Received June 25, 2019, accepted July 7, 2019, date of publication July 23, 2019, date of current version August 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2930627

Hierarchical Edge Computing: A Novel Multi-Source Multi-Dimensional Data Anomaly Detection Scheme for Industrial Internet of Things

YUHUAI PENG[1,2], AIPING TAN[1], JINGJING WU[1], AND YUANGUO BI[1]

[1]School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China

[2]Key Laboratory of Vibration and Control of Aero-Propulsion System, Ministry of Education, Northeastern University, Shenyang 110819, China

Corresponding author: Aiping Tan (aipingtan@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61871107, Grant 61501105, and Grant 61701102, and in part by the Fundamental Research Funds for the Central Universities under Grant N171612014, Grant N181613002, Grant N180708009, Grant N170308028, and Grant N180703018.

ABSTRACT Every year, many people around the world die because of mining accidents. Industrial Internet of Things (IIoT) can be employed to sense public safety hazards and provide early warning of accidents, thereby ensuring safe operations at underground mining, personnel positioning, and specific items supervision and emergency response. Real-time data anomaly detection can predict the probability of occurrence of the abnormal event. However, massive heterogeneous monitoring data, poor wireless environment and data spatio-temporal association have posed a serious challenge to data anomaly detection for underground mining. Existing methods are mostly concerned about single data or processing at cloud platform, with little regard for the time and space association. Focus on the accuracy and timeliness of data anomaly detection, a novel multi-source multi-dimensional data anomaly detection scheme based on hierarchical edge computing model is presented in this paper. Firstly, a hierarchical edge computing model is proposed to realize load balance and low-latency data processing at the sensor end and base-station end. Then a single-source data anomaly detection algorithm is designed based on fuzzy theory, which can comprehensively analyze the anomaly detection results of multiple consecutive moments. Finally, a multi-source data anomaly detection algorithm executed at the base-station end is designed to consider the sensing data associated attributes of time and space. Experimental results reveal that the proposed scheme has higher detection accuracy and lower processing delay compared with traditional solutions.

INDEX TERMS Industrial Internet of Things (IIoT), underground mining, anomaly detection, multi-source multi-dimensional data, edge computing.

I. INTRODUCTION

The Internet of Things (IoT) can connect various things to the network through information sensing devices and communication protocols, and conduct information exchange and communication through information media, to implement intelligent identification, positioning, tracking, supervision and other functions [1]–[3]. In the past few years, the IoT has been widely popularized in transportation, security, medical, industrial manufacturing and other fields, of which Industrial Internet of Things (IIoT) customized for manufacturing processes has received increasing attention from academia and industry [4]. Facing frequent mine accidents that cause a large number of casualties and property losses, it is imperative to construct a public safety monitoring IIoT to sense public safety hazards and provide early warning of accidents, thereby ensuring safe operations at underground mining, personnel positioning, and specific items supervision and emergency response, etc. Unlike other industrial scenarios [5], mining operations require construction workers to work in underground tunnels. However, the construction environment has the characteristics of closed and narrow space, which

The associate editor coordinating the review of this article and approving it for publication was Anfeng Liu.

may cause hidden dangers such as insufficient oxygen and temperature, to seriously threaten the personal safety of underground construction workers. At present, underground mining mainly includes operations for metal mines and coal mines. Under the conditions of weak light, unstable temperature, release of toxic gases, lack of oxygen, frequent collapse and explosion, the safety of mining workers poses a serious challenge to mining. For the complex mining environment, there is no perfect solution for mine safety monitoring and emergency treatment currently. In particular, traditional preventive measures still require manual treatment without achieving expected goal [6]. Every year, many people around the world die because of mine disasters. In 2018, there were more than 200 accidents in China's coal mines alone, and more than 300 people died [7]. This figure is still distressing. Therefore, employing IIoT to carry out safety monitoring and accident warning is of great practical significance for ensuring the safe production of underground mining.

There are still many challenges to accurately and timely warn safety: unreliable wireless transmission, abnormal detection of multi-source data, and energy consumption in harsh environments [8], [9]. Real-time data abnormal detection has important practical applications. In underground mining, different sensor nodes periodically collect information such as temperature, humidity, harmful gas concentration and personnel sign parameters, etc. By detecting the abnormality of the sensory data stream, the probability of occurrence of the abnormal event can be predicted, to conduct timely response and handling. The particularity of mining, as well as the inherent characteristics of sensory data and sensors make anomaly detection face many challenges. First of all, the mining environment needs to deploy multiple types of sensors to collect data, mainly including temperature, humidity, gas, wind speed, stress, displacement, etc. So how to perform anomaly detection on distributed sensor data is a problem. Secondly, the data anomaly detection for safety warning needs to consider the low-latency demand. The traditional methods mostly perform anomaly detection and decisions in the cloud, and it is difficult to meet the real-time of anomaly detection. Finally, data collection is geographically relevant and time-sensitive, which relies heavily on the time and geographic location information of the collection node. How to determine their relevance and accurate data anomaly detection is a challenge.

At present, Research of IIoT is mainly focused on mainly focus on energy consumption [10], delay problem [11], [12], channel access [13], [14], industrial applications [15]–[17] and so on. Until now, there has already some research about data anomaly detection in IIoT application scenarios. Wang *et al.* in [18]adopts Bayesian network to detect anomaly of wireless sensor network in coal mine. This method can learn the data of gas concentration, grasp the periodical change rule, and reduce the fault caused by individual error data. Moreover, the method can learn the correlation between different types of gases and concentration changes at multiple locations. Oliver Obst *et al.* in [19] studies the anomaly detection of gas concentration in coal mine using wireless sensor networks. In addition, a method based on echo state network is proposed, which uses machine learning theory to train normal data and is used to detect data anomalies. Compared with Bayesian method, this method has higher detection accuracy. Chen *et al.* in [20] proposes a fully distributed and general data anomaly detection method based on graph theory, which analyzes the temporal and spatial correlation of industrial field data and proposes an anomaly detection model for large-scale data. This method can effectively realize data anomaly detection in building construction and smart grid monitoring. Soydan *et al.* in [21] used image analysis method to monitor the mining process of a coal mine in Turkey. However, this method is mainly oriented to underground structure and does not analyze the environmental data, such as gas and temperature. Tan et al. proposed a multi-channel TDMA scheduling method for IIoT in underground mining environment, which can ensure the reliability of wireless transmission. And the network topology adopted in this paper is the same as in [22].

Existing methods of data anomaly detection are mostly concerned about single data and decision at cloud, with little regard for the time and space association. Focus on the accuracy and timeliness of anomaly detection problem, a novel multi-source multi-dimensional data anomaly detection scheme based on hierarchical edge computing model is proposed for early safety warning of underground mining. The main contributions of this paper can be listed as follows:

1) According to the special hybrid topology of the IIoT for mining operation monitoring, a hierarchical edge computing model is proposed, which can perform multi-source data anomaly detection at different ends: collection end (sensors) and sink end (base-stations), to realize load balance of the whole system and low-latency data processing on the premise of ensuring low energy consumption.
2) A single-source data anomaly detection algorithm is designed which would be executed at both the sensor end and base-station end. Based on fuzzy theory, this algorithm establishes an anomaly detection function, and considers the data monitoring values at adjacent time to comprehensively analyze the abnormal data detection results of multiple consecutive moments, so as to avoid the error of the anomaly detection results due to the one-sided estimation of the single time data.
3) Also, a multi-source data anomaly detection algorithm executed at the base-station end is presented. According to the special tunnel structure in underground mining, the definition of sensors location correlation based on distance is proposed. When the anomaly detection algorithm works, sensors with location correlation would be selected for detection and comprehensive analysis to avoid errors caused by a certain sensor fault.

The reminder of the paper is organized as follows: Section II reviews the related work in data anomaly detection.

Section III described network model and problem formulation. In section IV, a multi-source multi-dimensional data anomaly detection scheme is presented, which includes design of hierarchical edge computing model, a single-source data anomaly detection algorithm and a multi-source data anomaly detection algorithm. Section V carries on the experimental verification and results analysis. Finally, section VI gives the conclusion for this paper.

