Distributed Intrusion Detection Model in Wireless Sensor Network

http://dx.doi.org/10.3991/ijoe.v11i9.5067

Hanqing Zhang
Hebei Institute of Foreign Languages Qinhuangdao, HeBei Province, China

***Abstract*—The security issues of Wireless Sensor Network (WSN) are significant, among which intrusion detection can improve the defense detection performance of WSN, and also balance the security and energy-saving accurately and efficiently. In this paper, we focus on the intrusion detection problem in WSN. Specifically, we propose a cluster-based collaborative detection structure, and the detection algorithm is based on immunity system and Ant Colony Optimization (ACO). The basic idea is to formulate intrusion detection as an optimization problem and introduce immune mechanism into ACO during iterations. Finally, the experiment shows that proposed algorithm outperforms other methods.**

***Index Terms*—Wireless Sensor Network (WSN), intrusion detection, immunity system, Ant Colony Optimization (ACO).**

I. INTRODUCTION

As a new generation sensor network, Wireless Sensor Network (WSN) is an information acquisition system composed of many wireless sensor nodes distributed in a given region. The wireless sensor nodes are typically limited in energy, bandwidth, and storage and computing. Compared with traditional wireless networks, WSN is self-organized, highly fault tolerant and reliable, low-cost, and easy to deploy, and has been applied to many areas, such as environment monitoring, disaster response, intelligent building, health care, housekeeping, business and industry [1].

However, the security issues of WSN have drawn significant attentions [2]. For example, the WSN equipment is typically exposed in severe environment, uninhabited areas or enemy positions; besides, wireless network is inherently vulnerable.

Currently there are many works on the security of WSN, including firewalls, data encryption and access control. However, works on intrusion detection among WSN nodes is not sufficient. WSN nodes are independent, and the behavior abnormal based on neighbor nodes can help to detect possible intrusions. Indeed, intrusion detection provides protection against internal and external attacks by properly distributing nodes in networks for important data, resources and networks in key fields. Intrusion detection can improve the defense detection performance of WSN, and also balance the security and energy-saving accurately and efficiently. Specifically, intrusion detection technology can discover and report unauthorized and abnormal behaviors of the system, and actively provide dynamic monitoring and protection as a part of the security mechanism.

In this paper, we propose a distributed intrusion detection model based on immunity principle [3] and Ant Colony Optimization (ACO) [4] for intrusion detection in WSN. The basic idea is to combine the variety of immune system and the fast detection algorithm based on ACO to deal with different types of intrusions in WSN. Accordingly, the security of WSN can be ensured, and also the energy consumption can be reduced, so that the lifetime of WSN can be improved. Besides, to improve the performance of ACO based intrusion detection, we employ K-Nearest Neighbor (KNN) algorithm [5] to eliminate the redundancy of data set at the initialization step. Our experiments show that proposed method exhibits good performance in intrusion detection in WSN.

II. RELATED WORK

Intrusion detection problem has been well studied. For example, Chen et al. [6] proposed an intrusion detection algorithm based on kernel Fisher discriminant analysis in WSN. Salmon et al. [7] designed an intrusion detection model using Danger Theory (DT), which utilized distributed collaborative mechanism and therefore improves the detection performance and reduces the energy cost. Sommer et al. [8] proposed an intrusion detection model with context using environmental context, weakness context, feedback context and abnormal context, etc. Zhang et al. [9] used statistics method for intrusion detection based on threshold. Misra et al. [10] detected intrusion with the objective of balancing energy cost. Huang et al. [11] designed a cluster-based model to deploy detectors onto nodes. Bao et al. [12] proposed a cluster-based hierarchical trust management protocol for WSN to detect malicious nodes. Butun et al. [13] provide a survey of intrusion detection systems in WSN.

Another category of related work is the improvement efforts of ACO. For example, Akay et al. [14] combined honeybee colonies with ACO to efficiently solve optimization problems. Tuba et al. [15] improved ACO with pheromone correction strategy for TSP problem. Yoo et al. [16] designed a new cooperation mechanism between ants and provided a new definition of pheromone. Ciornei et al. [17] combined Genetic Algorithm (GA) with ACO for global continuous optimization. Hao et al. [18] designed an immune ACO algorithm for path planning. In this work, we introduce immunity system into ACO for intrusion detection in WSN.

