Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 5, pp. 3633–3640, DOI 10.12694/scpe.v25i5.3195

STATE MONITORING AND ANOMALY DETECTION ALGORITHMS FOR ELECTRICITY METERS BASED ON IOT TECHNOLOGY

CHUNGUANG WANG*, TIANFU HUANG†, ZHIWU WU‡, YING ZHANG§, AND HANBIN HUANG¶

Abstract. In response to the practical application of the electricity consumption information collection system in the online monitoring business of measuring equipment, the author introduces a method for analyzing the abnormal flying away of electricity meters based on the IoT technology LOF local anomaly detection algorithm. This method can effectively determine whether the abnormal energy representation value belongs to accidental or trend anomalies by calculating the abnormal factor of the energy representation value. After excluding the influence of accidental data, perform a secondary judgment on the abnormal flight of the energy meter. The experimental results show that when calculating the LOF factor of the electricity meter, it can be found that the LOF curve data range is mainly concentrated in the range of 0.8 to 1.3, and there is no significant change in the LOF factor near the mutation point. This proves that this method can effectively improve the accuracy of anomaly detection, avoid misjudgment of faults, and improve the efficiency of on-site fault handling.

Key words: Smart energy meters, Electricity information collection system, LOF, Outliers

1. Introduction. The electricity settlement of power enterprises is mainly completed through energy metering devices. Electricity metering management is an important link in the production and operation management of power enterprises and the safe operation of the power grid. Its technology and management level not only affect the development and corporate image of power enterprises, but also affect the accuracy and fairness of trade settlement, involving the interests of a large number of power customers [1]. Therefore, in order to ensure the accuracy and reliability of the energy meter, it is necessary to reduce the error of the energy meter and make direct payment. fair and reasonable. In order to ensure the accuracy and reliability of the electricity meter, it is necessary to ensure the accuracy and reliability of the standard electricity measurement equipment first [2]. Electrical measuring instruments are distributed only to various state electrical testing centers and power plant states, and currently, the measurements of measuring instruments in various modes are monitored in the laboratory by directly connected computers by the measuring staff in electronic measuring instruments. Since the power meter cannot be monitored remotely, if any abnormality or problem is detected during the measurement, the calibration staff can only remind the management to go to the laboratory for maintenance immediately, which will delay the detection. problems, long-term solutions, low performance, high cost, and difficulty adapting to the growing demand for energy metering.For example, a three-phase energy meter standard device in the metering center of a certain power supply bureau has a total of 16 calibration meter positions [3]. From the surface analysis, it can be judged that the crimping of positions 3, 4, and 11 is damaged and cannot properly calibrate three-phase energy meters, the remaining meter positions are working normally, and the calibration data also meets the requirements of the electric energy meter calibration regulations. There is a special case where the calibration data of the 16 meter positions is better than that of other normally working meter positions. It took three working days for maintenance personnel to check and find that the compression joint of the 16 meter positions was severely damaged, resulting in abnormal data. Due to the large workload of calibration, calibration personnel can only pay attention to whether the calibration data is out of tolerance and whether the voltage connection of the energy meter is normal for most of the time during the calibration

*Marketing Service Center of State Grid Fujian Electric Power Co., Ltd., Fuzhou, Fujian, 350013, China (Corresponding author, lunwen_zy@163.com)
†Marketing Service Center of State Grid Fujian Electric Power Co., Ltd., Fuzhou, Fujian, 350013, China
‡Marketing Service Center of State Grid Fujian Electric Power Co., Ltd., Fuzhou, Fujian, 350013, China
§Marketing Service Center of State Grid Fujian Electric Power Co., Ltd., Fuzhou, Fujian, 350013, China
¶Marketing Service Center of State Grid Fujian Electric Power Co., Ltd., Fuzhou, Fujian, 350013, China

process. Such problems require a long period of time to be discovered, which can have a long-term impact on the accuracy of energy measurement.