II. RELATED WORK

At present, the research on IIoT is mainly focused on the real-time processing, security and reliable transmission [23]–[27]. Most research of data anomaly detection is aimed at wireless sensor networks (WSN). Oluwasanya *et al.* in. [28] conduct a survey of data anomaly detection in WSN. Although many literatures studies anomaly detection problem, there are not many solutions for practical application scenarios. Many methods are based on mathematical models, and whether they can effectively solve practical problems remains to be verified. Rassam *et al.* in [29] propose a new anomaly detection model, in which before the sensor data is sent to the base-station, the local anomaly detection of sensor measurements is carried out. This method not only ensures certain detection accuracy, but also reduces energy consumption. However, this solution has higher CPU requirements for the sensor. Salem *et al.* in [30] proposed a data anomaly detection method for medical WSN, which can effectively detect anomaly changes, and timely warn dynamic changes. However, this method requires high reliability and real-time transmission in wireless environment. The research progress of anomaly detection in WSN can be summarized as follows:

1) *Anomaly Detection Method Based on Statistics:* Anomaly detection method based on statistics is old and mature, which creates a distribution model for the dataset and fits the target data object. It is assumed that the normal data falls in the high probability interval while the outliers are relatively in the low probability interval. Finally, the abnormality is judged according to the probability of the object in the target data set falling in the model. Rajasegarar *et al.* in [31], [32] propose a classification method for anomaly detection models and established two detection models: statistical detection models and non-parametric detection models. These two models can be applied in different scenarios, where the statistical model is suitable for applications with data types and sampling periods pre-determined, while the non-parametric model conducts detections by the behavior of current data and adjacent data with no prior knowledge. Fei et al. in [33] propose a multi-source data anomaly detection method, which performs detection by statistical ways. This method is mainly applied to the platform space, and determines the relationship between two nodes by two-dimensional coordinate positions. Ren et al. in [34] study the time series based anomaly detection method and proposed an anomaly detection method based on probability interval statistics. The algorithm has higher data recognition than the aggregation algorithm. Djenouri *et al.* in [35] investigate the application of anomaly detection methods in urban traffic analysis, with a focus on detection methods based on outliers. According to the survey, the current outlier detection method can effectively analyze traffic data, but in cities with complex traffic conditions, the application effect is limited.
2) *Anomaly Detection Method Based on Distance:* Distance-based methods are usually built on the same basic assumption that normal data objects are closer to each other and abnormal data objects and normal data objects are far apart. In the case where the attribute variables of the data object are continuous, the Euclidean distance is usually used to measure the near-far relationship between the data objects. Sricharan et al. in [36] proposed a model for determining the relationship between adjacent nodes. Based on the statistical properties of K-NN density estimates, they derived the deviations and variances of the insertion estimates in terms of sample size, sample dimensions, and potential probability distributions. Bosman *et al.* in [37] proposed a data anomaly detection method based on neighbor node information. This method uses a machine learning algorithm to implement data anomaly detection through distributed processing, thereby reducing communication overhead. Xie *et al.* in [38] proposed a distributed anomaly detection method, which analyzes the data of adjacent nodes and uses a distributed global probability density estimation method to measure the data values of adjacent time. This method effectively solves the problem of traditional single-source data detection, and can comprehensively analyze the information of adjacent nodes to perform data anomaly detection.
3) *Anomaly Detection Method Based on Clustering:* Clustering is to classify similar or related data objects into a cluster [39]. If a data object cannot be classified into any cluster during the clustering process, then the data object can be considered abnormal. Emadi and Mazinani in [40] analyzed the anomaly detection of data integrity in WSN. Through the characteristics of temperature, humidity and voltage, they adopt clustering method to analyze and complete data anomaly detection. This method can guarantee high detection accuracy. Seo et al. in [41] proposed a data anomaly detection method for micro-clustering, and designed a method for detecting and specifying outliers using a local outlier as a center of a micro-cluster in offline components. Sricharan *et al.* in [42] defined the outliers of data in WSN, and proposed a method based on data classification to estimate and calculate by probability density function. This method has been proven to be applicable to different types of data testing, including

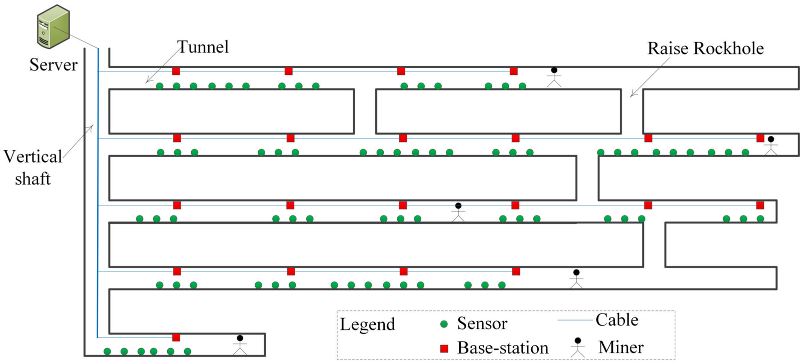


FIGURE 1. System topological structure.

Gaussian distribution. According to the dynamic characteristics of WSN, Rassam *et al.* in [43] analyzed the data anomaly detection models in static and dynamic environments respectively, and measured the similarity of sensor data by One-Class Principal Component Classifier (OCPCC). Through the incremental learning method, it can dynamically detect data changes. Ahmad *et al.* in [44] proposed an anomaly detection method based on K-Medoid custom clustering technology, which detects behaviors such as misguided attacks. By defining the detected parameters, a data anomaly detection model is established. The method realizes dynamic detection by setting the threshold, and is mainly applied in the field of remote sensing.

4) *Anomaly Detection Method Based on Artificial Intelligence:* At present, artificial intelligence theory has been applied in data anomaly detection. Among them, related algorithms represented by deep learning and machine learning have solved some problems. Through the artificial intelligence method, the detection accuracy can be improved by training of big data sets. Ramotsoela *et al.* in [45] conducted a survey of data anomaly detection methods in underwater WSN, and focus on machine learning-based methods, which can effectively improve the accuracy of detection, but usually need to be deployed in the cloud. Kwon *et al.* in [46] investigated the application of deep learning in anomaly detection. At present, the use of deep learning or machine learning methods for anomaly detection can effectively improve the accuracy of anomaly detection, and plays an important role in artificial intelligence and image recognition.

In addition, Pham *et al.* in [47] proposed an anomaly detection method for large-scale data mining applications based on spatial analysis and spectral anomaly detection to detect the original loss. This method can be applied to continuous big data case, such as video streaming. There are many data anomaly detection methods [48] in current WSN, and they have achieved very good results. However, there is still few research for special fields. In particular, the abnormality detection for underground mining scenario remains to be further studied.

III. SYSTEM MODEL AND PROBLEM FORMULATION

The underground mine can be divided into several tunnels according to the specific mining requirements, and each lane is connected by a vertical shaft. Therefore, the deployment of mining safety monitoring network adopts a hybrid network topology, in which each lane consists of a chain structure of a wired network composed of several base-stations, and each base-station contains a number of sensors, which constitutes a star topology of a wireless sensor network. Different tunnels are connected to remote cloud servers via the main optical fiber of vertical shaft. Therefore, the overall topology is a hybrid bus network with star topology and chain topology. Fig. 1 depicts the overall topology of the system. Each sensor collects environmental data, periodically sends wireless data to the base station, and the base station sends it to the server through the wired link. Finally, the server analyzes and processes the environmental safety monitoring data. In general, the structure of each roadway is basically the same and independent of each other, so an anomaly detection scheme based on a roadway's chain and shape hybrid topology is designed. Suppose there are a total of $N(N \in \mathbb{N}^+)$ base stations (sink nodes) in a tunnel, and there are several sensors under each base-station, assuming the number of sensors is $c^i(1 \leqslant i \leqslant N)$. The sensor collects the monitoring data, and periodically sends the wireless data to the base-station. There are many types of sensors in underground mining, and the data of different types of sensors are heterogeneous and have different data transmission periods (assuming that different sensors of the same type have the same data transmission period), the data transmission period of the j-th sensor of the i-th base-station is denoted by p_i^j. The data of the j-th sensor of the i-th base station at time $t(t \in \mathbb{N}^+)$ is denoted by $d_i^j(t)\,(t \in \mathbb{N}^+, i \in [1, N], j \in [1, c^i])$. A continuous time data stream $d_i^j(1), d_i^j(2), \ldots, d_i^j(T)$ is formed over a period of time $T(T \in \mathbb{N}^+)$, while all sensors of the i-th base station form a matrix D_i^T of data streams over a period of time 1 to T, which can be expressed as:

$$D_i{}^T = \begin{bmatrix} d_i^1(1) & d_i^1(2) & \ldots & d_i^1(T) \\ d_i^2(1) & d_i^2(2) & \ldots & d_i^2(T) \\ \ldots & \ldots & \ldots & \ldots \\ d_i^{C_i}(1) & d_i^{C_i}(2) & \ldots & d_i^{C_i}(T) \end{bmatrix} \quad (1)$$

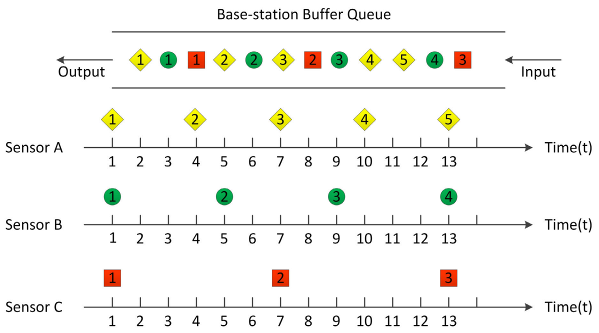


FIGURE 2. Base-station data buffer queue.

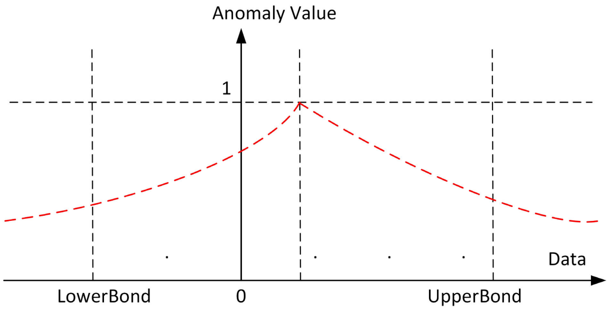


FIGURE 3. Data anomaly detection based on fuzzy theory.

