SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

October 2012 Vol. 55 No. 10: 2358–2368
doi: 10.1007/s11432-011-4451-0

An immune-theory-based model for monitoring inter-domain routing system

GUO Yi* & WANG ZhenXing

National Digital Switching System Engineering & Technological R&D Center, Zhengzhou 450002*, China*

Received February 11, 2011; accepted August 8, 2011; published online January 2, 2012

Abstract The inter-domain routing system faces many serious security threats because the border gateway protocol (BGP) lacks effective security mechanisms. However, there is no solution that satisfies the requirements of a real environment. To address this problem, we propose a new model based on immune theory to monitor the inter-domain routing system. We introduce the dynamic evolution models for the "self" and detection cells, and construct washout and update mechanisms for the memory detection cells. Furthermore, borrowing an idea from immune network theory, we present a new coordinative method to identify anomalous nodes in the inter-domain routing system. In this way, the more nodes working with their own information that join the coordinative network, the greater is the ability of the system to identify anomalous nodes through evaluation between nodes. Because it is not necessary to modify the BGP, the ITMM is easy to deploy and inexpensive to implement. The experimental results confirm the method's ability to detect abnormal routes and identify anomalous nodes in the inter-domain routing system.

Keywords border gateway protocol, inter-domain route, immune theory, anomalous node, coordinative identification

Citation Guo Y, Wang Z X. An immune-theory-based model for monitoring inter-domain routing system. Sci China Inf Sci, 2012, 55: 2358–2368, doi: 10.1007/s11432-011-4451-0

1 Introduction

As important infrastructure of the Internet, the inter-domain routing system is composed of many self-governing autonomous systems (ASs). Each AS is identified by a unique numerical ID obtained from regional Internet registries and inter-connected by the border gateway protocol (BGP), which is used to exchange reachability information and ultimately perform path selection. However, the BGP has some design flaws which result in many serious security issues for the inter-domain routing system [1–4], such as prefix hijacking and AS_PATH tampering attacks. When an AS announces a prefix that does not belong to the AS, data packets, whose destination addresses are included in the prefix, are routed to the malicious AS. This type of event happens frequently; for example, AS 174 hijacked the prefix of Google on May 7, 2005 [5], while Pakistan Telecom hijacked the prefix of YouTube on February 24, 2008 [6]. An AS_PATH tampering attack occurs when a route announced by an AS violates a certain BGP policy or contains an invalid AS number. This can be further classified as a forge shortest-path attack, route redistribution attack or forge AS number attack. Many solutions have been presented to enhance the

*Corresponding author (email: guoyi2006@yeah.net)

security of the inter-domain routing system. These solutions fall into two categories: BGP extended approaches and BGP security monitoring.

BGP extended approaches, such as secure BGP [7] and secure origin BGP [8], use public-key infrastructure to ensure creditability of BGP routes exchanged between different ASs. However, because of the high computation overhead and the high cost of updating equipment, these methods have not been accepted by many Internet service providers (ISPs). Consequently, various improved solutions have been developed, such as inter-domain route validation [9], Listen and Whisper [10], origin authentication [11], pretty secure BGP [12], and the security path vector [13]. However, these improved solutions have not been widely deployed.

Given the difficulty of deploying a secure inter-domain routing protocol [14], security monitoring systems have been developed in recent years. These systems significantly enhance the security of the inter-domain routing system through merely deploying several monitoring nodes. Compared with previous approaches, they are cheaper and easier to deploy because there is no need to construct public-key infrastructure or modify the BGP.

The PHAS system developed by Lad et al. [15] addressed the prefix hijacking problem. It first analyzes the BGP routing data provided by RouteViews (or RIPE) and then generates a report on the use of the specified prefix of concern to customers. If any unauthorized AS announces the prefix, the PHAS system alerts the owner that it has suffered prefix hijacking.

Liu et al. [16] analyzed the hierarchical characteristics of the inter-domain routing system and proposed a security evaluation model. In their paper, they described the hierarchical relationship of various routing entities based on constructing a route status tree. From analysis of known abnormal routes, the model can quantify the security state of each entity. However, this method requires users to provide abnormal BGP data, which is the most difficult procedure in security monitoring. Therefore, the method is only fit to quantify the security threat status of the inter-domain routing system.

From the above analysis, it is clear that existing solutions cannot support the detection of abnormal inter-domain routes nor recognize malicious BGP nodes. Thus, a new model based on immune theory for monitoring the inter-domain routing system, termed the ITMM, is proposed. The proposed model has greater ability to detect abnormal inter-domain routes and identify malicious nodes. First, we introduce dynamic evolution models for the "self" and detection cells, and construct washout and update mechanisms for the memory detection cells. Following an idea from immune network theory, a new method to identify anomalous nodes is then presented. In this way, the more nodes joining the coordinative network, the greater is the ability of the system to identify anomalous nodes through evaluation between nodes.