As the core infrastructure equipment of ubiquitous power Internet of Things, the demand for smart energy meters is constantly increasing. In terms of electricity metering, the accuracy of smart energy meters is the core element of fairness in electricity trade; In the field of big data research, accurate and qualified smart energy meters are one of the foundations for building ubiquitous power Internet of Things [4]. As of the end of 2020, more than 300 million electricity meters have been put into operation in China, and this number will continue to grow in the future. Traditional manual calibration methods cannot meet the growing demand for calibration, and the transformation from manual calibration to automated systems is imperative. As a result, intelligent electricity meter automated calibration assembly lines have emerged. During the long-term operation of automated calibration assembly lines, frequent connection of smart energy meters to meter positions can cause deformation of the mechanical crimping terminals at the meter positions; Long term live operation can accelerate the oxidation rate of the surface material of mechanical crimping terminals, leading to terminal corrosion [5]. The deformation and corrosion of the mechanical crimping process of the meter will directly affect the reliability of the error test results, thereby affecting the calibration quality of the smart energy meter. At present, the provincial metrology centers under the State Grid Corporation of China generally adopt the method of regular verification to inspect and repair the meter positions on the calibration assembly line. This method cannot detect meter position faults in a timely manner and relies on manual troubleshooting. Its reliability is insufficient and labor costs are high. Therefore, achieving online anomaly detection of calibration assembly line meter positions is of great significance [6].

The abnormal occurrence of the meter flying away is relatively accidental, and generally it will not occur repeatedly on the same meter; Due to external factors such as communication interference, there is a certain amount of data noise in the electricity consumption data. Therefore, using the general analysis method of threshold judgment can cause a large number of misjudgments of runaway anomalies, which affects the discovery and handling of actual meter runaway faults. The LOF algorithm is a classic density based time series anomaly diagnosis algorithm. The author used this algorithm to establish an intelligent diagnostic analysis method for the flying away of electric energy meters, which can effectively remove the influence of outliers, improve the accuracy and timeliness of detecting and judging the flying away anomalies of electric energy meters.

2. Construction of an accuracy analysis platform for electricity metering.

2.1. System Design. The author uses Hadoop distributed technology to build a massive data storage and high-performance parallel computing cluster that covers various power supply bureaus in the province to cope with large-scale data parallel processing. The author also uses Nigera load balancing technology and Redis as a caching component to improve throughput and system availability, meeting the needs of high concurrency requests [7]. Due to the need to process real-time collection of massive power grid measurement data, the system adopts a distributed architecture based on Hadoop and applies the HDFS distributed file system, from the monitoring terminal of measurement standard devices, laboratory calibration control system, measurement automation system, marketing system, on-site inspection business system, semi-structured or unstructured data, then it is stored in the non relational database HbaSe, and some of the data is cleaned and analyzed offline through the Hive data warehouse. The overall architecture of the accuracy analysis platform for electricity metering is shown in Figure 2.1.

2.2. System Software Architecture. In order to meet the requirements of high concurrency, high reliability, and system scalability for data collection, the system is planned to be constructed in MVC mode. The data collection part adopts Nginx+Tomcat cluster to achieve high concurrency and high reliability of web servers, and the backend uses queues+process pools to achieve multi-channel processing of various protocols and concurrent links; The system adopts a B/S architecture based on J2EE for construction. The core technology framework of the platform adopts JAVA as the development language, based on mainstream open-source J2EE frameworks, including Struts, Spring, Hibernate, Jquery, JBOSSSOA, JBPM, Druid, and other frameworks. Capable of supporting multiple heterogeneous databases and compatible with mainstream web containers. The technical structure is divided into four layers: basic environment, DAO layer, logic layer, and presentation layer [8].

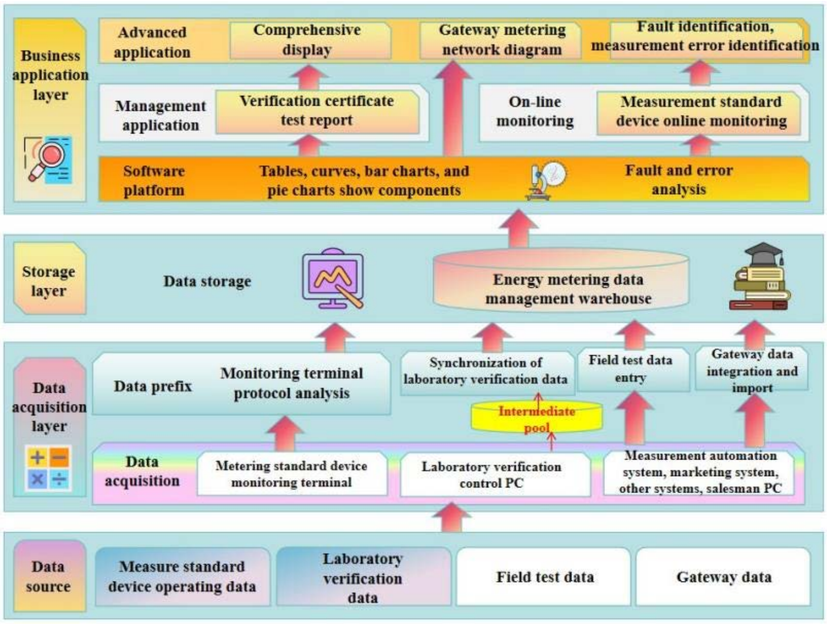

Fig. 2.1: Overall architecture of the accuracy analysis platform for electric energy metering

