Cluster Computing
https://doi.org/10.1007/s10586-018-1847-2

Mobile network intrusion detection for IoT system based on transfer learning algorithm

Lianbing Deng[1,2] · Daming Li[3,4,5] · Xiang Yao[2] · David Cox[6] · Haoxiang Wang[7,8]

Received: 10 October 2017 / Revised: 10 January 2018 / Accepted: 12 January 2018

Abstract
The open deployment environment and limited resources of the Internet of things (IoT) make it vulnerable to malicious attacks, while the traditional intrusion detection system is difficult to meet the heterogeneous and distributed features of the Internet of things. The security and privacy protection of IoT is directly related to the practical application of IoT. In this paper, We analyze the characteristics of networking security and security problems, and discuss the system framework of Internet security and some key security technologies, including key management, authentication and access control, routing security, privacy protection, intrusion detection and fault tolerance and intrusion etc. This paper introduces the current problems of IoT in network security, and points out the necessity of intrusion detection. Several kinds of intrusion detection technologies are discussed, and its application on IoT architecture is analyzed. We compare the application of different intrusion detection technologies, and make a prospect of the next phase of research. Using data mining and machine learning methods to study network intrusion technology has become a hot issue. A single class feature or a detection model is very difficult to improve the detection rate of network intrusion detection. The performance of the proposed model is validated through the public databases.

Keywords Intrusion detection · Internet of things · Information security · Pop learning

1 Introduction

With the development of wireless sensor network (WSN), communication technology and Internet of things (IoT) technology, more application environments use sensors and wireless communication devices, which consist of a large-scale IoT communication system. The application of the IoT will involve military, people's livelihood, industry and Commerce and other fields, and the importance of its network security is self-evident. Once the problem of information security appears, such as virus damage, hacker intrusion, malicious code attacks and other issues, the damage and losses will be very serious. The IoT is based on the construction of computer network, WSN, and the traditional mobile communication network, due to the inherent vulnerabilities of these networks, the safety of network information is also facing great challenges. Therefore, the study of the security technology of IoT is particularly important.

Intrusion behavior mainly refers to any attempt to destroy the integrity, confidentiality and access of target resources, which is the main aspect of the Security Research Institute of Internet of things. The traditional security defense mechanisms, such as encryption and authentication, are relatively passive. No matter how to update or update, they will always be attacked by intruders. Intrusion detection is a new security technology in recent years, can be relatively active for the network security detection and take corresponding measures

✉ Daming Li
lidaming@yahoo.com

1 Huazhong University of Science and Technology, Wuhan, China

2 Zhuhai Da Hengqin Science and Technology Development Co., Ltd., Hengqin, China

3 The Post-Doctoral Research Center of Zhuhai Da Hengqin Science and Technology Development Co., Ltd., Hengqin, China

4 City University of Macau, Taipa, Macau

5 International Postdoctoral Science and Technology Research Institute Co., Ltd, Wuhan, China

6 Harvard John A. Paulson School of Engineering & Applied Sciences, Harvard University, Cambridge, MA, USA

7 Cornell University, Ithaca, NY, USA

8 GoPerception Laboratory, New York, NY, USA

Published online: 31 January 2018

to compensate for the shortcomings of traditional security defense technology to a great extent [1].

Denning had defined intrusion detection model [2]. He pointed out that, intrusion detection should be based on the network packet information collection, and analyze the collected information, intrusion detection potential, and can be timely sent to the system management alert, to provide the corresponding treatment measures. Therefore, a typical intrusion detection system should include at least the necessary functions such as information collection, analysis, detection and alarm.

Sybil intrusion uses communication link vulnerabilities to invade computer system, which brings network security risks. Dealing with intrusion detection problems in IOT, improving networking security capabilities, has important application value in network security and networking network design, and the design method of intrusion prevention system received much attention [3–5].

Sybil intrusion prevention system for the IoT uses Sybil statistical characteristics of intrusion signal and high-order spectrum focusing characteristics, signal detection and recognition, and combined with the high-speed digital signal processing chip development and design of intrusion prevention system, and achieved certain results. The architecture of IoT is divided into three layers: the perception layer of data acquisition, the network layer of data transmission and the content application layer. Here we will briefly introduce the functions implemented by each layer and analyze the security threats existing in each layer.

It can be said that the Internet today is omnipresent and efficient and versatile. While the Internet has brought convenience and convenience to people's life and work, how to ensure the security of network information and the security of network equipment has increasingly become the focus of attention. According to incomplete statistics, hackers will occur every 20 s worldwide, and because of network security problems, the annual economic losses caused to the United States as much as more than 10 billion U.S. dollars [6]. According to reports, information theft is increasing at a rate of 250%, and 98% of well-known companies have experienced network intrusion events. Therefore, how to prevent network security has become an urgent and important problem in academia and industry, and one of the key technologies is network intrusion detection technology [7,8].

The IoT is a concept put forward in recent years, in the practical application, the sensor, processor and wireless communication module embedded or equipment to the power grid, railways, bridges, tunnels, highways, buildings and other objects, which are mutually connected, constitute the IoT. In 1999, Massachusetts Institute of Technology (MIT) established the automatic identification technology center, conceived the concept of IoT based on RFID, and put forward the concept of product electronic code. EPC system can track the goods in real time, and can optimize the whole supply chain to provide support to users, thus promoting the rapid development of automatic identification technology, and can greatly improve the quality of life of consumers [9].

From the point of view of information and network security, networking is a multi network integration of heterogeneous networks, not only the existence and sensor network, mobile communication network and the Internet the same security issues, and its particularity, such as privacy protection issues, heterogeneous network authentication and access control, information storage and management etc. In [10], the author believes that data and privacy protection is one of the challenges in the IoT applications. In the IoT, the RFID system realizes the perception of terminal information, Wang and Wei [11] discusses the cryptographic algorithm of data transmission in RFID system, and uses the logic encryption module in IC card to encrypt the information. Weber [12] designs an information service system based on RFID, mainly aiming at the application of logistics management. For the problems of security and privacy protection of Internet of things, in [13], a service security model of IoT is proposed, and the functions of each module in the model are analyzed. Mulligan [14] analyzes the status of IoT, and discusses the security issues.

This paper attempts to start from information security, confidentiality, integrity and availability. We analyze the security features and security issues of IoT, and discuss the security architecture of IoT, and some security key technologies, especially discuss the key management and routing technology. In Fig. 1, we present the IoT framework.

2 The composition of the Internet of things

2.1 Perception layer

The perceptron layer is the basis for the development and application of the IoT, which includes sensors for data acquisition, end devices, and sensor networks before data access to the Internet gateway. The perception layer is attacked, and usually the perception layer nodes are hijacked, including ordinary nodes and gateway nodes. Ordinary nodes, once controlled by the attacker, security risks is not only the information is stolen, the attacker can also tag on the goods handling, such as circuit interruption, clock failure etc. This type of attack can make legitimate common nodes impossible to identify, resulting in no corresponding service. Attackers can even separate the manipulated tags and objects from the nodes and associate them with other objects. The joint point of the network also has the potential of malicious manipulation. Once the attacker achieves this goal, he can broadcast a large number of interference signals to cause persistent congestion to the network. In addition, the IoT is to realize

Fig. 1 Internet of things framework

the object at any time any place connected network and ultimately to access the Internet, thus inevitably suffered attacks from the Internet, the more common is the illegal access and denial of service attacks. Because the nodes of sensor networks have single structure, small resources and low energy carrying capacity, they are vulnerable to attack, resulting in node collapse and even paralysis of sensor networks [15].

2.2 Network layer

The IoT is built on the basis of the existing communication network and the Internet, and combines the existing communication technology with the combination of the perception network and the communication network. The main job of this layer is to reliably receive data from the perception layer and then process it according to different application requirements. This layer mainly considers security threats and security architecture issues, and it can be transplanted or referenced to the existing research results of Internet security. To sum up, the security requirements of network layer include data confidentiality, integrity, detection and prevention of attacks, etc. [16].

