# Safety model construction for a complex automatic transportation system

**Alexey V. Ozerov[1]*, Alexey M. Olshansky[1]**
[1]*JSC NIIAS, Moscow, Russian federation*
*a.ozerov@vniias.ru

*Alexey V. Ozerov*

*Alexey M. Olshansky*

**Abstract.** *The* **Aim** *of the paper is to consider approaches to the analysis of a safety model of complex multi-loop transportation systems comprising not completely supervised subsystems.* **Method.** *For the description of a safety model, the paper uses systems theoretic process analysis (STPA) methods and the principles specified in ISO/PAS 21448:2019 (SOTIF).* **Result.** *The paper shows drawbacks of the FTA and FMEA local risk analysis methods and demonstrates a demand for some universal approach based on the combination of STPA and control theory. It gives an overview of the major stages of such analysis for the safety model of complex transportation systems exemplified by the Moscow Central Circle, which provide a feedback for safety evaluation of a transport control system under development. The paper analyzes the feasibility of using a virtual model for control purposes in the form of a so-called "supervised artificial neural network".* **Conclusion.** *Today, railways are actively testing autonomous systems (with no driver onboard) that apply as their subsystems automatic perception modules using machine learning. The introduction of the latter into the control loop complicates the task of hazard analysis and safety evaluation of such systems using conventional FTA and FMEA methods. The construction of a safety model of such complex multi-loop transportation systems comprising not completely supervised subsystems that use machine learning methods with not completely predictable behavior requires the application of a systems approach to the analysis of unsafe scenarios along with the compilation of a scenario library and the formalization of a hazard model's description, pertaining to the boundaries of various control loops as well, in order to reduce the regions of unknown unsafe scenarios for autonomous transportation systems under development.*

**Keywords:** *railway transport, autonomy, safety model, STPA, machine learning, artificial neural network (ANN).*

**For citation:** *Ozerov A.V., Olshansky A.M. Safety model construction for a complex automatic transportation system. Dependability 2021; 2: 31-37. https://doi.org/10.21683/1729-2646-2021-21-1-31-37*

**Received on:** *23.03.2021* / **Upon revision:** *12.05.2021* / **For printing:** *21.06.2021*

## 1. Introduction

Today, many countries, including Russia, are testing automatic solutions in passenger rail transportation that aim for autonomy. Currently, full automation of passenger train control (with no driver or personnel onboard trains) has been only achieved for subways. According to UITP [1], 64 metro lines in 42 cities of the world operate in that mode.

The IEC 26690:2014 standard [2] specifies general requirements for an automatic control system for urban rail transport and proposes the following Grades of Automation (GoA) of systems (Fig. 1):

It is obvious that along with the increase of the GoA and shift to full automation of control, there appear additional safety risks that require evaluation and consideration in the process of developing the functional safety concept of this complex control system comprising a large number of subsystems.

Compared to subway systems where access to track is restricted and the boarding/disembarking process is eased up by using platform screen doors, urban railways have to resolve the issue through different means. Those include trackside and onboard perception (automatic obstacle detection) subsystems that use machine learning in decision making. Their introduction into the control loop significantly complicates the already complicated overall task of hazard analysis and safety evaluation of the multiple-loop control system associated with the safety of people. This task cannot be solved by means of the conventional FTA and FMEA hazard analysis methods only.

## 2. Problem definition

The aim of the paper is to outline a new analysis method for a safety model of complex multi-loop transportation systems comprising not completely supervised control loops, subsystems and modules. In a practical sense, this method could be used for safety evaluation of a driverless control system planned to be deployed on the Moscow Central Circle (MCC).

The key factors threatening the functional safety of a complex system may be described by the following list:

– Lost control commands or errors in transmission of external incoming information;

– Incomplete, incompatible, incorrect process model;

– Control algorithm errors (generation defect, errors of process scenario changes, problems of adaptability and trainability, inappropriate changes, errors in system state evaluation, system identification errors);

– Invalid, incorrect or missing control commands;

– Target or mechanism actions unfit for the process;

– Inadequate responses from sensors and observers;

– Invalid, incorrect or missing feedback;

– Feedback inaccurate measurements or delays;

– Delayed delivery of commands, input losses or errors;

– Component failures, unrecognized external noise/commands, their possible overlapping.

