Cloud Computing for Network Security Intrusion Detection System

Jin Yang
[1] School of Information Science & Technology, Southwest Jiaotong Univ., Chengdu, China
[2] Department of Computer Science, LeShan Normal Univ., LeShan 614000, China
Email: jinnyang@163.com

[1,] Cilin Wang, [2,] Caiming Liu, [3,] Le Yu
[1, 2,] Department of Computer Science, LeShan Normal Univ., LeShan 614000, China
[3,] Military Representative Office of PLA, Guiyang, China
Email: bigluckboy@163.com

***Abstract*—In recent years, as a new distributed computing model, cloud computing has developed rapidly and become the focus of academia and industry. But now the security issue of cloud computing is a main critical problem of most enterprise customers faced. In the current network environment, that relying on a single terminal to check the Trojan virus is considered increasingly unreliable. This paper analyzes the characteristics of current cloud computing, and then proposes a comprehensive real-time network risk evaluation model for cloud computing based on the correspondence between the artificial immune system antibody and pathogen invasion intensity. The paper also combines assets evaluation system and network integration evaluation system, considering from the application layer, the host layer, network layer may be factors that affect the network risks. The experimental results show that this model improves the ability of intrusion detection and can support for the security of current cloud computing.**

***Index Terms*—Cloud Computing; Artificial Immune Systems; Network Security**

I. INTRODUCTION

Cloud computing is the delivery of computing as a service rather than a product, whereby shared resources, software, and information are provided to computers and other devices as a metered service over a network. It has so many advantages such as economy, complex calculations, agility, high scalability, high reliability, easy maintenance. The concept of cloud computing was born in the 1960s from the ideas of American computer scientist J.C.R. Licklider and John McCarthy stated that computing will become a publicly available service in the future[1]. In 1983, Sun Microsystems bring forward a singular vision that "the network is the computer" [2]. On August 09, 2006, the CEO Google, Eric Schmidt, firstly mentioned the concept of Cloud computing on SES San Jose 2006. On Jan.30th, 2008, Google declared "Cloud Computing Research Plan" in Taiwan and will promote the advanced technology in Taiwan's colleges. Feb 1th 2008, IBM (NYSE: IBM) announced it will establish the first Cloud Computing Center for software companies in China, which will be situated at the Wuxi-TaiHu New Town Science and Education Industrial Park, China. On March 5, 2010, Novell and CSA released a supplier neutral plane, named as Trusted Cloud Initiative. May 22, 2009, China's first Cloud Computing Conference held in Beijing China World Hotel. January 22, 2010, China cloud computing technology and industry alliance (CCCTIA) announced in Beijing.

Cloud computing through the network environment make the complex computational processing program to split into numerous smaller subroutines, and then handed over the analyzed results back to the user. It is a kind of typical network computing mode, which emphasizes on large-scale virtual computing environment to run application scalability and availability. It has become the focus of great concern to the industry, the academia, and even the government. But now, with the increasing popularity of cloud computing, the importance of network security in cloud computing is rising, which has become the important factor of restricting its development. The traditional network security approaches include virus detection, frangibility evaluation, and firewall etc., e.g., the Intrusion Detection System (IDS) [3]. They rely upon collecting and analyzing the viruses' specimens or intrusion signatures with some traditional techniques. Moreover, being lack of self-learning and self-adapting abilities, they can only prevent those known network intrusions, and can do nothing for those variety intrusions.

Recent years, the artificial immune system has the features of dynamic, self-adaptation and diversity [4-7] that just meet the constraints derived from the characteristics of the grid environment, and mobile agent has many same appealing properties as that of artificial immune system. Negative Selection Algorithm and the

This work was supported by China Postdoctoral Science Foundation (No.2011M501419), and the National Natural Science Foundation of China (No.61003310, No.61103249) and the Scientific Research Fund of Sichuan Provincial Education Department (No. 10ZB005).
† Corresponding author.

doi: 10.4304/jnw.8.1.140-147

concept of computer immunity proposed by Forrest in 1994 [8-9]. In contrast, the AIS theory adaptively generates new immune cells so that it is able to detect previously unknown and rapidly evolving harmful antigens [10]. However, much theoretical groundwork in immunological computation has been taken up, but there is a lack of perfectly systems based AIS of dynamical immunological surveillance for network security [11, 12]. Based on the correspondence between the artificial immune system antibody in the artificial immune systems and pathogen invasion intensity, this paper is to establish a network risk evaluation model [13, 15]. We built a hierarchical, quantitative measurement indicator system, and a unified evaluation information base and knowledge base. This model will help the network managers evaluate the possibility and the graveness degree of the network dangerous quickly, ease the pressure of recognition, to get targeted immediate defense strategy of the strength and risk level of the current network attacks.

