https://doi.org/10.33271/nvngu/2022-2/141

A. Ursarova[1],
orcid.org/0000-0002-2879-8857,
R. Mussaliyeva[1],
orcid.org/0000-0001-8867-9932,
B. Mussabayev[1],
orcid.org/0000-0002-1794-7554,
D. Kozachenko[2],
orcid.org/0000-0003-2611-1350,
R. Vernyhora[2],
orcid.org/0000-0001-7618-4617

1 – Academy of Logistics and Transport, Almaty, the Republic of Kazakhstan
2 – Ukrainian State University of Science and Technologies, Dnipro, Ukraine, e-mail: dmytro.kozachenko@outlook.com

MULTI-CRITERIA EVALUATION OF PROFESSIONAL QUALITIES OF RAILWAY DISPATCHING PERSONNEL USING COMPUTER SIMULATIONS

Purpose. Development of a formal method to evaluate professional training of operating and dispatching personnel of railway stations.

Methodology. Simulation approaches to generate performance indicators of a human dispatcher under the specified operational conditions are selected as well as techniques of cluster and discriminant analysis to evaluate the obtained results.

Findings. A human factor is one of the main factors determining the performance of technical systems in general and the efficiency of a railway transportation system in particular. Analysis of the transport accident causes demonstrates that from 63 to 95 % of them are related to some extent to a human factor. In this regard, a problem of objective assessment of dispatching staff training quality and its effect on the indices of railway transportation functioning, including the operation safety is reflected in this article. Based on investigations carried out and the results obtained the operator training simulator of railway station was developed.

Originality. A simulator is improved of the railway station functioning while using deterministic finite automata formalizing the procedures. The proposed approach helps simulate processes taking place at the railway stations and resulting from the correct actions by operators as well as the incorrect ones. In addition, the paper proposes the improved method to evaluate professional training of station operators based upon the cluster and discriminant analysis providing HF impact minimization in the process of the problem solving.

Practical value. The approaches have been proposed to apply while developing training equipment for station operators. On the one hand, formal methods of cluster and discriminant analysis help provide human participation in the evaluation of professional training of station operators at the stage of training sample formation to be analysed. On the other hand, they make it possible to identify the trainee affiliation with a group of the defined training level and prevent from HF impact while evaluating training results of certain applicants.

Keywords: *railway transport, traffic safety, human factor, professional selection, training facilities, simulation modelling*

Introduction. Provision of safe train movement is one of the key tasks of both mainline [1] and industrial railway transportation [2] being a complex, dynamic, stochastic, and controlled system. A subsystem of the operating dispatch control of train movement and switching operation in terms of railway blocks and stations is one of the most important railroad subsystems, whose functioning influences significantly the indicators of safe functioning, use of rolling stock, and infrastructure along with the economic indices. Consequently, a problem of objective assessment of the quality of dispatching staff training and its effect on the indices of railway transportation functioning, including the operation safety, is rather topical.

Literature review. Maintenance of safe train movement is one of the main requirements for the functioning of railroad transportation systems. The train safety control systems available in terms of railroad transportation means annual analysis of train movement safety for enterprises of different management levels [3]. These analytical reports concerning Ukrainian railroads are formed separately in terms of six railways (regional branches of *Ukrzaliznytsia*); in particular, they contain information concerning the number of transport accidents (TA) on each road and the amount of recoveries from staff for the traffic safety violation. Fig. 1 represents dynamics of the number of transport accidents for the period of 2008–2020.

As a comparison, Table 1 shows data concerning the number of transport accidents on the railroads of Kazakhstan which are rather similar to the Ukrainian ones in terms of infrastructure scales and operation volumes. Thus, the length in use of the *National Company "Kazakhstan Temir Zholy"* (NC "KTZ") railway lines is about 16 thousand km, and its total working railway mileage is 21 thousand km. In case of *Ukrzaliznytsia* (UZ), those are 19.8 thousand km and 27 thousand km, respectively. In 2020, NC "KTZ" transported 287 million tons of freight at the freight turnover being 232 billion t-km; as for UZ for the same period, the number is 306 million tons at the freight turnover being 176 billion t-km. However, the number of railway staff in Kazakhstan is quite smaller – 120 thousand people compared to 180 thousand people in Ukraine.

As the data in Table 1 demonstrate, in terms of approximately similar transportation volumes, the number of transport accidents at NC "KTZ" shows stable decreasing tendency – by 2.5 times for 7 years, while in case of Ukrzaliznytsia this index remains the same from year to year (with some exceptions) at the level of 500–700 accidents. The relative number of TA per unit of transport operation (tonne-kilometre) is also illustrative (Fig. 2).

If average value of that parameter for *Ukrzaliznytsia* is 3.5, NC "KTZ" has the value of only 1.4 with the clear tendency to its reduction.

Analysis of the transport accident causes demonstrates that from 63 to 95 % of them are related to some extent to a human factor.

Along with that, attribution of those transport accidents to a human factor is quite subjective and differs both in terms of railroads and time scale. Namely, in 2020 the average number of penalties per one transport accident on Lviv railroad was 0.5 while the same on Prydniprovska railway was 2.9. The overall dynamics of this index across *Ukrzaliznytsia* is shown in Fig. 3.

Analysis of the trends represented in Fig. 3 shows that within the period of 2008–2020 the estimate of a human factor in transport accidents at *Ukrzaliznytsia* has changed considerably: in 2009, the average 13.4 penalties were for violations related to transport safety per one transport accidents, but in 2020, the value of that factor was only 1.9, i. e. it decreased by 7 times.

