An adaptive method based on contextual anomaly detection in Internet of Things through wireless sensor networks

Xiang Yu¹, Hui Lu², Xianfei Yang¹, Ying Chen¹, Haifeng Song¹, Jianhua Li¹ and Wei Shi³

Abstract
With the widespread propagation of Internet of Things through wireless sensor networks, massive amounts of sensor data are being generated at an unprecedented rate, resulting in very large quantities of explicit or implicit information. When analyzing such sensor data, it is of particular importance to detect accurately and efficiently not only individual anomalous behaviors but also anomalous events (i.e. patterns of behaviors). However, most previous work has focused only on detecting anomalies while generally ignoring the correlations between them. Even in approaches that take into account correlations between anomalies, most disregard the fact that the anomaly status of sensor data changes over time. In this article, we propose an unsupervised contextual anomaly detection method in Internet of Things through wireless sensor networks. This method accounts for both a dynamic anomaly status and correlations between anomalies based contextually on their spatial and temporal neighbors. We then demonstrate the effectiveness of the proposed method in an anomaly detection model. The experimental results show that this method can accurately and efficiently detect not only individual anomalies but also anomalous events.

Keywords
Internet of Things, anomaly detection, anomaly event, clustering, grid

Date received: 11 February 2020; accepted: 30 March 2020

Handling Editor: Xiaojiang Du

Introduction
As a widely used technology in information security, anomaly detection is generally used to detect the anomalous behaviors or anomalous data that violate statistical rules. Other than misuse detection, which is considered as a rule-based method, anomaly detection refers to detect not only the behaviors that are different from the current stored normal patterns but also the unknown anomalous behaviors which are newly emerging.¹ With the advent of Internet of Things (IOT) era, people can conveniently obtain a large amount of various data by millions of sensors according to their requirements, especially in the case of IOT through

¹School of Electronics and Information Engineering, Taizhou University, Taizhou, China
²Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou, China
³School of Information Technology, Carleton University, Ottawa, ON, Canada

Corresponding authors:
Hui Lu, Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510006, China.
Email: luhui@gzhu.edu.cn
Xianfei Yang, School of Electronics and Information Engineering, Taizhou University, Taizhou 318000, China.
Email: yangxianfei@dlu.edu.cn
wireless sensor networks (WSNs) with good mobility and flexibility. Currently, more than 200 on-orbit satellites are capturing remote sensor (RS) data from earth observations, which are further used to forecast anomalies, such as climate extremes or geological extremes. And almost all automobile production lines detect anomalous auto parts according to the sensed data sent by the sensors.

However, most traditional anomaly detection methods rely on the patterns, both normal and anomalous, greatly. A behavior might be considered anomalous only if the detected behavior matches an existing known anomalous pattern. Nevertheless, a newly emerging anomalous behavior, which does not match any existing anomalous pattern, might be incorrectly detected as normal. In view of this, some anomaly detection methods based on machine learning are proposed, which can learn from newly emerging behaviors and update the existing anomalous patterns. However, the changes and correlations of the data from the same sensor are always ignored, so does the changes and correlations of the data from different sensors. In the IOT through WSN, due to the dynamic and self-organizing characteristics of WSNs, the difficulty of anomaly detection in such scenario is further increased, thus it is particularly necessary to take the changes of sensed data into account.

In IOT through WSNs, all the deployed sensors constantly send data back to central aggregators or management nodes, as the quantity of data increases, so does the complexity of data analysis and processing. In addition, as the data from the sensors change all the time, it is hard to detect anomalous events (temporal or spatial) accurately just on the basis of stored normal patterns. Considering a possible scenario, as follows: a temperature sensor deployed in a forest area sends back an anomalous temperature value to central aggregator at a certain time, but then it sends back a series of normal temperature values. This anomaly might be caused by the direct sunlight focus landed on the temperature sensor during dewdrop time. When the angle of sunlight changes, the temperature values drop naturally. Obviously, this should not be considered, which consequently triggers a fire alarm. However, if the temperature sensor constantly sends back anomalous temperature values to the central aggregator or its neighbor sensors send anomalous temperature values as well to the central aggregator at the same time period, this is more likely a fire alarm. Hence, it is necessary to propose a heuristic method to better detect both anomalous behaviors and anomalous events accurately and efficiently.

