Hindawi
Scientific Programming
Volume 2022, Article ID 4671381, 7 pages
https://doi.org/10.1155/2022/4671381

Research Article

Detecting Anomaly Data for IoT Sensor Networks

Zhe Wei [1] and Fang Wang [2]

[1] *School of Computer Science, Civil Aviation Flight University of China, Guanghan 618307, China*
[2] *School of Science, Civil Aviation Flight University of China, Guanghan 618307, China*

Correspondence should be addressed to Zhe Wei; findwei@foxmail.com

Received 21 January 2022; Revised 5 April 2022; Accepted 4 May 2022; Published 21 May 2022

Academic Editor: Le Sun

The Internet of Things, or IoT, has been widely recognized as a new perception paradigm for interacting between the digital world and the physical one. Acting as the interface and integral part of the Internet of Things, sensors embedded within the network are the principal components that collect the unprocessed data, and these sensors are usually deployed in unattended, hostile, or harsh areas, which inevitably makes the sensor readings prone to faults and even anomalies. Therefore, the quality of sensor readings will ultimately affect the quality of various data-oriented IoT services, and the sensor data are of vital importance affecting the performance of the system. However, the data anomaly detection is a nontrivial task for IoT because sensors are usually resource-constrained devices with limited computing, communication, and capacity. Therefore, an efficient and lightweight detecting method is needed to meet the requirements. In this study, we deal with the anomaly data by detecting the source sensor nodes through combination methods of the local outlier factor and time series. Simulations show that the proposed method can effectively detect the anomaly data and presents a better normal data rate.

1. Introduction

The Internet of Things, or IoT, has received extensive attention in the past few years by the research community owning to the progress of computing and real-time connections between data and devices and has been used in many application fields such as smart home/office, automobile, and medical assistance to solve practical problems [1–3]. The IoT depicts a future computing scenario where everyday physical objects will be connected to the Internet and identify themselves [4]. Nowadays, the Internet of Things is becoming a more and more important infrastructure component, and due to the heterogeneity of IoT devices, the data exchange between IoT sensors and various applications achieves rapid growth.

However, with the wide spread applications of the IoT, security threats are also becoming increasingly prominent [5–8]. The IoT is vulnerable to attacks from communication channels, which has become a common security problem [9]. IoT sensor failure may interrupt the system control [10], thus interrupting the services provided by the IoT system. The distributed deployment of IoT sensors makes networking more convenient, but it also brings more difficult risks [11]. In addition, many IoT applications are of heterogeneous components such as different sensors, services, protocols, and communication technologies like Zigbee, WiFi, and Bluetooth, which generate the complexity of the network management [12]. Therefore, anomaly data come into existence accompanied by the threats, attacks, risks, integration of heterogeneous technologies, and various IoT applications.

In the field of data mining and statistics, anomalies are usually referred to as either deviants or outliers [13]. The definition of anomaly data given in [14] is that anomaly data are data points that behave very differently from others or conform to some predefined abnormal behaviors. The definition of anomaly given in [15] is that it is an observation deviating so much from others to generate uncertainties. According to [16], the main causes of anomaly data are of two aspects: (1) internal malfunction, i.e., noise and fault caused by sensor hardware and software failure; (2) external influence, i.e., specific events occur in the places where nodes are deployed. Essentially, anomaly-based detection is an intrusion detection mechanism, it can be used to perceive important network mode attacks [17], and anomaly detection refers to identifying suspicious data items, events, or observations that are significantly different from most other data [18]. The application of anomaly detection can make certain contributions to the IoT data protection [19], and it is resource friendly and provides more extensive detection, which is very suitable for IoT sensor network applications [20].

Nevertheless, the anomaly data detection is a nontrivial task for IoT. The IoT nodes are usually resource-constrained sensors with low-cost embedded systems [21], and traditional anomaly detection solutions cannot be directly applied into IoT [22]. For this reason, it is important that a trade-off solution be found to the problem with decent accuracy while bringing minimum overhead.

