INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2021, VOL. 67, NO. 4, PP. 603–608
Manuscript received February 23, 2021; revised November, 2021. DOI: 10.24425/ijet.2021.137852

Conceptual Model Creation for Automated Self-training System of Functional Control and Detection of Railway Transport

Ayaulym Oralbekova, Marzhana Amanova, Kamila Rustambekova, Zhanat Kaskatayev, Olga Kisselyova, and Roza Nurgaliyeva

***Abstract*—In order to improve the operational reliability and service life of the main systems, components and assemblies (SCA) of railway transport (RT), it is necessary to timely detect (diagnose) their defects, including the use of the methods of intellectual analysis and data processing.**

One of the promising approaches to the synthesis of the SCA functional control system is the use of intelligent technology (INTECH) methods. This technology is based on maximizing the information capacity of an automated decision support system for detecting faults during its training.

***Keywords*—Railway transport, a system of components and assemblies, decision support systems, intelligent technology, automatic detection systems, an object used for training**

I. Introduction

ENSURING reliable and trouble-free operation of all railway and high-tech railway systems remains one of the priority tasks in the segment of scientific developments related to the implementation, operation and modernization of such high-tech complexes. As has been shown by many scientists who have been engaged in research in this direction, in order to increase the operational reliability and service life of the main systems, components and assemblies (SCA) of locomotives, timely identification (detection) of their defects is necessary, even before an emergency occurs. This problem is solved by the functional control system directly during the operation of the control system. In addition, in practice, the recognition classes characterizing the possible functional states of SCA intersect in the feature space, that requires defuzzification of fuzzy data. When using a quantitative scale for measuring detection features, an effective method of such defuzzification is the use of machine learning. This approach makes it possible to transform the a priori fuzzy partition of the space of features for detecting anomalies in operation and faults of SCA into a clear set [1-5].

One of the promising approaches that need further development of the synthesis of a functional control system of SCA of RT is the use of ideas and methods of information-extremal intelligent technology (IEI-technology), based on maximizing the information capacity of the decision support system in the training process of the automated detection system of SCA.

Ayaulym Oralbekova, Marzhana Amanova, Kamila Rustambekova, Zhanat Kaskatayev, Roza Nurgaliyeva are with Kazakh University Ways of Communications, Almaty, Kazakhstan (e-mail: ayaulym83@mail.ru, marzhan.a@mail.ru, kamila8888@mail.ru, Zhanchost1@mail.ru, roza.Nur@mail.ru).

Olga Kisselyova is with Kazakh Academy of Transport and Communications named after M. Tynyshpayev, Almaty, Kazakhstan (e-mail: K.Olga964@mail.ru).

II. Purpose of the research

1) The solution of the problem, in particular, is associated with the need to form on the basis of known and new features (i.e., which were not initially entered into the knowledge base of the automated detection system - ADS) of the so-called object used for training (OUFT). This object is a matrix based on the realizations of the features of anomalies in the operation of SCA.

Within the framework of the article, the following tasks are solved:

- to form a glossary of feature implementations for each class of anomalies or faults, as well as an alphabet of classes in terms of fault detection objects (FDO)

- to determine the minimum size of the matrix that is used in the training process of the ADS (OUFT) (subject to the requirements for its representativeness);

- to determine the normalized permissible deviations for the implementation of the features of recognition (detection) of anomalies or faults in the operation of the control system of railway rolling stock.

III. Methods and models

In order to obtain an input mathematical description of the ADS, it is necessary to study and analyze in detail the features of operation of the primary sources of information, from which the ADS system receives data on the certain realizations of the features of faults. For example, in existing methods and means of ND, there are used devices as primary sources of information. The mathematical model of ADS in general form as a set-theoretical structure can be represented as follows [14, 15]:

$$\Delta_B = \langle IS, T, RS, SS, OS, , \Phi \rangle , \quad (1)$$

where IS– a set of input signals that are processed in ADS;

T – moment of time to obtain information about the state of the detected system, node or assembly;

RS – a set of feature implementations that are used in the process of detecting faults;

$SS-$ a space of possible states for a system, node or assembly that are subject to the NDC procedure**;**

OS – a set of data that is obtained at the output from the module for primary processing of signals (information), for example, from the NDC tools. Or the module of primary data processing - PDPM;

$\Pi : IS \times T \times RS \to SS$ – transition quantifier (used to record changes in the state of SNA, which are subject to detection during their operation. It is assumed that a change in states can occur under the influence of internal or external factors)

$\Phi : IS \times T \times RS \to LM$ – formation quantifier of the set LM (learning matrix - m).

