# Model of risk management in railway transport

*Pavel* Burdyak[1*]

[1]Siberian Transport University 630049 Novosibirsk, Russia

**Abstract.** The article discusses approaches to risk management models based on modern information approaches. The described approach to the formation of assumptions arises with the provision of ubiquitous operation of railway transport based on the analysis of events that were not previously widely disseminated due to the occurrence of negative consequences. One of the promising methods for predicting risks is a machine study with the formation of a region of probabilistic results. The use of machine research and application is based on its combination of artificial intelligence, mathematical statistics, numerical methods, probability theory, graph theory. A necessary dimension of machine learning is working with a large amount of data, including data that was not previously represented in statistics. A scheme for working with data on all African realities, strategic risk management is described. The result of using models is the creation of technologies that reduce infrastructure investments, reduce the risk of risks when using transport processes, by taking into account the effects and predicting risks. The trend of "fine dynamics" repeats the processes and, as a result, they probably get a positive effect on the transfer process.

## 1 Introduction

Risks in the organization of train traffic can be divided into external and internal. As a rule, all attention is paid to internal risks that make unforeseen adjustments to the transportation process. These include, first of all, failures of technical means, technological violations and infrastructural restrictions. External risks are considered as more systemic phenomena: fluctuations in traffic volumes, seasonality, economic factors, etc. At the same time, there are external risks of a higher level, as an example - dominance in the industry. The loss of control over dominance in the industry, as practice shows, for a long time pushes the player into the position of a follower.

External and internal risks in the organization of train traffic are connected and should not be separated. The strategy for managing them cannot choose to avoid or pass them on. The only thing that will make it possible to control the transportation process and translate it outside is the way to reduce and accept internal and external risks. This approach is the most mature and correct.

The size of train traffic on the networks of Russian railways of foreign countries is huge, both in terms of transportation of goods and in terms of the number of transport units (trains). Despite the difficulties in organizing such a process, opportunities for obtaining a

* Corresponding author: BurdyakPS@mail.ru

competitive advantage should be considered. And first of all, this is getting a large amount of data (bigdata) about the progress of train flows, which can be collected and used for analysis. The volume of such data is many times greater than that of the European rail networks, where freight traffic is in the background after high-speed and high-speed traffic. Leading the way in using the results of big data analysis will make it possible to obtain positive results and offer them to the transport services market (Figure1)

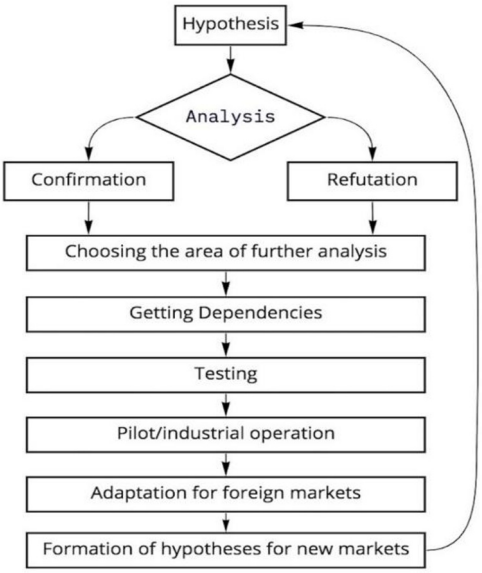


**Fig. 1.** Scheme of using data, them analysis and formation of a hypothesis

Figure 1 summarizes hypothesis testing and its use:
- a hypothesis is put forward about the presence of dependence;
- the data are analyzed to confirm or refute the hypothesis;
- the choice of the area of further analysis is set to obtain dependencies;
– testing, comparison of results with already known experience;
- Trial operation at the landfill or within the railway;
– industrial operation;
– adaptation for foreign markets, hypotheses for new markets and a new round of development.

## 2 Research methods

The unstoppable development of the transportation process, taking into account its own and "foreign" database, profit, technical and technological dominance will be the result. This approach solves the problem of continuous improvement of the transportation process due to the synergistic effect of input data from several sources.

The main stage in this sequence should be the stage of putting forward a hypothesis about the existing dependence and the problem being solved. To form a pool of hypotheses, it is proposed to process materials in digital form using artificial intelligence (AI) algorithms. Here, working materials, scientific research, conferences, library materials, etc. should be analyzed [1-14].

Several dozen systems are examples of actually used technologies for the use of automated text processing. Below are three examples: Tomita-parser, RCO Fact Extractor SDK, ABBYY FlexiCapture.

The formation of hypotheses about possible problems is perhaps a more important task than the use of already known and applied technologies for collecting and processing data. Including human errors. These errors can also be described, for example, using models:

– model by J. Rasmusen;

– J. Reason model.

