Mining Network Security Holes Based on Data Flow Analysis in Smart Grid

Yang Li[1(✉)], Xiaohua Liu[2], Lixin Zhang[2], Wenbin Guo[2], and Qian Guo[3]

[1] State Grid Xinjiang Electric Power Research Institute, Xinjiang, China
455508995@qq.com
[2] State Grid Xinjiang Electric Power Co., Ltd., Xinjiang, China
{liuxiaohua,zhanglixin,guowenbin}@xj.sgcc.com.cn
[3] Global Energy Internet Research Institute Ltd., Nanjing, China
guoqian@geiri.sgcc.com.cn

Abstract. With the popularity of mobile terminals and the sharp increase in network data traffic, the problem of security loopholes has become increasingly prominent. The traditional vulnerability detection methods can no longer meet the demands for detection efficiency. In order to satisfy the high requirements on network security in the era of big data, the vulnerability mining technology is extremely urgent. This paper describes the current situation and introduces relevant security technology and algorithm in smart grid. The decision tree algorithm is selected as the basic algorithm of big data security technology. Through the test, the missing alarm rate and false alarm rate are simulated experimentally. We obtain the results of experiments by controlling variables, which proves that our algorithm can effectively detect IP scanning, Port scanning and other attacks.

Keywords: Big data · Data analysis · Vulnerability detection · Smart grid

1 Introduction

With the popularity of mobile terminals, the network data traffic has increased dramatically, information transmission rate is faster, and security vulnerabilities have become increasingly prominent. In order to satisfy the needs of network security in the era of big data and the requirements of in-depth mining of security vulnerabilities, it is necessary to design a system model combined with big data analysis in smart grid.

This article mainly describes vulnerability mining and detection technology, including three main points: common types of vulnerabilities, exploits and classification of exploits [1]. It also includes big data security analysis technologies. In addition, this paper also introduces the relevant algorithm in detail. During the use of the system, it is inevitable to encounter related attacks, and the security

Supported by State Grid Xinjiang Electric Power Co., Ltd.

D.-J. Deng et al. (Eds.): SGIoT 2019, LNICST 324, pp. 57–71, 2020.
https://doi.org/10.1007/978-3-030-49610-4_5

scanning is a method to simulate it. The security scanning is tested with the data set, and the system tested by this method has a certain defense performance, which has been able to preliminarily complete the relevant security work.

If the source code of the software can be known, we can use this method. By scanning the source code, we find the existence of security vulnerabilities, which can be targeted at the vulnerability of the source code modification. Source code scanning is a significant part of the programmer's job. However, a system code may be too large, manual operation may have considerable difficulty. The main contributions of this paper are listed as follows.

(1) We show the detailed description of the big data security analysis platform, as well as related algorithms, such as decision tree, and the relevant algorithm simulation test.
(2) We test the accuracy of the decision tree and a series of simulation of it.
(3) We present a bold conjecture of the future development of the related technologies mentioned in this paper and improvements to the subsequent related work of the algorithm are proposed.

The remainder of this paper is organized as follows. The Sect. 2 briefly introduces the related technologies involved in the main content of this paper. The Sect. 3 explains the relevant knowledge and algorithms of the big data security analysis platform. The Sect. 4 introduces the system model in detail. The Sect. 5 explains the design idea of the model. The Sect. 6 introduces the function of decision tree in mining security vulnerability based on data flow. The Sect. 7 analyzes the accuracy of the decision tree and carry out a series of simulation tests for it. The Sect. 8 concludes this work and make plans for the future.

2 Related Work

The process of exploiting vulnerability is to exploit the vulnerability in the computer system. Attackers use such vulnerabilities to achieve the desired purpose. The process and classification of vulnerability exploitation are described in detail.

In order to discover network vulnerabilities, attackers can use vulnerability information to obtain website user information and gain the entire website [2]. The basic process of exploitation is as follows: Scan the target site; Scan the target website with professional vulnerability detection software; Scan for known vulnerabilities [3].

The main detection content includes the system version type, relevant data services, and the system external port Settings. The obtained scan results are analyzed, which find the location of the vulnerability and its use.

Classification of common exploitation is shown as follows:

SQL Injection: A common approach is to inject SQL commands into the relevant pages. In this way, the original instruction can't be executed and the injected instruction can be run, thus entering into the vulnerability of the database, destroying the content of the database and the structure of the database [4].