This research aims to detect the anomaly sensor data under the IoT environment. Different from other works that focus on analyzing the sensor data and evaluating the analyzed results by certain rules, according to the dynamic features of IoT networks, our proposed method tries to trace and identify the source nodes that generate or cause the anomaly data, and by identifying and confirming the suspected nodes, the anomaly data can be eventually deleted from the network. This study uses combination methods of local outlier factor and time series. The local outlier factor is applied to check the abnormal data so as to label the related sensor node as suspected one. Due to the dynamic changing environment of the IoT network, nodes might be influenced by the temporary error or communication interference resulting in being labeled as suspected nodes. Thus time series method is used to evaluate and further confirm these potential suspected nodes from the perspective of time windows.

The main contribution of this work is to provide a lightweight yet effective method for anomaly data detection in IoT sensor networks, due to the fact that sensors in IoT are usually resource-constrained devices in terms of computing, communication, and capacity. To this end, both the methods used in this study are of ease computing. Specifically, the local outlier factor is a density-based detecting algorithm which is simpler and not particularly picky about the distribution of datasets. By contrast, most of the other anomaly detection algorithms are based on statistical methods or borrow some clustering algorithms for anomaly identification; besides, the datasets obeying a specific probability distribution are usually assumed. Further, in the time series, instead of directly using traditional methods, Chebyshev polynomials are applied for the approximation of time series especially in comparing the similarity of two time series, which is also lightweight in computing and can be done through the related polynomial coefficients.

The remainder of this study is organized as follows. Section 2 presents the related works about the anomaly detection for sensor data and related definitions and classifications. Section 3 introduces the local outlier factor and time series methods in detail, on which our proposed method is based. Simulations are presented in Section 4, and Section 5 concludes this work.

2. Related Works

Anomaly detection comes from the data mining and statistics field, and it establishes a standard model to judge whether the relevant data match the model. According to [23], the anomaly detection algorithm for sensor data is classified as three aspects, i.e. (1) statistical method, which classifies the anomalies by measuring the probability of the measured data relative to the model; (2) proximity-based method, which relies on the distance between measured data to distinguish abnormal data from correct data; (3) prediction-based method, in which the past measurement data are used to train the model and it can predict the next measured value in the sensor data.

Machine learning-based anomaly detection methods have received much attention in the research community. The machine learning algorithms are used on the interested data and train the related models through the pattern extraction [24], based on which the anomaly detection uses machine learning technology to detect abnormal activities in network traffic packets [25]. Pathak et al. [4] applied supervised and unsupervised machine learning to solve the tampering problem of sensors in the Internet of Things. In [4], the real-time view of traffic pattern is considered to train the unsupervised machine learning method based on isolated forest for anomaly detection; it creates labels according to the traffic pattern, uses the decision tree supervision method to monitor all Internet of Things traffic on the gateway, and sends an alarm to the administrator when an anomaly is detected.

Kim et al. [13] provided a method for real-time detection and notification of abnormal conditions through machine learning by generating synthetic datasets for learning real-time data anomaly detection algorithms and by testing models based on gated recursive units and long-term and short-term memory for predicting time series data anomalies. Kim et al. [13] detected and notified abnormal conditions in the worker environment through sensor data. The method is based on the prediction-based anomaly detection method and neural network and is evaluated using synthetic data generated from time series with trend, season, and noise components. Kim et al. [13] further explained how to use neural networks to detect anomalies and how to evaluate the proposed model. Based on the combination of machine learning and visual data analysis, Vasily et al. [26] proposed an anomaly detection method in wireless sensor networks. Taking a water management system as an example, the method is tested, and the necessary datasets are generated by using the software model for testing anomaly detection.

Different from other works that mainly focus on the network layer and application layer, an adaptive context-aware anomaly detection method is proposed in [10] which centers on the physical properties of the IoT sensor system and identifies anomalous incidents in the environment properties of the system. This method uses a sensor association algorithm which can generate sensor fingerprints, cluster these fingerprints, and extract the context of the system. Then, according to the contextual information and through long short-term memory neural network and Gaussian estimator, the anomalies in the system together with the source could be detected.

Through distinguishing hostile events about the traffic pattern in the distributed smart space orchestration system, Reddy et al. [9] presented an anomaly detection system with characteristic examination. Reddy et al. [9] used a method based on supervised meta algorithm called bagging (which is one of the ensemble meta estimator learning technologies [27] and considers multiple predictors to calculate the aggregation predictor) to classify and process malicious operations and train the classifier-based anomaly detection to build a clarification model according to the intrusion data and predict the system by identifying when the system is in an abnormal state.