1. The basic premises for shaping a new approach to the construction of a safety model of complex transportation systems may be as follows:

2. Division into basic subsystems and error tree analysis for each subsystem does not take into account the interaction of these subsystems.

3. In a complex system, there may occur an event, when despite the constituent subsystems being operable, there may be incomplete interaction or multiple simultaneous delays due to external factors, which will cause an unintended reaction of the system in question.

| Grade of automation | Train Operation | Setting Train in Motion | Driving and Stopping the train | Door Closure | Operation in event of disruption |
|---|---|---|---|---|---|
| GoA1 | ATP with Driver | Driver | Driver | Driver | Driver |
| GoA2 | ATP and ATO with Driver | Driver / Automatic | Automatic | Driver | Driver |
| GoA3 | Driverless (DTO) | Automatic | Automatic | Attendant / Automatic | Attendant |
| GoA4 | Unattendend (UTO) | Automatic | Automatic | Automatic | Automatic |

Fig. 1. Grades of Automation (GoA) of operational modes in railway transportation

4. Complicated and time-consuming task of complete analysis of events in the system.

Insufficiency of a conventional redundant 2ooX system safety model when using ANNs in one or several subsystems. Necessity of applying additional safety measures, e.g., the implementation of a decision-making algorithm based on a digital twin. At the same time, the introduction of a digital twin (or virtual model) into a safety-critical system is an absolutely new and not yet well-proven approach to functional safety that is subject to further research (see Shubinsky et al., 2021 [3]).

## 3. STPA-based safety evaluation methodology

According to Qi Y. et al. (2020 [4]), the construction of a complex system safety model involves the development of a multi-level control system that includes the descriptions and apportionment of functional responsibilities between the system's components. The upper hierarchical level is a controller (control element) with a process model. The process model generates control commands through relations in the state space and a calculated control algorithm that is transmitted to the lower structures (target actuators). Through feedback devices, targets and other lower-level devices report about the execution of higher-level commands. The upper-level controller refers to the safety model and by comparing it with the received feedback, corrects the internal state of the model.

For such safety model, the probability of incidents comes down to situations where the internal state and feedback in the process model do not match. Such model is relevant to the functional structure of the system in question, while taking into account the relationship between subsystems as a sort of extension of multi-level control circuits.

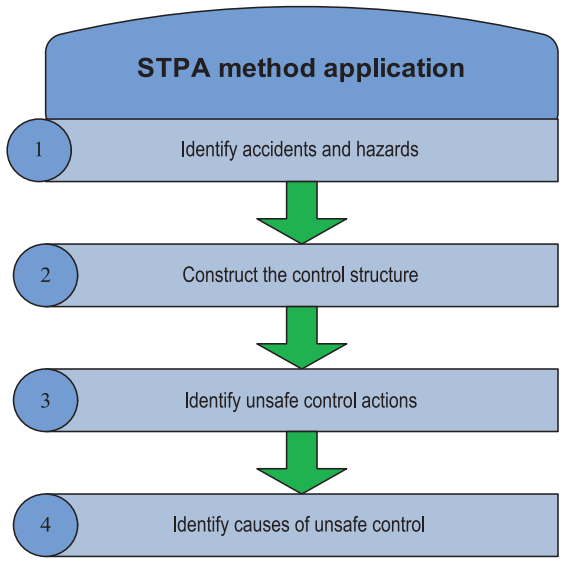


Fig. 2. STPA method application procedure

The proposed methodology is based on STPA methods assuming that we construct control and feedback circuits, target actuators, sensors and control processes and establish relationships between them that can be safety restrictions designed as systemically predefined cases (by the design and structure of such subsystems). By directly analyzing risks through an appropriate control process model, one has to evaluate safety requirements and all possible control solutions for each part of the system to identify potentially hazardous control actions and to improve the level of safety and restrictions that prevent hazardous behaviour caused by such control actions.