2 BGP routing event

2.1 Finite-state machine (FSM) for BGP nodes

Definition 1. Let s describe the running state of a BGP node at different times and $s_i(t)$ denote the state of node i at time t. For any node, Figure 1 shows its state machine. A node has one of three running states: the initial state where $s_i(t) = 0$, the normal state where $s_i(t) = 1$, and the failed state where $s_i(t) = -1$. The failed state means a node cannot exchange information with other nodes owing to overload or invalidity.

As shown in Figure 1, a BGP node moves from one state to another when satisfying specific conditions. For example, a node in the initial state needs to load the static routing table and accept the routing information of all its peers. It then enters its normal state. During this period, it receives and calculates routing UPDATE packets from its neighbors (including outdated routes) to maintain the timeliness of the routing table. Meanwhile, it has an obligation to send its routing information to its neighbors. If the workload becomes greater than its rated load, the node becomes overloaded and moves into the failed state. If the problem cannot be resolved, the node fails and then restarts in the initial state; otherwise, it moves into the normal state.

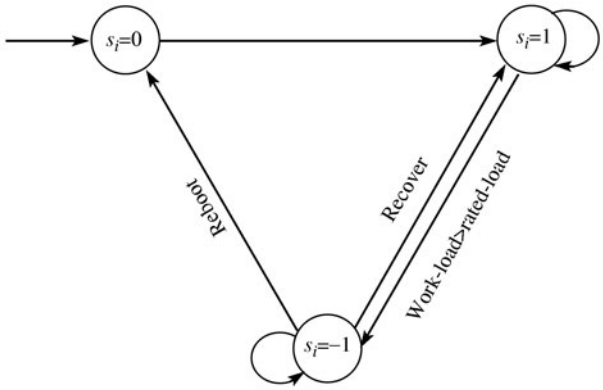

Figure 1 FSM for BGP nodes.

2.2 Classification of routing events

The inter-domain routing system is composed of many interconnected ASs. For the sake of simplicity, each AS is modeled by a single BGP node in this paper. These nodes exchange routing information within the BGP protocol, which contains four kinds of messages: the OPEN message, the KEEPALIVE message, the UPDATE message, and the NOTIFICATION message. Of these, the UPDATE message is the most important message used to announce new routes or outdated routes; that is, the exchange of information among ASs is realized mainly through UPDATE messages. When a BGP node fails, all its peers withdraw the related routing information. Let Γ_i denote the set of nodes that are adjacent to i. The above process can then be described as $\forall j \in \Gamma_i$, if $Failure(i)$, then $SentUpdate(j,k)$, where $k \in \Gamma_j$ and $k \neq i$. Moreover, if the node restarts, all its adjacent nodes send their own routing information to the node, which is expressed as $\forall j \in \Gamma_i$, if $Enable(i)$, then $SentUpdate(i,j)$. The node then updates its routing table and sends its routes to all its peers in turn, which is expressed as $\forall((k \in \Gamma_i) \wedge (k \neq j))$, if $Receive(i,j)$, then $SentUpdate(i,k)$.

Given an unweighted undirected graph $G = (AS, E_{AS})$, AS is the set of ASs and E_{AS} is the set of AS links. Each AS is identified by its unique AS number and an AS link describes a route between two ASs. Furthermore, the inter-domain routing system can be described as a graph $G = (R, E_{\text{router}})$, where $R = \{R_1, R_2, \ldots, R_N\}$ is the set of routers and $E_{\text{router}} = \{e^i_j | i \in R, j \in R \text{ and } i \neq j\}$ is the set of links between different routers. In general, a route between router i and router j at time t can be defined as $e^i_j(t) = (nextHop, flag)$, where $nextHop$ denotes the next hop router and $flag$ is used to denote whether the route is studied from the inter-domain (o) router or the intra-domain (i) router. The BGP routing events can be classified into five categories.

• If $e^i_j(t-1).nextHop = e^i_j(t).nextHop$ and $e^i_j(t-1).flag = e^i_j(t).flag$, then the event is a no-change (NC) routing event.

• If $e^i_j(t-1).nextHop \neq e^i_j(t).nextHop$ and $e^i_j(t-1).flag = e^i_j(t).flag = i$, then the event is an internal path change (IPC) routing event.

• If $e^i_j(t-1).nextHop \neq e^i_j(t).nextHop$ and $e^i_j(t-1).flag = e^i_j(t).flag = o$, then the event is an external path change (EPC) routing event.

• If $e^i_j(t-1).nextHop = e^i_j(t).nextHop$, $e^i_j(t-1).flag = o$ and $e^i_j(t).flag = i$, then the event is a loss of egress point (LEP) routing event.

• If $e^i_j(t-1).nextHop = e^i_j(t).nextHop$, $e^i_j(t-1).flag = i$ and $e^i_j(t).flag = o$, then the event is a gain of egress point (GEP) routing event.