2.3 Application layer

Application layer is the social division of labor of the Internet of things, combined with specific industries, to achieve a wide range of intelligent. This layer receives the information reliably from the network, and carries on the corresponding information processing and management by some middleware system. One point to note is that the received information needs to be judged first, and useful data, garbage data and malicious data are identified. The security challenges faced by the application layer is the first to face recognition and processing of massive data processing, the platform may also be distributed, how to allocate and coordinate and efficient intelligent processing of data but also need to consider the issue. In addition, the intelligent automatic process also has an attacker to bypass or tampering with the risks, once the automatic process is being attacked or have been attacked and lead to disaster, it should have a corresponding mechanism to guarantee can be controlled timely and effective interrupt and self protection, can recover from a disaster. Finally, in the era of personal and business information networking, it is necessary to establish a corresponding

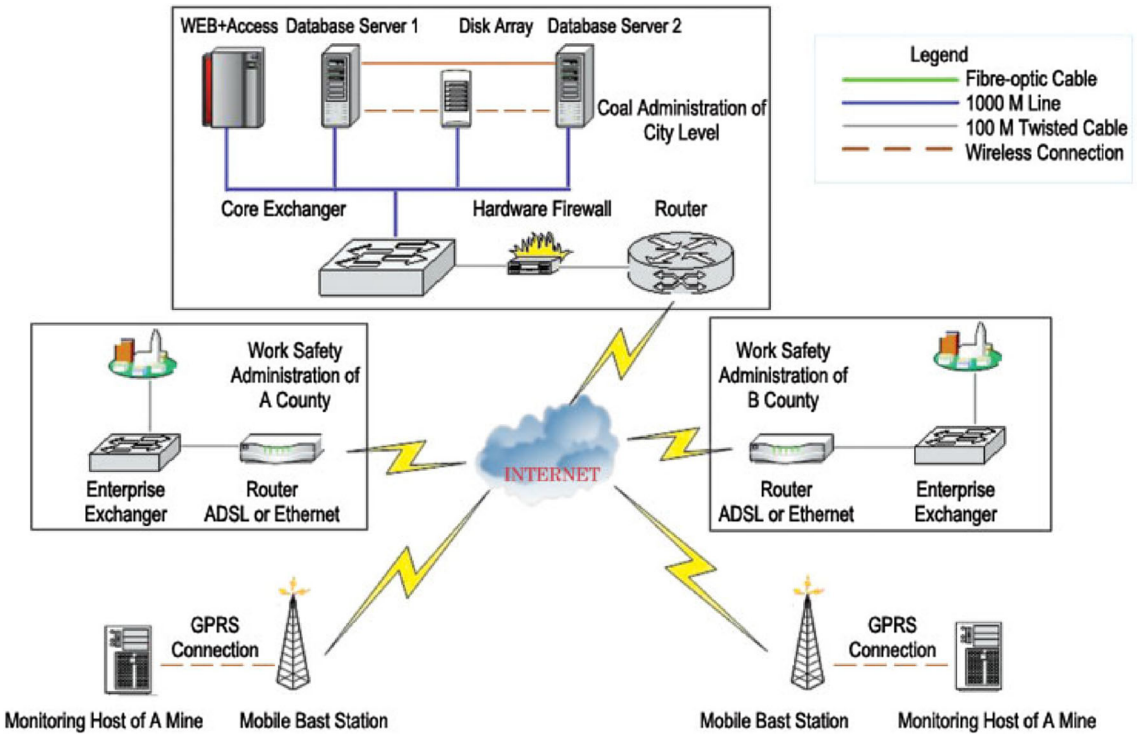


Fig. 2 Intrusion Detection System

security protection mechanism for privacy information. In Fig. 2, we present the architecture of the intrusion detection system.

3 Application of intrusion detection technology in Internet of things architecture

In the face of the security threats in the IoT, the effective intrusion detection technology needs to be simple, real-time and accurate. The following part describes the current intrusion detection technology applied to the IoT.

3.1 Intrusion detection technology based on multi-agent model

Agent is a software entity with independent logic processing capability and continuous operation under given conditions. Agent is autonomous, mobile, and Agent can communicate with each other to accomplish tasks. According to the features of the Internet of things, we consider the application of multi agent intrusion sensing layer detection technology, the intrusion detection system has to reduce network load and delay, dynamically adapt to network changes and rapid real-time response [17].

The multi-agent intrusion detection system based on the perception layer is composed of detection agent, host agent and network agent [18]. Intrusion detection systems are deployed on multiple network terminals in the sense layer. Each network terminal has a plurality of detection agents to monitor events occurring on the local machine. Each terminal is equipped with a host agent, whose main function is to manage all the test agents on the corresponding terminals. It is responsible for checking the operation of the detection agent and filtering the data reported by the detection agent. In a certain network latitude, the corresponding network agent will be set up. The host agent will report the filtered data to the network agent within the network. Between the host agent and agent network is many to many relationship, to avoid the system because of a network failure and downtime agent. The network agent will form a hierarchical structure, and the high-level network agent will be responsible for reporting the results to the console.

3.2 Intrusion detection technology based on game theory model

In the IoT perception layer, intrusion detection system needs to rely on its own behavior and the behavior of the intruder to take relevant actions. Any change in the policy of either side of an intruder and intrusion detection system can lead to changes in the other party's strategy. The distribution of

Table 1 Application result comparison among intrusion detection technologies

Evaluating index	Accuracy index		Efficiency index		System index	
	Detection rate	False positive rate	Maximum processing capacity	System robustness	Failure rate	Deployment cost
Multi agent based	High	Low	High	Low	High	High
Game theory based	Medium	Medium	Low	High	High	High
Bayes based	Medium	Medium	Low	High	Low	Low
Machine learning based	High	Low	Low	High	Low	Low

intrusion detector deployment in the perception layer network terminal will use a detection method of network data to distinguish between normal audit data and attack data, intrusion data filtering the detected data into simplified summary report submitted to the game model. By simulating the interaction and comparison of both offensive and defensive balance detection results and efficiency, thus the Nash equilibrium theory is obtained.

3.3 Intrusion detection technology based on machine learning

Using machine learning methods to achieve intrusion detection, the main methods include inductive learning, analytical learning, analog learning, genetic algorithm, etc. Genetic algorithms are good at solving global optimization problems, and they can jump out of local optima to find the global optimum. Moreover, the genetic algorithm allows the use of very complex fitness functions and limits the range of variables that can be changed. Under the guidance of no definite rules, the search direction can be adjusted adaptively [19].

The genetic algorithm only needs to search a few structures, plus the fitness information of the population, in order to figure out the good solution quickly through selection, crossover and mutation. Thus, in the complex information of the network layer, the intrusion attack information can be identified in time.

4 Application of Internet of things architecture

Multi agent based intrusion detection technology can be well applied to the IoT perception layer and combined with network terminal because of its autonomy and mobility. Through the host agent and the detection agent on the terminal, the operation and behavior of the sensing terminal are monitored, so as to monitor and prevent the hijack attack against the terminal node.

The intrusion detection technology based on game theory model collects and distinguishes the normal data and attack data from the detection terminal deployed on the network layer, and then gives the game model to balance, and obtains the reasonable corresponding strategy.

Machine learning based intrusion detection technology captures the network information in the network layer and detects intrusion behavior through machine learning, and it can provide good adaptive ability.

