*Proceedings of 26th International Scientific Conference. Transport Means 2022.*

# Signal Analysis of the Armature Rotation Irregularities in the Traction Electric Motor by Unsupervised Anomaly Detection Methods

**B. Bodnar[1], O. Ochkasov[2], V. Serdiuk[3], M. Ochkasov[4]**

[1]*Ukrainian State University of Science and Technologies, Lazaryana St., 2, 49010 Dnipro, Ukraine*
[2]*Ukrainian State University of Science and Technologies, Lazaryana St., 2, 49010 Dnipro, Ukraine,*
*E-mail: abochkasov@gmail.com*
[3]*Ukrainian State University of Science and Technologies, Lazaryana St., 2, 49010 Dnipro, Ukraine*
[4]*Taras Shevchenko National University of Kyiv, Bogdana Gavrilishina str., 24, Kyiv, Ukraine, 02000*
*E-mail: misha.ochkasov@gmail.com*

**Abstract**

Reducing the cost of maintenance and repair of vehicles is possible due to the early detection of malfunctions development and more complete use of the equipment resource. The introduction of contemporary information technologies allows real-time monitoring of the technical parameters for the equipment. Anomaly detection methods are used to process monitoring results. The paper presents the results of using the Unsupervised Anomaly Detection methods to analyze the signal of the rotation velocity irregularity of the armature shaft in the traction electric motor of a locomotive. When analyzing signals corresponding to faulty electric motors, anomalous components were identified. As the example of analyzing the signal of an electric motor with an increased radial clearance, the possibility of detecting the development of a malfunction at an early stage has been confirmed. The conducted research confirms the possibility of using anomaly search methods to control the technical condition of the traction electric motor in a locomotive during bench tests.
**KEY WORDS:** *technical monitoring, electric motor, unsupervised anomaly detection, DBSCAN, OneClassSVM, Elliptic Envelop, Isolation Forest*

## 1. Introduction

Monitoring the technical condition of a vehicle is a process of continuous supervision of its technical condition and processes occurring in operation. At the same time, monitoring can be implemented both continuously during operation using built-in (on-board) control systems, and periodically during bench tests, equipment performance testing, etc. Specifics of using intelligent technologies when choosing locomotive maintenance strategies are given in [1].

The main advantage of monitoring is the ability to detect an incipient malfunction early without exclusion of the locomotive from service. The use of information obtained from monitoring tools in the planning of repairs and the development of a maintenance system makes it possible to avoid the contradictions set forth in the works [2, 3]. The main contradictions are the insufficient amount of collected statistics, the constant change in operating conditions, the inadmissibility of failure occurrence of critical objects, and the reduction of the service life for the controlled object in the event of a failure. Thus, during the operation of the locomotive, monitoring its technical condition makes it possible to increase the efficiency of the maintenance system without the introduction of more complex diagnostic systems. Processing the results of monitoring the technical condition of locomotive nodes is a simpler task than diagnosing. At the same time, the results of monitoring are the initial data for analysis by diagnostic systems. In this regard, the introduction and improvement of methods for analyzing data from monitoring systems is an actual task.

## 2. Literature Review

The application of monitoring to control the technical condition of locomotives is more relevant for locomotives equipped with on-board diagnostic systems [4], as it can identify hidden faults and monitor the state of locomotive nodes in all operating modes. It is also possible to use monitoring during tests and break-in tests of nodes and assemblies of locomotives.

The main task of monitoring is to identify anomalies in the operation of equipment. An anomaly in the operation of the equipment is the appearance of unexpected values (patterns, ratios of values, etc.) in the data set [5]. The concept of an anomaly in data is closely related to such concepts as Novelty and outlier. The difference between these concepts is considered in detail in [6]. The difference between novelty and anomaly is that an anomaly is the appearance of non-typical features in the behavior of equipment during typical operating conditions. Novelty is the appearance of new signs in the behavior of equipment when operating conditions, external factors, or any other reasons change. Novelty, as a rule, manifests itself as a result of a fundamentally new behavior of the object. At the initial stage, the values corresponding to the novelty will be considered as an anomaly, later, with the help of experts in the maintenance of this type of equipment, the classification of the values is either an anomaly or a normal state.