XSS Cross-Site Scripting Attack: Similar to SQL injection. It is a method of injecting HTML. XSS can modify web pages and implant malicious scripts. The altered web page controls the browser with HTML statements when the user is using it [5]. The attacker can use the modified browser to steal the data stream related to the session, and through the data stream analysis, obtain the user's account password and change the appearance of the page.

Cross Site Request Forgery CSRF: The attacker interferes with the browser by forgery, so that the browser considers the intrusion site to be the previous authentication site, and the intrusion site obtains the corresponding operation authority. Authentication requests can only be made by the browser, but they may not be made by the user himself, and may be forged browser requests [6].

Click Hijacking: This is an attack method that uses the limitation of human senses to achieve the purpose. The main approach is to overlay the web page with an invisible HTML frame called an iframe. Since the iframe is not visible, the user mistakenly thinks it is a normal website for operation. In the operation process, the iframe will obtain the user's relevant account and password data through the operation of identifying the user [7].

Any File Upload Vulnerability: File upload vulnerability is mainly caused by the system is not strict screening. When uploading a file, the attributes of the file are not specified in the code [8]. An attacker can access any file on the web by changing its properties.

3 Preliminaries

In this section, we will present some preliminaries used in big data security analysis.

3.1 Big Data Security Analysis Technology

Big data analysis technology has gradually entered the field of security. We can use the big data analysis platform to collect data. In addition, we can use it for further analysis. By collecting and following up, we build a complete set of patterns [9]. This mode can effectively solve related security problems and maintain network security. Big data security analysis can also be combined with a variety of other technologies such as artificial intelligence [10].

In order to enable the platform to detect security issues more efficiently [11]. The shortcoming of the traditional detection method is solved and we need to continue to innovate better detection technology. We can adopt the related technology of artificial intelligence to improve such defects.

It is obviously to see from this diagram that the platform is an evolving hierarchy and the data flow is a gesture of directional movement. At the same time, it has feedback and adjustment in the platform. This also makes our platform more adaptable to some other scenarios, which means that the platform is more resistant and has a wider scope. In addition, big data security analysis platform

needs to do some conditions to be able to be an excellent system [12]. First, you need a quick fix. Large data backlogs can cause serious network problems if they are not processed in a timely manner [13]. Then, accuracy is also important. If there are a lot of data errors, the security will not be guaranteed. At the same time, there will be problems in related data transmission and subsequent processing will be affected.

3.2 Application and Platform of Big Data Security

With the development of big data analysis technology, in the field of network security, big data analysis technology has been applied to a number of specific security work [14].

Application of big data analysis based on security log. The core idea of this application is the use of logs. Identify a potential common denominator through relevant statistical behavior [15]. According to the statistical results to find out the relevant safety rules, according to the rules to establish abnormal behavior model. Through the security model to apply to the actual, found the relevant holes, and can be improved model processing [16].

Advanced persistent threat attack is also known as Advanced Persistent Threat (APT) attack [17]. The APT attack has the characteristics of continuous attack and the main function is to steal the core data. It can bypass the detection of traditional protection methods and exist in the system for a long time. We can detect the data in the system through big data analysis technology, and obtain relevant abnormal data for analysis [18]. Through big data security analysis, threat perception can be improved and APT attacks can be detected and organized in a timely manner. Detection methods based on traditional feature analysis and insufficient to find new vulnerability attacks. So we need to use big data security analysis technology to protect the system. And through the intelligent system platform mining vulnerability use behavior. Currently, there are several commonly used security analysis algorithms in the field of big data security [19]:

(1) **Decision Tree Algorithm**. Decision tree algorithm is a method to approximate the value of discrete function. It is a typical classification method, the steps are divided into two steps. The first step is to generate a decision tree by processing sample data and induction algorithm based on corresponding rules. The second step is to classify and process the data by using the generated decision tree. The essence of it is to represent a series of rules for data processing in the form of binary tree. Data is traversed from the root node of the decision tree to the leaf node until it reaches the terminal, and classification is completed.

(2) **Naive Bayesian Model**. Naive bayesian model is a very accurate model in mathematics. However, the naive bayesian algorithm is simpler than the decision tree algorithm and the data integrity cannot be guaranteed. But naive bayesian model is more accurate in classification. Naive bayesian model is only suitable for small amount of data, and is inferior to decision tree model in large data analysis.

4 System Model

In this section, we will analyze system model for mining security holes in smart grid environment based on data flow analysis.