3 Immune model

Both the structure and function of an inter-domain routing security monitoring system are similar to those of an immune system. First, both are composed of many independent objects that interact with each other in various ways [17,18]. The independent objects in a biological immune system are a variety

of immunocytes, whereas they are BGP monitoring nodes in an inter-domain routing monitoring system. Second, the goal of both systems is to make the protected system more secure. The main function of an immune system is to identify and restrain malicious antigens according to the principle that what is not a "self" must be a "non-self". Accordingly, the goal of deploying the monitoring system is to monitor the progress of exchanging routing information among ASs, and to detect abnormal routing events including abnormal routes and anomalous nodes, thereby ensuring the security of the inter-domain routing system.

As an immune system is good at self-learning and is adaptive, we propose the ITMM based on immune theory for monitoring the inter-domain routing system. Specifically, the anomaly detection borrows immunity mechanisms used to identify "self" and "non-self", such as immune memory and negative selection. Furthermore, based on dynamic immune network theory, a new method for identifying anomalous nodes is presented. In this way, the system can identify anomalous nodes through the mutual evaluation between nodes.

3.1 Definitions

Definition 2. In the ITMM model, $Self$ denotes the set of normal routes and $Nonself$ is the set of abnormal routes. Let Ag be the set of antigens (routes); then $Nonself \subset Ag$, $Self \cup Nonself = Ag$ and $Self \cap Nonself = \emptyset$. For any antigen $x \in Ag$, $x.b$ is the initial UPDATE message and $x.a$ is the characteristic presented by the UPDATE message; that is, $x.a = APC(x.b)$, where the function $APC()$ describes the process of antigenic presentation. Additionally, if $x.a \cap APC(Nonself) \neq \emptyset$, then $x \in Nonself$; else $x \notin Nonself$.

Definition 3. The detection cell that is used to detect abnormal routes simulates a lymphocyte, and the set of detection cells is defined as $D = \{\langle d, age, count\rangle | d \in X, age \in N, count \in N\}$, where age is the age of a detection cell and $count$ is the number of abnormal routes successfully recognized by a detection cell. Moreover, we define the set of memory cells as $D_{\text{memory}} = \{x | x \in D, x.count > \eta\}$, where η is the threshold value of $count$.

Definition 4. The information used to identify abnormal routes is called the $proof$, described as $P = \{\langle t, x\rangle | t \in N, x \in Ag\}$, where t denotes the time at which the proof is obtained and x is an abnormal UPDATE message.

Definition 5. Affinity describes the power of interaction between the antibody and the antigenic determinant in immunology. Correspondingly, it is used to describe the power of interaction between detection cells and detecting routes. Let d be a detection cell and r be the detecting route; then their affinity can be described as $f_{\text{affinity}}(d, r)$. If the value of $f_{\text{affinity}}(d, r)$ is greater than a certain threshold value ε, we say that d matches with r; that is, if $f_{\text{affinity}} \geqslant \varepsilon$, then $f_{\text{recognize}} = 1$; else $f_{\text{recognize}} = 0$.

3.2 Evolution model

To avoid having normal routes identified as abnormal routes, the detection cell must go through a self-tolerance process [19]. The goal of this process is to eliminate cells that can recognize self-cells. It is described as

$$f_{\text{tolerance}}(D) = D - \{d | d \in D \wedge \exists r \in Self(f_{\text{recognize}}(d, r) = 1)\}, \tag{1}$$

where D is the set of initial detection cells and $Self$ is the self-set. The evolution of D can then be defined as

$$D(t) = \begin{cases} D_{\text{initial}}, & t = 0, \\ D_{\text{tolerance}}(t) - D_{\text{aged}}(t) + D_{\text{new}}(t), & t \geqslant 1, \end{cases} \tag{2}$$

where

$$D_{\text{tolerance}}(t) = D(t-1) - \{d | d \in D(t-1) \wedge \exists r \in Self(f_{\text{recognize}}(d, r) = 1)\} \tag{3}$$

and

$$D_{\text{aged}}(t) = \{d | d \in D_{\text{tolerance}}(t) \wedge d.age > \beta\}. \tag{4}$$

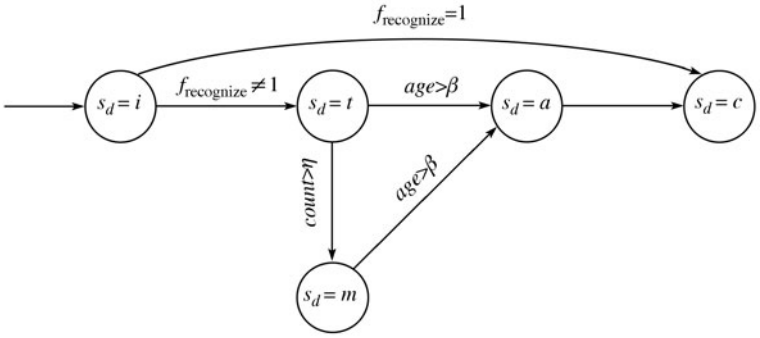

Figure 2 FSM of detection cells.

In Eq. (2), D_{initial} denotes the set of initial immature cells, which are randomly generated from some known rules, and D_{new} is the set of new detection cells. Thus, D is mature only after a period of tolerance in which cells that can recognize self-cells are deleted. For any detection cell, Figure 2 shows its state machine. A detection cell has one of five states: i (immature), t (tolerant), a (aged), c (cross-out) or m (memonic); this is described as $S_d = \{i, t, a, c, m\}$.