III. PRELIMINARIES

In this section, we briefly introduce the background on some algorithms that are employed in this paper.

A. *Immunity system*

Immunity system is a self-organized, distributed, self-adapted, self-learning and diversified system. Immunity system has been widely employed in information security [19]. Indeed, the characteristics of WSN such as dynamic topology structure, easy failure of nodes, and diversity of intrusions, make immunity system perfect fit for WSN.

Generally, immunity system simulates the antigen process in biological immunity. As shown in Figure 1, the process of immunity algorithm can be described as follows:

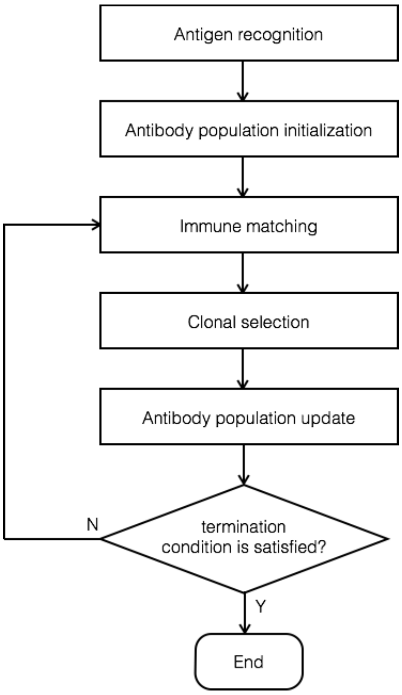
Antigen recognition
Antibody population initialization
Immune matching
Clonal selection
Antibody population update
termination condition is satisfied?
N
Y
End

Figure 1. Workflow of immunity algorithm

(1)Antigen recognition: or objective function and restrictions definition. Suppose a immunity based detection system $D=\{f,M\}$, where f is the classification function, M is the pattern set of intrusion detection. The model can be represented as:

$$f(M,s)=\begin{cases} normal & s\in M \\ abnormal & s\notin M \end{cases} , \qquad (1)$$

where s is the set of self.

(2)Antibody population initialization: the set of antibody corresponds to the solution of problem, and the affinity between antigen and antibody corresponds to the evaluation of solution. The lower the affinity is, the poorer the solution is.

(3)Immune matching: generate the detection component, and differentiate between self and non-self. Typically, immune matching can be evaluated by Euclidean distance, Manhattan distance, etc. The larger the distance is, the lower the affinity is.

(4)Clonal selection: various immune operators such as immune selection, cloning, mutation, clone suppression and population refresh are employed in this step.

(5)Antibody population update: if the termination condition is satisfied, the algorithm ends; otherwise, return to Step 3 until the current antibody population is the best solution.

B. *ACO algorithm*

ACO is one of the most popular swarm intelligence algorithms for optimization problems. Basically, ACO includes: (1) next step selection and (2) pheromone update. Suppose τ_{ij} is the amount of pheromone between nodes i and j, which is initialized as $\tau_{ij}(0)=B$, where B is a constant. At iteration t, the transfer probability of ant k from i to j can be calculated as:

$$p_{ij}^{k}(t)=\begin{cases} \dfrac{(\tau_{ij}(t))^{\alpha}(\eta_{ij}(t))^{\beta}}{\sum\limits_{u\in allowed_k}(\tau_{iu}(t))^{\alpha}(\eta_{iu}(t))^{\beta}}, & \text{if } j\in allowed_k; \\ 0, & \text{otherwise.} \end{cases} \qquad (2)$$

where $allowed_k$ denotes the set of qualified nodes for next step, α is the heuristic of pheromone indicating the importance of path, and β is the heuristic of path. $\eta_{ij}(t)=1/d_{ij}$ denotes the expectation of ant moving form i to j, where d_{ij} is the distance between nodes.