The vulnerability detection technology currently used in the society has great limitations in the era of big data explosion. Faced with massive data in the era of big data, we are bound to add network interface and network equipment to meet the detection requirements of massive data. This paper focuses on the above problems and proposes detection methods for big data security analysis [20]. And it mainly describes the construction of algorithm environment required by relevant experiments and the configuration of relevant parameters. The relevant steps of the experiment and parameters are introduced in detail, which is apparently to find that in the analysis of data flow, the original algorithm can achieve less false positives. It shows that the algorithm can satisfy a preliminary analysis function of data flow.

Big data analysis application based on data flow. We know that traffic is generated in network communication. Our application is built on the idea that behavior generates traffic. We set up a detection system for the nodes of the traffic flow in the network, through the analysis of the platform. Identify data streams that differ from normal traffic to complete the search for potential vulnerabilities. This model distinguish abnormal data and normal data, then carry out multidimensional analysis of abnormal data and establish relevant exception handling model to deal with the subsequent attack. This method can be used to exploit Web vulnerability, attack detection, scan attack and denial of service attack detection. Traffic analysis also allows for other aspects of the system, such as user behavior analysis. Figure 1 shows a system model for mining security holes in smart grid.

Big data analysis technology has gradually entered the field of security. We can use the big data analysis platform to collect data. In addition, we can use it for further analysis. By collecting and following up, we build a complete set of patterns. This mode can effectively solve the relevant security problems, so as to maintain network security. Big data security analysis can also be combined with a variety of other technologies such as artificial intelligence.

Current big data analysis methods are mainly based on distributed architecture to process massive data, and then match with the known database contents. After matching, detection, analysis and early warning of stored attacks in the database, the core of big data security analysis method based on self-updating threat intelligence database is to collect valuable intelligence to make up for the defects of traditional database.

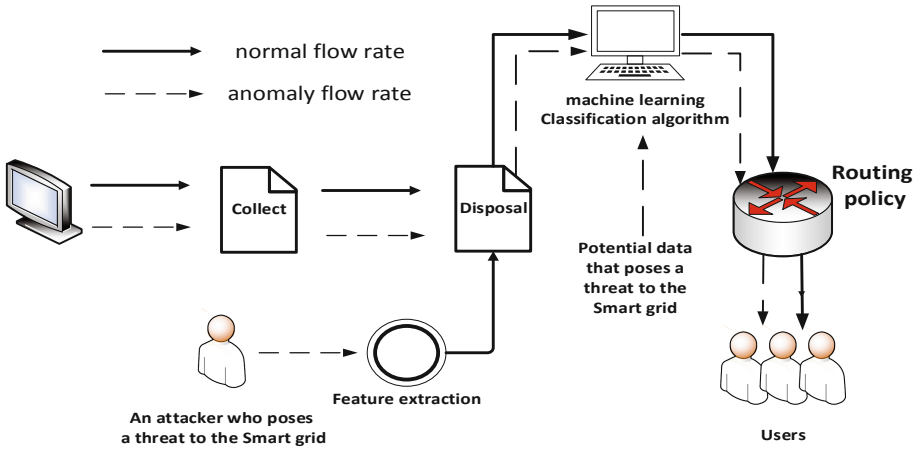

Fig. 1. A system model for mining security holes in smart grid.

5 System Design

The detection system is mainly composed of two parts: feature extraction and big data security analysis platform.

5.1 System Design Based on Big Data Detection

Feature Analysis: The main content of feature analysis is to analyze the related attributes of known data. After extraction, the attack behavior is classified from a series of factors such as traffic, behavior and attack mode. Feature extraction is the first step of data analysis [21]. Through feature analysis, we can more easily find the same rule of data flow. We use this common ground to further process the data. This process involves the screening of relevant redundant invalid data and classification of important data.

Big Data Security Analysis Platform: This module is the core module of big data security analysis [22]. It utilizes Spark-streaming platform and adopts decision tree classification. The algorithm analyzes the data obtained in one step, establishes the abnormal mechanism, and classifies the massive data. The classified data will distinguish between normal traffic and abnormal traffic, which will be assigned to different big data analysis modules for processing.

