# An Anomaly Pattern Detection Method for Sensor Data

Han Li[1,2], Bin Yu[1,2(✉)], and Ting Zhao[3]

[1] College of Computer Science, North China University of Technology, Beijing, China
lihan@ncut.edu.cn, yubin0574@qq.com
[2] Beijing Key Laboratory on Integration and Analysis of Large-Scale Stream Data, Beijing, China
[3] Advanced Computing and Big Data Technology Laboratory of SGCC, Global Energy Interconnection Research Institute, Beijing, China
zhaoting@geiri.sgcc.com.cn

**Abstract.** With the development of the Internet of Things (IOT) technology, a large number of sensor data have been produced. Due to the complex acquisition environment and transmission condition, anomalies are prevalent. Sensor data is a kind of typical time series data, its anomaly refers to not only outliers, but also the anomaly of continuous data fragments, namely anomaly patterns. To achieve anomaly pattern detection on sensor data, the characteristics of sensor data are analyzed including temporal correlation, spatial correlation and high dimension. Then based on these characteristics and the real-time processing requirements of sensor data, a sensor data oriented anomaly pattern detection approach is proposed in this paper. In the approach, the frequency domain features of sensor data are obtained by Fast Fourier Transform, the dimension of the feature space is reduced by describing frequency domain features with statistical values, and the high-dimensional sensor data is processed in time on the basis of Isolation Forest algorithm. In order to verify the feasibility and effectiveness of the proposed approach, experiments are carried out on the open dataset IBRL. The experimental results show that the approach can effectively identify the pattern anomalies of sensor data, and has low time cost while ensuring the high accuracy.

**Keywords:** Sensor data · Fourier Transform · Frequency domain feature · Isolation Forest · Anomaly pattern detection

## 1 Introduction

In recent years, as an information carrier based on Internet and traditional telecommunication network, the Internet of Things (IOT) enables all ordinary physical objects to interconnect and exchange information, and supports more intelligent physical object management. The rapid development of IoT technology can provide more abundant data. Based on these data, more abundant data analysis can be carried out to provide more accurate services. The environment of IoT is extremely complex, and there are many problems such as equipment failure, signal interference, abnormal transmission,


W. Ni et al. (Eds.): WISA 2019, LNCS 11817, pp. 270–281, 2019.
https://doi.org/10.1007/978-3-030-30952-7_28

etc. Therefore, the sensor data inevitably have anomalies. For abnormal sensor data, if it is not processed and are directly analyzed, there are two potential problems. Firstly, abnormal data will affect the accuracy of data analysis results, resulting in invalid decision-making. Secondly, if the anomalies hidden in the sensor data can't be identified as early as possible, it is not conducive to timely discovery of the physical world problems, and may cause unnecessary losses. Therefore, anomaly detection for sensor data is particularly important. Reliable anomaly detection can not only make data analysis decision more effective, but also detect anomaly sensors in time to reduce losses.

Hawkins [1] defines anomalies as distinct data in a data set, which makes one suspect that these data are not random deviations, but are generated by completely different mechanisms. Anomaly detection is a process of discovering abnormal data in data resources by using various data processing models and technologies. It is the premise and necessary link of discovering data anomalies and improving data quality. Traditional anomaly detection mostly aims at outliers, and is based on statistics, distance, density and clustering methods. But sensor data is mostly time series data, and there are not only outlier abnormalities, but also timing fragment abnormalities, that is, pattern abnormalities. Among them, anomaly pattern means that there is no anomaly at any data point of time series data (such as the data value is no more than the threshold), but the trend of this data fragment is obviously different from that of other similar data fragments. In addition, it is common to use multiple sensor devices to monitor the same physical entity. For example, various types of sensor data are usually combined to describe the operating conditions of industrial equipment or the environment. Thus, sensor data are always high dimensional. Similar to the single-dimensional time series, high-dimensional time series also has the problem of pattern anomalies, and efficiency is considered as one of the most important issues. If the traditional outlier detection methods are used to these high-dimensional time series, there will be a problem of high time cost, and it is difficult to judge the overall abnormal situation according to the abnormal situation of a single point.