After each ant moves a step or finishes the traversal of all n nodes, the pheromone is updated as follows:

$$\tau_{ij}(t+n)=(1-\rho)\cdot\tau_{ij}(t)+\Delta\tau_{ij}(t) \qquad (3)$$

where $\rho\in(0,1)$ is the evaporation coefficient of pheromone, and $\Delta\tau_{ij}(t)=\sum_{k=1}^{m}\Delta\tau_{ij}^{k}(t)$, and

$$\Delta\tau_{ij}^{k}=\begin{cases} \dfrac{Q}{L_k}, & \text{if ant } k \text{ passes } (i,j); \\ 0, & \text{otherwise.} \end{cases} \qquad (4)$$

IV. OVERALL FRAMEWORK

In this section, we present the overall framework and data preprocessing procedure. The details of intrusion detection algorithm will be discussed in the next section.

We employ the cluster-based WSN structure [11] in this work, which is composed of sink node, cluster head nodes and cluster member nodes, as shown in Figure 2. This cluster-based structure helps to concentrate the communication control within a smaller range (i.e. cluster), and reduce the communication cost between aggregation nodes. The intrusion detection process utilizes the collaborative detection mechanism. First, as the global control node of the whole structure, sink node globally determines if intrusion is detected based on its detector. Then, detectors on cluster head nodes are wake up for collaborative detection. If the intrusion cannot be determined, collaborative detection on cluster member nodes is later wake up. Based on above detailed information, sink node can accurately justify the intrusion detected.

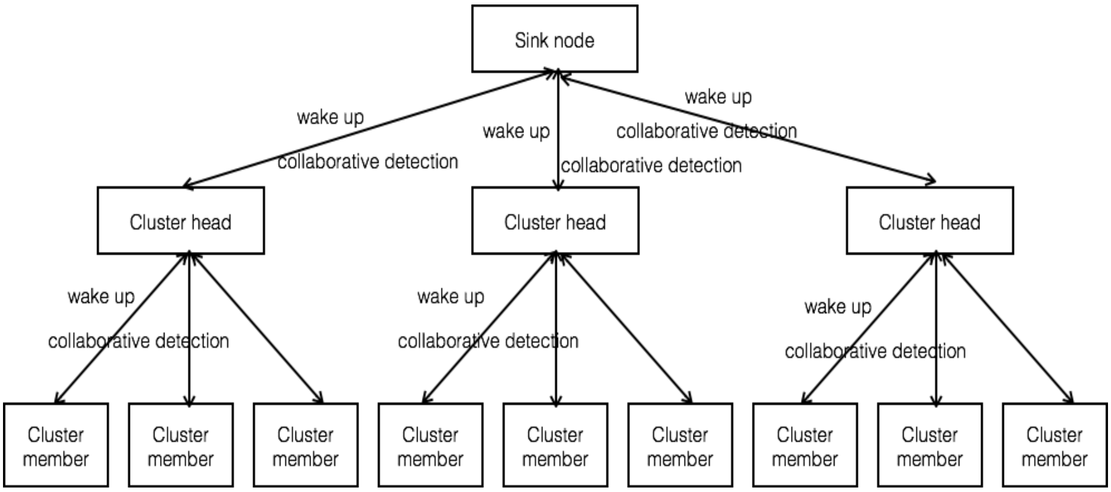

Figure 2. Overall structure of distributed intrusion detection in WSN

Specifically, the intrusion detection includes four steps: (1) data collection, (2) data process, (3) intrusion detection, and (4) response. Data collection in WSN intrusion detection is typically achieved by gathering data from various sensors and nodes. Data process includes feature selection and data preprocess.

In the process of intrusion detection, feature selection directly influences the performance of intrusion detection in terms of accuracy and speed. Typically, the original WSN dataset includes lots of redundant and useless information, and therefore brings issues such as curse of dimensionality. Besides, the dynamic and massive characteristics of WSN dataset also make it more difficult for intrusion detection. Therefore, before applying intrusion detection algorithm, we first eliminate redundant information using KNN algorithm, and then the ACO based algorithm is employed on the cleaned dataset.

Suppose the original number of features is m, and the feature set is represented as $O=\{F_i, i=1,2,...,m\}$. The feature set after redundancy elimination is R. Let $\lambda(F_i, F_j)$ denote the correlation between features F_i, F_j, and r_i^k denote the correlation between feature F_i and the k-th nearest neighbor in reduced feature set R. The process of dataset initialization based on KNN is as follows.

(1)Initialize $R=O$ and k.

(2)For each feature in R, calculate F_i.