The STPA method (systems theoretic process analysis) appeared as a further development of the STAMP model (systems theoretic accident model and processes) proposed

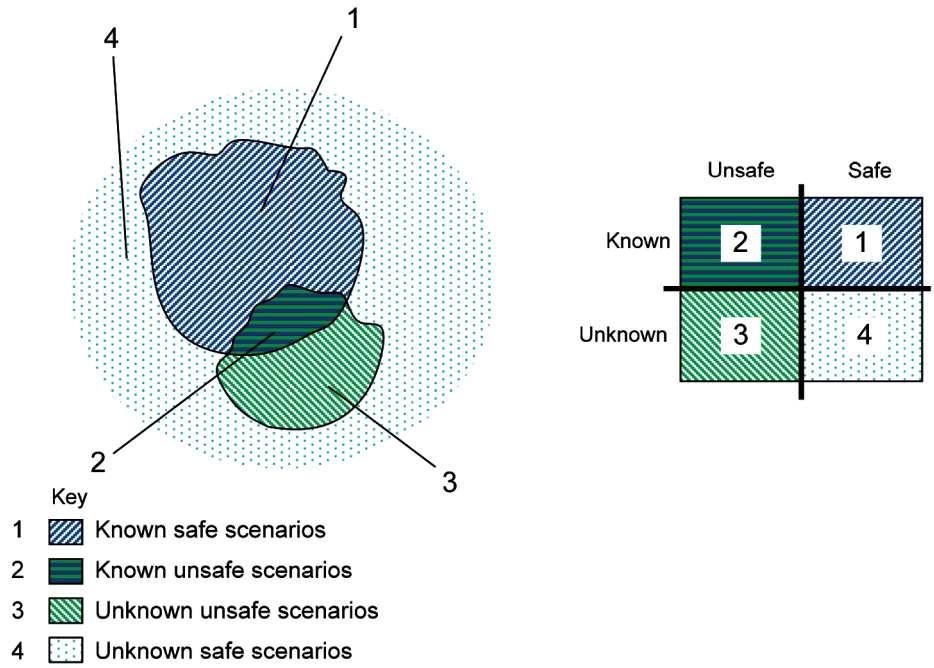


Fig. 3. Types of operational scenarios taken into account for a system's safety evaluation

