Propagation of Active Worms in P2P Networks: Modeling and Analysis

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

[1]Yukui Lu, [2]Shitong Zhu and [3]Jun Huang
[1]Yiling Pharmaceutical Co, LTD, Shijiazhuang, P.R.China
[2,3]School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, P.R.China
Email: [1]xiaolu912@163.com; [2]zstchina1993@gmail.com; [3]xiaoniuadmin@gmail.com

***Abstract*—Active worms, a category of self-replicating malicious programs which could spread in an automated fashion and flood particular Peer-to-Peer (P2P) networks within very short time, have drawn significant attention. However, only limited number of studies focus on propagation model of active worms with fair consideration of P2P nodes' dynamic features consisting of P2P churn, random quarantine, regular immunization, dynamic fragmentation and etc. This paper proposes three propagation models of active worms under dynamic P2P environment, conducts a mathematical analysis on the propagation of active worms under presented models and provides extensive numerical studies to the impact of relevant parameters on active worms' propagation speed under dynamic P2P environment. Models presented in this paper are simple, effective and thus applicable for defending against active worms in real P2P networks.**

***Index Terms*—active worm, dynamic feature, propagation model, P2P networks**

I. Introduction

As a future technology of Wireless Broadband Internet, the transmission mode of P2P networks has turned into a hot spot. Fortune Magazine acclaims P2P network one of the four technologies that will shape the Internet's future. Nowadays, P2P networks account for 37% of overall Internet traffic, and this number is even as high as 90 in multimedia data sharing services. P2P networks provide great convenience for resource sharing systems and fast routing mechanism, yet they also give it a rise to Internet worms' fast spread and massive invasion.

Based on the Computer Network Emergency Response Technical Team Coordination Center's statistics, the number of cyber security incidents has grown exponentially at a rate of 50% every year since 1988. Among them, malicious code over Internet always ranks the first place due to their fast diffusion, wide range of victims and strong penetrability. Currently active worms make the greatest potential pitfall for their self-propagation with no human invention.

Research into propagation model of active worms enables insights of worm behavior and features, helps detect and defense active worms. A great deal of research has been done on active worms' propagation model and defensive measure within recent years. For instance, Chen *et al*. developed an active worm propagation model on the basis of discrete time [1]; Yu *et al*. analyzed the propagation strategy and propagation process of P2P worms based on simple epidemic model applied to static network topology [2]; they also compared P2P worm propagation performance in four different attack strategies, indicating P2P worms based on hit-list scanning strategy best attack-efficient [3]; A method of building secure P2P networks using benign worms against malicious worms was introduced by Jia *et al*. [4]; Wang *et al*. presented a propagation model of active P2P worms under Chord networks [5]; Moreover, Ravikumar *et al*. modeled the spread of malware in decentralized, Gnutella type of peer-to-peer networks [6]; And propagation procedure of active worm in P2P networks based on topology scanning strategy using of logic matrix was addressed by Fan [7] [8]; Additionally Zhang *et al*. completed a static model on active worms within unstructured P2P system [9]; Besides a dynamic quarantine protocol towards defending active worms in P2P networks was designed by Yang *et al*. quarantining the suspicious host, and they developed a corresponding mathematical model of PWPQ to prove the effectiveness of this method [10]; Chen *et al*. brought forward a repair-and-patch approach to quarantine malicious worms quickly in unstructured P2P networks [11]; In addition, Feng *et al*. [12] and Li *et al*. [13] respectively addressed two propagation models of active worms with the reference to degree difference of nodes under unstructured P2P networks; Suto *et al*. proposed a method constructing network matching bimodal-degree distribution [14], thus being naturally robust against both attack and failure, and they obtained simulation results proving it eligible for higher resilience; A membership function of trusted set was established by Zhou *et al*.

Manuscript received August 6, 2013; revised March 15, 2014; accepted June 10, 2014.

doi:10.4304/jcp.9.11.2514-2524

according to the trusted level in the P2P trust model and maximum membership principle [15], the simulation results showed that the method was strong applicability in the event that the granularity of trusted level was great; Guo *et al*. proposed a peer classification method based on fuzzy clustering [16], and proved their method could effectively avoid false recommendation, and enhanced the accuracy of trust evaluation in P2P networks; Meng *et al*. came up with a hierarchical clustering P2P network model based on user interest in [17], and the simulation results showed this method could form cluster more rapidly and gain the appropriate resources faster than traditional algorithm.