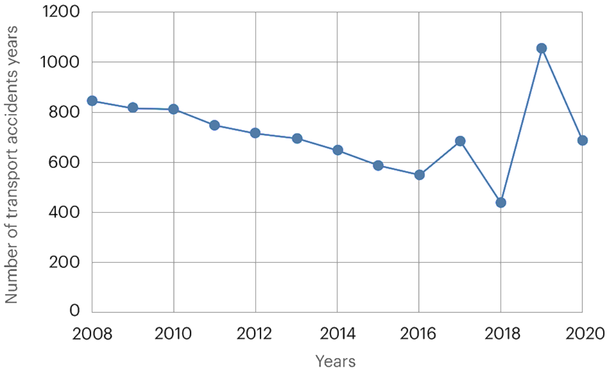

Fig. 1. The number of transport accidents on the railroads of Ukraine for the period of 2008–2020

Table 1

Indices of railway transportation safety in Ukraine and Kazakhstan

Year	Freight turnover, billion t-km		Number of TA		Average number of TA per 1 billion t-km	
	UZ	KTZ	UZ	KTZ	UZ	KTZ
2013	219.6	231.2	697	403	3.2	1.7
2014	210.2	216.5	650	386	3.1	1.8
2015	195.1	189.8	589	361	3.0	1.9
2016	187.6	188.2	550	355	2.9	1.9
2017	191.9	206.3	686	248	3.6	1.2
2018	186.3	220.0	442	180	2.4	0.8
2019	181.8	224.0	1054	178	5.8	0.8
2020	175.6	232.0	688	158	3.9	0.7

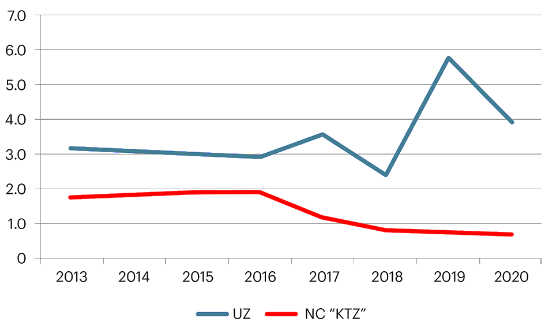

Fig. 2. The relative number of transport accidents on the railroads of Ukraine and Kazakhstan per 1 billion t-km

Meanwhile, it is hard to suppose that during that period the quality of Ukrainian railway staff performance, aimed at maintaining safe traffic, has increased by 7 times as well. The authors of the paper consider that one of the reasons of such significant change in this parameter is the subjective way of estimation while analysing possible causes of transport accidents. It is necessary to highlight that any violation of train movement safety is related to a human as it is the human having the functions of parrying the factors being their causes. Nevertheless, human activity is not always the main cause of transport accident.

Due to that, the issues of high-quality training and professional education of railroad staff is always topical. The comparative analysis of training and improvement of professional level of both Ukrainian and Kazakh railway staff shows the following: in 2020, NC "KTZ" spent about USD 6 million for the staff training, 28 % of all the workers were trained to one degree or another; but, in case of *Ukrzaliznytsia* and at total

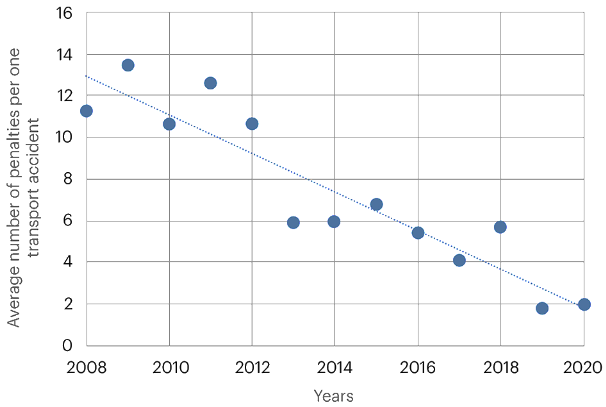

Fig. 3. The average number of staff penalties for traffic safety violation per one transport accident for the period of 2008–2020

costs of USD 1.5 million, about 17 % of all staff were covered by the training activities (without considering the planned technical lessons). In this context, average training time for one Kazakh railway worker in 2020 was 35 hours while the Ukrainian one was given 10–20 hours.

From the viewpoint of increasing level of traffic safety and reducing negative effect of a human factor, the important task is to get objective estimate of the level of professional readiness of the railway staff, first of all – operating and dispatching personnel, and its influence on the indices of transportation system functioning. This problem is impossible to solve without application of scientific approach and modern information technologies [4].

Arising of a problem concerning estimation of human influence on the conditions of technical system functioning goes back to the advent of industrial revolution in the Western European countries in the *18th–19th* centuries. The beginning of the *20th* century gave active development for different scientific methods to solve this problem in the framework of engineering psychology, management, and technical sciences. A great amount of historical material concerning the "human factor" problem is represented in review [5] carried out by M. Derksen.

According to the definition given in [6], human factors are the characteristics of a person (or a group of people) and a technical system revealed in some specific conditions of their interaction in the systems called "human – technical system", whose functioning is determined by achievement of the preset objective. It should be noted that applied studies of the "human factor" problem were connected, above all, with aviation and provision of flight safety. In 1910, professor H. Munsterberg formulated the concept of "personal factor" that was reduced to human errors while aircraft controlling. Dependence of the reliability of the person, responsible for the technical system control, on his/her psychological features was taken as the basis of the professional selection methodology. Further studies on the "human factor" problem demonstrated that quite often the people involved in flight accidents were the best pilots, whose qualification was undisputable. As a result, the conclusions were made in 1930s as for the following: the specialists' reliability is to be determined not only by their personal characteristics but also by disadvantages of their labour instruments; consequently, for people not to make any mistakes, any drawbacks of the labour instruments, reducing the reliability, should be excluded. Correct combination of human abilities and technical system capacities increases efficiency and reliability of the "human – technical system – environment" system, stipulates optimal use of technical facilities by people, and reduces both expenditures for creation of technical systems and costs for their operation.

