the incoming information flows and accumulated databases.
7. Branching software of both the generic part of the system, and, especially, rolling stock detection and control facilities.

Braband and Schäbe [2] note that due to the specificity of the ATO it requires special methods for proving the FS. It should be noted that one of the key features of the system is that, along with its distributed architecture, the connections within the system change significantly. The latter noticeably reduces the options to prove the safety of such a system.

Given the great uncertainty associated with the operation of ATO it is quite difficult to prove its FS using conventional methods, i.e., those set forth in STO RZD 1.19.009-2009 [5] that were largely applied to devices and simple systems with a known and limited number of vulnerabilities. The recommendations of IEC 61508-1-2012 (sections 6, 8) [6], IEC 61508-2-2012 [7], IEC 61508-3-2018 [8], and IEC 62279-2016 [9] may prove to be very helpful in this situation. Along the conventional methods of safety case preparation, the above standards suggest taking into account the design and manufacture process, quality and functional safety assurance organisation of complex hardware and software systems and their components for the purpose of evaluating the functional safety level of such systems. Such measures and procedures jointly solve the problem of *safety justification*. One of the components of a safety justification involves confirming the compliance with the specified requirements, which is largely ensured using the safety case. A development of this approach combined with the guidance material accumulated by the railway industry is reflected in GOST 33432-2015 [10].

As regards intelligent systems with the above distinctive features, the standard recommends the following scope of safety justification:

1. Development of an FS policy;
2. Development of an FS program;
3. Development of a safety case.

An *FS policy* is to be in place at the ATO system manufacturer and is to be generally applied to all the products developed by such an organisation. It is to make provisions for solving the following main problems:

- tasks and objectives of FS assurance;
- principles and approaches to ensuring FS;
- principles of FS-related risk management;
- organisation of FS assurance.

In [11], Braband and Schäbe suggest using the outputs of the ATO-RISK project ordered by the Deutsche Zentrum für Schienenverkehrsforschung for the purpose of managing risks. The project aims to define the criteria of risk acceptability as regards automatic train operation. As described in [11], the risk level is evaluated through a function-specific explicit risk analysis or using reference systems. Explicit risk analysis is performed by evaluating various scenarios using the semi-quantitative approach and a risk score matrix. The matrix qualitatively differentiates the expected severity of harm depending on the category of the accidents. That approach can be recommended for the purpose of system safety policy definition.

FS assurance and FS case program are developed for each product autonomously and are intended to be supplied to the customer as proof of the product being of high quality in accordance with the requirements of the FS standards and corresponds to the declared safety integrity level (SIL). The ultimate goal of the ATO FS measures consists in the preparation of a safety case.

2. Characteristic features of the functional safety case of automatic train operation systems

The scope of FS case preparation includes reports on not only the FS status, but on the measures taken by the ATO manufacturer for managing quality and ensuring FS. Those reports allow the customer to evaluate the engineering level and manufacturing quality of the system, including the supply of components, organisation and process quality of the FS assurance activities, risk evaluation results, depth and quality of the FS requirements verification and validation activities.

A conclusion of an ATO's compliance with the FS requirements is built upon the FS status report taking into consideration the above reports on the quality and FS management measures. That is a very important consideration. The matter is that the distributed system architecture, changing parameters of the control algorithms as the result of neural network training and other functional features of intelligent systems do not contribute to a guaranteed evaluation of their FS status. The use of reports of the adopted quality management and FS measures significantly enhance the informational description of the system and corroborates the confidence in the assessment of its FS state.

Confirmed compliance with the specified FS requirements plays a crucial role in the system FS case document. To that end, the following methods are used: statistical, experimental, expert, simulation, analytical, process.

The statistical and experimental methods enable the most objective, quantitative evaluation of a system's FS as long as their feasibility and reliability are ascertained. The matter of feasibility directly depends on the FS requirements. The required safety integrity level of an ATO system with continuous performance requests is typically SIL 2 [6, 12], which corresponds to the required range of a system's dangerous failure rates $\lambda_{WS} = (10^{-7}10^{-6})/h$. The probability of the system's dangerous failures within an hour of operation should be within the range $Q_{WS}(1) = 10^{-7}10^{-6}$ [6]. Under the above requirements, an experimental identification of a single dangerous failure would require at least $N \geq \frac{1}{Q(1)} = (10^{7}10^{6})$ tests, taking into account a statistical confidence level of 90% this will be even 3 106…3 107 hours.. As the duration of each test should be at least one hour, identifying a single

dangerous failure would take over 100 years. Even if the testing is carried out on many systems in parallel, it is complicated to accumulate a significant testing time is needed,

In principle, *experimental methods* allow indirectly confirming or disproving an ATO's compliance with the specified FS requirements. Naturally, an indirect estimate can only be used as an addition to other estimates rather than individually. The process of expert evaluation of complex system parameters is mature, as is the algebra of processing of experts' opinions. However, applying such methods to the ATO FS estimation has a number of difficulties. To begin with, the experience of ATO operation is still insignificant. The accumulated knowledge is clearly insufficient. Subsequently, it is difficult to presume an acceptable level of subject-area competence in the experts. Additionally, in various industries, including railway transportation, the number of FS experts is limited. Therefore, it is very difficult to involve a sufficient number of experts and evaluate the coherence of their opinions. However, we must strive for a situation, whereas expert methods can, to a certain extent, be used for confirming ATO compliance.

