

Modern Machine Learning Methods for Telemetry-Based Spacecraft Health Monitoring

P. A. Mukhachev^{1*}, T. R. Sadretdinov^{1**}, D. A. Pritykin^{1***},
A. B. Ivanov^{1****}, and S. V. Solov'ev^{2*****}

¹*Skolkovo Institute of Science and Technology, Moscow, 121205 Russia*

²*Korolev Rocket and Space Corporation Energia, Korolev, Moscow oblast, 141070 Russia*

e-mail: **petr.mukhachev@skoltech.ru*, ***tagir.sadretdinov@skoltech.ru*, ****d.pritykin@skoltech.ru*,
*****anton.ivanov2@skoltech.ru*, ******sergey.soloviev@scsc.ru*

Received December 4, 2020; revised January 4, 2021; accepted March 16, 2021

Abstract—We survey the progress in data mining methods for spacecraft health monitoring. The main emphasis is placed on the analysis of telemetry data enabling the identification of spacecraft states that are atypical during normal operation and the prediction of possible failures in the operation of the spacecraft or its components. The main stages required for the creation of general-purpose spacecraft state monitoring systems are considered; methods for detecting anomalies in telemetry data taking into account the specific features of the spacecraft are presented in detail; and publications on this topic known to the authors are analyzed. Examples of the implementation of such systems in flight control centers of various countries are given. The promising areas of development of methods for analyzing the technical state of complex systems relevant for solving problems in space technology are discussed, and the main factors that hinder the development of machine learning methods for analyzing telemetry data are noted.

Keywords: data mining, anomaly detection, flight control, technical diagnostics, telemetry data

DOI: 10.1134/S0005117921080014

1. INTRODUCTION

Spacecraft and their components are designed considering a large number of factors, including the flight plan, environment, interaction with other systems, and controls available. However, the inevitable design, modeling, and verification errors prevent one from creating accurate models of each of the components and the spacecraft as a whole as well as of all of its possible failures in actual space flight. Thus, various unforeseen situations leading to component malfunction, as well as situations that were not analyzed when creating the spacecraft and making the flight plan, may arise in spacecraft operation.

The general approach to fault detection, isolation, and recovery (FDIR) is designed to ensure the uninterrupted performance of spacecraft functions [1]. This approach involves ensuring redundancy, using automatic control systems that carry out diagnostics and start recovery procedures, and also thoroughly studying the identified failures and unforeseen situations and adjusting the flight plan appropriately.

Many authors note that at present the main method for spacecraft health monitoring is threshold value monitoring and constant monitoring of key telemetry parameters by the operator [2, 3]. However, specialist's time is expensive, and only a few telemetry parameters can constantly be in his/her field of vision. At the same time, the number of telemetry parameters in spacecraft often reaches several thousand. For example, GOCE spacecraft had 1300 telemetry parameters, and the number of telemetry parameters in larger spacecraft can reach 50 thousand [4]. This makes it impossible for

the operator to control the parameters manually. Generally speaking, threshold value monitoring does not permit one to detect the onset of an emergency situation or abnormal spacecraft behavior. With the development of projects for the deployment of mega-constellations, the automation and intellectualization of satellite health monitoring becomes even more topical [5, 6].

Practice shows that atypical behavior of telemetry parameter values is of interest in that it often precedes failures in spacecraft component operation. This abnormal behavior is usually not detected by standard threshold monitoring methods. Anomaly detection in telemetry is a very difficult task. As a rule, there is not enough historical information about anomalies, and a failure is not always preceded by the appearance of anomalies. Anomalies can be caused by random failures in the operation of sensors or the information transmission channel and by atypical situations, which may be either harmless or potentially dangerous for the spacecraft operation; anomalies do not have a clear threshold, the nominal state can evolve during the spacecraft operation, and the concept of anomaly is different for different components. Nevertheless, detecting anomalies in telemetry is important for the subsequent diagnostics and analysis of the spacecraft operation.

There are two main approaches to anomaly detection. The first approach assumes the presence of a quantitative model explicitly describing the system operation. In this case, the discrepancy between the behavior of the system and that of the model can serve as an indicator of a malfunction. On the one hand, this approach permits explicitly modeling faults and their consequences. On the other hand, it has the obvious disadvantage that creating the underlying models is difficult. The second approach is to use machine learning methods, which help one to create system operation models based on available data. The importance of developing such systems for spacecraft health analysis is indicated in [3, 6]. Spacecraft diagnostic expert systems based on such models can be developed provided that there is a broad database available. The present paper gives a survey of the recent progress in the application of machine learning methods to spacecraft health analysis.

Many recent survey papers covering both general and more narrow topics are devoted to anomaly detection in various fields. Of these surveys, we note the papers by Chandola [7], Pimentel [8], and Wang [9], which group the methods according to the main assumptions about the nature of the data to be analyzed, estimate the computational complexity of the algorithms, and list their main practical applications. Chalapathy et al. [10] take a closer look at methods based on artificial neural networks and deep learning. Zimek [11] and Thudumu [12] discuss the specific features of very high-dimensional data analysis. Gavrilovski [13], Khan and Yairi [14], and Basora [15] survey the applications, including those in the aviation industry, of anomaly detection methods.

Despite above-mentioned general surveys and aviation application surveys, to the best of the author's knowledge, there is no review paper on space applications of data-driven anomaly detection. The present survey bridges the gap, takes into account the specific features of space technology and possible methods for telemetry information analysis, singles out directions necessary for further progress, and indicates the most promising methods.

The survey is divided into eight sections. Section 2 describes the main definitions and problem statements from the viewpoint of data analysis. Various feature extraction methods for algorithm performance improvement are discussed in Sec. 3. Section 4 is devoted to the classification of anomaly detection methods and basic examples of their usage. Section 5 discusses how decisions are made to classify a system state as abnormal if the data analysis methods used do not classify the states but map them onto a contiguous interval of the real line. Section 6 describes methods for determining the performance of data-driven algorithms. Section 7 discusses the use of data-driven systems in various flight control centers on an ongoing basis, lists open data sources that can be used to test the operation of some algorithms, and discusses promising data analysis methods. Finally, Sec. 8 summarizes the main results of the paper.

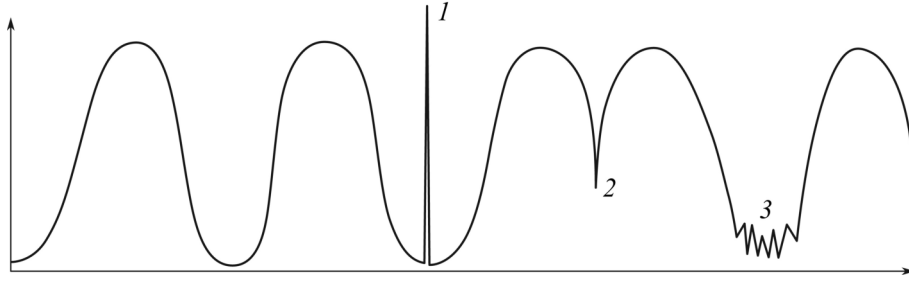


Fig. 1. Anomaly types: (1) point, (2) context, (3) collective.

2. MAIN NOTIONS

2.1. Anomalies and Outliers in Telemetry Information

In the literature, the terms *outlier* and *anomaly* are often used in the same sense, and there is no clear distinction, because researchers from different fields use different terminology. However, as noted by Pimentel [8], outliers are often used to mean a small fraction of normal data that is far from most truly normal data. Thus, it is the isolation of data that can significantly degrade the performance of some algorithms that the search for outliers is focused on.

Some researchers use the term “*new information*” (novelty) to refer to data not fitting a model based on data a priori considered to be normal. This class will no longer include random outliers described in the previous paragraph if they are present in the model of “normal” data.

Many authors divide anomalies into point, context, and collective ones [7]. Point anomalies are characterized by a single value outlying from the neighboring values and are easiest to detect. Context anomalies are situations whose occurrence depends on some external context or other variables without which they would not be considered anomalous. Collective anomalies are characteristic of time series or other ordered data. They occur if several nearby points exhibit abnormal behavior but cannot be considered abnormal individually (Fig. 1).

Telemetry information is the readings of spacecraft onboard sensors and software events circulating in the control loop and transmitted to a ground control center. The general structure of information flows in the spacecraft control loop is given in [16]. Sensor readings are most often structured into time series, which can be uneven, i.e., with a variable time step; sometimes, step can be different for different telemetry parameters. Telemetry is the main source of information about the spacecraft state, and in this sense, telemetry is most complete and reliable. Based on telemetry analysis, conclusions are drawn on the health of the spacecraft and its components and further flight planning decisions are made. In what follows, we focus on the analysis of telemetry presented in the form of multidimensional time series and do not consider program event analysis.

Telemetry is received and processed by a monitoring system implemented by onboard and ground systems. This system implements various algorithms for extracting useful information about the spacecraft state and further, if necessary, making a corrective action or alerting the operator to make a decision about such an action.

2.2. Possible Settings

The following large classes of algorithms for analyzing the state of technical systems can be distinguished: data-driven and model-based. Model-based algorithms are characterized by the fact that they require an explicit scheme created by the developer in advance and, unlike data-based algorithms, do not have the broad ability to adapt based on historical information about the system [17]. Data-driven approach, in turn, requires a sufficiently large body of historical information to select the coefficients of the model (training) and validate the correctness of its operation.

In general terms, the technical diagnostics problem can be stated as follows. For brevity, we introduce the following notation: $Z(t)|t = t_0$ is the state of the system at a certain moment in time, $Y = y(Z(t))$ is some discrete set of classes for assessing the performance of the system or such an assessment on a continuous scale, and $X(t) = f(Z(t)) + \xi$ is the set of visible observations of readings of various sensors. As a rule, the function f is known with some accuracy according to the specification of the sensor in question, but this is not always the case. The vector ξ is the set of residual measurement errors. The general setting of the problem is to reconstruct the value of Y based on X , i.e., to find the most accurate approximation to some function $g : Y = g(X)$. The function f is not always invertible, and the errors ξ cannot always be accurately described; therefore, the problem has no rigorous solution in general.

In what follows, we discuss data-driven methods in detail. They are of great practical importance, because they can be applied to systems in which a formal description of the operation encounters significant difficulties. The most general statement of the machine learning problem can be given as follows [18]:

- Supervised learning. Let some training set $D = \{(x^n, y^n), n = 1, \dots, N\}$ be known. The task is to reconstruct a function f such that the prediction $f(x^*) = y^*$ is accurate on a new instance of (x^*, y^*) in the test sample. In this case, the accuracy criterion is determined by the given error function.
- Unsupervised learning. Let some dataset $D = \{x^n, n = 1, \dots, N\}$ be given. The task is to describe this data in some more compact way determined by some objective function.

The collection of training and test samples is called a dataset. The collection of vectors y^n is called a markup. A supervised learning problem will be called *classification* if the range of y^n is discrete and *regression* if the range of y^n is continuous.