The Cartesian product of sets IS, T, RS, SS is used as a universe UT of tests during ADS testing

$$UT = IS \times T \times RS \times SS. \qquad (2)$$

III.I Problem statement.

The ADS scheme, which includes a software module with self-training elements, is shown on Figure 1.

The quantifier $O\Theta : CL^{|2|} \to RC^{|2|}$ is used to divide the space of realizations of the *OR* features (faults or anomalies in the operation of SCA or - faults detecting objects - FDO) into two recognition classes.

The classification parameter OC is used to test the statistical assumption (that is, the hypothesis) that the FDO belongs to a certain class of faults or anomalies in the operation of SCA.

After assessing the hypotheses using the quantifier hy, there is formed a set AR^{is} that characterizes the accuracy of detecting the corresponding faults or anomaly in the operation of SCA (i.e., FDO).

It is accepted that ζ- the number of statistical assumptions, $is = \zeta^2$– the number of characteristics that can be processed in the ADS for SCA of the railway rolling stock.

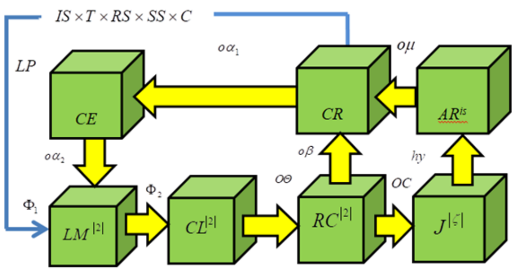


Fig. 1. Schematic diagram of the ADS software module

The quantifier $o\mu$ forms a set CR that allows to implement the procedure for assessing the efficiency of detecting faults or anomalies in the operation of SCA within a class.

The quantifier $o\beta$ closes the detection loop and is used to optimize the system of control deviations from templates (norms) that are stored in the ADS repository.

Quantifiers $\Phi_1 : IS \times T \times RS \times SS \times C \to LM^{|2|}$ and $\Phi_2 : LM^{|2|} \to L^{|2|}$ are used to form the input matrix used in the training process of the ADS (ITM) and in the organization of the binary training matrix (BTM), respectively. Here C - a fragment of data for detection.

The set CE is closed sequentially by quantifiers $o\alpha_1 : CR \to CE$ and $o\alpha_2 : CE \to LM^{|2|}$. These quantifiers allow to change the realizations of FDO features for different classes in the process of training ADS.

A quantifier $LP : CR \to IS \times T \times RS \times SS \times C$ is used to regulate the ADS training process.

Based on the conceptual scheme of the ADS operation, presented on Figure 1, we will formulate the following formalized statement of the problem of information synthesis of the ADS elements. Let the alphabet of FDO classes $\left\{CL_s^0 \left| s = \overline{1,S}\right.\right\}$ and a multidimensional binary matrix used for training (MBMT of FDO) which, accordingly, characterizes the $m-$th functional state of SCA for a specific recognition class CL_s^o, be known:

$$\left\| m_{s,i}^{(j)} \right\| = \left| \begin{array}{cccccc} m_{s,1}^{(1)} & m_{s,2}^{(1)} & \cdots & m_{s,k}^{(1)} & \cdots & m_{s,N}^{(1)} \\ m_{s,1}^{(2)} & m_{s,2}^{(2)} & \cdots & m_{s,k}^{(2)} & \cdots & m_{s,N}^{(2)} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ m_{s,1}^{(j)} & m_{s,2}^{(j)} & \cdots & m_{s,k}^{(j)} & \cdots & m_{s,N}^{(j)} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ m_{s,1}^{(n)} & m_{s,2}^{(n)} & \cdots & m_{s,k}^{(n)} & \cdots & m_{s,N}^{(n)} \end{array} \right| . \qquad (3)$$

In expression (3), the following designations are adopted: row of the matrix — implementation of the FDO $\left\{m_{s,i}^{(j)} \left| i = \overline{1,N}\right.\right\}$ representation, N — the number of informative realizations of features used to detect SCA; column - a stochastic sample $\left\{m_{s,i}^{(j)} \left| j = \overline{1,n}\right.\right\}$ that is used during training for a sample of size n.