The big data security analysis platform is mainly divided into three levels: the collection layer is mainly responsible for sorting out the data in the network and extracting characteristic information, that is, useful information. Through the collection layer, the data to be processed by the big data platform is preliminarily reduced to facilitate the subsequent data processing. While the acquisition layer is working, we can do some preliminary processing on the data package. Based on the properties of the packet, the core data stream is saved and invalid data

is discarded. After the data is preliminarily processed, the processed data is sent to the next layer for sorting. The main software of the finishing layer is distributed kafka system. The system receives the request from the upper level, organizes and stores the data stream, and waits for the data to make the next application call. In addition, the collation layer can also act as a buffer for the computing layer. The computing layer is mainly responsible for the calculation of mass data, and the classification algorithm is used here. This layer adopts Spark-streaming platform for streaming processing requirements. Decision tree algorithm is simple and easy to understand, which processes a large number of data, efficient and so on. By analyzing the potential security vulnerability of smart grid, we analyze and design the vulnerability mining model. Figure 2 shows the potential security threats in smart grid.

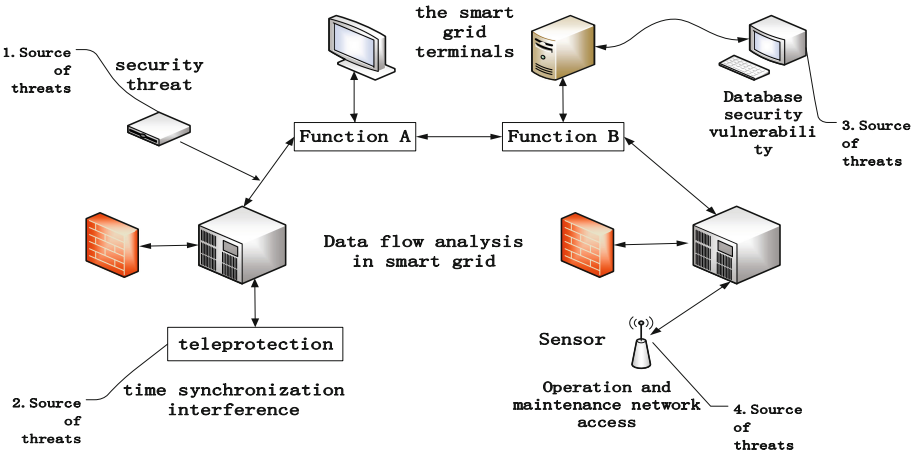

Fig. 2. Potential security threats in vulnerability mining design in smart grid.

5.2 Feature Extraction

Due to the massive data, we need to classify and process the big data first, so as to distinguish the normal traffic from the abnormal traffic. The steps for data classification generally have two parts. The first step is to classify data of known types according to relevant attributes and category relationships to obtain relevant data model. The data model is the feature model. The attributes used in the relevant classification are called eigenvalues. Second, after the establishment of the feature model, we can use the model to classify the new data and use the created vector model to classify and process data of unknown types. The accuracy of classification model is an embodiment of the accuracy of feature extraction. Common classification algorithms include decision tree classification and neural network classification [23].

6 Decision Tree Algorithm

Big data analysis technology has gradually entered the field of security. We can use the big data analysis platform to collect data. In addition, we can use it for further analysis. By collecting and following up, we build a complete set of patterns [24]. This mode can effectively solve the relevant security problems so as to maintain network security. Big data security analysis can also be combined with a variety of other technologies such as artificial intelligence. Decision tree algorithm is an important embodiment of big data security analysis and artificial intelligence [25]. In order to make the analysis more accurate, data analysis needs the support of multiple data. The required content is data traffic and relevant local log. For the obtained data, we can use Spark-streaming for processing. The calculations are then performed using Spark's own system. Flume used in the follow-up experiment is a data acquisition and simple processing system.

Decision tree algorithm is a method to approximate the value of discrete function. It is a typical classification method, the steps are divided into two steps. The first step is to generate a decision tree by processing sample data and induction algorithm based on corresponding rules. The second step is to classify and process the data by using the generated decision tree. The essence of decision tree is to represent a series of rules for data processing in the form of binary tree. Data is traversed from the root node of the decision tree to the leaf node until it reaches the terminal, and classification is completed.

Decision tree algorithm finds classification rules in data by constructing it. How to construct decision tree with high precision and small scale is the core step of decision tree algorithm [26]. After the above steps, we need to perfect the decision tree, that is, the pruning algorithm of it. By validating it, the rules that affect accuracy are removed, i.e. the corresponding branches of the decision tree are removed. Figure 3 shows the corresponding structure of the decision tree.