Researchers in [5-7] distinguish three types of anomalies: point, contextual, and group. A point anomaly when one or more measurement results differ significantly from the rest of the measurements. As a rule, such anomalies arise as a result of failures in the work of measuring instruments. For example, a one-time signal loss is the result of the impact of external electrical interference. Detecting this type of anomaly is easily done with the simplest methods. To eliminate the impact of point anomalies on the operation of recognition algorithms, the measurement results are pre-processed by filtering algorithms (cleared). The complexity of data cleaning is that it is necessary to prevent the data deletion characterizing the development of a malfunction.

Contextual anomalies are those that manifested in the discrepancy between the values of interrelated parameters. The fact of an anomaly is determined not by the value of the sensor signal, but by the correspondence of the sensor signal to the operating conditions of the equipment at the given moment. For example, the temperature of the armature winding of a traction electric motor may be 120°C at a certain moment, while after one second the temperature of the winding cannot be 60°C. The discrepancy lies in the rate of temperature change. At the same time, values of winding temperatures of 60 and 120°C are permissible from the viewpoint of boundary exceedances.

The collective anomaly manifests itself in interconnected data. These are measurement results that are considered to be anomalies when analyzed simultaneously with other data points. The deviation of the temperature value of one axle box unit from the temperature of the other axle box units of the locomotive in operation is an example of a collective anomaly. Similarly, the deviation of the cooling air temperature at the outlet of one of the electric traction motors from the air temperature at the outlet of the other motors can be considered an anomaly.

Currently, many methods and algorithms for anomaly search for various types of data have been developed [8-11]. Anomaly search methods are widely used in various branches of science and technology.

Let us consider the experience of using anomaly search methods for monitoring the technical condition of vehicles and equipment in the industry. According to [8], the search for anomalies in the data set that have been received from the sensors installed on the equipment is a classic task, which is solved using model methods (Digital twin) [9], machine learning methods [10], statistical analysis methods [11], neural networks [6] and a number of others. Both parameter values without additional transformations and values after additional processing can be used as initial data for anomaly search. An example of additional processing is the use of methods of data dimensionality reduction to simplify the process of interpreting the results [12, 13]. In addition to using sensor signals as control values, the concept of the index for the equipment technical condition (equipment health index) is pointed out in a number of studies [14, 15]. Using the index for the equipment technical condition allows you to control it in a generalized form, without being bound to specific parameters.

## 3. Methodology

In this article, we will consider the application possibility of using anomaly search methods to control the technical condition of the locomotive electric traction motor at bench tests. The signal of the rotation velocity sensor of the armature shaft is used as a controlled parameter. A technique for diagnosing an electric traction motor by the rotation velocity irregularity is given in [16]. Diagnosis of rotation velocity irregularity is performed using a high-precision incremental optical sensor of angular displacement (encoder). The sensor is installed on the technological cover of the bearing shield from the side of the collector. Using an adapter, the encoder shaft is rigidly connected to the armature shaft. The number of pulses (resolution) during measurements is 625 pulses per one complete revolution of the shaft. By processing the signal from the incremental sensor, we obtain information about the current value of the angle of a shaft rotation relative to the reference index mark (by the method of chained additions), as well as about its angular velocity. Diagnosis by the armature rotation irregularities is carried out when the electric motor is idling. Such a diagnostic system is available and does not require significant capital investments for its implementation.

The encoder signal is a time series with a frequency of 625 points since 625 values of the instantaneous rotation velocity of the armature correspond to each revolution of the motor armature. The methodology of time series analysis is used to analyze data on the technical condition of equipment and detect anomalies [17, 18]. The use of methods for working with time series is quite effective if there are measurement results gathered over a significant period of time, which makes it possible to trace the development of anomalies from a good state to a possible failure. The purpose of this publication is to demonstrate the possibilities of using unsupervised anomaly search methods (unsupervised learning) for signal analysis of the armature rotation irregularities in the traction electric motor. For processing, the results of tests for four traction electric motors in various technical conditions were used.

The most common anomaly detection methods using machine learning approaches, according to [8], are: Support Vector Machines for one class (OneClassSVM), Isolation Forest, and Ellipsoidal Data Fitting (Elliptic Envelope). The clustering algorithms DBSCAN, k-NN, and a number of others are also successfully used to detect anomalies. All methods described are unsupervised learning methods. The task of all anomaly detection methods is to answer the question of which class the analyzed data belongs to: normal or anomalous.