The authors in [16] studied the sensor fault and external event detection scheme in wireless sensor networks. Based on spatiotemporal fusion, Chen et al. [16] proposed a distributed anomaly detection method for wireless sensor networks based on one-class quarter-spherical support vector machine (QSSVM). In this method, the QSSVM model is trained to obtain relevant parameters, then the trained model is used to classify the streaming data in the network, and then the abnormal data types are determined and divided into noise, faults, and events. By converting unsupervised time series data into supervised form, Das et al. [28] proposed a segmentation-based anomaly detection method of IoT sensor data. In order to ensure that the data are not affected by inherent noise, the method performs Holt–Winters exponential smoothing [29] on the dataset and then trains a long-term memory neural network for the anomaly detection. Blockchain is a distributed network with some unique functions such as decentralization, transparency, and system autonomy [30], which can enhance network security and cooperation in the Internet of Things.

3. Preliminaries

In this research, our proposed method is mainly composed of two components, i.e., local outlier factor and Chebyshev polynomials-based time series, and as mentioned previously, both the components or methods are lightweight in computing and are applicable to resource-constrained IoT sensor nodes when dealing with anomaly data. They are introduced as follows.

Local outlier factor, or LOF, is a nearest neighbor algorithm. It attributes a fault or outlier score to each sensor reading based on the number of measurements around its K-nearest neighbors and the number of measurements around the sensor reading. Sensor readings with high scores are flagged as abnormal [31]. Anomalous data or outliers are treated as sensor data streams that are significantly different from normal behavioral data, and outlier detection can detect a high probability of false reads or data corruption, thereby ensuring the quality of data collected by sensors [32].

LOF is density-based outlier detection which has a fundamental assumption that the density around a nonoutlier object should be similar to the density around its neighborhood, while the density around an outlier object should be significantly different from that around its neighborhood. By assigning each data point an outlier factor that depends on the neighborhood density, it then evaluates whether the data point is an outlier. The larger the anomaly factor of the data is, the more likely the data are anomalous. The advantage of LOF is that it gives the degree to which a data point is an outlier [33].

The LOF algorithm is constructed on two main components, namely, the reachable distance and the local reachable density [34–36]. Based on the distance between the object p and each point in its k neighbors, the reachable distance is defined as

$$reach - dist_k(p, o) = \max\{k - dist(o), d(p, o)\}. \tag{1}$$

According to the mean distance of each data object in the neighborhood, a density parameter can be obtained, called the local reachable density, which is defined as

$$Lrd_k(p) = \frac{1}{\sum_{o \in N_k(p)} reach - di\ st_k(p, o) / |N_k(p)|}. \tag{2}$$

Through the mean value of the ratio of the local reachable density of p to the local reachable density of its nearest k neighbors, the local outlier factor of p is defined as

$$LOF_k(p) = \frac{\sum_{o \in N_k(p)} Lrd_k(o) / Lrd_k(p)}{|N_k(p)|}. \tag{3}$$

The idea behind LOF algorithm is to calculate outliers by drawing a circle centered on any but specific data point p, so that at least k data points are in the circle, and see how dense the neighborhood around p is [33].

A feature of the data generated by the IoT sensors is that due to the observed changes in the nature of the phenomenon, the data distribution may change in the network life cycle, and the anomaly detection technology must be able to adapt to the nonstationary data distribution to achieve the best performance [37]. To this end, the time series approach is applied in our proposed method.

As mentioned earlier, instead of directly using traditional time series methods, Chebyshev polynomials are applied here as a lightweight method for approximating the time series especially in comparing the similarity of two time series, in which it is not necessary to calculate all polynomials, and the similarity can be observed by comparing the related Chebyshev coefficients [38]. In [39], let $P_m(t)$ be a polynomial of t with degree m and $P_m(t) = cos(mcos^{-1}(t))$, where $t \in [-1, 1]$. Because of $cosm\theta + cos(m - 2)\theta = 2$ $cos\theta cos(m - 1)\theta$, $P_m(t) = cos(mcos^{-1}(t))$ can be rewritten into a recurrence relation, i.e., $P_m(t) = 2tP_{m-1}(t) - P_{m-2}(t)$, where $m \geq 2$. Due to the characteristics of Chebyshev polynomials, for an arbitrary function $f(t)$, it can be approximated as $f(t) \approx c_0P_0 + c_1P_1 + \ldots + c_mP_m$, where $c_0, \ldots c_m$ denote the coefficients of Chebyshev polynomials. Further, according to the Gauss–Chebyshev formula [40], the coefficients are defined as