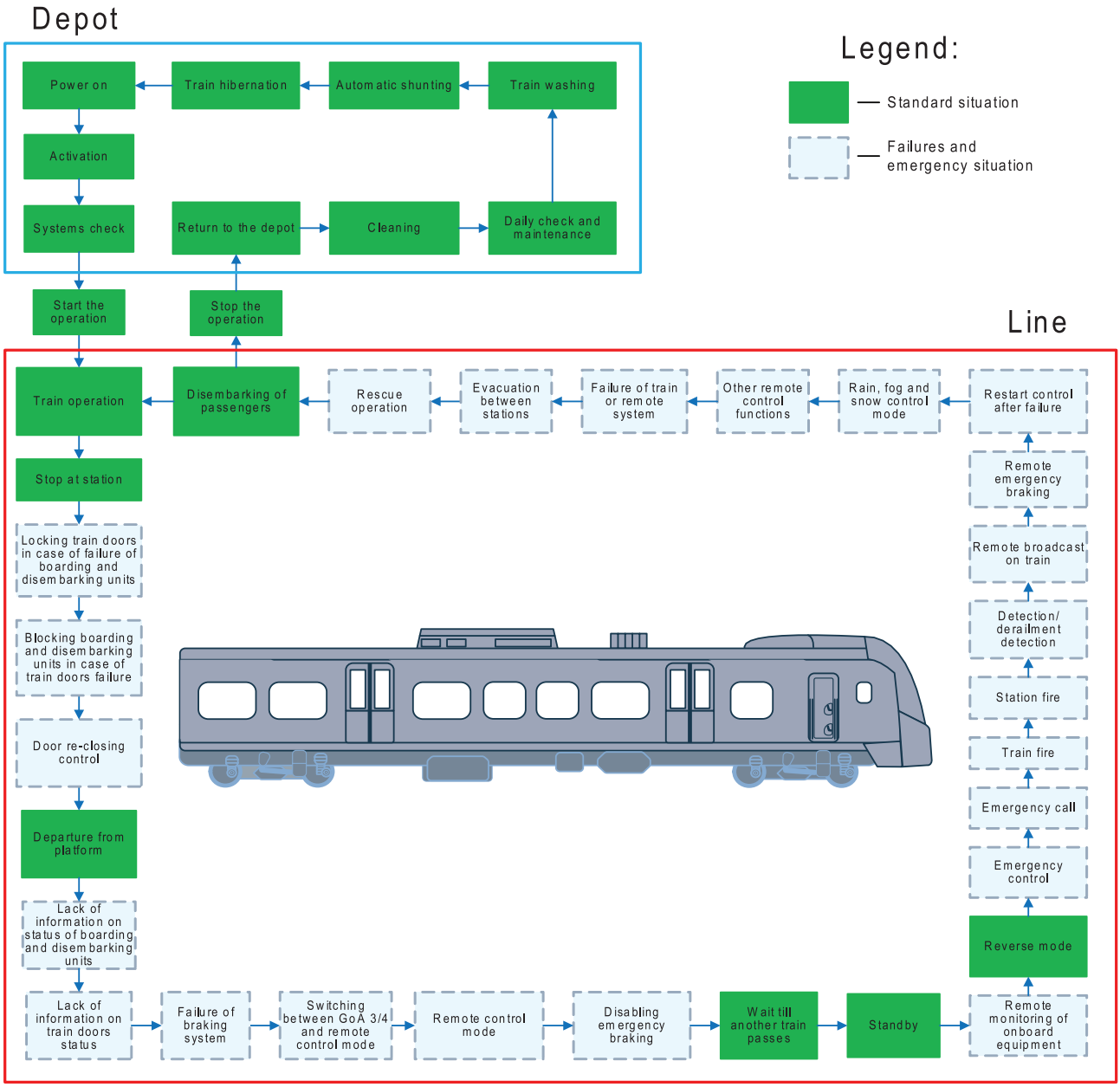


Fig. 4. Basic operational scenarios on urban railway such as the MCC

by Leveson (2004, [5]) and based on the control theory. The method is actively used in aviation, nuclear power and other industries associated with special safety requirements and complex systems. The method application procedure consists of 4 steps shown in Fig. 2 (see Chaima Bensaci et al., 2018 [6]):

Obviously, *at the first step*, we have to construct a scenario map for the entire complex system with scenario-to-scenario transition rules. Such scenarios could include all trigger events that lead to damage. In compliance with the ISO/PAS 21448:2019 (SOTIF) standard [7], one should take into account 4 scenario types presented in Fig. 3:

When constructing a safety model for a complex system, the objective is to get the maximum coverage for all scenarios and to bring the number of unsafe control scenarios to an acceptable level. As regards the MCC transportation system, we may propose a basic set of 1 and 2 type operational scenarios, which must be taken into account when constructing a safety model and compiling a general library of operational scenarios (Fig.4).

*At the second step,* it is required to construct a complete structural diagram of the control system under consideration. For instance, the MCC control system is designed as a multi-loop control system that implies two control modes, i.e., "autonomous" and remote ("remote control") (see Popov, 2020 [8]). In addition to the conventional track circuit-based train protection system, the control loop also includes radio communication between trackside and onboard train control and protection systems, as well as automatic obstacle detection by means of onboard and trackside perception modules that use ANNs and transmit relevant information to the remote control and supervision centre (RCSC). The overall