This article applied AIS technique in the field of network security situation awareness, designed and established an immune network security situation awareness system. It is aimed to carry out in cloud computing environment on real-time monitor the network security situation, realize real-time and quantitative awareness of network security situation before malicious network behavior becomes out of control, and help make timely and effective network security strategy adjustment for better general security safeguard of system.

II. THE IMMUNE-BASED INTRUSION DETECTION

Biological Immune System (BIS) is a complicated system with the ability of self-adapting, self-learning, self-organizing, parallel processing and distributed coordinating, and it also has the basic function to distinguish self and non-self and clean non-self. The problems in the field of computer security and Artificial Immune Systems have the astonishing similarity of keeping the system stable in a continuous changing environment. Artificial Immune System can use biological immune theoretic for references to search and design relevant models and algorithms to solve the various problems occurred in the field of computer security. Technology for Network Security Situational Awareness, which is a positive defense technology, has become the orientation of research in the field of network security. Based on the analysis of the papers from domestic and foreign on technologies for network security situational awareness, this paper designs and builds a network security situational awareness system based on the profound research of BIS. The system uses network intrusion detection, which based on the theory of biological immunity as the base of situational awareness, to detect known and unknown intrusions with the help of biological technology such as self/non-self discrimination, self-tolerance, self-learning, evolution mechanism, immunological surveillance, etc. According to correspondence relations of density change of antibody in the artificial immune systems and pathogen invasion intensity, a novel network risk evaluation model is also established. Based on the current real-time network risk evaluation, the thesis also makes risk evaluation on short-term, medium-term、long-term network in different span. These methods make overall and quantitative identification about network security, and it is also helpful to resemble network security tragedy effectively, therefore, protect network infrastructure greatly.

Simulating creatural immune system, we place a certain amount of immune cells into the network, and perceive the surrounding environment of the detectors. As soon as the immune detectors detect an attack, the detectors begin clone and generate a mass of similar detectors in order to defend from fiercer network attacks and warn the dangerous level of the network. The network security situation awareness agent in cloud computing environment shown in Figure 1 is itself a sub-network security situation awareness system, defined by recursion, and it mainly monitors on the sub-network security situation within its control, and specifically speaking, real-time monitor on the type, strength and harmfulness of attacks suffered by sub-network. Because there might as well be subnets under subnets, sub-network security situation awareness agents may be composed of sub-network security situation awareness agents at lower level. Eventually, security situation awareness agents, that monitor the specific host computer, are made up of intrusion detection and security situation evaluation of the host computer.

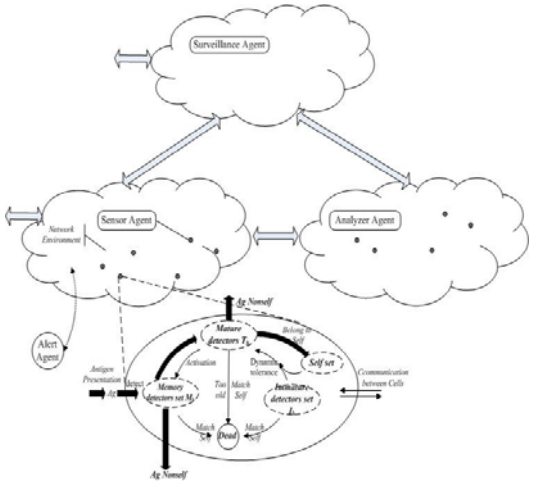

Figure 1. Architecture of the Evaluation of the Network Danger

TABLE I.
THE RELATIONSHIP BETWEEN THE BIS AND THE OUR MODEL

Biological immune system	Network intrusion detection system
Organism	Network
Organ	Network segment
Cell	Host computer
Vaccine distribution	The transmission of intrusion information
Antigen	The binary character string feature-extracted from IP packets
B Cell, Antibody	Antibodies represented in binary character string
Cell clone	Duplication of antibody

Combined with multi-agent and AIS technology, the detection model constitutes a multi-direction and multi-level intelligent network security model, with its mapping relationship with BIS model as shown in Table 1, and its system structure diagram as shown in Figure 1.

Agent of intrusion behavior distilling use vector space model and present the received datagram in discrete characters. Agent of Training generates various immature detectors from gene library to distinguish self and Non-self. According to immune principle, some of these new immature detectors are false detectors and they will be removed by the negative selection Agent, which matches them to the training datagram. If the match strength between an immature detector and one of the training datagram is over the pre-defined threshold, this new immature detector is considered as a false detector. Agent of intrusion surveillance matches the received datagram to the mature detectors. If the match strength between a received datagram and one of detectors, the behavior will be consider as the intrusion. The detail training phases are as following.

