Proceedings of CECNet 2022
A.J. Tallón-Ballesteros (Ed.)

doi:10.3233/FAIA220547

Efficient Security Detection Strategy Based on K-Means Algorithm in LPWAN

Minhao JIN[a,1] , Shaojie LUO[a], Jiabin HUANG[a] and Xincheng ZHOU [b]
[a] *State Grid Zhejiang Electric Power Company, China.*
[b] *Beijing University of Posts and Telecommunications, China.*

Abstract. With the development of multi-element integration and high elastic power grid construction, digital sensing of transmission, transformation and distribution equipment in many new energy sources, virtual power plants and load aggregators will be widely used. In this environment, it is urgent to build a reliable sensing network to support its digital sensing applications such as operation state sensing, data collection and demand management. However, due to the large number and wide distribution of distribution equipment, there may be errors in the uploading results of sensors. This paper proposes an efficient security detection strategy based on K-Means algorithm, which greatly improves the information integration and decision ability of the data center.

Keywords. Digital perception, efficient, security, LPWAN

1. Introduction

As the traditional transmission, transformation and distribution equipment is designed and deployed with less consideration of upgrading and transformation needs, problems such as inability to supply power for new modules, complex electromagnetic environment and limited internal space of equipment appear in the process of transformation and upgrading. As pointed out in the vision reference model IEEE smart grid control vision: reference model for 2030 and beyond, in order to realize this vision, research needs to be conducted to address all these challenges[1]. At present, there is an urgent need for a communication means with extremely low power consumption, strong anti-interference ability and flexible deployment to transform the traditional transmission, transformation and distribution equipment and help improve the power system.

Wireless sensor network has been widely used in recent years. It can be applied to military[2], environmental monitoring and forecasting[3], building state monitoring [4]and other fields. At the same time, the new generation of living smart home system developed based on smart home wireless technology makes the green home network gradually form a complex system to handle various tasks[5]. Scientific research using wireless sensor networks for habitat or animal monitoring is used to generate a large amount of data that needs to be analyzed and normalized to help researchers and other people interested in information. Managing these data and comparing it with information from other sources and systems can help the decision-making process and trigger other

[1] Corresponding author: Jin Minhao, State Grid Zhejiang Electric Power Company; E-mail: zhouxc@bupt.edu.cn

experiments again[6]. A large number of wireless sensor networks are also used in the fields of warehouse management, security monitoring of airports and large industrial parks[7].

LPWAN (low power wide area network) with characteristics of low bandwidth, low power consumption, long distance and large capacity meets the requirements of Internet of things applications[8]. The two major LPWAN technology representatives, NB-IoT and LoRa, are better known. Developing low-power wide area network (LPWAN) solutions for efficient adoption, deployment and maintenance is crucial for smart cities[9][10].

In [11], the author reviewed the rapid growth of research subjects using different graph based deep learning models (such as graph convolution and graph attention networks) in various problems of different types of communication networks (such as wireless networks, wired networks and software defined networks). In [12], the author proposes two new graph based intrusion detection solutions, which integrate the remaining learning into GNN using the available graph information. In[13], the author proposes a novel hierarchical adversarial attack (HAA) generation method for GNN based intrusion detection in IoT system. An E-GraphSAGE algorithm is proposed for network intrusion detection of the Internet of Things in[14].

This paper studies the efficient and accurate information collection of power transmission and distribution equipment nodes based on LPWAN, and proposes an efficient security detection strategy called ESDS based on K-Means algorithm, which greatly improves the information integration and decision ability of the data center.

The chapter II of this paper briefly describes the network model of the proposed algorithm strategy application, the chapter III describes the details of the algorithm in detail, the chapter IV shows the simulation results of the algorithm, and the chapter V is the summary of this paper.

