An Improved V-detector Algorithm for Wireless Sensor Network Intrusion Detection Technology based on Immune System Principle

Xiaohong Hao [1,a], Wanfei Jiang[1,b] and Yifang Yuan[1,c]

[1]School of Computer & Communication, LanZhou University of Technology, Lanzhou 730050,China

[a]316475958@qq.com, [b]136172464@qq.com, [c]1987701140@qq.com

Keywords: Wireless sensor networks, Immune system, Intrusion detection, V-detector algorithm

Abstract. This paper is based on the problem of Wireless Sensor Networks (WSN) being easily invaded, but using the WSN intrusion detection model that is similar to the biological immune system the problem can be solved. After the working mode of the immune system is introduced and applied to WSN, the way to improve it can be found. A modular architecture is used for immune system used for WSN. This paper will reduce node energy consumption improve the V-detector algorithm by reducing the complexity of the algorithm and the amount of data storage. Because of WSN limitations, this algorithm is more suitable for WSN.

Introduction

With the rapid development of modern society, Wireless sensor networks (WSN) are becoming more and more widely used. In particular, the rise of the Internet of things let the WSN, the source of the initial data, playing an increasingly important role. But the wireless sensor network, as the network, is also affected by a wide variety of network attacks. These attacks can cause different degrees of impact on WSN and even direct damage to the entire WSN and bring huge losses to users. So it is very important to study the intrusion detection technology for WSN. And the immunologic mechanism of biological immune system can be very good to help biometric identification of self and non-self, and destroy the non-self (also known as pathogens) cells or debris. It has very strong information processing capability, and has the characteristics of recognition, memory, regulation, learning and so on. These characteristics also determine the biological immune system in the field of information processing has a very good application prospect.

The principle of immune system

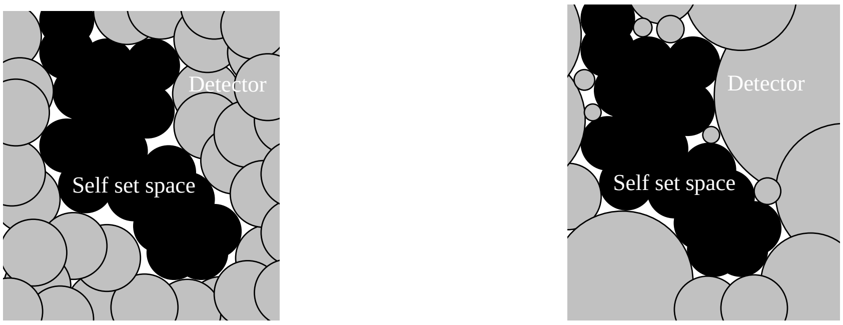
Immune system is an important system for the immune response and immune function in the body. It’s comprised of immune organs, immune cells and immune substances, and it’s the most effective weapon to defend against pathogens. It can find and remove the foreign body, foreign pathogenic microorganisms and other factors causing internal environmental fluctuations, but the hyperfunction of the immune system will damage their own organs or tissues[1].

An improved V-detector algorithm for Wireless Sensor Network Intrusion Detection Technology

Wireless sensor network intrusion detection system based on immune principle includes many aspects. First of all, the imitation of the gene pool requires a knowledge base, called the autologous library[2]. Secondly, it needs a detector generating module just like the lymphocyte, randomly generated candidate detectors by negative selection algorithm. Then the detectors compared with the self-set by the detection module, and remove the detectors that matches the self-set, the detectors which is not matches the self-set will be activated (It like the negative selection process in the immune principle). Finally, similar to the clone selection stage, in the clone selection module in

Intrusion Detection System, evolution to a memory detector with a longer life cycle and a smaller matching threshold, so it can achieve rapid attack detection[3].

In 2004, Zhou J according to the detector module set up a V-detector algorithm based on the negative selection algorithm[4]. With respect to the improved real valued negative selection algorithm, the V-detector algorithm can with fewer detectors cover a greater sample space, to reduce the black hole and improve the efficiency of detection. As shown in Figure 1, in the V-detector algorithm, all detectors are tangency with a sample in the self-set. A large radius of the detector can cover a large number of non-self-sets and thereby reduce the number of detectors and improve detection efficiency. While the small radius of the detector can cover the small gap between the sample and thus reduce the black hole. The V-detector algorithm of Zhou J is a pattern recognition algorithm based on distance computation. Each randomly generated detector needs to calculate the distance from the entire set of samples, the overhead of the algorithm is too large. And because of the stochastic nature of detector generation mechanism, leading to the late stage of the detector generation, majority of non-self-set space has been covered by the detector, the probability of generating effective detector is also greatly reduced. But this algorithm has very good improvement space. It can be applied to the intrusion detection of wireless sensor network by optimizing the algorithm. So this paper presents a detector optimization algorithm based on V-detector algorithm, further reducing the number of detectors generated, to some extent, it reduces the computation and data storage. This algorithm also improves the detection efficiency. Specific optimization methods will be described in detail below.