(3)Find feature i' with minimum F_i value, and delete its k nearest neighbors. Let $\varepsilon = r_{i'}^k$.

(4)If $k > card(R)-1$, then set $k = card(R)-1$. If $k=1$, it indicates that there is no feature or its neighbor whose value is smaller than ε to delete, and the process stops.

(5)If $r_{i'}^k \geq \varepsilon$, then $k=k-1$, $r_{i'}^k = \inf_{F_i \in R} r_i^k$. If $k=1$, the process stops.

(6)If $r_{i'}^k \leq \varepsilon$, return to Step 2.

Now we have the cleaned dataset without data redundancy. Then, we perform data normalization as follows.

Calculate the mean for each feature:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{5}$$

where n is the scale of dataset, and x_i denotes each value of specific feature.

Calculate the standard deviation for each feature:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2} \tag{6}$$

.

Normalize data based on mean and standard deviation:

$$x_i' = \frac{x_i - \mu}{\sigma} \tag{7}$$

.

Then, the preprocessed data is fed into the intrusion detection algorithm, which will be discussed in the next section. Once an intrusion is detected, corresponding response measure is made by sink node. For example, broadcast the intrusion to the whole network to remind other nodes to reduce or avoid connections with the intrusion node.

V. INTRUSION DETECTION ALGORITHM

In order to formulate the intrusion detection as optimization problem, we need to define the objective function. There are two objectives of intrusion detection: (1) the accuracy of detection should be as high as possible; and (2) the number of features selected should be as small as possible. Therefore, the objective function is defined as:

$$f(x_i) = \lambda P_a(i) - (1-\lambda)\frac{d}{D} \tag{8}$$

where d is the number of selected features, D is the number of features of the original dataset, P_a is the accuracy of intrusion detection, and λ is the weights of above two components. Therefore, the objective of ACO is to maximize Equation (8).

In this paper, we combine ACO with immunity principle. Figure 3 shows the workflow of proposed immunity based ACO algorithm. The steps of proposed algorithm are described as follows:

(1)Initialization: initialize parameters of immunity system and ACO. Suppose $a_i(t)$ denotes the number of ants on node i at iteration t , and $\omega_{ij}(t)$ is the amount of pheromone on path (i,j) at iteration t . Let m be the total number of ants and n be the number of nodes, and we have $m=\sum_{i=1}^{n} a_i(t)$. Distribute ants to nodes evenly, and the number of ants on each node is $\left\lfloor \frac{m}{n} \right\rfloor$.

(2)Each ant chooses to transfer to the next node by Equation (2), and then update the pheromone on path by Equation (3).

(3)If all ants complete the iteration, go to Step 4. Otherwise, return to Step 2.

(4)Set the current solution of ACO as the initial antibody population, and calculate the affinity between antigen and antibody.

(5)Perform immunity operators such as immune matching, clonal selection and then update the antibody population.

(6)If the maximum number of immune iterations is achieved, go to Step 7. Otherwise, return to Step 5.