The task of the OneClassSVM method [19] in the classical version is to determine the hyperplane that divides the measurement results of the normal and anomalous classes. If, as a result of measurements, m points in n-dimensional space are obtained.

$$x_1, x_2, \ldots x_m \in R^n \,. \tag{1}$$

Each of the points describes the state of objects belonging to one of the classes, and the coordinates of the points in space act as signs. The result of applying the OneClassSVM method is a function that allows you to divide classes from each other. With a linear separation of space $R^n$, the task of the method is to calculate the parameters of the hyperplane that divides the anomalous and normal classes:

$$\langle w,x\rangle - b = 0, w \in R^n, b \in R, \tag{2}$$

where w is the vector, normal to the hyperplane, $x$ is the point of the hyperplane, $b$ is the real number constant.

If for some point $p \in R^n$ inequality is satisfied (3), then it is considered that point p belongs to the first class, if the condition (3) is not satisfied, then the point belongs to the second class:

$$\langle w,p\rangle - b \geq 0. \tag{3}$$

Thus, the task of the OneClassSVM method is to construct a linear classifier, which is specified by a vector $w \in R^n$ and a number $b \in R$.

Modern implementations of the OneClassSVM method allow using not only a linear space division but also a polynomial one, the Gaussian RBF function. Examples of using the OneClassSVM method to determine anomalies are described in [19, 22].

Isolation Forest is the unsupervised learning belonging to the family of decision trees. The concept of using the method is that anomalous values are closer to the root of the decision tree because anomalous values are easier to distinguish from other ones. In line with [20], when processing the data flow from sensors, the data set Z is divided into a sequence of n-dimensional vectors:

$$Z = \{z(1), z(2), ...z(t), z(t+1), ...\}, \tag{4}$$

where $z(t) \in R^n, t \geq 1$.

First of all, the data set $Z$ is divided into blocks of equal length. In our case, each of these blocks consists of 625 points corresponding to one revolution. The main task of using the anomaly search method is to perform block analysis $z(t)$ and to single out anomalous values from this block. The complexity of singling out anomalous values requires the use of some criterion S by the value of which one or another value belongs to the class of anomalies. According to [21, 22], in the Isolation Forest, the condition is used (5):

$$S(x,n) = 2^{-\frac{E(h(x))}{c(n)}}, \tag{5}$$

where $h(x)$ is the path length of the observation, $E(h(x))$ is the average of values $h(x)$ from all sets of isolation trees. $c(n)$ is the average path length of an unsuccessful search in the binary search tree, n is the number of external nodes.

For each value from the block $z(t)$, the value of the criterion S is calculated. A value close to one indicates an anomaly, and a value significantly less than 0.5 indicates the absence of an anomaly. If all values of the $S$ criterion for values from the block $z(t)$ are close to 0.5, then the entire block does not have clearly defined anomalies.

The elliptic Envelope is based on an evaluation of the conformity of the data distribution to the normal (Gaussian) distribution law. The basic idea is that the distribution of the measured data is represented as a Gaussian distribution while singling out a group of unlikely data for the model. If something is unlikely within the model, then it is probably an anomaly. An ellipse is generated around the central data cluster and outliers are detected using the minimum covariance determinant. Mahalanobis distance is used as a metric (6):

$$d(x,\mu,C) = \sqrt{(x-\mu)^T C^{-1} (x-\mu)}, \tag{6}$$

where $x$ is input data, $\mu$ is the average value, C is the covariance matrix.

DBSCAN refers to unsupervised learning methods. The method performs clustering based on the packing density of data points. The result of the method is the selection of groups and clusters in the data set. The algorithm groups points together that are closely spaced (points with many close neighbors), marking those points as outliers that are lonely in areas of low density (the nearest neighbors are located far). The DBSCAN parameters are: minPts is the minimum number of points (threshold) grouped in place for the area to be considered dense; eps (ε) is the distance measure that will be used to define the location of points in the vicinity of any point.

The construction of an anomaly detection system consists of four stages: the collection of initial data, the systematization of the collected data, the formation of additional numerical features, and the construction of an object behavior model [23, 24]. In the simplest version, the stage of generating additional numerical features can be skipped.

Anomaly detection is performed using the behavior model of the monitored object.

**4. Results**

The results of measuring the rotation velocity of the armature shaft of four traction electric motors that were in various technical conditions were processed using the above-considered methods for detecting anomalies without a teacher. At the same time, the engine m1 was in good condition after the overhaul. The results of measuring the rotation velocity of the armature in this traction electric motor are used as a reference signal when training the anomaly detection models.