(a) Detector generating graph with a fixed radius

(b) Detector generating graph with variable radius

Figure.1 Generating detector by negative selection algorithm and V-detector algorithm

In wireless sensor network, the sample collected in normal communication data is not shown in Figure 3. Individual normal sample will be scattered, it may have a long distance with the other normal samples, even close to the attack sample. This creates a non-self space that is separated by these points, and lead the radius of the detector is not too large and the number of the detector will increased, even cause missing the attack sample which close to the self sample in the detection stage. Therefore, it can be used to select the training samples and remove the isolated samples of the scattered points, but the correct rate of the classification of normal samples can be reduced in the same time.

So set the following detector generation rule:

In the V-detector algorithm, the value of the parameter t is used as the criterion for the early stage and the late stage of the detector generation, when t reached $1/(1-c_0)$, entering the late stage of the detector generation and using the late stage detector generation rule, otherwise its may using the early stage detector generation rule.

The detector generating rule of early stage using the method of literature [6], the late stage of detector generation rule can set the random value range x-range of the center point of the detector, defined as:

$x-range=\{hypercubelhypercube=([low_1,low_2,...,low_d],[high_1,high_2,...high_d])\}$, $low_i=s_i-ran, high_i=s_i+ran$, ran is a fixed value by setting, s_i is the i-dimensional coordinates of any self samples. Randomly generated new detector center point in x-range, and then get a smaller radius of the detector to fill the loopholes.

Set the detector set D comes from V-detector algorithm, include n detectors. Each detector is composed of the d-dimensional coordinates and radius, as shown below:

$$D=\begin{bmatrix} x_{11},x_{12},...x_{1d},r_1 \\ x_{21},x_{22},...x_{2d},r_2 \\ \vdots \quad \vdots \qquad \vdots \quad \vdots \\ x_{n1},x_{n2},...x_{nd},r_n \end{bmatrix} \tag{1}$$

There is a high overlap problem exists in the mutual coverage of the D, the nodes in wireless sensor networks are limited in storage space and computing power, the redundancy of the detector greatly increases the amount of computation of the node's information storage and detection. Therefore, this paper proposes a detector optimization algorithm to remove the redundancy of the detector, reduce the number of detectors to speed up the detection efficiency.

Define detector cover matrix Overlap:

$$Overlap=\left|\begin{array}{cccccc} 0 & A_{1,2} & A_{1,3} & \cdots & & A_{1,n} \\ A_{2,1} & 0 & A_{2,3} & \cdots & & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ A_{i,1} & \cdots & A_{i,j} & \cdots & 0 \cdots & A_{i.n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{n,1} & A_{n,2} & A_{n,3} & & \cdots & 0 \end{array}\right| \tag{2}$$

Among them $A_{i,j}$ represents the coverage degree of the detector d_j to d_i, when the detector d_i is fully covered by the detector d_j, $A_{i,j}=e^0=1$, shows the coverage degree is 1.

$$A_{i,j}=e^{-(Dis_{x_ix_j}+r_i-r_j)^d/r_i^d)} \tag{3}$$

The totality coverage degree $TotalA_i$ is the computational detector d_i covered by the other detectors:

$$TotalA_i=\sum_{j=1}^{n}A_{ij} \tag{4}$$

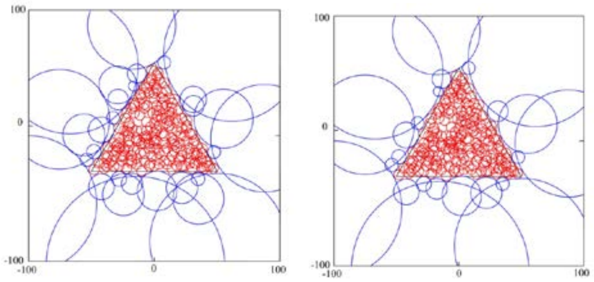
Set threshold σ_1, σ_2, develop detector optimization rules:

$$f\,\exists A_{i,j}>\sigma_1 \; or \; TotalA_i>\sigma_2 \; then \; delete \; d_i \tag{5}$$

After each detector is removed, it will recalculate the detector's coverage matrix and the totality coverage degree covered by the other detectors, loop operation optimization, until all the coverage degree of the detectors is within the threshold. At this time, the remaining detectors constitute the final mature detector set D, and then published from the base station to the detection node.

