A Defense Model of Reactive Worms Based on Dynamic Time

Haokun Tang
College of Economics and Management, SouthWest University, Chongqing, P.R.China
HanHua Financial Holding, P.R.China
Email: tanghaokun@hanhua.com

[1]Shitong Zhu, [2]Jun Huang and [3]Hong Liu
[1,2]School of communication and information engineering, Chongqing University of Posts and Telecommunications, Chongqing, P.R.China
[3]School of Software, Chongqing University of Posts and Telecommunications, Chongqing, P.R.China
Email: [1]zstchina1993@gmail.com; [2]xiaoniuadmin@gmail.com; [3]liuhong1@cqupt.edu.cn

***Abstract*—The popularity of reactive worms, whose attacking behavior inherits characteristics from both active worms and passive worms, has brought great threat to P2P networks in recent years. Most existing defense models only focus on the effects of P2P churn on reactive worm's propagation, but neglect the impact of user behaviors on the spread of worms. This paper proposes a defense model of reactive worms based on dynamic time with full consideration of various dynamic factors that restrict the propagation of reactive worms in real networks; then compares major distinctions of several key parameters in worms' propagation between models based on mean-field theory and the presented dynamic-time-based one; and deduces the crucial periods of time within a particular 24-hour day for defending against reactive worms' attack. Eventually simulation experiment shows this defense model is feasible and effective.**

***Index Terms*—dynamic time, probability theory, reactive worm, defense model, mean-field theory**

I. Introduction

Based on distributed system and computer network, Peer-to-Peer (P2P) is currently the most popular networking technology for data sharing, instant messaging, enterprise collaboration, etc. However, current P2P networks are facing serious security threats since they show some facilities towards worm attack and propagation. With the emergence of P2P worms, they bring harm to P2P networks and even pose an underlying threat to Internet.

P2P worms can be generally categorized into three groups: passive worms, reactive worms and active worms. Unlike passive worms that hide themselves in malicious files and trick users into downloading and executing them for propagation, active worms that automatically connect to potential targets by using topological information for propagation, reactive worms propagate themselves with legitimate network activities using security vulnerabilities on particular P2P nodes. This type of normal network connections can be initiated by a user or one reactive worm personating a legitimate user. If an infected node finds exploitable security vulnerabilities in a connection, it will pass a worm's copy to the uninfected node. Then once the worm copy is executed by the uninfected node, it will be injected and the newly-infected node will continue infecting its neighborhood nodes. Since reactive worms can propagate through normal connections, they are relatively tough against detection or common firewalls. Characteristics above make the propagation of reactive worms in P2P networks speedy and conceal.

The transmission mode of reactive worms can be generally divided into three categories due to different infection directions: source infection, target infection and mixed infection. Among these three, mixed infection reactive worms infect not only the source host (download port) but also the target host (upload port) in one connection, making it the most harmful one to P2P networks. This paper mainly lays emphasis on this kind of reactive worms.

Most research into reactive worm defense at present is mostly based on the epidemic model. Early in 1926, McKendrick et al. first modeled the spread of biological viruses by the means of mathematics [1], then proposed the epidemic mathematical model. From then on, the epidemic model was widely introduced into the process of modeling on computer viruses and reached some truly remarkable achievements. Some representative ones of them are listed as follows: Kephart et al. first introduced the epidemic model into computer virus modeling process in the early 90's [2]. Yu et al. developed a P2P worm propagation model on the basis of simple epidemic model, and discussed the prevention strategy of P2P worms [3]; Sun et al. studied the propagation process of active P2P worms by a dynamic model [4], took both the entire

Supported by China Postdoctoral Science Foundation funded project (Grant No. 2014M552308), NCET, NSFC (Grant No. 61272400,Grant No. 61309032), Chongqing Innovative Team Fund for College Development Project (Grant No. KJTD201310), Foundation of CEC (Grant No. KJ130525, KJ130523), and CQUPT Research Fund for Young Scholars (Grant No. A2012-79).

© 2014 ACADEMY PUBLISHER
doi:10.4304/jsw.9.11.2778-2788

network's status and each single node's action into consideration; Li et al. presented a stochastic model of worm propagation on the basis of the epidemic model, analyzed the process of state transition of nodes by using the state space of a Markov chain [5]; Yang et al. proposed a new method to integrate the delivery predictability of ProPHET-Routing and verified their proposed OOPProPHET-Routing method was better than Epidemic-Routing method by NS2 network simulator [6]; Xie et al. built a new multi-agents risk assessment model based on attack graph(MRAMBAG) and had shown that the MRAMBAG was a more feasible and effective way for evaluate the network security risk [7]; Xing et al. analyzed the security threat to the virtual network and brought forward a security guarantee embedding algorithm for virtual network [8], the simulation results showed that the algorithm was effective; Wang et al. gave a simulation analysis to reactive worms [9], and simulated the defense process of reactive worms under Internet environment and P2P environment. Also, the keys to defense in these environment had been proposed. Qin et al. and Feng et al. respectively modeled the propagation process and immune process of reactive worms on the basis of Kermack-Mckendrick model [10] [11], and also the prevalence condition of reactive worms was presented in these papers, pointing out the direction for defending against reactive worms in P2P environment. Yang et al. designed a dynamic quarantine protocol to defend active worms in P2P networks by quarantining the suspicious hosts, and he developed a mathematical model of PWPQ to prove the effectiveness of this defense method [12]. Ouyang et al. analyzed the trust mechanism and application model, set up a new kind of trust model based on P2P network [13], the simulation experiment showed this trust model helps improve the success rate of transaction.

The referred works above described the propagation process of worms to some extent and provided some valuable references for establishing the corresponding defense system of worm in P2P networks. But we notice quite few works focus on the defense model of reactive worms especially in modeling reactive worms' propagation with considerations of the dynamic environment such as user behavior, network size, and network bandwidth, etc. In some degree they are basically untouched. This paper attempts to address this issue, and mainly makes the following four contributions.

(1) We present a propagation strategy of reactive worms in dynamic environment, and provide the dynamic process of state transition of nodes when reactive worms spread in accordance with the strategy.

(2) On the basis of analyzing pros and cons of existing defense models of reactive worms, we develop an improved defense model based on mean-field theory and deduce a number of key parameters affecting propagation speed of reactive worms in dynamic environment.

(3) We analyze the shortages of the foregoing model, put forward some improvement methods. That is, we analyze factors including network size and user behavior at different time periods and simulate network size using probability theory, finally propose a defense model of reactive worms based on dynamic time.