Because different sensors differ in type, start-up time, and transmission period, it is possible that at time $t(1 \leqslant t \leqslant T)$, a sensor does not transmit data, that is, the value of $d_i^j(t)$ is empty. So the actual data stream is not a time-consecutive matrix. Therefore, as shown in Fig.2, the base-station uses a data buffer queue to receive sensor data. Sensors A, B, C under one base-station i have different data transmission periods, i.e. $p_i^A = 3$, $p_i^B = 4$, $p_i^C = 6$. At time 1, the three sensors initiate transmission of data, after which the amount of data received by the base-station is different at the same time, and reaching a maximum at time 13. As the receiving end, the base station will store the received data in a buffer queue at each time, waiting for anomaly detection.

Assume that the buffer queue of the base-station $i(1 \leqslant i \leqslant N)$ is represented as $Q_i = \{d_i^j(t) | i \in [1, N], j \in [1, c^i], t \in \mathbb{N}^+\}$, the objective of anomaly detection is to detect the data in the queue in turn. The anomaly detection is divided into two parts: single-source data anomaly detection and multi-source data anomaly detection.

A. FORMULATION OF SINGEL-SOURCE DATA ANOMALY DETECTION

The goal of single-source data anomaly detection is to analyze the data of one sensor. This kind of detection is only based on the normal value range of a certain type of sensor to detect anomalies, which can detect sensor data anomalies at a certain time. Assume that the lower bound of the normal data of the j-th sensor of the i-th base-station is $\varphi_{i,j}^L$, the upper bound is $\varphi_{i,j}^U$, so,

$\forall i \in [1, N], j \in [1, c^i], t \in \mathbb{N}^+$,we have

$$Dec(i,j,t) = \begin{cases} 1 & d_i^j(t) \in [\varphi_{i,j}^L, \varphi_{i,j}^U] \\ 0 & Otherwise \end{cases} \tag{2}$$

In the underground construction environment, different data within the normal range also have different meanings. For example, there are two temperature data 15 and 23, assuming they are both normal temperature values, but it is clear that these are two different working environments for the workers. Therefore, defining a data anomaly as yes or no can not reflect the actual situation. In order to improve the accuracy of data anomaly detection, this paper uses fuzzy theory to analyze the anomaly detection results. For the j-th sensor of the i-th base station whose anomaly detection result is between 0 and 1 at time t, this value represents the degree of data anomaly, where 1 represents the lowest degree of anomaly and 0 represents the highest degree of anomaly, we redefine Eq. (2) as:

$\forall i \in [1, N], j \in [1, c^i], t \in \mathbb{N}^+$,we have

$$Dec(i,j,t) = 1 - |2^{\frac{2d_i^j(t) - \varphi_{i,j}^L - \varphi_{i,j}^U}{\varphi_{i,j}^L - \varphi_{i,j}^U}} - 1| \tag{3}$$

Equation (3) may be represented by the curve of Fig.3, in the lower bound and upper bound intervals, the value of anomaly detection is between [0.5, 1], and when it exceeds upper bond or falls below the lower bound, it is less than 0.5 and approaches zero indefinitely (note that no value of anomaly detection result is defined as 0 in this paper, in order to avoid anomaly detection errors due to individual data as much as possible).

In some special cases, this detection result according to Eq. (3) is effective for single source data. However, because a certain data detection result can not reflect the overall data anomaly, this detection method can not effectively evaluate the overall data anomaly results. In order to improve the accuracy of single-source data anomaly detection and avoid the deviation caused by the detection results of single data, it is usually necessary to analyze the data results at adjacent times. As shown in Fig.4, suppose at time t, the result of the data anomaly detection need to analyze the data values at time $t - w + 1, \ldots, t - 2, t - 1, t$, respectively. In Fig.4, the red dot and the green matrix represent the data at time $t - w + 1$ to t of the two sensors A and B, respectively. According to Eq. (3), the values of the sensors A and B are abnormal at time t. However, according to the data change trend from time $t - w + 1$ to t, the data value of the sensor A is tending to be normal, and therefore, it is not possible to make an abnormal judgment at time t. Similarly, for sensor B, the data values show a linear increasing trend in the same time period. Although the value anomaly offset of sensor B at time t is smaller than that of sensor A, the data of node B is anomalous.

Obviously,the larger the value of w is, the more accurate the change trend of sensor outliers will be. However, too large a

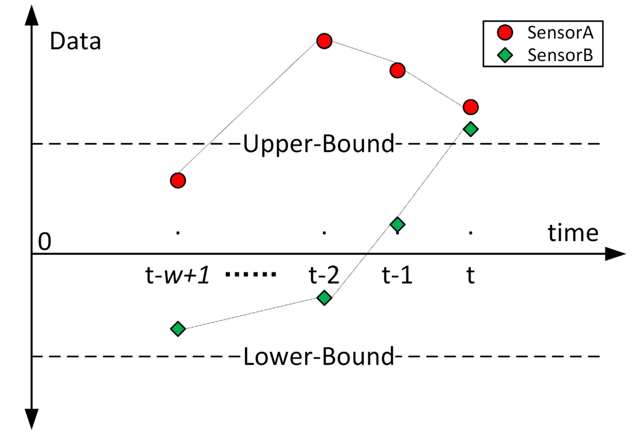


FIGURE 4. Variation of abnormal data at multiple times.

value of w will increase the processing delay of data anomaly detection, so it is necessary to set a reasonable value of w under the condition of delay constraints. Assuming that the data processing speed is a constant σ, the delay of the data processing is ξ. Then, for the j-th sensor of the i-th base-station, the processing delay at time t is $d_i^j(t)/\sigma$, and for the preceding w_i^j times including time t, the total processing delay satisfies:

$\forall i \in [1, N], j \in [1, c^i], t \in \mathbb{N}^+, w_i^j \in \mathbb{N}^+ \wedge w_i^j < t$,we have

$$w_i^j * \min_{x=1}^{t} \frac{d_i^j(x)}{\sigma} \leqslant \xi$$
$$\Rightarrow w_i^j * \min_{x=1}^{t} \frac{d_i^j(x)}{\sigma} \leqslant \xi$$
$$\Rightarrow w_i^j \leqslant \frac{\xi * \sigma}{\min_{x=1}^{t} d_i^j(x)} \tag{4}$$

Since the base-station uses the buffer queue to store the sensing data, the storage amount of the sensing data cannot exceed the length of the buffer queue. Assuming that for the base-station i, and the function $Cnt(x)$ represents the number of elements of the set x. So the queue length of the buffer queueQ_i is $Cnt(Q_i)$. For the j-th sensor of the base-station i, it can store $\lfloor Cnt(Q_i)/p_i^j \rfloor$ data simultaneously in the buffer queue, we have

$$w_i^j \leqslant \left\lfloor \frac{Cnt(Q_i)}{p_i^j} \right\rfloor \tag{5}$$

According to equation (4) and (5), we have

$$w_i^j = \left\lfloor \min(\frac{Cnt(Q_i)}{p_i^j}, \frac{\xi * \sigma}{\min_{x=1}^{t} d_i^j(x)}) \right\rfloor \tag{6}$$

According to equation (6), in this paper, w datas before t time are analyzed when abnormal data are detected. For the j-th sensor of the i-th base-station, some data may be useless in w_i^j candidate data at t time. In this paper, the number of valid data is defined as follows:

Definition 1 (Number of Valid Data): The value $\eta_i^j(t)$ is the number of valid data of the $j(j \in [1, c^i])$-th sensor in base-station $i(i \in [1, N])$ at time $t(t \in \mathbb{N})$,iff the three following conditions both hold:

Condition 1: The number of valid data cannot exceed the number of candidate data.

$$1 \leqslant \eta_i^j(t) \leqslant w_i^j \tag{7}$$

Condition 2: Valid data is incremented over time, that is $\forall k \in [t - \eta_i^j(t) + 2, t]$, we have

$$Dec(i, j, k) \geqslant Dec(i, j, k - 1) \tag{8}$$

Condition 3: $\eta_i^j(t)$ is the maximum number of valid data.

$$Dec(i, j, t - \eta_i^j(t)) > Dec(i, j, t - \eta_i^j(t) + 1)) \tag{9}$$

For the j-th sensor of the i-th base-station, if the larger the value of $eta_i^j(t)$, the change of the data value of the sensor tends to be normal. Therefore, the ratio of $eta_i^j(t)$ to w_i^j is used to represent the abnormal change of the single-source data in this paper. According to the value of $\eta_i^j(t)/w_i^j$ and formula (3), we get the calculation method of single source data anomaly detection based on time variation, which is expressed by $DecT(i, j, t)$.

