Research of distributed IDS based on mobile agent and genetic algorithm

Weimin Gao[1]*, Lizhen Xiao[2]

[1]*School of computer and information science, Hunan institute of technology, Heng yang 421002, China*

[2]*School of electronics information, Hengyang finance economics and industry polytechnic, Heng yang 421001, China*

Received 1 July 2014, www.cmnt.lv

Abstract

The special radiation and openness of the propagation channel during wireless communication will lead great threats to security of network management and communication. In recent years, there are more and more application of genetic algorithm and the mobile agent in IDS. As traditional knowledge based IDS has to build artificial rules and patterns from expert of field with human interventions, limitations of expert rules will be highlighted with the change of time and space, resulting in unsatisfied detection correctness and effectiveness. As a result, we need to optimize the performance of IDS. In this paper, we first introduce the mobile Internet network architecture and security problems, and put forward a general IDS model and classification, then design a intrusion detection system based on mobile Agent and genetic algorithm, with flexibility, scalability and strong adaptability and low error rate, which meets the needs of mobile IPv6 environment to use. Experimental results show that the proposed design model has advantages in the performance of the detection efficiency, which is suitable for mobile network.

Keywords: mobile internet network, intrusion detection system, mobile agent, genetic algorithm

1 Introduction

With the rapid development of the information network technology, people are no longer satisfied with using a fixed terminal, or a single mobile terminal which is connected to the Internet, but want to move subnets in a relatively stable and reliable style, and dynamically access information from the internet, which promote the development of wireless mobile Internet. However, the special radiation and openness of wireless communication in the wireless space communication channel will lead to: counterfeit attacks, network fraud, theft of information and other unsafe behaviours, making the security of network operations and communications suffer a great threat. To prevent the information of wireless communications to be easily intercepted by some others outside the receiver, or intruders to deceive access etc, it is required to take a series of security measures.

An intrusion detection system (IDS) is a universal technology of dealing with security issues, such as tampering, deleting and stealing acts. There have been many scientific research institutions and manufacturers in the intrusion detection system to carry out the work of the practice feasible research and development work. For example: the Iowa State University development of MAIDS (Mobile Intrusion Detection System), Cisco company NetRanger, Network Associates company CyberCOP, "Ice Eye" of NSFOCUS, and so on. The traditional knowledge-based IDS needs to manually create rules and models by experts in the field, and complex network over time and space of the changes will result in the limitations of highlighting of the expert rules database, which reduces the validity and correctness of the IDS.

The distributed IDS which is suitable to modern network feature is a research hot spot in the information area, but there are some usual problems that exist in modern distributed performance IDS. The defaults of modern IDS are as follows: bad real-time of the intrusion detection and response, bad system extension and flexibility, lack of studying and dynamic configuration ability and easily to cause network blocking.

At present, the genetic algorithm is widely applied to IDS research, and the methods are also emerging endlessly. From some references we can see: genetic algorithm can improve the detection efficiency of IDS, reduce the rate of false positives, as well as remove the useless items of analysis, which makes the optimized performance of IDS.

This paper proposed an IDS detection scheme for the security problems of the mobile internet network, which described a system structure, communication theory of implementation, and deployment and intrusion detection process in the network layer. And the performance and the execution efficiency of the scheme are analyzed through building experimental platform.

2 The network structure and security issues of mobile internet

Usually it can be considered that mobile internet network uses mobile phones, personal digital assistants, portable computers, dedicated mobile Internet network equipment, etc as terminal, uses the mobile communication network or

* ***Corresponding author' s*** e-mail: gwmhy@163.com

wireless local area network (LAN) as a means of access, and visits the internet through the Wireless Application Protocol (WAP) and the Internet business. Mobile Internet network structure is shown in Figure 1.Due to the core of the mobile Internet network is wireless network technology, wireless network openly transmit data in the air through radio waves, almost all the mobile users in the coverage area of the data transmitter can contact the data, and therefore it is vulnerable to malicious attacks.