(4) We conduct mathematical analysis to study the improved defense model, compare the difference of key parameters that affect reactive worm defense between the defense model that is based on mean-field theory and the one based on dynamic time, and deduce the most crucial period within a day for defending against reactive worm attack.

The rest of this paper is organized as follows. Section 2 presents a propagation strategy of reactive worms in dynamic environment, and elaborates the dynamic process of state transition of nodes when reactive worms spread in accordance with the strategy, section 3 develops an improved defense model of reactive worms in P2P networks based on mean-field theory, section 4 analyzes shortages of the foregoing model and puts forward some improvement methods, section 5 compares the difference of key parameters that affects reactive worm defense between the defense model based on mean-field theory and the one based on dynamic time, deduces the most crucial period of a particular day for defending against reactive worm attack, section 6 proposes the conclusion and some future work directions, the acknowledgment is put forward in section 7.

II. A Propagation Strategy for Reactive Worms in Dynamic Environment

A. State Transition of Nodes When Reactive Worms Spread

A P2P nodes has least six states in different stages of reactive worms' propagation, The summary of these states are listed as follows:

(1) Susceptible infected state (S state): This is the state when an online node is vulnerable to worm attack for its secure vulnerability. Yet it hasn't downloaded the worm file.

(2) Latent state (L state): This is the state when a node in S state has downloaded a worm file from another online worm node but the worm file hasn't been executed. At this stage, the node cannot be invaded by the same type of reactive worm infection. It is not contagious either.

(3) Infected state (I state): Once a worm file is executed by a node in L state, the state of this node changes from latent state to infected state. At this stage, the node is contagious and has already become a worm node.

(4) Quarantined state (Q state): Once a node in I state is detected by monitoring software for transmitting reactive worms, it will be quarantined and its state will be converted into quarantined state. At this stage, the node is no longer contagious.

(5) Immune state (R state): This is the state when an online node has been patched by security software. At this stage, the node cannot be infected by reactive worms anymore, nor is contagious.

(6) Offline state (O state): This is the state when the node has left P2P networks.

State transition of nodes is shown in “Fig. 1,” the description is as follows:

When a benign node containing security vulnerabilities joins P2P networks, it is in susceptible infected state(S); if it has been patched, it is in immune state(R); when an infectious malignant node just joins P2P networks, it is in infected state(I); when a node in I state connects to a node in S state, the infected node would inject a worm file into the uninfected one with a probability φ, when the node in S state has downloaded the worm without execution, the state of this node would be converted into latent state(L); and a node in L state would execute worm files with a probability η, then its state will be converted into infected state(I); when a node in I state is detected by monitoring software with the probability χ for transmitting reactive worms, it will be quarantined and its state would be converted into quarantined state (Q); if an online node in S state, L state, I state, or Q state is found with security vulnerabilities by security software in periodic inspection, it would be patched and its state would be converted to immune state(R) at a probability r_1 , r_2 , r_3 and r_4 respectively; all online nodes would choose to leave P2P networks with a probability α , and if so, their states would then be converted into offline state (O); meanwhile, all offline nodes would choose to join P2P networks with a probability β , and their states would then return to original states before being offline. And users of some offline nodes will reinstall their operation systems with a probability δ , thus their states would be converted into susceptible infected state (S) when being back online.

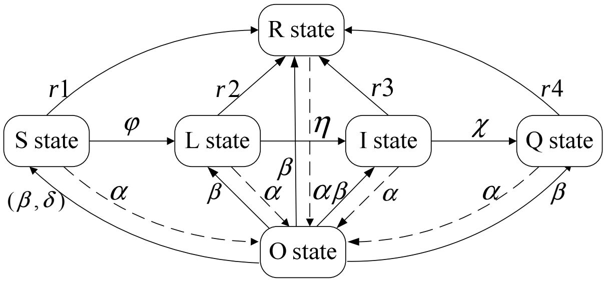

Figure 1. State transitions of nodes

III. A Defense Model Of Reactive Worms Based On Mean-Field Theory

A. Assumptions and Parameters in the Defense Model Based on Mean-field Theory

In order to simplify the modeling process of reactive worms based on mean-field theory, the following assumptions could be made:

(1) The number of nodes in P2P networks keeps constant.

(2) Each node has the same online rate and offline rate in a unit of time whichever its state is. Each offline node will be back to its original state if its operation system hasn’t been reinstalled.

(3) A node in L state can finish downloading all worm fragmentations from other infected nodes in a unit of time.

(4) Only infected node can spread worm fragmentations, which also possibly make it quarantined.

(5) Although all the worm fragmentations in those nodes in Q state have been cleaned up, there are still some vulnerabilities in them and some of them may go back to S state before they are patched.

(6) Reactive worms will launch the attack based on the mixed infection strategy.

Table I lists all variables in the model.

TABLE I.
VARIABLES IN THE DEFENSE MODEL OF REACTIVE WORMS BASED ON MEAN-FIELD THEORY

Variable	Description	Initial value
N	The total number of nodes in P2P network	$N=100000$
λ_d	Downloading rate of a node (The number of files that any P2P node can download in a unit of time)	$\lambda_d=20$
λ_e	Execution rate of a node (The number of files that any P2P node can execute in a unit of time)	$\lambda_e=8$
φ_d	Downloading infection rate (The probability of a node in S state that gets infected by downloading a file)	$\varphi_d=0.3$
φ_u	Uploading infection rate (The probability of a node in S state that gets infected by uploading a file)	$\varphi_u=0.2$
α	Offline rate of an online node	$\alpha=0.01$
β	Online rate of an offline node	$\beta=0.9$
δ	The probability for an offline node that will be back online after reinstalling OS	$\delta=0.05$
η	Execution infection rate (The probability of a node in L state that is infected by executing a download file and will be converted into infected state)	$\eta=0.05$
χ	Detection rate (The probability of a node in I state is that is detected by monitoring software for transmitting reactive worms. Then it will be converted into quarantined state)	$\chi=0.03$
λ	The probability of a node in Q state that will go back to S state after clearing up worm fragmentation	$\lambda=0.3$
r_1	The probability of a node in S state found to contain security vulnerabilities by security software. Then it will be patched and its state will be converted into immune state	$r_1=0.01$
r_2	The probability of a node in L state found to contain security vulnerabilities by security software. Then it will be patched and its state will be converted into immune state	$r_2=0.05$
r_3	The probability of a node in I state found to contain security vulnerabilities by security software. Then it will be patched and its state will be converted into immune state	$r_3=0.08$
r_4	The probability for a node in Q state is found to contain security vulnerabilities by security software, it will be patched and its state will be converted into immune state	$r_4=0.1$
$S_N(t)$	The number of online nodes in susceptible infected state at the time where $S_N(0)$ indicates the total number of nodes in susceptible infected state in P2P networks initially	$S_N(t)=9900$
$S_O(t)$	The number of offline nodes in susceptible infected state at time t	$S_O(t)=0$