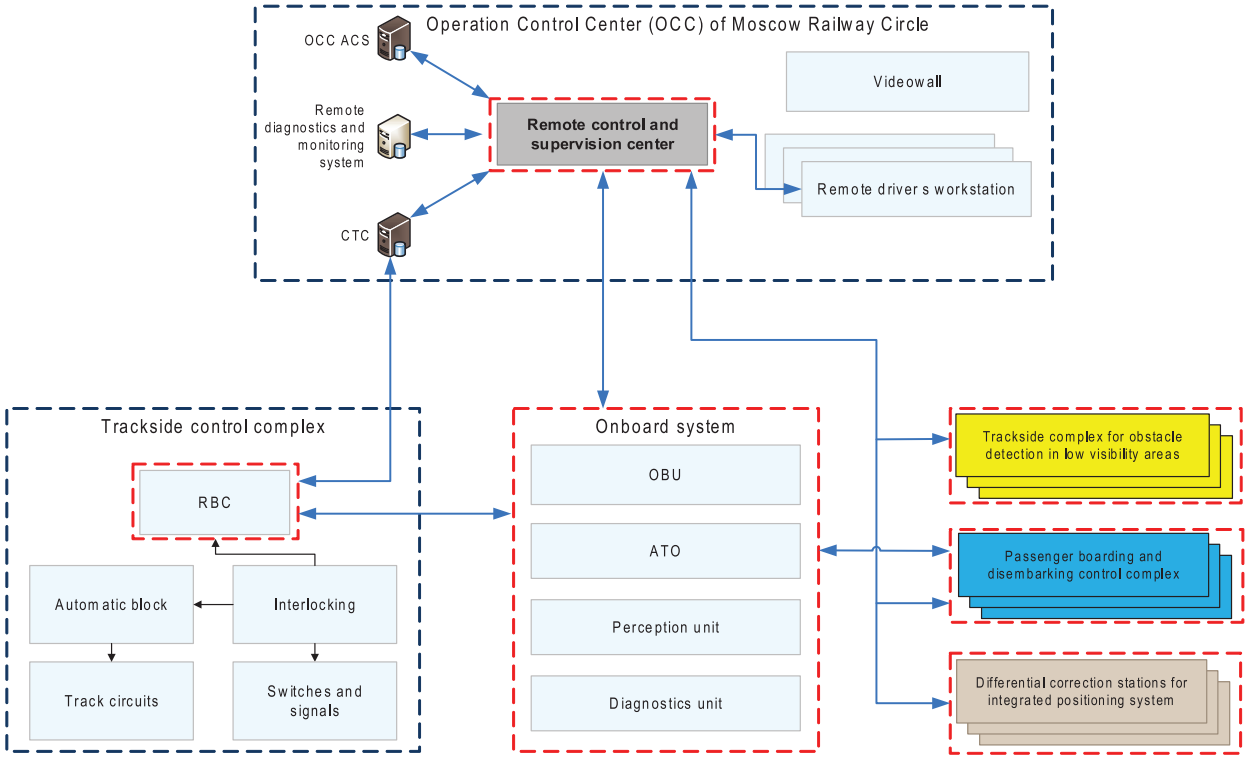


Fig. 5. The overall architecture of the MCC command and control system

architecture of the proposed MCC GoA3/4 control system is shown in Fig. 5 (red dash line indicates the subsystems making up the GoA3/4 control loop):

The presented control structure reduces the number of control layers as early as at the design stage and introduces some hierarchical order making the number of layers equal to two. The paper by Arnold (2004 [9]) clearly demonstrated that the systems with the number of layers equal to two can be sustainable provided that the upper-level control circuits are designed in a correct way. However, there should be further research and optimization of this structure to normalize relationship between control system complexes.

*The third step* of the study is the most time-consuming involving the definition and description of functional safety hazards according to the list of operational scenarios for each unit of the system at different hierarchical levels. Let us introduce the following notations for the purpose of analysis of the identified hazards: *Sc* is the total number of causal scenarios obtained through combinatorial means (which ensures 100% coverage of all devices and their combinations), *Mod* is the set of devices in the control loops that affect the functional safety of the system, *F* is the set of unsafe modes, *R* is the matrix of relationship between devices and unsafe modes, assuming that each device is incident with itself, i.e. the minimum sum of points in each device's line is 1.

In this case, subject to small modifications in terms of implementation in a particular programming language, the algorithm proposed by Yan F. et al. (2019 [10]) is applicable for the purpose of building a library of causal scenarios.

Therefore, with the introduced notation taken into account, we obtain the following sequence of actions to describe the functional safety hazards:

0. As the result of processing of a complete library of scenarios constructed according to the above syntax rules, one forms *Mod, F* sets.

1. *R* shall be constructed as a matrix (|*Mod*|, |*F*|). Note that the power of *F* set exceeds the total number of failure modes, since the same failure mode can be present in several scenarios. At the first stage, |*F*Yan F|>>|*M*| inequality shall be satisfied.

2. If *R* matrix line contains more than one entity, this means that at least one device out of *M* recorded in this line is involved in several failure modes.

3. Then identical columns shall be searched for. Their presence means that failure modes in them are the same. They can be included into one final scenario.

4. Thus, we have a library of relevant scenarios.

Such scenarios can be defined at all structural levels of the system under consideration. The general approach involves the following main stages of functional safety analysis:

1. Compilation of standard scenarios (see above), design of a hierarchical control structure, information flow diagrams.

2. Identification of hazard causes.

3. Development of safety measures.

The hierarchical control structure is a graphic representation of control layers, control commands from upper layers to lower layers and signals from lower layers, taking into account in the limit of sensors, doors, humans and microcontrollers. The selected units and devices are then described in

terms of normal and emergency behaviour as follows: "under normal conditions, N unit of X system provides (guarantees) a given property within the given range for the item."

A set of such statements as regards the elements of a hierarchical structure allows easily building a table of unsafe control actions. The table defines the description format: systems level hazards/control actions/not executed/executed incorrectly/control action is too early or too late/control action execution time is too short or too long.