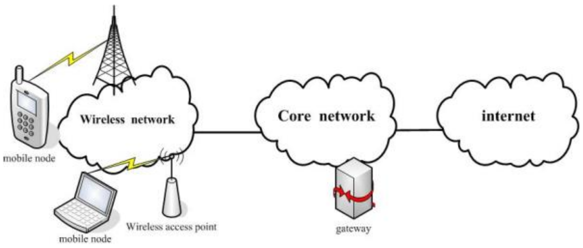

FIGURE 1 Mobile internet network structure

With the development of 3G and mobile Internet network, all the security issues on the Internet are likely to reappear in the mobile Internet;

On the other hand, the characteristics of the mobile Internet, the unique development way and the dissemination capability make a lot of new security problems to appear constantly [1]. As shown in MCAfee's "mobile security research report in 2008" more than 80% of the users concerned the mobile terminal security issues. Basically, the security problems of the current mobile Internet have mainly the following four aspects:

1) Information interrupt: the use of illegal means to attack the network availability; damage the software and hardware resources in the mobile Internet system, making the network work improperly.

2) Tampering: attacking the integrity of network, tampering with the core network elements of the mobile network and business database contents, modify the order of the message to delay or change the message.

3) Eavesdropping: through the wireless network transmission link of line and electromagnetic leakage, etc attack the confidentiality of network, causing leaks, or to the business flow into the information is extracted by precise, causing some privacy information disclosure and illegal criminal activities. Line analysis, to obtain useful information. There are many criminals eavesdropping and other available technical means.

4) Forgery: attacking the authenticity of the network, the forgery, false information injects into the network, counterfeit legitimate users access to mobile networks, resetting the intercepted legal information to achieve illegal purposes, transplanting malicious programs such as worms, Trojan, logic bombs, etc. to the mobile network to disrupt the normal working of the mobile network, denied the messages receiving or sending in the network.

3 Model and classification of the intrusion detection system

Intrusion Detection System (IDS) [2] is a network security technology of initiative protecting themselves against attacks; it is a kind of complement of firewall technology. Every complete IDS must support two functions: information control and information capture. Information control represents a rule, which is security and defence personnel who must be able to determine where the IDS itself packets can be sent. Its purpose is that after the decoy environment within the IDS was invaded, it will not be used to attack outside decoy environment of the machine and the organization. Information capture is to capture all the data flow of the invaders. Intrusion Detection System is considered to be the second security gate behind the firewall, without causing network performance to monitor the network, preventing or mitigating the threat of intruders to the network.

Intrusion detection is the detection of computer network and system to discover the process of violating the security policy events [3]. An intrusion detection system should include at least three functional modules: providing information sources of the event record flows, finding analysis engine of the signs of intrusion and the response component which based on the analysis engine. Figure 2 describes a general model of intrusion detection system, which is constituted with three parts: the event generator, activity logger and rule sets. Event generator produce activities in IDS model, the audit tracking network packets. Activity monitoring of recorder in the current state of the network. Rule set is a common engine that inspects and verifies event and state, which uses the model, rules, patterns, and the results to judge the intrusion behaviour. In addition, the feedback is an important part of the IDS model. Rule sets assembly according to generator feedback current events, trigger a system study to join the new rules or amend rules.

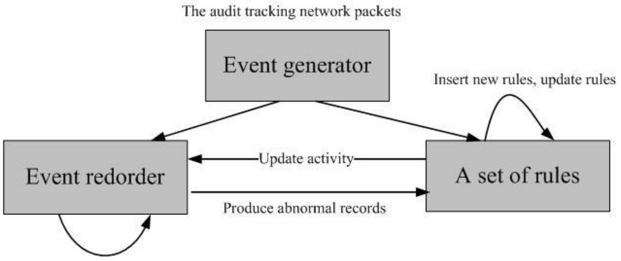

FIGURE 2 Universal model of IDS

From the perspective of data communication and processing, IDS is a typical data processing, which through a large number of audit data to analyze to determine the detected system whether is suffered the invasion. Its detection mechanism is a classification process of the main body behavior of the system, namely the malicious behaviour of the system which will be separated from the large amount of system behavior. Obviously, the key technology to solve the problem is how to obtain the normal behavior knowledge or knowledge about invasion behavior from the known data.

The intrusion detection systems have different classification criteria, according to the data source, can be divided into: host-based IDS with network-based IDS. According to the system structure is divided into: centralized intrusion detection and distributed intrusion detection. According to the principle of detection is divided into: anomaly intrusion detection and misuse intrusion detection. In the IDS study, the key technologies involved are: pattern matching, data mining, neural network, protocol analysis, data fusion and immune and various classification algorithms, but generally start from abnormal intrusion detection model and misuse intrusion detection model. Abnormal intrusion detection model uses the feature matching method to determine the attack, and the detection process is using a quantitative method to describe the acceptable behavior characteristics, to detect intrusion with characteristics of abnormal behavior. According to the pre-defined invasive pattern, misuse intrusion detection model means to observe the invasion situation and pattern matching, and generally, detect intrusion using the known system and application software weakness-attack modes [4-6].