$L_N(t)$	The number of online nodes in latent state at time t	$L_N(0)=0$
$L_O(t)$	The number of offline nodes in latent state at time t	$L_O(0)=0$
$I_N(t)$	The number of online nodes in infected state at time t, where $I_N(0)$ indicates the total number of nodes in infected state in P2P networks initially	$I_N(0)=1000$
$I_O(t)$	The number of offline nodes in infected state at time t	$I_O(t)=0$
$Q_N(t)$	The number of online nodes in quarantined state at time t	$Q_N(0)=0$
$Q_O(t)$	The number of offline nodes in quarantined state at time t	$Q_O(0)=0$
$R_N(t)$	The number of online nodes in immune state at time t	$R_N(0)=0$
$R_O(t)$	The number of offline nodes in immune state at time t	$R_O(0)=0$
$A(t)$	The number of additional online nodes whose states have converted from susceptible infected state to latent state at time t	$A(0)=0$
$O(t)$	The number of nodes in offline state at time t	$O(0)=0$

B. A Defense Model of Reactive Worms Based on Mean-Field Theory

The defense model of reactive worms based on mean-field theory should meet the following theorems:

Theorem 1:

$$S_o(t)=\alpha\sum_{i=0}^{t-1}S_N(i)(1-\beta)^{t-i}$$

$$E_o(t)=\alpha\sum_{i=0}^{t-1}E_N(i)(1-\beta)^{t-i}$$

$$I_o(t)=\alpha\sum_{i=0}^{t-1}I_N(i)(1-\beta)^{t-i}$$

$$Q_o(t)=\alpha\sum_{i=0}^{t-1}Q_N(i)(1-\beta)^{t-i}$$

$$R_o(t)=\alpha\sum_{i=0}^{t-1}R_N(i)(1-\beta)^{t-i}$$

Theorem 2:

$$A(t)=S_N(t)\{\varphi_d\lambda_d I_N(t)/(N-O(t)) \\ +\varphi_u[1-(1-1/(N-O(t)))^{\lambda_d I_N(t)}]\}$$

Theorem 3:

$$dS_N(t)/dt=(1-\delta)\beta S_o(t)+\delta\beta O(t)+\lambda Q(t) \\ -A(t)-(\alpha+r_1)S_N(t)$$

Theorem 4:

$$dL_N(t)/dt=\beta(1-\delta)L_o(t)+A(t)-(\alpha+r_2) \\ \bullet L_N(t)-L_N(t)[1-(1-\eta)^{\lambda_e}]$$

Theorem 5:

$$dI_N(t)/dt=\beta(1-\delta)I_o(t)+L_N(t)[1-(1-\eta)^{\lambda_e}] \\ -(\alpha+r_3+\chi)I_N(t)$$

Theorem 6:

$$dQ_N(t)/dt=\beta(1-\delta)Q_o(t)+\chi I_N(t) \\ -(\alpha+r_4+\lambda)Q_N(t)$$

Theorem 7:

$$dR_N(t)/dt=\beta(1-\delta)R_o(t)+r_1S_N(t)+r_2L_N(t) \\ +r_3I_N(t)+r_4Q_N(t)-\alpha R_N(t)$$

Theorem 8:

$$dO(t)/dt=\alpha[S_N(t)+L_N(t)+I_N(t) \\ +Q_N(t)+R_N(t)]-\beta O(t)$$

Due to space limitation, we leave their proof omitted. And we advise interested readers pay attention to the author’s follow-up papers for it.

C. Numerical Simulation and Analysis of Defense Model Based on Mean-field

The defense model of reactive worms based on mean-field theory should meet the following theorems:

Having developed the defense model of reactive worms, the next stage is to conduct simulation experiments by MATLAB. Some important experimental results are listed as follows.

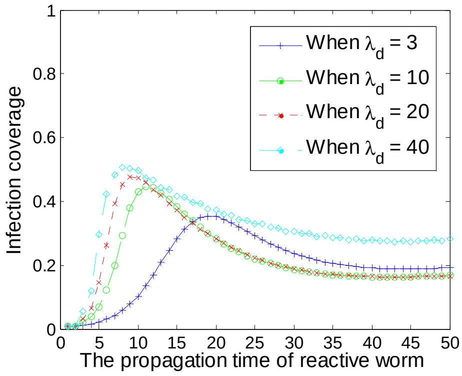

Figure 2. How the downloading rate of a node affects reactive worm propagation.

“Fig. 2” shows the influence of downloading rate of a node on the propagation speed of reactive worms. It can be seen from the figure that the higher the downloading rate of a node is, the faster reactive worms will spread. When the downloading rate of a node exceeds 10, the propagation speed of reactive worms will not be significantly increased.

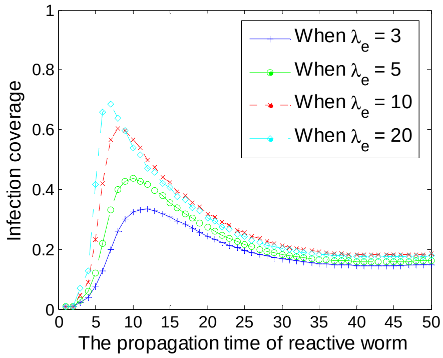

Figure 3. How the execution rate of a node affects reactive worm propagation

“Fig. 3” shows the influence of execution rate of a node on the propagation speed of reactive worms. It can be seen from the figure that the higher the execution rate of a node is, the faster reactive worms will spread. When the execution rate of a node is limited to be less than 3, the propagation speed of reactive worms can be effectively delayed.