A. Definition of Antigen, Antibody, Self and Non-self

Definition: Antigens (Ag , $Ag \subset U$, $D = \{0,1\}^l$) are fixed-length binary strings extracted from the Internet Protocol (IP) packets transferred in the network. The antigen consists of the source and destination IP addresses, port number, protocol type, IP flags, IP overall packet length, TCP/UDP/ICMP fields [16], etc. The structure of an antibody is the same as that of an antigen. For virus detection, the oneself set (Nonself) represents IP packets from a computer network attack, while the self set (Self) is normal sanctioned network service transactions and nonmalicious background clutter. Set Ag contains two subsets, $Self \subseteq Ag$ and $Nonself \subseteq Ag$ such that

$$Self \cup Nonself = Ag \qquad Self \cap Nonself = \Phi$$

B. The Dynamic Equations of the Mature Cells

Let G_b represent the set of all the mature cells, and G the amount of the mature cells in the set at some time. Then the dynamic equation of the mature-cell set is formulated as following:

$$G(t+\Delta t) = G(t) + g_{new} \cdot \Delta t - (\frac{\partial G_{active}}{\partial x_{active}} + \frac{\partial G_{death}}{\partial x_{death}}) \cdot \Delta t .$$

It implies that the changing process of the set G_b is separated into two stage: one is "flow into", the other is "flow out". The 1st stage when mature cells come into the set, $G = g_{new} \cdot \Delta t$ (where g_{new} is the new mature cells in a unit time interval), the 1st stage shows the amount of mature cells which have flew into the set G_b in the time interval Δt ; the 2nd stage when mature cells flow out from the set, including two aspects: the mature cells which have been stimulated to be active and the dead mature cells. We denote $\partial G_{active} / \partial x_{active} \cdot \Delta t$ the differential amount of activated cells and $\partial G_{death} / \partial x_{death} \cdot \Delta t$ the differential amount of dead cells.

$G_{dead} = \{x \mid x.age > \lambda, x.cout < \beta\}$

During the course of the mature cells activated and dying, the following course is happening

$$age(t+1) = age(t) + 1 \qquad t < \lambda$$
$$affinity(t+1) = affinity(t) + 1 \qquad t < \lambda$$

where λ means the affinity accumulating cycle.

Equation implies that mature cells must accumulate affinity with the antigen Ag. In one cycle λ , with one unit of time interval, the age of the mature cell is added 1; if the matching of affinity is successful, that is, the function $f_{match}(x, y)$ holds, the *affinity* of the mature cell is added 1, Therefore, there must be one of following three cases:

① If $affinity \geq \theta \wedge t < \lambda$, the mature cell is stimulated to be active and become the memory cell.

② If $affinity \geq \theta \wedge t < \lambda$, the affinity of the mature cell isn't sufficient, the cell needs to go on accumulating affinity.

③ If $affinity < \theta \wedge t > \lambda$, the cell can't accumulated enough affinity in one cycle λ , so it has to die.

In the course, θ is the threshold of the affinity for the activated cells. The affinity function $f_{mathch}(x, y)$ may be any kind of Hamming, Manhattan, Euclidean, and r-continuous matching, etc. In this model, we take r-continuous matching algorithm to compute the affinity of mature cells. The matching functions utilize the following definitions:

$$f_{\text{math}}(x, y) = \begin{cases} 1 & \exists i, j, j - i \geq r \wedge 0 < i < j \leq l, \\ & x_i = y_i, x_{i+1} = y_{i+1}, \cdots, x_j = y_j \\ 0 & otherwise \end{cases}$$

The r-continuous matching is commonly used method for measuring the distance between bit strings with the goal of producing a better similarity coefficient.

C. The Dynamic Equation of Memory Cells

Let M_b the set of memory cells, M the amount of the memory cells in the set at some time. Because memory cells are more difficult to come into being, in this paper, the changing process of the set M_b only includes stage "flow into" except the dying stage. Memory cells totally come from the activated mature cells G_{active} , that is, the dynamic equation of memory cells is:

$$M(t+\Delta t) = M(t) + \frac{\partial M_{active}}{\partial x_{active}} \cdot \Delta t + M_{other\_host}(\Delta t) - M_{dead}(\Delta t)$$

$$M_{dead}(t) = \{x \mid x \in M(t), f_{match}(x, Self(t-1)) = 1\}$$

Equation describes the dynamic process of the memory cell set. (where $\frac{\partial M_{active}}{\partial x_{active}} \cdot \Delta t = \frac{\partial G_{active}}{\partial x_{active}} \cdot \Delta t$)

$$M(t+\Delta t) = M(t) + M_{new}(\Delta t) + M_{from\_other}(\Delta t) - M_{dead}(\Delta t),$$
$$when \quad f_{match}(M(t), Ag(t)) \neq 1 \quad , \quad t > 1,$$