Leaf node N represents the relevant category. M_0 is the root node, and data processing starts from M_0 and $cond_{ij}$ is the relevant classification rule. In other words, we need to construct it with three conditions: data set, attribute set and attribute selection rule.

The construction of the decision tree is shown as follows:

Step 1: Construct a root node M. The classification or regression target of all nodes should be defined in the root node. The root node and the child node are opposite, that is, the child node is split from the root node according to a certain rule, and then the child node continues to split as a new root node until it cannot be split. The root node has no parent.
Step 2: If all the data in data set A have the same class index (denoted as class B), that is, all the sample data conform to rule B according to rule B. We'll just label M as B and return M.
Step 3: If all the data in data set A do not conform to the corresponding rules, that is, the class mark is different from all the class marks, then we mark M as the leaf node as the class mark with the most data, and A no longer performs classification calculation and returns M [27].

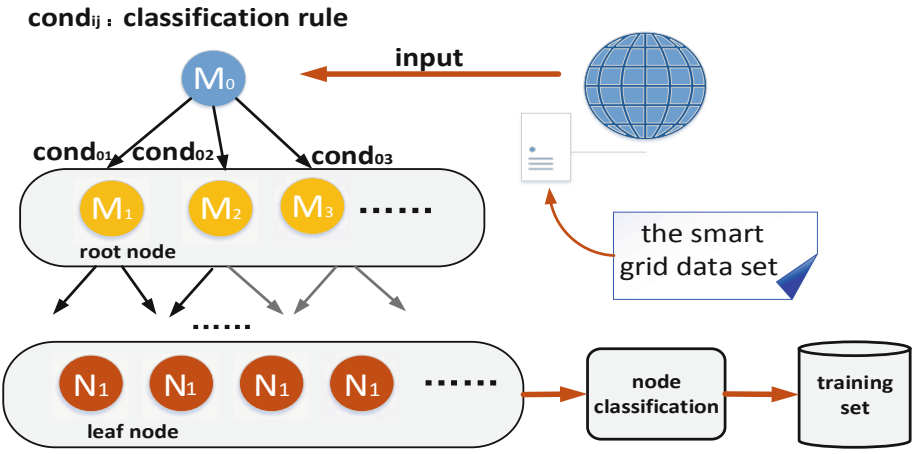

Fig. 3. Decision tree architecture applied to smart grid environment.

Step 4: The attribute selection algorithm is invoked to select the most accurate classification rule. Through this algorithm, the most accurate attributes of the classification are obtained, and subtrees of this category are required to belong to only one category as far as possible [28]. Entropy and index are commonly used to test the method of impurity. The test method of entropy is shown in formula:

$$Info(D) = -\sum_{i=1}^{n} p_i \log_2(p_i), \tag{1}$$

where n represents the total number of class B, P_i represents the probability that the data in A belongs to B_i, and the value of i ranging from 1 to n indicates the number of classes of these discrete variables. P_i is the probability that any sample data is class B. And Info of D is what we call entropy. If we have a property called R, R has n equals k values. Then we can divide the samples in D into $n = k$, and the entropy value formula after classification is shown in formula:

$$Info_R(D) = -\sum_{j=1}^{k} \frac{|D_j|}{|D|} Info(D_j). \tag{2}$$

This section mainly introduces the relevant big data security analysis platform program overview, including the big data security analysis platform related to some of the structure. At the same time this section chooses the decision tree algorithm to carry on the simulation experiment. The construction of decision tree is completed by simulation experiment, and the accuracy of the decision tree is also calculated. It also describes how to use entropy to calculate the accuracy of the relevant tree.

7 Performance Evaluation

7.1 Analysis of Decision Tree Environment

To build the Spark cluster, we use the Yarn-based farthest scheduler [29]. We use A total of three machines during the experiment, one as the host and we name it C, the other as the attack machine and we name it A, and the last one as the normal communication machine and we name it B. Brief experimental steps: we use attack machine A to attack host C, and machine B has normal communication with host C. In the interaction between the host and two machines, we use packet capture software to grab the interactive data packets, and then analyze the data packets through the built platform. The false alarm rate of abnormal data packets during communication and attack are calculated, and the average running time of three groups of independent experiments is calculated.

The simulation steps is shown as follows.

Step 1: Attack machine A and normal communication machine B access host C at the same time to ensure smooth communication.
Step 2: We use the packet capture software for packet capture.
Step 3: Data collation and start Kafka module.
Step 4: Import Kafka package into the program to be run and submit it to spark cluster for calculation in the form of spark-submit.