Simulation methods are widely used in the course of development and testing. They are based on the Monte Carlo method. The Monte Carlo simulation method allows using pseudorandom number generators to simulate practically the entire known spectrum of input, intermediate, and disturbance effects on a system. They are processed using software simulation of the system to generate outputs depending on the simulated data. However, that method has a serious drawback, i.e., the results contain a spread between the outputs of different simulations. Reducing the spread, i.e., reducing the dispersion, requires a large number of executions of the model, which, in turn, causes a sharp increase in the duration of the simulation. A number of methods of reduction of dispersion has been developed for the purpose of cutting the simulation time. Those include the following: Monte Carlo simulation (e.g., data and output value simulation), method of augemented variables, stratified sampling method, etc. Weighted sampling provides better results in terms of dispersion reduction. Drawing from that method, we have developed a simulation method based on semi-field testing [13] by means of artificial introduction of malfunctions (faults, perturbations, program errors) into the system. Despite the obvious advancements in simulation, those methods have a number of significant drawbacks that restrict their applicability in ATO research.

The main factors that restrict the application of simulation in ATO research are as follows:

1. A detailed description of the system and its features is required, which, for a system as complex as an ATO, requires significant efforts and associated large scope of work. Additionally, due to the complex system architecture, a clear description of such system is extremely complicated.
2. The high cost of developing a simulation model for the system.
3. Evidence of adequacy of the model to the actual system is required.
4. Each update of the system's structure and improvement of its algorithms require to repeat the activities specified above in Items 1 and 3. Practically, that comes down to the development of new simulation models.

Analytical methods are the main tool for safety case preparation. However, their applicability to ATO safety justification raises certain doubts. That is primarily due to the distributed architecture of such systems and, subsequently, the difficulty (or sometimes impossibility) to formalise the task of safety justification. In order to solve that problem, we propose the following. ***Heuristic semi-Markov (Markov) graph methods***. The matter is that non-formalised problems of safety justification of systems with complex architectures are solved using heuristics, i.e., *a person's own ideas, rules that allow reducing the scope of potential solutions.* The essence of the developed methods [4, 12] consists in a combination of heuristics in the data representations and mathematical models of the system's safety and dependability with strict mathematical methods of analysis.

Under uncertainty or absence of certain data, an analytical estimation of system safety indicators is achieved through multi-stage calculations that consist in the implementation of the following sequence of actions:

1. Construction of the Markov graph of the ATO.
2. Definition of the mathematical models of the graph's edges and nodes.
3. Definition of equations for FS.
4. Expert evaluation of initial data.
5. Calculation, analysis of the results.
6. Determination of the most significant factors.
7. Simplification of the obtained calculation formulas maintaining an acceptable error.
8. Analytical evaluation of compliance with the required SIL.
9. Finalisation of the procedure in case of confidence in the results of evaluation or improvement of the examined ATO functional model (FS graph). If necessary, the model will be refined and steps 1-9 of the analytical estimation of safety parameters for above for the updated model will be repeated.

If reliable information and data are available, individual actions will suffice, e.g., graph construction, definition of formulas, calculation and analysis of the results. Other actions, e.g., expert evaluation of the initial data, identification of the most significant factors, simplification of calculation formulas, improvement of graph construction conditions, repeated construction(s) of a FS graph arise as needed depending on the availability or non-availability of information to the system's safety analyst.

Due to the above distinctive features of an ATO and in order to improve the confidence in the FS examination results along the recommendations of EN 50129 [25] *the process methods of FS compliance assurance* should be widely used.

Evaluating the achieved SIL for each safety function of an ATO's hardware components is possible based on the recommendations of EN 50129 [25] chapter 7..

The applied methods and means of failure management are evaluated based on the recommendations of EN 50129 [25] annex B..

The applied methods and means for preventing systematic errors can be evaluated based on annexe E of EN 50129 [25].

Regarding the software of an ATO, EN 50128 [26] recommends a number of procedures (annexe A) whose application significantly improves the confidence in the FS state estimate.

3. Methods of ensuring safe and uninterrupted operation of a shunting locomotive control system with machine vision

3.1. Introduction

Railway signalling systems are undergoing a new stage of their development in order to solve one of the key problems in railway transportation that consist in the creation of unattended train operation. Along with the conventional means of functional safety, they include sufficiently complex ATO systems [16]. Now, using control algorithms based on logical and certain arithmetical operations is insufficient to ensure safe rain operation. The technological development of control systems is associated with solving complex mathematical problems and eventually with the use of neural networks for information processing.