“Fig. 4” shows the influence of offline rate of an online node on the propagation speed of reactive worms. It can be seen from the figure that the higher the offline rate of

© 2014 ACADEMY PUBLISHER

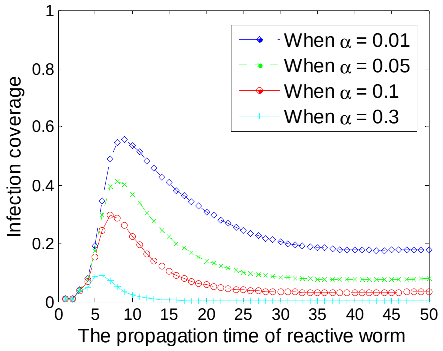

Figure 4. How the offline rate of a node affects reactive worm propagation

an online node is, the slower reactive worms will spread. When the offline rate of a node is greater than 0.1, the propagation speed of reactive worms will be effectively delayed.

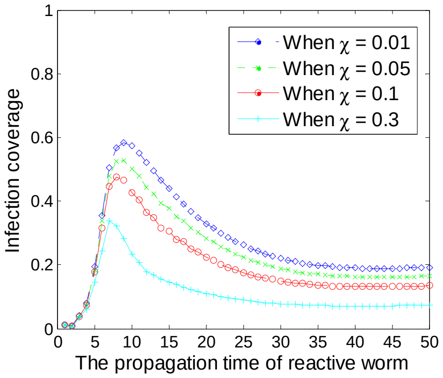

Figure 5. How the detection rate of a node affects reactive worms' propagation

"Fig. 5" shows the influence of detection rate of monitoring software on the propagation speed of reactive worms. It can be seen from the figure that the higher the detection rate of monitoring software is, the slower reactive worms will spread. When the detection rate exceeds 0.3, the propagation speed of reactive worms can be obviously delayed. This suggests that the propagation speed of reactive worms can be effectively delayed by improving detection density of monitoring software.

IV. Researches And Modification To Defense Model Of Reactive Worms Based On Mean-Field Theory

Although defense models of reactive worms based on mean-field and epidemiologic theories can roughly predict the infection ratio, the spread trend and the key points to defense reactive worms, they do not accurately match the defense process of reactive worms in dynamic environment. Because lots of parameters that influence the accuracy of these defense models are estimated under ideal conditions, these estimations are not adequate in reality. To address this issue, this paper presents a defense model based on dynamic time, then makes some improved methods to estimate the key parameters for ensuring the reliability and validity of our model. This section first analyses the deficiency of the foregoing defense model.

A. Deficiency of Defense Models Based on Mean-Field Theory

Such foregoing defense models based on mean-field theory studied the effect of P2P churn on defense effect of reactive worms under the hypothesis that the number of nodes in P2P networks remains basically unchanged within 24 hours, a day. This assumption is obviously unsuited to users' online habits.

Such foregoing defense models based on mean-field theory assume that all the nodes in L state can finish downloading each worm fragmentation within a unit of time. This assumption fails to consider the impact of the fragmentation size, network bandwidth, security awareness of user nodes and the number of seed nodes that probably provide worm fragmentations for user nodes to download during propagation.

Such foregoing defense models based on mean-field theory define the execution infection rate η as a constant. This is obviously inaccurate. Similarly, parameters including λ_e, χ, r_1, r_2, r_3 and r_4 should not be defined as constants either.

B. Improvements to Defense Models Based on Mean-Field Theory

Realistically, the size of online nodes is considerably different within a particular day. Most worms will take long time to reach the maximal infection peak from the beginning of attacks, the propagation of reactive worms is much more influenced by the network scale change, which heavily depends on users' habits during this period. Therefore the change rate of network scale has also been taken into consideration in our improved defense model.

In reality, the bigger and the smaller the worm fragmentation size and the network bandwidth are respectively, the lower the security awareness of user nodes is; the fewer the number of seed nodes providing worm fragmentation is, the longer a node in L state will take to download all the worm fragmentation and the larger the probability that a node in L state is detected by monitoring software is. Hence four parameters *WormSize* (represents the average size of worm fragmentation), *Bandwidth* (represents the average network bandwidth), *SecAw* (represents the security awareness of user nodes) and *SeedNum* (represents the number of seed nodes) are added in our improved defense model. Meanwhile, the probability θ that a node in L state is converted into S state has also been added.

In our improved defense model, those dynamic parameters such as η λ_e χ r_1 r_2 r_3 and r_4 are defined as mathematical functions associated with *SecAw* and ρ (represents infection coverage of reactive worms).

Two given parameters δ and λ have limited effect on the propagation of reactive worms, they are ignored in our improved defense model. In the same way, offline state will not be considered in our improved model.

© 2014 ACADEMY PUBLISHER

V. A Defense Model Of Reactive Worms Based On Dynamic Time

A. Analysis on the Network Size with Consideraton of Users' Online Habits

Time consumed on Internet of the public shows certain regularity in real world. "Fig. 6" shows the distribution within 24 hours a day according to the CNNIC's statistics in the twentieth statistical report of Internet development in China [14].

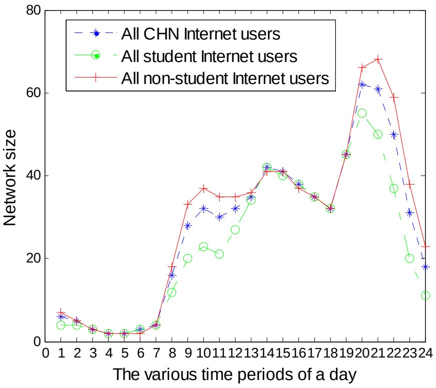

Figure 6. Online user distribution within 24 hours a day

As we see from "Fig. 6", the number of online users stays lowest from 1 a.m. to 7 a.m., and this number will gradually increase after 8 a.m. The trend will continue ascending until it reaches its first local maximum at 10 a.m. with roughly 30 percent of user nodes online. Then the number of online users will slightly drop down at around 11 a.m., while the figure keeps rising from 12 a.m. to 3 p.m. and reach the second local peak of a day with around 40 percent of user nodes online. Then the percentage falls slowly again after 3 p.m. There is a sharp rise in the number of online users around 6 p.m., and the figure will reach its third peak, the global maximum of a day at 9 p.m. with about 60 percent of user nodes online. Later on, the percentage falls rapidly and this trend will continue until 5 a.m. with only 2 percent of user nodes online still, also the minimum of a day. Moreover, the statistical report notes only 13 percent of user nodes have no fixed Internet time, while the remaining 87 percent does and follows the pattern mentioned.

On the basis of the above analytic results, we make the following assumptions about the dynamic changing regularity of the network size within 24 hours in a day.