The m2 and m3 motors required the overhaul and had malfunctions. When disassembling faulty motors, the following defects and malfunctions were found: loosening in mount and, as a result, rolling the front pressure washer on the shaft, microcracks in the welded connection of the front pressure washer with the armature shaft (motor m2), surface corrosion of the pump bodies, exceeding the allowable radial clearance of the anchor roller bearings (motor m3). The motor m4 had a radial bearing clearance of 0.13 mm and 0.17 mm with a limiting value of 0.20 mm.

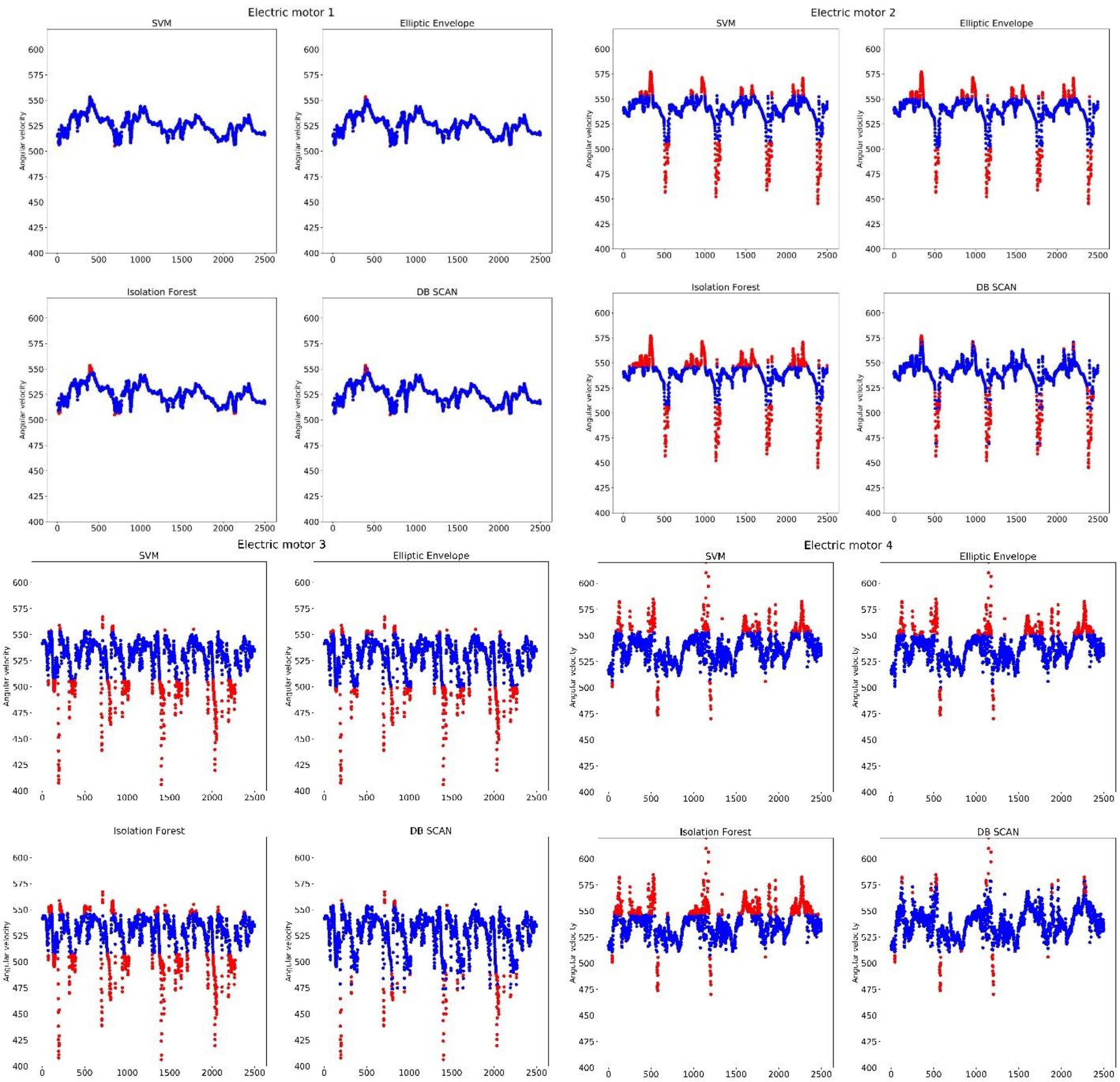


Fig. 1 Determination of anomalies in the signal of armature rotation irregularities (electric motors m1 - m4)

The results of processing the signal of the rotation velocity of the armature shaft in the traction electric motor by anomaly detection methods are shown in Fig. 1.