The above describes the propagation process of active worms to some extent, provides valuable reference material for establishing corresponding defense system of active worms. Whereas existing propagation models more or less ignore some major or minor behavior characteristics of P2P nodes within P2P networks, lack consideration of dynamic characteristics' effects on P2P nodes in which active worms propagate. Hence they have their limitations. This paper attempts to address the issue, makes the following three major contributions.

- We study three common attack strategies of active worms in P2P networks and provide states transition process of nodes when active worms spread in accordance with these strategies.
- We present three propagation models on the basis of above attack strategies and also describe the propagation process of active worms considering the dynamic features of P2P nodes comprising of random stirring, dynamic quarantine, regular immunization, data partition transmission, retarded growth of worms and execution in sequential order of downloaded files.
- We conduct mathematical analysis to study models proposed, deduce a number of key parameters affecting propagation speed of active worms in P2P networks, making it applicable for defending against active worms in real P2P networks.

The rest of this paper is organized as follows. Section 2 studies three common attack strategies of active worms in P2P networks and elaborates states transition process of nodes when active worms spread in accordance with these strategies; section 3 presents three propagation models of active worms in P2P networks based on different strategies and describes the propagation process of active worms with consideration of dynamic features of P2P nodes; section 4 analyzes the defined propagation models by mathematical analysis and deduces a number of key parameters affecting propagation speed of active worms in P2P network; section 5 proposes conclusions and future work; eventually the acknowledgment is put forward in section 6.

II. Three Attack Strategies for Active Worms in P2P Networks

In P2P networks, there are three major attack strategies of active worms, they are listed as below.

A. Random Scanning Attack Strategy

Based on this attack strategy, an infected node will pick one address randomly each time and attempt to attack without collecting routing information of other nodes in advance. If the selected IP address has been assigned to a P2P node with security vulnerabilities, the P2P node will be infected with a probability φ. Once the P2P node completes downloading all the worm file fragments, it will turn to be a new worm-spreading node and will infect other uninfected nodes in the same way. Otherwise this attack strategy will fail.

B. Hit-list Scanning Attack Strategy

Applied this attack strategy, a worm node collects routing information of other online P2P nodes in advance, then automatically creates a scan list (Hit-list) and attacks destination nodes based on the Hit-list. When any healthy nodes get infected by an infected node, the previously infected node will upload non-scanned portion of the Hit-list to them. Successively these newly infected nodes continue attacking uninfected nodes on list in the same pattern, until each node in Hit-list has been scanned.

C. Topological Scanning Attack Strategy

Once active worms are released into P2P networks using this attack strategy, they will scan all their neighbor nodes in accordance with network topology to find uninfected P2P nodes with secure vulnerability and employ attacks. If they have completed, newly infected nodes will search for their next targets in the same pattern.

D. States Transition Process of Nodes When Active Worms Spread

P2P nodes have some particular characteristics in dynamic environment, therefore indicate different states in different phases of active worms' propagation, The summary of these states is listed as follows:

- Infection-susceptible state (state S): This is the state in which one online node is vulnerable to worm attacks due to its security vulnerability, however worm file has not been downloaded.
- Latent state (state L): This is the state in which a previous S state node has completed downloading worm file from another online worm node, but the file has not been executed yet. At this phase, the node carrying worm file cannot be infected by the same type of active worms; nor is it contagious.
- Infected state (state I): Once a worm file has been executed by a L-state node, it transforms into I-state. Now, the node is contagious and a worm node.
- Quarantined state (state Q): Once an I-state node has been detected by monitoring software while transmitting active worms, it will be quarantined, and converted into the quarantined state. At this stage, the node is no longer contagious.
- Immune state (state R): This is the state when an online node has been patched by security software. At this stage, the node is immune to active worms

and contagious.

- Offline state (state O): This is the state where the node has left P2P networks.