$$M(t+\Delta t)=M(t)+M_{clone}(t)+M_{new}(\Delta t)+M_{from\_other}(\Delta t)$$
$$-M_{dead}(\Delta t), when \quad f_{match}(M(t),Ag(t))=1$$
$$M_{clone}(t)=\frac{\partial M_{clone}}{\partial x_{clone}}\cdot\frac{\partial M_{active}}{\partial x_{active}}\cdot\Delta(t-1),$$
$$when \quad f_{match}(M(t),Ag(t))=1$$
$$M_{clone}(t+\Delta t)=M_{clone}(t), M.\rho(t+\Delta t)=M.\rho(t)+V_p\cdot\Delta t,$$
$$M.count(t+\Delta t)=M.count(t)+1$$
$$M.\rho(t+\Delta t)=\frac{1}{2}\cdot M.\rho(t), M.age(t+\Delta t)=M.age(t)+1,$$
$$when \quad f_{match}(M(t),Ag(t))\neq 1$$
$$M_{new}(\Delta t)=\frac{\partial M_{new}}{\partial x_{new}}\cdot\Delta t=\frac{\partial T_{active}}{\partial x_{active}}\cdot\Delta(t-1)$$
$$M_{new}.\rho(t)=\rho_0$$
$$M_{dead}(\Delta t)=\frac{\partial M_{death}}{\partial x_{death}}\cdot\Delta t, when \quad f_{match}(M(t-1),Self(t-1))=1$$
$$M_{from\_other}(\Delta t)=\sum_{i=1}^{k}(\frac{\partial M^{i}_{from\_other}}{\partial x_{from\_other}}\cdot\Delta t)$$

Equations depict the dynamic evolution of memory detector. $M(t+\Delta t)$ simulates the process that the memory detector evolve into the next generation ones. $M_{new}(t)$ is the set of memory detector that are activated by antigens lately. These mature detector matched by an antigen will be activated immediately and turn to a memory detector. $M_{dead}(t)$ is the memory detector that be deleted if it matches a known self antigen. $M_{clone}(t)$ is the reproduced memory detector when the detector distinguish a antigens. $M_{from\_other}(t)$ is the memory detector that transformed from other computers. The k indicates that the ID number of the computer. Therefore, dynamic model of immune is to generate more antibodies and enhance the ability of self-adaptation for the system.

D. The Dynamic Model of Self

In a real-network environment some network services and activities are often change, which were permitted in the past but may be forbidden at the next time.

$$I(t)=I(0)=\{x_1,x_2,...,x_n\}, \qquad t=0$$
$$I(t+\Delta t)=I(t)+I_{new}(\Delta t)-I_{match\_self}(\Delta t)-I_{maturation}\cdot\Delta t, t>1$$
$$I.age(t+\Delta t)=I.age(t)+1$$
$$I_{match\_self}(t+\Delta t)=I(t), when \quad f_{match}(I(t-1),Self(t-1))=1$$
$$I_{maturatiio\,n}(t+\Delta t)=I(t), \qquad I.age(t+\Delta t)>\alpha$$
$$I_{new}(\Delta t)=(\xi_1\cdot\frac{\partial I_{\text{random}}}{\partial x})\cdot\Delta t+(\xi_2\cdot\frac{\partial I_{\text{inherit}}}{\partial x})\cdot\Delta t$$

Equation stimulates the dynamic evolution of self-antigens, where $x_i\in\Re(i\geq 1, i\in N)$ is the initial self element defined. Inew is the set of newly defined elements at time t, and Imaturation is the set of mutated elements. $f_{match}(y,x)$ is used to classify antigens as either self or nonself: if x is a self-antigen, return 0; if x is a nonself one, return 1; if x is detected as nonself but was detected as a self-antigen before, then it may be a nonself antigen (needs to be confirmed), and return 2. There are two advantages in this model. (1) Self immune surveillance: The model deletes mutated self-antigens (Imaturation) in time through surveillance. The false-negative error is reduced. (2) The dynamic growth of Self: The model can extend the depiction scope of self through adding new self-antigens (Inew) into Self. Therefore, the false-positive error is prevented.

E. The Antibody Variation

In order to prevent algorithm from converging prematurely, we take variation operation to the gene set $G_1=\{g_1,g_2,\cdots,g_i,\cdots g_n\}$ after the cross process. Based on the analysis of premature convergence of traditional evolutionary programming, a novel multi-subgroup evolutionary programming algorithm is proposed. This set G consists in accordance with a certain percentage composition of random, variation resulting from the combination of antibody fragments produced from gene sets, in part generated by the inherited characteristics of their parents. Let $i\in I$ be expressed as immature individual detector and I be the space for the individual detectors.