$\forall i \in [1, N], j \in [1, c^i], t \in \mathbb{N}^+$, we have

$$DecT(i, j, t) = \begin{cases} 1 & \mathit{Condition} \\ 0 & \mathit{Otherwise} \end{cases} \tag{10}$$

The *Condition* in Eq. (10) is described as:

$\forall i \in [1, N], j \in [1, c^i], t \in \mathbb{N}^+$, the *Condition* in Eq. (10) is hold iff one of the following two conditions hold:

C1: $0.5 \leqslant \mathrm{Dec}(i, j, t) \leqslant 1.$

C2: $\eta_i^j(t)/w_i^j \geqslant \zeta.$

ζ is the valid detection coefficient, set according to the specific requirements of underground construction.

B. FORMULATION OF MULTI-SOURCE DATA ANOMALY DETECTION

The goal of multi-source data anomaly detection is to analyze multi-sensor data of the same sensor's type. Unlike single-source data anomaly detection, this approach does not rely on the data of a single sensor, but analyzes the data of multiple sensors at different locations. Thus, the spatial dimension is added to the temporal-dimension of equation (10). Multi-source data anomaly detection needs to determine the location of multiple sensors, and get the node set that is closer to a sensor according to its location. This paper defines the distance relationship between different sensors by location correlation, which is used to decide the candidate anomaly detection queue, and comprehensively analyzes the anomaly data value at a certain time.

A method of node correlation calculation based on plane coordinate position is proposed in [33], but the calculation of the method is complicated, and the underground construction environment is in the tunnel, which the space is narrow, so the plane two-dimensional coordinate can be simplified to one-dimensional. As shown in the topology of Fig.1, the network belongs to a hybrid network topology, which consists of a chain structure between base-stations and a star network between base-stations and sensors. Therefore, in the tunnel, the position of the sensor is determined only by the distance from the laneway entrance. It is assumed that the distance between any two adjacent sensors is the same, and the base-stations and sensors are numbered according to their location. The location correlation of sensors is defined as follows:

Definition 2 (Location Correlation of Sensors): The location correlation between the *j*-th sensor of the *i*-th base-station and the *n*-th sensor of the *m*-th base-station is denoted by $Nbr_i^j(m, n)$, we have:
$\forall i, m \in [1, N], j \in [1, c^i], n \in [1, c^m]$

$$Nbr_i^j(m,n) = \chi - (\sum_{k=i}^{m} c^k - j - c^m + n) \qquad (11)$$

$\chi(\chi \in \mathbb{N}^+ \wedge \chi >= 2)$ is the correlation coefficient, set according to the specific requirements of underground construction.

If the correlation value is greater than 0, it indicates that there is correlation; if it is less than 0, it indicates that there is no correlation. Therefore, this paper selects data with correlation greater than 0 as candidate data for comprehensive analysis, and uses $H(i, j)$ as the candidate correlation node set of the *j*-th sensor of the *i*-th base-station, we have
$\forall m \in [1, N], n \in [1, c^m]$

$$H(i,j) = \{ \text{(m,n)}| i \leqslant m \wedge Nbr_i^j(m,n) > 0\} \qquad (12)$$

The method of multi-source data anomaly detection in this paper is based on the Eq. (11) to analyze the anomaly value of the node set whose position correlation is greater than 0 at time t. The multi-source data anomaly detection result of the *j*-th sensor of the *i*-th base-station at time t is denoted by $DecTL(i, j, t)$, and we have
$\forall i \in [1, N], j \in [1, c^i], t \in \mathbb{N}^+$

$$DecTL(i,j,t) = \begin{cases} 1 & Condition \\ 0 & Otherwise \end{cases} \qquad (13)$$

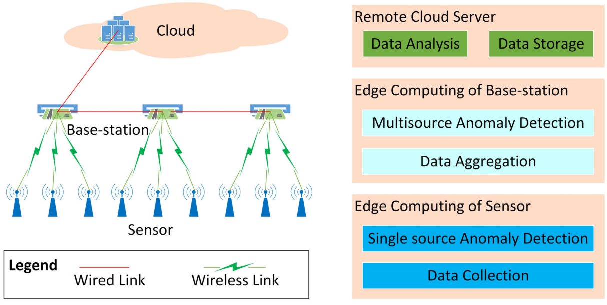
The *Condition* in Eq. (13) is described as:

$\forall i \in [1, N], j \in [1, c^i], t \in \mathbb{N}^+$, the *Condition* in Eq. (13) is hold iff one of the following two conditions hold:

C1: $$\frac{\sum_{(m,n)\in H(i,j)} DecT(m,n,t)}{Cnt(H(i,j))} \geqslant \psi.$$

C2: $$DecT(i,j,t) = 1.$$

ψ is the valid detection coefficient, set according to the specific requirements of underground construction.



(a) Physical structure (b) Logical structure

FIGURE 5. Hierarchical edge computing model.

IV. DESIGN OF ANOMALY DETECTION SCHEME

According to the special topology in IIoT for underground the mining, the data anomaly detection operation is distributed in different node units for processing. Taking into account the requirements of accuracy and efficiency in multi-source multi-dimensional data anomaly detection, the scheme consists of three parts: hierarchical edge computing model, single-source data anomaly detection algorithm and multi-source data anomaly detection algorithm. The traditional data anomaly detection algorithm usually performs abnormal decision in the remote cloud, and uses the big data storage and analysis capability on the cloud platform to realize intelligent data anomaly detection and analysis. However, the underground mining operation is complex and variable, and so abnormal events may occur at any time. This requires data anomaly detection to meet both accuracy and real-time requirements. In many mine accidents, due to the lack of timely prediction of environmental anomalies, it is impossible to effectively arrange evacuation and disaster relief when the disaster arrives. To this end, this paper proposes to use edge computing to transfer the cloud's anomaly decision to the edge side of the base station and sensor.

A. HIERARCHICAL EDGE COMPUTING MODEL

Fig. 5 shows the overall architecture of the hierarchical edge computing model. According to the functional division of anomaly detection, the IIoT system for mining operation safety monitoring and early warning mainly consists of three parts: remote cloud server, base station (aggregation node) and sensor. Fig. 5(a) shows the physical architecture of the early warning system, where the remote cloud is responsible for storing and analyzing the data uploaded by the base station. The base station is responsible for aggregating the data collected by the sensor and forwarding it through a wired link. The sensor is responsible for periodically collecting data such as temperature, humidity, gas concentration, etc., and transmitting it to the corresponding base station via wireless medium. Fig. 5(b) shows the logical model of the early warning system. The model distributes the edge computing in two units: the base station edge and the sensor edge. The hardware devices of the base station are superior to the sensors in terms of processing power and storage capacity. In addition,

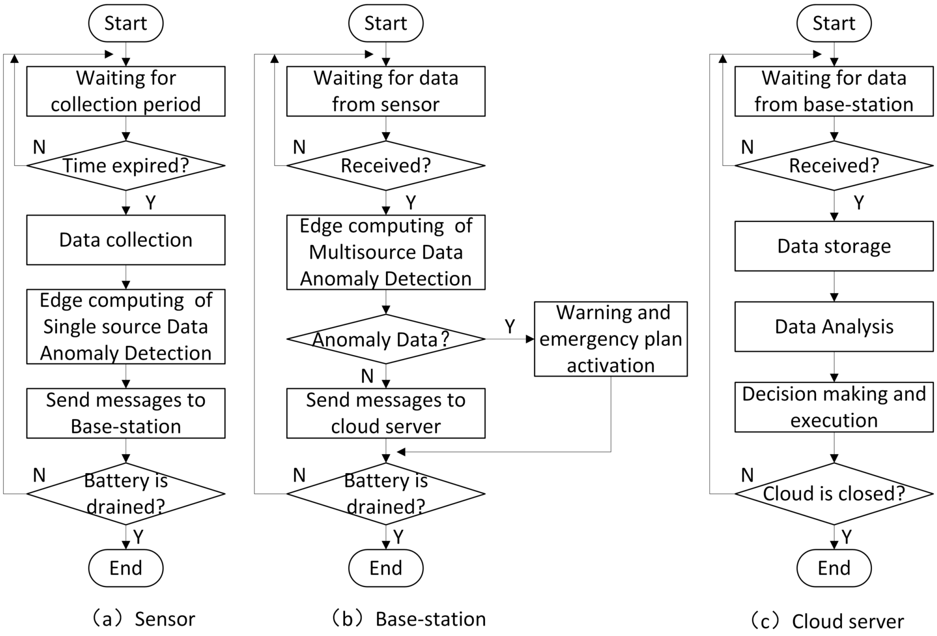

Start
Waiting for collection period
N
Time expired?
Y
Data collection
Edge computing of Single source Data Anomaly Detection
Send messages to Base-station
N
Battery is drained?
Y
End
（a） Sensor
Start
Waiting for data from sensor
N
Received?
Y
Edge computing of Multisource Data Anomaly Detection
Anomaly Data？
Y
Warning and emergency plan activation
N
Send messages to cloud server
N
Battery is drained?
Y
End
（b） Base-station
Start
Waiting for data from base-station
N
Received?
Y
Data storage
Data Analysis
Decision making and execution
N
Cloud is closed?
Y
End
（c） Cloud server

FIGURE 6. Flow chart of data anomaly detection in underground mining.

the base station equipment is usually powered by a wired power source, and the backup battery capacity is also large. Therefore, the task of edge computing in the base station is mainly to perform the execution of the multi-source data anomaly detection algorithm, and the task of edge computing in the sensor node is to perform the execution of the single source data anomaly detection algorithm (only the abnormal data detection at a single moment).