The evolution of $Self$ can be defined as

$$Self(t) = \begin{cases} Self_{\text{initial}}, & t = 0, \\ Self(t-1) + \varepsilon Ag_{Self}(t-1), & t \geqslant 1, \end{cases} \tag{5}$$

where the value of ε is -1 (withdrawing dated routes) or 1 (updating new routes) and Ag_{Self} is defined by Eq. (6). When detecting the abnormal routes, the initial self-set $Self_{\text{initial}}$ is the static routing table of a node (router) $route(0) = \{static\_r_0, static\_r_1, \ldots, static\_r_m\}$; that is, $Self(0) = route(0)$.

$$Ag_{\text{self}}(t) = Ag(t) - Ag_{\text{Nonself}}(t), \tag{6}$$

$$Ag_{\text{Nonself}}(t) = \{r | r \in Ag(t) \wedge \exists d \in D(t) \wedge f_{\text{recognize}}(d, r) = 1\}. \tag{7}$$

As shown in Eqs. (6) and (7), when a node receives an UPDATE message ($Ag(t)$) from a peer node at time t, it first uses the mature set to detect the antigen set $Ag(t)$ and then places the antigens that are recognized into the non-self-set $Ag_{\text{Nonself}}(t)$, while the other antigens are accepted into the self-set $Ag_{\text{Self}}(t)$.

3.3 Immune memory mechanism

A biological immune system can remember the characteristics of invading antigens [20]. If the system is invaded by kindred antigens, it can quickly identify and restrain the antigens. In a similar manner, for each detection cell, once $d.count > \eta$, we put it into D_{memory} as a memory detection cell. We then first utilize D_{memory}, which is a small set with high matching probabilities, to detect abnormal routes, thus improving the efficiency of detection.

The inter-domain routing system is a giant dynamic complex system [21,22]. A route that was identified as normal at time i may be abnormal at time j. In addition, the scale of D_{memory} increases over time if inapplicable and rarely used cells still exist, which would result in the detection of an error or a sharp decline in detection efficiency. In other words, not only will $Self/Nonself$ evolve, but also the detection cells and memory detection cells need to be updated [23,24]. Thus, we establish a washout and update mechanism for the memory detection cells as follows.

(1) $\forall d \in D \wedge \exists r \in Nonself$, if $f_{\text{recognize}}(r, d) = 1$, then $(d.count++)$;

(2) $\forall d \in D$, if $d.count > \eta$, then $D_{\text{memory}} = D_{\text{memory}} + \{d\}$;

(3) $\forall d \in D_{\text{memory}} \wedge \exists r \in Self$, if $f_{\text{recognize}}(r, d) = 1$, then $D_{\text{memory}} = D_{\text{memory}} - \{d\}$;

(4) $\forall d \in D_{\text{memory}}$, if $d.count \leqslant \eta$, then $D_{\text{memory}} = D_{\text{memory}} - \{d\}$;

(5) $\forall d \in D_{\text{memory}}$, if $d.age > \beta$, then $\{d.age = 0, d.count = 0\}$.

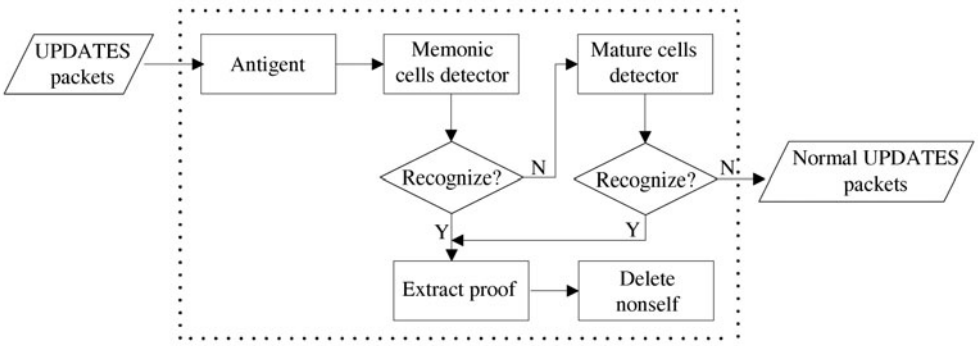

Figure 3 Workflow for detecting UPDATES.

3.4 UPDATE anomaly detection

Figure 3 shows the flow of detecting UPDATE messages. Let $UP = \{up_1, up_2, \ldots, up_n\}$ denote the input UPDATE messages. After detection, the output is $UP' = \{up_i | up_i \notin Nonself\}$, which does not contain any abnormal routes. The detection process can be divided into four steps: antigenic presentation, anomaly detection with D_{memory}, anomaly detection with D_{mature}, and deletion of abnormal routes.

• Antigenic presentation: $APC(\{up_1, up_2, \ldots, up_n\}) \to \{Ag|(x_1.a, x_1.b), (x_2.a, x_2.b), \ldots, (x_n.a, x_n.b)\}$.