7.2 Simulation Results

We set the time window to 160 s and recorded the missing alarm rate of IP scanning, Port scanning and FIN scanning (Table 1).

Table 1. The set of parameter results.

Parameter	IP	Port	FIN
False Alarm Rate	0.5%	1.7%	1.7%
Missing Report Rate	0.7%	1.3%	2.4%

It is obviously to see from the above tables that the alarm rate and false alarm rate related to the decision tree algorithm are simulated by means of data simulation experiment. We get three sets of experimental results and find that the data obtained from the experiment showed that the alarm rate is about 1% and the false alarm rate is about 1.5%. This proves that our algorithm can be capable of big data security analysis of data screening.

7.3 Performance Analysis

We simulate the missing alarm rate and false alarm rate of decision tree algorithm. The grab results are shown in the Figs. 4, 5, and 6:

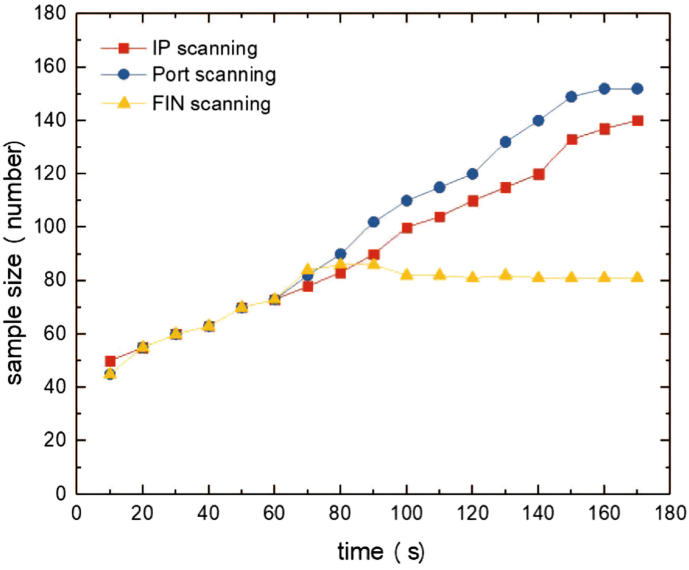

Fig. 4. Performance comparison when the time window is 20 s.

By means of data simulation experiment, this paper makes a simulation experiment on the missing alarm rate and false alarm rate of decision tree algorithm.

Through controlling variables, three groups of experimental results are obtained. The data obtained from the experiment shows that the alarm rate of IP scanning and Port scanning is about 1% and the false alarm rate is about 1.5%. We find that the missing alarm rate of the system is slightly higher than the false alarm rate, which is caused by the limitation of classification in the algorithm. Failure rate can be effectively reduced if the content of the original data set can be enlarged.

Through the integrated data of the three groups of tables, we can find that IP scanning attack has a better effect than other port attack and FIN scanning attack. The accuracy of Synchronize (SYN) attack and Domain Name System (DNS) attack needs to be improved. This proves that our algorithm can be capable of big data security analysis of data screening.

In conclusion, the experimental algorithm has a low false alarm rate, fast processing time and good performance against this kind of attack. Through this experiment, we find that big data has a high detection capability against several

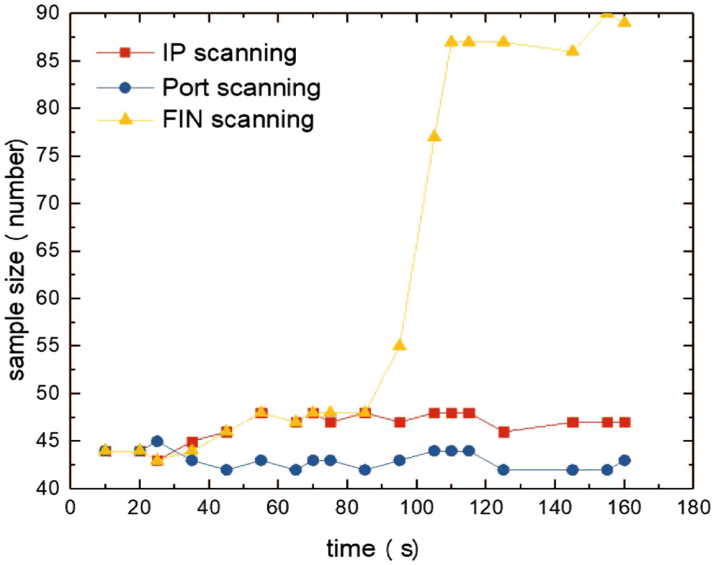

Fig. 5. Performance comparison when the time window is 40 s.