The remainder of this article is organized as follows. In section “Related work,” the previous related work on anomaly detection are introduced. In section “AEDTS model,” together with the anomaly detection method, the AEDTS (Anomalous Events Detection based on Temporal dimension and Spatial dimension) model is presented. In section “Experiments,” the experiments conducted to evaluate the performance of AEDTS on both real dataset and synthetic dataset are discussed. In section “Conclusion and future work,” we further summarize the article and discuss the directions of our future work.

Related work

Traditional anomaly detection methods

Anomaly detection is generally considered as the problem of finding the patterns in data which do not conform to priori expected behaviors. So far, researchers have done lots of related work on anomaly detection and made some achievements. Meanwhile, some general principles have been derived and some valuable methods have been proposed.

Generally, traditional anomaly detection methods can be roughly classified as three categories, that is, methods based on distance, methods based on density, and methods based on model. As to the methods based on distance, some common distance calculation methods, such as Euclidean distance, are often adopted as the measure for anomaly detection. In most distance-based methods, an object is considered as anomalous when the distance between the object and the center of detection model exceeds a certain threshold. Angiulli proposed an algorithm for mining distance-based outliers in very large datasets whose I/O cost corresponds to the cost of reading the input dataset twice. Similar to most distance-based algorithms, it is greatly influenced by data distribution. Moshtaghi et al. developed a theory underlying three measures of similarity which can be used for anomaly detection. However, the algorithm is constrained to high-dimensional data space. Zhao et al. proposed an anomaly detection method which is based on adaptive fuzzy clustering. However, it suffers from too many false alarms. Shaikh and Kitagawa analyzed the problem of anomaly detection from uncertain data which is based on the computation of top-k distance. Although the complexity of the anomaly detection algorithm based on distance computation is relatively low, when processing real-world datasets, the accuracy of the algorithm cannot be guaranteed. Shaikh and Kitagawa proposed a cell-based approach, aiming at solving the outlier detection on uncertain datasets of Gaussian distribution. But the algorithm is sensitive to parameters and cannot remain stable in anomaly detection.

As to the methods based on density, an object is considered as anomalous if the density of the area where it locates is less than the predefined threshold. In most local anomaly detection cases, density-based
methods perform better than the methods of other category. Liu et al.\textsuperscript{24} proposed an anomaly detection method based on the concept of isolation without employing any distance or density measure. As a result, the method still suffers from too many false alarms. Gupta et al. proposed an anomaly detection method focused on the individual data points which are considered as globally anomalous. However, it ignores the correlation among these data points.

As to the methods based on model, an anomaly detection model is first established, and then the model can be further trained by historical data.\textsuperscript{25–35} From the perspective of methods involved in the model, there exist the anomaly detection models based on statistical methods such as Gaussian distribution, machine learning methods such as artificial neural network, or overlapped methods such as support vector machine (SVM). Schlkopf et al.\textsuperscript{36} proposed a method for anomaly detection which is based on a natural extension of SVM. However, the method relies heavily on prior knowledge. Subsequently, some other anomaly detection methods, which are based on SVM, are put forward. However, the performance of these SVM-based anomaly detection methods suffer from the sensitivity to missing data. Huang et al.\textsuperscript{37} proposed two unsupervised anomaly detection algorithms which can distinguish local anomalies from normal instances. But it neglects the changes of anomalous behaviors. Ma et al.\textsuperscript{38} proposed a framework, which can detect anomalies in uncertain data stream. However, it cannot detect the correlation among anomalous behaviors effectively.

**Anomaly detection in IOT through WSN**

With the widespread application of IOT through WSNs, traditional anomaly detection methods cannot fully meet the requirements of detecting the anomalies in IOT through WSN accurately and efficiently. In view of this, researchers have developed some new methods for anomaly detection in IOT through WSN.

Chen et al.\textsuperscript{39} proposed a fully distributed general anomaly detection (GAD) scheme, which uses graph theory and exploits spatiotemporal correlations of processes for real-time anomaly detection. However, GAD cannot detect anomalous event which consists of a series of related anomalous behaviors and reflects the changes of anomalous behaviors.

Liu et al.\textsuperscript{40} proposed an unsupervised anomaly detection framework, which is capable of detecting anomalous behaviors and events from the perspective of spatial dimension and temporal dimension. Although the contextual neighbors of each sensor are taken into account, the sensor data change and the correlation among them are ignored.

It can be seen that most existing anomaly detection methods focus on individual anomalous behavior (data) or the anomalous event which consists of both anomalous behaviors and their spatial or temporal anomalous neighbors.\textsuperscript{41–43} However, most previous work has overlooked the changes in anomalous behaviors and the correlations among them.