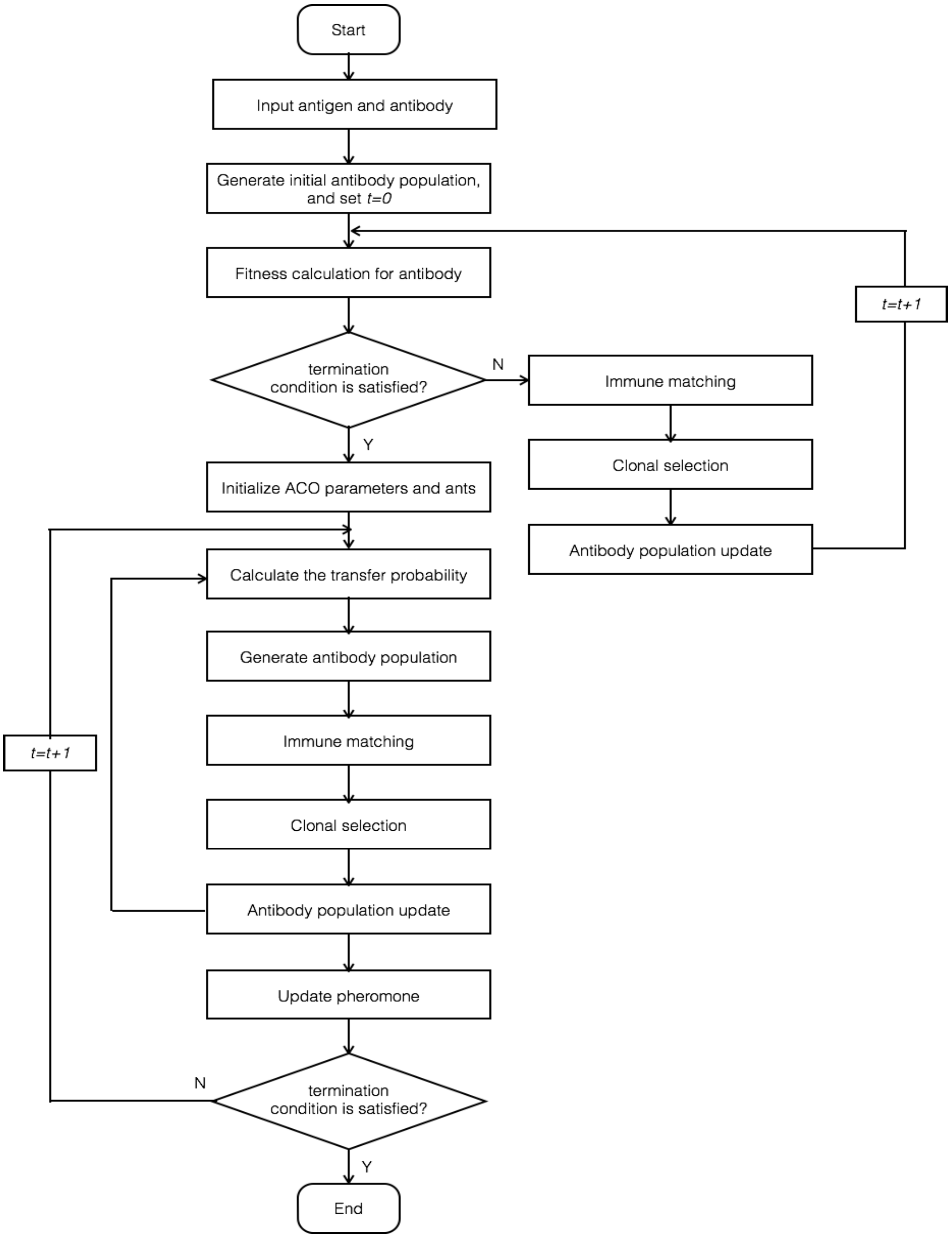
Start
Input antigen and antibody
Generate initial antibody population, and set t=0
Fitness calculation for antibody
termination condition is satisfied?
N
Y
Immune matching
Clonal selection
Antibody population update
t=t+1
Initialize ACO parameters and ants
Calculate the transfer probability
Generate antibody population
Immune matching
Clonal selection
Antibody population update
Update pheromone
termination condition is satisfied?
N
Y
t=t+1
End

Figure 3. Workflow of proposed immunity based ACO algorithm

(7)If the affinity value of the new solution is better than former solution, and the similarity between two solutions is small enough, add the new solution into pool. Otherwise, discard the new solution.

(8)Update the pheromone by:

$$\tau_{ij} = \begin{cases} (1-\rho)\tau_{ij} + \rho \cdot affinity, & if\ i,j\ are\ neighbors; \\ (1-\rho)\tau_{ij} + c \cdot \rho \cdot affinity, & otherwise. \end{cases} \quad (9)$$

where $\rho \in (0,1)$ is the evaporation coefficient of pheromone, $c \in (0,1)$ is a constant, and $affinity$ is the affinity value between antigen and antibody.

If the maximum number of ACO iterations is achieved, the algorithm ends. Otherwise, return to Step 2.

VI. EXPERIMENT

We simulate the experiment using Matlab software. The dataset used in this experiment is obtained from KDD CUP99, which includes four types of intrusions: DOS (denial of service), Probe (surveillance or probe), U2R (user to root) and R2L (remote to local). The training data includes 23 attacks in total, and the test data includes 38 attacks in total.