The simplicity of process-related tasks for the first safety integrity level allowed using mature methods to ensure compliance with functional safety requirements by means of hardware and software redundancy [6]. A clear advantage of using simpler technology in the form of hardwired logic and microcontrollers was the simplicity of built-in online testing and, subsequently, to achieve the required rate of dangerous failures [17].

In the process of automatic train operation system development, it became clear that their rate of dangerous failures will not be below the SIL2 threshold. The application of prototypes of such systems in railway transportation has shown that, in principle, they are man-machine systems, in which automatic train operation facilities cannot be fully trusted with ensuring train protection without an operator's involvement.

Systems use to ensure operational safety in stations as part of shunting operations require a SIL 2. That is due to the fact that the speed of train movement in the course of shunting operations is significantly lower than in the course of mainline operations [18]. Meanwhile, the demanding work performed by a driver in the course of shunting operations should be taken into consideration. When and where possible such operations should be automated. Thus, even if a driver is present onboard, the future requirements must be close to SIL3 or a new, more detailed SIL 2+ classification of safety is to be introduced. For information processing facilities, this level can be achieved with the help of a real-time operating system and high-performance microprocessors. In this context, the matters of validity of information processing and completeness of online tests arise. The tendency for using complex computer-based systems, on the one hand, and the expectation of their high redundancy, on the other hand, complicate such control. Indeed, within the information processing circuit, a small amount of memory and limited number of commands are used. In this context, a high test coverage cannot be guaranteed, as many elements of the information processing structure are not utilised. That, in turn, causes limitations in the assurance of an acceptable level of correct detection of failures of the automatic shunting cab signalling system (ASCSS).

3.2. Problem definition

Currently, the ASCSS shunting locomotive control system is single-channel, system, which does not allow to raise its safety integrity level above SIL 2. By using information redundancy, a virtual second channel can be created to ensure additional monitoring of this computer-based system [19]. That will enable a high probability of correct detection of ASCSS failures. The monitoring process is to be designed in a way as to not affect the operation of the control algorithm of a shunting locomotive. The safety device (SD) software generates an ordered sequence of computer instructions that, within the ASCSS system, are implemented as a series of reference signatures, which allows additionally monitoring of the operation of complex ASCSS devices, thus enhancing its SIL.

By building upon that principle, such SIL2 and SIL3 devices can be used for monitoring even more complex devices with video cameras and neural networks. It must be noted that such a complex device itself, e.g., ASCSS, implements functions of the type *"prevention passing shunting signal at danger"* that align with the typical purpose of ASCSS and additionally detect obstacles using machine vision. After identifying the basic technical function of information processing, it can be used as a functional test for the equipment of a complex unattended system. That can be seen as a segment within the space of possible solutions of unattended systems.

Thus, the shunting signals themselves become a functional test of an even more complex system [20]. Additionally, within the examined system, the hardware and software machine vision facilities are to be additionally monitored by comparing the readings of onboard and trackside machine vision sensors [21]. Such a hierarchy may prove to be useful in reducing the cost of hardware and simplifying the safety case preparation as compared with the situation when all functions are implemented within a single processor [22]. It must be noted that, if no innovative solutions are used, ensuring system dependability becomes an issue, as machine vision facilities significantly increase the scope of system hardware, which causes a reduction of its dependability.

3.3. Research model and findings

Any information that can be depicted as objects and connections can be conveniently represented in graph form. Graphs are commonly used for visualising information, involving the transformation of large amounts of complex types of abstract information into a user-friendly visual form.

The authors built the model based on the following criteria:

a safe failure involving the failure of ASCSS and MV facilities, control of the locomotive is assigned to the driver; *a dangerous failure* involving the failure of ASCSS, MV facilities and SD, the shunting locomotive is brought to a complete stop. The question of the criticality of a dangerous failure is not discussed in this paper. It should be examined individually.

Figure 1 shows the state graph of functional safety of interaction between ASCSS and the SD and MV facilities.