Consideration of the "human factor" problem on railway transport is related mostly to the issues of train movement safety and civil defence. A great number of scientific papers deals with the analysis of dynamics and specification of the trends in train movement safety indices. The examples here are the works by Evans [7] for the conditions of the European Union and by Tokmurzina-Kobernyak, et al. [8] for the conditions of Kazakhstan. Both papers indicate the fact that the staff errors are the main cause for violating the train movement safety. Analysis of a human factor in the provision of safe railway functioning is performed in paper [9].

It should be emphasized that railway disasters and accidents are usually the results of conjunction of numerous unfavourable factors. The paper by Wang, et al. [10] proposes to use a learning algorithm for singling out the factors being the transport accident cause. The research by Vorobyov, at al. [11] offers an algorithm for calculating losses due to a human factor taking into account the number and duration of its resulting breakdowns. The study by Samsonkin V. & Goretskyi O. [12] represents a statistic method to evaluate train movement safety aimed at reducing a role of subjective human factor while solving this problem.

Analysis of technological processes of railway functioning shows that the requirements for the provision of train movement safety can often restrict the conditions of operations and reduce the processing and carrying capacity of railway transport. Moreover, there is a connection between the flow intensity of trains processed at a station and the number of transport accidents. According to [13], in most cases the connection is of linear nature. The research carried out in [14] demonstrates that with the course of time a gap between the conditions, being the basis for calculation of complex time norms for technological railway-transport operations, and requirements of train movement safety is coming up. Thus, a problem arises concerning the analysis of work of the railway-transport operating and dispatching personnel. The publication by Kyriakidis [15] touches upon a problem of assessing the quality of railway staff operation. A method proposed in [15] makes it possible to identify the factors favouring accidents or incidents and resulting in the deteriorated performance of dispatching personnel.

Transport accidents are quite rare events; as a rule, they are the result of conjunction of several unfavourable factors. In such conditions, it is quite difficult to estimate the efficiency of measures for improving train movement safety basing on the results of observations or physical experiments. Numerous scientific papers connect solution of this problem with the use of simulation modelling methods. In this context, two trends of studies can be highlighted:

- development of simulation models of railway transportation systems making it possible to model human activity as one of the elements of such systems;
- development of simulation models making it possible for people to participate in the modelling process.

Papers [16, 17] are the examples of the studies where the first approach is implemented. In particular, the paper by Boudnaya, at al. [16] represents a model of railway signalling system developed on the basis of Stochastic Petri Nets Analysis and taking into account all possible equipment failures and an operator's errors. The paper by Shalyagin & Nichiporuk [17] gives the analysis of approaches to modelling the operation of a human operator of railway transportation systems. Advantages of the simulation models, which include a human operator model, are in the fact that they allow modelling the long-term periods of time. However, considering the fact that the main task of a human operator is to solve complexly certain formalized problems, adequate modelling of the human operator's behaviour is rather a complexly formalized problem as well.

The second approach is implemented in training facilities for the operating and dispatching personnel. The current models use the methods of virtual reality [18], network learning technologies etc. The facilities help simulate real operating conditions of station officers and evaluate the influence of real people on the performance of the objects and systems under their control. A disadvantage of this approach is in the fact that the modelling speed is limited, and its results depend considerably on subjective decisions made by a human involved in modelling. In additions, one of the trends of the second approach application is the implementation of simulation models in the automated systems of railway transport control [19], which makes it possible to forecast the effect of decisions made by a human on the efficiency of operating transportation systems.

Consequently, the performed analysis demonstrates that a human factor is one of the main factors determining the performance of technical systems in general and the efficiency of a railway transportation system in particular. Currently, the problems as for development of methods for formal evaluation of training of the railway operating and dispatching personnel and estimation of its effect on the railway transport operation have not been solved completely, requiring further studies.

This paper deals with the problem of evaluation of professional training of station officers. Incorrect actions of dispatching staff result not only in decreased operating efficiency of stations; they also cause frequent transport accidents [20]. At the same time, it is a human in the system of railway station management that helps provide proper functioning of a station under dynamically changed conditions: in terms of changing volumes and nature of work and environmental conditions.

Purpose. The purpose of the paper is to develop formal methods for evaluating professional training of the railway-station operating and dispatching staff. The selected research methods are as follows: a method of simulation modelling to obtain performance measures of a human operator under the specified operating conditions; methods of cluster and discriminant analysis to evaluate the obtained results.

Methods. Computer-based simulation modelling was selected as the main method for obtaining the indices of railway station operation. In this context, a station is considered as a complex, dynamic, stochastic, and ergatic (man-controlled) system [21]. General structure of the developed model is represented in Fig. 4.

A simulation model of a station is based on the functional model represented in [21]. In this model, a process of station functioning is analysed as a sequence of transitions between separate states of technological process of station functioning. In this context, each of these states is characterized by the location of a rolling stock on tracks and a degree of completion of technological operations with different objects (trains, locomotives, train cars). A procedure of implementation of technological operations in a model is preset with the help of finite deterministic automata. Internal and external signals effect the transitions of automata, corresponding to the objects, from state to state. Internal signals come from the model itself upon the completion of technological operations in terms of changes in the rolling stock location on tracks as well as at certain moments of time in response to a timer. External signals (commands) come from a human operator performing the modelling. The commands make it possible for a human operator to select an order and priority of servicing certain objects. A list of possible commands, corresponding to the operator's correct or incorrect actions and resulting in changes in a current state, is specified preliminarily for each model state. All other commands of an operator's are considered to be inadmissible. Issuing inadmis-