A clear organization of the glossary of realizations of recognition (detection) features (or GFDO) is a prerequisite for ADS. The procedure for filling GFDO $\sum^{|N|}$, where $N = DS\sum^{|N|}$,is implemented as a sequence of actions aimed at formalizing the realization of features. In this case, the primary features characterize directly the failure of SCA, and the secondary features are derived from the primary ones.

As primary realizations of features, you can use parameters that are read from certain sensors or experimental data obtained directly, for example, during the implementation of the methods of the NDC SCA of railway rolling stock.

Various statistical characteristics can be used as secondary realizations of FDO features, for example, vectors of realization of a certain class $\left\{m_{s,i}^{(j)} \left| i = \overline{1,N}\right.\right\}$, training set $\left\{m_{s,i}^{(j)} \left| j = \overline{1,n}\right.\right\}$ for OUFT, etc.

The alphabet of FDO classes $\left\{m_s^0\right\}$ for ADS is formed at the first stage by the system developer with the involvement of specialists in diagnosing faults of the railway rolling stock.

At the second stage, the synthesis of the alphabet continues with the help of intelligent systems and technologies, for example, with the help of DSS or expert systems (ES) [14–17], which are capable of directly operating in the mode of cluster analysis of input data.

As was shown earlier in [3–10], in case of the invariability of the glossary of the realizations of features of FDO and an increase in the capacity of the alphabet, it is possible to change the asymptotic characteristics of ADS. Accordingly, this factor can significantly affect the functional efficiency of the training procedure for such systems. This is, in particular, due to an increase in the degree of intersection of classes of faults or anomalies in the operation of ADS, which are subject to detection.

III.II Formation of a binary matrix for training ADS

For a more convenient implementation of the procedure for creating a container, the following assumption is made: there is a container (CON) [14, 15], which allows consideration of the optimization parameters of CON in a binary feature space (BFS - RS_b), for some standard vector, for example, $cl_s \in CL_s^o$. The vertex of the vector defines the geometric center of CON - C_s^o. To calculate the radius (container radius), taking into account the works [14, 15, 16, 17], the following expression was used:

$$r_s = \sum_{i=1}^{N} (cl_{m,i} \oplus \zeta_i), \tag{4}$$

where $cl_{s,i}$ – i-th coordinate of the standard vector cl_s;

ζ_i – $i-$th coordinate of the vector ζ for FDO realization, the vertex of which refers to the container $C_s^o \in CL_s^o$;

N – number of realizations of FDO features in the ADS knowledge base.

The advantage of the K-means algorithm is its low computational complexity, while the algorithm works well when a large amount of data is processed. The DBSCAN algorithm is rather slow with a large amount of data. In the tasks of automatization of the SCA diagnosis, as a rule, the data volumes are small and the use of the K-means algorithm does not give a tangible effect [12].

Let's suppose that a series of measurements of the values of the controlled realizations of features in FDO for SCA of the railway rolling stock has been carried out, and the resulting matrix has the following form:

$$m = \begin{pmatrix} 0 & 1 & \dots & 1 & \dots & 1 \\ 1 & 0 & \dots & - & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ - & 1 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & - & \dots & 0 \end{pmatrix}.$$

Thus, the set of objects to be checked belonging to a class is specified by binary features. A dash indicates the uncertainty of the realization of the feature for FDO.

More detailed research results of the procedures for the formation of binary matrices used for training in recognition systems are given in [14-17], as well as in US patents: US 2011/0208714 A1; US9294502 and others.

We will assume that in order to assess the functional efficiency of a self-training ADS, it is also necessary to take into account the effect on the parameters of its operation of a structured vector of space-time parameters $v =< v_1, \dots, v_k, \dots, v_{RS} >$. We also take into account the corresponding restrictions $RC_k(v_1, \dots, v_{RS}) \le 0$. Using the technology of training ADS, the task of determining the optimal parameters of the vector $\{v_c^*\}$ in the field of determining max of the information criterion of the effectiveness of training ADS is set as the resulting goal of the training procedure [14, 15]:

$$CR_s^* = \max_V CR_s, \tag{5}$$

where CR_s – information effectiveness criterion (IEC) of the ADS training procedure during the detection of the FDO class CL_s^0;

V – permissible values of the parameters of ADS functioning.