Since in practice the true spacecraft state is often unavailable for direct observation, it is usually assumed that the sensor readings accurately determine the state, and the tasks of monitoring the technical state are reduced to the following settings:

1. Classification of states and their sequences. In this case, $Y \in \mathbb{N}$ is discrete, and there is a certain number of observations $Y = g(Z(t_i))$.
2. Anomaly detection. In this setting, $Y = \{0, 1\}$ and the observations of only one of the two classes are known, $Z : g(Z) = 0$.
3. Outlier selection. In this case, it is necessary to estimate the value of $g(Z)$ on the basis of observations of X alone in the absence of observations of Y .
4. Ranking of results by importance (degree of abnormality), also called scoring. This problem is characterized by a continuous range of values $Y \in \mathbb{R}$.
5. Forecast of the spacecraft state: finding the values of the telemetry parameter $X(t_1 + n)$, assessing the operation performance $Y = y(X(t_1 + n))$ or directly the system state $Z(t_1 + n)$ based on the history of telemetry parameter values $\{X(t) : t_0 \leq t \leq t_1\}$.
6. Estimation of the remaining useful life of the spacecraft or its parts is similar to the previous setting; however, in this case, as a rule, a longer-term forecast is of interest.

In the outlier detection problem, the algorithm is given an unlabeled dataset in which it is required to isolate unlikely states or their sequences; i.e., this setting is an unsupervised learning problem. However, these methods can produce a large proportion of false positives and false negatives in anomaly detection, because not every anomaly is an outlier and not every outlier can be considered an anomaly. Many authors note that this statement is mainly used to get rid of random rare noises and spikes, since in this case the initial data contain very little information about the correct spacecraft operation.

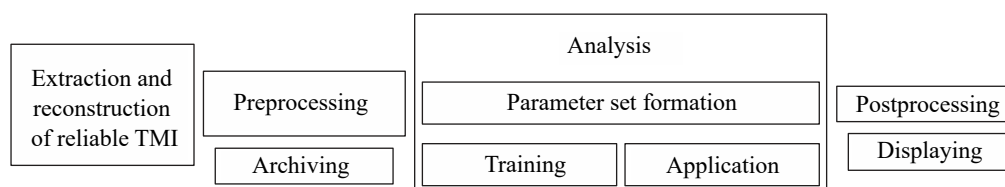


Fig. 2. Main stages of telemetry analysis.

In anomaly detection, this setting can be used to localize the found anomaly or to provide the operator with more detailed information about the normal spacecraft operation. As a rule, this setting uses supervised learning methods. Supervised learning consists in finding an approximation to an unknown function from the set of its observed values. The classification problem in this case is to find a function with a discrete range of values.

In the classification problem, the training set X consists of prelabeled data: each state or situation X_i belongs to one or more of separate classes Y_i . In the process, the algorithm assigns each new situation to a class to which it belongs with the greatest probability.

In anomaly detection, the training sample, as a rule, consists only of representatives of one class—the normal operation of the system. In some setting, the training sample may contain a relatively small number of representatives of anomalous operation. During operation, the algorithm detects situations that are not similar to those encountered in the training set.

In addition, algorithms capable of active learning are beginning to emerge. These algorithms are trained in a semisupervised mode, i.e., when a data markup is not available for all X_i , and are able to prompt the user for markup when needed [19, 20]. Such algorithms are suitable for very large databases, when complete manual markup is not possible when creating the dataset. This class of algorithms appeared relatively recently and requires further study.

2.3. Main Stages in Telemetry Data Analysis

Despite the recent progress in the creation of mathematical apparatus and algorithms for data mining, a designer still has to choose an algorithm and adapt it to the task at hand and its features. The ASUM-DM procedure [21] identifies six main stages in a data project: analysis of topical area; development; creation and customization of a working model; commissioning; operation and optimization; and project management.

When creating the working model and adjusting its parameters, it is important to know the objective performance of its operation in comparison with other models and the operator. To assess the performance, a reference data markup is required, which is often absent in real-world problems. Usually, in the absence of an objective criterion for the performance of a model or a data-based algorithm, the justification for the correctness of its operation is speculative. This phenomenological principle of verification is fundamentally different from the verification of model-based systems, where the correctness of operation is shown by the correctness of the initial assumptions and logical inferences.

At the stage of operation of the monitoring system (Fig. 2), the analysis is usually preceded by the isolation and, if mechanisms are available, reconstructing reliable telemetry information (screening out random outliers) and changing the sampling frequency for which knowledge about the characteristics of the sensor and the nature of the measured quantity is required. After that, an analysis is performed using various algorithms to be discussed below. The analysis can also be divided into several stages:

1. Formation of a set of parameters (features) with which the data mining algorithm will work directly.

2. Building or refining the model.
3. Applying the model to obtain the result.

The features can be extracted both by experts and automatically; however, automatic feature extraction, as a rule, requires a larger amount of training data, and the constructed model is more difficult to interpret. In this case, the stage of feature extraction is not explicitly expressed as is the case during the construction and application of algorithms. This, for example, happens in the case of using deep neural networks.

The last stage is postprocessing the results of the analysis to generate reports and display them on the operator's screen.

3. FORMING THE COLLECTION OF PARAMETERS FOR ANALYSIS

In the context of machine learning problems, any measurable or computable characteristics of the considered object that help in solving the problem are usually called features. For example, when determining the current spacecraft state, the telemetry parameter values, the space weather parameters, etc. can be considered as features. In complex problems, it is important to single out auxiliary features that are some functions of the original ones and help to improve the result and/or reduce the computational complexity. With automatic feature extraction, as is the case with the use of deep neural networks, features can be devoid of specific physical meaning and be only a mathematical function of the initial data.

Feature extraction is not an independent method for anomaly detection or spacecraft health analysis, but it is of great importance, especially when analyzing multidimensional time series in complex systems, where there is a large number of related parameters due to sensor redundancy or parameters that are not related to studied anomalies and thus introduce noise to the data and make it difficult to find them. The correct feature extraction from the initial data can greatly facilitate anomaly detection and improve the performance of the algorithm and its interpretability. In most cases, the choice of telemetry parameters according to which the condition will be monitored is done by an expert.

Suo et al. [22] introduced the concept of fuzzy Bayesian risk to select a subset of variables for further classification. In [22], they have shown that when using a certain kind of loss function, it is possible to construct a greedy algorithm for choosing a subspace in which the separation of classes from each other will be optimal in terms of this metric. The computational complexity of the algorithm is $O(knm^2)$, where k is the number of selected telemetry parameters, n is the number of classes, and m is the number of points. The authors used the marked-up results of the numerical simulation of an energy supply system to demonstrate the algorithm and compare it with the algorithms in ReliefF [23], mRMR [24], NRS [25], etc.

For the analysis of periodic signals, especially high-frequency ones, the discrete Fourier transform and the wavelet transforms are often used [26] as a signal representation. When analyzing multidimensional time series, dimensionality reduction techniques such as principal component analysis and factor analysis, as well as expert selection and the creation of informative features, are of great importance, since they can significantly reduce the computational resources required for the analysis.

Barreyre [27] carried out a large-scale research of the anomaly detection performance depending on the selected features on univariate time series. The authors of the present paper studied the features identified using the Fourier transform, wavelet transforms, Gaussian kernel eigenfunctions, principal component analysis, periodograms, and anomaly detection methods based on distance and local density (Sec. 4.1) as well as the support vector machine (Sec. 4.4). The authors used artificially generated data as well as real data to validate the results (1700 telemetry parameters

collected over 12 years). The article provides practical conclusions about the methods and features used.

On the other hand, O'Meara and Schlag [28] have shown that the use of manually specified features based on the physics of the observed process can significantly improve the quality of predictions and anomaly detection. The same is concluded by Gowda in [29]. "Feature extraction is almost critical. It is more important than the use of one or another machine learning model," the author writes. Gowda used telemetry parameter data from Mars Express [30] to build a predictive consumption model of a temperature control system (TCS).

4. METHODS FOR TELEMETRY DATA ANALYSIS AND APPLICATION EXAMPLES

Many publications [7, 8, 15] propose the following algorithm taxonomy, which is based on basic assumptions about normal and anomalous data:

1. Distance-based, including
 - (a) based on the analysis of a neighborhood graph.
 - (b) clustering-based.
2. Static, including
 - (a) parametric ones:
 - a mixture of distributions.
 - stochastic processes.
 - (b) nonparametric ones.
3. Classification-based.
4. Reconstruction-based.
5. Ensemble-based.

Figure 3 shows the classification and typical examples of the algorithms in each class. Note that despite the fact that the presented classification includes many algorithms, it cannot be viewed as complete and taking into account all aspects of existing algorithms.

Most of these methods are not specifically designed for time series analysis; however, they can be adapted by appropriate choice of features or applied to individual states without using the sequential time series structure.

In addition, in many systems, only some of telemetry parameters carry useful information for finding certain types of failures. For example, the telemetry parameters of one of the spacecraft components will carry little information about the performance of the other components. Therefore, to reduce the potentially required resources and increase the search accuracy, various feature selection techniques can be used (see Sec. 3).

When choosing an algorithm and its parameters, it is necessary to use an objective criterion for evaluating the algorithm performance as well as compare it with the performance of the operator's work (see Sec. 6).

For convenience, the articles reviewed here are collected in the [Appendix](#).

4.1. Methods Based on Neighborhood Graph Analysis

Methods based on nearest neighbors are based on the hypothesis that points characterizing normal states are located close to each other, while abnormal data are far from normal in a certain space of features. Initially, the distance to the K th nearest neighbor was used as an anomaly score. Methods in this group differ from each other in that they differently define the concept of neighborhood.

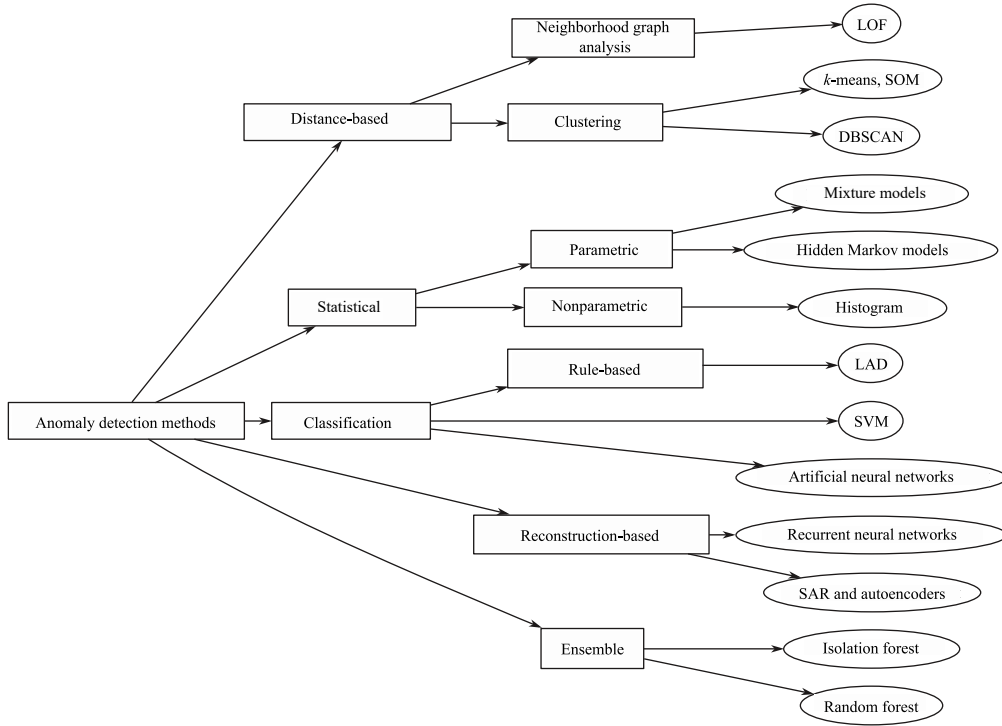


Fig. 3. Classification of algorithms used to find anomalies, and some examples of corresponding classes.