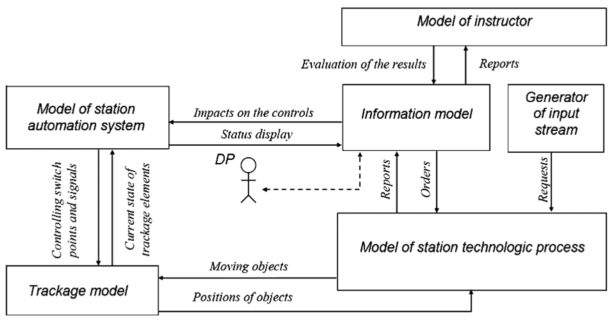

Fig. 4. General structure of a railway station model in the operator training simulator of railway station

sible commands is taken as the operator's error; it does not cause model changing. Command transmission to a model from a human operator and reflection of the current model state are performed with the help of an information model, which provides the operator with all information necessary for station control – both visual (location of a rolling stock on the station tracks, condition of points and traffic lights, block occupancy and others) and verbal (reports on completion of technological operations, commands from a dispatcher, requests for operation completions, etc.). The information model is implemented in the form of mnemonic diagram of a railway station with the necessary controlling elements (buttons, handles, communication switches) based on the station operator's workplace. Fig. 5 shows the example of such an information model for the training facility of intermediate station.

The instructor model provides organization and support of the modelling and training processes as well as the evaluation of the performance measures of the station and efficiency of the human operator's actions for controlling its technological process.

The modules are synchronized at the discrete moments of time upon the system timer command. While working on a training facility, all human actions are recorded; upon the completion, the initial indices are determined which characterizes the quality of station operation control such as the number of processed trains and cars within a certain station, downtime of trains and cars at this station, their downtime at neighbouring stations and blocks, list of errors made by the operator and so on.

In terms of the available training facilities, in most cases a level of the operator's professional training is assessed by a human instructor basing on the performance analysis of the completed training. However, the availability of numerous factors complicates their unambiguous interpretation. Moreover, such an approach does not exclude the influence of a subjective factor while assessing the worker's actions according to his/her results of working on the training facility. Thus, the authors have developed a methodology of integral evaluation of the level of professional training of station operators according to their results while working on the training facility. The methodology relies on mathematical methods of classification that make it possible to include an applicant into one of the worker groups, each of which is characterized by certain professional qualification.

The methodology essence is as follows. The results of facility-based training of a human operator (applicant) are assessed to form a vector of classifying variables $\boldsymbol{X}^* = \{x_1, x_2, ..., x_N\}$. Basing on this vector, an applicant is included into one or another group using one of the discriminant analysis methods. Each of those groups includes the workers of certain professional level. Discriminant analysis is the method of classification with learning, i.e. it means the availability of certain database of already classified workers (learning sample). The learning sample is the matrix **S** with dimension $M \times N$, where M is sample volume. Vector $\boldsymbol{X}_i = \{x_{i1}, x_{i2}, ..., x_{ij}, ..., x_N\}$ $(I = 1 - M, j = 1 - N)$ is set for each matrix line (sample object); the vector includes

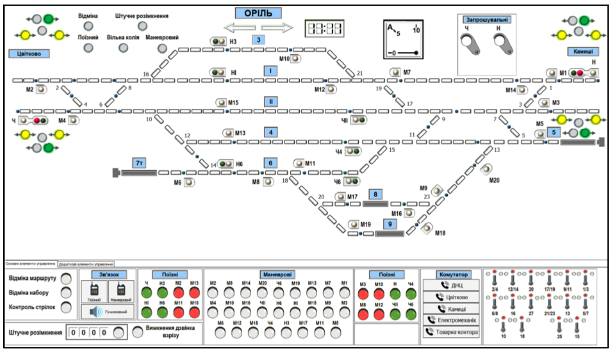

Fig. 5. Information model of an intermediate station

values of the simulator training performance for one of the previously classified workers and may be represented by a point in the N – dimensional space. In this context, each vector X_i of the learning sample belongs to a certain qualification group g_k, ($k = 1 - G$, where G is the total number of groups). The initial classification of workers to get learning sample **S** can be performed with the help of a cluster analysis method taking into account expert estimates. To implement the indicated approach to the evaluation of applicants' professional qualification, the following problems were set and solved:

- selection of informative parameters to form a classifying vector;
- standardization of the values of informative indices;
- generation of the learning sample for classification.

Selection of the indices to evaluate the worker's qualification is quite an important task since the excessive number of N indices usually results in the reduction of the obtained estimates. Thus, it is necessary to single out the subset of the most significant (informative) indices among numerous measures. This problem has been solved involving a disperse analysis method and Scheffe's multiple-comparison approach. As a result, the following informative indices have been specified: the number of trains processed at station M, average duration of train servicing at a station – T_t, average downtime of trains at neighbouring stations – T_n, average downtime of trains before the entrance signal – T_e, total of penalty points per one train – Z. Performance indices M, T_t, T_n, Te are specified immediately after the training results. Taking into consideration the fact that the significance of possible errors, made by station operators during their work, can differ greatly, each error is estimated by certain number of penalty points depending on its significance, while forming Z index. To obtain these estimates, the representatives of auditing and operating-dispatching staff of Prydniprovia railway were questioned to use the results for determining the significance (according to a 10-point scale) of each possible error.

The indices used for classification are measured in different units and have different value ranges. Due to that, indices with a wide value range exert dominant influence on the classification results. To eliminate this influence, one should standardize the values of both indices of learning sample **S** and vector X^* used while classifying a concrete object. Selection of the method can change substantively the geometry of the initial N-dimensional space; consequently, development of the methodology for assessing the training results involved analysis of the efficiency of different data standardization methods.