The last 4 categories constitute unsafe scenarios (control actions). For each unsafe control action, the "cause – constraint" system is described, whereas the constraint describes the principles of safe behaviour in a particular situation under the selected unsafe control actions. For instance, the loop "train – Remote Control and Supervision Center (RCSC) – trackside obstacle detection system (TODS)" contains at least 2 sources of unsafe control actions, i.e., the signal from RCSC that may not arrive to the train, and TODS that may not send a request or send it too late. As the result, communication becomes a critical source of risk for the entire MCC transportation system.

A separate research will presumably be needed to cover unsafe scenarios that may take place at the boundary or at the overlap of the identified complexes "cause – constraints", or control loops. Special attention will have to be paid to the "overlap" of RCSC – TODS and "onboard perception unit (OPU) – RCSC" loops, as there is a probability of unsafe control actions from both loops when a train is in a low visibility area (TODS responsibility zone). Also, it should be kept in mind that neither loop is completely observable since both OPU and TODS use machine learning algorithms (VoVNet family convolutional ANNs), whose behaviour cannot be considered completely predictable.

It may result in a further review and change of the safety model of a transportation system under design by means of introducing an additional component in the model taking on the supervision and constraining function. As a constraining element, there are various alternatives being researched – from final state machine based on "hard" logics to supervising network. Fig. 6. shows a simplified control structure, a virtual model ("digital twin") that can be implemented as "supervised ANN".

Unfortunately, the supervisor in the form of the so-called "supervised ANN" has a delay (if an adequate solution does not appear at the second step, its search can last longer than two steps and even infinitely, till it is not stopped by a decision maker). Moreover, the supervisor's algorithms of acceptability estimation and decision-making must be fast enough in order that a total delay could be reasonable. It should be kept in mind that the confidence level *P* returned by a decision-making algorithm will always be less than 100%. We hope that further research will help solve these issues.

## 4. Conclusion

With the increase of the GoA and shift to full automation of control, for a transportation system there arise additional safety risks related to not completely predictable behaviour of the constituent subsystems due to the application of machine learning methods in them. The introduction of ANN-based perception modules into the control loop significantly complicates the task of hazard analysis and safety evaluation of such systems using conventional FTA