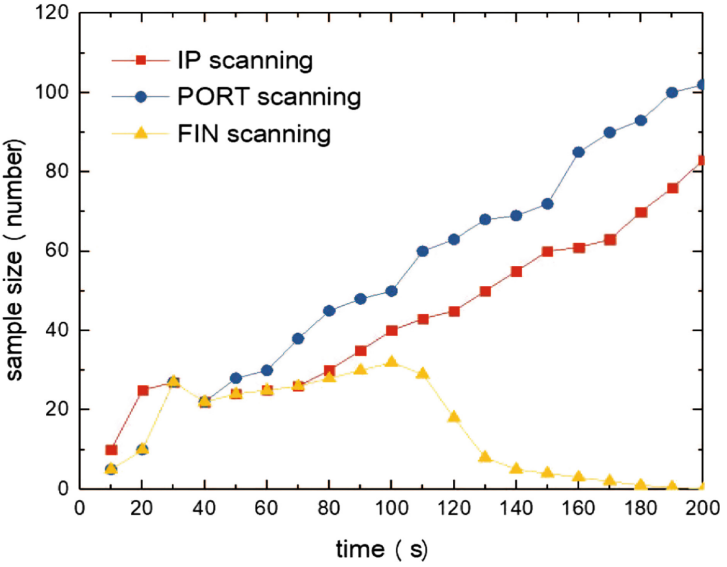

Fig. 6. Performance comparison when the time window is 120 s.

attack technologies in the above table. Compared with traditional vulnerability detection technology, big data security analysis technology has lower alarm rate and false alarm rate.

Through the experimental simulation, it is obviously to find that the original algorithm can achieve less false positives and less false positives. This shows that the algorithm can satisfy a preliminary analysis function of data flow.

This section conducts a simulation experiment on the missing alarm rate and false alarm rate related to decision tree algorithm by means of data simulation experiment. After controlling for variables, we got three sets of experimental results. We find that the data obtained from the experiment showed that the alarm rate was about 1% and the false alarm rate was about 1.5%. We are able to distinguish normal data and abnormal data accurately through the test of sample data.

8 Conclusion and Feature Work

In this paper, we propose an effective method of mining network security vulnerabilities based on data flow analysis technology. In order to improve the security of smart grid data transmission, we introduce the decision tree algorithm into the data transmission environment of smart grid and adopt more accurate vulnerability detection mechanism to ensure network security. In addition, we also demonstrate the characteristics of low alarm rate and false alarm rate by using decision tree algorithm for data flow analysis. This proves that our algorithm has the performance of big data security analysis. We also have some shortcomings which will be corrected in the future work. For example, if a user provides wrong information, they cannot be effectively found, so we need to design a scheme to distinguish them.

References

1. Hu, R.: Key technology for big visual data analysis in security space and its applications. In: 2016 International Conference on Advanced Cloud and Big Data (CBD), vol. 3, no. 4, p. 333 (2016)
2. Yu, J., Wang, K., Li, P., Xia, R., Guo, S., Guo, M.: Efficient trustworthiness management for malicious user detection in big data collection. In: IEEE Transactions on Big Data, October 2017
3. Albu, A.: From logical inference to decision trees in medical diagnosis. In: E-Health and Bioengineering Conference (EHB). Sinaia **2017**, 65–68 (2017)
4. Park, K.: On the effectiveness of route-based packet filtering for distributed DoS attack prevention in power-law internets. In: 2018 ACM SIGCOMM Computer Communication Review, December 2018
5. laski, J., Stanley, W.: Software verification and analysis. In: 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Springer, London (2018) https://doi.org/10.1007/978-1-84882-240-5
6. Wang, K., et al.: Wireless big data computing in smart grid. IEEE Wireless Communi. **24**(2), 58–64 (2017)
7. Wang, K., Du, M., Maharjan, S., Sun, Y.: Strategic honeypot game model for distributed denial of service attacks in the smart grid. IEEE Trans. Smart Grid **8**(5), 2474–2482 (2017)