The standardization methods were studied in combination with different methods of cluster and discriminant analysis while classifying the training facility performance of a control group of 50 people that included 15 professional station operators and 35 students. In this context, first of all, the "reference" classification of the trainees into three groups was carried out by expertise: "high level of training" (group A), "medium level" (group B), and "low level" (group C). The efficiency of each standardization method was estimated by two criteria: probability of erroneous classification P_{err} and value of total intragroup dispersion D_{ig}, which determines magnitude of dispersion of certain index values in the qualification groups g_k after classification. Criterion P_{err} for each method of standardization was defined by comparing the classification results obtained with the use of this method with the "reference". The studies showed that the best results are obtained by using a method of "maximum standardization" according to the formula

$$x_{i,j} = \frac{x^*_{i,j}}{x^{\max}_j}, \qquad (1)$$

where $x^*_{i,j}$ is the value of the j^{th} index for the ith object of the sample before standardization; $x^{\max}_j$ is the maximum value of the j^{th} index among all sample objects.

While applying this standardization method, error probability P_{err} was 0.03–0.05; and the value of intragroup dispersion

D_{ig} accepted substantially smaller values than while using other methods of data standardization. As a comparison, it should be stressed that application of the poorest standardization methods for solving this problem stipulates probability of erroneous classification at the level of 0.30 the entrance signal 0.40.

To obtain the initial learning sample **S** used for further evaluation of the worker's professional qualification by a discriminant analysis, methods of cluster analysis are used. The purpose of cluster analysis methods is to single out such homogeneous spaces (clusters) in the initial multidimensional data, so that the objects inside a cluster were "similar" in a peculiar way to each other while the objects from different clusters were "not similar". Here, the "similarity" means closeness of the objects in a multidimensional space.

A task of getting learning sample **S** is in division of some initial set **X** of vectors $X_i = \{x_{i1}, x_{i2}, ..., x_{ij}, ..., x_N\}$ with the training indices of the control group of people into clusters (groups) g_k ($k = 1 - G$) so that each cluster g_k would include vectors X_i characterizing the people with similar level of professional training.

Among numerous available algorithms of cluster analysis, we have selected and studied the algorithms of one-way classification as they are oriented to highlighting the clusters with preset properties that is in compliance with the conditions of a specified problem. It should be noted that one-way classification algorithms are divided into agglomerative (strategy of cluster unification) and divisive (strategy of division). In case of agglomerative algorithms, each classified object is a separate cluster when classification starts. Each algorithm step involves unification of two closest clusters. That continues up to the moment when the number of clusters reaches certain value G (in most cases, the value is preset). As for divisive algorithms, all objects belong to one cluster for the beginning. Next, the initial totality experiences gradual division up to the moment when the necessary division degree is obtained (the preset number of clusters k is achieved). As a rule, divisive algorithms are applied when the initial totality should be divided into a rather large number of clusters (more than 10) that does not correspond to the conditions of a problem of workers' classification in terms of their facility training results. Moreover, divisive algorithms are quite unstable while classifying "disputable" objects being at the boundaries of clusters. Thus, only agglomerative algorithms were studied for the most efficient method for obtaining the initial sample.

While selecting a classification method for generating the learning sample, four measures of inequality for separate objects were compared to identify the most suitable one for solving the specified problem. A measure of inequality between two objects a and b is the value (distance) d_{ab}, which is the greater, the fewer similarity is between objects a and b. Formulas for calculation of values d_{ab} for each of considered measures of inequality are given in Table 2. Here, x_{aj}, x_{bj} are values of the j^{th} index for objects a and b, respectively ($x_{aj} \in X_a$, $x_{bj} \in X_b$, $j = 1 - N$).

The efficiency of each measure of inequality was estimated in terms of probability of erroneous classification P_{err} and value of total intragroup dispersion D_{ig} while comparing with the "reference" classification of a control group. Studies on the efficiency of application of inequality measures while forming the learning sample by the cluster analysis methods showed that the best quality for classification of the training facility performance is obtained while using Canberra measure ($P_{err} = 0.04–0.06$). This measure of inequality appeared to be the most stable one while classifying "disputable" objects. The poorest results were obtained while calculating distances between the objects with the use of Manhattan measure ($P_{err} = 0.25–0.30$).

As it has been mentioned earlier, agglomerative algorithms of cluster analysis mean gradual unification of clusters until their number reaches the specified value. Each step of agglomerative algorithm involves unification of two clusters, the distance between them being minimal. In this regard, there arises a problem of determining distances between certain clusters. A technique of calculation of this distance is determined by the adopted unification strategy.

Assume that there are two groups of objects (clusters) U and V, the number of elements in which is n_u and n_v, respectively; a distance between these clusters is d_{uv}. Suppose that d_{uv} is minimal among all possible distances between the rest of the clusters. In this case, clusters U and V are united into one cluster W with the number of elements $n_w = n_u + n_v$. Consider some cluster Y that includes n_y elements. If distances d_{uy} and d_{vy} are known prior to the unification of clusters U and V, then a distance between new cluster W and cluster Y will be as follows

$$d_{wy} = \alpha_u \cdot d_{uy} + \alpha_v \cdot d_{vy} + \beta \cdot d_{uv} + \gamma \cdot |d_{uy} - d_{vy}|, \qquad (2)$$

where α_u, α_v, β, γ are parameters determining the essence of unification strategy.

While determining the most efficient classification algorithm when getting the learning sample, six unification strategies represented in Table 3, were studied. In addition, Table 3 represents the parameters for determining distances d_{wy} (2) between separate clusters for different unification strategies.

Comparative analysis of application of the indicated strategies to classify the training results of a control group has shown that the "flexible" unification strategy is the one that makes it possible to obtain the results being the closest ones to the "reference" learning sample; in terms of this strategy, the classification errors were not more than 4–6 %. The poorest results while solving the problem under consideration are observed at using the strategies of "nearest" and "far" neighbour (35–40 % of errors).