Simulation and analysis

In order to verify the effectiveness of the proposed detector optimization algorithm in this paper, simulation experiment a two-dimensional was used. In a space of 200*200, in order to simulate the complexity of the self-set, 200 self-set individuals were put. The center points of the individuals were randomly distributed in a regular triangle with sides measuring 100. Simulation experiments were carried out using the original V-detector algorithm and the optimized algorithm. The maximum number of the randomly generated detectors, the detector coverage was $t_{\max}=50$, the radius of the self-set sample r_s was 4, the coverage degree of the detectors was $C_0=99.99\%$. Thresholds σ_1 and σ_2 were 0.95 and 1.2. The result of the simulation is shown in figure 2. From the figure, using a similar self-set coverage rate, the number of detectors is reduced by 20% from 30 to 24.



(a) Detector generation in V-detector algorithm (30 detectors)

(b) Detector generation in improved algorithm (24 detectors)

Figure.2 The collation map of detector generation before and after improving the V-detector algorithm

In order to further verify the superiority of the improved algorithm, the multi dimensional simulation analysis is carried out. Simulation platform using Matlab2014a, simulation experiment data comes from the 10% intrusion detection data set KDD CUP99, parameter settings as shown in table 1.

Table 1 Parameter setting of multi dimensional simulation experiment

Parameter	$T_{\max}$	r_s	c_0	η	σ_1	σ_2	σ_3
Set value	150	0.02	99.99%	2	0.85	3	100

In this experiment, selecting 5000 samples include 2000 Normal samples, 2000 Dos samples, 500 Probe samples, 400 R2L samples and 100 U2R samples from 10% intrusion detection data set KDD CUP99 for the test samples. The data in Table 2 is the training sample before and after screening, the correct rate of the detection algorithm for the different types of samples. With the Table 2, Most of the correct judgment rate are increased, but as some Normal samples were cut, the correct judgment rate dropped from 99.55% to 95.75%. However, compared to the other attack sample detection rate, the Normal detection rate is reduced in the range of acceptance.

Table 2 The correct judgment rate of each type of sample before and after screening

	Normal	Dos	Probe	R2L	U2R
Without screening training sample	99.55%	100%	68.4%	80.25%	64%
Screening training sample	95.75%	100%	100%	86.75%	76%

Table 3 data is the simulation results contrast before and after improved V-detector algorithm. As shown in Table 3, the new detector generation rule based on the improved algorithm reduces the total time of the formation stage by 38.11% from 622.41s to 385.20s. Improve the efficiency of generating an effective detector in the late stage of the detector generation. The improved algorithm also reduces the number of detectors from 162 to 120, and 16 of them are memory detectors. At the same time, with the decrease of the number of detectors, the calculation of the detection phase is reduced, so the time of the detection is reduces by 22.39% from 23.58s to 18.30s.

Table3. The simulation results contrast before and after improved V-detector algorithm

	Before improve	After improve
Detector generation time	622.41s	385.20s
Generated detector number	162	120
Detection time	23.58s	18.30s

Conclusion

This paper, describes an improved V-detector algorithm which reduces the amount of calculation of the algorithm and data storage and is more suitable for application in wireless sensor networks. In practical application it is difficult to complete the definition of the self-set, and the algorithm, even after optimization, still needs a large amount of calculation and computation time, so it can only be applied with some limitations to resource abundant base stations. The author will continue to focus on the application of the immune system principle in the research of wireless sensor network intrusion detection, hoping for research progress together with other research scholars.

Reference:

[1] Boquan Jin． Medical immunology [M]. Beijing: People's Health Publishing House, 2008.6: 12-74.

[2] P. M. Lydyard, A. Whelan, M. W. Fanger. Instant Notes in Immunology [M]. BIOS Scientific Publishers, 2000.

[3] Guangwei Liang. Artificial Immune based Intrusion Detection in Wireless Sensor Networks [D]. WuXi: Jiangnan University. 2014.6.21-28.

[4] Zhou J，Dasgupta D. Real-valued negative selection algorithm with variable-sized detectors [C]. In: Proc. of 6th Conference on Annual Genetic and Evolutionary Computation. Seattle，WA: 2004. 287-298.