4 Intrusion detection system based on the mobile agent and GA

4.1 MOBILE AGENT AND GA THEORY

The traditional intrusion detection method has some limitations duo to the use of specific detection methods and models for different environments. For example: update the attack feature database is not timely, the different between IDS and collaboration of other security technology are not strong, the high rate of false positives and false negatives, structure is unitary and so on, as a result, it is difficult to determine the real intrusion behaviour.

Mobile Agent technology [7, 8] can independently migrate between homogeneous or heterogeneous network hosts, which can blend distributed object technology, software Agent technology and mobile code technology in one body. It has the Agent's autonomy, responsiveness, intelligence and mobility, the mobile based IDS will make the Agent to distribute in the key points of the system, complete the distributed data collection, detection and real-time response. In addition, the mobile Agent is an independent functional entity and has good expansibility.

Mobile Agent can change the method of detection technology application, Increase the efficiency and effectiveness of diction. Mobile agent executing their tasks in a network and cooperate with other agents, so reduce the mobile giant data. It can be used to make a special agent for the intrusions of a specific type, because the agent can duplicate themselves, can the movement detection in multi-platform. At the same time, it can make the IDS to a maximum speed of response to intrusion detection, can improve the performance of system from many aspects, and improve the detection capability. The combination of mobile agent and intrusion detection system which can make IDS has the following capability:

- Responding to target machine;
- Collecting information;
- Reducing network traffic.

Mobile Agent is an actual program which can migrate independently from the source host to the target host in the running process and can interact with other Agents or resources. Mobile Agent model is shown in Figure 3.

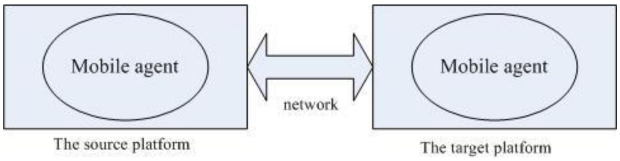

FIGURE 3 Mobile agent model

GA(Genetic Algorithm) is a highly developed parallel randomized, adaptive search algorithm which references biological laws of evolution, simulates Darwin's evolutionary mechanism of natural selection and genetic mechanism of the biological evolution. It uses the data sequence to represent the genetic; breeding is divided into two independent steps: crossover and mutation to proceed. The algorithm process is as follows:

1) Initialization: determine the population size N, gene crossover probability PC, gene mutation probability PM and terminate the evolution rules, generating random N individuals as the initial population X(0), and the evolution algebra counter t=0.

2) Individual assessment: calculating the fitness of specific individual N(i) (0<=i<N), the fitness of the individual refers to in certain environmental conditions, a known genotype individual will pass its genes to its offspring genes which the library relative capacity is a measure of individual survival and reproductive opportunities. Fitness is larger, the higher the survival and reproductive opportunities.

3) The process of population evolution:

- Select the father generation: the purpose of choice is to inherit directly the optimization of individual to the next generation or through matching cross to produce new individual and then genetic to the next generation. Select the N/2 pairs from the population to cross-breeding the next generation.
- Cross breeding is the core of the genetic algorithm: the so-called cross is the part of the structure of the two parent individuals to replace the restructuring to generate new individuals operation. From the paired parent, parts of them contribute a portion of the gene fragments consisting individuals. Parent gene fragments have a variety of combination ways and produce M progeny.
- Genetic variation: copy the parent to offspring and intercross the two parent genes, which produce the mutation offspring.
- Elimination choice: according to the fitness, select the number moderate offspring from the formed progeny group to a new generation of parent groups. As well as the genetic algebra counter also plus 1.If termination test did not pass, it should continue to the next generation evolution.

4) Termination of inspection: if termination conditions have been already met, the evolutionary process will be terminated. Termination conditions is to terminate

algorithm when the fitness of best individual reaches a given threshold, or the fitness of the best individual and group fitness rise no more, or the number of iterations reaches a preset algebra.

Genetic algorithm in IDS is mainly anomaly detection, and most is using training data to automatically calculate a threshold of the according network, as a condition to determine whether it is invaded.

4.2 SYSTEM DESIGN

Mobile Agent and GA based IDS establishes specific model based on some intrusion behaviors, designs multiple mutual coordinative Agent, analyzes the data from data collection Agent, collaboratively learns and trains rule data sets between different Agent, infers whether intrusion behavior appears with attack script. The structure of the IDS is shown in Figure 4.