State transition of nodes is shown in Figure 1, the description is as follows:

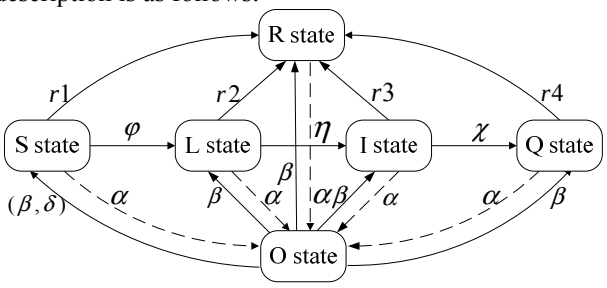

Figure 1. States transition process of nodes

When a benign node containing security vulnerability joins P2P network, it is in infection-susceptible state (S); if it has been patched, it is in immune state (R). When a malevolently infectious node joins P2P network, it is in infected state (I); when a node in I -state scans another node in S -state, the infected node will inject a worm file into the uninfected node with a probability φ ; after the S -state node completes downloading worm file without execution, the state of this node changes into the latent state (L); A L -state node executes worm files with a probability of η , then its state will be converted into infected state (I); when an I -state node is detected by monitoring software with a probability χ while transmitting active worms, it will be quarantined and successively its state will be converted into quarantined state (Q); if an online node been found alternatively in S -state, L -state, I -state, or Q -state with security vulnerability in periodic inspection, security software will patch it converting its state into immune state(R) with probabilities r_1 , r_2 , r_3 , and r_4 respectively; all online nodes have a probability α to choose leaving P2P network, and their states would be converted into offline state(O); meanwhile, all offline nodes have a probability β to choose joining P2P network, and then their states will return to their original states before being offline ; And some offline nodes will reinstall their operation systems with a probability δ , so when they resume to be online and their states will be convert into infection-susceptible state (S).

III. Three Propagation Models of Active Worms in P2P Networks

A. Parameters and Hypotheses

We develop three propagation models of active worms in P2P networks based on attack strategies mentioned before. In order to simplify modeling process, we suppose worm nodes based on Hit-list scanning strategy have collected all the routing information of online P2P nodes in advance; no matter which attack strategy worm nodes apply, they will not attack infected nodes for a second time; addtionally, a worm node can finish injecting all of worm file fragments into an uninfected node within a unit duration; the number of P2P nodes is set to be 10000. Table I. lists all variables in models.

TABLE
VARIABLES IN MODELS

Variable	Definition
T	Total number of nodes in networks
λ	Scanning rate of a worm node (the number of nodes that can be simultaneously scanned by one infected node)
φ	The probability that a node in state S (already scanned by a worm node) will be infected.
α	Offline rate of an online node
β	Online rate of an offline node
δ	The probability that an offline node resume being online after reinstalling OS
a	Average bandwidth of P2P networks
W	Size of an integrated worm file
η	Download rate of a worm node(number of worm files that can be downloaded by a L -state node within a unit of time, $\eta = a/W$)
χ	Detection rate (the probability that a node in state I is detected by monitoring software when transmitting worm file fragments, and then its state will be turned into quarantined state)
γ	The ratio of valid addresses to total $IPv4$ address space in P2P networks, $\gamma < 24\%$ [17]
r_1	The probability that a node in state S is found containing security vulnerability by security software, then it will be patched and its state will be transited into immune state
r_2	The probability that a node in state L is found containing security vulnerability by security software, then it will be patched and its state will be transited into immune state
r_3	The probability that a node in state I is found containing security vulnerability by security software, then it will be patched and its state will be transited into immune state
r_4	The probability that a node in state Q is found containing security vulnerability by security software, it will be patched and its state will be transited into immune state
ϖ	Mean degree value of a node in unstructured P2P networks
θ	Degree value of a node in structured P2P networks
C_1	A constant corresponding to ϖ
ε	The extent to which topologies of P2P networks meet power law model, $\varepsilon \in [1,8]$
k	Degree of any node in P2P networks
$\Theta M(t)$	Average probability of a node in state S connecting to a I -state node at time t
$S_N(t)$	Number of online nodes in infection-susceptible state at time t , where $S_N(0)$ indicates the total number of nodes in infection-susceptible state in P2P networks initially
$S_O(t)$	The number of offline nodes that were transited from the nodes in state S at time t
$S(t)$	Total number of nodes in infection-susceptible state at time t , $S(t) = S_N(t) + S_o(t)$