2.3. System Hardware Architecture. Establish a dedicated network for the metering center on the internal power network and install encrypted communication. VPN gateway achieves remote access by encrypting data packets and converting the target address of the data packets. Implement VPN proxy servers through various methods such as servers, hardware, and software. By using virtual private networks, the measurement center network is made more humanized, software oriented, and intelligent, providing secure, controllable, and flexible resource scheduling capabilities to meet the dynamic restructuring needs of power metering data communication. Addressing network security and monitoring, MAC address tracking, and security vulnerabilities in virtual machine management programs [9].

3. LOF anomaly detection algorithm.

3.1. Abnormal point detection. Outlier detection (also known as outlier detection) is an important part of data mining and refers to the process of identifying objects whose behavior is significantly different from what is expected. These items are called outliers. Vulnerability detection is critical to many applications, including healthcare, public safety, fault detection, image processing, sensor/video networks, and testing see. Misdiagnosis can be divided into controlled and uncontrolled [10]. If the analyst can find registered original and unusual items, they can be used to create an abnormality detection sample before using control methods to detect them. In some applications, unsupervised learning techniques are used when there are no objects labeled as "normal" or "abnormal". Undoubtedly detection assumes that the product contains some kind of impurity.

3.2. Method Comparison. Due to the different electricity consumption patterns of different users, it is not possible to use a unified method for state labeling of energy representation values, and it is not possible to provide a learning set of supervised methods. Therefore, unsupervised anomaly detection algorithms are more suitable for the analysis of energy representation value anomalies. In unsupervised methods, statistical, distance, and density based methods are currently the main methods for anomaly detection [11].

The statistical anomaly detection method mainly analyzes the dispersion of data, analyzes the distribution of data, extracts the variation indicators of data, commonly used include standard deviation, interquartile spacing, etc., and extracts outliers through the variation indicators. This type of algorithm requires prior knowledge of the distribution characteristics of the data, as well as parameter determination of mutation indicators. The algorithm has poor universality and is mainly used in scientific research fields.

The distance based anomaly detection method considers the neighborhood of an object with a given radius. If it is an anomaly, there are not enough other points in its neighborhood. Compared to statistical algorithms, it does not require users to have any domain knowledge and is more intuitive in concept. However, due to the algorithm detecting outliers from a global perspective by calculating the distance between objects, the detection effect is not good when there are multiple distributions or subsets with different densities in the dataset [12]. Density based detection methods mainly examine the density of the object and its neighbors. If its density is much lower than its neighbors, it is considered an outlier. In this type of algorithm, each point will calculate an outlier degree, overcoming detection errors caused by mixing different density subsets, and the detection accuracy is relatively high. The author mainly adopts density based anomaly detection methods.

3.3. LOF algorithm. The key of density detection is to compare the density around an object with the surrounding density. There is a significant difference between the density of non-outlier objects and the surrounding neighborhood, and there is a significant difference between them and their surroundings.In this paper, a Local Outlier Factor (LOF) is proposed based on the combination of density and anomaly detection. The algorithm is defined as follows:

For a given set of objects D, the objects contained in D are denoted as o. The k-distance of object o is denoted as $dist_k(o)$, which is the distance dist(o, p) between o and another object $p \in D$, such that:

*At least k objects $o \in D - \{o\}$, such that dist $(o, o) \geqslant dist(o, p)$

*At least k-1 objects $o \in D - \{o\}$, such that $dist(o, o) < dist(o, p)$

$dist_k(o)$ is the distance between o and its k-th nearest neighbor. Therefore, the k-distance neighborhood of o includes all objects whose distance to o is not greater than $dist_k(o)$, denoted as:

$$N_o = \{o' | dist(o, o') dist(o)\} \tag{3.1}$$

If the average distance between objects in $N_k(o)$ and o is used as the local density measure of o, a problem arises. When a very close neighbor o is encountered in o, causing dist (o, o') to be very small, the statistical fluctuation of the distance measure will be unexpectedly high. In order to solve this problem, a smoothing effect can be added to convert it into the following reachable distance[13].