4.1 Comparison of application effect

In traditional network environment, the performance indexes of IDS are usually divided into three categories, namely, accuracy index, efficiency index and system index. Among them, the accuracy index is used to describe the accuracy of intrusion detection technology for intrusion detection, including detection rate, false positive rate and false negative rate [20]. Efficiency index describes the execution ability of intrusion detection technology, usually including the maximum processing capacity, the number of concurrent TCP per second, the maximum number of concurrent TCP and so on. But the efficiency index depends on the environment and equipment, and the selection of efficiency index between different levels of equipment is also very different. The system index usually embodies the stability and ease of use of the intrusion detection system, which usually includes the maximum number of rules and the average failure free interval. Table 1 presents the application result comparison among intrusion detection technologies.

5 Improved learning model

Combating cyber crime, ensuring the normal operation order of the network and reducing the economic losses caused by cyber crimes have become a major task of the reconnaissance authorities. Network forensics was first proposed by Marcus Ranum in 1990s of last century. Using data mining and machine learning methods to study Network Intrusion Forensics has become one of the focuses of network forensics research.

Mukkamala and Sung [21] proposed the use of artificial neural networks and support vector machines (SVM) to analyze the importance of network traffic characteristics. This method can improve the automatic process of network forensics to some extent and improve the accuracy of event detection.

In network forensics analysis technology, there are many new methods. But they all have their own applicable environment. A single class feature or a detection model is very difficult to improve the detection rate of network intrusion detection. We propose an integrated scheme based on multiple features or multiple models.

Network intrusion detection is a typical artificial intelligence problem. Specifically, it is a problem of knowledge discovery, so data mining can be used to solve the problem of network intrusion detection. The main methods of data mining are association rule analysis, classification analysis and association analysis. This method is suitable for extracting decision model, the method is applied to computer network forensics data analysis, mining data acquisition phase data, discover in time and space on the relationship between events, extracting relevant features and rules in using the features and rules of the user definition of normal and abnormal patterns stored in knowledge library. Mining the current user's behavior data, generating user behavior characteristics and rules, and comparing with the model in the known knowledge base, and analyzing the data of network forensics to determine whether the user behavior is an attack.

At present, most of the intrusion detection techniques are based on feature matching, and the updating of the knowledge base often requires additional human and material resources. After Lee has applied data mining to intrusion detection, a variety of machine learning methods have also been introduced. The intrusion detection model based on data mining can machine learning and model expansion, and update the knowledge base automatically according to the network environment, which greatly reduces the manual and empirical components in the system.

5.1 Rough set based

The rough set classification method is applied to intrusion detection. Firstly, the data is discretized, and then the hybrid genetic algorithm is used to compute the attributes reduction of the rough sets. The main advantage of rough sets is that they can obtain a minimum set of prediction rules, which can improve the detection speed. The resulting rule is the production rule in the form of "if-then", which is highly explanatory.

5.2 SVM based

SVM is a new type of statistical learning model. It has many unique advantages in solving nonlinear and high-dimensional pattern recognition problems. In 2002, Eskin and Honig used an SVM for unsupervised learning.

Table 2 Experimental results of different intrusion detection algorithms

Algorithms	Correct classification rate of DOS (%)	Correct classification rate of Probe (%)
Rough set based	95.8	81
SVM based	78	84.4

We select MIT Lincoln Laboratory 1999 DARPA intrusion detection evaluation plan to collect the original traffic data, which has 494,021 data records. Each record consists of 41 feature sets and one classification attribute. The classification attribute has only five values: normal and DOS, R2L, Probe, and U2R. Training data has 9775 connections, 1928 normal connections, 7847 attacks. In the test data, there are 10952 connections in all, including 2251 normal connections. Table 2 present the experimental results of different intrusion detection algorithms.

As can be seen above, rough sets are more accurate for detecting DOS attacks. But because of the complexity of the network environment, almost impossible to design a perfect classifier or learning algorithms in various event classification performance, so all kinds of learning algorithms, and design a new algorithm to solve the existing calculation method in analyzing the performance of various complex events in instability. Based on this idea, this paper proposes to integrate traditional popular learning algorithms to improve the adaptability and stability of the original learning algorithm.

5.3 A model based on ensemble learning

Classification is an important task in data mining. It classifies the sample data according to the discriminative feature of the training samples. There are many kinds of single classification technology, such as decision tree technology, AQ method, rough set method, SVM and so on. However, single classification technology is often limited by certain conditions in application. Therefore, constructing a good combination classifier has become one of the new research fields in data mining.

The combination of classifiers is to combine the prediction of multiple single classifiers by a combination technique to generate a new classifier and classify the training samples with the new classifier. Voting classifier integration scheme is a static voting method, this method does not change with the change of the training data set. Using analysis algorithm to detect different types of attacks, the adaptive model detection model automatically generated system, detection results using the ensemble learning method in machine learning

fusion test results. The model can achieve real-time and complementary data mining, and improve the speed and accuracy of network intrusion analysis.

5.4 Method of feature vector selection based on PCA

With the increasing amount of data transmitted by the Internet of things, one of the main problems faced by the traditional intrusion detection systems is the limited performance of the sensor nodes in the IoT. In the process of transmission, if the transmission networks of the IoT will be facing a large amount of data, large load capacity, low detection rate, then caused the consequences of delay, omission, resulting in the face of network intrusion behavior is not timely take necessary safety measures, resulting in information security problems.

Principal component analysis is introduced in 1901 by Pearson. PCA can reduce the dimensionality of the sample space, the purpose is to use less variables to represent more variables in the data, and to ensure the integrity of the data information. Only a small part of the mass data transmitted by the network is utilized in the intrusion detection process. Thus, in the intrusion detection feature selection, the PCA algorithm will greatly reduce the time spent by the system [22].

The principal component analysis maps the sample matrix X to the vector Y in the lower dimensional space in accordance with the linear transformation in Eq. 1

$$Y = W^T X, \tag{1}$$

where X is sample matrix, W represents projection matrix which is composed of sample covariance matrix, defined as:

$$C = \frac{1}{N}\sum_{i=1}^{N}[(X_i - u_i)^2] \tag{2}$$

$$u_i = \frac{1}{N}\sum_{i=1}^{N} X[m, n]. \tag{3}$$

Thus we have:

$$CW_i = \lambda_i W_i. \tag{4}$$

The massive data information in the network is only used in the process of intrusion detection, and the system needs to spend a great deal of computation when removing the useless information. PCA is a dimension reduction algorithm. In the face of more complex systems and large amounts of data, it can eliminate the unwanted information contained in it and retain the necessary information. The PCA algorithm is applied to the feature selection of intrusion detection, which can reduce the number of features of the system processing, thus shortening the detection time. The PCA algorithm is designed to project high dimensional data to the low dimensional space, the multi-dimensional vector discriminative variables is not removed, leaving large variance feature, with fewer new variables instead of more variables, and to ensure that the data is not distorted [23].

5.5 Fuzzy C means clustering algorithm

K-means is a distance based clustering algorithm. Euclidean distance is used as a similarity evaluation index, and the closer the distance between 2 objects, the greater the similarity between them. For a set of data, it will be seen as some clusters, forming a cluster by near objects, each cluster is a cluster, because objects between the same clustering distance, so it will be possible to use the similarity, the K-means algorithm can be all objects according to similarity clustering, and calculate each clustering center a class [24].

FCM is a kind of clustering algorithm proposed by Bezdek in 1981, the algorithm uses the membership matrix is given for each sample belonging to a certain degree of clustering, even though some difficult to classify the sample variables, FCM can achieve better clustering results, clustering is the most widely used algorithm of data analysis.

Assume we have a data sample set, denoted as:

$$X = \{x_1, x_2, \ldots, x_n\}. \tag{5}$$

FCM divides the data set into c fuzzy clustering sets, so that the objective function of the non similarity index is minimum.