Immediate evaluation of the trainee's professional level involves a discriminant analysis method. The essence of the method is in the fact that a vector of classifying variables $X^* = \{x_1, x_2, ..., x_N\}$ of a trainee is used to relate him/her to one of the worker groups, the data about whom is in the training facility database. In this context, the worker being assessed is sent to the qualification group g_i ($I = 1 - G$), for which a value of discriminant function of normal distribution is reduced to maximum

$$f_i(X^*) = (2\pi)^{-N/2} \cdot (\det \mathbf{S}_i)^{-1/2} \times \\ \times \exp\{-1/2 \cdot ((X^* - \mathbf{m}_i) \cdot \mathbf{S}_i^{-1} \cdot (X^* - m_i)')\} \to \max, \qquad (3)$$

where $\mathbf{m}_i$ is vector of mathematical expectations of the parameters $x_1, x_2, ..., x_N$ for the i^{th} qualification group; $\mathbf{S}_i$, $\mathbf{S}_i^{-1}$ are covariational and its reverse matrices with the geometry of $N \times N$ for the i^{th} group.

Along with that, as the studies demonstrate, application of function (3) for classifying the training results leads to erroneous classification (error probability is 0.20–0.25), which is connected with the assumption concerning normal distribution of data in the classification groups. Therefore, Fisher's linear discriminant function was used to evaluate the training results. As is known, if all groups g_i, among which classification of object X^* is performed, have statistically similar covariational matrix **S**, then expression (3) can be represented in the form of linear function

Table 2

Measures of inequality in the process of classification

No.	Measure	Formula
1	Euclidean	$d_{ab} = \sqrt{\sum_j^N (x_{aj} - x_{bj})^2}$
2	Manhattan	$d_{ab} = \sum_j^N \lvert x_{aj} - x_{bj} \rvert$
3	Bray-Curtis	$d_{ab} = \frac{\sum_j^N \lvert x_{aj} - x_{bj} \rvert}{\sum_j^N x_{aj} + \sum_j^N x_{bj}}$
4	Canberra	$d_{ab} = \sum_j^N \frac{\lvert x_{aj} - x_{bj} \rvert}{\lvert x_{aj} \rvert + \lvert x_{bj} \rvert}$

Table 3

Parameters of unification strategies in the agglomerative algorithms of classification

No.	Strategy name	Strategy parameters
1	"Closest neighbour"	$\alpha_u = \alpha_v = 0.5 \quad \beta = 0 \quad \gamma = -0.5$
2	"Far neighbour"	$\alpha_u = \alpha_v = 0.5 \quad \beta = 0 \quad \gamma = 0.5$
3	"Group mean"	$\alpha_u = n_u/n_w \quad \alpha_v = n_v/n_w \quad \beta = \gamma = 0$
4	"Centroid"	$\alpha_u = n_u/n_w \quad \alpha_v = n_v/n_w \quad \beta = -\alpha_u \cdot \alpha_v$ $\gamma = 0$
5	"Quadratic"	$\alpha_u = (n_u + n_y)/(n_w + n_y)$ $\alpha_v = (n_v + n_y)/(n_w + n_y)$ $\beta = -n_w/(n_w + n_y) \quad \gamma = 0$
6	"Flexible"	$\alpha_u = \alpha_v = 0.625 \quad \beta = -0.25 \quad \gamma = 0$

$$h_i(X^*) = (X^* \cdot \gamma_i) \cdot \lambda_i \to \max,$$

where

$$\gamma_i = \mathbf{S}^{-1} \cdot \mathbf{m}_i; \quad \lambda_i = 0.5 \cdot (\mathbf{m}_i \cdot \mathbf{S}^{-1} \cdot \mathbf{m}/i). \tag{4}$$

Fisher's linear discriminant (4) has a series of advantages compared to (3) as, owing to its linearity and better reliability of estimates of covariational matrices, it gives a smaller error in the process of classification.

One of the methods of discriminant analysis – method of coalitions – means uniting the groups g_i, which covariational matrices are statistically equal, into so-called coalitions. The data of the groups included into one coalition are united; and they are used to calculate general covariational matrix **S**. Next, classification of the object X^* is done in terms of maximum of Fisher's linear discriminant function (4). Hypothesis of the equality of covariational matrices of different groups g_i is checked involving Box or Bartlett-Anderson method.

The studies carried out showed that, while assessing the training-facility performance of station operators, Fisher's linear function demonstrates a rather low level of erroneous classification (5–7 %), and it is more stable while classifying "disputable" objects.

Results. To obtain the initial set of applicant's performance, a mathematical model of the station operator's training facility has been proposed; the model is based on representation of technological process of station functioning in the form of deterministic finite automata. While training, such a model admits issuing of either correct or incorrect commands by a human; the effect of these commands change adequately a technological process of the station functioning. Use of Scheffe's multiple-comparison approach helped identify the informative indices that are determined according to the training results. The performed studies allow adapting the methods of cluster and discriminant analysis to solve the applied problem of evaluation of the station operators' professional training on the basis of the computer training facility results. It has been identified that the best results of the performance evaluation are reached while using:

- method of "maximum standardization" while standardizing the values of informative indices;

- Canberra measure and "flexible" strategy of unifying while forming the learning sample to classify by the cluster analysis methods;

- Fischer's linear discriminant function for the immediate evaluation of the trainee's professional level by the discriminant analysis methods.