Since most ordinary users are sleeping from 4 a.m. to 6 a.m., we define these nodes who stay online during these hours as "forever online nodes", and the number of this kind of user nodes remains constant of a day.

User nodes in our improved model are classified into two categories: working nodes and leisure nodes. The time consumed on Internet of working nodes is mainly during the working hours between 9 a.m. to 5 p.m. and the online time of leisure nodes is mainly spent during leisure hours which we assume as 7 p.m. to 12 p.m.

To simplify this model, we assume that all the working nodes or leisure nodes are online at their peak of a day.

As addressed before, there are only 8 working hours in a day and most users surf the Internet using leisure nodes at home. We set the number of leisure nodes equal to be 1.5 times than the amount of working nodes.

Both the number of working online nodes and leisure online nodes of a day are subject to the distributing rules mentioned above.

B. Assumptions and Parameters in the Defense Model Based on Dynamic Time

In order to simplify the modeling process of reactive worms, following assumptions are made:

(1) The number of online nodes in P2P networks is a dynamic variable that varies with time.

(2) P2P nodes are divided into two categories: one is forever online nodes that account for 4% of the total; the other is temporarily online nodes accounting for 96% of the total. As noted above, the temporarily online nodes can be further classified as working online nodes and leisure online nodes. The former accounts for 38.4% of the total, while the latter occupies 57.6%. Besides, the number of online nodes in various states is subject to this proportion.

(3) In various periods of a day, the number of temporarily online nodes is subject to a range of Poisson distribution with different parameters.

(4) The probability θ and r_2 are directly proportional to parameters $WormSize$ and $SecAw$, and inversely proportional to parameters $Bandwidth$ and $SeedNum$. While the probability η follows the opposite law to θ or r_2 with parameters.

Table II lists all variables in our improved defense model.

TABLE II.
Variables In The Defense Model Of Reactive Worms Based On Dynamic Time

Variable	Description	Initial values
$N(t)$	The total number of nodes in P2P networks at time t	$N(1) = 16417$
N_d	The total number of forever online nodes in P2P networks	$N_d = 10417$
$N_t(t)$	The total number of temporarily online nodes in P2P networks at time t	$N_t(1) = 6000$
$N_{tw}(t)$	The total number of working online nodes at time t	$N_{tw}(1) = 0$
$N_{tr}(t)$	The total number of leisure online nodes at time t	$N_{tr}(1) = 6000$
P_{tw}	The maximum peak of working online nodes of a day	$P_{tw} = 100000$
P_{tr}	The maximum peak of leisure online nodes of a day	$P_{tr} = 150000$
$WormSize$	The average size of worm fragmentation	$WormSize = 6MB$
$SeedNu$ $-m(t)$	The number of seed nodes at time t	$SeedNum(1) = 8.2$
$\rho(t)$	The infection coverage of reactive worms at time t, and $\rho(t) = I(t) / N(t)$	$\rho(1) = 0.01$
$SecAw(t)$	The security awareness function of user nodes at time t, and $SecAw(t) = 0.3 + 1.1 \times \rho(t)$	$SecAw(1) = 0.311$

© 2014 ACADEMY PUBLISHER

$\theta(t)$	The probability for a node in L state to be converted into S state at time t	$\theta(1)=0.0724$
$\eta(t)$	The probability for a node in L state infected by executing a download file at time t	$\eta(1)=0.051$
λ_d	Downloading rate of a node (The number of files that any P2P node can download in a unit of time)	$\lambda_d=20$
$\lambda_e(t)$	Execution rate of a node at time t (The number of files that this P2P node can execute in t unit of time)	$\lambda_e(1)=20$
φ_d	Downloading infection rate (The probability of a node in S state infected by downloading a file)	$\varphi_d=0.3$
φ_u	Uploading infection rate (The probability of a node in S state infected by uploading a file)	$\varphi_u=0.2$
$\chi(t)$	The probability of a node in I state that is converted into Q state at time t, where $\chi(t)=0.1\times SecAw(t)$	$\chi(1)=0.0311$
$r_1(t)$	The probability of a node in S state found to contain security vulnerabilities by security software, then its state will be converted into immune state at time t. $r_1(t)=0.08\times SecAw(t)$	$r_1(1)=0.0249$
$r_2(t)$	The probability of a node in L state found to contain security vulnerabilities by security software, then its state will be converted into immune state at time t. $r_2(t)=0.08\times SecAw(t)\lg[SeedNum(t)$ $+Bandwidth/WormSize]$	$r_2(1)=0.0241$
$r_3(t)$	The probability of a node in I state found to contain security vulnerabilities by security software, then its state will be converted into immune state at time t. $r_3(t)=0.15\times SecAw(t)$	$r_3(1)=0.0466$
$r_4(t)$	The probability for a node in Q state found to contain security vulnerabilities by security software, then it will be patched and its state will be converted into immune state at time t. $r_4(t)=0.2\times SecAw(t)$	$r_4(1)=0.0622$
$S(t)$	The number of online nodes in S state at time t, where $S(1)$ indicates the total number of nodes in S state in P2P networks initially	$S(1)=16253$
$L(t)$	The number of online nodes in L state at time t	$L(1)=0$
$I(t)$	The number of online nodes in I state at time t, where $\mathrm{I}(1)$ indicates the total number of nodes in I state in P2P networks initially	$I(1)=164$
$Q(t)$	The number of online nodes in Q state at time t	$Q(1)=0$
$R(t)$	The number of online nodes in R state at time t	$R(1)=0$

C. A Defense Model of Reactive Worms Based on Dynamic Time

Having proposed model assumptions and parameter elucidations, our next stage is to develop the improved defense model based on dynamic time. The modeling methodology based on dynamic time is similar to the one based on mean-field theory. The key of the improved modeling methodology is how to estimate the number of online nodes in various states at different times within a day. This section will focus on resolving this problem.

In this model, variable t represents different hours of a day, and the number of online nodes at different times of a day can be discussed the way as follows:

(1) The discussion on the number of working online nodes at different times of a day.