Bay and Schwabacher [31] proposed the ORCA method, showing that in some cases the nearest neighbor method can be optimized to almost linear scaling by partial truncation of the search.

This class also includes methods that use local density of points. The Local Outlier Factor (LOF) method [32] is based on calculating the local density around a point and around its neighbors,

$$LOF_{kNN}(p) = \left(\sum_{o \in kNN(p)} \frac{lrd_{kNN}(o)}{lrd_{kNN}(p)} \Big/ |N_{kNN}(p)| \right)^{-1}.$$

The LOF value is the reciprocal of the local reachability density of a point relative to the average local reachability density of its neighbors, where the local reachability density

$$lrd_{kNN}(p) = |N_{kNN}(p)| \Big/ \sum_{o \in kNN(p)} RD_{kNN}(p, o)$$

is the reciprocal of the reachability distance $RD_k(p, o) = \max(d_k(o), d(p, o))$ of a point averaged over k of its nearest neighbors. Note that $N_{kNN}(p)$ can be greater than k if there are several points equidistant from the point p .

The value $LOF_{kNN}(p)$ can be interpreted as the isolation of a considered point p in comparison with the isolation of its k nearest neighbors. For this method, it is necessary to empirically set the number of nearest points that are considered to be neighbors of the considered one. The local correlation integral (LOCI) method works in a similar way and partially solves the problem of manually specifying the number of neighbors, replacing this parameter with a distance within which points are deemed neighbors. The k th neighborhood relationship is generally not symmetric, so many methods use the inverse and symmetric neighborhood relationships to determine outliers.

Von Brünken [33] reworked the LOF method, suggesting to use a local estimate of the intrinsic dimension of the data [34, 35] as a density estimate. The new method was coined as IDOS (Intrinsic Dimension Outlier Score) and showed very good results in comparison with other modern methods—LOF, SOD [36] and FastABOD [37]—on various datasets, while requiring less computing resources.

Some algorithms also use kernel-based methods to estimate the density [38], as it happens in statistical methods to estimate the distribution density. For example, one of the recent publications [39] proposes to estimate the outlier score of a point using a kernel probability density estimate based on the neighbors of the point in question.

Most often, the Euclidean distance is chosen for these methods in a certain space of features. However, there are other distance options, such as the Manhattan distance, or, equivalently, the L_1 metric, or the dynamic time warping (DTW) [40] that is used to measure the distance between time series.

The advantages of neighborhood graph analysis methods are as follows:

- The operation of these methods is relatively easy to interpret.
- Many algorithms operate without a supervisor and can be adapted to semisupervised learning.
- They are easy to adapt to new data types—to achieve this, it suffices to define a metric on them.
- Some modifications can be scaled well in terms of the time required for computations over a large-dimensional space.

The drawbacks of neighborhood graph analysis methods are as follows:

- These algorithms rely on a metric, so it can be tricky to deal with categorical variables and complex data types like graphs.
- They do not scale well for large amounts of data.
- They are difficult to adapt to time series.
- Many methods rely on distance, so their efficiency decreases in a high-dimensional space.
- If anomalous points have many neighbors of any type, we obtain a false positive result.

The main disadvantage of these methods is the need to compose the neighborhood graph, which requires calculating the distance between every two points in the dataset, and therefore most algorithms in this class scale as $O(N^2)$. This dependence does not allow these methods to be used effectively on large samples.

Application of neighborhood graph analysis methods

Based on the IDOS method, O’Meara et al. [41] created the ATHMoS (Automated Telemetry Health Monitoring System) used in the German Space Operations Center (GSOC). This system supplements the IDOS method with automatic selection of the point anomaly threshold value for a given base of nominal telemetry parameter values. The authors of [41] called the resulting method OPVID. The method permits one to analyze many subsystems and a large number of telemetry parameters without much of the operator’s intervention in the system training process. In the example considered in [41], the authors showed the ability of the algorithm to detect anomalies in the original data and compared the operation of the OPVID with the algorithm proposed by Martinez-Heras [42], having received much fewer false positives. Later, Schlag and O’Meara [43] carried out a detailed empirical verification of the performance of the OPVID on various datasets using both manually created features and an autoencoder. The OPVID demonstrated good performance compared to the LOF and the LoOP.

4.2. Clustering-Based Methods

Clustering in general is used to group similar items together. Most clustering algorithms are not specifically designed to search for outliers or anomalies; however, they can use the structure of the available data and, on its basis, detect anomalies according to a specific criterion. Often, these methods can be trained in an unsupervised or semisupervised mode. Chandola [7] distinguishes three classes of clustering-based anomaly detection algorithms based on hypotheses about the nature of the original data and potential anomalies in it.

1. Normal data belong to some cluster, while anomalies do not. In this case, we can use clustering algorithms that do not require each point belonging to a cluster, for example DBSCAN, ROCK [44], SNN (shared nearest neighbours) [45], and their modifications.
2. Normal data lie closely to one of found centroids, while the anomalous ones are located far from them. In this class, one can use methods that calculate the centroid of the cluster such as the method of k -means and self-organizing maps (SOMs) [46]. Clustering by the method of k -means is the most popular algorithm in this class.
3. Normal data belong to large dense clusters, while anomalies form small or sparse clusters. This subclass accommodates such algorithms as FindCBLOF [47] and CD-trees [48].

The DBSCAN, ROCK, and SNN algorithms are based on the study of the neighborhood graph. The DBSCAN algorithm requires setting the neighborhood distance and is based on the study of the direct reachability of the points in the resulting neighborhood graph. In this sense, the DBSCAN is closely related to the previous group of methods. The algorithm is capable of identifying the core of the cluster (points that have at least n neighbors), points that are related to the cluster but are not the core, and the outliers.

The ROCK and Shared Nearest Neighbors (SNN) algorithms use the concept of a link, which is the number of potential common neighbors between two clusters, or, equivalently, the number of possible links of length two between clusters in the neighborhood graph. The ROCK algorithm is designed for cases where it is necessary to cluster vectors from categorical variables and can use a nonmetric similarity function that is based on a similarity table. In this method, for clustering, the authors propose to consider each point as a separate cluster and use a recurrent mechanism to combine them. Thus, unlike the k -means, the ROCK does not require specifying the number of clusters. Using the number of common neighbors to compute the similarity of two vectors alleviates the problems associated with large data dimension and decreasing density.

The methods of k -means, k -medoids, and self-organizing maps are widely known. They use the concept of a centroid and group elements based on the distance from the centroid, after which they change the coordinates of the centroid to the center of mass of the elements belonging to it. When using these algorithms, data that lies far from the detected centroids are considered anomalous.

Algorithms in the third subclass perform clustering with one of the existing clustering algorithms, after which some of the clusters are declared anomalous based on some criteria such as the density of the cluster and the number of points in it, or according to some existing markup. The FindCBLOF algorithm [47] is a modified clustering algorithm based on the local density of points.

The advantages of clustering-based methods are that

1. They admit unsupervised learning.
2. They can operate various data types.
3. They work well with small spherical datasets.

The shortcomings of clustering-based methods are that

1. The anomalies are binary; i.e., the algorithms do not assign the values of abnormality.

2. Many methods are sensitive to the initial choice of the number of clusters; however, when analyzing the time series in real time, the number of clusters cannot be known in advance.
3. Many methods rely on distance, and therefore, their efficiency decreases in a high-dimensional space.
4. Methods based on centroids are sensitive to their initial locations, suffer from the curse of dimensionality on very large datasets, and are sensitive to outliers.

Application of clustering-based methods

Iverson [49] presented an algorithm that he called an inductive monitoring system (IMS). The author took the k -means algorithm as a basis and proposed a mechanism for increasing the number of allocated clusters in the process of data analysis. This mechanism is based on thresholding the distance between a newly observed point to already existing clusters. The algorithm is trained on the data of nominal operating modes, and it is assumed that the points lying far from the selected centers of the clusters are anomalous. IMS-based systems have been used to monitor the health of gyroscopes and the temperature control system at the International Space Station [50]. Singh [51] offers similar principles for constructing a system for monitoring the state of a nanosatellite.

Chen et al. [52] used laboratory test data of satellite flywheel bearings to distinguish normal bearings from bearings with problems in the outer ring or cage. In this study, the authors used the correlation dimension of accelerometer readings as features. To improve the results, they applied the k -medoid clustering successively two times: on the data of the correlation dimension and on the cluster centers obtained as a result of the first clustering. The results were validated on experimental data and showed a good discriminating ability of the algorithm.

Suo et al. [53] proposed a neighborhood-based grid clustering method and verified the method using numerical simulation data for the electrical power system (EPS). The algorithm uses a pre-determined grid spacing to determine neighboring points and employs the resulting grid density to optimize the complexity of classical distance-based methods, quadratic in the number of points.

4.3. Statistical Methods

Statistical (or probabilistic) methods are used under the assumption that normal and abnormal data have statistically different probability densities, based on which they can be distinguished. The main advantage of these methods is that, in addition to the probability of attributing a situation to one class or another, they also allow calculating the confidence intervals of the resulting estimates. However, in the case of the anomaly detection problem, there is a problem of determining the threshold probability starting from which the state should be considered anomalous.

There are the following subclasses of statistical methods:

- Parametric methods:
 1. Gaussian mixture model (GMM).
 2. Kalman filters.
 3. Hidden Markov models.
 4. Autoregressive integral moving average (ARIMA).
- Nonparametric methods:
 1. Kernel density estimates.
 2. Histogram methods.

Parametric methods make some assumptions about the distribution of observables. Kalman filters assume the existence of a fixed set of parameters that form the estimated state vector of the system and evolve according to a known law (up to a normally distributed random variable that is

additively included in this law). The observation model is also assumed to be given. In accordance with this model, the measurement data are associated with the state vector by a known function (again, up to a normally distributed random variable). Various modifications of the Kalman filter are also used to estimate the covariance matrix of the error in predicted parameters. If the actual measurement took place in an area with a low probability density, the point is marked as anomalous.

Hidden Markov models work in a similar way in a discrete variable space and allow one to define functions in a tabular form.