$$\begin{cases} c_0 = \dfrac{1}{m}\displaystyle\sum_{j=1}^{m} f\left(t_j\right)P_0\left(t_j\right) = \dfrac{1}{m}\displaystyle\sum_{j=1}^{m} f\left(t_j\right) & i = 0 \\ c_i = \dfrac{2}{m}\displaystyle\sum_{j=1}^{m} f\left(t_j\right)P_i\left(t_j\right) & 1 \leq i \leq m \end{cases}. \tag{4}$$

Equation (4) is only applicable to interval functions and cannot be directly applied to the time series of coefficient calculation. Discrete sequences need to be extended to interval functions. Assume $T = \{(t_1, v_1), \ldots, (t_N, v_N)\}$ is a time series where $-1 \leq t_1 < \ldots < t_N \leq 1$ and time t is normalized in $[-1, 1]$ resulting in the division of interval $[-1, 1]$ into N disjoint subintervals as follows [39].

$$I_i = \begin{cases} \left[-1, \dfrac{t_1 + t_2}{2}\right] & if\ i = 1 \\ \left[\dfrac{t_{i-1} + t_i}{2}, \dfrac{t_i + t_{i+1}}{2}\right] & if\ 2 \leq i \leq N-1 \\ \left[\dfrac{t_{N-1} + t_N}{2}, 1\right] & if\ i = N \end{cases}. \tag{5}$$

Map v_i into an interval function denoted by $g(t) = v_i$ where $t \in I_i$ $1 \leq i \leq N$. Being extended into an interval function, the time series is defined by

$$f'(t) = \frac{g(t)}{\sqrt{w(t)\left|I_i\right|}}, \tag{6}$$

where $t \in I_i$, $1 \leq i \leq N$, $|I_i|$ is the length of subinterval I_i, and $w(t)$ is the weight function defined as $w(t) = 1/\sqrt{1 - t^2}$[39]. The Chebyshev coefficients of time series are now calculated as follows, and the details of the above calculation steps can be referred to [39].

$$\begin{cases} c_0 = \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} f'\left(t_j\right)P_0\left(t_j\right) = \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} f\left(t_j\right) & i = 0 \\ c_i = \dfrac{2}{N}\displaystyle\sum_{j=1}^{N} f'\left(t_j\right)P_i\left(t_j\right) & 1 \leq i \leq N \end{cases}. \tag{7}$$

4. Simulations

An IoT sensor cluster is formed to sense the oxygen content of a certain workshop. The oxygen content in an ordinary workshop environment shall be 18%~21%, or else ventilation measures shall be taken if it is not within this range. Each sensor forwards related data to a base station (BS). The cluster consists of 80 nodes, of which 20 are malicious nodes and 60 are normal nodes. Assume that normal sensor nodes correctly sense and forward the data, while malicious nodes selectively falsify or modify the normal oxygen content data into the range rather than 18% ~ 21% so as to damage the system. Besides, several other assumptions are also made: (1) all nodes form a star topology and are evenly distributed in a circular area centered on a base station; (2) each node has a unique ID, and the header of each packet includes source node ID, packet group length, and packet sequence number; (3) sensor nodes have 1% communication error and the direct communication between each sensor node and the base station is also assumed.

The base station regularly sends oxygen content requests to the sensor nodes in the network. When a request ends, the BS node firstly uses the LOF algorithm to analyze the data received from the network, checks the abnormal data, and labels its corresponding sender nodes as suspected ones. Secondly, the time series method is used to track and analyze the suspected nodes with regard to their subsequent data in the following requests. When the analysis result is greater than the given threshold, these nodes are confirmed to be malicious. Then, the BS will not accept the data from these malicious nodes any more.