When $t=1-6$, sleeping time for ordinary users, the number of working online nodes at this period is set as 0. That is $N_{tw}(t)=0\ (t\in[1,2,..6])$ (1)

When $t=7-15$, the number of working online nodes gradually increases because most users begin working and major stock exchanges throughout the world open one after another. The number of working online nodes keeps growing within this period, and this figure will reach the peak at 3 p.m. The changing trend of working online nodes during this period obeys the Poisson distribution of parameter 2.6. That is

$$N_{tw}(t)=P_{tw}\bullet \sum_{k=0}^{t-7}(\lambda^k/k!)\bullet e^{-\lambda}\ (\lambda=2.6,t\in[7,8,...,15]) \tag{2}$$

When $t=16-24$, almost the end of daily working hours, the number of working online nodes continue to retreat from its peak at 3 p.m. until it finally reduces to be zero. Given individual users might do extra work at night, the descending trend of the number of working online nodes will continue to 12 p.m. The changing trend of working online nodes during this period also conforms to the Poisson distribution of parameter 3.3. That is

$$N_{tw}(t)=P_{tw}(1-\sum_{k=0}^{t-16}(\lambda^k/k!)\bullet e^{-\lambda}) \quad (\lambda=3.3,t\in[16,17,...,24]) \tag{3}$$

When $t>15$, $P_{tw}\sum_{k=0}^{t-7}(\lambda^k/k!)\bullet e^{-\lambda}=1$. Combining the three Equations (1) (2) (3), the change of working online nodes within 24 hours a day can be calculated as follows:

$$N_{tw}(t)=P_{tw}(\sum_{k=0}^{t-7}(2.6^k/k!)\bullet e^{-2.6}-\sum_{k=0}^{t-7}(3.3^k/k!)\bullet e^{-3.3}), \quad \text{And } 1\le t\le 24 \tag{4}$$

(2) The discussion on the number of leisure online nodes at different times of a day.

When $t=1-3$, most ordinary users are sleeping, while only a few users are playing online games or watching online videos. Therefore the number of this part of users is on the decline, the change of leisure online nodes during this period conforms to the Poisson distribution of parameter 1.3. That is

$$N_{tr}(t)=P_{tr}(1-\sum_{k=0}^{t+2}(\lambda^k/k!)\bullet e^{-\lambda})\ (\lambda=1.3,t\in[1,2,3]) \tag{5}$$

When $t=4-18$,most users are either resting or working, few of leisure online nodes are being used at this period, so the number of leisure online nodes at this period is set as 0. That is $N_{tr}(t)=0\ (t\in[4,5,...,18])$ (6)

When $t=19-22$, the leisure online nodes are being used in large amounts, so the number of leisure online nodes will rapidly climb to the maximum peak of a day. The change of leisure online nodes during this period also conforms to the Poisson distribution of parameter 0.5. That is

$$N_{tr}(t) = P_{tr}\sum\nolimits_{k=0}^{t-19}(\lambda^k / k!)\bullet e^{-\lambda}, \quad \lambda = 0.5, \quad t \in [19, 20, 21, 22] \tag{7}$$

When $t = 23-24$, it's again time for bed, causing the number of leisure online nodes rapidly decreases. The change of leisure online nodes during this period obeys the Poisson distribution of parameter 1.3. That is

$$N_{tr}(t) = P_{tr}(1-\sum\nolimits_{k=0}^{t-23}(\lambda^k / k!)\bullet e^{-\lambda}) \; (\lambda = 1.3, t \in [23, 24]) \tag{8}$$

When $t > 3$, the equation $1-\sum\nolimits_{k=0}^{t+2}(1.3^k / k!)\bullet e^{-1.3} = 0$ is valid, no matter what value the variable t is. And when $t > 22$, the equation $\sum\nolimits_{k=0}^{t-19}(0.5^k / k!)\bullet e^{-0.5} = 1$ is valid, no matter what value the variable t is. Combining the four Equations (5) (6) (7) (8), the change of leisure online nodes within 24 hours a day can be calculated as following equation (9).

$$N_{tr}(t) = P_{tr}(1-\sum_{k=0}^{t+2}\frac{1.3^k}{k!}e^{-1.3} + \sum_{k=0}^{t-19}\frac{0.5^k}{k!}e^{-0.5} - \sum_{k=0}^{t-23}\frac{1.3^k}{k!}e^{-1.3})$$

And $1 \le t \le 24$. (9)

(3) In conclusion, the change of all the online nodes within 24 hours a day can be calculated as following Equation (10).

$$\begin{aligned}N(t) &= N_d + N_t(t) = N_d + N_{tw}(t) + N_{tr}(t)\\ &= N_d + P_{tw}(\sum_{k=0}^{t-7}\frac{2.6^k}{k!}e^{-2.6} - \sum_{k=0}^{t-7}\frac{3.3^k}{k!}e^{-3.3}) + P_{tr}(1\\ &\quad -\sum_{k=0}^{t+2}\frac{1.3^k}{k!}e^{-1.3} + \sum_{k=0}^{t-19}\frac{0.5^k}{k!}e^{-0.5} - \sum_{k=0}^{t-23}\frac{1.3^k}{k!}e^{-1.3})\end{aligned}$$

And $1 \le t \le 24$ (10)

In this improved defense model, the change of nodes in S state in a unit of time is composed of four parts. The first part will be converted into L state for downloading worm fragmentation from infected nodes; the second part will be converted into L state for uploading some resource to infected nodes; the third part is converted from nodes in L state because these latent nodes are found to contain worm fragmentation by security software before their states are converted into S with worm fragmentation being removed; the fourth part will be converted into R state because these nodes in S state are found to contain security vulnerabilities by security software, they will be patched and their states will be converted into immune state. Given the above, the change rate of nodes in S state satisfies the following Equation (11).

$$\begin{aligned}dS(t)/dt = \theta(t)L(t) - [&\varphi_d\lambda_d I(t)/N(t) +\\ &\varphi_u\{1-[1-1/N(t)]^{\lambda_d I(t)}\} + r_1(t)]S(t)\end{aligned} \tag{11}$$

The same theory proves that the change rate of nodes in L state satisfies the following Equation (12).

$$\begin{aligned}dL(t)/dt = [\varphi_d\lambda_d I(t)/N(t) + \varphi_u\{1-[1-1/N(t)]^{\lambda_d I(t)}\}]\\ S(t) - \{\theta(t) + r_2(t) + \{1-[1-\eta(t)]^{\lambda_e(t)}\}\}L(t)\end{aligned} \tag{12}$$

The change rate of nodes in I state satisfies the following Equation (13).

$$dI(t)/dt = \{1-[1-\eta(t)]^{\lambda_e(t)}\}L(t) - [r_3(t) + \chi(t)]I(t) \tag{13}$$

The change rate of nodes in Q state satisfies the following Equation (14).

$$dQ(t)/dt = \chi(t)I(t) - r_4(t)Q(t) \tag{14}$$

And the change rate of nodes in R state satisfies the following Equation (15).

$$dR(t)/dt = r_1(t)S(t) + r_2(t)L(t) + r_3(t)I(t) + r_4(t)Q(t) \tag{15}$$

For the sake of brevity, we leave their proof omitted.