The ARIMA model approximates experimental data by the dependence

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) (1 - L)^d X_t = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t,$$

which is concisely denoted by $ARIMA(p, d, q)$, where X_t is a time series; L^i is the evolution operator; ε_t are independent random variables taken from $N(0, \sigma)$; p , d , and q are the parameters of autoregression (AR), integral (I), and moving average (MA) models, respectively; and ϕ_i and θ_i are model parameters. Special cases of this model are

- $ARIMA(0, 0, n)$ is the moving average in a window of width n .
- $ARIMA(0, n, 0)$ means that the n th derivative of the time series is constant.
- $ARIMA(n, 0, 0)$ is the autoregression over n previous states.

Histogram methods are a way of estimating the probability function for categorical or discrete variables that work by simple counting and normalization. Kernel probability density estimate is a natural generalization of this method for estimating continuous distribution functions.

The advantages of the statistical methods are

- Good mathematical soundness.
- The model does not require much memory for storage.

The disadvantages of the statistical methods include

- The emergence of problems when the experimental data do not correspond to the assumptions made about the distribution function.
- Low efficiency in the case of small amount of data or large dimension.
- Nonparametric methods cannot efficiently handle relationships between variables.

Applications of statistical methods

Azevedo et al. [54] apply k -means clustering and expectation maximization algorithms, assuming that the normal data are a mixture of normal distributions. They show that both algorithms are sensitive enough to detect an overcharge anomaly six hours before the threshold is exceeded.

Rahimi et al. [55] presented an approach based on unscented Kalman filter for detecting anomalies in flywheel control motors (FCMs). The abnormal system operation is detected when the values of the monitored telemetry parameters go beyond three standard deviations calculated in a window of some predetermined width. In the case of anomalous operation, the authors of [55] propose to adapt the filtering algorithm by changing the diagonal elements of the error covariance matrix corresponding to the anomalous values of the controlled parameters. The operation of the algorithm on the FCM model is demonstrated in the case of various scenarios of abnormal operation (with model anomalies in the FCM input voltage). It is shown that the developed algorithm is capable of detecting the indicated model anomalies, while the assessment of the controlled parameters of the FCM remains satisfactory and makes it possible to maintain the system operation.

Yairi et al. [56] proposed an algorithm that combines probabilistic PCA for continuous variables as well as a probabilistic model for discrete state variables, thus extending the distribution modeling

approach proposed by Tipping and Bishop in [57]. As a result of the analysis, a time series is obtained with an anomaly score of the initial data. The paper [56] uses data from the Japanese satellite JAXA SDS-4, including its attitude determination and control system (ADCS), electrical power system, and thermal control system (TCS). As a result of the analysis, it was possible to distinguish different modes of operation of the spacecraft from each other as well as to identify an anomaly in the ADCS operation. The same paper [56] discusses various options for raw data preprocessing.

Adnane et al. [58] demonstrate a modification of the generalized Kalman filter to improve the state estimation in the ADCS loop under various scenarios of errors in magnetometer measurements. It is shown that the proposed algorithm is able to determine and evaluate the magnetometer errors online. The algorithm was tested on high-precision microsatellite simulation data.

Ahmed et al. [59] proposed to apply the Kaplan–Meier method [60] after a preliminary classification based on logical analysis of data (LAD). The authors of [59] used the NASA Ames Prognostics dataset to classify batteries according to their condition based on capacity and showed that this method improves the estimate of the probability of failure compared with the Kaplan–Meier method alone.

4.4. Classification Methods

Classification methods are based on the general assumption that a classifier can be created that can reliably separate normal states from abnormal ones in a given space of features. These methods are mainly used to identify known anomalies. The following types of algorithms belong in classification methods:

1. Methods based on the support vector machine, including those with the use of kernel functions.
2. Methods based on neural networks.
3. Rule-based methods.
4. Methods based on Bayesian networks.

All of these algorithms work in supervised learning mode and therefore require labeled data for training.

The support vector machine (SVM) looks for a hyperplane in the feature space that can divide the training sample into two classes. This method can be adapted to detect a nonlinear boundary using a kernel function mechanism. Classification by the method of k nearest neighbors (kNN) is one of the most popular in the class.

Various types of neural networks can also be used for classification, such as feedforward networks (also known as the multilayer perceptron (MLP)), recurrent networks, including those with gated recurrent unit (GRU) [61] or long short-term memory (LSTM) [62] neurons, convolutional neural nets (CNN), and other types used alone or sequentially with other algorithms.

Rule-based methods include association rule mining [63], logical analysis of data (LAD), and methods based on them. These are sometimes referred to as decision-tree-based methods. In this case, a certain interpretable rule is constructed on the training sample. Data that do not obey it are considered anomalous. All Bayesian-network-based methods estimate the posterior probability of seeing a label of a particular class. The simple Bayesian classifier is the simplest and most popular method in this group. The basic technique assumes that variables are independent of each other. However, it can be generalized using more complex Bayesian networks.

Single-class classification methods are often ascribed to classification methods. They are trained only on the samples of one class, and in the case of anomaly detection, these are representatives of the normal operation of the system. The algorithm builds the boundary of the area in which all these representatives fall, and objects outside this area are considered anomalous. The most commonly used method in this category is one-class support vector machine (OC-SVM) [64].

Das et al. [65] developed a multiple kernel anomaly detection (MKAD) method, which combines in a single kernel the longest common subsequence (LCS) metric [66] for discrete variables and a kernel based on symbolic aggregate approximation (SAX) [67] for representation of continuous variables for the OC-SVM single-class classification. The paper [65] demonstrates the advantage of the proposed scheme over the ORCA and SequenceMiner algorithms when searching for significant deviations in flight data.

The advantages of classification methods are as follows:

- Ready-made classification algorithms can be used.
- The created models are fairly fast.

The drawbacks of classification methods are as follows:

- Classification methods rely on a reference markup in the training set, which is far from being always available.
- They only assign a discrete label, making it difficult to rank the results.

Applications of classification methods

Ke Li et al. investigated various methods for classifying time series. In [68], they used the fuzzy c-means algorithm and feature extraction using principal component analysis (PCA), which can be used to facilitate data markup by the operator, after which they applied a modification of the weighted proximal support vector machine (WPSVM) to classify the signals of the power supply system of the spacecraft into normal and abnormal. The authors studied the performance of the algorithm for different approaches to multiclass classification by the support vector machine—differentiation of classes in pairs and one of the classes with the rest—and also compared the WPSVM with the operation of the naive Bayesian classifier and the k nearest neighbors (kNN) method. The authors have shown that the performance of the WPSVM allows it to be used in online setting. In [69], the authors investigated a feature extraction method based on a deep belief network with subsequent classification using a random forest. The data was prefiltered using a wavelet transform. For comparison, the authors cited the results of classification by other algorithms: kNN, SVM, naive Bayesian classifier using several feature extraction methods (PCA, stacked autoencoder), and also on the initial data.

Vorontsov [70] suggested using a feedforward network to classify the modes of operation of temperature control and power supply systems. On a sample of 50 tests, the operation of the algorithm was demonstrated and the frequencies of 20% and 10% were reached for false positive and false negative responses, respectively.

Nassar et al. [71, 72] compared two supervised algorithms: partial least squares discriminant analysis (PLS-DA), kernel support vector machine (KSVM), and the results of calculations in a commercial package. The authors state that the PLS-DA based approach separates anomalous data better than the PCA, which only looks for the direction of maximum variability in the data. The system was tested on 16 telemetry parameters of the spacecraft attitude determination and control system.

Fuertes et al. [73] presented the limitations of classical approaches to anomaly detection by threshold values and described the created NOSTRADAMUS software product, which was used to monitor the telemetry of the CNES spacecraft. This product downsizes data and produces one-class support vector machine classification. The authors described examples of using this system based on TCS errors and oscillations of the star tracker readings. In addition, they compared the OC-SVM, kNN, and LOF methods and briefly described the advantages and disadvantages of these methods. In conclusion, the authors described the concept of operations, functionality, and architecture of the selected solution for implementation and greater automation of the operators' work.

Galal et al. [74] applied the principal component method and Bayesian classifier to the telemetry of the spacecraft battery. To mark the abnormal data, they used a routine check of the power supply operability, since in this mode the battery experiences loads that are uncharacteristically large for normal operation. For validation, the authors used data from a degraded battery after an accident. In [74], the authors managed to reduce the dimension of the original data from 30 to 5 without significant loss in the quality of classification and achieve an accuracy of 0.95 according to the AUC ROC metric.

Using the telemetry information data of the electrical power system of the EgyptSat-1 spacecraft, Ibrahim et al. [75] applied the SVM regression [76] to predict the behavior of the time series and continued to work on the interpretability of the result using the method of logical analysis of data (LAD) [77], which was trained on the results of k -means clustering. Combining this approach with the analysis of the fault tree, the authors estimated the most probable cause of the failure to investigate the possible causes of the loss of the EgyptSat spacecraft.

Unfortunately, in many publications, the authors provide only the classification accuracy metric and do not indicate the accuracy and completeness of recognition for each class or matrix of errors. It is these indicators that are critical for practical applications and the analysis of results with unbalanced samples that occur when working with failures and anomalies. In addition, these metrics are usually applied to unstructured data, and their application to time series has some peculiarities.

4.5. *Reconstruction-Based Methods*

Reconstruction-based techniques assume that it is possible to build some kind of a data model. By comparing the current readings of the sensors with the data calculated using this model, an anomaly can be detected. In other words, reconstruction-based methods use reconstruction error to separate the data into normal and abnormal. An algorithm trained on normal data will yield significant reconstruction errors on abnormal data.

These methods can be roughly divided into methods based on signal reconstruction and methods based on prediction.

The methods of the first class, in fact, assume that in each state, the readings of the sensors are related by some dependence, and therefore, they can be fully described by fewer variables. One of the most popular methods is the principal component analysis (PCA), including that with the use of kernel functions. This class also includes methods based on autoencoders.

Prediction-based methods assume that based on the data, one can construct a model of their evolution under nominal operating conditions. These methods include various types of neural networks, including recurrent, feedforward, and generative ones.

The advantages of reconstruction-based methods are as follows:

- They make problems with multidimensional data easier by automatically downsizing them.
- They can be used in the unsupervised learning mode.

The drawbacks of reconstruction-based methods are as follows:

- They are useful only if anomalies can be separated in some feature space.
- They usually have high computational complexity.

Applications of reconstruction-based methods

Abramov et al. [78] suggested employing a feedforward neural network for recurrent time series prediction but used threshold methods to detect anomalies in the predicted time series. Using this method, the authors predicted the system operation for four minutes and achieved an emergency prediction accuracy of 0.887 with a completeness of 0.938.

O'Meara et al. [28] investigated the capabilities of feedforward neural networks to predict the time series from orbit-aggregated telemetry parameter statistics and detected an anomaly based on the mismatch between the measured and predicted parameters. In addition, in [28], the authors investigated the anomaly detection capabilities of the ATHMoS system using features extracted with an autoencoder as well as those in combination with manually specified features. In the first part of [28], the authors managed to correctly predict the behavior of aggregated parameters three orbit cycles ahead. In the second part, the authors showed that manually selected features play a large role in the reliable detection of anomalies.