In this article, we present AEDTS, an anomaly detection model, which takes the changes of sensor data into account and is capable of detecting both anomalous behaviors and anomalous events accurately and efficiently. In AEDTS, an anomalous behavior detected will not be considered as anomalous event first, and the changes of the anomalous behavior and its contextual neighbors are further taken into account on both temporal dimension and spatial dimension. If the behaviors (data) sent from the same sensor constantly show anomaly, these temporal sensor data might trigger an anomalous event. If most spatial neighbors of a series of sensors have shown anomalies in succession within a predefined time period, an anomalous event might be triggered.

**AEDTS model**

In this section, the problem formulation of anomalous event detection is presented, and then follows the definitions, analysis, and our solution in detail. The proposed AEDTS consists of the following phases: data preparation phase, initial clustering and grid partition phase, and anomaly detection phase.

**Problem formulation**

In most cases, data extremum attract more attention than normal data such as financial crisis warning, climate extremes, and bank loan. In this section, anomaly and anomalous event are first formally defined, and then the unsupervised density-based clustering algorithm, the grid partition strategy, and the notion of spatial–temporal anomaly are introduced. Based on these definitions, the unsupervised contextual anomaly detection model AEDTS is presented.

In Figures 1 and 2, the temporal anomaly and spatial anomaly, both which are involved in contextual anomaly, are shown, respectively. As shown in Figure 1, the sensor data, which are sent from a certain sensor, change over time, and there exist a series adjacent anomalies and two separate anomalies which are surrounded by a rectangular dotted box and solid circles, respectively. Among the anomalies, the separate and adjacent anomalies are considered as anomalous behaviors and an anomalous event, respectively. As shown in Figure 2, the data from all sensors are first aggregated by DBSCAN (Density-Based Spatial Clustering of Applications with Noise), a density-based clustering
algorithm, and then the outliers, which are separated from the clusters, are identified as anomalies, which are represented by red solid points. In order to clearly reveal the anomalies in grid cells, all normal data are not shown in Figure 2. As a widely cited clustering algorithm in scientific literature, DBSCAN is awarded the Test of Time Award in 2014. When clustering, other than most distance-based algorithm, DBSCAN does not need to specify the number of clusters in advance and it can find the clusters of arbitrary shape. Next, the monitored spatial region, where all sensors are deployed, is partitioned into grid cells, where a grid cell is considered to be anomalous only when the number of anomalies, which are generated from the clustering process and located in the grid cell, exceeds a certain proportion. Then, the anomalous grid cells are aggregated by the grid merging strategy. The generated clusters of anomalous grid cells, marked as circles 1 and 2 in Figure 2, are considered as spatial anomalous events (SAEs) that must trigger an alarm, while each individual sparsely situated grid cells are considered as anomalous behaviors. In addition, if a spatial anomaly event remains for a certain period of time, it is then considered as a temporal anomalous event (TAE) as well.

**Definition 1: Anomalous behavior.** An anomalous behavior occurs when, at a certain point in time, an input from a sensor exceeds a threshold or is significantly different from the inputs of its neighboring sensors.

**Definition 2: Temporal anomalous event.** A temporal anomalous event occurs when, within a given time period, a cluster of inputs from a sensor exceeds a given threshold.

**Definition 3: Spatial anomalous event.** A spatial anomalous event occurs when two anomalous clusters, generated by anomalous grid cells during the clustering process, are spatial adjacent.

**Grid partition and grid cells merging**

In this section, we describe in detail the key steps of grid partition and anomalous grid cell merging in the anomaly detection model. In order to detect the distribution of anomalies, we first partition the region where all sensors are deployed into grid cells, and the statistics of each grid cell can be represented by a 5-tuple. As shown in Figure 3, in each 5-tuple \((X_i, Y_i, N_{Ano}, N_{Sum}, t)\), \(X_i\) and \(Y_i\) denote the order number of the grid cell in \(X\) and \(Y\) directions, respectively; \(N_{Ano}\) and \(N_{Sum}\) denote the number of the sensors which sent anomalous data back and the total number of sensors in the grid cell, respectively; and \(t\) denotes the time stamp of the grid cell.

Suppose the region where all sensors in a WSN are deployed is partitioned into \(m \times n\) grid cells with parameter \(\delta\). In each grid cell, the white solid point and red solid point represent normal data and anomalous data,
respectively. As to grid cell gc_1 in Figure 3, its 5-tuple is represented as (6, 7, 3, 5, t). As to the same grid cell with different time stamp, the N_{Ano} and N_{Sum} might be different.