The proposed methodology for evaluating the level of station operators' professional training helps eliminate human participation while assessing the performance of certain applicants. At the same time, as it has been mentioned earlier, a discriminant analysis is the method of classification with learning. It means that the learning sample can be constantly updated with the expansion of its volume at the expense of data of the objects being classified correctly. Such an organization also admits possibility of correcting erroneous classification. Thus, if classification errors are identified, a human instructor can correct the score received by a trainee and record the values of his/her performance into the necessary qualification group in the learning sample. In this context, the decision made by a human instructor is recorded, and the results are reflected on the estimates of other applicants' performance.

Basing on the developed method, a model of instructor of station operator's training facility has been implemented; the model evaluates a level of professional training of the station operators.

Originality of the research is in the improvement of a simulation model of railway station functioning on the basis of deterministic finite automata to formalize technological processes. The proposed approach makes it possible to simulate the processes taking place at railway stations as a result of both correct and erroneous actions of the railway operations. In addition, the paper offers the improved method for assessing professional qualification of station operators based on the application of cluster and discriminant analysis, which provides reduced influence of subjective factors while solving this problem. The proposed methods were used while developing training simulator for the station operators. On the one hand, use of formal methods of cluster and discriminant analysis allows providing human participation in the evaluation of professional qualification of station operators while forming the learning sample for further analysis. On the other hand, it helps identify formally the trainee's belonging to a group of certain level of qualification and eliminate the effect of subjective factors while evaluating the training results of certain applicants.

Differences between the classification of training results performed automatically or by a human instructor are within the range of 5–7 %. The indicated value of difference can be reduced at the expense of replenishment; and correction of the learning sample will help lessen this difference. Moreover, use of training facilities makes it possible to evaluate the work of different station operators under the specified conditions and establish the relations between operating conditions characterized, first of all, by the intensity of ingoing train flow as well as intensity of failures of technical means, on the one hand, and indices describing safety of train movement, on the other hand.

Conclusions. The studies carried out help draw the following conclusions:

- a human dispatcher is one of the key elements of railway transportation systems, which stipulates objectively considerable influence of a human factor both on the indices of safe train movement and the efficiency of railway transport use;

- there is a connection between the intensity of train and car flows processed at a station and the number of transport accidents. Accordingly, the human factor impact on the operations at railway stations as well as the professional level of their operating and dispatching staff should be evaluated basing on the experiments concerning station management by different people in terms of preset parameters of the arriving units. Taking into consideration the fact that experiments at real railway stations are practically impossible (as they put safety of train movement at great risk), the main method to get quantitative estimation of station operations is simulation modelling of technological processes;

- application of the methods of cluster and discriminant analysis helps develop a method for formal evaluation of the professional level of station operators. The proposed approach allows both providing objective assessment of the indices demonstrated by different station operators under similar working conditions at stations and giving the possibility for a human instructor to participate in evaluation at the stage of learning sample formation for the analysis.

References.

1. Kozachenko, D. R., Vernigora, R., Balanov, V. N., Berezovy, N., Yelnikova, L., & Germanyuk, Y. (2014). Evaluation of the Transition