2. System Model

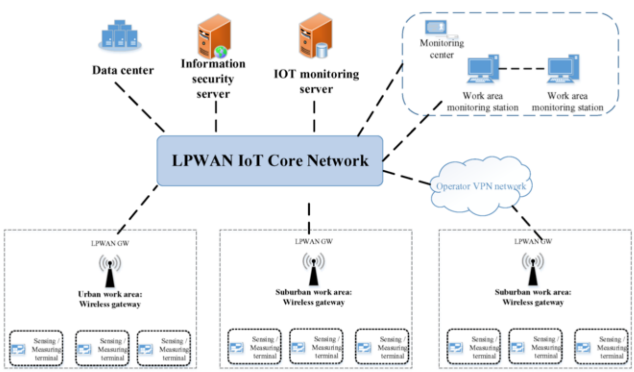

Figure 1. Schematic diagram of power Internet of things

Figure 1. is a basic schematic diagram of a city level power IoT network, which divides different production work areas, deploys sensor network terminals respectively, and plans technical coverage and human management according to the coverage radius of the wireless gateway and the management area of the work area. The LPWAN wireless gateway is connected to the LPWAN IoT core network through the internal transmission

network of the power grid. In some remote areas, such as covering the edge suburbs, where the internal transmission network cannot reach, the VPN network of the operator can be connected to the LPWAN control network through 3G/4G wireless VPN tunnel technology. And establish the IoT monitoring server to manage the data collected by the IoT, and the information security server to manage all the data information to ensure the information security of the monitoring network. The monitoring center accesses the control network through the internal network to obtain the concerned equipment state information and environment state information in real time.

The power system digital sensing network is composed of a large number of wireless sensor nodes deployed in the monitoring area. It is a multi hop self-organizing network system formed by wireless communication. Its purpose is to monitor, sense and collect the information of various sensing objects in the network coverage area and process the information. Because there are many nodes in the wireless sensor network and the setting range is wide, in the power grid application, the data transmission needs to ensure the accuracy of the data, and at the same time, considering the impact of security, the data judgment center is designed at the central node to comprehensively judge the data uploaded by each node, so as to eliminate the sensor nodes that have failed or been attacked.

3. Data Fusion Algorithm Based on K-Means Introduction

Because there are many nodes in the power system digital sensor network, and the setting range is wide, and it is necessary to maintain the long-term stable operation of the system. At the same time, in power grid applications, the data transmission equipment is required to work normally in extreme weather environments such as cold or hot weather, so the accuracy of the new power system digital sensor network is extremely high, especially in the case of equipment overheating, In order to ensure the efficiency and accuracy of the system, the system classifies and eliminates the sensor terminals that have failed and have low accuracy in the process of information collection and transmission.

3.1. Reputation value establishment

Firstly, the detection level of each sensor node is obtained by comparing the results uploaded by each sensor node with the actual results, and the beta reputation system is used to evaluate it. Beta reputation system allocates reputation to members through beta distribution function. Use x to indicate that the sensor has correctly transmitted the collected information, y to indicate that the sensor has not successfully collected the information or transmitted the wrong collection results, let r denote the number of times of event x and s denote the number of times of event y, then the probability of occurrence of x in the two events follows the beta distribution, and the probability density function is:

$$f(p \mid \alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1} \tag{1}$$

where, $\Gamma(\cdot)$ is the gamma function, $\alpha=r+1$, $\beta=s+1$ in the beta distribution, and the expectation of event x is

$$E[x]=\frac{\alpha}{\alpha+\beta} \tag{2}$$

3.2. The implementation of K-Means algorithm

The historical information uploaded by each sensor node is summarized and counted through the beta reputation system, and then the K-Means clustering algorithm is used to classify each sensor node, so as to eliminate the sensor nodes that have failed and have low accuracy. Figure 2. shows the implementation steps of K-Means algorithm.