When detecting SAEs, we first calculate the ratio of anomalous data in each grid cell. In Figure 3, we use different scales of gray to indicate the level of anomalous data in a grid cell. If the ratio exceeds a predefined threshold \( \theta \), the grid cell is considered as anomalous. Then, the anomalous grid cells are aggregated according to the following rules: for a given pair of anomalous grid cells \((X_i, Y_i, N_{Ano}, N_{Sum}, t)\) and \((X_j, Y_j, N_{Ano}, N_{Sum}, t)\), if \(|X_i - X_j| \leq 1\) or \(|Y_i - Y_j| \leq 1\), namely, the two grid cells are adjacent, then they will be merged. After all adjacent anomalous grid cells are merged, clusters (SAEs) of anomalous grid cells are generated. The details of how grid cells are merged is described in the following section.

**Anomaly detection model**

In our model, a series of sensor data from the sensors in a WSN are taken as input, the data from the same sensor are analyzed based on time dimension, and the data from different sensors are analyzed synchronously. Generally, the anomalies detected are further classified into individual anomalous behaviors and anomalous events. As such, our proposed anomaly detection model consists of data preparation phase, initial clustering and grid partition phase, and anomaly detection phase. The proposed anomaly detection model is shown in Figure 4.

**Data preparation.** In the data preparation phase, the sensor data from WSNs are first cleaned and normalized, because it is almost impossible to avoid the situation of data missing or appearing noise, which might lead to skewed data distribution in the process of data acquisition. Then, the important features are further extracted, which might improve detection efficiency and reduce the interference from irrelevant features.

**Initial clustering and grid partition.** In the initial clustering phase, all the data from the sensors with the same time stamp are aggregated by DBSCAN. And the generated clusters are identified as normal data and distinguished from the anomalies. Then, the region where the sensors are deployed is partitioned into grid cells according to the predefined \( \delta \). Furthermore, each anomaly identified by DBSCAN is merged into the corresponding grid cell according to the location of the sensor that send the anomaly. At different times, individual spatial anomalous data and the clusters of spatial anomalous data can be identified by merging the adjacent anomalous grid cells. Evidently, there may exist both anomalous data and normal data from the same sensor with different time stamp. Then, the collected anomalies from the same sensor with different time stamp are analyzed.

**Anomaly detection.** In the anomaly detection phase, not only individual anomalies but also anomalous events are detected by temporal and spatial analysis. When detecting TAEs, the anomalies over a period of time are analyzed, and each series of adjacent anomalies on time dimension, which meet the predetermined temporal requirements, is identified as TAE. When detecting SAEs, the anomalous adjacent grid cells are aggregated by grid-based clustering algorithm, each generated clusters of certain scale are identified as SAEs, and the rest of the anomalies in the other individual anomalous grid cells are identified as anomalous behaviors.

**Anomaly detection algorithm**

In this section, we describe the key steps in the anomaly detection model in detail. As mentioned above, the complete algorithm of AEDTS consists of three key steps: the step of data preparation, which includes the process of filtering missing data, cleaning the sensor data in WSN, and normalizing the data to be analyzed; the step of initial clustering and grid partition, which includes the process of initial temporal anomalous data detection, spatial grid partition, and calculating the statistics of grid cells; and the step of anomalous event detection.
Algorithm 1. AEDTS

Input: \( D = D_1, D_2, \ldots, D_p \) — sets of all sensor data at different \( p \) times
\( d_q = d_1, d_2, \ldots, d_p \) — sensor data from sensor \( q \) at all \( p \) times
\( \delta \) — grid partition parameter; \( \theta \) — predetermined ratio threshold; \( \eta \) — time period parameter