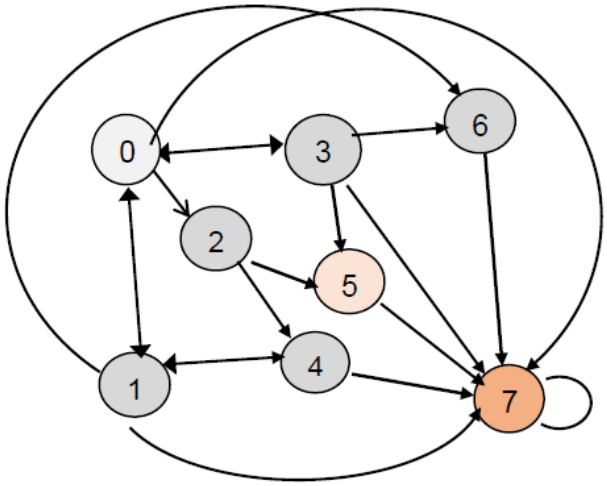

Figure 1. FS state graph of the interaction between ASCSS and the safety device and machine vision facilities

States of the system model:

1. All objects of the control system are up;
2. SD has failed and is recovering; all the other system facilities are up;
3. ASCSS has failed, all the other system facilities are up;
4. MV facilities have failed and are recovering; all the other system facilities are up;
5. MALS and SD have failed; ASCSS is recovering;
6. MV and ASCSS have failed, locomotive control is assigned to the driver (safe failure);
7. MV and SD have failed;
8. All three systems have failed.

Dangerous failure

System safe states are marked with the following colours:

SIL 3, safe failure (SIL1)

SIL 2, dangerous failure

In the system model, the following transitions are provided for:

0-1, SD failure;

0-2, ASCSS failure detected using built-in tests and/or signature analysis;

0-3, MV failure detected using built-in tests and/or by comparing the readings with the ASCSS program;

0-5, undetected MV failure;

0-7 and 1-7, undetected ASCSS failure;

1-4, ASCSS failure subject to SD failure;

1-6, MV failure subject to SD failure;

1-0, SD repair;

2-4, SD failure subject to ASCSS failure;

2-5, MV failure subject to ASCSS failure;

3-5, ASCSS failure subject to MV failure;

3-6, SD failure subject to ASCSS failure;

3-0, MV repair;

4-7, MV failure subject to MALS and SD failure;

4-1, ASCSS repair;

5-7, SD failure subject to MALS and MV failure;

6-7, MALS failure subject to MV and SD failure;

7-0, transition into the original state as the result of possible modification of ASCSS, if the risk of dangerous failures is acceptable.

The adopted premises and assumptions, defined mathematical models of the graph's edges and

nodes, the FS formulae, as well as the expert evaluation of the initial data are set forth by us in [23].

That model allows examining the FS of an ATO layer by layer, i.e., evaluating the time to safety integrity level degradation depending on failures of components and onset of a complete dangerous failure (dangerous failure of the second type in Fig.1). Thus, in particular, in [23] it was established that

- *the mean time of the system being in SIL3* (state 0 in the graph in Fig.1) is described with the formula

$$T_0 = \frac{1}{\lambda_{SD} + \lambda_{\text{M}} + \lambda_{MV}} \tag{1}$$

where λSD, λM, λMV are the failure rates of the safety device, locomotive control system and machine vision, respectively;

- *the mean time of faultless system operation at a level at least as high as SIL2, the mean time to a safe failure of type 1*

$$T_{\geq SIL2}^{system} \approx \frac{\lambda_{SD}(2\lambda_{SD} + 3\lambda_{MV}) + \lambda^2{}_{MV}}{(\lambda \quad + \lambda \quad)\lambda \quad \lambda_{\text{M}}} \tag{2}$$

- *the mean time of faultless system operation to a dangerous failure (complete stop of the shunting locomotiv)*

$$T_{wrong-side} \approx \frac{\lambda_{SD}(2\lambda_{SD} + 3\lambda_{MV}) + \lambda^2{}_{MV} + 2\lambda_{\text{M}}\lambda_{MV}}{(\lambda \quad + \lambda \quad)\lambda \quad \lambda_{\text{M}}} \tag{3}$$

Formulae (2) and (3) were obtained with an error not exceeding the first order of magnitude assuming that the failure detection parameters of ASCSS and machine vision facilities are close to one. That assumption is based on the fact that the monitoring of ASCSS operation using additional signature analysis procedures, as well as regular comparison of the ASCSS outputs with the machine vision outputs ensure complete and reliable performance monitoring of both the ASCSS control system, and the machine vision facilities.

Figure 2 shows the time of faultless operation of the shunting locomotive to a complete stop-vs-the failure rate of machine vision and ASCSS equipment curve. The failure rate of safety devices is taken equal to λ_{SD}=1*10^{-8}, which corresponds to the safety integrity rate of SIL3.