to the Organization of Freight Trains Traffic by the Schedule. *Transport Problems*, *11*(1), 41-48. https://doi.org/10.20858/tp.2016.11.1.4.
2. Taran, I.A., & Klimenko, I.Yu. (2014). Transfer ratio of double-split transmissions in case of planetary gear input. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, (6), 60-66.
3. Naumov, V., Taran, I., Litvinova, Z., & Bauer, M. (2020). Optimizing resources of multimodal transport terminal for material flow service. *Sustainability* (*Switzerland*), *12*(16), 6545. https://doi.org/10.3390/su12166545.
4. Samorodov, V., Bondarenko, A., Taran, I., & Klymenko, I. (2020). Power flows in a hydrostatic-mechanical transmission of a mining locomotive during the braking process. *Transport Problems*, *15*(3), 17-28. https://doi.org/10.21307/tp-2020-030.
5. Derksen, M. (2014). Turning men into machines? Scientific management, industrial psychology, and the "human factor". *Journal of the History of the Behavioral Sciences*, *50*(2), 148-165. https://doi.org/10.1002/jhbs.21650.
6. *GOST 26387-84 Man-machine system. Terms and Definitions* (n.d.). Retrieved from http://helpnik.college.ks.ua/standart/gost/Catalog/Index/1/1563.htm.
7. Evans, A.W. (2010). Fatal train accidents on Europe's railways: 1980–2009. *Accident Analysis and Prevention*, *43*, 391-401. https://doi.org/10.1016/j.aap.2010.09.009.
8. Tokmurzina-Kobernyak, N., Vakhitova, L., & Ursarova, A. (2020). To the question of ensuring traffic safety by railway transport. *Bulletin of KazATC*, *4*(115), 137-144. Retrieved from https://vestnik.alt.edu.kz/index.php/journal/issue/view/1/1.
9. Di Grazia, G., Vittorini, B., Carlizza, L., Lamedica, R., & Fabbri, G. (2014). Human factor relevance in safety assurance of railway operations. *2014 AEIT Annual Conference – From Research to Industry: The Need for a More Effective Technology Transfer (AEIT)*, 1-6. https:// doi: 10.1109/AEIT.2014.7002059.
10. Wang, Y., Zheng, W., Dong, H., & Gao, P. (2020). Factors correlation mining on railway accidents using association rule learning algorithm. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems*, 1-6. https://doi.org/10.1109/ITSC45102.2020.9294317.
11. Vorobyov, V., Manakov, A., & Repina, I. (2018). Economic assessment of the control of human factor impact on faults of technical facilities in railway-transport technological processes. *MATEC Web Conf. Siberian Transport Forum – TransSiberia 2018*, 239. https://doi.org/10.1051/matecconf/201823908011.
12. Samsonkin, V., & Goretskyi, O. (2020). Control Technology of Railway Traffic Safety: A System Approach and Digitalization. TRANSBALTICA XI: Transportation Science and Technology. TRANSBALTICA 2021. *Lecture Notes in Intelligent Transportation and Infrastructure*. Cham: Springer. https://doi.org/10.1007/978-3-030-38666-5_66.
13. Zavyalov, A., Aksenov, V., Siniakina, I., Volkova, E., & Gulyaev, A. (2019). Analysis of the Human Factor Influence on the Quality of Work of Railway Station. *IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, *2019*, 2311-2316. https://doi.org/10.1109/EIConRus.2019.8657229.
14. Kozachenko, D.N., Verlan, A.I., & Korobiova, R.H. (2021). Development of Analytical Methods for Calculating Time Standards for Shunting Operations. *Science and transport progress*, *1*(91), 51-64. https://doi.org/10.15802/stp2021/228097.
15. Kyriakidis, M., Majumdar, A., & Ochieng, W.Y. (2018). The human performance railway operational index – a novel approach to assess human performance for railway operations. *Reliability Engineering & System Safety*, *170*, 226-243. https://doi.org/10.1016/j.ress.2017.10.012.
16. Boudnaya, J., Mkhida, A., & Idrissi, B. B. (2015). Dependability Assessment of the Railway Signalling Systems Based on the Stochastic Petri Nets Analysis. *International Journal of Intelligent Systems and Applications in Engineering*, *3*(2), 90-96. Retrieved from https://www.ijisae.org/IJISAE/article/view/547.
17. Shalyagin, D., & Nichiporuk, V. (2012). Simulation of the work of a human operator of a train traffic control system. *Nauka i tehnika transporta,* (2), 16-21.
18. Ferlin, A., Qiu, S., Bon, P., Sallak, M., Dutilleul, S.C., Schon, W., & Cherfi-Boulange, Z. (2018). An Automated Method for the Study of Human Reliability in Railway Supervision Systems. *EEE Transactions on Intelligent Transportation Systems, 19*(10), 3360-3375. https://doi.org/10.1109/TITS.2017.2779565.
19. Kyrychenko, H., Berdnychenko, Yu., Strelko, O., & Shcherbyna, R. (2021). Application of the Automated System at the Change of Technology of Work of Reference Stations on the Railway (2021). *Proceedings of 25th International Scientific Conference. Transport Means 2021.* Retrieved from https://transportmeans.ktu.edu/wp-content/uploads/sites/307/2018/02/Transport-Means-2021-Part-II.pdf.
20. Sabraliev, N., Abzhapbarova, A., Nugymanova, G., Taran, I., & Zhanbirov, Z. (2019). Modern aspects of modeling of transport routes in Kazakhstan. *News of the National Academy of Sciences of the Republic of Kazakhstan, Series of Geology and Technical Sciences*, *2*(434), 62-68. https://doi.org/10.32014/2019.2518-170X.39.
21. Bobrovskiy, V.I., Kozachenko, D.N., & Vernygora, R.V. (2014). Functional simulation of railway stations on the basis of finite-state automata. *Transport Problems*, *9*(3), 57-66. Retrieved from http://transportproblems.polsl.pl/pl/Archiwum/2014/zeszyt3/2014t9z3_06.pdf.

Багатокритеріальна оцінка професійних якостей диспетчерського персоналу залізниць із використанням комп'ютерних тренажерів

А. Урсарова[1], *Р. Муссалієва*[1], *Б. Муссабаєв*[1], *Д. Козаченко*[2], *Р. Вернигора*[2]

1 – Академія логістики та транспорту, м. Алмати, Республіка Казахстан
2 – Український державний університет науки і технологій, м. Дніпро, Україна, e-mail: dmytro.kozachenko@outlook.com

Мета. Розробка формальних методів оцінки професійної підготовки оперативно-диспетчерського персоналу залізничних станцій.

Методика. В якості методів дослідження обрані методи імітаційного моделювання для отримання показників роботи людини-диспетчера в заданих експлуатаційних умовах, а також методи кластерного й дискримінантного аналізу для оцінки отриманих результатів.

Результати. Людський фактор є одним з основних чинників, що визначає ефективність функціонування технічних систем у цілому та залізничної транспортної системи зокрема. Аналіз причин транспортних пригод показує, що від 63 до 95 % із них тим чи іншим чином пов'язані з людським фактором. У зв'язку з цим проблема об'єктивної оцінки якості підготовки оперативно-диспетчерського персоналу та його впливу на показники якості й функціонування залізничного транспорту, у тому числі показники безпеки, розглянуті у цій статті. На основі виконаних досліджень та їх результатів був розроблений тренажер чергового по залізничній станції.

Наукова новизна. Полягає у вдосконаленні імітаційної моделі функціонування залізничної станції на основі використання детермінованих кінцевих автоматів для формалізації технологічних процесів. Запропонований підхід дозволяє імітувати процеси, що відбуваються на залізничних станціях як у результаті правильних, так і помилкових дій чергових по станціях. Також у роботі запропоновано удосконалений метод оцінки професійної підготовки черговим по станціях, заснований на застосуванні методів кластерного й дискримінантного аналізу, що забезпечує зниження впливу суб'єктивних факторів при вирішенні даного завдання.

Практична значимість. Запропоновані методи використані при розробці тренажерів чергових по станціях. Використання формальних методів кластерного й дискримінантного аналізу дозволяє, з одного боку, забезпечити участь людини в оцінці рівня професійної підготовки чергових по станціях на етапі формування навчальної вибірки для аналізу, а з іншого – формально встановлювати належність тренованого до групи певного рівня підготовки та виключити вплив суб'єктивних факторів при оцінці результатів тренування окремих претендентів.

Ключові слова: *залізничний транспорт, безпека руху, людський фактор, професійний відбір, тренажер, імітаційне моделювання*

The manuscript was submitted 13.07.21.