In order to measure the performance of intrusion detection, we employ two metrics, i.e., detection rate DR and error rate ER:

$$DR = \frac{\text{the number of detected abnormal}}{\text{total number of abnormal in sample data}} \quad (10)$$

$$ER = \frac{\text{the number of normal error reported as abnormal}}{\text{total number of normal in sample data}} \quad (11)$$

We compare the performance of proposed algorithm with ACO [20] and BP network [21] based intrusion detection method. Figures 4 and 5 give the results of detection rate and error rate respectively. We can observe that our method outperforms other two for all four types of attacks in terms of detection rate and error rate. Specifically, for DOS, Probe and R2L attacks, all three methods can detect intrusions efficiently and the error rate is relatively low. However, the performance of U2R attack is relatively poor. But still, our method can improve the detection compared to other methods.

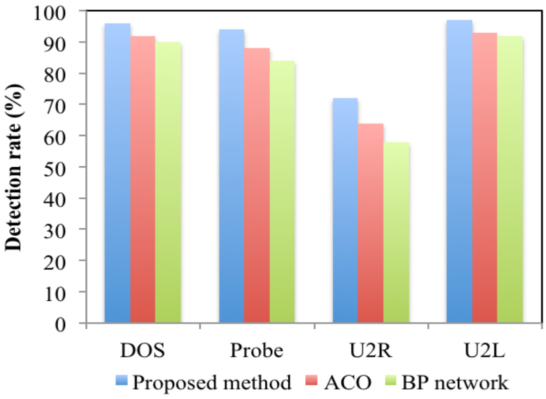

Figure 4. Detection rate comparison

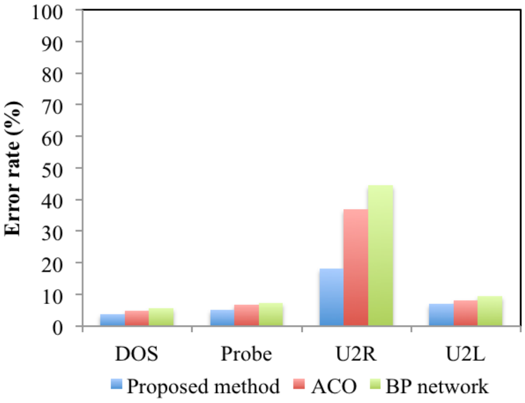

Figure 5. Error rate comparison

VII. CONCLUSION

In this paper, we design a cluster-based intrusion detection structure in WSN, and introduce immunity principle and ACO algorithm for intrusion detection. Experimentally, proposed method exhibits high accuracy in intrusion detection. However, in this work, we simplify the problem by ignoring the communication cost between sensor nodes in intrusion detection. In future works, we would like to extend the intrusion detection to deal with various scenarios and conditions.

REFERENCES

[1] Yick, Jennifer, Biswanath Mukherjee, and Dipak Ghosal. "Wireless sensor network survey." Computer networks 52.12 (2008): 2292-2330. http://dx.doi.org/10.1016/j.comnet.2008.04.002

[2] Perrig, Adrian, John Stankovic, and David Wagner. "Security in wireless sensor networks." Communications of the ACM 47.6 (2004): 53-57. http://dx.doi.org/10.1145/990680.990707

[3] ZENG Peng, Liang W, et al. Research on Security System of Wireless Sensor Network Based on Biological Immunity Principle [J]. Mini-micro Systems, 2005.

[4] Dorigo M, Birattari M, Stutzle T. Ant Colony Optimization [M]// Wiley Encyclopedia of Operations Research and Management Science. John Wiley & Sons, Inc., 2010:36-39.

[5] Liao Y, Vemuri V R. Use of K-Nearest Neighbor Classifier for Intrusion Detection [J]. Computers & Security, 2002, 21(2):439-448. http://dx.doi.org/10.1016/S0167-4048(02)00514-X

[6] Chen Z G, Dong-Yan L I. Intrusion Detection Based On Kernel Fisher Discriminant Analysis and Minimax Probability Machine Classifier [J]. Journal of University of Electronic Science & Technology of China, 2007, 36(6):1192-1194.