$t=0$;
initialize $I(0):\{i_1(0),i_2(0),\cdots,i_\mu(0)\}$;
evaluate $I(0):\{\Phi(i_1(0)),\Phi(i_2(0)),\cdots,\Phi(i_\mu(0))\}$;
while $(l(I(t))\neq True)$ do
evaluate: $I''(t):\{\ \Phi(I''(t))\ \}$
crossover: $I'(t):=c(I(t))$;
mutation: $I''(t):=m(I'(t))$

selection: If (μ,λ)—selection
then $s(I''(t))$
else $s(I'(t)\cup I''(t))$;
t=t+1;
end

where c is the crossover operator and m is the mutation operation, and s stands for selecting into the next generation of groups. Normally in order to maintain the diversity of individuals, we make the composition of immature detector diversity, reflecting the characteristics of immune diversity.

F. The Process of Network Security Surveillance

The self-adaptability, distributed character and quantization of antibody concentration that biological immune system bears are just the effective method to solve the technological problems of network security situation awareness, thus this article applies the characteristics of biological immune mechanism in the field of network security situation awareness research, and establishes immune network security situation awareness system to make real-time and quantitative analysis on network security condition and its changing trend. Therefore, system evaluates the network security by perceiving the danger around of them. The values of G_b and M_b reflect the intensity of intrusion in current network. The bigger the value G and M are, the more serious the network intrusion degree is. And the bigger

the value $\partial M_{active} / \partial x_{active}$ is, the quicker the change about network situation is. Through distinguishing the type of G and M, we can know different kinds of network intrusion. The values of λ and θ reflect the activity degree of the mature cell. Therefore, with immune cells' status parameters and the parameters weight, we can get network danger situation and evaluate network security at real time. The following contents will elaborate how to establish this model.

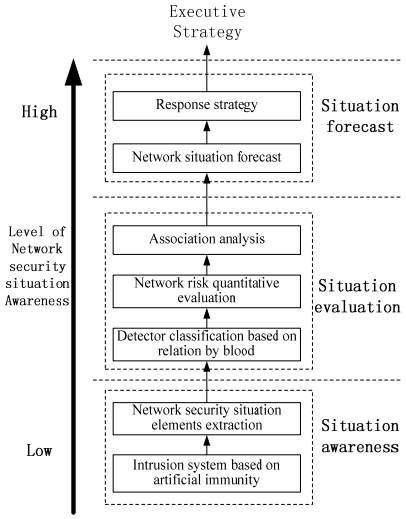

Figure 2. the structure of network security situation Surveillance

III. EVALUATION OF NETWORK RISK

After we describe the network attacking actions, it is necessary to evaluate the dangerous degree of the network, and judge the severity of the attacking actions. Thus, evaluation is a process involving numerous complicated factors. Owing to the fact that our model relates to enormous factors for evaluation, on purpose of reasonably and entirely measuring the network dangerous status, we classify the involved factors as host dangers, area dangers, detectors dangers, and special dangers. Afterwards, we subdivide and arrange all the factors which influence the network dangers, in order to let them locate on different layers, forming a structure model with identify matrix.

Here we quantify these indicators, associated with the use of multi-level gray scale model. Suppose there are n types of indicators which can impact the Importancce-indicator in the network. And each Importancce-indicator has m kinds of attributes. In other words, we can use m kinds of attributes to measure and influence the values of the Importancce-indicator. To determine the evaluation indicator system based on the evaluation object, we use the following set of indicators to describe.

$$X_i' = \left(x_i'(1), x_i'(2), \cdots, x_i'(m)\right)^T, \quad i = 1, 2, \cdots, n$$

The n types of indicators sequence formed the following matrix:

$$\left(X_1', X_2' \cdots, X_n'\right) = \begin{pmatrix} x_1'(1) & x_2'(1) & \cdots & x_n'(1) \\ x_1'(2) & x_2'(2) & \cdots & x_n'(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_1'(m) & x_2'(m) & \cdots & x_n'(m) \end{pmatrix}$$

The reference data column should be an ideal standard of comparison, which can be composed of optimal value of each index (or the worst value). Depending on the purpose of the evaluation we can also choose another reference data. The reference data column recorded as

$$X_0' = \left(x_0'(1), x_0'(2), \cdots, x_0'(m)\right)$$

The indicators data sequence after dimensionless formed the following matrix.

$$\left(X_0, X_1, \cdots, X_n\right) = \begin{pmatrix} x_0(1) & x_1(1) & \cdots & x_n(1) \\ x_0(2) & x_1(2) & \cdots & x_n(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_0(m) & x_1(m) & \cdots & x_n(m) \end{pmatrix}$$

Here, the dimensionless method we adopted is the mean-quantization way.

$$x_i(k) = \frac{x_i'(k)}{\frac{1}{m}\sum_{k=1}^{m} x_i'(k)}, i = 0, 1, \cdots, n; \quad k = 1, 2, \cdots, m.$$