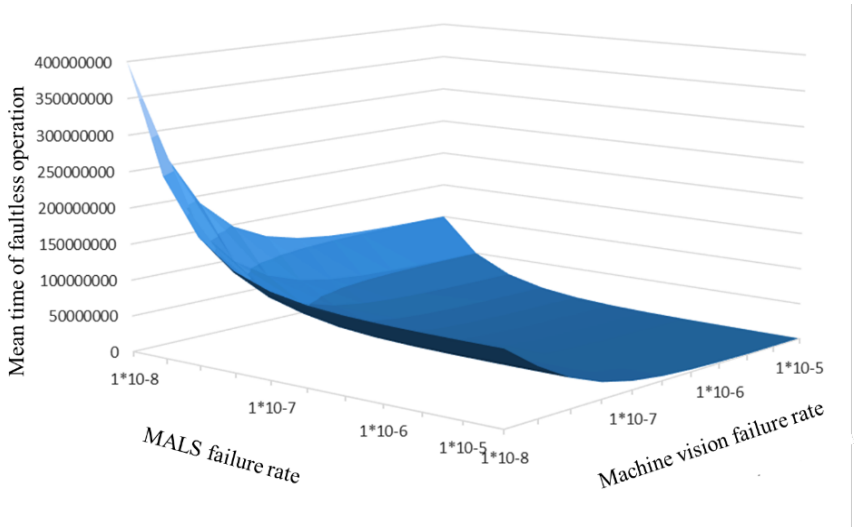

Figure 2. Time of faultless shunting locomotive operation to a complete stop-vs-the failure rate of machine vision and ASCSS equipment curve ("MALS failure rate")

The key task of the research consists in evaluating the level of functional safety of the automatic train operation system of a shunting locomotive. Such a comprehensive assessment can be enabled by a research of the system's safety coefficient. The probability of an opposite event, i.e. a dangerous failure, is the system's danger coefficient. That coefficient, under the same assumptions that were used for deducing formulas (2) and (3), was obtained in [23] with an error not exceeding the first order of magnitude. It was established that the hazard coefficient significantly depends on the repair rates of facilities μ and ASCSS repair rate upon a hazardous failure μ_1.

The three-dimensional graphs of a system's hazard coefficient against parameters μ and μ_1 subject to $\lambda_{\text{M}} = 10^{-5}\,1/h$ and $\lambda_{MV} = 10^{-5}\,1/h$ are shown in Fig. 3.

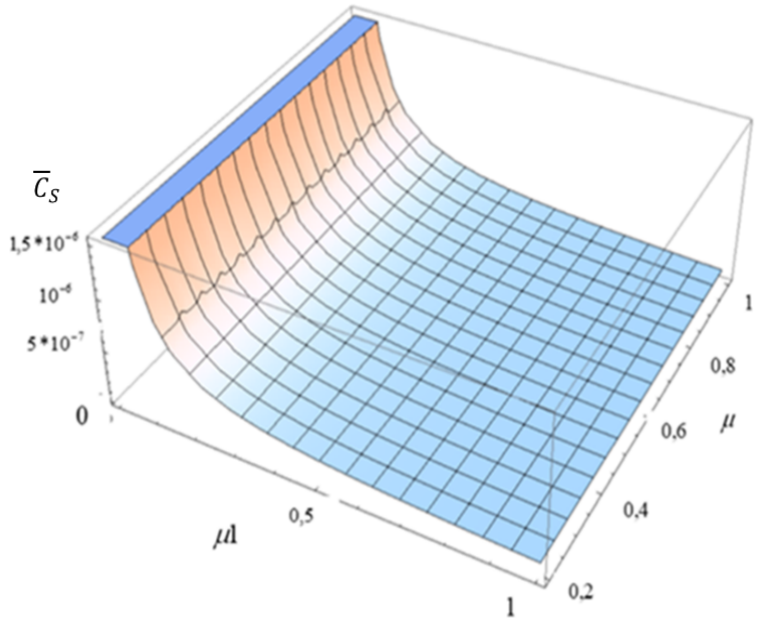

Figure 3. Graphs of hazard coefficient $\overline{C}_S$ against repair rates μ_1 and μ.

From these graphs it follows that as the system's repair rate μ_1 increases from 0.0059 /h to 1 /h the hazard coefficient decreases almost 30 times. The chosen limit values of the repair rate correspond to the system's repair times from an hour to a week. That range was chosen based on the nature of the malfunction. Thus, if a set of spare parts is available, hardware failures can be rectified within an hour, while rectifying software errors may take up to 7 days. Therefore, timely and prompt rectification of malfunctions may significantly improve a system's safety indicators.

4. A general approach to designing the functional safety of automatic train operation systems

4.1. Methods for designing the functional safety of automatic train operation systems

The main problem in the development of that approach consists in the fact that such a system has many distinctive features associated with the complex architecture and information processing algorithms, the incompleteness and fuzziness of initial data. Therefore, it is difficult to apply classical methods of probabilistic evaluation in the form of two or more independent hardware and software

information processors [20]. That is the exact reason why the redundancy of such information processors in the form of onboard machine vision cameras for the purpose of safe detection of obstacles on the track is unlikely to achieve the required safety level due to the unknown testing time of such a learning, i.e., constantly changing vital information processing system.

Braband and Schäbe [2] assumed the mandatory presence, , of an additional device within the processing system, whose safety could be proven using conventional methods owing to its constant structure.