Output: \( AnoE \) — set of anomalous events
1. \( AnoTE = AnoSE = \phi \) // sets of temporal and spatial anomalous events
2. \( AnoB = \phi \) // set of anomalous behaviors
3. \( Attr = \phi \) // set of extracted features
4. for (each \( D_q \) in \( D \))
5. { Clean missing data in \( D_q \)
6. Filtering noises in \( D_q \)
7. }
8. for (each \( d_j \) in \( D \))
9. { result = DBSCAN(\( D_j \)) // initial clustering
10. \( AnoB_{ni} = Anomaly\_Detect(result) \) // anomalies detected at \( t_i \)
11. \( AnoB = AnoB \cup AnoB_{ni} \)
12. }
13. GC = Partition\_region(\( \delta \)) // partition the region into grid cells
14. for (each \( d_j \) in \( D \))
15. { Locate\_GC(GC, \( d_j \)) // locate the corresponding grid cell for \( d_j \)
16. Cal\_GC(AnoB, \( D_{ti} \)) // calculate the statistics for the grid cells in \( GC \)
17. Update\_Grid\_Ano(GC, \( t_i \)) // update the status (anomalous or normal) of grid cell
18. GC\_ano = Filter\_Nor(GC) \( \phi \) // select the anomalous grid cells
19. for (each \( d_j \) in \( D \))
20. { \( Seg\_Ano = Seg\_Ano \cup Filter\_Nor(d_q, \theta) \) // select the anomalous data segments
21. AnoTE = AnoTE \cup Seg\_Ano
22. }
23. }
24. }
25. }
26. GC = GC\_ano
27. for (each \( g_k \) in \( GC \))
28. { \( Clu = Clu \cup g_k \)
29. \( GC' = GC' \cup \{ g_k \} \)
30. while (exist \( g_k \in Clu \&\& g_k \in GC' \))
31. { if ((|g_k, X - g_k, X| \leq 1) \&\& (|g_k, Y - g_k, Y| \leq 1))
32. \( Clu = Clu \cup g_k \) // merge grid cells
33. }
34. }
35. AnoSE = AnoSE \cup Clu
36. AnoE = AnoE \cup AnoTE
37. return AnoE

Experiments

In this section, the performance of AEDTS is evaluated based on the experiments on several UCI datasets and is also compared with three state-of-the-art anomaly detection methods. In section “The datasets,” the datasets used in this article are presented. In section “The baseline methods,” the methods which are employed to compare with AEDTS are introduced. In section “The evaluation,” the evaluation indexes which are widely employed for the experiments of anomaly detection are presented. In section “Experimental results and analysis,” the experimental results are presented in detail.

The datasets

In this article, the performance of AEDTS is evaluated based on three UCI datasets, as shown in Table 1, which include the subset of KDD CUP’99 network intrusion (Http), El Nino, and the Electricity Board Hourly Readings. The KDD CUP’99 network intrusion dataset contains the network intrusion data, the El Nino dataset contains oceanographic and surface meteorological readings taken from nearly 70 moored buoys positioned throughout the equatorial Pacific, and the Electricity Board Hourly Readings dataset contains the real-time readings from Tamil Nadu. As for the data in the datasets, since we need only the attributes with continuous values, the binary attributes and nominal attributes are removed. In this article, With Mulcross data generator, a synthetic dataset with 100,000 data points and 9% anomalies is generated, and we further add 10 anomalous events, each of which includes 100 continuous anomalies, to the synthetic dataset.

The baseline methods

In this article, we compare the performance of AEDTS with three baseline methods, which perform well in anomaly detection or can be used to evaluate the performance of anomaly detection methods. The first baseline method employed in the experiments is ORCA (ORCA is a program for mining outliers in large multivariate data sets) which is based on the well-known k-nearest neighbor (KNN) method. Meanwhile, ORCA employs a sample randomization strategy and a simple pruning rule and almost has a linear time complexity. The second baseline method employed in the

| Datasets            | Instances | Dimensions | Anomaly |
|---------------------|-----------|------------|---------|
| Http (KDD CUP’99)   | 567,497   | 3          | Attack  |
| El Nino            | 178,080   | 12         | Temperature |
| Electricity Board Hourly Readings | 45,781 | 5 | Data |
| Mulcross           | 100,000   | 4          | Data    |

| Table 1. Properties of datasets used in the experiments. |
experiments is one-class SVM. Since one-class SVM with radial basis function kernel performs well in anomaly detection, it is generally used as the baseline method for evaluation. The third baseline method employed in the experiments is local outlier factor (LOF), a density-based method which can detect outliers effectively in high-dimensional dataset.