As shown in Fig. 6, the edge computing process of the early warning system is summarized as follows:

1) The sensor program periodically collects environmental state data according to requirements. After collecting the data, it performs single source data anomaly detection according to Eq. (3). When the detection is completed, the original data and detection results are sent to the base station. The sensor's processor generally only has simple information processing and wireless transmission functions, so only simple data anomaly detection can be performed.
2) The base station program waits to receive data transmitted by the sensor side. After receiving the data, the base station performs multi-source heterogeneous data detection, and combines the received single-source data anomaly detection result with other detection results generated at sensors that are correlated in time and space, and performs comprehensive analysis. The final anomaly detection result is obtained according to Eq. (10) and Eq. (13), and it is sent to the cloud with the original data. A base station device generally has a relatively powerful processor, such as an MSP430 and an ARM. Therefore, a multi-source abnormality detection program is deployed at a base station for execution. Moreover, when the data detection triggers an abnormal event, the system will start an emergency warning and treatment plan according to the safety prevention and early warning level in underground mining.
3) The cloud platform side waits to receive data sent by the base station. When the data is received, it is stored in the database of the cloud platform. Then, the decision center uses data mining, artificial intelligence and other algorithms to analyze and make decisions on the original data, and implement the corresponding decision processing scheme.

B. SINGLE-SOURCE DATA ANOMALY DETECTION ALGORITHM

Single-source data anomaly detection algorithm uses Eq. (3) and (10) to detect anomalies in one sensor data. For the j-th sensor of the i-th base-station, the algorithm for detecting anomaly data at time t is divided into two parts: Firstly, according to Eq. (6), the anomaly value of single-source data at time t is calculated, and the algorithm of this part is execute at the sensor. Secondly, according to the Eq. (10), the outliers of the single-source data at several times before t are synthetically analyzed, and the algorithm performs edge computing at the base-station. For the j-th sensor of the i-th base-station at time t, the process of single-source data anomaly detection algorithm is shown as follows:

1) At the sensor side, according to Eq. (3), the abnormal data value $dec(i, j, t)$ at time t is calculated, and the result is transmitted to the base-station together with the original data and stored in the buffer queue Q of base-station.

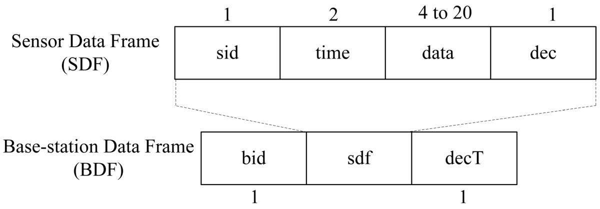


FIGURE 7. Data structure.

2) At the base-station, according to the Eq. (6), the number w_i^j of candidate data sets for data anomaly detection is calculated, the data at time $t - w_i^j + 1, \ldots, t-2, t-1, t$ are sequentially traversed, and the number $\eta_i^j(t)$ of valid data is determined according to the definition 1.
3) Return the anomaly detection result according to Eq. (10).

The base-station stores the data transmitted by the sensor through a buffer queue, and for the base station i, the storage queue Q_i stores up to $Cnt(Q_i)$ data. The sensor data includes the sensor number, time, raw data, and simple anomaly detection results. The base-station stores this information in a buffer queue, including the base-station number, data information from sensor and the final result of the single-source data anomaly detection. In this paper, the storage structure of the data is represented by Sensor Data Frame (SDF) and Base-station Data Frame (BDF), respectively, as shown in Fig.7.

Single-source data anomaly detection algorithm is divided into two parts: sensor's Fuzzy Theory Anomaly Detection Method (FTADM) and base-station's Single-source Anomaly Detection Method (SDADM).

FTADM algorithm is mainly based on simple anomaly detection of the data collected, according to equation (3) to determine the preliminary detection results, the algorithm is described in Fig. 8.

Sensor sends SDF data to base-station, and the base-station stores the SDF data in the local cache queue, traverses the SDF data in the buffer queue in turn, and performs a single-source anomaly detection algorithm for multiple times. The SDADM algorithm is executed in the base-station and ultimately encapsulates the detection results and SDF data into BDF data, as shown in Fig. 9.

C. MULTI-SOURCE DATA ANOMALY DETECTION ALGORITHM

After executing the two-part single-source data anomaly detection algorithm of sensor and base-station, we can get the single-source data anomaly result at a certain time. According to section.III, this anomaly detection result does not take into account the anomaly detection result of the nearer sensor node, so the results of Fig. 8 and Fig. 9 may be incorrect. For example, the data value collected by a sensor is abnormal because of equipment failure, but this kind of anomaly is not real data, so we need to consider the data anomaly value of multi similar nodes.

```
FTADM($J, d[\,], T, Lb, Ub, x$) {
//   $J$ is the number of sensor.
//   $d[\,]$ is an array of data collected by sensor,$d[i](i =$
     $1, 2, \ldots)$ represents data at time $i$.
//   $T$ is the length of the superframe in which the sensor
     collects data, that is, the length of array $d[\,]$.
//   $Ub$ and $Lb$ are the upper bound and lower bound of
     the normal data value, respectively.
//   $x(x > 0)$ is the invalid parameter of data
     SDF $s[T]$; // array of SDF
     for (int $i = 0; i < T; i{+}{+}$){ //start of $T$-loop
         if ($(d[t] > (Ub + x)$ or $d[t] < (Lb - x)${
             //Invalid Data
             $s[i].dec = 0;$
         } else {
             //according to Eq. (3)
             $s[i].dec = 1 - abs($
                 $(2 * d[t] - Lb - Ub)/(Lb - Ub) - 1);$
         } // end of if
         $s[i].data = d[t];$
         $s[i].time = i;$
     } //end of $T$-loop
     $s[i].id = J;$
     return $s[\,]$;
}
```

FIGURE 8. Pseudocode for sensor's fuzzy theory anomaly detection algorithm.

The single source data anomaly detection results can not guarantee the detection results must be correct. The following describes how to perform multi-source data anomaly detection on different sensors of the same type.

After the base-station completes the single-source data anomaly detection, it will produce BDF data, which contains the raw data and single-source anomaly detection at a certain time. The Multi-source Data Anomaly Detection Method (MDADM) is described in Fig.10.

V. EXPERIMENTAL VERIFICATION AND RESULT ANALYSIS

In order to evaluate the effectiveness of the anomaly detection method proposed in this paper, we built a verification platform and conducted a large number of experiments. The experimental hardware platform uses TI's MSP430 and CC2530 for base station and sensor programming. The algorithm program is written in C language, and the data results are analyzed by Python. To verify the performance of detection algorithms under different experimental conditions, the accuracy and delay indicators under different data scales were measured. There are 10 sensors under each base station, and there are a total of 60 sensors. There are three types of sensors that collect temperature, wind speed, and gas parameters. The sampling period of each type of sensor

```
BFTADM($I, s[\,], L, p[\,], Di, Ri, Fi$){
//  $I$ is the number of base-station.
//  $s[\,]$ is an array of base-station's buffer queue, $s[i](i =$
    $1, 2, ...)$ represents $i$-th SDF data in queue.
//  $L$ is the length of the buffer queue, that is, the length of
    array $s[\,]$.
//  $p[\,]$ is an array of sensor's data acquisition period, $s[j]$
    $(j \in [1, c^I])$ represents the peroid of $j$-th sensor of
    base-station $I$
//  $Di$ is the data proces delay constraint, and $Ri$ is the
    data processing rate
//  $Fi$ is the valid detection coeffcient in Eq. (10).
    BDF $b[L] = 0$; // array of BDF
    for (int $i = 0; i < L; i++$){ //start of $L$-loop
        if ($s[i].dec >= 0.5$){ //claimed normal data
            $b[i].decT = 1$;
            continue;
        } // end of if
        int $k = s[i].data$;
        for (int $j = i; j >= 0; j--$){ //start of $i$-loop
            if ($s[j].sid == s[i].sid$ and $k > s[j].data$)
                $k = s[j].data$;
        } // end of $i$-loop
        int $w =$min$(L/p[s[i].sid], Di * Ri/k)$;
        int $m = i, count = 0$;
        for (int $j = m; j >= i - w + 1; j--$){//$m$-loop
            if ($s[j].sid == s[m].sid$){
                if ($s[j].dec <= s[m].dec$){
                    $m = j$;
                    $count++$;
                }else
                    break;
            } // end of if
        } // end of $m$-loop
        if ($count/w >= Fi$)
            $b[i].decT = 1$;
    } //end of $L$-loop
    return $b[\,]$;
}
```