$$reach - dist(o, o') = MAX\{k - distance(o), d(o, o)\} \tag{3.2}$$

The local reachable density of object o is the reciprocal of the average reachable distance between object o and its MinPts neighborhood.

$$lrd_{MinPts}(o) = \frac{1}{\frac{\sum reach-distMinPts(o', o)}{|N_{MinPts}(o)|}} \tag{3.3}$$

The local anomaly factor of object o is defined as:

$$LOF_{MinPts}(o) = \frac{\sum_0 \in N_{MinPts}(o) \frac{lrd_{MinPts}(o')}{lrd_{MinPts}(o)}}{|N_{MinPts}(o)|} \tag{3.4}$$

The degree of anomaly in object o can be evaluated through its local anomaly factors; The anomaly factor is directly proportional to the degree of anomaly of the object. If the factor value is larger, the likelihood of anomaly increases; If the factor value is smaller, the likelihood of anomalies decreases, and the LOF factor level of normal values is generally within 1.

4. Example Analysis.

4.1. Calculation steps. In the analysis of electricity meter runaway, we can use the LOF algorithm to filter the original data, effectively identify abnormal points in electricity meter readings, and eliminate the impact of abnormal points in electricity meter readings on the judgment of electricity meter runaway. At the same time, due to the introduction of table readings before and after abnormal points in the judgment, it is necessary to improve the threshold judgment. The author mainly judges based on the daily maximum electricity consumption and daily average electricity consumption [14].

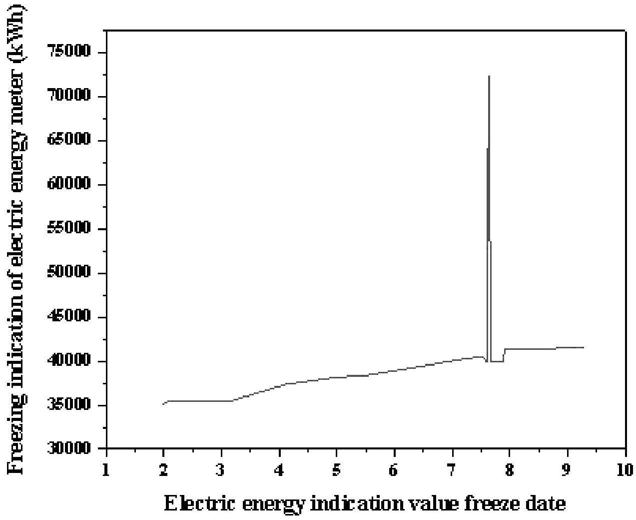

Fig. 4.1: Daily freezing curve of electric energy representation (accidental mutation)

Table 4.1: Daily freezing readings and corresponding LOF factors (accidental mutation) for the three days before & after the electricity meter trip

Date	July 18th	July 20th	July 21st	July 22nd	July 23rd	July 25th	July 26th
Indicating value	40647	39908	39908	72300	39908	39908	39908
LOF	1.501447	0.984361	0.984361	6041.262	0.984361	0.984361	0.984361

(1) Preliminary judgment of abnormal flight according to the flight formula of the electric energy meter;
(2) Extract suspected daily data of the flight, generally including at least 7 days before and after the flight date;
(3) Extract daily data from one of the meters, organize the data and sort the dates;
(4) Calculate the LOF factor and count the number of abnormal factors;
(5) Remove the daily data corresponding to LOF abnormal factors;
(6) Calculate daily electricity consumption;
(7) Calculate the improvement N value.

$$N = \frac{\text{Maximum daily electricity consumption}}{\text{Average daily electricity consumption}} \tag{4.1}$$

4.2. Calculation results. Perform LOF factor calculation on the energy meter in Figure 4.1, and the calculation result is shown in Figure 4.2. It can be observed that the position of the mutation point on the electric energy representation curve is the same as that on the LOF curve [15,16]. We extracted the daily data of the mutation point and the three days before and after, as shown in Table 4.1. It can be found that the LOF of the mutation point is 6041.262, while the LOF of the other points is basically around 1. The outlier can be filtered out. After elimination, calculate the daily electricity consumption and calculate the N value for flight judgment.

The LOF factor calculation was performed on the electricity meter, and the results are shown in Figure 4.3. It can be found that the LOF curve data range is mainly concentrated in the range of 0.8 to 1.3, and there is no significant change in the LOF factor near the mutation point, as shown in Table 4.2 [17,18,19,20].