These models and their basic principles make it possible to use them to find weaknesses in new, more complex technical and technological systems, taking into account possible human errors.

With the development of information technologies for the collection and storage of data - databases (databases), data warehouses (data warehousing), it became possible to process large amounts of information without human intervention. Using Data Mining, it is possible to process information sources (materials of scientific research, conferences, library materials, newspaper articles) in order to extract and collect data that are relevant to the topic of interest in predicting risks in the organization of train traffic. Data extraction is the first and preparatory stage for knowledge extraction (analysis).

At this stage, with the help of Data Mining, the problem of determining the impact of an event on the process of train traffic and further classifying input events according to known characteristics (the influence of weather conditions, geopolitical events in the world, recession or growth of the economy, new ideas of competing players in the transportation market, etc.) is solved. Data after this stage in an ordered form, according to the classification, are placed in data warehouses (Figure 2).

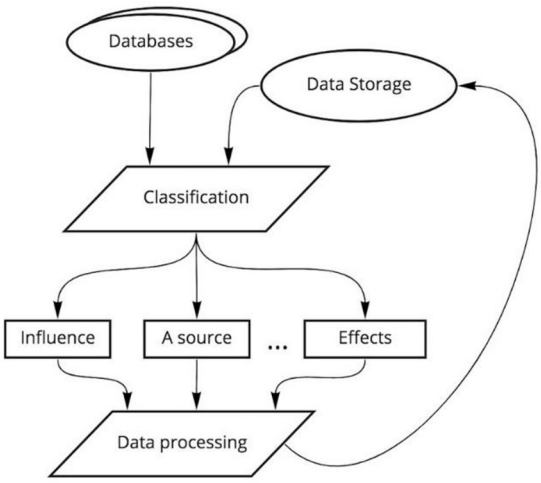


**Fig. 2.** Scheme of using data, them analysis and formation of a hypothesis

The use of Machine Learning for learning and further application is based on the use of:

– artificial intelligence;

– mathematical statistics;

– numerical methods;

– probability theory;

– graph theory, etc.

The main task of such a system is to learn to solve problems similar to a pool of classified events.

To solve this problem, there is no restriction in the application of the teaching method with or without a teacher. If we apply the Supervised learning method, for each precedent a pair is set: situation - required solution. When choosing a teaching method without a teacher (Unsupervised learning) - only the situation. In this case, the system learns to perform the task without outside interference. For the unsupervised learning method, the problem of clustering and identifying its structure is formulated. The solution methods can be graph clustering algorithms or statistical clustering algorithms.

Dividing the sample into groups of similar objects according to risks, significant events of competing players, and identifying trends will simplify further data processing and decision-making on strategic actions in risk management in the organization of train traffic.

As an example of the operation of the model, a transportation process is considered, which has fluctuations associated with various factors. These factors are mostly random and unpredictable. They are incorporated in the calculation of the capacity of railway lines and are intended to restore the system's performance. However, this reserve has limitations. Therefore, part of the interruptions in the movement of trains can be reduced and, accordingly, the quality indicators of railways can be increased...

Of all the possible dependencies between transportation processes and its execution, it is worth highlighting the most significant. To abandon the technical means (the full set of data is in the fixations of technological and special private ones), which led to a failure in the movement of trains, the initial data can be:

1. The effectiveness of the actions of employees to eliminate the violation (number, experience, provision of the necessary resources).
2. The state of the object of failure (running time in hours, frequency of maintenance, manufacturer).
3. Natural conditions (temperature, precipitation, time of day, wind speed) - weather station data.
4. Other data from the systems of the work performed.

Logical links should be drawn from the selected data on failures and systematized by place, time, quantity, frequency, distance from the objects of influence, etc. Bigdata and AI allow you to go beyond obvious connections, as in the example: a worker of service "A" and a worker of service "B" were able to reduce the failure time due to their random presence in the place of failure (performed other job duties). The delivery time to the point of failure will be the result of a cross-match of unrelated failures. This may affect the location of the locations of responsible employees.

Another non-obvious example: in a group of trains on a stage, the first train is under control

the most experienced (rested, prosperous, healthy, strong, etc. the number of criteria is not limited for bigdata) driver.

Suppose that the “most” driver goes first, and the “not the best” driver closes. At the same time, for the system, “not the best” does not mean “bad”, but only inferior to “most”. In the event of a failure associated with a malfunction of the locomotive, the probable exit time from it is less if the “most” is the first, and accordingly more if the group is led by the closing “not the most”. On fig. 3 shows a fragment of the schedule of the executed movement of trains with the situation described above. The data in the figure is anonymized.