Database module: it is used to store rule base, record events, original data and training data that system detected; to store variety of mobile Agent code bases as well as some specific events handler functions. Due to the large amount of network data, we need to create an identification labels for automatic classification, and to train data classification algorithm after the pattern matching.

Intrusion detection Agent module: capture the traffic packet on the network, then through protocol decoding and analysis to achieve the data packet statistics, operation logs and audit.

Response Agent module: it is used to respond the analysis results of the intrusion detection Agent. For example, intrusion is detected, the response agency to take corresponding measures, such as warning the intruder, prohibiting the connection or leading to the honey pot virtual system.

User interface Agent module: it is the interface of system and user, to be responsible for the user's commands and requests to send to the agent, and in the intuitive form to display alarms, and to put forward processed suggestions to users. The user interface Agent dynamically monitors the working states of kinds of the system components, allow the administrators to view, manage and maintain.

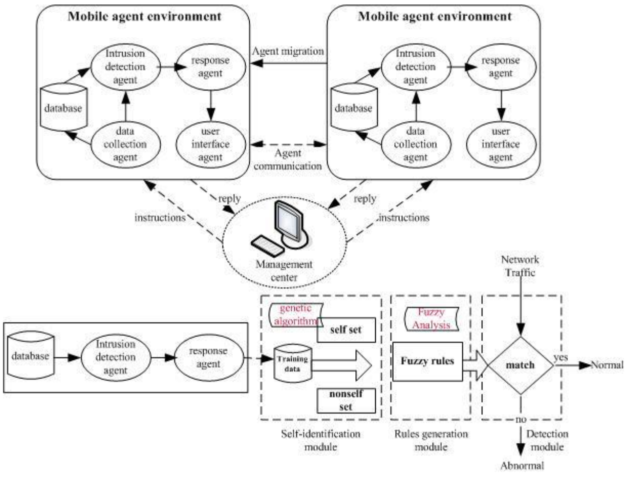

FIGURE 4 The IDS structure based on agent and genetic algorithm

Data collection Agent module: collects system and network data, user's activity state and behavior data. Generally it collects the information from several different key points in the network, which is to expand the detection range as far as possible, in addition, making that checking suspicious behavior and invasion is easily from several data source even if it can't been found from one data source. The data collection Agent specifically collects system data and network data for the monitored host, after the data and network interaction analysis between each Agent, the data will be sent to the intrusion detection Agent after pre-process.

Management center: communicates and releases instructions between mobile agents, and centralized manages the backstage database.

All above mentioned Agents are based on distributed mobile environment. The mobile Agent environment is the foundation of the invasion system, and it controls the mobile Agents' basic services such as the movement, establishment, and cancellation and so on. The mobile Agent environment proxy server could package the event handlers in the strategy library and produce mobile Agents which has been allocated to the correspondent nodes to implement the detection tasks.

4.3 SYSTEM DEPLOYMENT AND WORKFLOW

The mobile agent environment MAE should be installed on the workstations and servers of the network, and run different agents on them as needed. A management control center be configured on each segment to control the entire network segment.

System administrator can learn the running conditions of the whole system, and configure and control the invasion system through the user interface Agent - the IBM Aglets can be used as implementation platform, and then the predetermined security policy. Each tested mobile node or host runs different data collection Agent and analysis detection Agent. The data collection Agent runs on the key nodes of a network segment be used to collect the network packets, the deployment structure is shown in Figure 5.

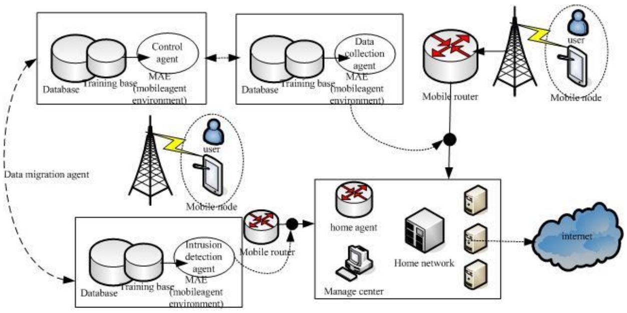

FIGURE 5 IDS deployment based on agent and genetic algorithm

Invasion detection Agent is the most important part of system deployment and design, it takes the task that checking whether there are invasions or not from the suspicious data information, and intercepting the network information, audit analysis, determine the suspicious degrees, then responding to the corresponding events. At the same time, the invasion detection Agent should keep communication with the other mobile agents, then broadcast the events which have reached a certain level of shadiness. These agents could migrate towards the heterogeneous network nodes, and analyze the invasion on the basis of the database information in the monitored host or information transmitted by other hosts, or rule and train the database and log by using genetic algorithm.