The evaluation

In order to evaluate the performance of AEDTS, three commonly used evaluation indexes are employed, which are detection rate (DR), false alarm rate (FAR), and run time. As described in Table 2, true positive (TP) denotes the anomalies which are correctly identified as anomalous, false positive (FP) denotes the normal data which are incorrectly identified as anomalies, true negative (TN) denotes the normal data which are correctly identified as normal, and false negative (FN) denotes the anomalies which are incorrectly identified as normal data. Based on the indexes described in the confusion matrix, the DR and FAR are further calculated by equations (1) and (2)

\[
\text{Detection Rate} = \frac{TP}{TP + FP} \quad (1)
\]

\[
\text{False Alarm Rate} = \frac{FP}{FP + TN} \quad (2)
\]

Experimental results and analysis

Performance of anomaly detection. In this experiment, we compare the performance of ORCA, LOF, one-class SVM, and AEDTS using both KDD CUP’99(Http) and El Nino datasets. As for the parameter settings, we set \(k = 15\) in both ORCA and LOF, in order to guarantee that \(k\) is larger than the size of the largest anomalous cluster. For one-class SVM, the commonly used radial basis function kernel is employed as suggested in Caputo et al. For AEDTS model, we set the ratio threshold \(\theta = 0.5\). It can be seen from the receiver operating characteristic (ROC) curves shown in Figure 5 that, both AEDTS and LOF perform better than one-class SVM, while ORCA demonstrates the worst performance. This is mainly because the performance of ORCA is greatly influenced by the shape of clusters. As shown in Figure 6, AEDTS and LOF have similar runtimes, which are slightly better than ORCA.

However, SVM needs the most running time and has the worst efficiency among these four methods.

Performance of AEDTS with different parameter settings. The following experiments investigate the efficiency and effects of AEDTS with different parameter settings on the synthetic dataset. Generally, the region where all sensors are deployed is partitioned into \(m \times n\) grid cells, each of which is a range of \(\delta \times \delta\) size. In this article, the region \(R\) is extended to \(p \times p\) grid cells and the size of each grid cell is still \(\delta \times \delta\), where \(p = \max(m, n)\) and \(\delta = \sqrt{R}/p\). And the time period \(\eta\) is set to be \(k\) times of sampling time interval \(\eta\). As shown in Figure 7, with the increase in \(p\), which is equal to the decrease in \(\delta\), less
sensor data are involved in each grid cell. When $p$ exceeds a certain threshold, the performance of AEDTS deteriorates greatly, which is mainly because very few grid cells can be identified as anomalous grid cells in this situation. As shown in Figure 8, with the increase in $k$, which is equal to the extending of time period, the performance of AEDTS improves first, and when $k$ exceeds a certain threshold, the performance of AEDTS deteriorates dramatically. This is mainly because that TAEs cannot be detected accurately with the time period which is too large or too small. Hence, the best parameters for AEDTS should be selected by the leave-one-out (LOO) cross-validation and a grid search approach.

**Conclusion and future work**

Although many anomaly detection methods have been proposed to date, very few can detect anomalous events
to adapt to the changes that exist in sensor data. In this article, we propose an adaptive method based on contextual anomaly detection in IOT through WSNs, which can not only detect individual anomalies but also detect TAEs and SAEs. The advantages of our proposed method are as follows:

1. Compared to traditional anomaly detection method, our proposed novel contextual anomaly detection model is combined with density-based clustering algorithm that can identify clusters of arbitrary shapes;
2. Based on AEDTS model, not only individual anomalies but also TAEs and SAEs that directly or indirectly reflect the correlation between anomalies can be detected efficiently and effectively;
3. The proposed model can be further improved to detect anomalies in different application scenarios such as remote sensing applications or IOT applications. Furthermore, advanced approaches, such as interactive data mining and machine learning techniques, can be integrated into the model to further improve the detection performance.

In our solution, the proposed model partially depends on the selection of clustering method, which has a certain influence on the detection results. As a future work, we will further improve on the selection strategy of the proposed clustering method and the adaptability of AEDTS. In addition, we will also work on improving the ability of AEDTS on processing real-time data streams and on anomaly detection in various application scenarios.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Guangdong Province Key Area R&D Program of China (Grant No. 2019B010137004), the National Natural Science Foundation of China (Grant Nos 61972108, U1636215, and 61871140), the National Key Research and Development Plan (Grant No. 2018YFB0803504), the National Arts Youth Fund (Grant No. 31901243), and the Science and Technology Project of Taizhou City (Grant No.1901gy19).