• Anomaly detection with D_{memory}: $\forall x_i \in Ag \wedge \exists d \in D_{\text{memory}}$, if $f_{\text{recognize}}(x_i, d) = 1$, then $Extract(P(x_i)) \to (P(x_i).t, x_i)$.

• Anomaly detection with D_{mature}: $\forall x_i \in Ag \wedge \exists d \in D_{\text{mature}}$, if $f_{\text{recognize}}(x_i, d) = 1$, then $Extract(P(x_i)) \to (P(x_i).t, x_i)$.

• Deletion of abnormal routes: $UP - \{x_i | x_i \in Nonself\} \to UP'$.

It is well known that the emergence of non-self antigens in a biological immune system generally has the characteristic that the majority of non-self antigens are of the same type in a particular period. Through long-term study, we have found that the occurrence of abnormal routing events in the inter-domain routing system has a similar feature. Therefore, we use D_{memory} to detect abnormal routes, which reduces the number of antigens that could move into the Mature Cells Detector, thereby improving the detection efficiency.

4 Coordinative identification model

The inter-domain routing system has a self-organizing property that once malicious behavior has occurred at a certain AS, the best way for the system to remain stable is not to eliminate the AS node, but to recognize and inhibit its activities until the AS node returns to normal [18]. As this behavior is similar to the immune mechanism for maintaining balance, we present a coordinative identification method based on dynamic immune network theory to identify anomalous BGP nodes in the inter-domain routing system.

4.1 Definition of the model

Definition 6. A BGP node that participates in coordinative identification is called a coordinative node $as_{ci} \in AS$, and the network consisting of various coordinative nodes is called a coordinative identification network (CIN). Given a CIN as an unweighted undirected graph $G_{CIN} = (AS_{ci}, E_{AS_{ci}})$, AS_{ci} is the set of coordinative nodes as_{ci}, and $E_{AS_{ci}}$ is the set of edges, with each edge e_{ij} denoting a link between $as_{ci}i$ and $as_{ci}j$.

Definition 7. A coordinative node can evaluate the operational status of its peer nodes. Let $evaluate_{ij}(t)$ denote the status assessment of $as_{ci}i$ to $as_{ci}j$. Then $evaluate_{ij}(t) = 1$ means $as_{ci}i$ thinks that $as_{ci}j$ is normal, while $evaluate_{ij}(t) = 0$ means $as_{ci}i$ thinks that $as_{ci}j$ is abnormal.

Definition 8. Eigenvalue r_i is used to quantify the operational status of $as_{ci}i$, and $R_i(0 \leqslant R_i \leqslant 1)$ denotes the standardized value of r_i. $R_i = 0$ means $as_{ci}i$ is normal, while $R_i = 1$ means $as_{ci}i$ is abnormal. There are three factors that may affect the value of r_i: $evaluate_{ji}$ $(j \in \Gamma_i)$, $evaluate_{ij}$ $(j \in \Gamma_i)$, and $r_i(t)$. Additionally, if $evaluate_{ij}$ differs from $evaluate_{kj}$ $(k \in \mathrm{T}_j \wedge k \neq i)$, then it can be assumed that $as_{ci}i$ is abnormal.

4.2 Adjustment mechanism

The key to the coordinative identification method is the adjustment of eigenvalues, a process that is continuous. There are many adjustment methods, such as gray model adjustment, black model adjustment, and white model adjustment.

Because BGP is a policy-based routing protocol, the routing strategy is determined by the relationship among nodes. There are two main relationships: provider-customer and peer-peer. If the relationship between two nodes is provider-customer, the provider must provide Internet access services to the customer and send all the necessary routes to the customer. On the contrary, if the relationship is peer-peer, each of nodes only sends its own routes and customers' routes to the other peering node. Additionally, the information in the inter-domain routing system includes not only known information but also unknown or uncertain information. Information such as the number of BGP nodes and their connections are known; however, some information such as exchange routes and the system topology may be uncertain owing to dynamic changes in the system. Thus, the inter-domain routing system is a gray system.

On the basis of the above analysis, we adopt the gray model as our adjustment method. The adjustment can then be described as

$$dr_i(t)/dt = \sum_{j \in \Gamma_j} T_{ji}^{+} R_j(t) - r_i(t), \tag{8}$$

where Γ_i denotes the set of peer nodes of i and $R_i(t) = 1/(1 + \exp(-r_i(t)))$. The value of T_{ij}^{+} is given by the following rules. If the relationship of i and j is peer-peer, then $T_{ij}^{+} = T_{ij} + T_{ji} - 2$; else if it is customer-provider, then $T_{ij}^{+} = T_{ij} + T_{ji} - 1$; else if it is provider-customer, then $T_{ij}^{+} = 0$. Additionally, if both i and j are normal, then $T_{ij} = 1$; else $T_{ij} = -1$. This ensures that only the high-level nodes are entitled to evaluate the operational status of lower level nodes; that is, the stability of nodes in a core layer cannot be affected by the lower nodes.