To test performance, seven datasets of UCI were used. For each data set, 20% of the data is reserved for the test

$$\sum_{i=1}^{c} u_{ij} = 1, \quad \forall j = 1, \ldots, n \tag{6}$$

$$\forall i, \quad \forall j, u_{ij} \in [0, 1] \tag{7}$$

$$\forall i, \quad \sum_{j=1}^{n} u_{ij} > 0 \tag{8}$$

In order to get the optimal fuzzy solution, the objective function is set as:

$$J_{FCM(U,V)} = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m d_{ij}^2, \tag{9}$$

where V_i is the clustering center, $V = \{v_1, v_2, \ldots, v_n\}$ is set of fuzzy group. We use the Lagrange multiplier method to solve the optimal solution:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}(\frac{d_{ij}}{d_{kj}})}, \quad \forall k, i; 1 \le k \le c; 1 \le c \le n \tag{10}$$

$$u_k = \frac{\sum_{j=1}^{n} u_{ij} * x_i}{\sum_{j=1}^{n} u_{ij}^m}, \quad \forall i \tag{11}$$

6 Security features and architecture of Internet of things

The goal of information and network security is to achieve the confidentiality, integrity, and availability of protected information (confidentiality). In the early stages of the Internet, more attention was paid to basic theory and applied research. With the increasing scale of network and service, security problems have been highlighted, and people have paid more and more attention to it. Some security technologies have been introduced, such as intrusion detection system, firewall, PKI and so on. The research and application of the IoT in the primary stage, the theory and key technology of many breakthroughs, especially compared with the Internet and mobile communication network, yet to show the practical application of convincing, we will from the development of the Internet to explore the process of networking security issues.

6.1 Internet of things security features

The process of information processing embodies the security characteristics and requirements of the IoT, and also reveals the security problems faced by the process of gathering, gathering, merging, transmitting, decision-making and controlling of the perceived information.

The perception nodes show multi source heterogeneity, and the sensing nodes usually have simple functions and low energy carrying capacity, which makes them unable to have complex security protection capabilities, and the perception network is diverse. Data transmission and messaging do not have specific standards, so there is no way to provide a unified security protection system.

The core network security protection capability is relatively complete, but due to the number of nodes in the IoT is huge, and exist in the cluster, thus will result in data transmission, the data sent a large number of machines to make the network congestion, resulting in denial of service attacks. In addition, the security architecture of the existing communication network is designed from the point of view of human communication.

Support business networking platform with different security strategies, such as cloud computing, distributed system, massive information processing, the supporting platform to establish an efficient, reliable and credible system for the service management and large-scale industry application. The encryption of information is an important means to realize confidentiality. Because of the heterogeneous and heterogeneous nature of the IoT, the key management is more difficult.

The integrity and usability of the IoT runs through the entire process of the IoT data streams, network intrusion, denial of attack service, Sybil attacks, routing attacks, etc., which undermine the integrity and availability of information. At the same time, the IoT perception interaction process also requires network with high stability and reliability, the Internet is a physical device and many applications related to network, to ensure stable and reliable, such as warehousing and logistics applications, things must be stable, to ensure network connectivity.

Therefore, the safety characteristics of things reflects the diversity of diversity, the network environment perception information and application requirements, showing the size of the network and the data processing capacity of large, complex decision control, put forward the new challenge.

6.2 Internet of things security framework

The perception layer can obtain all kinds of data, including the object properties, environmental conditions, behavior state of dynamic and static information, through the data on the perceived level of aggregation and transmission of sensor network or RF reader network and equipment. The transport layer is mainly implemented by mobile communication network, satellite network and Internet, so as to realize the access and transmission of the sensing layer information. The support layer provides an efficient and reliable supporting technology platform for the upper application services, and provides services for applications through parallel data mining and processing, shielding the heterogeneity of the underlying network and information. The application layer is based on the user's needs to establish the corresponding business model, and run the corresponding application system.

As a multiple networks fusion network, Internet safety involves the different levels of each network, in the practical application of these independent networks have a variety of security technology, especially mobile communications network and Internet security research has experienced a long time, but for sensor networks in the IoT, because of the limitations on resources, make the safety research is difficult.

Security defense technology ensures the security of information, the network and communication transmission security, the security technology of network environment, such as VPN, routing, etc., to realize the safety of network interconnection is designed to ensure confidentiality, integrity, and availability of the communication. The application environment mainly focuses on the user's access control and audit, and the security problems that the application system produces during execution.

6.3 Key management mechanism

Key system is the foundation of security and is one of the means to realize the privacy protection of perceived information. Restrictions on the Internet because there is no computational resources, asymmetric and symmetric key system can be applied, the Internet security is mainly comes from its original design, open management mode is a kind of no strict management center network. Mobile communication network is a relatively centralized management network, while WSNs and sensing nodes have more requirements for key systems due to the limitation of computing resources.

There are two ways to achieve a unified key management system. Centralized management of the Internet. The Internet key distribution center is responsible for the key management of the entire IoT, once the sensor network access to the Internet, through the key center and the sensor network convergence point to interact, realize the key nodes in the network management. Distributed management in the center of each network. The key management system design of WSN is limited by its own features to a great extent, therefore on the design requirements and the wired network and the traditional resources of the wireless network is different, there is no limit on the special should fully consider the limitation of sensor nodes and network characteristics.

6.4 Data processing and privacy

IoT data including information perception, acquisition, convergence, fusion, transmission, storage, mining, decision-making and control process, at the end of the sensor networks, it almost involves the whole process of information processing, just because of the sensor nodes and the focal point of resource constraints, in terms of information mining and decision does not occupy the main position. IoT application not only facing the security of information collection, but also considering the illicit close sex of the information transmission, request information could not be tampered with and unauthorized users, at the same time, also considering the network of reliable, credible and safe. Whether or not the Internet of things can be widely extended depends on whether it can secure user data and privacy.

The data processing involves location-based services and privacy protection during information processing. ACM established SIGSPATIAL in 2008, dedicated to spatial information theory and application research. Location-based services are basic functions of the IoT. Positioning technology currently has GPS positioning, mobile positioning, WSN positioning, etc. WSNs are located primarily in rfid, bluetooth and ZigBee.

The privacy content of location-based services involves two aspects, location privacy and query privacy. Location privacy refers to the location of the user in the past or present in the position, and query privacy refers to the sensitive information query and mining, such as a user query often one regional restaurant or a hospital, can analyze the user's location, income, life behavior, health and other sensitive information, causing leakage of personal privacy information, query of privacy is the privacy protection problem in the process of data processing.

Table 3 Sensor network attack and solution

Attack types	Solution
External attack and link layer security	Link layer encryption and authentication
Sybil attack	ID authentication
HELLO flood attack	Two-way link authentication
Wormholes and sewage pools	Design routing protocol
Selective forwarding attack	Multipath routing techniques
Certified radio and flood	Broadcast authentication

6.5 Secure routing protocol

At present, the domestic and foreign scholars put forward a variety of WSN routing protocol, the routing protocol of the original design goals are usually is the smallest communication, computing, storage cost to complete the data transmission between nodes, but most of these routing protocols without considering the security problems. In fact, due to the limited capacity of wireless sensor nodes, limited computing capacity, limited storage capacity and deployment of the field, it is highly vulnerable to all kinds of attacks.

Attacks on WSN routing protocol basically include the following categories: false routing information attack and selective forwarding attack, sewage pool attack, the witch attack, wormhole attack, Hello flooding attack, confirm the attack, etc. In Table 3, we present the sensor network attack and solution.

6.6 Intrusion detection and fault-tolerant technology

The intrusion means that in the case of malicious intrusion, the network can still function normally. The security of WSN is because of the open nature of network deployment area and radio characteristics of wireless network, attackers often use these two features, hinder the normal work of the network node, and destruction of the whole sensor network operation, reduce the availability of the network. At present, the technology of wireless sensor network is mainly focused on

the network topology, the security routing and the intrusion mechanism of data transmission.