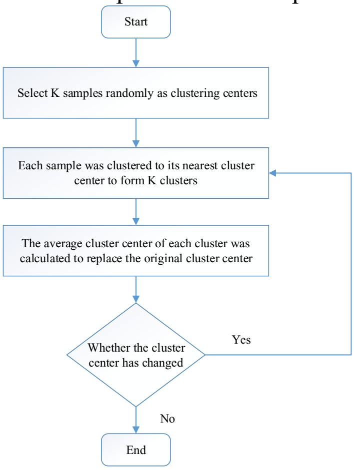

Figure 2. The implementation steps of K-Means algorithm

The specific description of the above steps is as follows:

Step 1: For spatial data set $X=\{X_1,...,X_i,...,X_n\}$, k values $C_1,C_2,...,C_j,...,C_k$ are randomly selected as the initial clustering center, and the data shall meet the following conditions:

$$X=\cup C_i(i=1,2,...,k) \tag{3}$$

$$C_i\neq\varnothing(i=1,2,...,k) \tag{4}$$

$$C_i\cap C_j\neq\varnothing(i=1,2,...,k;i\neq j) \tag{5}$$

Step 2: Calculate the Euclidean distance from sample X_i to each cluster center C cluster each sample to its nearest cluster center C_i, and form k clusters.

$$D_i=\min_{i<=j<=k}\{\sum_{i}^{n}(x_i-c_j)^2\} \tag{6}$$

Step 3 : Recalculate the average cluster center of K clusters, and replace the original cluster center with the new average cluster center.

$$C_j=\frac{1}{N(C_j)}\sum_{i\in C_j}X_i \tag{7}$$

Step 4 : Repeat steps 2 and 3 until the clustering center is unchanged or basically unchanged.

4. Simulation Result

In this paper, MATLAB software is used to simulate and design an ad hoc network with 10 wireless sensor nodes. Among them, 8 sensor nodes can still be used normally, but 2 sensor nodes cannot collect information. Due to the damage and transmission environment, the collected information transmitted by 2 sensor nodes to the data center is wrong. The data center collects the information uploaded by all sensor nodes, performs data fusion, compares it with the actual state of each environment, obtains the final detection probability, and evaluates the working state and detection level of each sensor. The simulation diagram is as follows:

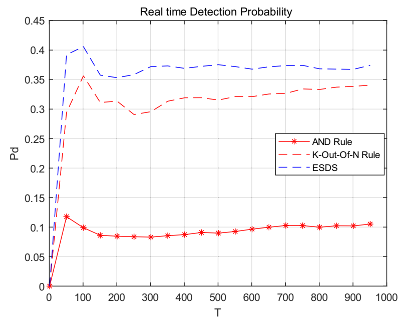

Figure 3. Real time Detection Probability Simulation Result(10 nodes)

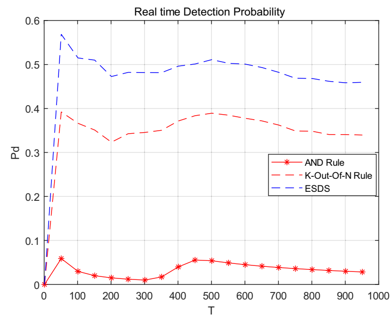

Figure 4. Real time Detection Probability Simulation Result(15 nodes

In Figure 3., due to the existence of sensing nodes that fail and upload erroneous data, when the traditional "AND" criterion is used to fuse the statistical data, the accuracy is low, and the detection probability is only 10%. When using the "K-Out-Of-N" criterion to judge in the data center, the detection level of the system can be greatly improved. The detection algorithm using K-Means algorithm for classification proposed in this paper is improved by 10% on the basis of "K-Out-Of-N" criterion algorithm. This is because before the data judgment, the K-Means algorithm is used to classify the failed and low accuracy perceptual nodes and normal perceptual nodes by summarizing and classifying the historical collected information of each perceptual node, and different judgment weights are given in the final data fusion, so as to achieve efficient and accurate digital perception.

It can be found from the simulation in Figure 4. that when the number of sensor nodes in the core network reaches 15, the complexity of information to be processed increases due to the large number of loaded nodes, the difficulty of data center decision increases. At this time, the algorithm proposed in this paper still has advantages in the final data decision.

5. Conclusion

In this paper, the K-Means algorithm is used to eliminate the sensor nodes that have failed or have low accuracy, thus increasing the decision level of the final data center. By using the digital sensing network model proposed in this paper, the operation state and information collection of transmission, transformation and distribution equipment in the power system can be effectively and accurately sensed, so as to improve the intelligent level of the power system.