FIGURE 9. Pseudocode for base-station single-source anomaly detection algorithm.

```
MDADM($I, b[\,], L, bm[\,][\,][\,], N, c[\,], X, Pi$){
//  $I$ is the number of base-station.
//  $b[\,]$ is an array of BDF data, $d[i](i = 1, 2, ...)$
    represents $i$-th BDF data .
//  $L$ is the length of array $d[\,]$.
//  $bm[\,][\,][\,]$ is an array of the result of single-source data
    anomaly detection, and $bm[i][j][t]$ represents the
    detection result of the $j$-th sensor of the $i$-th base-station
    at time $t$, which is normally 1 and the anomaly is 0.
//  $N$ is the total number of base-stations.
//  $c[\,]$ is an array of number of sensors in one base-station,
    and $c[i]$ represents the sensors' number in base-station $i$.
//  $X$ is the correlation coefficient in definition 2.
//  $Pi$ is the valid detection coefficient in Eq. (13).
    int $dt[M][L] = 0$; // array of result of MDADM, $M = c[I]$
    for(int $i = 0; i < L; i++$){ //$L$-loop
        int $j = b[i].sdf.sid$;
        int $t = b[i].sdf.time$;
        if ($b[i].decT == 1$) { // normal data
            $dt[j][t] = 1$;
            continue;
        } // end of if
        int $h, k = 0$;
        for (int $m = 0; m < I; m++$)} //$I$-loop
            for (int $n = 0; n < c[m]; n++$)} // $c[m]$-loop
                int $s =$sum$(c[n], c[m])$;
                int $d = X-$abs$(s - j - c[m] + n)$; // according
                    to definition 2
                if ($d > 0$){
                    $h++$;
                    if (($bm[m][n][t] == 1$))
                        $k++$;
                } // end of if
            } // end of $c[m]$-loop
        } // end of $I$-loop
        if ($k/h >= Pi$)
            $dt[j][t] = 1$;
    } // end of $L$-loop
    return $dt[\,][\,]$;
}
```

FIGURE 10. Pseudocode for multi-source data anomaly detection algorithm.

is different (1000, 2000, 3000ms), and each sensor sends 100 data periodically. The test data set randomly generates abnormal data by manual setting, and the overall data value obeys the normal distribution $N(\mu, \sigma^2)$, among which $\mu = (Ub - Lb)/2$ and $\sigma = \sqrt{(Ub + Lb)/2}$. In Eq. (10), (11) and (13), we set $\zeta = 0.5$, $\chi = 5$ and $\psi = 0.4$, respectively. The experimental platform environment and parameter settings are shown in Table. 1 below.

Different from simulation, the experiments mainly explore how to perform edge computing on the actual embedded system, so some theoretical parameters will be simplified according to the situation. For example, it is assumed that the CPU processing rate is constant 1 in Eq.(6). This setting is also in line with the actual situation. Although the processing speeds of different CPUs are different, the impact on test results can be ignored in actual experiments.

This paper mainly examines three performance indicators of the detection algorithm: detection accuracy, algorithm execution time and average delay. The detection accuracy is defined as the ratio of the number of results of data anomaly detection to the total number of anomalies. This indicator can reflect the execution efficiency of the detection algorithm. The algorithm execution time is defined as the total time required for data anomaly detection, which is used to measure the time complexity of algorithm execution. The average delay is defined as the average of the processing time of all data in the anomaly detection, which is used to measure the time sensitivity of the detection scheme. The experiments explored the performance of four anomaly detection schemes: Traditional Anomaly Detection Method(TADM) (Eq. (2)),

TABLE 1. Experiment parameters.

		Base-station	Sensor
Hardware	**CPU**	MSP430	CC2530(8051CPU)
	Wired link	RS485	No
	Wireless link	2.4-GHz IEEE 802.15.4	2.4-GHz IEEE 802.15.4
Programming language	C and Python		
The number of base-station	1,2,3,4,5,6		
Type of sensors	The ratio Temperature, wind speed and gas, which is 3: 3: 4		
The number of sensors	Each base-station contains 10 sensors		
Sensor data transmission period	1,2 and 3s for three types		
Upper-bound of normal data (Ub)	40,80 and 20 for three types		
Lower-bound of normal data (Lb)	-20,10 and 10 for three types		
Data size	Total 100 datas, each with a size of 10 Bytes		
Experiment number	20		

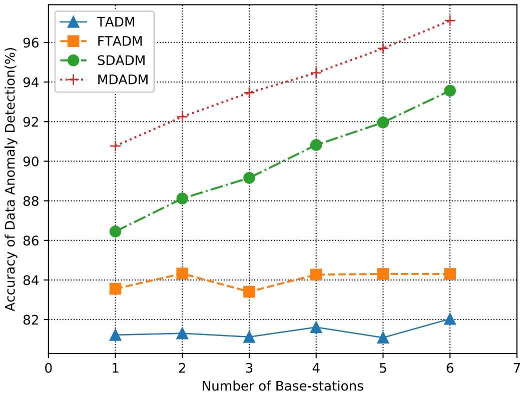


FIGURE 11. Accuracy analysis of different methods in this paper.

Fuzzy Theory Anomaly Detection Method (FTADM) (Eq. 3), Single-source Data Anomaly Detection Method (SDADM) (Eq. 10) and Multi-source Data Anomaly Detection Method (MDADM) (Eq. 13). In the case of the same data set, the experiment compares and analyzes the differences in detection efficiency and cost between different schemes.

Fig. 11 describes the accuracy comparison between the four detection schemes as the data amount increases. It can be seen from Fig. 11 that the MDADM has the highest detection accuracy while the TADM has the lowest. This shows that the proposed detection method can effectively analyze the time and space changes of the data and obtain reasonable results. In addition, we can see that the detection accuracy of both the SDADM and MDADM methods shows an approximately linear increase as the data amount increases, while the detection accuracy of both the TADM and FTADM methods does not change significantly as the data amount grows. This is because both the TADM and FTADM methods rely primarily on single-point data values, and changes in data scale do not affect their detection accuracy. When the data scale is large, the probability of detection failure of these two methods (especially TADM) will be greater. Therefore, the proposed detection algorithms have higher detection accuracy compared with TADM.

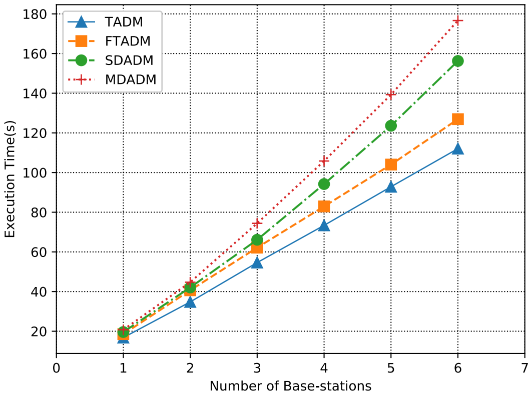


FIGURE 12. Accuracy analysis of different methods in this paper.

Fig. 12 shows the execution time comparison between the four detection schemes as the data amount increases. As can be seen from Fig. 12, the execution time of the MDADM method is greater than that of other methods. In particular, this trend is more pronounced as data scale continues to increase. This is because the MDADM method analyzes multiple data for continuous time and related location nodes, so execution time is more than other methods. In addition, it can be seen that when the data scale is not very large, the gap between different algorithms is small. This increase cost in execution time is worthwhile compared to the improvement in detection accuracy.

The MDADM method takes more time in total execution time. But if we calculate the average processing delay of the detection scheme, the gap will be small. Fig. 13 shows the average delays comparison between the four detection schemes as the data amount increases. As can be seen from Fig. 13, the average data processing delay does not increase linearly as the data scale increases. Moreover, the difference between the average delays of the various detection algorithms differs by a few milliseconds. For underground mining environment, the data processing delay of milliseconds is completely acceptable.

Among the current data anomaly detection methods, cluster-based methods are widely used, and the main methods

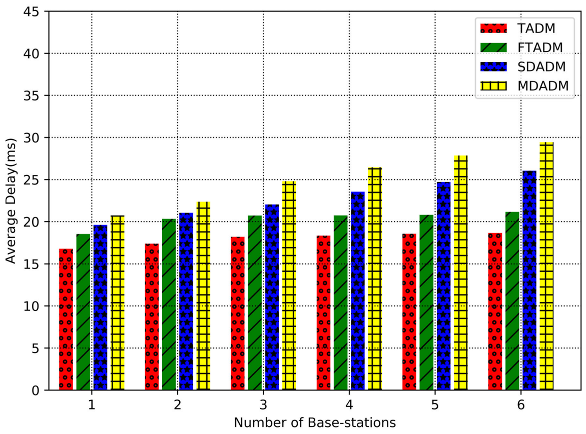


FIGURE 13. Average delay analysis of different methods in this paper.

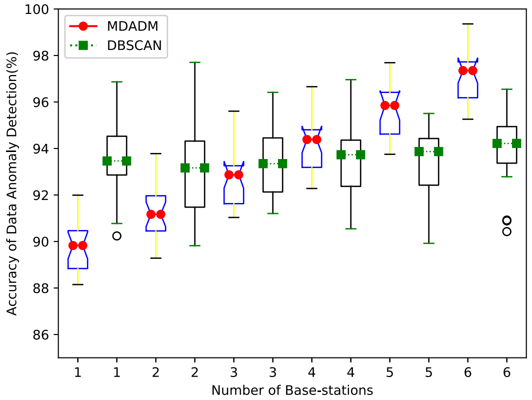


FIGURE 14. Accuracy analysis comparison with existing methods.