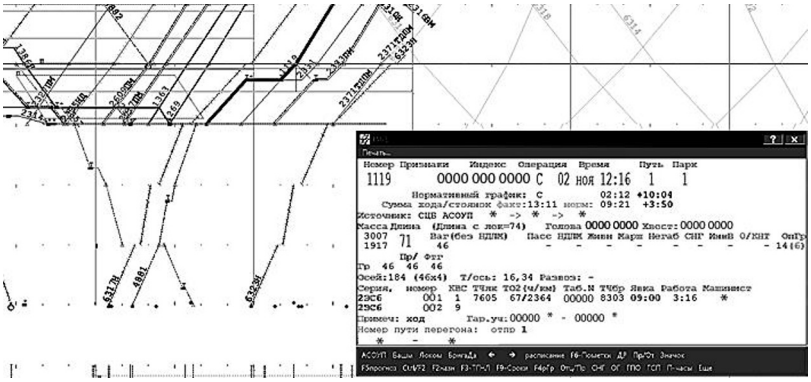


**Fig. 3.** A fragment of the schedule of the completed train movement on the stage

Figure 3 shows a real case of delay of train No. 1119 on the haul and subsequent delay of three freight trains. At the same time, an hour before the violation, trains of a larger mass passed the stage without failures. Elimination of delays on the stage and movement on schedule was restored only after 50 minutes.

This case was registered in the system on the note of the train dispatcher and is subject to investigation and attribution of guilt for the violation to the responsible service or unit. However, a group of six trains that passed without failure before the sent electric train will not be analyzed. There is no reason for this. At the same time, this group probably carries more information about the combination of trains and the order of their departure than "failed No. 1119".

From the analysis of the train situation, it is possible to extract information about the case according to the traffic schedule without using special algorithms. For example:

– driver's working hours;
– information about the locomotive;
– train data;
– the presence of restrictions on the stage and stations of the section;
– the presence of failures in the schedule at the previous sections;
– etc.

This data is mainly used to decide on liability for infringement. They are not even compared with the same violations in all the variety of combinations of factors surrounding the violation. It is worth noting that the trains that passed the place of violation before and after train No. 1119 had, according to Fig. 3, additional letters in the train number. This means that the conditions for running such trains were complicated by increased weight, oversized cargo, and long trains.

## 3 Research results

It can be assumed that any changed arrangement of the order of drivers will have a decrease in consequences compared to the original arrangement of drivers in the group. At the same time, the impact on the technology of operation of departure stations will also be positive. But this is only an assumption that needs to be confirmed. The effect of regrouping the order of departure will be in cases where trains have passed the section with a minimum number of failures. In cases where there was a failure in the train schedule, we will not see this.

If this order of things is followed by a data-supported analysis of a huge sample of failure cases, you can get an inconspicuous effect locally and a noticeable global one. The size of the group and the placement of drivers will give an unobvious result: increase the

throughput, section speed and launch a mechanism for a deeper study of the effect. This approach is of particular importance when planning the passage of barrier places and fulfilling the indicators of a variant train schedule.

Finland, Canada and other countries have already faced the problem of improving the quality of work of locomotive crews and are solving the problem in an extraordinary way: by monitoring the health, nutrition, rest, psychological state of the employee and his family members based on monitoring data, receiving them from different sources. The principle is the same - the analysis of many previously unaccounted for factors that can affect the process.

The result of applying this approach for the company will be:

– new data about users of the system and beyond;
– increasing confidence in AI tools;
– additional information about objects in the operational environment;
– individual recommendations as a decision support system;
– automation of routine actions;
– convenience and economy of the process of transportation management;
– simplification of decision-making;
– constant search for the best combinations and solutions;
– information about risk tolerance.

Figure 4 depicts a data management system for implementing a risk reduction strategy