In the inter-domain routing system, some events may affect the stability of the system, but after a period of adjustment, the system reaches a new equilibrium. The R value of a node at this moment represents whether it is normal or abnormal. For instance, at the beginning, the system is in a balanced state; that is, all nodes are running properly $(R = 1)$. Once the R value of a node is incorrectly modified to zero, the balance is destroyed, but after a period of adjustment, the system returns to the original equilibrium. On the contrary, if an exception occurs at a node, the balance is again broken. After a certain period of dynamic adjustment, the system moves into another state, and the R value of the abnormal node is adjusted to zero. In other words, at any time, it is known which nodes are abnormal by observing the state of the system.

5 Experiment and evaluation

5.1 Experimental data

To validate the effectiveness of the model in practical application, we selected as experimental data the BGP AS links data provided by the CAIDA project in January 2010. Utilizing our BGP Information Statistic System (ISS) to analyze the data, we obtain topological information of the inter-domain routing system, which can be used directly to simulate the ITMM. Figures 4 and 5 show the related information in detail.

# of ASs	33508
# of transit ASs	21262
# of stub ASs	12246
# of links	75001
Max degree	2631
Min degree	1
Mean degree	4.48

Figure 4 Related topological information.

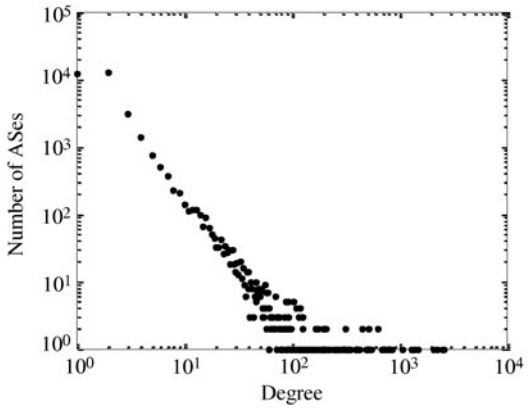

Figure 5 Degree distribution of the inter-domain routing system.

5.2 Performance evaluation

5.2.1 *Capability of detecting abnormal routing events*

Two metrics are used to assess the ability of the model to detect abnormal routes. The first is the true positive rate (TP), which is the ratio of correctly detected abnormal routes to the total number of abnormal routes. The other metric is the false positive rate (FP), which is the ratio of routes that are mistaken for abnormal to the total number of normal routes.

The experimental dataset has 103870 UPDATE messages randomly extracted from routes provided by RouteViews, 99% of which are normal UPDATE messages; that is, there are 102840 normal messages and 1030 abnormal messages. We divided the experiment into two parts: investigation of the effect of crucial parameters on model performance and investigation of the ability of the ITMM to detect different kinds of abnormal routing events.

Both η (matching threshold) and β (life cycle) are adjustable parameters. Their values can be set dynamically according to the required TP or FP, the application environment, and other specific conditions. Figures 6 and 7 show how the values of η and β affect the detection performance. It should be noted that the initial D_{memory} consists of elements that possess characteristics of typical abnormal routes.

Parameter η is used to determine whether a mature detection cell satisfies the condition to become a memory detection cell. Once a mature cell has identified k abnormal routes where $k \geqslant \eta$, the mature cell turns into a memory cell. As shown in Figure 6, the TP and FP values decrease with an increase in η. The reason for this is that the value of η directly affects the number of memory detection cells. If η becomes small, a considerable number of small groups generate memory detection cells. This results in the dispersion of a large number of memory detection cells, resulting in high values for TP and FP. On the other hand, if η becomes large, it is not easy for a mature cell to become a memonic, and the number of memory detection cells is thus small or even zero, which could lead to both TP and FP being low.

The role of β is to wash out the aging memory detection cells. As shown in Figure 7, as β increases, so do the values of TP and FP. The reason for this is that β determines the survival of memory detection cells. If β is low, most of the memory cells are eliminated in a very short period of time; that is, the total number of memory cells is small, which results in low values of TP and FP. On the contrary, if β is large, memory cells survive longer, which results in a sharp increase in the number of memory cells, causing both TP and FP to increase.

Table 1 lists the experimental results for the ITMM detecting different types of abnormal routes. It is seen that when detecting routes containing an illegal prefix, AS number or AS cycle, the method returns a high TP and low FP. Meanwhile, it has a great ability to detect abnormal routes of fake route type and multiple-origin AS (MOAS) conflict type.

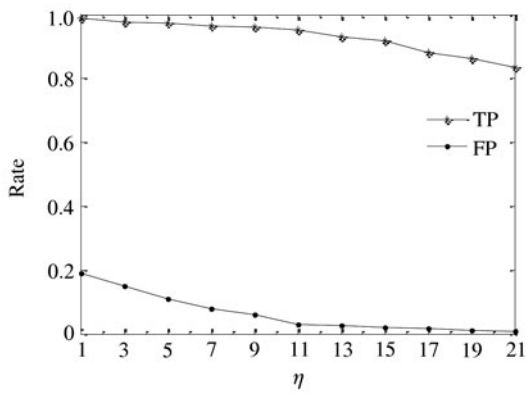

Figure 6 Effect of η on model performance.