Therefore, this paper considers the high-dimensional time series as the sensor data, and an anomaly pattern detection method is proposed for these sensor data to improve the efficiency under the premise of ensuring accuracy.

## 2 Related Work

In recent years, as a branch of data mining, anomaly detection is receiving more and more attention. Among them, according to the object of anomaly detection, it is mainly divided into anomaly detection of outliers and anomaly detection of time series data.

Outlier detection can be roughly divided into four categories: statistical-based [2], distance-based [3], density-based [4] and clustering-based [5] methods. The statistical-based approach is a model-based approach that firstly creates a model for the data and then evaluates it [6]. However, such methods need sufficient data and prior knowledge, and are more suitable for outlier detection of individual attributes. The distance-based method is similar to the density-based method. The distance-based method is based on the distance between the point and other points [7], and the density-based method is

based on whether there are enough points in the neighborhood of the point [8]. These two methods are simple in thought, but generally require $O(n^2)$ time overhead. The cost is too high for large data sets, and the methods are also sensitive to parameters. Without proper parameters, the performance of the algorithm will be worse. Clustering-based method regards data that do not belong to any cluster as outliers [9]. Some clustering-based methods, such as K-means, have linear or near-linear time and space complexity, so this kind of algorithm may be highly effective for outlier detection. The difficulty lies in the selection of cluster number and the existence of outliers. The results or effects produced by different cluster numbers are completely different, so each clustering model is only suitable for a specific data type.

In the aspect of time series data anomaly detection, pattern anomaly detection algorithms for time series data have also made some achievements. Chen [10] et al. proposed the D-Stream clustering algorithm for clustering of time series data. Its main idea is to divide the data space into a series of grids in advance. By mapping the time series data to the corresponding grids, the results of grid processing can be obtained. However, this algorithm requires users to set more parameters in advance and the accuracy is low. Yan [11] et al. used the probability density function of data to re-express Euclidean distance, and obtained a probability measure to calculate the dissimilarity between two uncertain sequences. However, the detection effect of this algorithm depends on the size of the detection window. There is no suitable method to find the appropriate detection window size, and there are certain requirements for data. Cai [12] et al. proposed a new anomaly detection algorithm for time series data by constructing distributed recursive computing strategy and k-Nearest Neighbor fast selection strategy. However, the algorithm is effective for one-dimensional sensor data and does not consider multi-dimensional data.

Considering the problems of the above methods, this paper proposes an anomaly pattern detection method for high-dimensional sensor data. The method uses Fast Fourier Transform to transform time domain data into frequency domain data, and then realizes data dimensionality reduction through feature extraction. Finally, the anomaly patterns of sensor data are detected by using the spatial-temporal correlation characteristics of sensor data and the Isolation Forest algorithm.

## 3 Anomaly Pattern Detection Method for Sensor Data

### 3.1 Sensor Data Characteristics

Sensor data refers to the data collected by sensor devices, which is often used to continuously perceive the information of the physical world. Therefore, sensor data has many special characteristics. In order to identify the pattern anomalies of sensor data more pertinently, the main characteristics of sensor data are analyzed as follows:

(1) Time continuity: Sensor data is generated continuously by sensors. Generally, sensor data acquisition is carried out according to a certain frequency, so time continuity is one of the most basic characteristics of sensor data.
(2) Spatial-temporal correlation: Sensor data is usually used to perceive the information of the physical world, so the sensor data will be associated with the

physical world it perceives, and the association is specifically expressed as spatial and temporal correlation. That is to say, in different time or space environment, the sensor data collected by the same sensor may also be different.

(3) Data similarity: Based on the spatial-temporal correlation of sensor data, the sensor data collected by the same type of sensors in the similar time and space range have similarity. For example, different sensors which are set up in the same environment detect the similar environmental indicators at the same time. In addition, if the monitoring object of the sensor has similar behavior, the sensor data used to describe the behavior should also have similarity. For example, if users have similar electricity consumption patterns, their meter data fluctuations should be roughly similar. If there are abnormalities, there may be problems such as electricity theft. Therefore, it can be considered that similar sensor data have data similarity under similar time, space or behavior conditions.