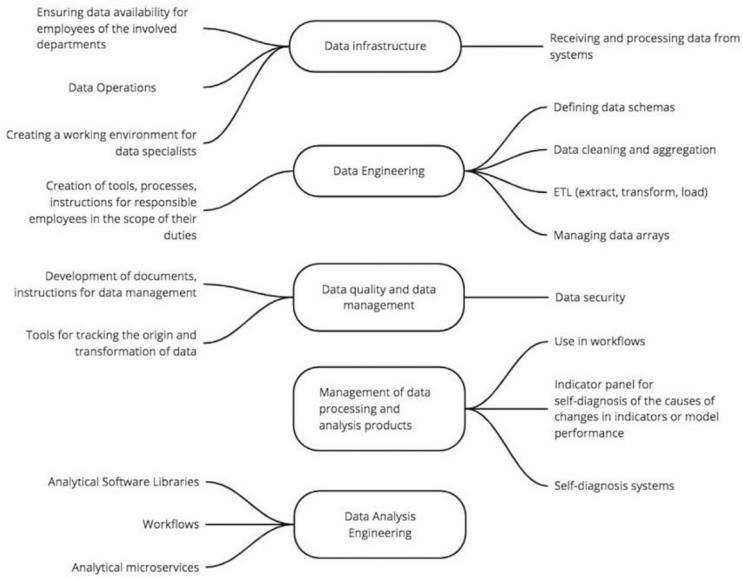


**Fig. 4**. Scheme of working with data at all stages of the implementation of the risk management strategy

## 4 Results discussion

The risk management strategy should have the goal of collecting hypotheses and their solutions in order to move the problem-finding function from the field of manual labor to automated using AI. The emergence of higher-level algorithms, with a supporting system consisting of data from existing automated control systems and modeling the development of events is a dominance trend.

In order to quickly solve the problems formulated above, it is necessary to address them not only to specialists on the ground, but also outside, to existing channels of communication with potential solutions: individual entrepreneurs, start-ups, platforms for finding ideas, etc.

For the first stage, data from existing automated control systems can be used, and then data from modified and supplemented automated control systems. These changes should be based on the obtained and predictive solutions from the areas of search for solutions outside the areas covered by the existing ACS.

The result of the above is a technological approach, not an infrastructural one.

The creation of a more advanced technology will make it possible to partially compensate for the creation of a new infrastructure and will allow the use of the existing infrastructure and its reserves by fine-tuning the transportation process. These results can be offered to the players of the transport market of railway transport everywhere and for small fees, since the main goal is to get even more data to improve their own processes.

## References

1. H. Rodemann, S. Templar, International Encyclopedia of Transportation, Eurasia Rail Freight: Enablers and Inhibitors of Future Growth, 435-455 (2021) DOI: 10.1016/B978-0-08-102671-7.10283-0
2. A.Di Graziano, V.Marchetta, Journal of Rail Transport Planning & Management 20, A risk-based decision support system in local railways management, 475-482 (2021), DOI: 10.1016/j.jrtpm.2021.100284
3. 3. Levin Dmitry Yuryev Bulletin of the Research Institute of Railway Transport 77, Adapting the organization of local work to modern conditions 18-26 (2018), DOI: 10.21780/2223-9731-2018-77-1-18-26
4. List of countries by rail usage - URL: https://en.wikipedia.org/wiki/List_of_countries_by_rail_usage (Date of the application: 22.01.2022)
5. Molchanov A. N., Skurnovich A. V., Online Journal of Science Studies **1 (26)**, Mathematical model of natural language text for estimating coherence, 1-14 (2015), DOI: 10.15862/70TVN115
6. J-L ZhouYi Lei, Reliability Engineering & System Safety, **204**, 107148 (2020), A slim integrated with empirical study and network analysis for human error assessment in the railway driving process DOI: 10.1016 / j.ress.2020.107148
7. J. Rasmussen, Cognitive systems engineering (1994) DOI: 10.1016 / S0005-1098 (97) 84591-8
8. J. Reason, Managing the risks of organizational accidents (2016) DOI: 10.4324/9781315543543
9. Huining Li, Huan Chen, Chenhan Xu, Anarghya Das, Xingyu Chen, Zhengxiong Li, Jian Xiao, Ming-Chun Huang, Wenyao Xu, Smart Health **23**, Privacy computing using deep compression learning techniques for neural decoding, 1-8, 2022, DOI:10.1016/j.smhl.2021.100229.
10. Komleva N., Zinovatnaya S., Liubchenko V., Applied aspects of information technology **1**, Methodology of information monitoring and diagnostics of objects represented by quantitative estimates based on cluster analysis, 376-392, 2020, DOI: 10.15276/aait.01.2020.1

11. Bessonenko, S., MATEC Web of Conferences. Siberian Transport Forum **239**, Issues of regulation of fillability of a train traffic schedule, 1–11 (2018) DOI:10.1051/matecconf/201823902004
12. V. Ponomarenko, Advances in Intelligent Systems and Computing, 823, 576-581 (2019) Generative games in aviation DOI: 10.1007/978-3-319-96074-6_59
13. A. Bieliatynskyi, K. Krayushkina, V. Breskich, M. Lunyakov, Basalt Fiber Geomats – Modern Material for Reinforcing the Motor Road Embankment Slopes, Transportation Research Procedia, Volume 54, 2021, Pages 744-757, https://doi.org/10.1016/j.trpro.2021.02.128
14. Wiebke Lenze, Nils Nießen, Journal of Rail Transport Planning & Management **20,** Quality in rail timetabling - A detailed review of stakeholders' interests compared with current practices in Germany, 21 p (2021) DOI: 10.1016/j.jrtpm.2021.100282.