Another requirement for WSN availability is the fault tolerance of the network. In general, fault tolerance refers to the characteristic that the system does not fail and still works properly when the fault exists. Fault tolerance refers to the WSN after the part of the node or link failure, network to transmit data recovery or self-healing network structure, so as to minimize the function of WSN node or link failure. As a result of sensor nodes in energy, storage, computing power and communication bandwidth are limited, many aspects, such as and usually work in the harsh environment, sensor nodes in a network is often there will be a failure. Therefore, fault tolerance becomes an important design factor in WSN, and fault tolerant technology is also an important field of WSN research.

7 Decomposition of social network learning elements based on Sakai

Social network learning emphasizes the integration of three networks, the human brain network, the social network, the Internet. Social networks and the Internet is a social network learning environment, to the survival of Sakai platform provides learners with an online social network learning environment, platform integration of a variety of social software can effectively support the development of the social network learning, contain a variety of network information resources (such as email, blog, BBS, Web pages, etc.). Social network includes interpersonal relationship network, real world requires learners to actively participate in and share, there are many kinds of communication and collaboration tools in Sakai platform, learners will be able to fully communicate and share, can both small group collaboration and discussion, also can undertake a wide range of learning and sharing, to form an atmosphere of mutual assistance and sharing of social network environment.

The social network learning environment must have three basic functions that allow users to create and maintain relationships with friends. Upload and download content information. Share information with other users through browsing, sharing and sharing. Sakai is a user-centered, integrating management, social, free content creation of mutual learning environment platform, and its diversity, project-based learning tool integration between the participants in the case of do not need to be familiar with the knowledge of HTML, through Sakai provide collaborative learning tools to meet their learning needs, make learning more closely interaction, cooperative learning more effectively.

In the Internet age, the amount of information is huge. For learners, there is an unlimited choice of space. The wider the range of choice, the stronger the learner's individuality, so learning resources are the basis of social network learning. By analyzing the resources required by learners, this paper divides learning resources into three categories: basic knowledge resources, knowledge use resources and interests, and expanding resources.

With the help of Sakai platform for social network learning, the source of learning resources is the first problem to be solved. The basic knowledge resource is the most important resource for learners to learn on the Sakai platform. It mainly comes from the excellent courses or open courses designed by the teachers according to the teaching plan and curriculum arrangement. Knowledge, using resources and interests, and expanding resources are the resources for learners to learn after completing the basic knowledge. On the one hand through the platform provided by the link and subscription tool on the Internet related learning resources linked to the course site, one can put the latest information and learning resources for learners, and that learners can more easily find the resources needed. On the other hand, through the platform provided by resources, tools, teachers, teaching assistants or students can upload and share useful resources, and constantly enrich and improve the curriculum learning resource library. By delivering box tools, teachers can be more targeted to push personalized learning resources and learning circles or groups for students. The learners do not have to worry about how to search for their own learning resources from the infinite cyber source, through a variety of ways from the previous learners, teachers or other learners get more targeted resources on the platform of learner.

With the arrival of the era of Web 2, social networking technology has gradually matured, social network learning has received more and more attention of scholars at home and abroad, on the social network learning learning mode research system becomes very necessary. Sakai platform is a very representative of the open source network teaching platform, its good usability, sociality and scalability, and cooperation and sharing platform design idea is very consistent with the core concept of social network learning, suitable as a social network learning platform. In this paper, through the analysis of specific elements of the social network learning to construct learning model with modern learning characteristics of the social network, the social network learning model in the specific process of detailed discussion, and further carries out an empirical study on the social network learning, effectively demonstrates the social network model based on Sakai learning the practical value of the [25].

7.1 Intrusion detection and fault-tolerant technology

According to the analysis of the data in different ways, intrusion detection is divided into misuse intrusion detection and

anomaly intrusion detection, because of its advantages and disadvantages, therefore, on this basis, this paper proposes hybrid intrusion detection technology, to achieve the two complement each other, to improve further.

The precondition of misuse detection is to analyze the known attack and intrusion methods, and then get the characteristics of attack behavior and intrusion behavior, so as to form the intrusion rule database. In intrusion detection, firstly, log data to the host or network audit data are pre processed, and then with the intrusion rules match, if the match is successful then an invasion. Most of this research focuses on how to improve the accuracy of rule base and improve the speed of rule matching. For the generation of rule base, language analysis method and expert system are adopted, while string matching is used in rule matching.

It can be seen from the process of misuse detection that its deficiency is that it can only detect the known attack pattern, and can not do anything about the unknown intrusion. When the emergence of new vulnerabilities, you can only use the machine learning system or manual methods, the new intrusion patterns to increase the intrusion rule base, so the need for regular updates of the rule base, which is intrusion detection will lag behind the new intrusion, intrusion behaviors are prone to false negative cases.

7.2 Anomaly detection technology

Anomaly detection is based on data analysis technology, detect unknown attack modes and aggressive behavior, is usually the degree of deviation from the normal behavior and judgment to determine whether an intrusion behavior, so there is no need to pre generate intrusion rule base. The techniques used in anomaly detection include statistical analysis model, neural network, immune system method, Bayesian inference theory, and document completeness check. From the analysis of the process of anomaly detection, we can see that there exists a false detection in the anomaly detection. In addition, the exception detection does not need to build up the intrusion rule base based on the intrusion behavior characteristics, but it needs to establish the rule base of the user's normal behavior. At the same time, the user's normal behavior also changes at the moment, so the abnormal detection will be faced with regular behavior patterns updated regularly.

7.3 Hybrid detection technique

Hybrid detection technology synthesizes the advantages of misuse detection and anomaly detection, overcomes their shortcomings, so it can detect known intrusion behaviors, also can detect unknown intrusions, greatly reduce the intrusion of false positive and false negative proportion.

8 System design and implementation

8.1 Hardware modular design of Sybil intrusion prevention system

Based on Sybil intrusion prevention system's overall design and the design of intrusion detection algorithm, we Sybil intrusion defense system hardware design and software development, system of modularization design mainly includes the filter circuit module, main control circuit module, AD module and detection circuit module. The Sybil intrusion detection feature is matched by the feedforward modulation filter, and the filter circuit is constructed.

The Sybil intrusion signal for the IoT is the original input and a simple filter form is given:

$$H(z) = \frac{N(z)}{D(z)}, \tag{12}$$

where $N(z)$ is a low-communication function of the Sybil intrusion defense system, $D(z)$ is the initial state of the equivalent low communication channel. The starting frequency and initial phase of the feedforward modulation filter can be calculated according to frequency parameter and bandwidth parameter of filter:

$$\omega_0 = \arccos(-a/2). \tag{13}$$

When the measurement noises are not correlated with each other, the Sybil intrusion detection feed forward filter is obtained by weighting, and the high-frequency response characteristic function is:

$$e^{j\pi} = V(e^{j\omega_0}) = \frac{\sin\theta_2 + \sin\theta_1(1+\sin\theta_2)e^{j\omega_0} + e^{j2\omega_0}}{1+\sin\theta_1(1+\sin\theta_2)e^{j\omega_0} + \sin\theta_2 e^{j2\omega_0}}. \tag{14}$$

As a result, the transfer function of the feedforward modulation filter of the Sybil intrusion prevention system is:

$$H(z) = \frac{1}{2}[1+V(z)]V(e^{j\omega}) + e^{j\varphi(\omega)} \tag{15}$$

The constraint condition is:

$$TW \ll \frac{c}{2|v|}, \quad \left|\frac{2v}{c}\right| \ll 1. \tag{16}$$

The largest output response feature can satisfy the performance requirements of Sybil intrusion detection at this time. Bayesian inference is an inductive reasoning, discovered by the British clergyman bayesian, as an inference method, and bayesian inference is expanded from the bayesian theorem of probability theory.