$L_N(t)$	Number of online nodes in latent state at time t
$L_O(t)$	The number of offline nodes transited from state L at time t
$I_N(t)$	The number of online nodes in infected state at time t, where $I_N(0)$ indicates the total number of nodes in infected state under P2P networks initially
$I_O(t)$	Number of offline nodes transited from state I at time t
$Q_N(t)$	Number of online nodes in quarantined state at time t
$Q_O(t)$	Number of offline nodes transited from state Q at time t
$R_N(t)$	Number of online nodes in immune state at time t
$R_O(t)$	Number of offline nodes transited from state R at time t
$A_N(t)$	Number of additional online nodes whose states have been transited from S to L at time t
$O(t)$	Total number of nodes in offline state at time t
$X_N^{(k)}(t)$	The number of nodes with k degree in various states at time t, the state which these nodes belonged to is determined by values of X, $X \in (S,L,I,Q,R)$
$A^{(k)}(t)$	The number of additional online nodes with k degree whose states have been transited from S to L at time t

B. Propagation Model Based on Random Scanning Strategy (PRS Model)

Propagation Model Based on Random Scanning Strategy (PRS Model):

Theorem 1:

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

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

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

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

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

Proof:

When $t=1$, $S_O(1)$ is $1-\beta$ times than $S_O(0)$, and when $t=0$, the number of offline nodes transited from nodes in state S is $\alpha \bullet S_N(0)$.

Therefore $S_O(1) = \alpha \bullet S_N(0) \bullet (1-\beta)$.

When $t=2$, $S_O(2)$ is $1-\beta$ times than $S_O(1)$. When $t=1$, the number of offline nodes transited from nodes in state S equals $S_o(1) + \alpha \bullet S_N(1)$.

Therefore $S_O(2) = \alpha \bullet \sum_{i=0}^{2-1} S_N(i) \bullet (1-\beta)^{2-i}$.

Similarly, the theorem $S_o(t) = \alpha \sum_{i=0}^{t-1} S_N(i)(1-\beta)^{t-i}$ holds.

Evidenced by the same token:

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

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

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

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

Theorem 2:

$$A_N(t+1) = \varphi \bullet S_N(t) \bullet [1-(1-1/T)^{I_N(t) \bullet \lambda \bullet (1-I_N(t)/S_N(0))}]$$

Proof:

The probability of an $S-state$ node that has not been selected by a worm node based on random scanning strategy in one-time scan equals $1-1/T$; there were $I_N(t)$ online infected nodes at time t, every online infected node will be scanned λ times in each round; given that several different active worms based on random scanning strategy are likely to pick up the same node to attack, the number of effective scans for such attack is approximately subject to the law of logistic block growth. The ratio of effective scans to all scans could be calculated as $1-I_N(t)/S_N(t)$, thus all online nodes in state I would conduct $\lambda \bullet I_N(t) \bullet [1-I_N(t)/S_N(t)]$ effective scans at time t; probability of a node in state S that would be scanned effectively at least once by a worm node is $1-(1-1/T)^{\lambda \bullet I_N(t) \bullet [1-I_N(t)/S_N(0)]}$; meanwhile, there are $S_N(t)$ online nodes in state S at time t, if these nodes are scanned by worm nodes, they would be infected with a probability φ. Therefore the number of additional online nodes whose states are transited from S to L at time $t+1$ is:

$$A_N(t+1) = \varphi \bullet S_N(t) \bullet [1-(1-1/T)^{I_N(t) \bullet \lambda \bullet (1-I_N(t)/S_N(0))}].$$

Theorem 3:

$$S_N(t+1) = (1-\alpha-r_1) \bullet S_N(t) + \beta \bullet S_o(t) + \delta \bullet O(t) - A_N(t+1)$$

Proof:

At time $t+1$, in addition to original $S_N(t)$ nodes in state S, there are $\beta \bullet S_o(t) + \delta \bullet O(t)$ nodes whose states have been transited from O to S; meanwhile, there are $A_N(t+1)$ nodes whose states have been transited from S to L, $\alpha \bullet S_N(t)$ nodes whose states have been transited from S to O, and $r_1 \bullet S_N(t)$ nodes whose states have been transited from S to R. Hence, theorem 3 holds.

Theorem 4:

$$L_N(t+1) = (1-\alpha-r_2-\eta) \bullet L_N(t) + \beta \bullet L_o(t) + A_N(t+1)$$

Proof:

At time $t+1$, in addition to original $L_N(t)$ nodes in state L, there are $\beta \bullet L_o(t)$ nodes whose states have been transited from O to L and $A_N(t+1)$ nodes whose states have been transited from S to L; meanwhile, there are $\alpha \bullet L_N(t)$ nodes whose states have been transited from L to O, $r_2 \bullet L_N(t)$ nodes whose states have been transited from L to R and $\eta \bullet L_N(t)$ nodes whose states have been transited from L to I. Hence, *theorem 4* holds.