8. Xu, C., Wang, K., Li, P., Xia, R., Guo, S., Guo, M.: Renewable energy-aware big data analytics in geo-distributed data centers with reinforcement learning. IEEE Tran. Network Sci. Eng. **7**, 205–215 (2018)
9. Amir, M.A.U.: Optimal specification of wave flume in confined space. In: 2016 International Conference on Information and Communication Technology (ICICTM), Kuala Lumpur, pp. 136–140 (2016)
10. Du, M., Wang, K., Xia, Z., Zhang, Y.: Differential privacy preserving of training model in wireless big data with edge computing. In: IEEE Transactions on Big Data, May 2018
11. Al-Shomrani, A., Yang, Y.M., Wang, X.: Quickest detection of false data injection attack in wide-area smart grids. IEEE Trans. Smart Grid **6**(6), 2725–2735 (2017)
12. Wang, K., Fathy, F., Jambi, K.: Policy enforcement for big data security. In: 2017 2nd International Conference on Anti-Cyber Crimes (ICACC), pp. 70–74. Abha (2017)
13. Burguera, I., Zurutuza, U.: Behavior-based Malware detection system for Android. In: Proceedings of the 1st ACM Workshop on Security and Privacy in Smartphones and Mobile Devices, pp. 15–26. ACM, Chicago (2011)
14. Xie, L., Zhang, X., Seifert, J.: A behavior-based malware detection system for cellphone devices. In: Proceedings of the Third ACM Conference on Wireless Network Security, pp. 37–48. New Jersey, ACM (2010)
15. Gavankar, S.S., Sawarkar, S.D.: Eager decision tree. In: 2nd International Conference for Convergence in Technology (I2CT). Mumbai **2017**, 837–840 (2017)
16. Wang, K., Ouyang, Z., Krishnan, R., Shu, L., He, L.: A game theory based energy management system using price elasticity for smart grids. IEEE Trans. Ind. Inform. **11**(6), 1607–1616 (2015)
17. Al-Hoqani, W.M., Giampapa, J.A.: Difficulties of marking decision tree diagrams. Computing Conference. London **2017**, 1190–1194 (2017)
18. Wang, Y., Wang, K., Huang, H., Miyazaki, T., Guo, S.: Traffic and computation co-offloading with reinforcement learning in fog computing for industrial applications. IEEE Trans. Ind. Inform. **15**(2), 976–986 (2019)
19. Ignatov, D., Ignatov, A.: Decision stream: cultivating deep decision trees. In: 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), Boston, MA, pp. 905–912 (2017)
20. Gao, M., Wang, K., He, L.: Probabilistic model checking and scheduling implementation of energy router system in energy internet for green cities. IEEE Trans. Ind. Inform. **14**(4), 1501–1510 (2018)
21. Kumar, R.: Development of synchronization system of two spark gaps. In: 2012 IEEE 5th India International Conference on Power Electronics (IICPE), Delhi, pp. 1–3 (2012)
22. Revathy, P., Mukesh, R.: Analysis of big data security practices. In: 2017 3rd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Tumkur, pp. 264–267 (2017)
23. Wang, K., Shao, Y., Xie, L., Wu, J., Guo, S.: Adaptive and fault-tolerant data processing in healthcare IoT based on fog computing. IEEE Transactions on Network Science and Engineering, July 2018
24. Lighari, S.N., Hussain, D.M.A.: Hybrid model of rule based and clustering analysis for big data security. In: 2017 First International Conference on Latest trends in Electrical Engineering and Computing Technologies (INTELLECT), Karachi, pp. 1–5 (2017)

25. Aljuhani, A., Alharbi, T.: Virtualized network functions security attacks and vulnerabilities. In: IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC). Las Vegas, NV **2017**, 1–4 (2017)
26. Wang, K., Yu, J., Liu, X., Guo, S.: A pre-authentication approach to proxy reencryption in big data context. In: IEEE Transactions on Big Data, May 2017
27. Li, P., Min, X.: Accurate marking method of network attacking information based on big data analysis. In: 2019 International Conference on Intelligent Transportation, Big Data Smart City (ICITBS), Changsha, China, pp. 228–231 (2019)
28. Wang, K., Li, H., Maharjan, S., Zhang, Y., Guo, S.: Green energy scheduling for demand side management in the smart grid. IEEE Trans. Green Commun. Network. **2**(2), 596–611 (2018)
29. Nethercote, N., Seward, J.: A Framework for Heavyweight Dynamic Binary Instrumentation (2017)