Suppose d_i and d_j are two data time series, and let c_{d_i} and c_{d_j} be the corresponding vectors of Chebyshev coefficients with $c_{d_i} = [x_0, \ldots, x_m]$ and $c_{d_j} = [y_0, \ldots, y_m]$. By comparing the corresponding Chebyshev coefficients, the similarity of the two time series can be obtained. In consideration of computing simplicity and node energy saving, the Euclidean distance is applied here and it is defined as

$$Dist\left(c_{d_i}, c_{d_j}\right) = \sqrt{\frac{\pi}{2}\sum_{k=0}^{m}\left(x_k - y_k\right)^2}. \tag{8}$$

A threshold μ is set in this method for measuring the distance, e.g., if $Di\ st(c_{d_i}, c_{d_j}) \geq \mu$ is established, it indicates that the result of time series comparison is abnormal, and the related sensor node is considered malicious and should be isolated from the network.

In Figures 1–3, we test the normal data rate, or NDR. NDR is the ratio of the amount of normal data received to the amount of normal data the base station should receive. m is the number of Chebyshev coefficients and μ is the threshold for similarity calculation. len represents the length of the time series, making it equal to a certain number of queries initiated by the BS in the different tests. We divide the simulations into three groups to test the influence of these three parameters on NDR, respectively.

In Figure 1, as the number of base station queries increases, the NDR values of the three sets of comparison parameters continue to increase, among which ($m = 4$, $\mu = 0.9$, len = 10) is the fastest followed by both ($m = 3$, $\mu = 0.9$, len = 10) and ($m = 2$, $\mu = 0.9$, len = 10), which are relatively the slowest. For example, for the 100th query, the NDR values of the three groups are about 0.85, 0.84, and 0.82, respectively. This is because the time series and related similarity computations help to expand the anomalous patterns of anomaly nodes and identify them efficiently. It can also be noticed in Figure 1 that different coefficients have different effects on NDR. For example, the NDR at $m = 4$ is significantly higher than that at $m = 2$. But the increase of m means that more coefficients are needed which will definitely make the calculation more complicated. It can be seen that

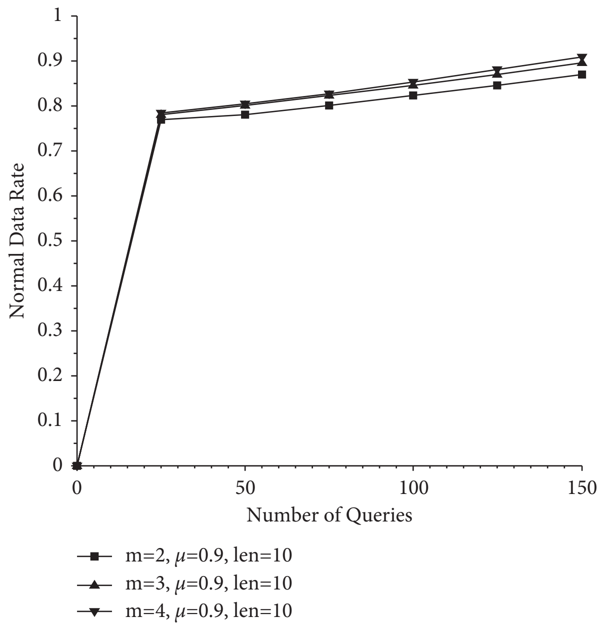

FIGURE 1: Normal data rate ($m = 2/3/4$, $\mu = 0.9$, len = 10).

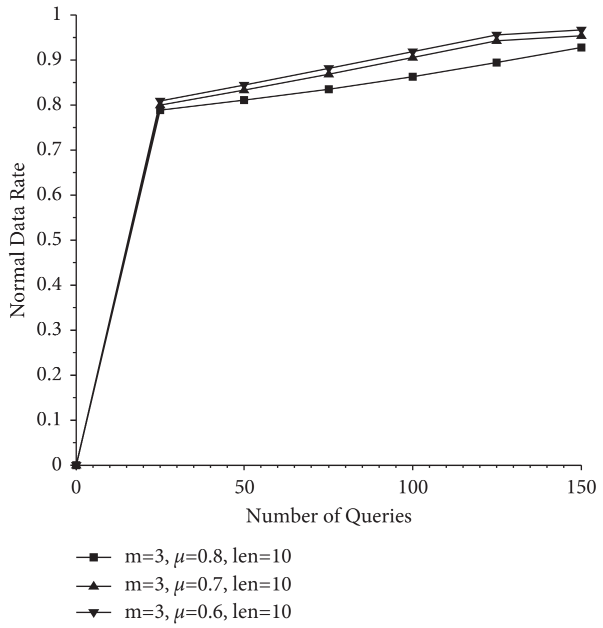

FIGURE 2: Normal data rate ($m = 3$, $\mu = 0.8/0.7/0.6$, len = 10).

the NDR at $m = 4$ is similar to the NDR at $m = 3$. Therefore, this study recommends $m = 3$.