D. Numerical Simulation and Analysis of Defense Model Based on Dynamic Time

There are three steps to count the number of infected nodes:

The first step is to initialize the number of online nodes in all states during a first time period, the second step is to calculate the change in numbers of online nodes in all states during the same period of time according to the formula (11-15), the third step is to reckon the actual number of online nodes during a second time period according to the formula (10), the fourth step is to initialize the number of online nodes in all states during a second time period according to the proportion of different states of online nodes that has been calculated in second step and the actual number of online nodes that has been reckon in third step, then the change in numbers of online nodes in all states during the second period of time can be calculated in the fifth step, The rest can be done in the same manner, until the number change of online nodes in all states for 24 hours within a day has been calculated.

Key parameters affecting reactive worm defense in real environment can be deduced by adjusting the parameters in our improved defense model.

"Fig. 7" shows the influence of downloading rate of a node on the defense of reactive worms in this improved model. It can be seen from the figure that the higher the downloading rate of a node, the larger the number of online nodes infected by reactive worms will be. This

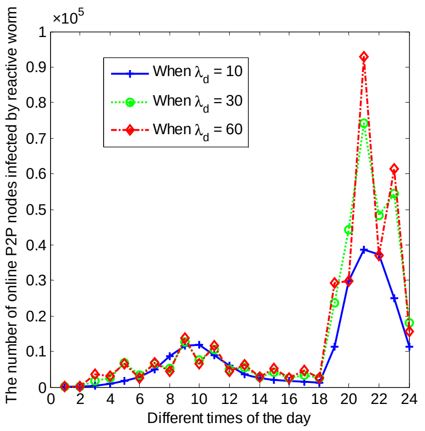

Figure 7. How the downloading rate of a node affects reactive worm propagation in improved model

parameter has great effects on the defense of reactive worms, if large amount of reactive worms have been found in P2P networks, the propagation of reactive worms can be effectively controlled by restricting the downloading rate of each node.

“Fig. 8” shows the influence of downloading infection rate of a node on the defense of reactive worms in this improved model. It can be seen from the figure that the higher the downloading infection rate of a node stays, the larger the number of online nodes infected by reactive worms will be.

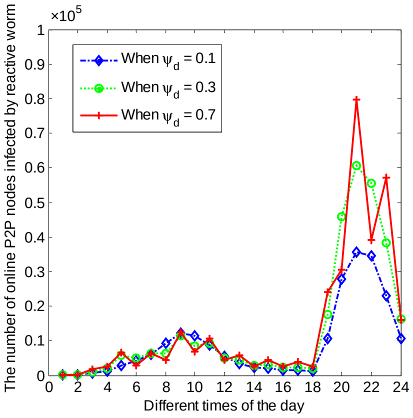

Figure 8. How the downloading infection rate of a node affects reactive worm propagation in improved model

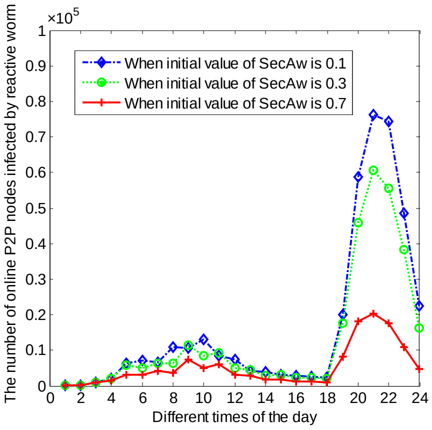

Figure 9. How the initial value of SecAw affects reactive worm propagation in improved model

“Fig. 9” shows the influence of the initial value of the security awareness function of user nodes on the defense of reactive worms in this improved model. In this improved model, those parameters such as $\theta(t)$, $\lambda_e(t)$, $\chi(t)$, $r_1(t)$, $r_2(t)$, $r_3(t)$, $r_4(t)$ are all related to the security awareness function of user nodes. It can be seen from the figure that the higher the security awareness of a user node is, the fewer the number of online nodes infected by reactive worms will be and also the better the defense effect of reactive worm can be obtained. Therefore cyber-safety education should be expanded to all ranges of Internet users for raising their knowledge level and safe consciousness, which effectively help defense against reactive worms in P2P networks.

“Fig. 10” shows the influence of the initial value of immunity system on the defense of reactive worms in this improved model. It can be seen from the figure that the higher the initial value of immunity system is, the fewer the number of online nodes infected by reactive worm will be and the better the defense effect of reactive worm can be.

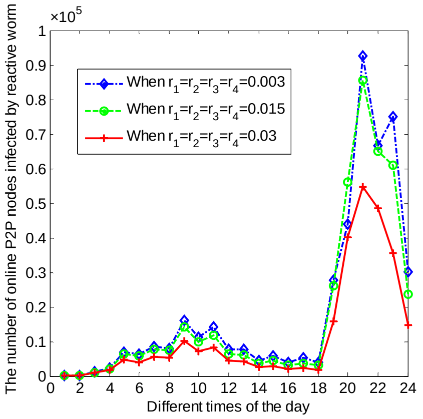

Figure 10. How the initial value of immunity system affects reactive worm propagation in improved model

In conclusion, user's online habits give it a rise to the most significant impact on the worm attack according to these simulation results. As you can see from these figures, only few of online P2P nodes are infected by reactive worms between 1 a.m. and 4 a.m. because most users are sleeping; the number of online P2P nodes infected by reactive worms begin ascending between 4 a.m. and 11 a.m. because most users go to work during this period; however the number of online infected nodes remains very limited even during peak hours. The major reason for this is that most leisure nodes that occupy the mainstream of P2P networks are not at working hours during this period, meanwhile only few working nodes join P2P networks, seriously restricting the developing speed of reactive worms; the number of online nodes in P2P networks further reduces form 11 a.m. to 6 p.m. because major stock exchanges close and most users go home from working; and the number of P2P nodes infected by reactive worms will fall from the previous peak to relatively lower level; the peak of infection of a day occurs during 6 p.m. and 12 p.m.; the maximum of the number of infected nodes appears about 8 p.m. because most leisure nodes have joined to P2P networks by that time; the peak hour of surfing on the internet also appears at 8 p.m., offering an opportunity for reactive worms’ sudden spread, meanwhile the number of online P2P nodes infected by reactive worms is booming; the number of infected P2P nodes will decrease rapidly within bed time after 10 p.m. Obviously the most crucial time of defending reactive worms is from 6 p.m. to 10 p.m. In order to effectively guard against reactive worms’ attack in P2P networks and ensure the availability of normal operations of P2P networks, we should increase strength on key nodes’ supervision, speed up the frequency of scanning vulnerability during particular periods within a day.