Theorem 5:

$$I_N(t+1) = (1-\alpha-r_3-\chi) \bullet I_N(t) + \beta \bullet I_o(t) + \eta \bullet E_N(t)$$

Proof:

At time $t+1$, in addition to original $I_N(t)$ nodes in state I, there are $\beta \bullet I_o(t)$ nodes whose states have been

transited from O to I and $\eta \bullet E_N(t)$ nodes whose states have been transited from L to I; meanwhile, there are $\alpha \bullet I_N(t)$ nodes whose states have been transited from I to O, $r_3 \bullet I_N(t)$ nodes whose states have been transited from I to R and $\chi \bullet I_N(t)$ nodes whose states have been transited from I to Q. Hence, *theorem 5* holds.

Theorem 6:

$$Q_N(t+1) = (1-\alpha-r_4) \bullet Q_N(t) + \beta \bullet Q_o(t) + \chi \bullet I_N(t)$$

Proof:

At time $t+1$, in addition to original $Q_N(t)$ nodes in state Q, there are $\beta \bullet Q_o(t)$ nodes whose states have been transited from O to Q and $\chi \bullet I_N(t)$ nodes whose states have been transited from I to Q; meanwhile, there are $\alpha \bullet Q_N(t)$ nodes whose states have been transited from Q to O and $r_4 \bullet Q_N(t)$ nodes whose states have been transited from Q to R. Hence, *theorem 6* holds.

Theorem 7:

$$R_N(t+1) = (1-\alpha) \bullet R_N(t) + \beta \bullet R_o(t) + r_1 \bullet S_N(t) + r_2 \bullet L_N(t) + r_3 \bullet I_N(t) + r_4 \bullet Q_N(t)$$

Proof:

At time $t+1$, in addition to original $R_N(t)$ nodes in state R, there are $\beta \bullet R_o(t)$ nodes whose states have been transited from O to R and $r_1 \bullet S_N(t) + r_2 \bullet L_N(t) + r_3 \bullet I_N(t) + r_4 \bullet Q_N(t)$ nodes whose states have been transited from various states to state R; meanwhile, there are $\alpha \bullet R_N(t)$ nodes whose states have been transited from R to O. Hence, theorem 7 holds.

Theorem 8：

$$O(t+1) = (1-\beta) \bullet O(t) + \alpha \bullet [S_N(t) + E_N(t) + I_N(t) + Q_N(t) + R_N(t)]$$

Proof:

At time $t+1$, all nodes in offline state are composed of two portions, one is offline nodes that are not yet online in the previous unit of time and their number is $(1-\beta) \bullet O(t)$; another is additional offline nodes whose states have been transited from various states inherited from previous online nodes in the previous unit of time and their number is

$$\alpha \bullet [S_N(t) + E_N(t) + I_N(t) + Q_N(t) + R_N(t)].$$

Hence, *theorem 8* holds.

C. Propagation Model Based on Hit-list Scanning Strategy (HLS Model)

Propagation model of active worm based on this attack strategy should meet basically same theorems as PRS Model. To save space, we only outline different theorems with respect to this strategy:

Theorem 9:

$$A_N(t+1) = \varphi \bullet S_N(t) \bullet \{1 - [1 - 1/(S(t) - A_N(t))]^{I_N(t) \bullet \lambda}\}$$

where $S(0) = T \bullet \gamma$, $A_N(t) = 0$

Proof:

At time t, there are $S(t) - A_N(t)$ P2P nodes left in Hit-list that have not been scanned by worm nodes, so the probability of an uninfected S-state node no to be selected by a worm is $1 - 1/[s(t) - A_N(t)]$; meanwhile, there are $I_N(t)$ online infected nodes, each of them will be scanned λ times for each round. Therefore the probability of a node in state S that would be scanned at least once at time t is $S_N(t) \bullet \{1 - [1 - 1/(S(t) - A_N(t))]^{I_N(t) \bullet \lambda}\}$; besides probability of a node that would be infected by worm node is φ. Hence, theorem 9 holds.