To make the similarity calculation more strict, different thresholds μ are set in Figure 2. As is shown in the test, more anomaly nodes are being identified, resulting in higher NDR. For instance, for the 100th query, the NDR values of the three groups are about 0.91, 0.90, and 0.86, respectively. Note that μ should be appropriately reduced; otherwise, normal nodes will be misjudged as anomaly nodes. This is because normal nodes cannot guarantee 100% trouble-free operation in the event of a communication error or even a dead battery. As can be seen from Figure 2, the difference between the NDR at $\mu = 0.7$ and NDR at $\mu = 0.6$ is small, so $\mu = 0.7$ is recommended. Similar test result trends can also be seen in Figure 3, where different len values are set. For example, still for the 100th query, the NDR with len = 25 is around 0.96 and is much higher than the others. Higher len values indicate larger time windows and the higher the len value is, the better the NDR effect becomes. However, the increase of len will also increase the complexity of calculation, so it should be chosen wisely.

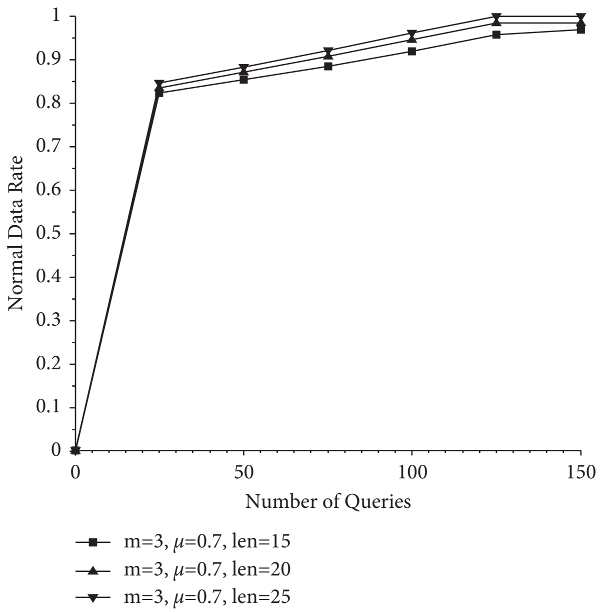

FIGURE 3: Normal data rate ($m = 3$, $\mu = 0.7$, len = 15/20/25).

5. Conclusions

The data quality collected by sensor nodes is affected by anomalies like abnormal events and malicious attacks, and when the anomaly datasets enter the system, the overall system performance would be affected making the system unreliable. Therefore, anomaly detection is a necessary process to ensure the quality of sensor data before it is used for analysis and decision making. In the field of Internet of Things, anomaly detection is an ongoing research field aiming to provide protection against abnormal sensor readings. In addition, due to its low price and commercial attraction, security has not been given much priority. Therefore, it is necessary to protect Internet of Things devices and smart homes from potentially destructive abnormal data and related source sensor nodes.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to express their gratitude to the Civil Aviation Information Technology Research Center of CAFUC for the support and help. This study was partially supported by the Scientific Project of CAFUC under grant nos. JG2022-06 and J2022-042, the Central University Teaching Reform Project under grant nos. E2020044 and E2021038, Civil Aviation Professional Project under grant nos. 0252109 and MHJY2022038, and the Sichuan Science and Technology Program under grant nos. 2022YFG0190 and 2022JDR0116.

References

[1] Y. Wang, L. Sun, and S. Subramani, "CAB: classifying arrhythmias based on imbalanced sensor data," *KSII Transactions on Internet & Information Systems*, vol. 15, no. 7, pp. 2304–2320, 2021.

[2] L. Sun, Q. Yu, D. Peng, S. Subramani, and X. Wang, "Fogmed: a fog-based framework for disease prognosis based medical sensor data streams," *Computers, Materials & Continua*, vol. 66, no. 1, pp. 603–619, 2020.