Rozenberg and Shubinsky [12, 15] suggested using the so-called multi-level structures in order to ensure FS. This approach has shown good results in the development of advanced onboard and trackside safety systems. Additionally, an extremely important property of a system's safety evaluation was used that consists in obtaining reliable information on a facility's state history in terms of safety.

As regards the safety cases of neural network-enabled automatic train operation systems the principles of multi-level safety should be used. The difference consists in the fact that a complete set of technical equipment within a locomotive's operating environment is to be examined rather than an individual smart device, e.g., a machine vision camera on such a locomotive.

Indeed, a camera with a predesigned program for processing information on obstacles on the track does not depend only on the previously taken neural network training measures, but on the specific factors that affect the operability of the camera's hardware, software faults, etc. Additionally, it should be noted that the effect of the environment, i.e., snow, fog, and rain influences the obstacle detection zone, which directly affects the safety, as it is associated with the braking distance.

Under such conditions, the situation ahead of the train is additionally monitored from a special control centre, where an operating driver monitors several locomotives [21].

The complexity of this method consists in the fact that the reaction of the operating driver becomes a critical component, while he/she depends on the stable onboard camera image and dependability of the broadband communication at a particular location.

On the other hand, the division of the information processing process into two subprocesses, i.e., internal intelligent information processing onboard the locomotive for the purpose of decision-making regarding track vacancy and communication of initial visual information to the operating driver for decision-making allows improving safety. The criterion in this case is that the onboard system should have a high probability of false alarm, while the operating driver can rectify this situation using a special command transmitted to the locomotive by radio. In practice, if this principle was not used, an ATO system would stop, for instance, because of a plastic bag on the track.

It should be noted that the system includes trackside devices that monitor track vacancy in places with poor visibility [20]. Information on such fixed systems is communicated to the locomotive in real time, which significantly improves train traffic safety. Thus, the used model is simplified, but it enables an analytical study of the problem. That constitutes the advantage of this approach to developing the research model over more complex models. An interesting feature of the interaction between the fixed and onboard machine vision facilities is that, under identical environmental conditions, they can see the same objects, within the line of sight or under various, interesting angles.

The availability of objects detected by two independent systems allows using this property for cross-comparison of intelligent technical facilities, especially for the purpose of making correct decisions

by intelligent onboard systems that operate in more difficult operating conditions (traffic speed, limited visibility zone, etc.). An object comparison can be in the form of images processed by fixed and onboard cameras or it can contain the expected inversion of an image of the same object if two machine vision cameras point at it from opposite directions. Such a predefined property for a system for safe comparison of results enables better independence of information processing. Each technical facility, including video cameras, contains self-testing features that necessarily contribute to the calculation of their safe operation. Given that, as regards an intelligent system that employs neural networks, it is difficult to talk about complete testing, self-diagnostics using observed objects known beforehand should be used. For instance, next to the railway track, within the scanning zone of machine vision cameras or lidars, there are signals, control cabinets, catenary masts, and communications posts that are strictly referenced to the track coordinates, which is even more relevant if a 3D map of the infrastructure facilities is used onboard.

Thus, capturing such objects allows testing onboard cameras and sensors taking into account the detection distance and identification of the type of objects. If the frequency of object acquisition is high enough within the distance between such locations, the probability of no failure or no error of the information processing algorithm can be calculated for a moving object. The advantage of this method consists in the complete processing of information, when, along internal testing of hardware components, the required level of system safety can be achieved. In that case, the system itself appears to be a "black box", but with perfectly known outputs at an absolutely known spatial coordinate.

4.2. Conceptual safety model of an automatic train operation system

An ATO system includes the following key facilities:

- onboard train control and protection equipment;
- monitoring centre equipment;
- trackside machine vision facilities;
- onboard machine vision facilities.

The conceptual safety model of an automatic train operation system contains a description of the dependability and safety states of the system's component facilities, their interrelations, as well as the effects of disturbing weather effects. This model is presented in the form of a system safety state graph (Fig. 4).

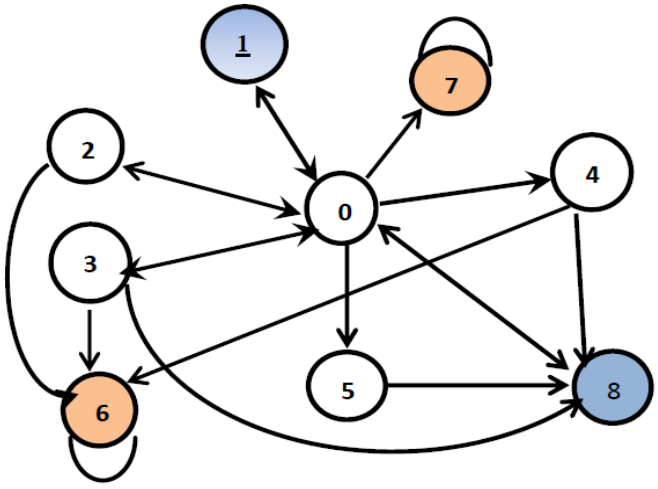

Figure 4. Safety state graph of an automatic train operation system

While building the system safety model, the following *criterion of dangerous failure* was adopted: failure of all machine vision facilities and monitoring centre or an undetected failure of a locomotive control and protection system. *Criterion of safe failure*: failure of fixed machine vision facilities, monitoring centre and effect of disturbing weather conditions or detected failure of the locomotive control and protection system.