(4) High-dimensionality: In the actual production environment, the single-dimensional sensor data has the problem of not being able to describe the complex physical world. Therefore, it is often necessary to combine various types of sensor data to describe the state of a physical entity, thus forming a high-dimensional sensor data. Taking the sensor data for monitoring the working conditions of thermal power generators as an example, the dimensions of the data are up to dozens.

### 3.2 The Proposed Approach

The goal of anomaly pattern detection method for sensor data is to quickly identify anomaly patterns in high-dimensional sensor data. The method is divided into two stages: data preprocessing and anomaly pattern detection.

The goal of data preprocessing stage is to reduce the data dimension for the accurate and fast detection of abnormal patterns on the premise of ensuring the characteristics of sensor data patterns. Briefly, the Fast Fourier Transform is used to transform the time series data into the frequency domain data, and then the characteristics of the frequency domain data are extracted for dimension reduction.

The goal of anomaly pattern detection stage is to improve the efficiency on the premise of ensuring accuracy. Briefly, based on the spatial-temporal correlation and data similarity of the sensor data, the sensor data with obvious difference from the pattern of the adjacent sensor data is found by comparing the time and space related sensor data. In order to solve the problem of fast processing of high-dimensional data, the method adopts the idea of ensemble learning and detects anomaly patterns based on Isolation Forest algorithm.

**Data Preprocessing.** Due to the time continuity and high-dimensionality of the sensor data, the sensor data segment consists of a large number of temporally consecutive high-dimensional data points. It is not only difficult to directly describe the pattern of the sensor data segments by these continuous high-dimensional data points, but also difficult to quickly identify the pattern abnormality. Therefore, the data preprocessing stage mainly focuses on the extraction of sensor data pattern features and the dimensionality reduction of data features. Figure 1 shows the workflow of the data preprocessing stage, which consists of feature extraction and feature reduction.

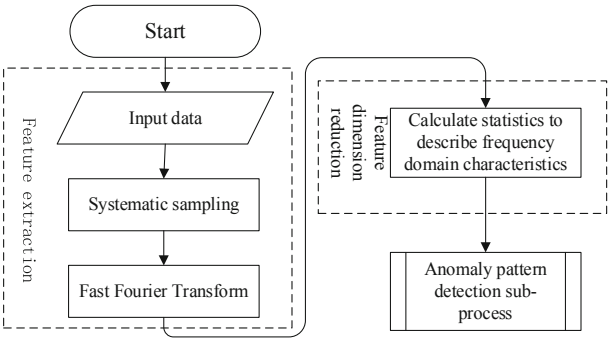


**Fig. 1.** Workflow of the data preprocessing stage

(1) Feature extraction

Time-frequency transform is used to discard the single difference of sensor data in time domain, and frequency domain features are used to describe the pattern characteristics of data fragments. In order to ensure the efficiency of anomaly pattern detection, Fast Fourier Transform (FFT) with low time complexity is used to transform time domain data to frequency domain data. Fast Fourier Transform makes full use of the symmetry and periodicity of exponential factors in Discrete Fourier Transform (DFT) formulas, and then obtains the results of Discrete Fourier Transform corresponding to these short sequences and combines them appropriately, so as to delete repetitive computation, reduce multiplication and simplify the structure. Compared with the time complexity of $O(n^2)$ in the Discrete Fourier Transform, the Fast Fourier Transform can reduce the time complexity to O (nlogn) level. In the case of a larger amount of data, the advantage of the Fast Fourier Transform in terms of time is more obvious. Assuming that the sensor data segment is $T = \{T_1, T_2, \ldots, T_i, \ldots, T_n\}$, and $T_i$ is the data segment of the i-th dimension data. The extraction step of the sensor data feature is as follows: Firstly, the data fragments of all dimensions in T are sampled by equidistant sampling method, requiring that each dimension has $2^m$ sample points $\{t_1, t_2, \ldots, t_j, \ldots, t_2^m\}$, and $t_j$ is the j-th sample point, and $2^m$ should be as close as possible to the number of original samples in this dimension. Secondly, performing Fast Fourier Transform on the sample data on the data segment of each dimension, and obtaining n frequency domain data sets $\{F_1, F_2, \ldots, F_i, \ldots, F_n\}$, and $F_i$ is the frequency domain data of the i-th dimension data fragment in sensor data.