[3] L. Sun, Y. Wang, Z. Qu, and N. N. Xiong, "BeatClass: A Sustainable ECG Classification System in IoT-Based eHealth," *IEEE Internet of Things Journal*, vol. 9, 2021.

[4] A. K. Pathak, S. Saguna, K. Mitra, and C. Åhlund, "Anomaly Detection Using Machine Learning to Discover Sensor Tampering in IoT Systems," in *Proceedings of the ICC 2021 - IEEE International Conference on Communications*, pp. 1–6, Montreal, QC, USA, June 2021.

[5] E. Borgia, "The Internet of things vision: key features, applications and open issues," *Computer Communications*, vol. 54, pp. 1–31, 2014.

[6] I. Alrashdi, A. Alqazzaz, E. Aloufi, R. Alharthi, M. Zohdy, and H. Ming, "AD-IoT: anomaly detection of IoT cyberattacks in smart city using machine learning," ", in *Proceedings of the IEEE 9th Annual Computing and Communication Workshop and Conference*, pp. 305–310, Las Vegas, NV, USA, January 2019.

[7] M. Wu, L. Tan, and N. Xiong, "A structure fidelity approach for big data collection in wireless sensor networks," *Sensors*, vol. 15, no. 1, pp. 248–273, 2014.

[8] S. Huang, A. Liu, S. Zhang, T. Wang, and N. N. Xiong, "BD-VTE: a novel baseline data based verifiable trust evaluation scheme for smart network systems," *IEEE Transactions on Network Science and Engineering*, vol. 8, 2020.

[9] D. K. K. Reddy, H. S. Behera, G. M. S. Pratyusha, and R. Karri, "Ensemble bagging approach for IoT sensor based anomaly detection," *Lecture Notes in Electrical Engineering*, vol. 702, pp. 647–665, 2021.

[10] R. Yasaei, F. Hernandez, and M. A. A. Faruque, "IoT-CAD," *Proceedings of the 39th International Conference on Computer-Aided Design*, vol. 9, pp. 1–9, 2020.

[11] X. Yang, Y. Chen, X. Qian, T. Li, and X. Lv, "BCEAD: a blockchain-empowered ensemble anomaly detection for wireless sensor network via isolation forest," *Security and Communication Networks*, vol. 2021, pp. 1–10, Article ID 9430132, 2021.

[12] F. Restuccia and S. T. D'Oro, "Securing the Internet of things in the age of machine learning and software-defined networking," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4829–4842, 2018.

[13] J. Kim, Y. W. Cho, and D. Kim, "Anomaly detection of environmental sensor data using recurrent neural network at the edge device," in *Proceedings of the International Conference on Information and Communication Technology Convergence*, pp. 1624–1628, ICTC), Jeju, Korea (South), October 2020.

[14] S. Sadik and L. Gruenwald, "Online outlier detection for data streams," in *Proceedings of the IDEAS 2011 Proceedings of the 15th Symposium on International Database Engineering & Applications*, pp. 88–96, Barcelona, Spain, December 2011.

[15] K. Haseeb and N. M. Q. O. E. K. N. W. T. Abbas, "RCER: reliable cluster-based energy-aware routing protocol for heterogeneous wireless sensor networks," *PLoS One*, vol. 14, no. 9, pp. e0222009–24, 2019.

[16] Z. Chen, H. Wu, H. Zhu, and Y. Miao, "Distributed anomaly detection method in wireless sensor networks based on temporal-spatial QSSVM," *Advances in Intelligent Systems and Computing*, vol. 905, pp. 934–943, 2020.

[17] E. Hodo and X. A. P.-L. E. C. R. Bellekens, "Threat analysis of IoT networks using artificial neural network intrusion detection system," *2016 International Symposium on Networks, Computers and Communications (ISNCC)*, in *Proceedings of the 2016 2016 International Symposium on Networks, Computers and Communications (ISNCC)*, pp. 1–6, Yasmine Hammamet, Tunisia, May 2016.

[18] T. Fernando and H. S. S. C. Gammulle, "Deep learning for medical anomaly detection - a survey," *ACM Computing Surveys*, vol. 54, no. 7, pp. 1–37, 2022.