Graph states:

0, up state, no disturbing weather effects;

1, detected locomotive control and protection system failure, *a safe failure*;
2, failure of monitoring centre facilities;
3, failure of trackside machine vision facilities;
4, failure of onboard machine vision facilities;
5, disturbing weather effects;
6, failure of all machine vision facilities and monitoring centre, *a dangerous failure* of the automatic train operation system;
7, non-detected locomotive control and protection system failure, *a dangerous failure*;
8, failure of trcakside machine vision facilities, monitoring centre and disturbing weather effects, ***a safe failure.***

The entire set of system states according to the state graph in Fig. 4 is divided into the following subsets: - the subset of up states $S_U = \{0,2,3,4,5\}$; - the subset of safety states $S_S = \{1,8\}$; - the subset of hazardous states $S_H = \{6,7\}$.

The up and safe states form the set of good states.

Given below are the model's good state transitions that need clarification: 1-0, 2-0, 3-0, 8-0, repair of facilities after failures; 3-8, monitoring centre failure subject to trackside machine vision facilities failure; 4-8, monitoring centre failure subject to onboard machine vision facilities failure; 7-8, failure of trackside machine vision facilities subject to disturbing weather conditions.

The mathematical formulation of the model takes into account the following considerations. The system is new and unique, no statistical information about it is available. Therefore, the distribution functions of system parameters are not established. Based on the existing experience in railway control and management systems, it can be safely assumed that failures of electronic devices, such as devices of a train control and protection system, monitoring centre facilities, and machine vision facilities are exponentially distributed. This assumption does not apply to random values of time to device repair estoration after failures, much less to random adverse weather effects. The problem of disturbing effects was theoretically examined by Schäbe and Viertl in [23]. Those models are also applicable to disturbing weather effects. In order to ensure adequate results, the authors were forced to use a complex mathematical description of the random process of adverse effects on the locomotive's control system. The above circumstances complicate their practical application in mathematical simulation of the safety of the automatic train operation system.

In the absence of practical information, it is very difficult to predict the quantitative safety indicators of the automatic train operation system. In this paper, in the context of great uncertainty, we aim to identify the most significant factors affecting the system's safety. The assumption of the Poisson process of random events in the automatic train operation system fits this purpose. The Poisson processes are ordinary, stationary and have no aftereffect. On the one hand, due to the significant uncertainty in the initial conditions, their application does not contribute to accurate prediction of the safety of a system's behaviour characteristics. On the other hand, the obtained outputs can be regarded as conservative bounds for constructing a safe ATO system by neutralising the identified

most significant adverse factors. Thus, the used model is simplified, but it enables an analytical study of the problem. That constitutes the advantage of this approach over more complex models.

The adopted assumptions defined mathematical models of the graph's edges and nodes, the FS formulas, as well as the expert evaluation of the initial data are provided by us in [25].

The limit value of an automatic train operation system's time to dangerous failure takes place subject to the absence of destructive disturbing weather conditions ($\gamma \to 0$) and compliance with the requirements of IEC 61508-2 [6] ($\alpha \to 0$).

Under those conditions, the probability of dangerous failure with an error not exceeding the first order of smallness tends to the following form:

$$G_{WS}(t) \cong \lambda_{WS} \cdot t \to \frac{\lambda}{2} t,$$

where λ is the failure rate of the machine vision facilities (it is assumed that the onboard and fixed facilities have about the same dependability).

5. Conclusion

Ensuring the FS of an automatic train operation requires not only developing or applying known methods of designing a safe system, but, most importantly, proving the acceptability of the achieved level of FS. In respect to automatic train operation systems, the conventional methods of proving the FS (statistical, experimental, expert, simulation) are of limited use. That is due to the distributed architecture of the systems, changing information processing algorithms in the course of training, large number of vulnerabilities, etc. For the purpose of improving the confidence in the FS evaluation results, it is proposed to focus on the technological methods and use the wideky applied analytical expert semi-Markov method, proposed here.