Assume we have a event set, denoted as:

$$B_i (i = 1, 2, \ldots, k) \tag{17}$$

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_n)P(A|B_n)}. \tag{18}$$

We can select the characteristic values of different aspects of the network system, denoted as B_i. By measuring the B_i variable values at different moments in the network system, the B_i variables have two values, 1 is abnormal and 0 is normal.

Reliability and sensitivity representation of each variable are denoted as $P(B_i|1/A)$ and $P(B_i|1/\neg A)$, thus we can calculate credibility:

$$P(A|B_1, B_2, \ldots, B_n) = P(B_1, B_2, \ldots, B_n|A) \frac{P(A)}{P(B_1, B_2, \ldots, B_n)} \tag{19}$$

$$P(B_1, B_2, \ldots, B_n|A) = \prod_{i=1}^{n} P(B_i|A) \tag{20}$$

$$P(B_1, B_2, \ldots, B_n|\neg A) = \prod_{i=1}^{n} P(B_i|\neg A). \tag{21}$$

Thus we have:

$$\frac{P(A|B_1, B_2, \ldots, B_n)}{P(\neg A|B_1, B_2, \ldots, B_n)} = \frac{P(A)\prod_{i=1}^{n} P(B_i|A)}{P(\neg A)\prod_{i=1}^{n} P(B_i|\neg A)} \tag{22}$$

According to the above mentioned, the probability of intrusion attack can be determined according to the value of various abnormal tests, the prior probability of invasion, and the anomalous probability of each measured value in intrusion. To test the accuracy of the result, also need to consider the various abnormal independence between measured value, at this time can through the network layer in the different characteristic values of correlation analysis, to determine the variable relationship with the invasion of attack.

8.2 CIDF standard framework

In order to make different research institutions and development company IT IDS intrusion detection system can be compatible with each other, and improve the scalability of IDS system, UC Davis security laboratory developed a general framework of intrusion detection system. The framework has become an implementation standard of intrusion detection research.

Event generator: event from intrusion detection system outside the computing environment, and provide this event to other components of the system to the CIDF *gidos* format, *gidos* are common intrusion detection object format, a *gido* can express some specific events in some specific time, also can express some conclusions drawn from a series of in the event, and even can be said to perform the action instruction.

Event analyzer: receive *gidos* from other components, analyze the resulting data, and generate new*gidos*.

Response unit: the function unit that makes a predetermined response to the analysis result can terminate the process, reset the connection, and change the file properties and so on, and also can only realize the simple alarm.

Event database: a storage unit that stores various intermediate and final data. It can be either a database format file or a simple text file.

8.3 Intrusion detection technology based on cloud computing platform

Because of the intrusion detection system needs a lot of computing resources and storage resources, computing resources and storage resources to become the bottleneck of the intrusion detection system, but cloud computing platform has a large number of computing resources and storage resources, to solve the bottleneck of resources in intrusion detection system, improve the real-time and efficiency of intrusion detection and intrusion; the detection system is the main focus of the attack, and the cloud computing system of distributed devices are hidden in the clouds, it is difficult to track, so the deployment of intrusion detection system can hide the actual position of the intrusion detection system in the cloud, the attacker is not easy to intrusion detection system for tracking attacks, intrusion detection system to ensure its safety.

8.4 Secure routing protocol for Internet of things

The routing of the IoT has to go through many kinds of networks. There are routing protocols based on IP addresses, routing algorithms based on identification of mobile communication networks and sensor networks, so we should solve at least two problems. Routing problems in multi network fusion and routing problems in sensor networks. The former can consider mapping identity tags into similar IP addresses, and implement an address based unified routing system. The latter is due to the limitations and vulnerable attacks of the computing resources of the sensor network, so an anti attack secure routing algorithm should be designed.

At present, domestic and foreign scholars have proposed a variety of routing protocols in WSNs, the routing protocol is usually the original design goals with minimal communication, computation and storage overhead to complete data transmission between nodes, but these routing protocols are not taking into account security issues. In fact, wireless sensor nodes are vulnerable to various attacks because of limited

power, limited computing power, limited storage capacity and deployment of the field.

8.5 Authentication and access control of Internet of things

Authentication in network mainly includes identity authentication and message authentication. Authentication enables both parties to trust each other's identity and exchange session keys. Confidentiality and timeliness are two important issues in authenticated key exchange. To prevent counterfeiting and session key leaks, important information such as user identification and session key must be transmitted as ciphertext, which requires prior master keys or public keys for that purpose. Because there may be message playback, timeliness is very important, and in the worst case, an attacker can use replay attacks to threaten session keys, or succeed in impersonation of another party.

The authentication process in the Internet, sensor network authentication mechanism is an important part of the study, in the WSN authentication technology mainly includes the public key authentication technology based on lightweight and pre shared key authentication, random key pre distribution authentication technology, using auxiliary information, authentication based on hash function authentication.

Authentication technology based on lightweight public key algorithm. In view of the classical public key algorithms require high computation in wireless sensor cyber source Limited is not operational, some current research is being devoted to the public key algorithm to optimize the design so that it can adapt to the WSN, but in the energy and resources still have a lot of room for improvement, such as authentication based on TinyPK scheme of RSA public key algorithm, and based on identity authentication algorithm.

Authentication method based on one-way hash function. This method is mainly used in broadcast authentication. A key chain is generated by one-way hash function, and the key can be unpredictable by the irreversibility of one-way hash function. In some way, the key in the key chain is published sequentially, so that the message can be authenticated. At present, the broadcast authentication method based on one-way hash function is mainly the improvement of TESLA protocol. The TESLA protocol is based on the TESLA protocol, and it improves the key update process and initial authentication process, so that it can be implemented effectively in wireless sensor networks [26–28].

9 Experiments

In order to evaluate the effectiveness of the proposed method, the KDD-CUP99 [29] data set is used and the Matlab simulation platform is used to evaluate the constructed intrusion detection model. The KDDCUP99 data set is divided into training data and test data, and the training set is classified according to Normal, DoS, Probing, R2L, and U2R5, wherein each instance contains 41 features. 1 training sets are randomly selected, including 10,000 instances, and then five test sets, each containing 10,000 instances.

Table 4 Components of dataset

Type	Mode	Meaning
DoS	Smurf, neptune, teardrop, land	Access denied server
U2R	buffer_overflow, load-module, p-erl	Illegal access
R2L	ftp_write, guess_pass-wd, imap, multihop, phf, warezmaster	Illegal access from remote machines
Probing	Ipsweep, nmap, satan, portsweep	Surveillance and other detection activities
Normal	N/A	Normal record

In order to cluster data with similar components into one class, the K-means clustering algorithm is used to cluster the samples, which involves the selection of k values. In Table 4, we present the components of dataset.

Silhouette values are important metrics to measure the effect of classification. When $k = 5$, the sample is divided into five classes, a considerable part of the silhouette data value is less than 0.6 or even 0.4, this shows that the clustering effect is poor, it will increase the amount of the main PCA after the treatment. The effect of reducing dimension effect, and increase the number of features need to deal with the intrusion detection process. When the k value is increased, the silhouette value and the clustering effect can be improved. But the bigger the k, the system calculated during the sample training volume increases, will reduce the system efficiency of computing and Realization of lightweight and with loss; the increase of k value will appear to have the same characteristics as other conditions, increasing the amount of calculation system. Figure 3 shows the results.

The overall clustering time increases with the increase of K value. The larger the k value is, the longer the computation time is, the larger the corresponding computation is. By repeatedly simulation comparison found that when $k = 12$ can also achieve good clustering time and clustering effect, will not be too low, clustering imprecision silhouette value, limited constraints combined with networking resources of sensor nodes, thus $k = 12$. Figure 4 shows the simulation set.