Acknowledgment

This work was supported by the project of State Grid Zhejiang Electric Power Company (B311HZ220001).

References

[1] A. Annaswamy, "IEEE Vision for Smart Grid Control: 2030 and Beyond Roadmap," in IEEE Vision for Smart Grid Control: 2030 and Beyond Roadmap , vol., no., pp.1-12, 24 Oct. 2013, doi: 10.1109/IEEESTD.2013.6648362.
[2] Lixianli, P. Wei, A. Jianyong and W. Ping, "The Application Research on Military Internet of Things," 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2020, pp. 187-191, doi: 10.1109/ICCWAMTIP51612.2020.9317321.
[3] F. Wu, C. Rüdiger, J. Redouté and M. R. Yuce, "Live Demonstration: An IoT Platform for Environmental Monitoring Using Self-Powered Sensors," 2018 IEEE SENSORS, 2018, pp. 1-1, doi: 10.1109/ICSENS.2018.8589904.
[4] S. Sarkar, K. U. Rao, J. Bhargav, S. Sheshaprasad and A. Sharma C.A., "IoT Based Wireless Sensor Network (WSN) for Condition Monitoring of Low Power Rooftop PV Panels," 2019 IEEE 4th International Conference on Condition Assessment Techniques in Electrical Systems (CATCON), 2019, pp. 1-5, doi: 10.1109/CATCON47128.2019.CN004.
[5] H. Wei-Dong and Z. Bo-Xuan, "Smart Home Wireless System Using ZigBee and IEEE802.15.4," 2016 Sixth International Conference on Instrumentation & Measurement, Computer, Communication and Control (IMCCC), 2016, pp. 858-863, doi: 10.1109/IMCCC.2016.168.
[6] R. A. G. da Costa and C. E. Cugnasca, "Use of Data Warehouse to Manage Data from Wireless Sensors Networks That Monitor Pollinators," 2010 Eleventh International Conference on Mobile Data Management, 2010, pp. 402-406, doi: 10.1109/MDM.2010.72.
[7] Wei Xia, S. X. Yang, Xiaofang Li, Huibing Wang, Xijun Yan and Xiaodong Wei, "Design of wireless sensor networks for safety monitoring at civil engineering construction sites," 2011 Second International Conference on Mechanic Automation and Control Engineering, 2011, pp. 6418-6421, doi: 10.1109/MACE.2011.5988511.
[8] S. Y. Mane, "LPWAN's – Overview, Market Scenario and Performance Analysis of Lora, Sigfox Using NB-Fi Range Calculator," 2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), 2021, pp. 1-4, doi: 10.1109/SMARTGENCON51891.2021.9645902.
[9] A. Ramoutar, Z. Motamedi and M. Abdulla, "Latency of Concatenating Unlicensed LPWAN with Cellular IoT: An Experimental QoE Study," 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall), 2021, pp. 1-6, doi: 10.1109/VTC2021-Fall52928.2021.9625483.
[10] M. L. Liya and D. Arjun, "A Survey of LPWAN Technology in Agricultural Field," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020, pp. 313-317, doi: 10.1109/I-SMAC49090.2020.9243410.
[11] Jiang, W. (2021). Graph-based deep learning for communication networks: A survey. Computer Communications.
[12] Chang, L., & Branco, P. (2021). Graph-based Solutions with Residuals for Intrusion Detection: the Modified E-GraphSAGE and E-ResGAT Algorithms. arXiv preprint arXiv:2111.13597.
[13] Zhou, X., Liang, W., Li, W., Yan, K., Shimizu, S., Kevin, I., & Wang, K. (2021). Hierarchical adversarial attacks against graph neural network based IoT network intrusion detection system. IEEE Internet of Things Journal.
[14] Lo, W. W., Layeghy, S., Sarhan, M., Gallagher, M., & Portmann, M. (2022, April). E-GraphSAGE: A Graph Neural Network based Intrusion Detection System for IoT. In NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium (pp. 1-9). IEEE.