The absolute difference $|x_0(k) - x_i(k)|$ is individually calculated between each target sequence (sequence comparison) to the references sequences corresponding to the elements. And the values of $\min_{i=1}^{n}\min_{k=1}^{m}|x_0(k) - x_i(k)|$ and $\max_{i=1}^{n}\max_{k=1}^{m}|x_0(k) - x_i(k)|$ are also get out by calculated the correlation coefficient between each comparison sequence to the reference sequence corresponding to the elements.

$$\zeta_i(k) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \rho \cdot \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \cdot \max_i \max_k |x_0(k) - x_i(k)|}$$

$$k = 1, \cdots, m$$

Where the ρ is the distinction coefficient, $\rho \in (0,1)$. The smaller the value of ρ is, the greater the difference between the correlation coefficient will make, and the stronger the ability of distinguish is. Here we let $\rho = 0.5$.

Evaluation of the object (more sequences) were calculated with reference to its target sequence of m elements corresponding to the mean correlation coefficient to reflect the evaluation of the object and the reference sequence of association. Since every indicators play different role in the comprehensive evaluation of this system, we used a weighted average of the correlation coefficient requirements, where the value of W_k is the weight of each index factors.

$$r_{0i}' = \frac{1}{m}\sum_{k=1}^{m} W_k \cdot \zeta_i(k) \qquad k = 1, \cdots, m$$

Eventually based on the observation of objects associated ordinal, the evaluation of the results is obtained.

Let $\rho_i(t)$ be the antibody concentration of jth host detect attacking at time t. Let u be the danger coefficient of the ith kind of attack in the network. Then, we can get the danger level value Ri, j (t) facing the ith kind of attack as follows:

$$R_{i,j}(t) = \tanh(\zeta \cdot u \cdot \sum_{x \in A_i(t)} \rho(t))$$

The results of different abnormity behavior harm to a same host are different. Therefore, the comprehensive danger level value rj (t) is the linear weighted sum of the jth host facing all of the attacks. Let $u_i (0 \le u_i \le 1)$ be the relative weight value of danger of the ith kind of attack in the network. Then, we can define the danger level value $r_j(t)$ of the ith kind of attack as follows:

$$R_j(t) = \tanh(\zeta \cdot \sum_{i=1}^{n} (u_i \cdot \sum_{x \in A_i(t)} \rho_i(t)))$$

The entire network of danger level should fully reflect the value of each of the host facing attacks. As the host of each position is not the same such as running a different system for different users and providing different services, influencing different economic, affecting different social and even political values, they are in possession of different essentiality.

Let $\text{Importance}_i = \sum_{k=1}^{8} (I_k \times W_k)$ be the importance coefficient of jth host in the network. Then, we obtain the network entire danger level value: R (t) =∑ (indicator value × indicator weight). Therefore, we can get network danger R (t) situation and evaluate network security at real time.

$$R(t) = \tanh(\sum_{m=1}^{N} (\sum_{i=1}^{n} (Host_i's \ \ danger \times \text{Importance}_j) \times LCRS\_Weight_m))$$

$$= \tanh(\sum_{m=1}^{N} (\sum_{i=1}^{n} (Host_i's \ \ danger \times \sum_{k=1}^{8} (I_{j,\,k} \times W_k)) \times LCRS\_Weight_m))$$

$$= \tanh(\sum_{m=1}^{N} (\sum_{i=1}^{n} (r_i(t) \times \sum_{k=1}^{8} (I_{j,\,k} \times W_k)) \times LCRS\_Weight_m))$$

By applying the basic principle of the artificial immune system in the domain of network security situation awareness, the architecture is established which includes network security situation detection, network security situation evaluation, network security situation prediction. Through detecting malicious intrusions, in plus with real-time and quantitative analysis, prediction according to the current security situation and the future tendency, so as to make the network information system be self-learning and self-adapting as BIS, thus, to improve immune ability and survivability for web system, as well as alleviate damage made by network attack and enhance the emergency response ability.

IV. Network Security Situation Forecast

Situation forecast is the highest level of situation awareness, it is based on historical and present network security situation information and makes quantitative prediction of the network security situation some period in the future so that decision-maker can have more complete network security situation and provides accurate grounds for reasonable response strategy to restrain network attack.

As to the fuzziness, randomness and uncertainty of future security situation change, it is put forth that gray theory can be adopted for establishing network security situation forecast model. Meanwhile, considering that network security situation awareness system is non-linear and the data is of high random fluctuation, Markov's state transition matrix is adopted to modify gray model's forecast results and make up for the limitation of gray forecast model. Therefore, gray theory and Markov theory are combined to bring the advantages of both to full play and overcome the defects of both, thus Gray Markov forecast model came into being. During the forecast process, the classical GM (1, 1) model of gray theory is adopted to make prediction of network security situation data and find out its changing trend, and then Markov theory is used to make modifications on model error to improve the forecast accuracy of network security situation changing trend.