Martinez-Heras and Donati [79] used a deep autoencoder for anomaly detection in Mars Express telemetry and suggested that recurrent neural networks could be used to determine threshold values.

Petković et al. [80] propose a batch processing method for predicting the power consumption of the temperature control system on the Mars Express spacecraft. They used manual feature extraction and ensemble techniques such as random forest and gradient boosting on decision trees to create a predictive model. Their model was used to distribute power to scientific experiments. The authors note that the model can be used to find anomalies in the mismatch between predicted and measured power consumption.

Ibrahim et al. [81] compared various deep learning methods on telemetry parameters obtained from one of the components of the EgyptSat spacecraft. They used various architectures such as ARIMA, MLP (multilayer perceptron), RNN, LSTM, GRU, DLSTM, and DGRU to predict time series behavior. The comparison was made according to the prediction accuracy performance metrics: mean square error, mean absolute error, correlation coefficient, r^2 metric, and training time. The authors concluded that simpler methods such as ARIMA are better suited for predicting the time series during regular operation of spacecraft with short active lifetime of approximately 3–5 years.

Omran and Murtada investigated the feasibility of feature extraction for classifying failures using Prony's method. In [82], the authors used a recurrent neural network trained on features extracted using Prony's method for anomaly detection in the operation of flywheels. The authors used the data of high-precision numerical modeling. After that, the authors used a simple feedforward network to classify the found anomalies into known classes of abnormal operation of the subsystem under consideration. As a result, the algorithm was able to correctly identify all types of simulated faults and correctly distinguish previously unknown operating modes with an accuracy close to 100% and exceeding the accuracy of algorithms such as ARMA, PCA, and SVM. In [83], the authors investigated the capabilities of Prony's method for feature extraction and subsequent identification of solar battery failures and the noise immunity of such an algorithm. Numerical modeling data were used for analysis. Using the feedforward network, the authors identified ten different failure types in noisy data.

Shin and Tariq [84] described the system implemented for monitoring the state of the KOMPSAT-2 spacecraft at the Korea Aerospace Research Institute (KARI). As in [28], the authors used the statistically aggregated features for each telemetry parameter for every 10 minutes—mean, maximum and minimum, median value, mean signal energy, etc. Unlike [28], the authors treated this data as a third-order tensor using CANDECOMP/PARAFAC decomposition in [85]. After that, the authors applied clustering on the rows of one of the matrices in the expansion to register the moment when the anomaly occurred. In the final part of the analysis, the authors applied the dynamic threshold method proposed by Hundman [86].

Pilastre [87] proposed an anomaly detection method based on a sparse linear representation of signal segments. The method operates with time series with discrete and continuous telemetry parameters and works in two stages. At the first stage, one finds a representation of the projection of only discrete variables and a dictionary of continuous variables corresponding to these discrete variables. In case such a representation can be found, the second stage constructs the continuous signal

from a linear combination of the components of the dictionary of continuous variables. Anomaly detection is possible at the first stage if the representation of discrete variables from the dictionary is impossible and at the second stage if the signal reconstruction error exceeds a predetermined threshold.

4.6. Ensemble Methods

Ensemble methods have been discussed in detail by Aggarwal [88]. These primarily include various meta-algorithms, i.e., algorithms that use and combine other algorithms. Two categorizations of ensemble methods can be distinguished—by the independence of the components and by their type.

In ensembles with independent components, different algorithms are applied on different parts of the original data or their different components. After this, they are combined to obtain a more stable result. The random forest is a typical representative of this class. In contrast, when constructing an ensemble sequentially, each subsequent algorithm depends on the result of the previous one. This happens, for example, when using gradient boosting.

In categorization by component type, Aggarwal highlights model- and data-oriented ensembles. The former aggregate the results of basic algorithms that are different in nature. This raises problems associated with different types of results for each of the algorithms. For example, the result of one algorithm can be a value from -1 to 1 , with the result of another being the conditional probability of an outlier.

Data-driven ensembles can use a subset of the data to build each predictive model and then aggregate their results for a more robust result. The random forest is a representative of this class of algorithms.

The advantages of ensemble methods are as follows:

- In some cases, it is possible to combine the advantages of different methods.
- In the case of independent components of the ensemble, the algorithms are easily scalable for multiprocessor computations.

The drawbacks of ensemble methods are as follows:

- Usually, the high computational complexity due to the need to build several models.

Applications of ensemble methods

Carlton et al. [89] used statistical approaches to search for points of interest in the time series, after which they grouped the results by the same components in different spacecraft or by different components of the same spacecraft based on the proximity of the found points in time. If there were a lot of points in one or several groups at some point in time, then this point in time received a large indicator of anomaly and the corresponding data was displayed on the operator's screen. The authors validated this approach using data from Intelsat and Inmarsat telecommunication satellites and the SpaceTrak database.¹

Nozari et al. [90] proposed to analyze the results of various classifiers using a feedforward neural network and demonstrated the network operation on the support vector machine, the PLS (partial least squares), and a random forest using data on the operation of an FCM.

5. DETERMINING THRESHOLD VALUES

As a result of applying many of the methods described, in particular, statistical or reconstruction-based methods, a time series is obtained with an estimate of the anomalousness of the corresponding

¹ <https://www.seradata.com/products/spacetrak/>.

state. However, this raises the aforementioned problem of determining the threshold value of anomalies starting from which the operator must be warned about the possible approach of an emergency situation. This often happens by empirically determining a threshold value based on information about already known failures and choosing a ratio between the number of false positives and false negatives.

As a rule, long-term operation in modes close to the threshold is a sign of abnormal processes. However, anomaly detection using thresholds is not capable of detecting such modes of operation. Hundman et al. [86] proposed a method for dynamically determining the threshold in time series analysis that can deal with this problem. The authors of [86] proposed to determine the threshold value depending on the duration and magnitude of the signal prediction error and suggested a mechanism for filtering false positives. The operation of the algorithm was demonstrated using the Numenta Anomaly Benchmark as well as data from the Soil Moisture Active Passive spacecraft (25 telemetry parameters) and the Mars Science Laboratory (55 telemetry parameters).

Pang et al. [91] proposed a coverage width criterion for calculating the anomaly threshold as a model hyperparameter using probabilistic models. The authors of this publication have tested the efficiency of this method on the telemetry parameters of the Fengyun spacecraft power supply system.

6. METHOD QUALITY METRICS

In classification and clustering problems, typical algorithm performance measures are precision and recall,

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}),$$

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}),$$

where TP, FN, and FP are true positive, false negative, and false positive classification results, respectively. Thus, precision reflects the proportion of correct answers from all points that the system attributed to a given class, and sensitivity, which is also called recall, reflects the proportion of correctly classified points from all points of a given class. As a rule, these indicators cannot be improved simultaneously, and a balance between them must be chosen for each task. For example, precision can be sacrificed in very large systems with great stability, whereas if the cost of an error is very high, then it is necessary to optimize the sensitivity at the expense of more false positives. Precision and recall are often combined into an F-measure, which is defined as their geometric mean.

These metrics are intended for analyzing the classification of individual objects, but they are not fully applicable to time series, since they ignore their structure. Lavin et al. [92] proposed a Numenta Anomaly Benchmark that, to an extent, allows for the sequential structure of the time series and rewards early anomaly detection. From the point of view of the authors of [92], an ideal detector has the following properties:

1. It detects all anomalies in the data.
2. It detects anomalies as early as possible, ideally, before they are noticed by a human.
3. It does not produce false positives.
4. It works with data online.
5. It is fully automated to work with various datasets.

Based on these requirements, one defines a window of given width around a point anomaly. If an algorithm finds any anomaly inside the window, the result is considered as true positive and rewarded, while the triggering outside the window is treated as false and is penalized. Thus, this metric treats anomalies as point anomalies but takes into account the temporal structure of the data. In addition to these methods for calculating metrics, the authors of [92] provide the

Numenta NAB Data Corpus,² which contains one-dimensional time series from various fields, such as information and communication technologies, air conditioning systems in buildings, etc. This data has a reference markup and algorithms to assess the anomaly detection efficiency.

Lee and Tatbul [93] proposed customizable time series anomaly detection performance metrics that generalize the concepts of precision and recall for sets. These metrics take into account the partial intersection of identified and actual anomalies by introducing functions that penalize different types of inaccuracies in determining anomalous intervals. This metric can take into account the specifics of the problem being solved by introducing other functions for the penalty for errors and rewards for the found anomalies. At the same time, by choosing certain functions of rewards and penalties, this metric can be reduced to the metric proposed in [92] and hence is more general.

Note that both metrics require a reference markup of anomalies for correct performance assessment and quantitative comparison of the operation of various algorithms.

7. APPLICATIONS AND PROSPECTS OF DEVELOPMENT IN SPACE TECHNOLOGY

In this section we describe the existing industrial monitoring systems based on data analysis that have been used or are being used by various space agencies to facilitate the operation of spacecraft, possible options for the development of the methods used, as well as open data sources that can be used to test algorithms developed by the open community.

One of the first systems based on data analysis was the Inductive Monitoring System developed by Iverson and colleagues at NASA [49]. Its principle of operation is based on the k -means algorithm; however, a significant difference is a refinement that allows one to increase the number of clusters dynamically. This method requires a database of nominal states and provides a selection of states that lie far enough from the known nominal states.

O'Meara [41] presented the system used at the German Space Operations Center (GSOC) DLR. The core of the system uses IDOS algorithm [33], which analyzes point density calculated as an estimate of local intrinsic dimensionality and supplements it with an algorithm for choosing a threshold value for each telemetry parameter. In addition, [28] briefly describes the basic infrastructure of the monitoring system.

Columbus Control Centre DLR implemented Project Sibyl [94] as a system for detecting anomalies when monitoring manned space flights. The system realizes the LoOP algorithm based on the ELKI framework [95]. Further plans are to use Sibyl's developments in the ATHMoS project to create a unified telemetry information monitoring system for all GSOC missions.

CNES jointly with TESA developed and implemented the NOSTRADAMUS system [73] to control the flight of its spacecraft. The authors described the main functionality of the system and the concept of its operation as well as the main intermediate steps of data analysis. The conclusions of [73] provide a description of the main drawbacks in the system revealed during operation, such as an increased sensitivity to input data (owing to peculiarities in the operation of the OC-SVM algorithm chosen as the basis of the system) and the difficulty of interpreting the result.

Data-driven monitoring systems also work at JAXA [56], NASA JPL [86], and Korea Aerospace Research Institute (KARI) [84]. However, the papers [56, 86, 84] describe only the mathematical method of information analysis and do not provide the solution infrastructure.

It can be seen that many space agencies are actively exploiting and implementing various mathematical methods to analyze telemetry information by organizing necessary computing infrastructure and access to the multiannual telemetry information for the necessary validation of the created algorithms.

² <https://github.com/numenta/NAB/tree/master/data/>.