[19] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: methods, systems and tools," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 303–336, 2014.

[20] 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, pp. 2491–2524, 2018.

[21] D. Widhalm, K. M. Goeschka, W. Kastner, and SoK, "A taxonomy for anomaly detection in wireless sensor networks focused on node-level techniques," *Proceedings of the 15th International Conference on Availability, Reliability and Security*, vol. 17, pp. 1–10, 2020.

[22] M. Rassam, A. Zainal, and M. Maarof, "Advancements of data anomaly detection research in wireless sensor networks: a survey and open issues," *Sensors*, vol. 13, no. 8, pp. 10087–10122, 2013.

[23] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection," *ACM Computing Surveys*, vol. 41, no. 3, pp. 1–58, 2009.

[24] A. A. Elngar and S. Bhatt, "IoT-based efficient tamper detection mechanism for healthcare application," *IJ Network Security*, vol. 20, no. 3, pp. 489–495, 2018.

[25] R. K. Shrivastava, S. Mishra, V. E. Archana, and C. Hota, "Preventing data tampering in IoT networks," in *Proceedings of the IEEE International Conference on Advanced Networks and Telecommunications Systems*, pp. 1–6, Goa, India, December 2019.

[26] A. A. Vasily, V. Desnitsky, I. Kotenko, E. Novikova, and A. Shulepov, "Combined approach to anomaly detection inWireless sensor networks on example of water management system," in *Proceedings of the 10th Mediterranean Conference on Embedded Computing*, pp. 7–10, Budva, Montenegro, June 2021.

[27] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[28] T. Das, R. M. ShuklaT, and S. Sengupta, "Imposters among us: a supervised learning approach to anomaly detection in IoT

sensor data," in *Proceedings of the IEEE 7th World Forum on Internet of Things*, pp. 818–823, New Orleans, LA, USA, July 2021.

[29] P. R. Winters, "Forecasting sales by exponentially weighted moving averages," *Management Science*, vol. 6, no. 3, pp. 324–342, 1960.

[30] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the Internet of things," *IEEE Access*, vol. 4, pp. 2292–2303, 2016.

[31] M. Ahmed, A. Naser Mahmood, and J. Hu, "A survey of network anomaly detection techniques," *Journal of Network and Computer Applications*, vol. 60, pp. 19–31, 2016.

[32] A. Gaddam, T. Wilkin, and M. Angelova, "Anomaly detection models for detecting sensor faults and outliers in the IoT - a survey," *2019 13th International Conference on Sensing Technology (ICST)*, in *Proceedings of the 2019 13th International Conference on Sensing Technology (ICST)*, pp. 1–6, Sydney, NSW, Australia, December 2019.

[33] D. McDonald and S. S. F. Sanchez, "A survey of methods for finding outliers in wireless sensor networks," *Journal of Network and Systems Management*, vol. 23, no. 1, pp. 163–182, 2015.

[34] Z. Gan and X. Zhou, "Abnormal network traffic detection based on improved LOF algorithm," *2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, vol. 1, pp. 142–145, 2018.

[35] M. M. Breunig and H.-P. R. T. J. Kriegel, "Lof," *ACM SIGMOD Record*, vol. 29, no. 2, pp. 93–104, 2000.

[36] F. Wang, Z. Wei, and X. Zuo, "Anomaly IoT node detection based on local outlier factor and time series," *Computers, Materials & Continua*, vol. 64, no. 2, pp. 1063–1073, 2020.

[37] C. OReilly, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Anomaly detection in wireless sensor networks in a non-stationary environment," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1413–1432, 2014.

[38] F. Wang and Z. Wei, "A statistical trust for detecting malicious nodes in IoT sensor networks," *IEICE - Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E104.A, no. 8, pp. 1084–1087, 2021.

[39] Y. Cai and R. Ng, "Indexing spatio-temporal trajectories with Chebyshev polynomials," *Proceedings of the 2004 ACM SIGMOD international conference on Management of data - SIGMOD '04"*, in *Proceedings of the 2004 ACM SIGMOD international conference on Management of data - SIGMOD '04*, pp. 599–610, Paris, France, June 2004.

[40] J. C. Mason and D. Handscomb, *Chebyshev Polynomials"*, Chapman & Hall, 2003.