Thus, during the information synthesis of ADS, it is possible to partially solve the problem by determining the optimal values of the parameter v_k^*:

$$v_k^* = \arg_{V_k} CR_s^*, \tag{6}$$

where V_k– the range of permissible values of the parameter v_k.

Let's consider the procedure for the DSS functioning as an element of ADS in the training mode "with a teacher", i.e. the case where there is a matrix used for training.

As a result of working with the multidimensional information space of the realization of features of FDO ADS for the railway rolling stock, it is possible to obtain a binary training matrix (BTM) $\{CL_s^j\}$, which consists of structured vectors-implementations of the image of the corresponding anomaly in the operation of SCA or their fault:

$$cl_s^j = \left\langle cl_{s,1}^{(j)}, ..., cl_{s,i}^{(j)}, ..., cl_{s,N}^{(j)} \right\rangle. \tag{7}$$

The binary training matrix is also used to assess the verification of permissible deviations in the detection process (or the permissible deviations verification/control system - PDCS). PDCS $\{\delta_{n,i} \,|i = \overline{1,N}\}$, as well as the parameters that determine the sample of coordinates of binary vectors for the standard classes of FDO, are subsequently stored in the ADS database.

For example, Figure 2 shows an example of converting component temperature measurement data in decimal and binary form for ADS.

№ п/п	Десятичное представление	Двоичное представление
C1	73	0100 1001
C2	98	0110 0010
C3	99	0110 0011
C4	115	0111 0011
C5	129	1000 0001
C6	140	1000 1100
C7	18	0001 0010
C8	200	1100 1000
C9	201	1100 1001
C10	240	1111 0000
C11	56	0011 1000
C12	236	1110 1100

Fig. 2. An example of the formation of a binary matrix for training ADS

Similar binary training matrices can be obtained for other units of measurement that are now used in the NDC tools. At the same time, the use of binary data representation is much more effective precisely in systems based on machine learning methods and intelligent technologies for data analysis, which was previously shown in [11–17].

In the DSS training mode as a part of ADS, a binary matrix OUFT is formed over a period of time τ, which is fed to the input of the DSS software component module responsible for the ADS training procedure. Note that the formation of OUFT should occur for a predetermined level of confidence to the generated OUFT matrices. Many works in the field of machine learning systems have been devoted to this issue [5, 13, 15, 17].

IV. Experiment

Figure 3 shows the process of forming the structure of the matrix used for training ADS. Moreover, the formation of the structure occurs in stages, including the vectors of realizations $\left\{cl_1^{(j)}\right\} \in CL_1^0$ and $\left\{cl_2^{(j)}\right\} \in CL_2^0$, respectively [14–16].

To create such a matrix, only the important characteristics of FDO should be selected. That is, the characteristics that uniquely distinguish some realizations of features of faults or anomalies in the operation of SCA within the class from others.

All possible values of each FDO property can be encoded in binary form [17], or using non-negative integers [14, 16], where zero corresponds to an undefined value of the FDO property. This, in particular, makes it possible to take into account the missing, new, or not yet provided values of the FDO properties.

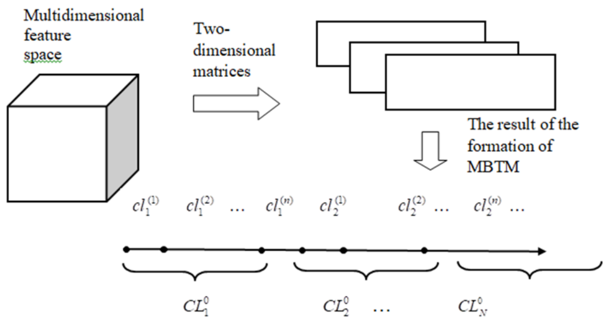


Fig. 3. Scheme of work with the multidimensional information space of the realizations of features of FDO ADS for the railway rolling stock

As an illustration, Table 1 shows an example of the process of forming a binary matrix and, accordingly, clustering the realizations of features in the process of detecting SCA using NDC tools based on the acoustic control of the components of the railway rolling stock [13–17].

This effect is used in various devices that make it possible to implement acoustic control of rolling stock components and assemblies, see Figure 4.

Eddy current NDC is widely used in various branches of the scientific and industrial complex of Kazakhstan and other countries, due to the high efficiency and reliability of solving problems of flaw detection, quality control of materials and products, determination of parameters and characteristics of objects of control for various purposes [12–16].