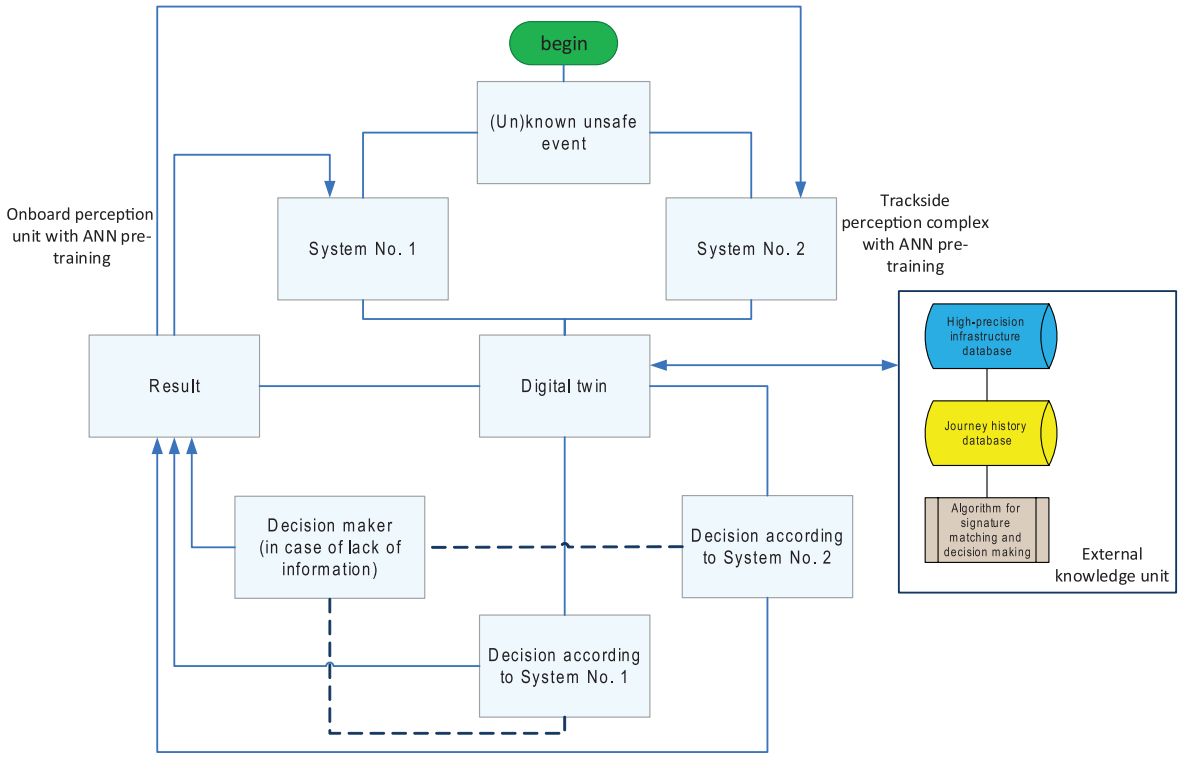


Fig. 6. Control structure with a virtual model

and FMEA methods. Evidently, the construction of a safety model of such complex multi-loop transportation systems requires the application of a comprehensive approach.

This approach must include a mandatory systems analysis of unsafe scenarios along with the compilation of scenarios library and the formalization of a hazard model's description, pertaining to the boundaries of various control loops as well. The systems analysis may result in a further review and change of the safety model of a transportation system under design and the conclusion about the necessity of having an additional component in the model taking on the supervision and constraining function – e.g., by implementing a decision making algorithm based on a digital twin. At the same time, the introduction of a digital twin (or virtual model) into a safety-critical system is an absolutely new and not yet well-proven approach to functional safety that is subject to further research. We can only hope that further research will make it possible to prove the feasibility of constructing a "supervised artificial neural network" complying with the conventional safety requirements applied to mass transportation systems, or to develop some other adequate supervision and constraining algorithm.

In turn, the proposed method based on STPA and control theory may become a universal methodological platform for the simulation and design of autonomous transportation systems. As a logical extension, based on the presented approach there may later also follow some design and development of a specialized software for automated risk evaluation of systems and technological process under construction.

## References

1. https://www.uitp.org/publications/world-report-on-metro-automation/.

2. IEC 26690:2014. Railway applications – Urban guided transport management and command/control systems – Part 1: System principles and fundamental concepts.

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

4. Qi Y., Cao Y., Sun Y. Safety analysis on typical scenarios of GTCS based on STAMP and STPA. *IOP Conference Series: Materials Science and Engineering* 2020;768(4):042042.

5. Leveson N.G. A systems-theoretic approach to safety in software-intensive systems. *IEEE*

6. *Transactions on Dependable and Secure Computing* 2004;1(1):66-86.

7. Bensaci C., Zennir Y., Pomorski D. A Comparative Study of STPA Hierarchical Structures in Risk Analysis: The case of a Complex Multi-Robot Mobile System. European Conference on Electrical Engineering & Computer Science. Bern (Switzerland); 2018.

8. ISO/PAS 21448:2019 (SOTIF). Road Vehicles – Safety of the Intended Function.

9. Popov P.A. [Development of Russian and foreign driverless operation technology]. *Automation, Communication and Informatics* 2020;9:6-12. (in Russ.)

10. Arnold V.I. "Hard" and "soft" mathematical models. MTSNMO Publishing house; 2004. (in Russ.).

11. Yan F., Zhang S., Tang T. Autonomous Train Operational Safety assurance by Accidental Scenarios Searching. IEEE Intelligent Transportation Systems Conference. IEEE; 2019. P. 3488-3495.

## About the authors

**Alexey V. Ozerov,** Head of Department, JSC NIIAS, 27, bldg 1 Nizhegorodskaya St., 109029, Moscow, Russian Federation; e-mail: A.Ozerov@vniias.ru

**Alexey M. Olshansky,** Head of Centre, JSC NIIAS, 27, bldg 1 Nizhegorodskaya St., 109029, Moscow, Russian Federation; e-mail: A.Olshanskiy@vniias.ru

## The authors' contribution

**Ozerov A.V.** analyzed major approaches, identified their benefits and drawbacks, developed a general scheme of command and control, basic operational scenarios, a control scheme with a virtual model.

**Olshansky A.M.** proposed the concept of "supervised artificial neural network" and the sequence of actions related to the description of functional safety hazards.

## Conflict of interests

The authors declare the absence of a conflict of interests.