The proposed process of monitoring the operation of ASCSS and machine vision facilities, creation of a second, virtual channel allow improving the FS of the shunting locomotive control system with machine vision from SIL 2 to SIL 3 and maintaining it over a sufficiently long period of time (over a quarter of the mean time to failure of the ASCSS). The mean time of faultless operation of the shunting locomotive control system may grow almost three times as long as the achieved level of the system's FS remains unchanged. Additionally, the time of faultless operation of the locomotive until it has to be brought to a complete stop for safety reasons can also increase over three times. This important result can be practically achieved despite the increased amount of the system's equipment due to the introduction of machine vision facilities.

A general approach to ensuring the FS of an ATO is proposed. It is based on the division of the information processing process into two subprocesses, i.e., internal intelligent information processing onboard the locomotive for the purpose of decision-making regarding track vacancy and communication of initial visual information to the locomotive driver for decision-making. The division of this process must be combined with redundant machine vision facilities, regular comparison of the outputs of the onboard and fixed machine vision facilities, redundant comparison outputs, smoothing of the outputs in the process of locomotive movement. The EN 50129 functional safety requirements for the locomotive control and protection system and SIL4 requirements for the machine vision facilities are to be fulfilled as well. Additionally, adverse weather effects are to be countered by improving the efficiency of machine learning of the machine vision software.

References

1. Vapnik V.N., Chervonenkis A.Ya. Pattern recognition theory (statistical problems of training). Moscow: Nauka; 1974. (in Russ.)

2. Braband J., Schäbe H. On safety assessment of artificial intelligence. Dependability 2020;4:25-34.

3. I.B. Shubinsky, E.N. Rozenberg, H. Schabe INNOVATIVE METHODS OF ENSURING THE FUNCTIONAL SAFETY OF TRAIN CONTROL SYSTEMS. Reliability: Theory & Applications. 2023, December 4(76): 909-920, DOI: https://doi.org/10.24412/1932-2321-2023-476-909-920

4. Shubinsky I.B., Rozenberg E.N. General provisions of the substantiation of functional safety of intelligent systems in railway transportation. Dependability 2023;23(3):38-45.

5. STO RZD 1-19.009-2009. Railway signalling devices and systems. Safety case.

6. IEC 61508-1:2012. Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems. Part 1. General requirements. Moscow: Standartinform; 2014.

7. IEC 61508-2:2012. Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems. Part 2. System requirements. Moscow: Standartinform; 2014.

8. IEC 61508-3:2018. Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems. Part 3. Software requirements. Moscow: Standartinform; 2018.

9. IEC 62279-2016. Railway applications. Communication, signalling and processing systems. Software for railway control and protection systems. Moscow: Standartinform; 2017.

10. GOST 33432-2015. Functional safety. Policy and programme of safety provision. Safety proof of the railway objects. Moscow: Standartinform; 2019. (in Russ.)

11. Braband J., Schäbe H. Risk analysis for automated driving – validation and findings. Signal+ Draht 2023;115(4).

12. Shubinsky I.B., Rozenberg E.N. Functional safety of railway control systems. Vologda: Infra-Inzheneria; 2023. (in Russ.)

13. Shubinsky I.B. Dependable, fault-tolerant information systems. Methods of synthesis. Moscow: Dependability; 2016. (in Russ.)

14. Shubinsky I.B. Structural dependability of information systems. Methods of analysis. Moscow: Dependability; 2012. (in Russ.)

15. Rozenberg E.N. Multi-level train control and protection system: Doctoral Thesis (Doctor of Technical Sciences): 05.13.06, 05.22.08. Moscow; 2004. (in Russ.)

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

17. Shvir V. Dependability of electronic circuits in railway signalling devices. Rail International 1986;1:59-67. (in Russ.)

18. Kalinin A.V. Driverless shunting locomotives control. Main principles and prospects of technology development. Intellektualnye IT oupravleniya ITNOU 2017;1:12-14. (in Russ.)

19. Shubinsky I.B., Rozenberg E.N., Korovin A.S., Penkova N.G. On a method for ensuring functional safety of a system with single-channel information processing. Dependability 2022;22(3):44-52.

20. Sapozhnikov V.V., Sapozhnikov Vl.V., Khristov Kh.A., Gavzov D.V. Methods for constructing safe computer-based railway signalling systems. Moscow: Transport; 1995. (in Russ.)

21. Mylnikov P.D., Okhotnikiv A.P., Popov P.A. Onboard information system. Pat. no. 2742960 dated 12.02.2021. Bul. no. 5 N. (in Russ.)

22. Shubinsky I.B. Functional dependability of information systems. Methods of analysis. Moscow: Dependability; 2012. (in Russ.)

23. Shubinsky I.B., Rozenberg E.N., Panfiorov I.A., Boyarinova N.A., Kuzmin A.I. Estimating the safety and reliability of the control system of a locomotive with machine vision. Dependability 2023;23(1):30-37. (in Russ.)

24. Schäbe H., Viertl R. An axiomatic approach to models of accelerated life testing. Eng. Fract. Mechanics 1995;50(2):203-217.

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