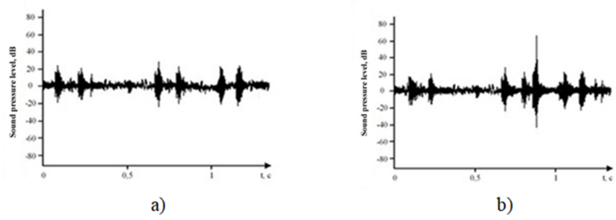
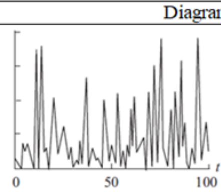
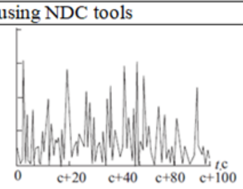
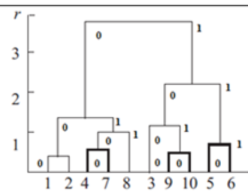
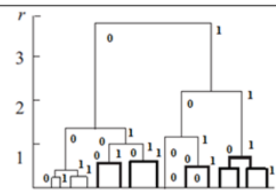


Fig. 4. Fragment of the automatic analysis of the sound accompanying the passage of the rolling stock of the undamaged (a) and damaged (b) rolling surface of the wheelset of the railway rolling stock

Table 1. An example of the formation of clusters and a binary matrix for OUFT for detecting the state of the rolling surface of a wheelset of the railway rolling stock (see Figure 4)

No acoustic anomalies	Acoustic anomalies recorded
Diagrams obtained using NDC tools	
	
Formation of clusters	
	
Formation of a binary matrix for OUFT	
Column - sample set for training ADS	Row - implementation of the FDO
$m = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$	$m = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & - & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & - & 0 & 0 \end{pmatrix}$

By combining the data into compact clusters, it is possible to analyze the typical representatives of each cluster and decide whether such data is a realization of a feature of a fault or anomaly in the operation of the SCA or not. Then this solution is transferred to all representatives of the studied cluster. This approach significantly reduces the amount of information required for the successful classification of FDO.

Since clusters can take complex forms in the multidimensional space of feature realizations, some authors have proposed various algorithms for clustering feature realizations. So, for example, in the works listed below, the application of the methods and algorithms K-means, DBSCAN, FDBSCAN [14-16], etc. is described, see Figure 5.

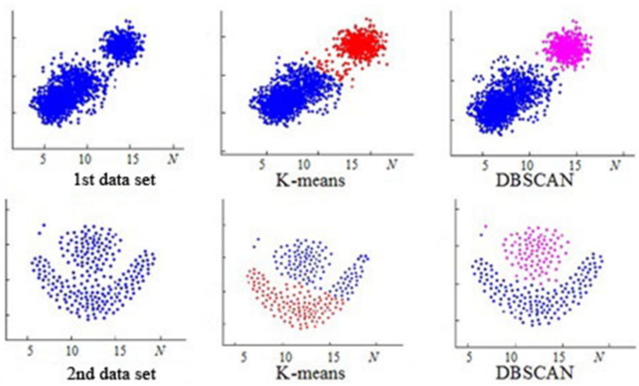


Fig. 5. Examples of data clustering based on DBSCAN and K-means algorithms in diagnostic systems

In [18–21], it was shown that the computational complexity of algorithms used in a binary space for realizations of the detection features of SCA (BSFR) of the corresponding class (classes) depends on the optimal container shape for the corresponding class of the object of detection.

After the formation of binary matrices that are used as objects in the process of learning of the automated system for diagnostics (detection) anomalies and failures in the operation of nodes and aggregates of the rolling stock, there are created binary trees of anomalies or failures signs clustering, see fig. 6,7.

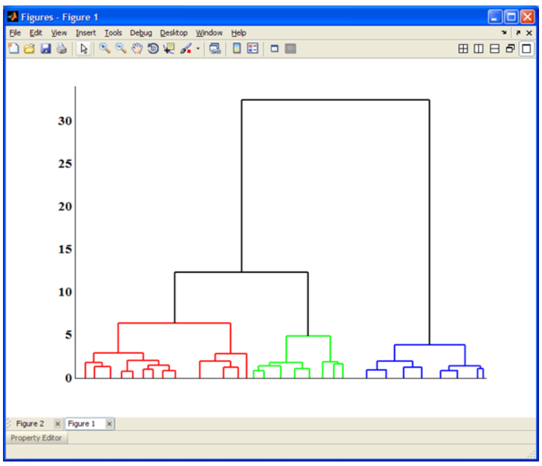


Fig. 6. Normal behavior of the detected nodes and aggregates of the railway transport rolling stock

In this case, the amount of recognition signs varied within $N = 9 - 15$. The optimal amount of clusters was selected at the maximum values of ICFE. As the analysis of the results showed the optimal amount of clusters is equal to $C = 3$.