The values of the instantaneous rotation velocity of the armature (in red) by machine learning algorithms, are determined to be anomalous. As it can be seen from Fig. 1, all four considered methods single out anomalous velocity values and can be used to control the technical condition of the electric motor at monitoring. However, each of the methods has its own characteristics.

To evaluate the accuracy of models used in anomaly recognition, anomalies were searched for the signal corresponding to the m1 motor. As a result of processing the signal in m1 motor, the OneClassSVM algorithm provided an anomaly recognition accuracy of 99.9%, the Elliptic Envelope algorithm – 99.5%, the DBSCAN algorithm – 99.1%, the Isolation Forest algorithm – 98%. By recognition error, we mean the assignment of normal values to an anomalous class. The m2 and m3 electric motors had pronounced malfunctions. In the signals corresponding to these motors, all algorithms have identified a clear allowable range in changing the rotation velocity of the armature. The technical condition of the bearings in the m4 electric motor was approaching the red lines. The rotational velocity signal does not have such pronounced peaks as in the m2 and m3 motors, while the algorithms have identified anomalous areas in the signal waveform. After comparing the results in the work of algorithms, we also see that the OneClassSVM and Elliptic Envelope algorithms have a narrower range of normal values. The DBSCAN algorithm has the widest range of acceptable (normal values), concurrently, the DBSAN algorithm defines some of the values from the anomalous zone as normal values. Examples of such misclassification are displayed on the graphs corresponding to the m2 and m3 motors. A feature of the Isolation Forest algorithm, in comparison with others considered, is that this algorithm tends to determine the anomalous values of the upper part of the graph. An example of such a classification is visible in the images corresponding to the m2 and m4 electric motors.

## 5. Conclusions

The conducted research confirms the possibility of using anomaly search methods to control the technical condition of the locomotive traction motor during bench tests. The signal of the armature shaft velocity sensor was used as a controlled parameter. All the considered methods for detecting anomalies make it possible to determine the presence of a malfunction in the bearings and the mechanical part of the engine. The detection of anomalous components in the signal of the m4 engine, the radial bearing clearance of which was approaching the maximum allowable limit, allows us to speak about the possibility of early detection of a malfunction. The considered methods differ in different levels of sensitivity, the width of the range of permissible deviations of the signal. The use of the considered methods for detecting anomalies makes it possible to determine the presence of deviations from the norm in the technical condition of the controlled unit or unit, while the considered methods do not allow determining the cause of the deviation. To determine the cause of the deviation, it is necessary to perform additional processing of the speed signal in order to highlight additional information features in the signal, as well as apply supervised machine learning methods.

## References


1. **Ochkasov, O.; Shvets, O.; Cerniauskaite, L.** 2017. Usage of intelligent technologies in choosing the strategy of technical maintenance of locomotives, Technologijos ir Menas = Technology and Art 8: 68-71.
2. **Resnikoff, H.L.** 1977. Mathematical aspects of reliability-centered maintenance, United Air Lines Inc San Francisco Ca.
3. **Moubray, J.** 2001. Reliability-centered maintenance, Industrial Press Inc.
4. **Bodnar, B.; Ochkasov, O.** 2017. System Choice of the Technical Maintenance of Locomotives Equipped with on-Board Diagnostic Systems, Transport Means: Proceedings of 21st International Scientific Conference Part I: 43-47.
5. **Shkodyrev, V.P.; Yagafarov, K.I.; Bashtovenko, V.A.; Ilyina E.E.** 2017. The Overview of Anomaly Detection Methods in Data Streams, Proc. of the Second Conference on Software Engineering and Information Management 1864: 7-11. (In Russ.)
6. **Paula Almeida Ribeiro, R.; Pereira, P; Gama, J.** 2016. Sequential anomalies: a study in the Railway Industry, Machine Learning 105(1): 127-153. doi:10.1007/s10994-016-5584-6
7. **Hayes, M.A.; Capretz, M.A.** 2015. Contextual anomaly detection framework for big sensor data, Journal of Big Data 2(2). https://doi.org/10.1186/s40537-014-0011-y
8. **Diakonov, A.G.; Golovina A.M.** 2017. Detection of anomalies in the work of mechanisms by machine learning methods, Proc. XIX Int. Conf.: Analytics and data management in areas with intensive data use, 469-476 (in Russ.)
9. **Venkatesan, S.; Manickavasagam, K.; Tengenkai, N.; Vijayalakshmi, N.** 2019. Health monitoring and 11 prognosis of electric vehicle motor using intelligent- digital twin, IET Electric Power Applications 13(9): 1328-1335. doi: 10.1049/ietepa.2018.5732
10. **Munir, M.; Siddiqui, SA; Chattha, MA; Dengel, A; Ahmed, S.; Fuse, A.D**. 2019. Unsupervised Anomaly Detection in Streaming Sensors Data by Fusing Statistical and Deep Learning Models, Sensors 19(11): 2451. https://doi.org/10.3390/s19112451
11. **Zope, K.; Singh, K.; Nistala, S.H.; Basak, A.; Rathore, P.; Runkana, V.** 2019. Anomaly Detection and Diagnosis In Manufacturing Systems: A Comparative Study of Statistical, Machine Learning And Deep Learning Techniques, Annual Conference of the PHM Society 11(1). https://doi.org/10.36001/phmconf.2019.v11i1.815
12. **Bodnar, B.; Ochkasov, O.; Bodnar, Ye.; Hryshechkina, T.; Keršys, R.** 2018. Safety Performance Analysis of the Movement and Operation of Locomotives, Transport Means 2018: Proc. of the 22nd Intern. Sci. Conf II: 839-843.
13. **Bodnar, B.; Ochkasov, O.** 2021. Devising a procedure to form the diagnostic parameters for locomotives using a principal components analysis, Eastern-European Journal of Enterprise Technologies 2(1(110)): 97-103. DOI: 10.15587/1729-4061.2021.230293.