There are also several open information sources:

- Nasa Prognostics Center of Excellence³ provides access to data from various organizations and universities that can be used to study and debug algorithms.⁴
- Access to the telemetry information of the GOCE spacecraft [96] was opened in 2019. It contains approximately 13000 telemetry parameters. However, it does not contain data markup indicating breakdowns and anomalies.
- In 2016, an open competition was held to create an effective predictive model of the thermoregulation system for the Mars Express Orbiter [30]. The data was an archive for three Martian years, including context (for example, angles to the Sun, logs of commands issued, etc.) and observational data (consumption of 33 TCS telemetry parameters). The task was to predict TCS consumption for the fourth year.
- SpaceTrak⁵ is a commercially maintained database of space accidents. This database does not contain telemetry information but can serve as a source of data for validation or other reliability studies.

Unfortunately, the authors of this survey were unable to find a dataset with satellite telemetry data with labeled anomalies. The limited availability of data significantly complicates the development and validation of both specialized and general-purpose spacecraft monitoring methods.

Despite considerable advances and widespread industrial applications, there is a significant demand for new developments. These include active learning technologies for anomaly detection. This request is due to the large amount of incoming information and the fundamental impossibility of complete manual labeling.

In addition, most machine learning algorithms assume that training and test data are from the same distribution. However, this condition is not met in practice owing to the instability of production and various conditions in which the spacecraft are located. Therefore, applying a model trained on one of the spacecraft in a constellation to another spacecraft will lead to inaccuracies. Recently, domain-adversarial training techniques [97] have been used to overcome these difficulties. This topic requires additional research for technical diagnostics problems.

The question remains about the use of expert knowledge and physical models and their integration into the process of automated technical diagnostics. Several studies [98] have shown that this can improve the accuracy and interpretability of the result. Tipaldi [3] suggested a possible architecture for such a solution. A detailed survey of deep learning perspectives for health management problems was given by Fink et al. [99].

8. CONCLUSIONS

The article discusses the main types of data analysis methods used to analyze telemetry information and detect anomalies in it and surveys the main publications on the application of these methods. The features of the problem statement and validation of the spacecraft state monitoring systems operating on this principle are described. Possible metrics of the algorithm performance and the issues of choosing the threshold values of anomaly score for the necessary balance between these metrics are discussed. Possible development prospects in this area are briefly outlined.

The authors of the present survey noticed that most algorithms quickly lose their efficiency with an increase in the number of telemetry parameters to be analyzed; for this reason, many authors in one way or another tried to identify features that have lower dimension than the original data. This problem is commonly known as the curse of dimensionality. To combat it, a reasonable segmentation

³<https://ti.arc.nasa.gov/tech/dash/groups/pcoe/>.

⁴<https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/>.

⁵<https://www.seradata.com/products/spacetrak/>.

of the monitoring task according to the parameters of individual components of the spacecraft with the subsequent aggregation of the results can be useful.

The most popular methods for detecting anomalies are methods based on density, which is calculated in one way or another, as well as methods based on reconstruction, and one of the most efficient ways to improve the performance of models is to study the features of specific components of the spacecraft and carefully prepare features. Often it is this factor that is decisive in the performance of the model rather than the algorithm used for data analysis. Despite this, it can be seen that most publications take into account only the telemetry received from a specific spacecraft, without taking into account the factors of space weather and its forecast, which is extremely important in the spacecraft operation.

Many authors use artificially generated data to validate their algorithms. At the same time, the question of the possibility of extending such results to real flight data remains open. Many space agencies that have access to data to create and validate algorithms are developing and implementing their own systems for health monitoring. These systems show efficient operation in industrial use and reduce the burden on operators.

Unfortunately, at the moment there is not a sufficient amount of various open data with spacecraft telemetry information and reference ground truth labels. This makes it impossible to reproduce and confirm published works, take into account various features of the spacecraft components, or compare the efficiency of experts and algorithms of various types. These factors significantly slow down the development and implementation of automated systems for spacecraft health monitoring. To adapt general-purpose systems to the specific features of the spacecraft operation, it is necessary to create an appropriate labeled dataset as well as objective metrics for assessing and comparing the performance of various algorithms and operators.

In addition, to improve the performance of forecasting the spacecraft health, it is necessary to take into account space weather conditions and integrate this information into decision support systems. On the other hand, we need methods for dividing the problem of monitoring the state of the entire spacecraft into problems of lower dimension, for example, by its component parts, and methods for reducing the dimension by more efficient presentation of information to improve the performance of algorithms, since their efficiency is directly related to the density of the initial data.

Despite the listed difficulties and unsolved problems, data-driven systems for monitoring the spacecraft health are already used in many space agencies around the world and have established themselves as an important technical diagnostics tool.

APPENDIX

Table 1. List of publications

Publication	Data	Feature extraction	Method
Iverson, 2004 [49]	Columbia Space Shuttle	N/A	Clustering
Azevedo, 2012 [54]	Brazilian spacecraft PSS	N/A	k-means, statistical (EM)
Li, 2015 [68]	PSS modeling data	PCA	Fuzzy C-means, WPSVM
Fuertes, 2016 [73]	CNES spacecraft	Statistical parameters of signal in a window	PCA + OC-SVM
O'Meara, 2016 [41]	TET-1 spacecraft flywheels	Vector of consolidated information for each telemetry parameter per turn	OPVID

Table 1. (Contd.)

Publication	Data	Feature extraction	Method
Verzola, 2016 [94]	Mass Memory Unit, Columbus Orbital Laboratory	Time of day, coordinates, sun position, space weather data	Probabilistic parametric
Nassar, 2017 [72]	16 ASS telemetry parameters	N/A	Classification PLS-DA, KSVM
Rahimi, 2017 [55]	Numerical modeling		P-adaptive Kalman filter
Yairi, 2017 [56]	JAXA SDS-4 spacecraft PSS		Bayesian classification + stochastic PCA
Chen, 2018 [52]	Laboratory experiment with ASS FCM	Correlation dimension	K-medoids (2 times)
Pang, 2018 [91]	PSS FENGYUN spacecraft	N/A	LS-SVM, Gaussian Process Regression, Restricted Vector Machine
Adnane, 2018 [58]	Numerical modeling of ASS		Generalized Kalman filter
Carlton, 2018 [89]	Intelsat, Inmarsat		Search for temporary phenomena and transition points by statistical methods, grouping of events
O'Meara, 2018 [28]	Several years of historical information on spacecraft operation	Extracted both manually and by autoencoder	Feedforward network
Martinez, 2018 [79]	991 telemetry parameters of Mars Express spacecraft	N/A	Determining threshold values using feedforward neural network
Gowda, 2018 [29]	Mars Express Challenge telemetry parameters	Manual feature extraction	XGBoost
Suo, 2019 [22]	PSS of geosynchronous spacecraft (artificial)	Selected by the proposed technique	SVM
Galal, 2019 [74]	30 PSS variables	SAR (7 components)	Bayesian classifier
Ibrahim, 2019 [81]	EgyptSat-1		ARIMA, MLP, RNN, LSTM, GRU, DLSTM, DGRU
Ahmed, 2019 [59]	NASA Ames Prognostics		Classification + Kaplan–Meier method
Petković, 2019 [80]	Mars Express PSS	Manual	Random forest, gradient boosting
Suo, 2019 [53]	PSS of geosynchronous spacecraft (artificial)	N/A	NGC network neighborhood-based clustering
Omran, 2019 [82]	High-precision numerical modeling of ASS flywheel	Prony's method	Elman RNN for discord detection and feedforward network for classifying anomalies
Murtada, 2019 [83]	Numerical simulation of solar panels with noise	Prony's method	Feedforward network

Table 1. (Contd.)

Publication	Data	Feature extraction	Method
Nozari, 2019 [90]	Numerical modeling	Results of other classifiers	Meta-classifier based on SVM, NB, PLS, RF
Ibrahim, 2020 [75]	EgyptSat-1 TCS		OC-SVM, K-means, LAD
Pilastre, 2020 [87]	18 days telemetry information: 1000 telemetry parameters (90 affected by anomalies)	Discrete and continuous parameters from a fixed-width window	Anomaly detection using dictionary (ADDICT)

REFERENCES

1. Zolghadri A., Advanced model-based FDIR techniques for aerospace systems: today challenges and opportunities, *Progr. Aerosp. Sci.*, 2012, vol. 53, pp. 18–29.
<https://doi.org/10.1016/j.paerosci.2012.02.004>
2. Abanin, O.I. and Solov'ev, S.V., Content and structure of anomaly detection problems in spacecraft onboard system operation, *Inzh. Zh.: Nauka Innovatsii*, 2019, no. 6(90), p. 7.
<https://doi.org/10.18698/2308-6033-2019-6-1890>
3. Tipaldi, M. et al., On applying AI-driven flight data analysis for operational spacecraft model-based diagnostics, *Annu. Rev. Control*, 2020, vol. 49, pp. 197–211.
<https://doi.org/10.1016/j.arcontrol.2020.04.012>
4. Solov'ev, S.V. and Mishurova, N.V., Analysis of the current control process state in spacecraft flight control, *Inzh. Zh.: Nauka Innovatsii*, 2016, no. 3(51), p. 3.
<https://doi.org/10.18698/2308-6033-2016-3-1474>
5. Solov'ev, V.A., Lyubinskii, V.E., and Zhuk, E.I., State of the art and development prospects of the spacecraft flight control system, *Pilotiruemye Polety Kosmos*, 2011, no. 1(1), pp. 27–37.
6. Balukhto, A.N. and Romanov, A.A., Artificial intelligence in space technology: state of the art and development prospects, *Raketo-Kosm. Priborostr. Inf. Sist.*, 2019, vol. 6, no. 1, pp. 65–75.
<https://doi.org/10.30894/issn2409-0239.2019.6.1.65.75>
7. Chandola, V., Banerjee, A., and Kumar, V., Anomaly detection: a survey, *ACM Comput. Surv.*, 2009, vol. 41, no. 3. <https://doi.org/10.1145/1541880.1541882>
8. Pimentel, M.A.F. et al., A review of novelty detection, *Signal Process.*, 2014, vol. 99, pp. 215–249.
<https://doi.org/10.1016/j.sigpro.2013.12.026>
9. Wang, H., Bah, M.J., and Hammad, M., Progress in outlier detection techniques: a survey, *IEEE Access.*, 2019, vol. 7, pp. 107964–108000. <https://doi.org/10.1109/ACCESS.2019.2932769>
10. Chalapathy, R. and Chawla, S., Deep learning for anomaly detection: a survey, 2019. [arXiv:1901.03407](https://arxiv.org/abs/1901.03407) [cs, stat]. Cited April 17, 2020.
11. Zimek, A., Schubert, E., and Kriegel, H.-P., A survey on unsupervised outlier detection in high-dimensional numerical data, *Stat. Anal. Data Min.: ASA Data Sci. J.*, 2012, vol. 5, no. 5, pp. 363–387.
<https://doi.org/10.1002/sam.11161>
12. Thudumu, S. et al., A comprehensive survey of anomaly detection techniques for high dimensional big data, *J. Big Data*, 2020, vol. 7, no. 1, p. 42. <https://doi.org/10.1186/s40537-020-00320-x>
13. Gavrilovski, A. et al., Challenges and opportunities in flight data mining: a review of the state of the art, in *AIAA Infotech @ Aerospace*, San Diego, California: AIAA, 2016.
<https://doi.org/10.2514/6.2016-0923>
14. Khan, S. and Yairi, T., A review on the application of deep learning in system health management, *Mech. Syst. Signal Process.*, 2018, vol. 107, pp. 241–265. <https://doi.org/10.1016/j.ymssp.2017.11.024>