(2) Feature dimension reduction

In order to ensure that the frequency domain data can depict the time domain data as accurately as possible, the sample density of the time domain data is high. Therefore, the frequency domain data obtained by time-frequency transformation also has a high data density, that is, a large amount of data. For the purpose of ensuring the fast processing of high-dimensional data fragments, the method has to reduce the dimensionality of data. Because the amplitude of the sensor data at a certain frequency in the frequency domain space is directly related to the modulus of the result value of the fast

Fourier transform at that frequency. In order to reflect the concentration trend, the degree of dispersion and the maximum amplitude of the sensor data in the frequency domain space, the mean, variance and peak value of each dimension frequency domain data module of the sensor data fragment are selected as the frequency domain characteristics of the sensor data fragment.

**Anomaly Detection.** In this paper, the anomaly pattern of sensor data is defined as a pattern with "few but different" characteristics compared with other patterns. Based on the definition, the anomaly pattern does not only refer to the wrong pattern, but mainly emphasizes the specificity of the pattern. The pattern anomaly of high-dimensional sensor data can appear in any dimension, so it is necessary to detect the data of each dimension to identify the anomalies of the sensor data. Obviously, it would be more accurate to detect pattern anomalies in each data dimension, but the computational efficiency will be very low. Therefore, the goal of anomaly detection in this paper is to improve the efficiency of anomaly detection on the premise of ensuring accuracy.

Isolation tree is a random binary tree, which is similar to decision tree, can be used for data classification. When constructing binary tree, the isolation tree randomly selects attribute values, and identifies anomalies by the depth (also known as isolation depth) of the data in the binary tree. Therefore, the algorithm not only makes full use of the "few but different" characteristics of anomalies, but also has good processing performance. However, due to the randomness of the construction process of isolation tree, it is a weak classifier. Therefore, it is necessary to adopt ensemble learning method to improve the generalization performance of anomaly detection. In this paper, the anomaly detection is implemented by the Isolation Forest algorithm [13] based on bagging ensemble learning method and isolation tree algorithm. The workflow is shown in Fig. 2.

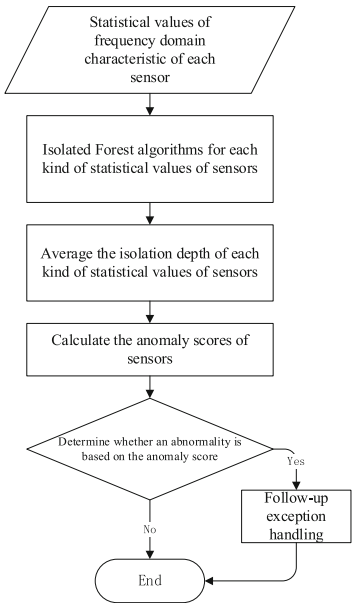


**Fig. 2.** Workflow of the anomaly detection stage

As shown in Fig. 2, anomaly detection consists of three steps: Firstly, the statistical values of each sensor are processed by the Isolation Forest algorithm to obtain the corresponding isolation depth $\{D_1, D_2, \ldots, D_i, \ldots, D_n\}$, where $D_i$ is the isolation depth of the i-th sensor, including the average isolation depth $\{m_i, v_i, p_i\}$ of mean, variance and peak value. Secondly, calculating the total average isolation depth $\{d_1, d_2, \ldots, d_i, \ldots d_n\}$ according to the statistical values, where $d_i$ is the isolation depth of the i-th sensor, and $d_i = (m_i + v_i + p_i)/3$, which is used to calculate anomaly score. Thirdly, the anomaly score is calculated for each sensor and the anomaly is judged according to the score. According to the definition of the anomaly score of the Isolation Forest algorithm, this paper defines the anomaly score s(i, M) of the i-th sensor as $2^{-n}$, where n is the ratio of the average isolation depth $d_i$ of the i-th sensor to the average path length c (M) of the M isolation trees used to construct the isolation forest. The calculation method is as shown in Eq. (1), and the value of c(M) is calculated by Eq. (2).