D. Propagation Model Based on Topological Scanning Strategy (TPS Model)

The mean degree value of a node in unstructured P2P networks meets power-law distribution according to [18], which means the probability of any node with k degree is $p(k) = C_1 \bullet (\varpi / k^{\varepsilon})$ in unstructured P2P networks; meantime, each node in structured P2P networks has the same degree.

To save space, we only outline different theorems for this attack strategy. The propagation model of active worms based on topological scanning strategy in unstructured P2P networks should meet following theorems:

Theorem 10: Average probability of a node in state S that was infected and transited to state I by connecting to an infected node at time t in unstructured P2P networks is $\varphi \bullet \Theta(M(t))$ and $\Theta(M(t)) = \sum (sP(s) \bullet I_N^{(k)}(t) / S_N^{(k)}(1)) / \sum kP(k)$

Proof:

At time t, the proportion of infected online nodes with k degree to all the nodes also with k degree equals $I_N^{(k)}(t) / S_N^{(k)}(1)$ and the average probability of a node in state S connected to a node in state I at time t in unstructured P2P networks is $\Theta(M(t))$ according to [18]; the probability of the node in state S to be infected by a worm node is φ. Hence, theorem 10 holds.

Theorem 11: $A^{(k)}(t+1) = S_N^{(k)}(t)[1 - (1 - \varphi \bullet \Theta(M(t)))^k]$

Proof:

The probability of a node with k degree in state S that is not infected by its neighbors at time t in unstructured P2P networks is $1 - \varphi \bullet \Theta(M(t))$ according to theorem 10. Since this node has k neighbors, the probability of a node with k degree in state S that

would be infected at least once by its neighbors equals $1-(1-\varphi\bullet\Theta(M(t)))^k$; meanwhile, there are $S_N^{(k)}(t)$ online nodes in state S, thus the number of additional online nodes with k degree whose states have been transited from S to L at time $t+1$ is $A^{(k)}(t+1)=S_N^{(k)}(t)[1-(1-\varphi\bullet\Theta(M(t)))^k]$. Hence, theorem 11 holds.

Theorem 12: $A(t+1)=S_N(t)[1-(1-\varphi\bullet\Theta(M(t)))^\theta]$

Proof:

The proof of this theorem is similar to the former one, except each node under structured P2P networks has the same degree. Therefore we only need to change the degree of all nodes from k to θ in theorem 11. Hence theorem 12 holds.

IV. Numerical Results And Performance Analysis

We studied the influence of various parameters in models proposed on the propagation speed of active worm in P2P networks by MATLAB. The platform we used is Microsoft Windows XP Professional SP3 co-working with 3.1 GHz Processor and 4 GB of memory built-in. The initial value of the P2P nodes is 10000. We observed the change tendency of infection ratio (infected host number / total vulnerable host number) by adjusting some particular variables in models to explore critical influencing factors applied on active worm propagation. Due to space limitation, we can only present a limited number of cases here. However, the conclusions we draw here generally hold for other cases we have evaluated.

A. Numerical Analysis Results of PRS Model

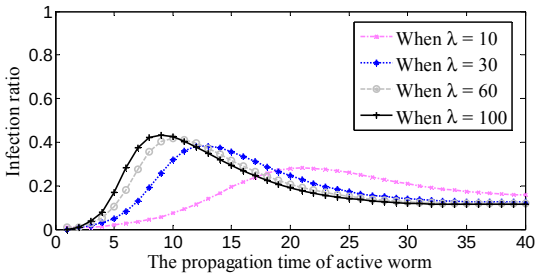

Figure 2. The scanning rate of a worm node in PRS model affects active worm propagation.

Fig. 2 shows the influence of the scanning rate of a worm node on the propagation speed of active worms. As the figure shown, the higher the scanning rate of a worm node is, the earlier the peak of infection ratio reaches and the greater infection ratio is in early stages of worm propagation; while at the middle or later stage of worm propagation, the lower the scanning rate of a worm nodes is, the greater the infection ratio is. The infection ratio caused by active worm propagation falls from the early peak and approximates successively to an equilibrium point within the middle and later stage.