15. Basora, L., Olive, X., and Dubot, T., Recent advances in anomaly detection methods applied to aviation, *Aerospace*, 2019, vol. 6, no. 11, p. 117. <https://doi.org/10.3390/aerospace6110117>
16. Solov'ev, V.A., Lysenko, L.N., and Lyubinskii, V.E., *Upravlenie kosmicheskimi poletami. Uch. pos. Ch. 1* (Space Flight Control. A Handbook. Part 1), Moscow: Mosk. Gos. Univ. im. Baumana, 2009.
17. Gao, Z., Cecati, C., and Ding, S., A survey of fault diagnosis and fault-tolerant techniques—Part I: fault diagnosis with model-based and signal-based approaches, *IEEE Trans. Ind. Electron.*, 2015, vol. 62, no. 6, pp. 3757–3767. <https://doi.org/10.1109/TIE.2015.2417501>
18. Barber, D., *Bayesian Reasoning and Machine Learning*, New York: Cambridge Univ. Press, 2012.
19. Pimentel, T. et al., Deep active learning for anomaly detection, 2020. [arXiv:1805.09411](https://arxiv.org/abs/1805.09411) [cs, stat]. Cited August 8, 2020.
20. Das, S. et al., Incorporating expert feedback into active anomaly discovery, *2016 IEEE 16th Int. Conf. Data Mining (ICDM)* (2016), pp. 853–858. <https://doi.org/10.1109/ICDM.2016.0102>
21. *IBM Analytics. ASUM Analytics Solutions Unified Method*, 2015.
22. Suo, M. et al., Data-driven fault diagnosis of satellite power system using fuzzy Bayes risk and SVM, *Aerosp. Sci. Technol.*, 2019, vol. 84, pp. 1092–1105. <https://doi.org/10.1016/j.ast.2018.11.049>
23. Kononenko, I., Estimating attributes: analysis and extensions of RELIEF, in *Machine Learning: ECML-94. Lecture Notes in Computer Science*, Bergadano, F. and De Raedt, L., Eds., Berlin–Heidelberg: Springer, 1994, pp. 171–182. https://doi.org/10.1007/3-540-57868-4_57
24. Hanchuan Peng, Fuhui Long, and Ding, C., Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2005, vol. 27, no. 8, pp. 1226–1238. <https://doi.org/10.1109/TPAMI.2005.159>
25. Hu, Q. et al., Neighborhood rough set based heterogeneous feature subset selection, *Inf. Sci.*, 2008, vol. 178, no. 18, pp. 3577–3594. <https://doi.org/10.1016/j.ins.2008.05.024>
26. Mallat, S., *A Wavelet Tour of Signal Processing: The Sparse Way. 3rd Ed.*, New York: Academic Press, 2008. <https://doi.org/10.1016/B978-0-12374370-1.X0001-8>
27. Barreyre, C. et al., Statistical methods for outlier detection in space telemetries, in *Space Operations: Inspiring Humankind's Future*, Pasquier, H. et al., Eds., Cham: Springer Int. Publ., 2019, pp. 513–547. https://doi.org/10.1007/978-3-030-11536-4_20
28. O'Meara, C., Schlag, L., and Wickler, M., Applications of deep learning neural networks to satellite telemetry monitoring, *15th Int. Conf. on Space Operations*, Marseilles, France: AIAA, 2018. <https://doi.org/10.2514/6.2018-2558>
29. Gowda, G.M. et al., The added value of advanced feature engineering and selection for machine learning models in spacecraft behavior prediction, in *2018 SpaceOps Conf.*, Marseilles, France: AIAA, 2018. <https://doi.org/10.2514/6.2018-2561>
30. Lucas, L. and Boumghar, R., Machine learning for spacecraft operations support—the Mars Express power challenge, in *2017 6th Int. Conf. Space Mission Challenges Inf. Technol. (SMC-IT)*, 2017, pp. 82–87. <https://doi.org/10.1109/SMC-IT.2017.21>
31. Bay, S.D. and Schwabacher, M., Mining distance-based outliers in near linear time with randomization and a simple pruning rule, in *Proc. 9th ACM SIGKDD Int. Conf. Knowledge Discovery Data Min. (KDD '03)*, Washington, D.C.: ACM, 2003, pp. 29–38. <https://doi.org/10.1145/956750.956758>
32. Breunig, M.M. et al., LOF: identifying density-based local outliers, in *Proc. 2000 ACM SIGMOD Int. Conf. Manage. Data (SIGMOD '00)*, Dallas, Texas: ACM, 2000, pp. 93–104. <https://doi.org/10.1145/342009.335388>
33. Von Brünken, J., Houle, M.E., and Zimek, A., Intrinsic dimensional outlier detection in high-dimensional data, *NII Tech. Rep.*, 2015, vol. 2015, no. 3, pp. 1–12.

34. Houle, M.E., Kashima, H., and Nett, M., Generalized expansion dimension, in *2012 IEEE 12th Int. Conf. Data Min. Workshops*, 2012, pp. 587–594. <https://doi.org/10.1109/ICDMW.2012.94>
35. Houle, M.E., Dimensionality, discriminability, density and distance distributions, in *2013 IEEE 13th Int. Conf. Data Min. Workshops*, 2013, pp. 468–473. <https://doi.org/10.1109/ICDMW.2013.139>
36. Kriegel, H.-P. et al., Outlier detection in axis-parallel subspaces of high dimensional data, in *Advances in Knowledge Discovery and Data Mining (Lecture Notes in Computer Science)*, Theeramunkong, T. et al., Eds., Berlin–Heidelberg: Springer, 2009, pp. 831–838. https://doi.org/10.1007/978-3-642-01307-2_86
37. Kriegel, H.-P., Schubert, M., and Zimek, A., Angle-based outlier detection in high-dimensional data, in *Proc. 14th ACM SIGKDD Int. Conf. Knowl. Discovery Data Min. (KDD '08)*, New York: ACM, 2008, pp. 444–452. <https://doi.org/10.1145/1401890.1401946>
38. Rosenblatt, M., Remarks on some nonparametric estimates of a density function, *Ann. Math. Stat.*, 1956, vol. 27, no. 3, pp. 832–837. <https://doi.org/10.1214/aoms/1177728190>
39. Tang, B. and He, H., A local density-based approach for outlier detection, *Neurocomputing*, 2017, vol. 241, pp. 171–180. <https://doi.org/10.1016/j.neucom.2017.02.039>
40. Dynamic time warping, in *Inf. Retrieval for Music and Motion*, Muller, M., Ed., Berlin–Heidelberg: Springer, 2007, pp. 69–84. https://doi.org/10.1007/978-3-540-74048-3_4
41. O'Meara, C. et al., ATHMoS: automated telemetry health monitoring system at GSOC using outlier detection and supervised machine learning, in *SpaceOps 2016 Conf.*, Daejeon, Korea: AIAA, 2016. <https://doi.org/10.2514/6.2016-2347>
42. Martinez, J., New telemetry monitoring paradigm with novelty detection, in *SpaceOps 2012 Conf.*, AIAA, 2012. <https://doi.org/10.2514/6.2012-1275123>
43. Schlag, L., O'Meara, C., and Wickler, M., Numerical analysis of automated anomaly detection algorithms for satellite telemetry, in *15th Int. Conf. Space Oper.*, Marseilles, France: AIAA, 2018. <https://doi.org/10.2514/6.2018-2534>
44. Guha, S., Rastogi, R., and Shim, K., Rock: a robust clustering algorithm for categorical attributes, *Inf. Syst.*, 2000, vol. 25, no. 5, pp. 345–366. [https://doi.org/10.1016/S0306-4379\(00\)00022-3](https://doi.org/10.1016/S0306-4379(00)00022-3)
45. Ertöz, L., Steinbach, M., and Kumar, V., Finding clusters of different sizes, shapes, and densities in noisy, high dimensional data, in *Proc. 2003 SIAM Int. Conf. Data Min.*, Soc. Ind. Appl. Math., 2003, pp. 47–58. <https://doi.org/10.1137/1.9781611972733.5>
46. Kohonen, T., Exploration of very large databases by self-organizing maps, *Proc. Int. Conf. Neural Networks (ICNN'97)*, IEEE, 1997, vol. 1, pp. PL1–PL6. <https://doi.org/10.1109/ICNN.1997.611622>
47. He, Z., Xu, X., and Deng, S., Discovering cluster-based local outliers, *Pattern Recognit. Lett.*, 2003, vol. 24, no. 9, pp. 1641–1650. [https://doi.org/10.1016/S0167-8655\(03\)00003-5](https://doi.org/10.1016/S0167-8655(03)00003-5)
48. Sun, H. et al., CD-trees: an efficient index structure for outlier detection, in *Advances in Web-Age Information Management (Lecture Notes in Computer Science)*, Li, Q., Wang, G., and Feng, L., Eds., Berlin–Heidelberg: Springer, 2004, pp. 600–609. https://doi.org/10.1007/978-3-540-27772-9_60
49. Iverson, D.L. and Field, M., *Inductive System Health Monitoring*, 2004.
50. Iverson, D.L. et al., General purpose data-driven monitoring for space operations, *J. Aerosp. Comput. Inf. Commun.*, 2012, vol. 9, no. 2, pp. 26–44. <https://doi.org/10.2514/1.54964>
51. Singh, S., A data-driven approach to Cubesat health monitoring, *Master's Theses and Project Reports*, 2017. <https://doi.org/10.15368/theses.2017.100>
52. Chen, C. et al., A fault diagnosis method for satellite flywheel bearings based on 3D correlation dimension clustering technology, *IEEE Access*, 2018, vol. 6, pp. 78483–78492. <https://doi.org/10.1109/ACCESS.2018.2885046>