© 2014 ACADEMY PUBLISHER

VI. Conclusion And Future Work

In the paper, firstly we proposed a propagation strategy of reactive worms in dynamic environment and provided the process of nodes' state transition when reactive worms spread in accordance with the strategy proposed; second we developed a defense model based on mean-field theory; third, we analyzed shortages of the foregoing model and proposed an improved defense model of reactive worms based on dynamic time; finally we compared the difference among key parameters that affect reactive worm defense between the model based on mean-field theory and the one based on dynamic time and deduced the most important period of a day for defending against reactive worm attack.

Future work will involve how to improve the detection rate of monitoring software according to the characteristics of reactive worms, how to improve the accuracy of the defense model of reactive worms by considering the trust relationship and social nature between P2P nodes upon the propagation of reactive worms and how to build an efficient defense system to prevent reactive worms based on these works.

Acknowledgment

This work is supported by China Postdoctoral Science Foundation funded project (Grant No. 2014M552308), NCET, NSFC (Grant No. 61272400,Grant No. 61309032), Chongqing Innovative Team Fund for College Development Project (Grant No. KJTD201310), Foundation of CEC (Grant No. KJ130525, KJ130523), and CQUPT Research Fund for Young Scholars (Grant No. A2012-79).

References

[1] A. G. McKendrick. "Applications of mathematics to medical problems," Proc. the 44th Edinburgh Mathematica Society. Edinburgh, pp. 98-130, 1926.

[2] J. O. Kephart, S. R. White. "Directed-graph epidemiological models of computer viruses". Proc. the IEEE Symp on Security and Privacy. Piscataway, NJ,pp. 343-359, 1991.

[3] W. Yu, "Analyze the Worm-Based Attack in Large Scale P2P Networks", Proc. the 8th IEEE International Symposium on High Assurance Systems Engineering (HASE'04), pp.308-309, 2004.

[4] Q. D. Sun, Q. Wang, J. Ren, "Modeling and Analysis of the Proactive Worm in Unstructured Peer-to-Peer Network", Journal of Convergence Information Technology, vol. 5, No. 5, pp.111-117, 2010.

[5] Q. R. Li, W. Han，"An Analysis for Stochastic Model of Worm Propagation"，Journal of IJACT, Korea, vol. 4, No. 3, pp. 156-164, 2012.

[6] S. Y. Yang, J. T. Jiang, P. Z. Chen, "OOPProPHET: A New Routing Method to Integrate the Delivery Predictability of ProPHETRouting with OOP-Routing in Delay Tolerant Networks, " Journal of Computers, vol.8, no.7, pp. 1656-1663, July 2013.

[7] L. X. Xie, X. Zhang, J. Y. Zhang, "Network Security Risk Assessment Based on Attack Graph, " Journal of Computers, vol.8, no.9, pp. 2339-2347, September 2013.

[8] C. Q. Xing, J. L. Lan, Y. X. Hu, "Virtual Network with Security Guarantee Embedding Algorithms," Journal of Computers, vol.8, no.11,pp. 2782-2788, November 2013.

[9] Y. W. Wang, J. W. Jing, J. Xiang, "Contagion worm propagation simulation and analysis", Journal of Computer Research and Development, China, vol. 45, pp.207-216, 2008.

[10] Z. G. Qin, C. S. Feng, F. L. Zhang et al. "Modeling propagation of reactive worm in P2P networks". Proc. Communications, Circuits and Systems 2009(ICCCAS 2009), Milpitas, CA, 2009, 335-340.

[11] C. S. Feng, Z. G. Qin, L. Cuthbert.. "Reactive Worms Propagation Modeling and Analysis in Peer-to-Peer Networks", Journal of Computer Research and Development, China, vol. 47, pp.500-507, 2010.

[12] W. Yang, M. G. R. Chang, Y. Yao, and X. M. Shen, "Stability Analysis of P2P Worm Propagation Model with Dynamic Quarantine Defense," Journal of Networks. Finland, vol. 6, pp. 153–162, January 2011.

[13] G. Ouyang, X. Chen. "Trust Model Based on P2P Network," Journal of Networks. Finland, vol. 8, pp. 2013–2020, September 2013.

[14] CNNIC, The twentieth statistical report on Internet development in China [EB/OL]. http://www.cnnic.cn/gywm/xwzx/rdxw/2007nrd/201207/t20120710_31532.htm, July 18, 2007.

Haokun Tang received B.S. degree in computer application from Southwestern Normal University, China, in 1999. M.S. degree in computer application from Southwestern Normal University, China, in 2003. And Ph.D. in Computer Systems Organization from the University of Electronic Science and Technology of China. China, in 2013. Now he is a researcher at postdoctoral workstation, HanHua Financial Holding. He has published more than 16 refereed journal/conference papers. His current research interests are in network security, P2P applications, cloud computing.

Shitong Zhu is currently a full-time sophomore student at School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, China. His current research interests include Device-to-Device communications, LTE-A etc.

Jun Huang received B.S. degree in computer science from Hubei University of Automotive Technology, China, in 2005. M.S. degree (with honor) in computer science from Chongqing University of Posts and Telecommunications, China, in 2009. And Ph.D. degree (with honor) from Institute of Network Technology, Beijing University of Posts and Telecommunications, China, in 2012. Now he is an Associate Professor at School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications. He was a visiting researcher at Global Information and Telecommunication Institute, Waseda University, Tokyo, from Sept. 2010 to Sept.

2011. He is a member of IEEE and IEICE. A best paper award winner of AsiaFI 2011. He has published more than 20 refereed journal/conference papers. His current research interests include network optimization, future Internet, and Cloud computing etc.

Hong Liu currently studies for her Ph.D. in Computer Engineering from the University of Chongqing. She is an associate professor in Chongqing University of Posts and Telecommunications in the Department of Software Engineering, China. She published several papers in the area of wireless networks and image security. Her research interests include sensor networks, image processing and image security.

© 2014 ACADEMY PUBLISHER