14. **Jürgensen, J.H.; Godin, A.S.; Hilber, P.** 2017. Health index as condition estimator for power system equipment: a critical discussion and case study, CIRED 1: 202-205. DOI: 10.1049/oap-cired.2017.1174
15. **Bodnar, B.; Ochkasov, O.; Ochkasov, M.** 2021. Devising a procedure for calculating the technical condition index of locomotive nodes based on monitoring results, Eastern-European Journal of Enterprise Technologies 5(3(113)): 37-45. DOI: 10.15587/1729-4061.2021.242478.
16. **Bodnar, B.Ye.; Ochkasov, A.B.; Chernyayev, D.V.; Shevchenko, Ya.I.** 2013. Diagnosis of Traction Electric Motor at Irregularity in Speed of Anchor Rotation. Science and Transport Progress, Bulletin of the Dnipropetrovsk National University named after Academician V. Lazaryan 3(45): 13-21. DOI: 10.15802/stp2013/14793.
17. **Weldezgina Asres**, **M. et. al** 2021. Unsupervised Deep Variational Model for Multivariate Sensor Anomaly Detection, 2021 IEEE International Conference on Progress in Informatics and Computing (PIC), 364-371, doi: 10.1109/PIC53636.2021.9687034
18. **Ding, N.; Gao, H.; Bu, H.; Ma, H.; Si, H.** 2018. Multivariate-timeseries-driven real-time anomaly detection based on bayesian network, Sensors 18(10): 3367.
19. **Sch¨olkopf, B.; Williamson, R.C.; Smola, A.J.; Shawe-Taylor, J.; Platt, J.C.** 1999. Support Vector Method for Novelty Detection, NIPS Proceedings 12: 582-588.
20. **Ding, Z.; Fei, M.** 2013. An Anomaly Detection Approach Based on Isolation Forest Algorithm for Streaming Data using Sliding Window, IFAC Proceedings Volumes 46(20): 12-17. doi.org/10.3182/20130902-3-CN-3020.00044.
21. **Liu, F.T.; Ting, K.M.; Zhou, Z.** 2008. Isolation Forest, 2008 Eighth IEEE International Conference on Data Mining, 413-422, doi: 10.1109/ICDM.2008.17.
22. **McKinnon, C.; Carroll, J.; McDonald, A.; Koukoura, S.; Infield, D.; Soraghan, C.** 2020. Comparison of New Anomaly Detection Technique for Wind Turbine Condition Monitoring Using Gearbox SCADA Data, Energies 13(19):5152. https://doi.org/10.3390/en13195152
23. **He, P.; Zhu, J.; He, S.; Li, J.; Lyu, M.R.** 2016. An evaluation study on log parsing and its use in log mining, in DSN’16: Proc. of the 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 654-661. doi: 10.1109/DSN.2016.66
24. **Zwietasch, T.** 2014. Detecting Anomalies in System Log Files using Machine Learning Techniques, University of Stuttgart. http://dx.doi. org/10.18419/opus-3454