53. Suo, M. et al., Neighborhood grid clustering and its application in fault diagnosis of satellite power system, *Proc. Inst. Mech. Eng. Part G: J. Aerosp. Eng.*, 2019, vol. 233, no. 4, pp. 1270–1283. <https://doi.org/10.1177/0954410017751991>
54. Azevedo, D.R., Ambrosio, A.M., and Vieira, M., Applying data mining for detecting anomalies in satellites, in *2012 9th Eur. Dependable Comput. Conf.*, 2012, pp. 212–217. <https://doi.org/10.1109/EDCC.2012.19>
55. Rahimi, A., Kumar, K.D., and Alighanbari, H., Fault Estimation of Satellite Reaction Wheels Using Covariance Based Adaptive Unscented Kalman Filter, *Acta Astronaut.*, 2017, vol. 134, pp. 159–169. <https://doi.org/10.1016/j.actaastro.2017.02.003>
56. Yairi, T. et al., A data-driven health monitoring method for satellite housekeeping data based on probabilistic clustering and dimensionality reduction, *IEEE Trans. Aerosp. Electron. Syst.*, 2017, vol. 53, no. 3, pp. 1384–1401. <https://doi.org/10.1109/TAES.2017.2671247>
57. Tipping, M.E. and Bishop, C.M., Mixtures of probabilistic principal component analyzers, *Neural Comput.*, 1999, vol. 11, no. 2, pp. 443–482. <https://doi.org/10.1162/089976699300016728>
58. Adnane, A. et al., Real-time sensor fault detection and isolation for LEO satellite attitude estimation through magnetometer data, *Adv. Space Res.*, 2018, vol. 61, no. 4, pp. 1143–1157. <https://doi.org/10.1016/j.asr.2017.12.007>
59. Ahmed, A.M. et al., Prediction of battery remaining useful life on board satellites using logical analysis of data, in *2019 IEEE Aerosp. Conf.*, 2019. P. 1–8. <https://doi.org/10.1109/AERO.2019.8741717>
60. Kaplan, E.L. and Meier, P., Nonparametric estimation from incomplete observations, *J. Am. Stat. Assoc.*, 1958, vol. 53, no. 282, pp. 457–481. <https://doi.org/10.2307/2281868>
61. Chung, J. et al., Empirical evaluation of gated recurrent neural networks on sequence modeling, 2014. [arXiv:1412.3555](https://arxiv.org/abs/1412.3555) [cs]. Cited January 9, 2020.
62. Hochreiter, S. and Schmidhuber, J., Long short-term memory, *Neural Comput.*, 1997. <https://doi.org/10.1162/neco.1997.9.8.1735>
63. Agrawal, R., Imielinski, T., and Swami, A., Mining association rules between sets of items in large databases, *ACM SIGMOD Record*, 1993, vol. 22, no. 2, pp. 207–216. <https://doi.org/10.1145/170036.170072>
64. Schölkopf, B. et al., Estimating the support of a high-dimensional distribution, *Neural Comput.*, 2001, vol. 13, no. 7, pp. 1443–1471. <https://doi.org/10.1162/089976601750264965>
65. Das, S. et al., Multiple kernel learning for heterogeneous anomaly detection: algorithm and aviation safety case study, in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Min. (KDD '10)*, New York: ACM, 2010, pp. 47–56. <https://doi.org/10.1145/1835804.1835813>
66. Budalakoti, S., Srivastava, A.N., and Otey, M.E., Anomaly detection and diagnosis algorithms for discrete symbol sequences with applications to airline safety, *IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev.*, 2009, vol. 39, no. 1, pp. 101–113. <https://doi.org/10.1109/TSMCC.2008.2007248>
67. Patel, P., et al., Mining motifs in massive time series databases, *Proc. 2002 IEEE Int. Conf. Data Min.*, 2002, pp. 370–377. <https://doi.org/10.1109/ICDM.2002.1183925>
68. Li, K. et al., A spacecraft electrical characteristics multi-label classification method based on off-line FCM clustering and on-line WPSVM, *PLoS ONE*, 2015, vol. 10, no. 11, p. e0140395. <https://doi.org/10.1371/journal.pone.0140395>
69. Li, K. et al., Multi-label spacecraft electrical signal classification method based on DBN and random forest, *PLOS ONE*, 2017, vol. 12, no. 5, p. e0176614. <https://doi.org/10.1371/journal.pone.0176614>
70. Vorontsov, V.A. and Fedorov, E.A., Development of a prototype of an intelligent system for operational monitoring and technical condition of the main onboard spacecraft systems, *Tr. Mosk. Aviats. Inst.*, 2015, no. 82, p. 35.

71. Nassar, B. and Hussein, W., State-of-health analysis applied to spacecraft telemetry based on a new projection to latent structure discriminant analysis algorithm, in *2015 IEEE Aerosp. Conf.*, 2015, pp. 1–11. <https://doi.org/10.1109/AERO.2015.7118887>
72. Nassar, B., Hussein, W., and Medhat, M., Supervised learning algorithms for spacecraft attitude determination and control system health monitoring, *IEEE Aerosp. Electron. Syst. Mag.*, 2017, vol. 32, no. 4, pp. 26–39. <https://doi.org/10.1109/MAES.2017.150049>
73. Fuertes, S. et al., Improving spacecraft health monitoring with automatic anomaly detection techniques, *SpaceOps 2016 Conf.*, AIAA, 2016. <https://doi.org/10.2514/6.2016-2430>
74. Galal M.A., et al., Satellite battery fault detection using naive Bayesian classifier, *2019 IEEE Aerosp. Conf.*, 2019, pp. 1–11. <https://doi.org/10.1109/AERO.2019.8741963>
75. Ibrahim, S.K. et al., Machine learning techniques for satellite fault diagnosis, *Ain Shams Eng. J.*, 2020, vol. 11, no. 1, pp. 45–56. <https://doi.org/10.1016/j.asej.2019.08.006>
76. Trafalis, T.B. and Ince, H., Support vector machine for regression and applications to financial forecasting, in *Proc. IEEE-INNS-ENNS Int. Joint Conf. Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, IEEE, 2000, vol. 6. pp. 348–353. <https://doi.org/10.1109/IJCNN.2000.859420>
77. Chikalov, I. et al., Logical analysis of data: theory, methodology and applications, in *Three Approaches to Data Analysis: Test Theory, Rough Sets and Logical Analysis of Data (Intelligent Systems Reference Library)*, Chikalov, I. et al., Eds., Berlin–Heidelberg: Springer, 2013, pp. 147–192. https://doi.org/10.1007/978-3-642-28667-4_3
78. Abramov, N.S. et al., High-performance neural network system for monitoring the state and behavior of spacecraft subsystems using telemetry data, *Program. Sist. Teor. Pril. (Rossiya, Ves'kovo)*, 2017, no. 3(30).
79. Martinez, J. and Donati, A., Novelty detection with deep learning, *2018 SpaceOps Conf.*, AIAA, 2018. <https://doi.org/10.2514/6.2018-2560>
80. Petković, M. et al., Machine learning for predicting thermal power consumption of the Mars Express spacecraft, *IEEE Aerosp. Electron. Syst. Mag.*, 2019, vol. 34, no. 7, pp. 46–60. <https://doi.org/10.1109/MAES.2019.2915456>
81. Ibrahim, S.K. et al., Machine learning methods for spacecraft telemetry mining, *IEEE Trans. Aerosp. Electron. Syst.*, 2019, vol. 55, no. 4, pp. 1816–1827. <https://doi.org/10.1109/TAES.2018.2876586>
82. Omran, E.A. and Murtada, W.A., Efficient anomaly classification for spacecraft reaction wheels, *Neural Comput. Appl.*, 2019, vol. 31, no. 7, pp. 2741–2747. <https://doi.org/10.1007/s00521-017-3226-y>
83. Murtada, W.A. and Omran, E.A., Robust anomaly identification algorithm for noisy signals: spacecraft solar panels model, *Neural Comput. Appl.*, 2019, vol. 32, pp. 12281–12294. <https://doi.org/10.1007/s00521-019-04407-2>
84. Shin, Y. et al., ITAD: integrative tensor-based anomaly detection system for reducing false positives of satellite systems, in *Proc. 29th ACM Int. Conf. Inf. & Knowl. Manage. (CIKM '20)*, New York: ACM, 2020, pp. 2733–2740. <https://doi.org/10.1145/3340531.3412716>
85. Kiers, H.A.L., Towards a standardized notation and terminology in multiway analysis, *J. Chemometrics*, 2000, vol. 14, no. 3, pp. 105–122.
86. Hundman, K. et al., Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding, *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Min.*, 2018, pp. 387–395. <https://doi.org/10.1145/3219819.3219845>
87. Pilastre, B. et al., Anomaly detection in mixed telemetry data using a sparse representation and dictionary learning, *Signal Process.*, 2020, vol. 168, p. 107320. <https://doi.org/10.1016/j.sigpro.2019.107320>
88. Aggarwal, C.C., Outlier ensembles: position paper, *ACM SIGKDD Explor. Newsl.*, 2013, vol. 14, no. 2, pp. 49–58. <https://doi.org/10.1145/2481244.2481252>

89. Carlton, A. et al., Telemetry fault-detection algorithms: applications for spacecraft monitoring and space environment sensing, *J. Aerosp. Inf. Syst.*, 2018, vol. 15, no. 5, pp. 239–252.
<https://doi.org/10.2514/1.1010587>
90. Nozari, H.A. et al., Novel non-model-based fault detection and isolation of satellite reaction wheels based on a mixed-learning fusion framework, *IFAC PapersOnLine*, 2019, vol. 52, no. 12 (21st IFAC Symposium on Automatic Control in Aerospace ACA 2019), pp. 194–199.
<https://doi.org/10.1016/j.ifacol.2019.11.222>
91. Pang, J. et al., Anomaly detection for satellite telemetry series with prediction interval optimization, *2018 Int. Conf. Sensing Diagn. Prognostics Control (SDPC)*, 2018, pp. 408–414.
<https://doi.org/10.1109/SDPC.2018.8664879>
92. Lavin, A. and Ahmad, S., Evaluating real-time anomaly detection algorithms—the Numenta Anomaly Benchmark, in *2015 IEEE 14th Int. Conf. Mach. Learn. Appl. (ICMLA)*, 2015, pp. 38–44.
<https://doi.org/10.1109/ICMLA.2015.141>
93. Tatbul, N. et al., Precision and recall for time series, in *Advances in Neural Information Processing Systems 31*, Bengio, S. et al., Eds., Curran Assoc., 2018, pp. 1920–1930.
94. Verzola I., et al., Project Sibyl: a novelty detection system for human spaceflight operations, *SpaceOps 2016 Conf.*, AIAA, 2016. <https://doi.org/10.2514/6.2016-2405>
95. ELKI Data Mining Framework. <https://elki-project.github.io/>. Cited November 2, 2020.
96. *GOCE Telemetry Data Collection*, Eur. Space Agency, 2019. <https://doi.org/10.5270/esa-7nc8pjp>
97. Ganin, Y. et al., Domain-adversarial training of neural networks, *J. Mach. Learn. Res.*, 2016, vol. 17, no. 59, pp. 1–35.
98. Von Rueden, L. et al., Informed machine learning—a taxonomy and survey of integrating knowledge into learning systems, 2020. [arXiv:1903.12394](https://arxiv.org/abs/1903.12394) [cs, stat]. Cited April 9, 2020.
99. Fink, O. et al., Potential, challenges and future directions for deep learning in prognostics and health management applications, *Eng. Appl. Artif. Intell.*, 2020, p. 103678.
<https://doi.org/10.1016/j.engappai.2020.103678>

This paper was recommended for publication by V.M. Glumov, a member of the Editorial Board