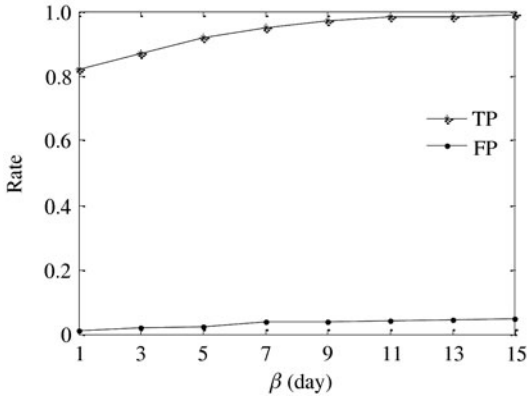

Figure 7 Effect of β on model performance.

Table 1 Capability of detecting different types of abnormal routes

Anomalous type	Illegal prefix	Illegal AS	Cyclic path	Fake route	MOAS conflict
# of events	253	319	102	189	167
TP (%)	97.63	96.84	100	87.31	89.37
FP (%)	3.07	4.19	0	7.68	9.12

5.2.2 *Capability of identifying anomalous nodes*

The coordinative identification method needs to deploy a certain number of BGP nodes as coordinative nodes to construct a CIN network. If an AS joins the CIN, it becomes a coordinative node. The system identifies existing anomalous nodes through evaluations of coordinative nodes on other nodes. To verify the ability of the coordinative method, we introduce a new evaluation variable φ_s, the definition of which is given below.

Definition 9. φ_s is used to describe the ratio of nodes that can be accurately identified to the total number of nodes. Its value is given by $\varphi_s = (\sum_{i \in IDRS} \phi(i))/N$, where N is the total number of ASs in the inter-domain routing system and $\phi(i)$ represents whether a node can be identified. The value of $\phi(i)$ is

$$\phi(i) = \begin{cases} 1, & \left(\sum_{j \in \Gamma_i} \tau_j\right) \Big/ M \geqslant \xi, \\ 0, & \left(\sum_{j \in \Gamma_i} \tau_j\right) \Big/ M < \xi, \end{cases} \tag{9}$$

where ζ is a control variable used to adjust the percentage of adjacent nodes joining the CIN to ensure that a node can be accurately evaluated, Γ_i is the set of nodes adjacent to AS i, M is the number of nodes adjacent to AS i, and τ_j represents whether AS j is a coordinative node (1) or not (0).

The ASs are ordered according to their node degree from large to small, and join the CIN network in this order. The experimental results are shown in Figures 8 and 9.

Figure 8 shows that the greater the value of ζ, the higher is the accuracy in identifying anomalous nodes. When ζ is 0.9, the accuracy is higher than 99%, and thus, as long as ζ is 0.9, reliability of the identification results can be ensured. Figure 9 shows the experimental results for testing the monitoring ability of the coordinative identification method. It is seen that the more nodes joining the CIN, the larger the part of the system that can be effectively monitored. Moreover, if just 1% of nodes become coordinative nodes, the effective monitoring scope rises to nearly 45%, while 13% of the nodes joining the CIN can extend the effective monitoring scope to more than 80%.

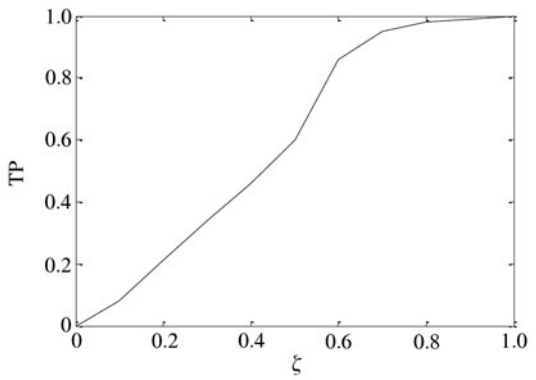

Figure 8 Effect of ξ on identification accuracy.

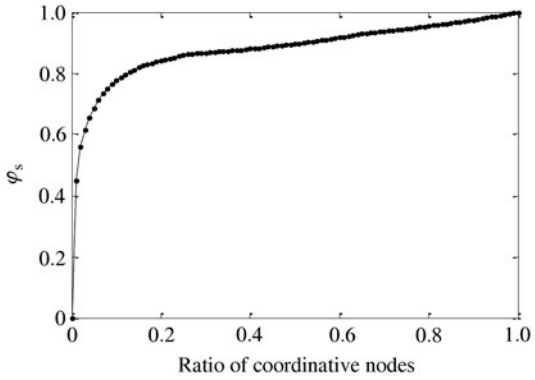

Figure 9 Monitoring capability versus number of coordinative nodes.