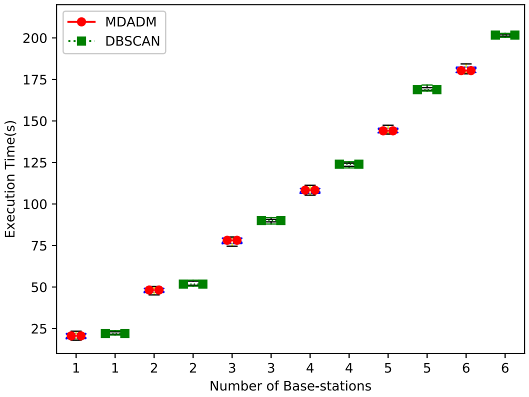


FIGURE 15. Execution time analysis comparison with existing methods.

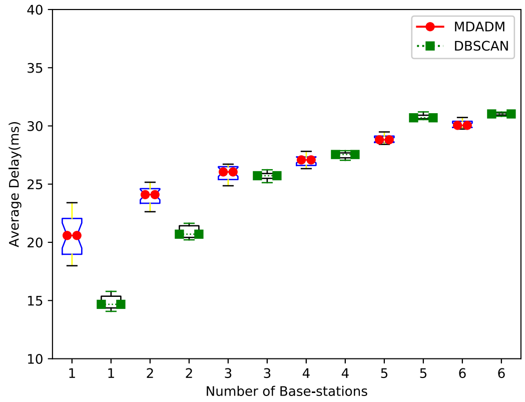


FIGURE 16. Average delay analysis comparison with existing methods.

are DBSCAN [49], k-means [50], and so on. The cluster-based method can realize the intelligent analysis of large-scale data sets through the artificial intelligence theory such as machine learning, and can obtain high anomaly detection accuracy. Among the many methods, DBSCAN is one of the most studied in the academic world. Therefore, this paper chooses the abnormal detection method based on DBSCAN for comparative analysis. In [40], an anomaly detection algorithm based on DSCAN is proposed. Here, we compare the performance of MDADM algorithm with this DBSCAN algorithm in terms of accuracy, execution time and average delay. We performed statistics on 20 experiments, and used box-plots to represent comparative analysis results.

Fig. 14 shows the detection accuracy comparison between the MDADM algorithm and the DBSCAN algorithm as the data amount increases. It can be seen from Figure 14 that with the increase of data scale, the MDADM algorithm has little difference in detection accuracy from DBSCAN. When the data scale is small, the detection accuracy of DBCAN is greater than that of MDADM. However, as the data scale increases, the difference in detection accuracy between the two algorithms becomes smaller and smaller. Moreover, in a large number of cases, the MDADM algorithm has higher detection accuracy than the DBSCAN algorithm. According to the data change situation in Fig. 14, the data scale has less influence on the DBSCAN algorithm, and has a greater influence on the MDADM algorithm. Therefore, the MDADM method is very suitable for the anomaly detection of large amounts of data in large-scale WSN for underground mining. According to the experimental analysis, the MDADM algorithm has reached the detection accuracy level of the DBSCAN algorithm. Moreover, when the data scale is large, the MDADM algorithm can obtain higher abnormality detection accuracy.

Fig. 15 shows the execution time comparison between the MDADM algorithm and the DBSCAN algorithm as the data amount increases. It can be seen from Fig. 15 that as the data scale increases, the execution time of the MDADM algorithm and the DBSCAN algorithm both show an approximately linear increase trend. Overall, the DBSCAN algorithm performs more time than the MDADM algorithm. Especially when the data scale is large, the difference in execution time between the two algorithms is obvious. From the experimental results, we can see the MDADM algorithm has less execution time than the DBSCAN algorithm, and the effect of data scale on execution time is linear. This is because the MDADM algorithm uses hierarchical edge calculation, which results in the algorithm execution time being dispersed on the base station and sensor nodes, achieving effective load balancing and reducing node energy consumption.

Fig. 16 shows the average delay comparison between the MDADM algorithm and the DBSCAN algorithm as the data amount increases. As can be seen from Figure 16, as the data scale increases, the average latency of the MDADM algorithm and the DBSCAN algorithm increases. Compared to Fig. 16, this change is gently rising. In particular, the average delay of the MDADM method is smaller than the DBSCAN algorithm when the data scale is large. Therefore, for the abnormal detection of large amounts of data in large-scale WSN in underground mining environments, it is appropriate to adopt the MDADM method.

VI. CONCLUTION

In this paper, a multi-source multi-dimensional data anomaly detection method based on hierarchical edge computing is proposed, which is aiming at early warning of an accident in IIoT for underground mining environment. According to the special hybrid topology of star network and chain network, a hierarchical edge computing model is first proposed, which can perform multi-source data anomaly detection at sensors end and base-stations end, to realize load balance and low-latency data processing. Based on fuzzy theory, A single-source data anomaly detection algorithm is then proposed, which takes into account the temporal correlation of monitoring data. At last, a multi-source data anomaly detection algorithm is designed, which considers the temporal and spatial correlation properties of multi-source data. Extensive experimental verification demonstrates that the proposed scheme performs better in detection accuracy and processing delay than traditional schemes.

REFERENCES

[1] J. Zhang, X. Hu, Z. Ning, E. C.-H. Ngai, L. Zhou, J. Wei, J. Cheng, and B. Hu, "Energy-latency tradeoff for energy-aware offloading in mobile edge computing networks," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2633–2645, Aug. 2018. doi: 10.1109/JIOT.2017.2786343.

[2] X. Wang, Z. Ning, X. Hu, L. Wang, L. Guo, and B. Hu, "Future communications and energy management in Internet of vehicles: Toward intelligent energy-harvesting," *IEEE Wireless Commun.*, to be published. doi: 10.1109/MWC.2019.1900009.

[3] Z. Ning, P. Dong, X. Kong, and F. Xia, "A cooperative partial computation offloading scheme for mobile edge computing enabled Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4804–4814, Jun. 2019.

[4] X. Liu and P. Zhang, "Data drainage: A novel load balancing strategy for wireless sensor networks," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 125–128, Jan. 2018.

[5] T. M. Fernández-Caramés, and P. Fraga-Lamas, "A review on human-centered IoT-connected smart labels for the industry 4.0," *IEEE Access*, vol. 6, pp. 25939–25957, 2018.

[6] U. I. Minhas, I. H. Naqvi, S. Qaisar, K. Ali, S. Shahid, and M. A. Aslam, "A WSN for monitoring and event reporting in underground mine environments," *IEEE Syst. J.*, vol. 12, no. 1, pp. 485–496, Mar. 2018.

[7] Y. Zhu, D. Wang, Z. Shao, C. Xu, X. Zhu, X. Qi, and F. Liu, "A statistical analysis of coalmine fires and explosions in China," *Process Saf. Environ. Protection*, vol. 121, pp. 357–366, Jan. 2019.

[8] M. Thibaud, H. Chi, W. Zhou, and S. Piramuthu, "Internet of Things (IoT) in high-risk environment, health and safety (EHS) industries: A comprehensive review," *Decis. Support Syst.*, vol. 108, pp. 79–95, Apr. 2018.

[9] S. Huang, B. Guo, W. Ju, X. Zhang, J. Han, C. Phillips, J. Zhang, and W. Gu, "A novel framework and the application mechanism with cooperation of control and management in multi-domain WSON," *J. Netw. Syst. Manage.*, vol. 21, no. 3, pp. 453–473, 2013.

[10] Z. Ning, J. Huang, X. Wang, J. J. P. C. Rodrigues, and L. Guo, "Mobile edge computing-enabled Internet of vehicles: Toward energy-efficient scheduling," *IEEE Netw.*, to be published. doi: 10.1109/MNET.2019.1800309.

[11] S. Huang, J. Li, Y. Ye, P. Shi, J. Zhou, B. Guo, and W. Gu, "The further investigation of the true time delay unit based on discrete fiber Bragg gratings," *Opt. Laser Technol.*, vol. 44, no. 4, pp. 776–780, 2012.

[12] W. Zhang, W. Liu, T. Wang, A. Liu, Z. Zeng, H. Song, and S. Zhang, "Adaption resizing communication buffer to maximize lifetime and reduce delay for WVSNs," *IEEE Access*, vol. 7, pp. 48266–48287, 2019.

[13] Z. Ning, X. Wang, J. J. Rodrigues, and F. Xia, "Joint computation offloading, power allocation, and channel assignment for 5G-enabled traffic management systems," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 3058–3067, May 2019.

[14] S. Huang, Y. Zhou, S. Yin, Q. Kong, M. Zhang, Y. Zhao, J. Zhang, and W. Gu, "Fragmentation assessment based on-line routing and spectrum allocation for intra-data-center networks with centralized control," *Opt. Switching Netw.*, vol. 14, pp. 274–281, Aug. 2014.

[15] X. Liu, "Node deployment based on extra path creation for wireless sensor networks on mountain roads," *IEEE Commun. Lett.*, vol. 21, no. 11, pp. 2376–2379, Nov. 2017.

[16] X. Wang, Z. Ning, M. C. Zhou, X. Hu, L. Wang, Y. Zhang, F. R. Yu, and B. Hu, "Privacy-preserving content dissemination for vehicular social networks: Challenges and solutions," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1314–1345, 2nd Quart., 2019.