$$s(i, M) = 2^{-\frac{d_i}{c(M)}} \tag{1}$$

$$c(M) = 2\ln(M-1) + \xi - \frac{2(M-1)}{M}, \text{where } \xi \text{ is Euler constant} \tag{2}$$

Based on the definition and calculation method of anomaly score, the anomaly score s(i, M) has the following properties: Firstly, the range of anomaly score is [0,1], and the closer to 1, the higher the probability of anomaly. Secondly, if the anomaly scores of all samples are smaller than 0.5, it can be basically determined as normal data. Thirdly, if the anomaly scores of all samples are around 0.5, the data does not contain significant anomalous samples. According to the above properties of anomaly score, the anomaly of sensor data can be determined. It should be noted that anomalies are relative, and the distribution of anomaly scores generated by different data sets is different, so the specific criteria for determining anomalies are also different.

## 4 Experiments and Results

In order to evaluate the effectiveness of the method, the IBRL [14] (Intel Berkeley Research Lab) data set is used for verification. The wireless sensor network is deployed at the Intel Research Lab at Berkeley University and contains 54 Mica2Dot sensor nodes. The sampling period is from February 28, 2004 to April 5, 2004, and sampling is performed every 30 s to obtain a set of data. Figure 3 shows the deployment diagram of the network. The location of each node in the network is represented by a black hexagon. The white number is the ID of each node. In this network, each node collects four types of values, namely temperature, humidity, light and voltage. Since the wireless sensor network is deployed in the same laboratory, except for the sudden change of illumination data caused by frequent switching operations, other data sampling values are relatively stable, and it can be considered that the data obtained by the sensors are similar. The experiment selects the data from 08:00:00-24:00:00 on February 28, 2004. However, based on data observation and analysis, some observation data is lost due to network packet loss and other reasons. In order to ensure the

reliability of the experiment, the missing values are interpolated based on the average or spatial correlation characteristics, and the values sampled in the time period of 15 s to 20 s per minute are used.

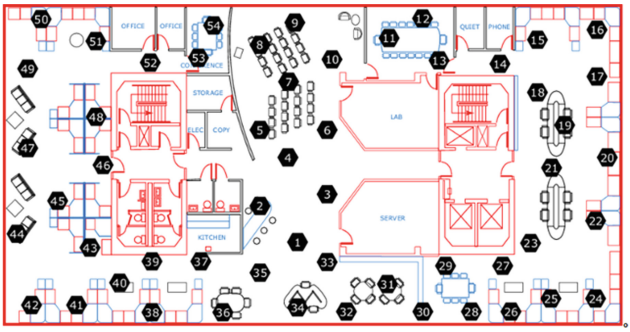


**Fig. 3.** Node distribution in the IBRL data set

To evaluate the performance of the abnormal pattern detection algorithm, some normal observation values were randomly modified to become abnormal data. To avoid the generality, the distribution of abnormal data should be different from the distribution of normal data, and the sample space should overlap as much as possible. In addition, the anomaly pattern should be a small probability event relative to the normal sample set collected by the non-faulty node. Therefore, the abnormal data generated by the simulation has a slight deviation from the normal sample data distribution. The details of the data set are shown in Tables 1 and 2. Table 1 shows the details of the 64-min data for a single experiment, and Table 2 shows the overall data details of this experiment.

**Table 1.** Data set per 64 min

| Time | Number of normal sensors | Number of abnormal sensors | Number of normal samples | Number of abnormal samples |
|---|---|---|---|---|
| Every 64 min | 52 | 2 | 3416 | 40 |

**Table 2.** Overall data set situation

| Time | Total normal sensor cumulative value | Total anomaly sensor cumulative value | Total number of normal samples | Total number of anomaly samples |
|---|---|---|---|---|
| 8:00-24:00 | 780 | 30 | 51240 | 600 |

In order to show the differences between the abnormal pattern and the normal pattern, the normal data of 8:00:17-8:30:17 is shown in Fig. 4.