6 Conclusions

There are many security problems that have not been adequately solved in the BGP-based inter-domain routing system. To this end, we have proposed a model to monitor the inter-domain routing system on the basis of immune theory. This model has great ability to detect abnormal routes and identify anomalous nodes. First, we introduced dynamic evolution models for self and detection cells, and constructed washout and update mechanisms for memory detection cells. Borrowing from the immune network theory, we then presented a new method to identify anomalous nodes. In this way, the more nodes working with their own information that join the coordinative network, the greater is the ability of the system to identify anomalous nodes through evaluation between nodes. Because there is no need to modify the BGP, the ITMM is easy to deploy and inexpensive to implement. The experimental results show the method's ability to detect abnormal routes and identify anomalous nodes in the inter-domain routing system. In summary, the ITMM provides strong support for the security of the inter-domain routing system.

Acknowledgements

This work was supported by National Basic Research Program of China (Grant No. 2007CB307102), and National High-Tech Research & Development Program of China (Grant No. 2007AA01Z2A1).

References

1 Rekhter Y, Li T, Hares S. A border gateway protocol 4 (BGP-4). http://www.ietf.org/rfc/rfc4271.txt
2 Siganos G, Faloutsos M. Neighborhood watch for internet routing: Can we improve the robustness of internet routing today? In: Proceedings of IEEE INFOCOM, Anchorage, Alaska, USA, 2007
3 Chavali S, Radoaca V, Miri M, et al. Peer prefix limits exchange in BGP IETF draft. http://tools.ietf.org/html/draft-chavali-bgp-prefixlimit
4 Barbir A, Murphy S, Yang Y. Generic threats to routing protocols. http://www.ietf.org/rfc/rfc4593.txt
5 Wan T, Oorschot C. Analysis of BGP prefix origins during Google's May 2005 outage. In: Spirakis P, ed. Proc. of the Security in Systems and Networks. Washington: IEEE Computer Society Press, 2006. 8–15
6 Karlin J, Forrest S, Rexford J. Autonomous security for autonomous systems. Comput Netw, 2008, 52: 2908–2923
7 Kent S, Lynn C, Seo K. Secure border gateway protocol (S-BGP). IEEE J Select Areas Commun (Special Issue on Network Security), 2000, 18: 582–592
8 White R. Securing BGP through secure origin BGP. Int Protocol J, 2003, 6: 15–22
9 Goodell G, Aiello W, Griffin T, et al. Working around BGP: An incremental approach to improving security and accuracy of inter-domain routing. In: Proc. Of the ISOC NDSS 2003, San Diego, 2003. 75–85
10 Subramanian L, Roth V, Stoica I, et al. Security mechanisms for BGP. In: Proc. of the 1st Symp. on Networked Systems Design and Implementation (NSDI 2004), San Francisco: USENIX, 2004. 127–140
11 Aiello W, Ioannidis J, McDaniel P. Origin authentication in Inter-domain routing. In: Proc. of the 10th ACM Conf. on Computer and Communications Security. Washington: ACM, 2003. 165–178

12 Wan T, Kranakis E, Oorschot P C. Pretty secure BGP (psBGP). In: Proc of the 12th Annual Network and Distributed System Security Symposium (NDSS'05). San Diego, California: Internet Society, 2005
13 Hu C, Perring A, Sirbu M. SPV: Secure path vector routing for securing BGP. In: Yavatkar R, ed. Proc. of the ACM SIGCOMM. Washington: ACM Press, 2004. 179–192
14 Gao L, Gong Z H, Liu Y P, et al. A TLP approach for BGP based on local speculation. Sci China Ser F-Inf Sci, 2008, 38: 1663–1678
15 Lad M, Massey D, Pei D. PHAS: a prefix hijack alert system. In: Proc of the 15th USENIX Security Symposium (USENIX-SS'06): Vancouver BC Canada USENIX Association, 2006. 18–l19
16 Liu X, Wang X Q, Zhu P D, et al. Security evaluation for inter-domain routing system in the Internet (in Chinese). J Comput Res Devel, 2009, 46: 1669–1677
17 Lu X C, Zhao J J, Zhu P D, et al. Self-organization of inter-domain routing system. Chin J Softw, 2006, l7: 1922–1932
18 Wang L, Liu X Y. A study on a coordinative immune-computing model. Acta Electr Sin, 2009, 37: 1739–1744
19 Esponda F, Forrest S, Helman P. A formal framework for positive and negative detection. IEEE Trans Syst Man Cybern B, 2004, 34: 357–373
20 Erica K. Inspired by immunity. Nature, 2002, 415: 468–470
21 Deng W P, Zhu P D, Lu X C. On evaluating BGP routing stress attack. J Commun, 2010, 5: 13–22
22 Prehofer C, Bettstetter C. Self-organization in communication networks: Principles and design paradigms. IEEE Commun Mag, 2005, 43: 78–85
23 Li T. An immune based model for network monitoring. Chin J Comput, 2006, 29: 1515–1522
24 Zhang P T, Wang W, Tan Y. A malware detection model based on a negative selection algorithm with penalty factor. Sci China Inf Sci, 2010, 53: 2461–2471