Under the mobile Agent and the genetic algorithm based distributed invasion detection system is safe, if the Agents are independent from each other under the premise to keep the Agents itself safe, but also move freely to coordinate in the network, then actualize the mobile strategy, and build a flexible and stable invasion detection system. The system work process is shown in Figure 6.

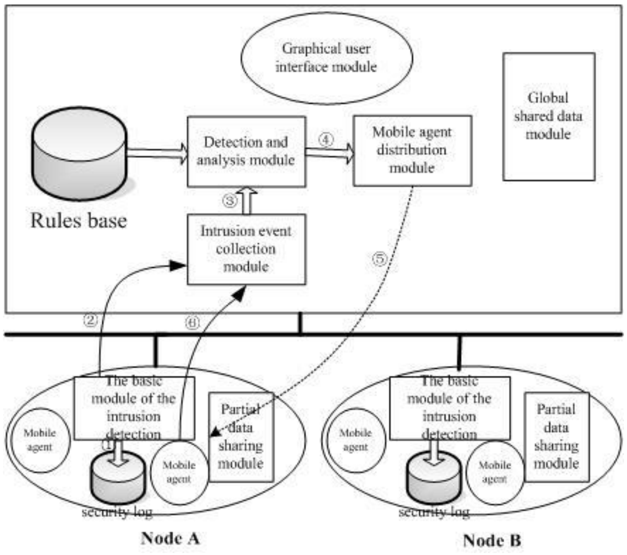
Graphical user interface module
Global shared data module
Detection and analysis module
④
Mobile agent distribution module
③
Rules base
Intrusion event collection module
⑤
②
⑥
The basic module of the intrusion detection
Mobile agent
①
Partial data sharing module
Mobile agent
security log
Node A
The basic module of the intrusion detection
Mobile agent
Partial data sharing module
Mobile agent
security log
Node B

FIGURE 6 IDS workflow

1) For each testing nodes in the network, the basic invasion detection modules detect the suspicious security incidents by checking the system logs and application logs.

2) The basic invasion detection module plays an important role in auditing and analyzing the standard data, detecting attacks or system abnormality, and reporting the result to the invasion collection module.

3) The invasion collection module sends all kinds of security events to the early-warning analysis module, so that the system can respond right.

4) The early-warning analysis module practice and deduce the invasion and its intention according to the invasion rule base.

5) The mobile Agent distribution module dispatches appropriate mobile agents to a specific detection node to collect information.

6) The mobile Agent gets local information from local share data module, and checks the specific data, and then the result will be send to the invasion collection module directly.

5 The experimental result and analysis

To test the model's performance and effectiveness, a testing environment based on the wireless local area network (WLAN) was built firstly, and there are 8 hosts in a WLAN, two installs data collection security tools snort and iptable, while one was the console host, one stored the central database host, one was target host, one was associate the engine host, one was policy server, one was data collector, and the last one was a collection agent. The internal network connect external network through a router, the mobile node in the external network was treated as an attack host, and the system was tested on a campus network owned by 3G polytechnic institute of Hengyang branch of China telecom. And software such as SQL server2009, WinPcap, Snort, Aglets, Visual studio 2005, Iptable were needed. Using the data of the first weekend in June, 2012 as the training data, and the data of the second weekend as the testing data, this contains the attack packets, and totally 7149800 data packets been collected. We encoded the data packets by genetic algorithm initialization coding method, there were 1243700 data packets left to be self set after the redundant data were removed. The data packets used in the experiment are as shown in Table 1.

TABLE 1 Experiment data

Data set	Training data	Testing data	
		normal	abnormal
IPv6 packets	7149800	5213489	11357

The standard to judge an invasion detection system is accuracy, which is presented by the false negative rate and false alarm rate. The parameters should be defined:

Detection rate=the number of invasion events been correctly detected/The total number of invasion events*100%;

False negative rate=the number of invasion events don't be detected/ the total number of invasion events*100%;

False alarm rate=the number of events been false reported/The total number of normal events*100%;

We compared the IDS' detection rate, false alarm rate with the open invasion detection system Snort, and the results are shown in Figure 7.

According to the results: the detection rate of IDS, which based on mobile Agent and genetic algorithm is superior to traditional Snort.