**ORCID iD**

Hui Lu https://orcid.org/0000-0002-4120-7716

**References**

1. Yin C, Zhang S and Kim KJ. Mobile anomaly detection based on improved self-organizing maps. Mobile Inf Syst 2017; 2017(1): 1–9.
2. Habeeb R, Nasaruddin F, Gani A, et al. Real-time big data processing for anomaly detection: a survey. Int J Inf Manag 2019; 45(2019): 289–307.
3. Liu Q, Klucik R, Chen C, et al. Unsupervised detection of contextual anomaly in remotely sensor data. Remote Sens Environ 2017; 202(1): 75–87.
4. Alvera-Azcárate A, Sirjacobs D, Barth A, et al. Outlier detection in satellite data using spatial coherence. Remote Sens Environ 2012; 119(1): 84–91.
5. Tan Q, Gao Y, Shi J, et al. Toward a comprehensive insight into the eclipse attacks of Tor hidden services. IEEE Internet Things J 2019; 6(2): 1584–1593.
6. Mart i L, Sanchez-Pi N, Molina JM, et al. Anomaly detection based on sensor data in petroleum industry applications. Sensors 2015; 15(2): 2774–2797.
7. Fateh B and Govindarasu M. Energy minimization by exploiting data redundancy in real-time wireless sensor networks. Ad Hoc Netw 2013; 11(6): 1715–1731.
8. Angiulli F. Dolphin: an efficient algorithm for mining distance-based outliers in very large datasets. ACM Trans Knowl Discov 2009; 3(1): 777–781.
9. Moshtaghi M, Havens TC, Bezdek JC, et al. Clustering ellipses for anomaly detection. Pattern Recogn 2011; 44(1): 55–69.
10. Zhao J, Liu K, Wang W, et al. Adaptive fuzzy clustering based anomaly data detection in energy system of steel industry. Inform Sci 2014; 259(3): 335–345.
11. Shaikh SA and Kitagawa H. Top-k outlier detection from uncertain data. Int J Autom Comput 2014; 11(2): 128–142.
12. Shaikh SA and Kitagawa H. Efficient distance-based outlier detection on uncertain datasets of Gaussian distribution. World Wide Web 2014; 17(4): 511–538.
13. Du X and Chen HH. Security in wireless sensor networks. IEEE Wirel Commun Mag 2008; 15(4): 60–66.
14. Tian Z, Li M, Qiu M, et al. Block-DEF: a secure digital evidence framework using blockchain. Inform Sci 2019; 491(2019): 151–165.
15. Cai D, He X, Han J, et al. Graph regularized nonnegative matrix factorization for data representation. IEEE Trans Pattern Anal Mach Intell 2011; 33(8): 1548–1560.
16. Huang J, Zou J and Xing CC. Competitions among service providers in cloud computing: a new economic model. *IEEE Trans New Serv Manag* 2018; 15: 866–877.

17. Du X, Xiao Y, Guizani M, et al. An effective key management scheme for heterogeneous sensor networks. *Ad Hoc New* 2007; 5(1): 24–34.

18. Xiao Y, Rayi V, Sun B, et al. A survey of key management schemes in wireless sensor networks. *J Comput Commun* 2007; 30(11–12): 2314–2341.

19. Shu X, Yao D and Bertino E. Privacy-preserving detection of sensitive data exposure. *IEEE Trans Inform Forensic Secur* 2015; 10(5): 1092–1103.

20. Li M, Sun Y, Lu H, et al. Deep reinforcement learning for partially observable data poisoning attack in crowdsensing systems. *IEEE Internet Things J*. Epub ahead of print 30 December 2019. DOI: 10.1109/JIOT.2019.2962914

21. Tian Z, Gao X, Su S, et al. Vcash: a novel reputation service providers in cloud computing: a new economic model. *IEEE Trans New Serv Manag* 2018; 15: 866–877.

22. Du X, Guizani M, Xiao Y, et al. A routing-driven elliptic curve cryptography based key management scheme for heterogeneous sensor networks. *IEEE T Wirel Commun* 2009; 8(3): 1223–1229.

23. Xiao Y, Du X, Zhang J, et al. Internet protocol television (IPTV): the killer application for the next generation Internet. *IEEE Commun Mag* 2007; 45(11): 126–134.

24. Liu FT, Ting KM and Zhou ZH. Isolation-based anomaly detection. *ACM Trans Knowl Discov* 2012; 6(1): 74–77.

25. Tian Z, Gao X, Su S, et al. Evaluating reputation management schemes of internet of vehicles based on evolutionary game theory. *IEEE Trans Veh Technol* 2019; 68(6): 5971–5980.

26. Alassi Alhajj D. Effectiveness of template detection on noise reduction and websites summarization. *Inform Sci* 2013; 219: 41–72.

27. Jiang F, Yu Y, Gupta BB, et al. Deep learning based multichannel intelligent attack detection for data security. *IEEE T Sustain Comput*. Epub ahead of print 15 January 2018. DOI: 10.1109/TSSUC.2018.2793284.