Figure 4 shows a histogram of the dependence of the ICFE value for variants of the dictionaries of the anomalies and failures signs of nodes and aggregates from the amount of steps of the SADS learning algorithm$\{w\}$, shows the dependence of ICFE from the amount of signs used to train the system for failure diagnostics and detection.

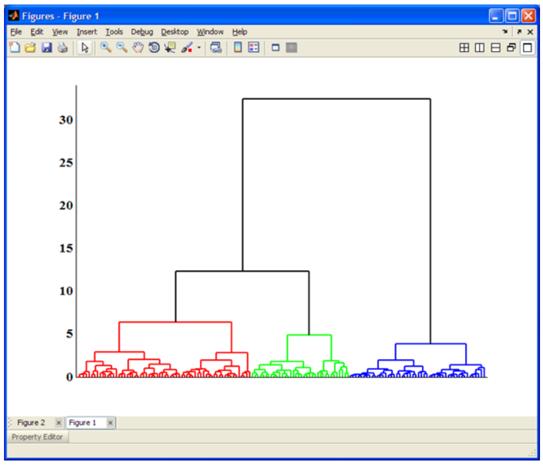


Fig. 7. Anomalies or failures signs clustering of the detected nodes and aggregates of the railway transport rolling stock

Analysis of simulation results showed that the use of an algorithm with 5–10 signs of learning is quite effective in SADS. That is, for this case, the ICFE reaches its maximum value. This, in turn, indicates the possibility of creation error-free decision rules in failure diagnostics and detection.

In the SADS testing mode a sufficient amount of steps $\{w\}$ for accurately determination of anomalies and failures classes were $w = 2500 - 3000$.

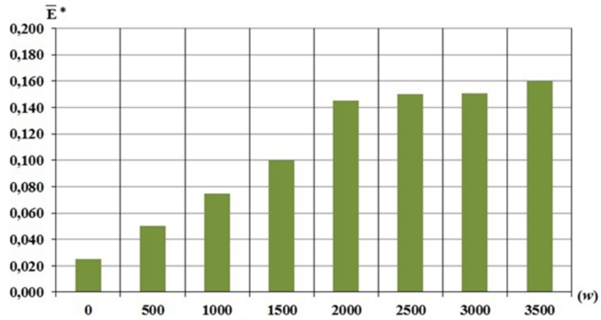


Fig. 8. Dependence of the max ICFE value for variants of the dictionary of anomalies and failures signs of nodes and aggregates of the simulated railways transport systems

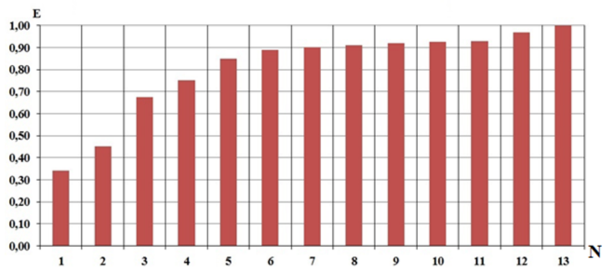


Fig. 9. Diagram of the dependence of ICFE from the amount of signs realizations used for SADS

In the situations when during the simulation process and at the creation of an algorithm for recognizing anomalies and failures of nodes and aggregates of railway transport there were added representative sets of greater length, the efficiency of the algorithm was the same. Adding representative sets of shorter length reduced the efficiency of the algorithm.

Conclusions. The article proposes clarifications and additions to the machine learning method of the automatic detection system (ADS) of the functional state of components and assemblies of railway transport. As well as the corresponding model and machine learning algorithm, which, in contrast to existing solutions, are implemented by parallel optimization of control permissions for the features of fault recognition of SCA. Such a solution allows, in the future, to create effective decision rules for intelligent decision support systems (DSS) and ADS of faults and for diagnostics of the state of components and assemblies of railway transport.