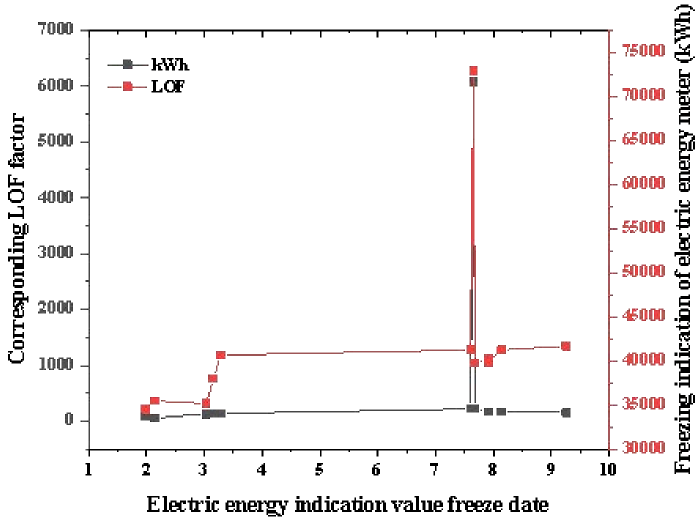

Fig. 4.2: Comparison of Daily Freezing Curve and LOF Curve for Energy Representation (Accidental Sudden Change)

Table 4.2: Daily freezing indication and corresponding LOF factor (indication flying away) for the three days before and after the jump of the electric energy meter

date	June 9th	June 10th	June 11th	June 12th	June 13th	June 14th	June 15th
indicating value	1913	1919	119614	119619	119627	119632	119638
LOF	1.0772	1.1393	1.1476	1.1077	1.0639	1.0495	1.0411

5. Conclusion. Through the above algorithms, it is possible to achieve a secondary judgment of the abnormal flight of the electric energy meter detected by the monitoring of the electricity consumption information collection system, which greatly improves the quality and efficiency of the judgment of the electric energy meter flight. It can timely detect the abnormal flight of the electric energy meter, effectively avoid fault misjudgment, and improve the efficiency of on-site fault handling; This method can be further extended to the analysis of other topics related to electricity consumption information collection, improving the accuracy of diagnosis and analysis of the operating conditions of measuring equipment, and further supporting the marketing business decision-making and implementation of power supply enterprises.

Meanwhile, the above is only a practical method for handling data outliers. With the deepening application of electricity information collection system data, how to handle anomalies in data will be the first problem that needs to be solved in future business applications. The rapid development of technologies such as statistical analysis, data mining, and machine learning has laid a solid theoretical foundation for solving the above problems. We need to further increase the introduction, understanding, and practice of data science in the power business.

REFERENCES

[1] Douiba, M., Benkirane, S., Guezzaz, A., & Azrour, M. (2022). Anomaly detection model based on gradient boosting and decision tree for iot environments security. Journal of Reliable Intelligent Environments, 85(7),1-12.

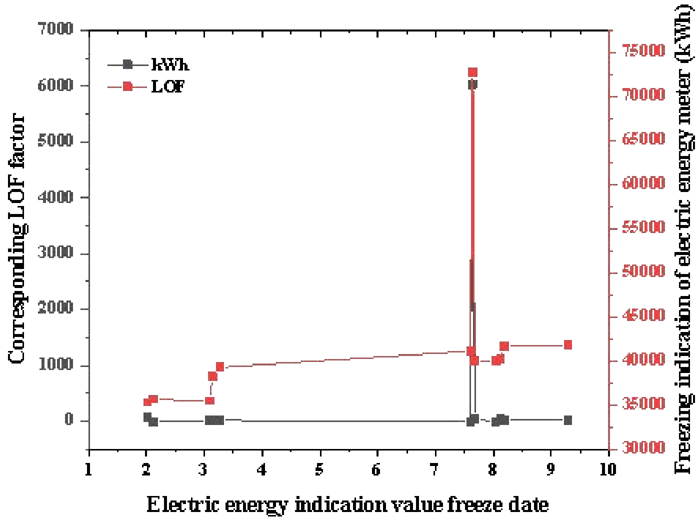

Fig. 4.3: Daily Freezing Curve of Electric Energy Representation Value (Display Value Flying)

[2] Shu, X., Zhang, S., Li, Y., & Chen, M. (2022). An anomaly detection method based on random convolutional kernel and isolation forest for equipment state monitoring. Maintenance and Reliability, 11(2), 72.