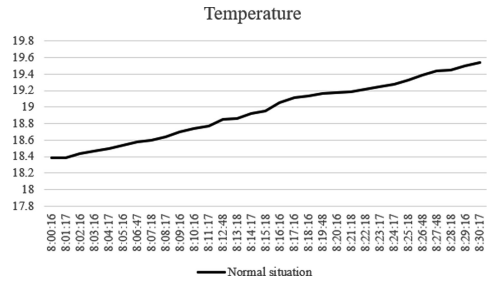


**Fig. 4.** Normal pattern

There are two kinds of abnormal patterns, which are mutation abnormalities and trend abnormalities, as shown in Fig. 5.

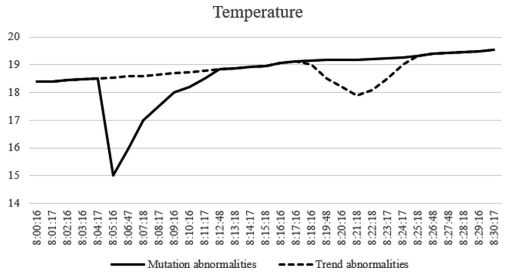


**Fig. 5.** Mutation abnormalities and Trend abnormalities

### 4.1 Evaluation Indicators

- Recall rate
  It is also known as True Positive Rate (TPR) or sensitivity (sensitivity). This indicator corresponds to the proportion of all abnormalities actually detected in the data set, that is, the ratio of correctly detected anomalies to all anomalies in the data set.
- False Positive Rate (FPR)
  It is also known as the false detection rate. The indicator corresponds to the ratio of the normal sample in the data set that is incorrectly judged to be abnormal, that is, the ratio of the sample misjudged as abnormal to all normal samples in the data set.
- Accuracy rate
  The indicator corresponds to the ratio of the correctly predicted samples in the data set, that is, the ratio of the correctly predicted samples to all samples in the data set.

### 4.2 Experimental Analysis

Since two sensors are selected to simulate 20 outliers every 64 min in the sensor data of 54 sensors, the experiment takes the two sensors with the largest abnormality score as sensors with abnormal pattern, and then observe whether the simulated anomaly sensor is correctly detected. Due to the randomness of the Isolation Forest, the experiment conducted 10 times on the data set. Table 3 shows the experimental results corresponding to each experiment.

**Table 3.** Experimental results

| Data set time | Recall rate | False Positive Rate | Accuracy rate |
|---|---|---|---|
| 8:00-24:00 | 0.8 | 0.0077 | 0.985 |
| 8:00-24:00 | 0.833 | 0.0064 | 0.987 |
| 8:00-24:00 | 0.8 | 0.0077 | 0.985 |
| 8:00-24:00 | 0.867 | 0.0051 | 0.99 |
| 8:00-24:00 | 0.833 | 0.0064 | 0.987 |
| 8:00-24:00 | 0.8 | 0.0077 | 0.985 |
| 8:00-24:00 | 0.8 | 0.0077 | 0.985 |
| 8:00-24:00 | 0.833 | 0.0064 | 0.987 |
| 8:00-24:00 | 0.8 | 0.0077 | 0.985 |
| 8:00-24:00 | 0.8 | 0.0077 | 0.985 |

The experimental results show that the recall rate of the results using the proposed method maintains above 80%, indicating that most sensors with abnormal pattern can be detected. At the same time, by observing the experimental data of the false-detection sensor, it is found that the misdetection of most sensors may be caused by the sudden change of illumination. Because the experimental data is collected under real conditions, there are cases of switching lights or curtains. For example, the No. 20 sensor by the window and the No. 11 sensor in the closed environment both have cliff-like rise or fall in the value of illumination attributes. As a result, the abnormal pattern is more obvious than the abnormal pattern simulated in this experiment, and the final calculated abnormality score is also higher, resulting in false detection.

### 4.3 Comparisons

In order to verify the efficiency of the proposed method, a comparative experiment is carried out between the commonly used anomaly detection algorithms DBSCAN, One Class SVM and the proposed method. In order to ensure that the anomaly can be detected normally, the statistical values after the fast Fourier transform are used as the data set of this comparative experiment. Table 4 shows the results.