[7] Salmon, Helio Mendes, et al. "Intrusion detection system for wireless sensor networks using danger theory immune-inspired techniques." International journal of wireless information networks 20.1 (2013): 39-66. http://dx.doi.org/10.1007/s10776-012-0179-z

[8] Sommer, Robin, and Vern Paxson. "Enhancing byte-level network intrusion detection signatures with context." Proceedings of the 10th ACM conference on Computer and communications security. ACM, 2003. http://dx.doi.org/10.1145/948109.948145

[9] Zhang, Zheng, et al. "HIDE: a hierarchical network intrusion detection system using statistical preprocessing and neural network classification." Proc. IEEE Workshop on Information Assurance and Security. 2001.

[10] Misra, Sudip, P. Venkata Krishna, and Kiran Isaac Abraham. "Energy efficient learning solution for intrusion detection in wireless sensor networks." Communication Systems and Networks (COMSNETS), 2010 Second International Conference on. IEEE, 2010. http://dx.doi.org/10.1109/comsnets.2010.5431976

[11] Huang, Yi-an, and Wenke Lee. "A cooperative intrusion detection system for ad hoc networks." Proceedings of the 1st ACM workshop on Security of ad hoc and sensor networks. ACM, 2003.

[12] Bao, Fenye, et al. "Hierarchical trust management for wireless sensor networks and its applications to trust-based routing and intrusion detection." Network and Service Management, IEEE Transactions on 9.2 (2012): 169-183. http://dx.doi.org/10.1109/TCOMM.2012.031912.110179

[13] Butun, Ismail, Salvatore D. Morgera, and Ravi Sankar. "A survey of intrusion detection systems in wireless sensor networks." Communications Surveys & Tutorials, IEEE 16.1 (2014): 266-282. http://dx.doi.org/10.1109/SURV.2013.050113.00191

[14] Akay, Bahriye, and Dervis Karaboga. "A modified artificial bee colony algorithm for real-parameter optimization." Information Sciences 192 (2012): 120-142. http://dx.doi.org/10.1016/j.ins.2010.07.015

[15] Tuba, Milan, and Raka Jovanovic. "Improved ACO algorithm with pheromone correction strategy for the traveling salesman problem." International Journal of Computers Communications & Control 8.3 (2013): 477-485. http://dx.doi.org/10.15837/ijccc.2013.3.7

[16] Yoo, Kwang-Seon, and Seog-Young Han. "A modified ant colony optimization algorithm for dynamic topology optimization." Computers & Structures 123 (2013): 68-78. http://dx.doi.org/10.1016/j.compstruc.2013.04.012

[17] Ciornei, Irina, and Elias Kyriakides. "Hybrid ant colony-genetic algorithm (GAAPI) for global continuous optimization." Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 42.1 (2012): 234-245. http://dx.doi.org/10.1109/TSMCB.2011.2164245

[18] Hao, Wei, and Xinying Xu. "Immune ant colony optimization network algorithm for multi-robot path planning." Software Engineering and Service Science (ICSESS), 2014 5th IEEE International Conference on. IEEE, 2014. http://dx.doi.org/10.1109/icsess.2014.6933762

[19] Victor A. Skormin, Jose G. Delgado-Frias, Dennis L. McGee, Joseph V. Giordano, Leonard J. Popyack, Vladimir I. Gorodetski and Alexander O. Tarakanov. BASIS: A Biological Approach to System Information Security. [M] Information Assurance in Computer Networks. Springer Berlin Heidelberg, 2001:127--142.

[20] Janakiraman S, Vasudevan V, Janakiraman S, et al. ACO based Distributed Intrusion Detection System.[J]. International Journal of Digital Content Technology & Its Applications, 2009.

[21] Bao H, Qin J, Zhang X H. The Application of BP-Neural Network in the Intrusion Detection [J]. Computer Security, 2010.

AUTHOR

Hanqing Zhang. He graduated from College of Information Engineering, Yanshan University and received his M.Sc. in Information Engineering in 2008. He began to work in Hebei Institute of Foreign Languages in 2009 and now is Lecturer. Since 2010 he has been Director of Office Affairs of English Department. His current research interests include different aspects of Information Technology, Network Security, Artificial Intelligence and Distributed Systems. (email: 763777988@qq.com)

Submitted 21 September 2015. Published as resubmitted by the authors 20 October 2015.