According to the theory of time series analysis, we propose a new algorithm for network risk prediction. We divide the non-stationary time series into the determine sequence which represents a trend or cyclical regularity and the random sequence. Network intrusion affected by the combined effects of complex factors such as the social development, individual behavior and equipment and technology updates, which show the network risk situation a clear trend and randomness. This network intrusions behaviors mostly follow certain cycle regular of fluctuation. For example, the day average intrusion behaviors follow every 24 hours for the fluctuations in cycle regularity. We use the ARMA model. The notation ARMA (p, q) refers to the model with p autoregressive terms and q moving-average terms.

$$X(t) = \phi_1 X(t-1) + \phi_2 X(t-2) + \ldots + \phi_p X(t-p) +$$

$$u(t) - \theta_1 u(t-1) - \theta_2 u(t-2) - \ldots - \theta_q u(t-q)$$

If q=0 → pure AR (p) process.

If p=0 → pure MA (q) process.

Introducing the lag operator B in our model, then, we can write:

$$\varphi(B)X(t) = \theta(B)u(t)$$

The predictive value of time series of the entire monitoring network risk equals $\{Y(t)\}$ the predictive of time series value of nonlinear fitting plus $\{X(t)\}$ the predictive value of the residual time series. We can write:

$$\hat{R}(t) = R(t) + X(t)$$

On the foundation of AIS based network risk evaluation, we also make evaluation towards short-term, medium-term, long-term network with different period, discuss the randomness of risk changes in real-time and short-time, as well as the periodicity in mid-term and long-term network risk changes. The model takes an overall overview on risk change tendency on every hierarchy of network from different viewpoints; therefore, it can build a safeguard system.

V. Experimental Results and Analysis

The following experiments were carried out in the Laboratory of Computer Network Security. Considering the preciseness and efficiency, we use 12 indicators to evaluate the network danger, which include host danger, area danger, cells danger, special danger etc. An antigen was defined as a fixed length binary string composed of the source/destination IP address, port number, protocol type, IP flags, IP overall packet length, TCP/UDP/ICMP fields, and etc. The network was attacked by 25 kinds of attacks, such as Syn Flood, Land, Smurf, and Teardrop. A total of 20 computers in a network were under surveillance. The task aimed to detect network attacks. Figure 3 illustrates the syn attacks. Figure 4 illustrates the land attacks. And Figure 5 depict the evaluation of the network danger in our model. And we developed some series experiments. Here are the coefficients for the model as the Table 2 showing. As is shown in Figure 5, R (t) changes when attack levels changes. The rise in attack levels is accompanied by a corresponding increase in R (t), as implies the bad network security. On the other hand, if attack levels decline, R (t) decreases accordingly after seconds of delay. Therefore, the network can stays on guard even when the attacks occur once again during a very short time.

TABLE II.
COEFFICIENTS FOR THE MODEL

– Parameter	– Value
– r-contiguous bits matching rule	– 8
– The size of initial self set n	– 40
– The Initial Scale of Detectors	– 100
– Match Threshold β	– 40~60
– Activable Threshold λ	– 50~150
– Clone Scale	– 20
– Mutation Scale	– 19
– The Life Cycle of the Mature Detectors	– 120s

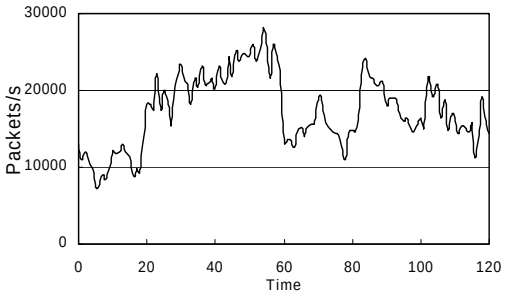

Figure 3. The network suffering from the syn incursions for instance

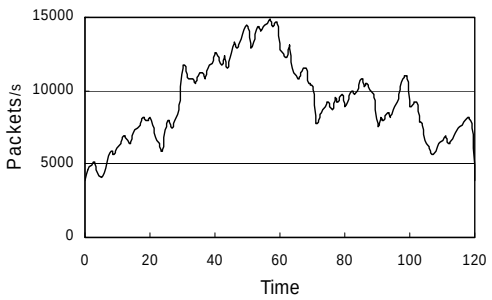

Figure 4. The network suffering from the land incursions for instance

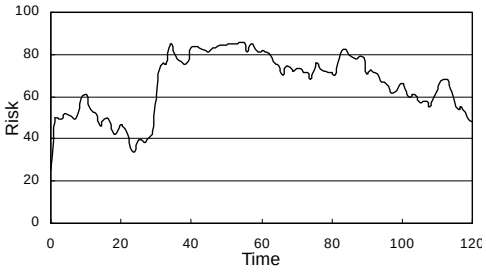

Figure 5. The line of the network dangers obtained by our model at these incursions