28. Tian Z, Luo C, Qiu J, et al. A distributed deep learning system for web attack detection on edge devices. *IEEE T Ind Inform* 2020; 16: 1963–1971.

29. Qiu J, Du L, Zhang D, et al. Nei-TTE: intelligent traffic time estimation based on fine-grained time derivation of road segments for smart city. *IEEE Trans Ind Inform* 2019; 16: 2659–2666.

30. Shen M, Ma B, Zhu L, et al. Cloud-based approximate constrained shortest distance queries over encrypted graphs with privacy protection. *IEEE T Inform Forensic Secur* 2018; 13(4): 940–953.

31. Dong P, Du X, Zhang H, et al. A detection method for a novel DDoS attack against SDN controllers by vast new low-traffic flows. In: Proceedings of the IEEE ICC, Kuala Lumpur, Malaysia, 22–27 May 2016. New York: IEEE.

32. Xiao L, Wan X, Dai C, et al. Security in mobile edge caching with reinforcement learning. *IEEE Wirel Commun* 2018; 25(3): 116–122.

33. Wu L, Du X, Wang W, et al. An out-of-hand authentication scheme for Internet of Things using blockchain technology. In: *Proceedings of IEEE ICNC 2018*, Maui, HI, 5–8 March 2018. New York: IEEE.

34. Qiu J, Tian Z, Du C, et al. A survey on access control in the age of Internet of Things. *IEEE Internet Things J*. Epub ahead of print 24 January 2020. DOI: 10.1109/JIOT.2020.2969326

35. Jiang N, Tian F, Li J, et al. MAN: mutual attention neural networks model for aspect-level sentiment classification in IoT. *IEEE Internet Things J* 2020; 7: 2901–2913.

36. Schikopf B, Platt J, Shawe-Taylor J, et al. Estimating the support of a high-dimensional distribution. *Neural Comput* 2001; 13(7): 1443–1471.

37. Huang H, Qin H, Yoo S, et al. Physics-based anomaly detection defined on manifold space. *ACM Trans Knowl Discov* 2014; 9(2): 1–39.

38. Ma J, Sun L, Wang H, et al. Supervised anomaly detection in uncertain pseudoperiodic data streams. *ACM Trans Internet Technol* 2016; 16(1): 1–20.

39. Chen PY, Yang S and McCann JA. Distributed real-time anomaly detection in networked industrial sensing systems. *IEEE T Ind Electron* 2015; 62(6): 3832–3842.

40. Liu Q, Klucik R, Chen C, et al. Unsupervised detection of contextual anomaly in remotely sensed data. *Remote Sens Environ* 2017. DOI: 10.1016/j.rse.2017.01.034

41. Jiang N, Xu D, Zhou J, et al. Toward optimal participant decisions with voting-based incentive model for crowd sensing. *Inform Sci* 2020; 512: 1–17.

42. Zhang W, Zhang B, Zhou Y, et al. An IoT honeynet based on multi-port honeypots for capturing IoT attacks. *IEEE Internet Things J*. Epub ahead of print 28 November 2019. DOI: 10.1109/JIOT.2019.2965173.

43. Zhang W, Wang H, He H, et al. DAMBA: detecting android malware by ORGB analysis. *IEEE T Reliab* 2020; 69: 55–69.

44. Khedo K, Doomun R and Aucharuz S. READA: redundancy elimination for accurate data aggregation in wireless sensor networks. *Wirel Sens Netw* 2010; 2(4): 300–308.

45. Chandola V, Banerjee A and Kumar V. Anomaly detection: a survey. *ACM Comput Surv* 2009; 43(3): 1–58.

46. Tian Z, Shi W, Wand Y, et al. Real-time lateral movement detection based on evidence reasoning network for edge computing environment. *IEEE Trans Ind Inform* 2019; 15(7): 4285–4294.

47. Tian Z, Su S, Shi W, et al. A data-driven method for future internet route decision modeling. *Future Gener Comput Syst* 2019; 95: 212–220.

48. Xiao L, Li Y, Huang X, et al. Cloud-based malware detection game for mobile devices with offloading. *IEEE Trans Mobile Comput* 2017; 16(10): 2742–2750.

49. Chen Y, Abraham A and Yang B. Hybrid flexible neural-tree-based intrusion detection systems. *Int J Intel Syst* 2007; 22: 337–352.

50. Caputo B, Sim K, Furesjo F, et al. Appearance-based object recognition using SVMs: which kernel should I use? In: *Proceedings of NIPS workshop on statistical methods for computational experiments in visual processing and computer vision 2002*. 