[17] Z. Ning, P. Dong, X. X. Wang, and J. Rodrigues, "Deep reinforcement learning for vehicular edge computing: An intelligent offloading system," *ACM Trans. Intell. Syst. Technol.*, vol. 25, p. 1, May 2019. doi: 10.1145/3317572.

[18] X. R. Wang, J. T. Lizier, O. Obst, M. Prokopenko, and P. Wang, "Spatiotemporal anomaly detection in gas monitoring sensor networks," in *Proc. Eur. Conf. Wireless Sensor Netw.* Berlin, Germany: Springer, 2008, pp. 90–105.

[19] O. Obst, X. R. Wang, and M. Prokopenko, "Using echo state networks for anomaly detection in underground coal mines," in *Proc. Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2008, pp. 219–229.

[20] P.-Y. Chen, S. Yang, and J. A. McCann, "Distributed real-time anomaly detection in networked industrial sensing systems," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3832–3842, Jun. 2015.

[21] H. Soydan, A. Koz, and H. S. Düzgün, "Spatio-temporal anomaly detection for environmental impact assessment: A case of an abandoned coal mine site in Turkey," *Proc. SPIE*, vol. 10405, Sep. 2017, Art. no. 104050B.

[22] A. Tan, Q. Wang, N. Guan, Q. Deng, and X. S. Hu, "Inter-cell channel time-slot scheduling for multichannel multiradio cellular fieldbuses," in *Proc. IEEE Real-Time Syst. Symp.*, Dec. 2015, pp. 227–238.

[23] Y. He, J. Guo, and X. Zheng, "From surveillance to digital twin: Challenges and recent advances of signal processing for industrial Internet of Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 120–129, Sep. 2018.

[24] S. Verma, Y. Kawamoto, Z. M. Fadlullah, H. Nishiyama, and N. Kato, "A survey on network methodologies for real-time analytics of massive IoT data and open research issues," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1457–1477, 3rd Quart., 2017.

[25] S. Huang, J. Li, J. Zhou, and W. Gu, "Novel spectrum properties of the periodic π-phase-shifted fiber Bragg grating," *Opt. Commun.*, vol. 285, no. 6, pp. 1113–1117, 2012. doi: 10.1016/j.optcom.2011.10.052.

[26] A.-R. Sadeghi, C. Wachsmann, and M. Waidner, "Security and privacy challenges in industrial Internet of Things," in *Proc. 52nd ACM/EDAC/IEEE Design Automat. Conf. (DAC)*, Jun. 2015, pp. 1–6.

[27] S. Huang, B. Guo, X. Li, J. Zhang, Y. Zhao, and W. Gu, "Pre-configured polyhedron based protection against multi-link failures in optical mesh networks," *Opt. Express*, vol. 22, no. 3, pp. 2386–2402, 2014.

[28] P. Oluwasanya, "Anomaly detection in wireless sensor networks," Univ. Edinburgh, Scotland, U.K., Tech. Rep. PGEE11110, 2017.

[29] M. A. Rassam, A. Zainal, and M. A. Maarof, "One-class principal component classifier for anomaly detection in wireless sensor network," in *Proc. 4th Int. Conf. Comput. Aspects Social Netw. (CASoN)*, Nov. 2012, pp. 271–276.

[30] O. Salem, Y. Liu, and A. Mehaoua, "Anomaly detection in medical WSNs using enclosing ellipse and chi-square distance," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 3658–3663.

[31] S. Rajasegarar, C. Leckie, and M. Palaniswami, "Anomaly detection in wireless sensor networks," *IEEE Wireless Commun.*, vol. 15, no. 4, pp. 34–40, Aug. 2008.

[32] S. Rajasegarar, C. Leckie, and M. Palaniswami, "Detecting data anomalies in wireless sensor networks," in *Security in Ad Hoc And Sensor Networks*. Singapore: World Scientific, 2010, pp. 231–259.

[33] H. Fei, F. Xiao, G. H. Li, and L. J. Sun, "An anomaly detection method of wireless sensor network based on multi-modals data stream," *Chin. J. Comput.*, vol. 40, no. 8, pp. 1829–1842, 2017.

[34] H. Ren, M. Liu, X. Liao, L. Liang, Z. Ye, and Z. Li, "Anomaly detection in time series based on interval sets," *IEEJ Trans. Elect. Electron. Eng.*, vol. 13, no. 5, pp. 757–762, 2018.

[35] Y. Djenouri, A. Belhadi, J. C.-W. Lin, D. Djenouri, and A. Cano, "A survey on urban traffic anomalies detection algorithms," *IEEE Access*, vol. 7, pp. 12192–12205, 2019.

[36] K. Sricharan, R. Raich, and A. O. Hero, III, "Empirical estimation of entropy functionals with confidence," 2010, *arXiv:1012.4188*. [Online]. Available: https://arxiv.org/abs/1012.4188

[37] H. H. Bosman, G. Iacca, A. Tejada, H. J. Wörtche, and A. Liotta, "Spatial anomaly detection in sensor networks using neighborhood information," *Inf. Fusion*, vol. 33, pp. 41–56, Jan. 2017.

[38] M. Xie, J. Hu, S. Guo, and A. Y. Zomaya, "Distributed segment-based anomaly detection with Kullback–Leibler divergence in wireless sensor networks," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1, pp. 101–110, Jan. 2017.

[39] S. Huang, W. Lian, X. Zhang, B. Guo, P. Luo, J. Zhang, and W. Gu, "A novel method to evaluate clustering algorithms for hierarchical optical networks," *Photonic Netw. Commun.*, vol. 23, no. 2, pp. 183–190, 2012.

[40] H. S. Emadi and S. M. Mazinani, "A novel anomaly detection algorithm using DBSCAN and SVM in wireless sensor networks," *Wireless Pers. Commun.*, vol. 98, no. 2, pp. 2025–2035, 2018.

[41] S. Seo, S. Park, I. Hwang, and J. Kim, "ADSTREAM: Anomaly detection in large-scale data streams using local outlier factor based on micro-cluster," *Adv. Sci. Lett.*, vol. 23, no. 10, pp. 10204–10209, 2017.

[42] K. Sricharan, R. Raich, and A. O. Hero, "K-nearest neighbor estimation of entropies with confidence," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul./Aug. 2011, pp. 1205–1209.

[43] M. A. Rassam, M. A. Maarof, and A. Zainal, "Adaptive and online data anomaly detection for wireless sensor systems," *Knowl.-Based Syst.*, vol. 60, pp. 44–57, Apr. 2014.

[44] B. Ahmad, W. Jian, Z. A. Ali, S. Tanvir, and M. S. A. Khan, "Hybrid anomaly detection by using clustering for wireless sensor network," *Wireless Pers. Commun.*, vol. 106, no. 4, pp. 1841–1853, 2019.

[45] D. Ramotsoela, A. Abu-Mahfouz, and G. Hancke, "A survey of anomaly detection in industrial wireless sensor networks with critical water system infrastructure as a case study," *Sensors*, vol. 18, no. 8, p. 2491, 2018.

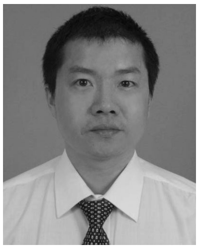
[46] D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim, "A survey of deep learning-based network anomaly detection," *Cluster Comput.*, pp. 1–13, Sep. 2017.

[47] D.-S. Pham, S. Venkatesh, M. Lazarescu, and S. Budhaditya, "Anomaly detection in large-scale data stream networks," *Data Mining Knowl. Discovery*, vol. 28, no. 1, pp. 145–189, 2014.

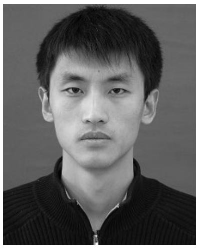
[48] S. Huang, B. Li, B. Guo, J. Zhang, P. Luo, D. Tan, and W. Gu, "Distributed protocol for removal of loop backs with asymmetric digraph using GMPLS in p-cycle based optical networks," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 541–551, Feb. 2011.

[49] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise," 1996, pp. 226–231.

[50] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, Oakland, CA, USA, 1967, pp. 281–297.



YUHUAI PENG received the Ph.D. degree in communication and information systems from Northeastern University, in 2013, where he is currently an Associate Professor. His research interests include the Internet of Things (IoT), industrial communication networks, and health monitoring.



AIPING TAN received the M.E. degree in computer science from Northeastern University, Shenyang, China, in 2010, where he is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering. His research interests include industrial wireless sensor networks and the Internet of Things.

JINGJING WU received the Ph.D. degree in communication and information systems from Northeastern University, Shenyang, China, in 2012, where she is currently an Associate Professor with the School of Computer Science and Engineering. Her research interests include survivability, optical networks, and wireless local area networks.

YUANGUO BI received the Ph.D. degree from Northeastern University, Shenyang, China, in 2010, where he joined the School of Computer Science and Engineering, as an Associate Professor, in 2010. His current research interests include fog computing, software-defined networking, QoS routing, multi-hop broadcast, mobility management, and vehicular networks.