In order to evaluate the detection effect, the detection time, detection rate and false positive rate are used as evaluation indexes. Detection time is used as a lightweight measure, and detection rate and false positive rate are used to evaluate the

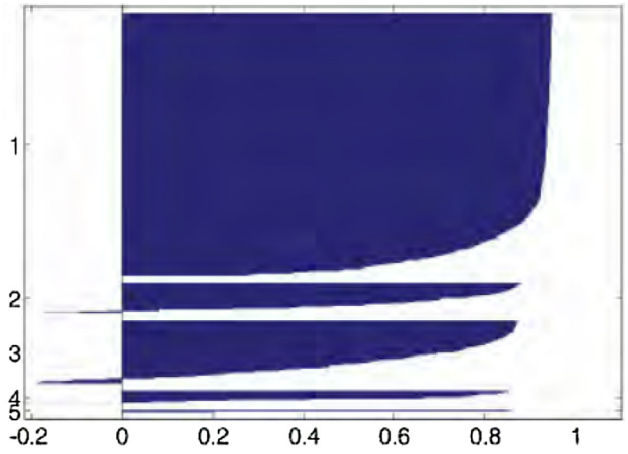


Fig. 3 Silhouette value when $k = 5$

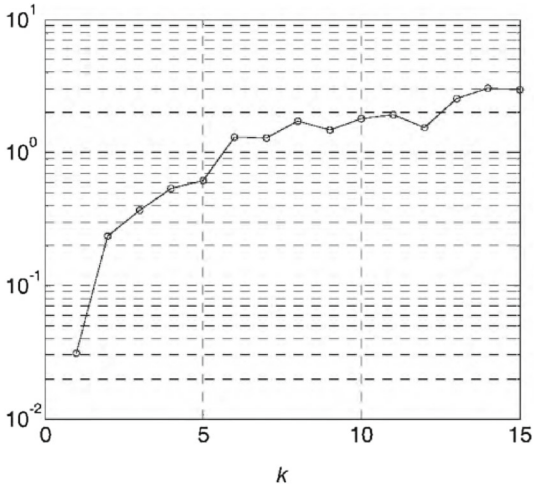


Fig. 4 Corresponding relation of clustering time and k

Table 5 Statistical results of detection accuracy

Algorithms	Characteristic number	Detection rate	False positive rate
Proposed method	8–16	96.8%	1.6%
Filter	41	92.5%	2.1%
Wrapper	41	95.3%	1.7%

detection accuracy of this method. This method compared with the traditional filter or wrapper model, will need to deal with the characteristics of a reduction from 41 to 8–16, the detection rate reached 96.8%, the error rate is 1.6%, which is also reduced in the number of features has maintained a high detection rate and low false alarm rate, and achieves the design goal of lightweight intrusion detection method, in Table 5, we list the statistical results of detection accuracy.

In order to test the performance of the designed IoT-oriented Sybil intrusion prevention system, system debugging and simulation experiments, the design of the detection

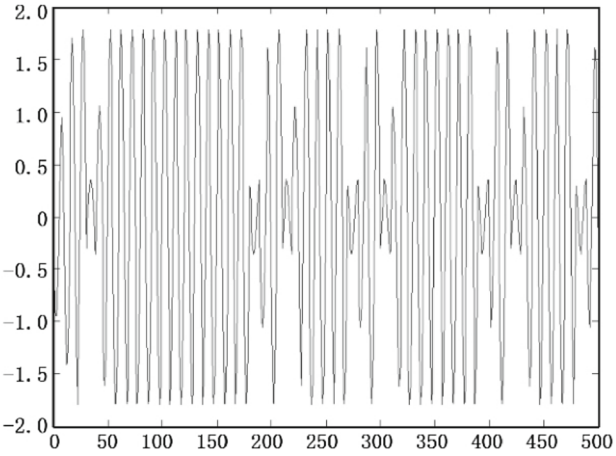


Fig. 5 Raw data waveform

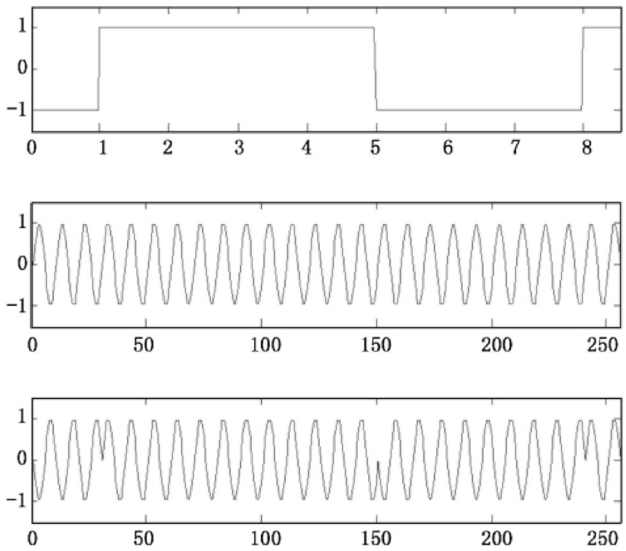


Fig. 6 Signal separation results of Sybil intrusion detection

algorithm using Matlab programming. The frequency of the invading signal center is: $f_0 = 1000$ Hz. The discrete sampling rate of intrusion data is $f_s = 10 \times f_0$ Hz = 10 kHz, the bandwidth of serial port control is $B = 1000$ Hz.The filter parameters of Sybil intrusion prevention system are chosen as: $\phi = \phi_0 = 0.5$, the order of the interrupt vector is 24, and the iterative step length of intrusion detection is 0.01. Time domain waveform of raw data for Sybil intrusion detection is shown as follows (Figs. 5, 6, 7).

It can be seen from the diagram, interfered in the network environment, it is difficult to detect Sybil intrusion signal, using this method to detect Sybil intrusion. The signal separation results of network intrusion detection are shown in following figures:

By using this method to design the intrusion prevention system, Sybil, embedded design to networking, data transmission, through accurate detection information, improve the

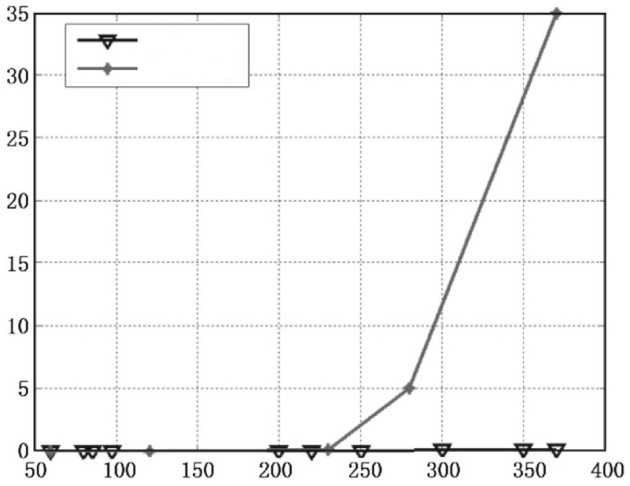


Fig. 7 Signal separation results of Sybil intrusion detection

network security performance, reduce the packet loss rate. Through tests, we know that the design of this system has strong compatibility.

10 Conclusion

The IoT is a new generation of information technology, which has become a hot research topic of governments and organizations all over the world, and its security problems have been attracted more attention. Intrusion detection technology is an important method to ensure network security, which has become a hot topic in the study of the security of IoT. Based on the technology framework of the IoT, this paper expounds the components of the IoT, and briefly analyzes the security threats and security requirements of each part, and probes into the intrusion detection technology deeply. Finally, some possible future research directions are proposed, and the future work is prospected.

The security and privacy protection of the IoT is the key to the IoT services, and the multi-source and heterogeneous nature of the IoT poses a huge challenge to the security of the IoT. The Internet and mobile communication networks and other social networks have established some effective mechanisms and methods, which provides abundant information resources for our life and work schedule, and changes people's way of life and work.