In order to prove that the population and the iteration time of genetic algorithm have influence on IDS, the simulation experiment is provided. As shown in Figure 8, with the increasing of population, the detection rate of the IDS system which based on mobile Agent and genetic algorithm will be improved, and at the same time, the false alarm rate will have a certain degree of decreasing.

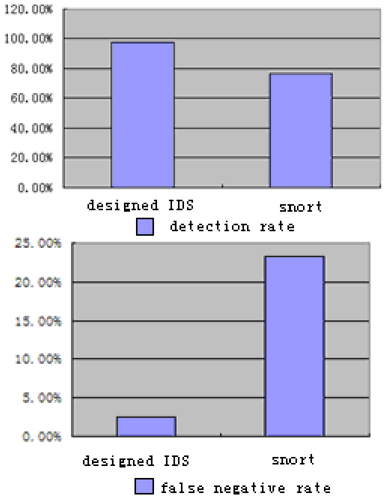

FIGURE 7 Comparison of experimental results

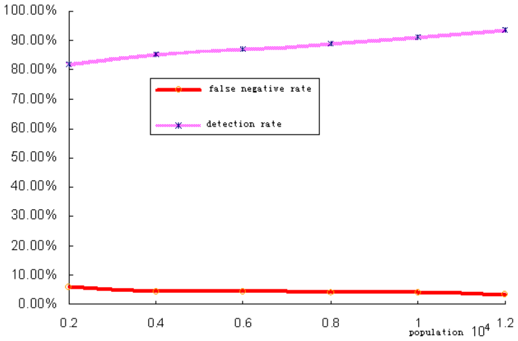

FIGURE 8 Influence properties of population

6 Conclusion

This paper put forward a method that applying mobile Agent and genetic algorithm to the distributed invasion detection in IPV6 environment, collecting security events by distribute the mobile agents dynamically, using the genetic algorithm to train data and calculate the network abnormal threshold adaptively. This proposed system has advantages like strong predictability, fast real-time response, high intelligence and fault-tolerant capacity, strong ability to resist attacks, and good cooperatively. The performance of the IDS system in the IPV6 network environment was validated, presenting advantages on algorithm performance and detection efficiency. And as a result, it is well suitable for the mobile internet.

Acknowledgment

This work was financially supported by National natural science foundation of china (61103245) Hunan province science and technology plan projects (2014FJ6026). Hunan institute of technology reform project (JY201425)

References

[1] Gao W, Hy X2006 The Research and Implementation of Network Invasion Induced Control System *Computer Measurement & Control* **14**(12) 1751-3

[2] Ma J, Ma H 2009 Analysis and Suggestion of Mobile Internet Security Problem *Modern Science & Technology of Telecommunications* **28**(7) 46-9

[3] Luo L, Zhou Z 2012 Research of network security invasion technology on IPV6 *Bulletin of science and technology* **28**(4) 114-5

[4] Chen J, He Z, Liang Y 2010 The design of IPV6 Network Invasion Detection System. Computer Technology and Development **20**(9) 123-6

[5] Zhao R 2009 Research and Implementation of Distributed Invasion Detection System on IPV6 network *Xi'an University of Electronic Science and Technology*

[6] Hou F 2005 Reserch and Design of Invasion Detection System in the Wireless Network *Shan Dong University*

[7] Xu S, Fu X 2007 The Application of Improved Genetic Algorithm in Invasion Detection System *Computer Systems & Applications*

[8] Zhao J, Gao Z, Jia S 2010 Improved Mobile IPv6 Switching Management Scheme *Communications Technology* **43**(12) 103-5

[9] Wang J 2008 Optimization Handover Scheme Based on HMIPv6 [D] *Wuhan, Wuhan University of Technology*

[10] Ren S, Cai R, Tang L, et al. 2010 Proxy mobile IPv6 based inter-domain mobility management approach and performance analysis *Application Research of Computers* **27**(3) 1118-21

Authors

	Weimin Gao, born in 1975, Hunan Province, Qidong country, China. **Current position, grades**: associate professor, Master, in School of Computer Science of Hunan Institute of Technology. **University studies**: Master degree in electronic engineering from Hunan University. **Scientific interest**: wireless sensor network, computer network and information security.
	Lizhen Xiao, born in 1967, Hunan Province, Hengyang City, City. **Current position, grades**: associate professor, Master, in School of Computer Science of Hunan Institute of Technology. **Scientific interest**: software engineering.