[3] Petrillo, A., Murino, T., Piccirillo, G., Santini, S., & Caiazzo, B. (2023). An iot-based and cloud-assisted ai-driven monitoring platform for smart manufacturing: design architecture and experimental validation. Journal of Manufacturing Technology Management, 34(4), 507-534.

[4] Lee, K. H., Lee, J. E., Seo, J. Y., & Kim, T. H. (2023). Health monitoring and anomaly detection of motor-driven equipment based on the tft model and vibration signal decomposition. The Journal of Korean Institute of Information Technology, 16(7), 203.

[5] Jang, Y. M. (2022). Real-time energy data acquisition, anomaly detection, and monitoring system: implementation of a secured, robust, and integrated global iiot infrastructure with edge and cloud ai. Sensors, 46(2), 1864-1881.

[6] Tang, M., Chen, W., & Yang, W. (2022). Anomaly detection of industrial state quantity time-series data based on correlation and long short-term memory. Connection Science, 34(1), 2048-2065.

[7] Kumar, Y. R. S., & Champa, H. (2022). Anomaly detection framework for efficient sensing in healthcare iot systems. 2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS),65(7), 81-85.

[8] Nayak, J., Naik, B., Dash, P. B., Vimal, S., & Kadry, S. (2022). Hybrid bayesian optimization hypertuned catboost approach for malicious access and anomaly detection in iot nomalyframework. Sustainable computing: Informatics and systems, 46(4), 4362-4379.

[9] Hu, K., Cai, Y., Cai, Z., Li, X., Cen, B., & Chen, Y. (2022). Fault location method based on structure-preserving state estimation for distribution networks. IET generation, transmission & distribution,74(15), 16.

[10] Omonov, A., Fitriyah, A., Kato, T., & Kawabata, Y. (2022). Improving monitoring technologies for the salinity of irrigated lands based on gis/rs and satellite imagery for the conditions of the syrdarya region, uzbekistan. Journal of Arid Land Studies, 32(3), 96-96.

[11] Zhou, Y., Liu, C., Yu, X., Liu, B., & Quan, Y. (2022). Tool wear mechanism, monitoring and remaining useful life (rul) technology based on big data: a review. SN Applied Sciences,96(8), 4.

[12] Kong, J., Jiang, W., Tian, Q., Jiang, M., & Liu, T. (2023). Anomaly detection based on joint spatio-temporal learning for building electricity consumption. Applied energy, 46(3), 2418-2437.

[13] Ayalew, L. G., Mattihalli, C., & Asmare, F. M. (2022). Wirelessly controlled plant health monitoring and medicate system based on iot technology. Communications in Computer and Information Science.

[14] Liu, W., Lei, S., Peng, L., Feng, J., Pan, S., & Gao, M. (2022). Active anomaly detection technology based on ensemble learning. Springer, Singapore, 20(1), 25.

[15] Jérémy Renaud, Karam, R., Salomon, M., & Couturier, R. (2023). Deep learning and gradient boosting for urban environmental noise monitoring in smart cities. Expert Systems with Applications, 218(32), 119568.

[16] Wang, N., Zhang, G., Ren, L., Pang, W., & Li, Y. (2022). Correction to: novel monitoring method for belt wear state based on machine vision and image processing under grinding parameter variation. The International Journal of Advanced Manufacturing Technology, 122(1), 103-104.

[17] Martins, I., Resende, J. S., Sousa, P. R., Silva, S., Antunes, L., & Gama, J. (2022). Host-based ids: a review and open issues

of an anomaly detection system in iot. Future generations computer systems: FGCS,96(133), 133.
[18] Mukherjee, I., Sahu, N. K., & Sahana, S. K. (2023). Simulation and modeling for anomaly detection in iot network using machine learning. International journal of wireless information networks, 47(3), 557-564.
[19] Hou, Y., Fu, Y., Guo, J., Xu, J., Liu, R., & Xiang, X. (2022). Hybrid intrusion detection model based on a designed autoencoder. Journal of Ambient Intelligence and Humanized Computing, 14(8), 10799-10809.
[20] Mukherjee, I., Sahu, N. K., & Sahana, S. K. (2022). Simulation and modeling for anomaly detection in iot network using machine learning. International Journal of Wireless Information Networks, 30(2), 173-189.

Edited by: Zhigao Zheng
Special issue on: Graph Powered Big Aerospace Data Processing
Received: Jan 17, 2024
Accepted: Mar 5, 2024