This paper proposes a new intrusion detection scheme for Internet of things, that is, lightweight intrusion detection method combined with FCM algorithm and PCA algorithm. Simulation results show that the proposed method can improve the detection efficiency and make the false positive rate lower.

Acknowledgements This paper is supported by the The Project of Macau Foundation (No. M1617): The First-phase Construction of Big-Data on Smart Macao.

References

1. Zhou, J.: Wireless sensor network intrusion detection model research. In: CIE 16th Information Theory Academic Conference Proceedings, pp. 799–804. Electronic Industry Press, Beijing (2009)
2. Denning, D.E.: An intrusion-detection model. IEEE Trans. Softw. Eng. **SE–13**(2), 222–232 (1987)
3. Xinghua, L., Xiaojun, Z.: Multi person cross region loss prevention method in personnel image tracking process. Comput. Simul. **31**(9), 243–246 (2014)
4. Tao, Y., Hu, B., Gao, X., Hou, X.: Research on dynamic tracking compensation method of hyperspectral interference imaging. Acta Photon. Sin. **45**(7), 716–723 (2016)
5. Xinghua, L., et al.: Research on software design of smart home control system based on Android. Internet Things Technol. **35**(5), 692–695 (2015)
6. Kun, Z., Meng, X.: Research and prevention measures of computer network security. Fujian **10**, 102–103 (2009)
7. United States General Accounting Office: Computer Attacks at Department of Defense Pose Increasing Risks. GAO/AIMD-96-84 Defense Information Security, Washington DC (1996)
8. United States General Accounting Office: Opportunities for improved OMB over sight of agency practices. GAO/AIMD Information Security, Washington DC (1996)
9. Conti, J.P.: The Internet of things. Commun. Eng. **4**(6), 20–25 (2006)
10. ITU: The Internet of Things. [EB/OL] (2005-12-17) [2010-07-03]. http://www.itu.int/internetofthings
11. Xiaon, W.A.N.G., Guiying, W.E.I.: Cipher algorithm in data transmission of RFID system on the Internet for things. J. Beijing Inf. Sci. Technol. Univ. **24**(4), 25–78 (2009)
12. Weber, R.H.: Internet of things—new security and privacy challenges. Comput. Law Security Rev. **26**, 23–30 (2010)
13. Leusse, P., Periorellis, P., Dimitrakos, T., Nair, S.K.: Self managed security cell a security model for the Internet of things and services. In: Proceedings of the First International Conference on Advances in Future Internet, pp. 47–52. IEEE Computer Society, Washington DC (2009)
14. Mulligan, G.: The Internet of things: here now and coming soon. Internet Comput. **1**, 36–37 (2010)
15. Zheng, S.Q., Han, Y.J., Zhang, Q.: Architecture and application of IOT. Softw. Ind. Eng. **6**(6), 27–31 (2010)
16. Wu, C.K.: Initially search on security architecture of IoT. Proc. Inf. Security **25**(4), 411–419 (2010)
17. Zhu, J.M., Ma, J.F.: An intrusion detection data collection model based on multi-agents. Comput. Appl. Res. **1**, 103–105 (2004)
18. Zhao, P., Wang, F., Xiao, X.C.: An intrusion detection technology based on multi-agents and its realization. Comput. Eng. Sci. **23**(6), 39–42 (2001)
19. Liu, W.T.: Network intrusion detection based on improved genetic algorithm. J. Chongqing Univ. Commerce (Nat. Sci.) **27**(5), 476–480 (2010)
20. Luo, H.W.: Network intrusion detection system and performance indicators. Telecommun. Netw. Technol. **11**, 24–26 (2005)
21. Mukkamala, S., Sung, A.H.: Identifying significant features for network forensic analysis using artificial intelligent techniques. Int. J. Digit. Evid. **1**(4), 1–17 (2003)
22. Lei, Y., Huan, L.: Feature selection for high dimensional data: a fast correlation based filter solution. In: Proceedings of 20th International Conference on Machine Learning

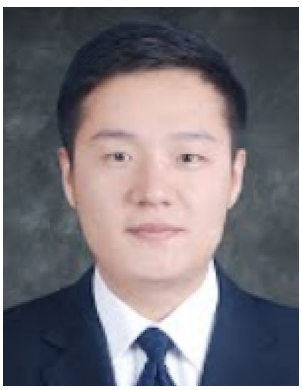
23. Yu, L., Liu, H.: Feature selection for high-dimensional data: a fast correlation-based filter solution. In: Proceedings of the 20th International Conference on Machine Learning, pp. 856–863. Morgan Kaufmann, San Francisco (2003)
24. Mac, Q.J.B.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, pp. 281–297. University of California Press, Berkeley (1967)
25. Martin, G., Smart, N.P.: Distributing the key distribution centre in Sakai-Kasahara based systems. In: Parker, M.G. (ed.) Cryptography and Coding. LNCS, vol. 5921, pp. 252–262. Springer, Heidelberg (2009)
26. Ozdem, I.R.S.: Secure and reliable data aggregation for wireless sensor networks. In: Ichikawa, H. et al. (eds.) UCS 2007. LNCS, vol. 4836, pp. 102–109. Springer, Berlin (2007)
27. Wood, A.D., Fang, L., Stankovic, J.A.: A Family of Configurable, Secure Routing Protocols for Wireless Sensor Networks, pp. 35–48. ACM, New York (2006)
28. Liu, D., Ning, P.: Multi level TESLA: a broadcast authentication system for distributed sensor net works. ACM Trans. Embed. Comput. Syst. (TECS) **3**(4), 800–836 (2004)
29. Hettich, S., Bay, S.D.: KDD cup 1999 data [EB/OL]. http://kdd.ics.uci.edu/databases/kdd-cup99/kddcup99.html. Accessed 23 Sept 2014

Lianbing Deng is the Director & General Manager of Zhuhai Da Hengqin Science and Technology Development Co., Ltd. And he is also the Director of Post-doctoral Programme of China (Hengqin) Pilot Free Trade Zone and the Director of Information Centre of Hengqin New Area. He has received his doctor degree in Huazhong University of Science and Technology. His researches are in the field of big data, project management, and economic research. He is the Vice Chairman of China Big Data Council of MIIT of the People's Republic of China.

Daming Li is currently a research fellow at Post-doctoral Research Center of China (Hengqin) Pilot Free Trade Zone and Information Centre of Hengqin New Area. He is also a course consultant of City University of Macau. He has received his doctor degree in City University of Macau in 2014. He is the membership of International System Dynamics Society of the State University Of New York. His research interests focus on the big data technology, smart city, and quantitative analysis.

Xiang Yao is currently an architect in China (Hengqin) Pilot Free Trade Zone and Information Centre of Hengqin New Area. He was received his bachelor degree in Heilongjiang Institute of Technology in 2007. His research interests focus on the Cloud, Big Data, IoT, Block Chain, ML technology and System analysis.

David Cox is an Assistant Professor of Computer Science, and is a member of the Center for Brain Science at Harvard University. He completed his Ph.D. in the Department of Brain and Cognitive Sciences at MIT with a specialization in computational neuroscience. Prior to joining MCB/CBS, he was a Junior Fellow at the Rowland Institute at Harvard, a multidisciplinary institute focused on high-risk, high-reward scientific research at the boundaries of traditional fields.



Haoxiang Wang is currently the director and lead executive faculty member of GoPerception Laboratory, New York, USA. His research interests include multimedia information processing, pattern recognition and machine learning, remote sensing image processing and data-driven business intelligence. He has co-authored over 50 journal and conference papers in these fields on journals such as Springer MTAP, Cluster Computing; Elsevier Computers & Electrical Engineering, Optik, Sustainable Computing: Informatics and Systems, Journal of Computational Science, Pattern Recognition Letters, Information Sciences, Future Generation Computer Systems and conference such as IEEE SMC, ICPR, ICTAI, CCIS, ICACI.