**Table 4.** Efficiency of the three methods

| Algorithm name | Run time/s | Accuracy rate |
|---|---|---|
| Isolation Forest | 1.716 s | 0.995 |
| DBSCAN | 4.400 s | 0.935 |
| One Class SVM | 33.114 s | 0.966 |

According to the above results, the advantage of the isolation forest algorithm in efficiency is relatively obvious on the basis of ensuring accuracy. Because DBSCAN and One Class SVM are density-based spatial clustering algorithms, which need to calculate the distance from the core point to the surrounding, and if the amount of data is larger, the computational complexity is higher.

## 5 Conclusions

Aiming at the pattern abnormality of sensor data, this paper analyses the characteristics of sensor data, including time continuity, spatial-temporal correlation, data similarity and high dimensionality. Considering these characteristics, an anomaly pattern detection method for sensor data is proposed under the consideration of efficiency and accuracy. In the approach, the frequency domain features of sensor data are obtained by Fast Fourier Transform, and then the dimension of feature space is reduced by describing frequency domain features with statistical values. Finally, the sensor with anomaly patterns is determined by anomaly detection using Isolation Forest algorithm. The experimental results show that the method not only has advanced accuracy, but also has obvious advantages in running time and can effectively identify the pattern anomalies of sensor data. In the future, the main work is to carry out more extensive experiments and to study anomaly pattern detection methods for more complex sensor data.

**Acknowledgements.** This paper is supported by the Scientific and Technological Research Program of Beijing Municipal Education Commission (KM201810009004) and the National Natural Science Foundation of China (61702014).

## References

1. Hawkins, D.M.: Identification of Outliers. Springer, Netherlands (1980). https://doi.org/10.1007/978-94-015-3994-4
2. Xi, Y., Zhuang, X., Wang, X., et al.: A research and application based on gradient boosting decision tree. In: 15th International Conference on Web Information Systems and Applications, pp. 15–26 (2018)
3. Ramaswamy, S., Rastogi, R., Shim, K.: Efficient algorithms for mining outliers from large data sets. ACM SIGMOD Rec. **29**(2), 427–438 (2000)

4. Frank, R., Jin, W., Ester, M.: Efficiently mining regional outliers in spatial data. In: Papadias, D., Zhang, D., Kollios, G. (eds.) SSTD 2007. LNCS, vol. 4605, pp. 112–129. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73540-3_7
5. Gaddam, S., Phoha, V., Balagani, K.: K-Means+ID3: a novel method for supervised anomaly detection by cascading K-Means clustering and ID3 decision tree learning methods. IEEE Trans. Knowl. Data Eng. **19**(3), 345–354 (2007)
6. Kasliwal, B., Bhatia, S., Saini, S., et al.: A hybrid anomaly detection model using G-LDA. In: 2014 IEEE International Advance Computing Conference, Gurgaon, pp. 288–293 (2014)
7. Zhang, Y., Du, B., Zhang, L., et al.: A low-rank and sparse matrix decomposition-based Mahalanobis distance method for hyperspectral anomaly detection. IEEE Trans. Geosci. Remote Sens. **53**(3), 1–14 (2015)
8. Huang, T., Zhu, Y., Zhang, Q., et al.: An LOF-based adaptive anomaly detection scheme for cloud computing. In: 37th Annual Computer Software and Applications Conference Workshops, Japan, pp. 206–211 (2013)
9. Münz, G., Li, S., Carle, G.: Traffic anomaly detection using K-Means clustering. In: 4th GI/ITG-Workshop MMBnet, Hamburg (2007)
10. Chen, Y.: Density-based clustering for real-time stream data. In: 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, California, pp. 133–142 (2007)
11. Yan, Q.Y., Xia, S.X., Feng, K.W.: Probabilistic distance based abnormal pattern detection in uncertain series data. Knowl. Based Syst. **36**(11), 182–190 (2012)
12. Cai, L., Thornhill, N., Kuenzel, S., et al.: Real-time detection of power system disturbances based on k-nearest neighbor analysis. IEEE Access **99**, 1–8 (2017)
13. Liu, F., Ting, K., Zhou, Z.H.: Isolation forest. In: 8th IEEE International Conference on Data Mining, Los Alamitos, pp. 413–422 (2008)
14. Intel Berkeley Research Lab dataset. http://db.csail.mit.edu/labdata/labdata.html. Accessed 18 Apr 2019