VI. CONCLUSIONS

The self-adaptability, distributed character and quantization of antibody concentration that biological immune system bears are just the effective method to solve the technological problems of network security situation awareness, thus this article applies the characteristics of biological immune mechanism in the field of network security situation awareness research, and establishes immune network security situation awareness system to make real-time and quantitative analysis on network security condition and its changing trend. This paper combines the risk evaluation methods with application security engineering principles, and can change current passive defense situation using traditional network security approaches, and is helpful to establish new generation proactive defense theories and realization techniques. At the same time, the work is of not only theoretic values to design proactive defense systems which have intrusion tolerant ability and survivability in any complex network circumstances, but also very significant to protect network infrastructure. The experimental results show that the proposed model has the features of real-time processing that provide a good solution for network surveillance.

ACKNOWLEDGMENT

This work was supported by China Postdoctoral Science Foundation (No.2011M501419), and the National Natural Science Foundation of China (No.61003310, No.61103249) and the Scientific Research Fund of Sichuan Provincial Education Department (No. 10ZB005).

REFERENCES

[1] http: //en.wikipedia.org/wiki/Cloud_Computing.
[2] http: //www.mysql.com/news-and-events/sun-to-acquire-mysql.html
[3] Phurivit Sangkatsanee, Naruemon Wattanapongsakorn, Chalermpol Charnsripinyo. Practical real-time intrusion detection using machine learning approaches, Computer Communications, vol. 34, pp. 2227-2235, 2011
[4] Huwaida Tagelsir Elshoush, Alert correlation in collaborative intelligent intrusion detection systems-A survey, Applied Soft Computing, vol. 11, pp. 4349-4365, 2011.

[5] Fatemeh Amiri, MohammadMahdi Rezaei Yousefi, Mutual information-based feature selection for intrusion detection systems, Journal of Network and Computer Applications, vol. 34 (4), pp. 1184-1199, 2011
[6] Vincent Toubiana, Houda Labiod, Laurent Reynaud. A global security architecture for operated hybrid WLAN mesh networks. Computer Networks, vol. 54 (2), pp. 218-230, 2010
[7] Kuby J., Immunology. Fifith Edition by Richard A. Goldsby et al.
[8] F.M.Burnet. The Clone Selection Theory of Acquired Immunity. Gambridge, Gambridge University Press, 1959
[9] S A Hofmeyr, and S Forrest. Architecture for an artificial immune system. Evolutionary Computation, vol. 8, pp. 443-473, 2000
[10] S Forrest, A S Perelson, L Allen, and R Cherukuri. Self-Nonself Discrimination in a Computer. Proceedings of IEEE Symposium on Re-search in Security and Privacy, Oakland, 1994
[11] Serap Atay, Marcelo Masera. Challenges for the security analysis of Next Generation Networks, Information Security Technical Report, vol. 16, pp. 3-11, 2011
[12] M. Mezmaz, N. Melab, Y. Kessaci. A parallel bi-objective hybrid metaheuristic for energy-aware scheduling for cloud computing systems, Journal of Parallel and Distributed Computing, vol. 71, pp. 1497-1508, 2011
[13] Md. Tanzim Khorshed. A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing, Future Generation Computer Systems, vol. 28, pp. 833-851, 2012
[14] Manel Bourguiba, Kamel Haddadou, Packet aggregation based network I/O virtualization for cloud computing, Computer Communications, vol. 35 (3), pp. 309-319, 2012
[15] Changbok, Hyokyung Chang, Filtering Technique on Mobile Cloud Computing, Energy Procedia, vol. 16, 1305-1311, 2012
[16] Tao Li. An immunity based network security risk estimation, Science in China Ser. F Information Sciences. vol. 48, no. 2005, pp. 557- 578

Jin Yang received his M.S. degree and the Ph.D. degree in computer science from Sichuan University, Sichuan, China. He is an Associate Professor in Department of Computer Science at LeShan normal university. His main research interests include network security, artificial immune, knowledge discovery and expert systems.

Cilin Wang received his M.S. degree in computer science from Sichuan University, Sichuan, China. He is an Associate Professor in Department of Computer Science at LeShan normal university. His main research interests include mathematics, knowledge discovery.

LeYu received his Bachelor's degree in computer science from Sichuan University, Sichuan, China. His main research interests include computer network, electronic technology.

Caiming Liu received his Ph.D. degree in computer science from Sichuan University, Sichuan, China. He is an Associate Professor in Department of Computer Science at LeShan normal university. His main research interests include network security and artificial immune.