REFERENCES

[1] M. Schickert, I. Aydin, M. Karaköse, E. & Akin, "A new contactless fault diagnosis approach for pantograph-catenary system," in *MECHATRONIKA*, 2012 15th International Symposium, pp. 1-6, IEEE, 2012.

[2] A. Le Mortellec, J. Clarhaut, Y. Sallez, D. Berger, & Trentesaux, "Embedded holonic fault diagnosis of complex transportation systems," *Engineering Applications of Artificial Intelligence*, 26(1), pp. 227-24, 2013.

[3] Ph M. Papaelias, C. Roberts, L.C. & Davis, "A review on non-destructive evaluation of rails: state-of-the-art and future development," *Proceedings of the Institution of Mechanical Engineers*, Part F: Journal of Rail and rapid transit, 222(4), pp. 367-384, 2008.

[4] J. Seeliger, Mackel, D. Georges, "Measurement and diagnosis of process-disturbing oscillations in high-speed rolling plants," in *Proc. XIV IMEKO World Congress*, 2002.

[5] X.S. Jin, J. Guo, X.B. Xiao, Z.F. Wen, & Z.R. Zhou, "Key scientific problems in the study on running safety of high speed trains," *Engineering Mechanics*, 26(Sup II), pp. 8–25, 2009.

[6] L. Mariani, F. Pastore, & M. Pezze, "Dynamic analysis for diagnosing integration faults," *IEEE Transactions on Software Engineering*, 37(4), pp. 486-508, 2011.

[7] Y.C. Huang, H.C. Sun, "Dissolved gas analysis of mineral oil for power transformer fault diagnosis using fuzzy logic," *IEEE Transactions on Dielectrics and Electrical Insulation*, 20(3), pp. 974-981, 2013.

[8] Orbán, Zoltán, and Marc Gutermann, "Assessment of masonry arch railway bridges using non-destructive in-situ testing methods," *Engineering Structures*, 31.10 (2009): pp. 2287-2298, 2009.

[9] Yella, Siril, M. S. Dougherty, and N. K. Gupta, "Artificial intelligence techniques for the automatic interpretation of data from non-destructive testing," *Insight-Non-Destructive Testing and Condition Monitoring*, 48.1: pp. 10-20, 2006.

[10] P.S. Bhowmik, S. Pradhan, M. Prakash, "Fault diagnostic and monitoring methods of inductionmotor: a review," *International Journal of Applied Control, Electrical and Electronics*, Vol. 1, pp. 1-18, 2013.

[11] Petr Dolezel, Pavel Skrabanek, Lumir Gago, "Pattern recognition neural network as a tool for pest birds detection," *Computational Intelligence*,SSCIIEEE Symposium Series on, pp. 1-6, 2016.

[12] V. Lakhno, Y. Tkach, T. Petrenko, S. Zaitsev, & V. Bazylevych, V. "Development of adaptive expert system of information security using a procedure of clustering the attributes of anomalies and cyber attacks," *Eastern-European Journal of Enterprise Technologies*, (6 (9)), pp. 32-44, 2016.

[13] V. Lakhno, "Creation of the adaptive cyber threat detection system on the basis of fuzzy feature clustering," *Eastern-European Journal of Enterprise Technologies*, 2.9 2016:18, 2016.

[14] V.A. Lakhno, P.U. Kravchuk, V.P. Malyukov, V.N. Domrachev, L.V. Myrutenko, O.S. Piven, "Developing of the cyber security system based on clustering and formation of control deviation signs," *Journal of Theoretical and Applied Information Technology*, Vol. 95, Iss. 21, pp. 5778-5786.

[15] A.S. Dovbish, "Osnovi proektuvannja ntelektualnih sistem" / A.S. Dovbish. Sumi: SumDU, 171 p., 2009.

[16] X. Zhang, N. Feng, Y. Wang, & Y. Shen, "Acoustic emission detection of rail defect based on wavelet transform and Shannon entropy," *Journal of Sound and Vibration*, p. 419-432, 2015.

[17] A. Giantomassi, F. Ferracuti, S. Iarlori, G. Ippoliti, & S. Longhi, "Electric motor fault detection and diagnosis by kernel density estimation and Kullback–Leibler divergence based on stator current measurements," *IEEE Transactions on Industrial Electronics*, 2015, 62(3), pp. 1770-1780, 2015.