The reason why Fig. 2 shapes this way is there are a large number of uninfected nodes with security vulnerability in early stages of worm propagation; thus the higher the scanning rate of a worm node based on PRS is, the more effective vulnerable nodes it will find in a single scanning round and the faster the infection ratio grows; however more and more nodes in state S have been transited to state L or state I with rapid growth of worm node numbers in P2P networks, when infection ratio reaches its maximum, the number of the effective vulnerable nodes that can be found by worm nodes begins dropping in each scanning round, leading to the decline of infection ratio at the middle and later stage. Once the incremental infected nodes equal to subtractive infected nodes in one scanning round, the infection ratio achieves the balance and tends to maintain it. Meanwhile, the probability of a worm node with lower scanning rate being detected and quarantined by monitoring software for its probing attack is lower than that of a worm node with higher scanning rate, leading to dropping of worm nodes' number with lower scanning rate whose states have been transited to state Q comparing with higher scanning rate ones. Therefore the lower the scanning rate of a worm node is, the greater the infection ratio is at the middle and later stage of worm propagation.

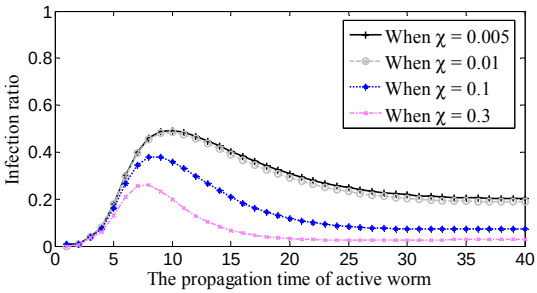

Figure 3. The detection rate of monitoring software in PRS model affects active worm propagation.

Fig. 3 manifests the influence of monitoring software's detection rate on the propagation speed of active worms in PRS model. Give the figure, it is obvious to find that the lower the detection rate of monitoring software is, the greater the infection ratio is. Likewise, the infection ratio caused by active worm propagation falls from the early peak and approximates successively to an equilibrium point at the middle and later stage.

This is basically because there are fewer worm nodes in early stages of worm propagation, the probability of monitoring software to find infected nodes spreading worm file fragments is relatively low. Hence active worms can spread quickly, and the infection ratio also increases rapidly; however then more and more worm nodes are detected and quarantined by monitoring software with rapid growth of worm node numbers in P2P networks, causing the decline of infection ratio at the middle and later stage. Likewise, when the incremental infected nodes are equal to the subtractive infected nodes in a single round, the infection ratio achieves the balance and maintains it. Meanwhile, the higher the detection rate of monitoring software is, the larger the probability of detecting worm behavior is and the better the inhibitory effect of active worm propagation is.

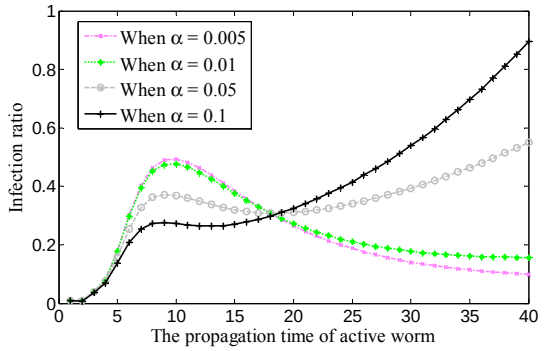

Figure 4. The offline rate of an online node in PRS model affects active worm propagation.

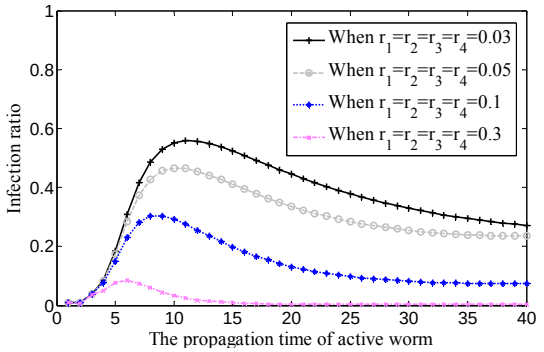

Figure 5. The level of immune response of an online node in PRS model affects active worm propagation.

Fig. 4 shows the influence of the offline rate on an online node towards propagation speed of active worms in PRS model. As the figure shows, the lower the offline rate of an online node is, the greater the infection ratio is in the early stages of worm propagation, while at the middle and later stage of worm propagation, the higher the offline rate of an online node is, the greater the infection ratio is.

This phenomenon indicates that there are a large number of uninfected nodes with security vulnerability in early stages of worm propagation, providing a lot of potential targets for active worms based on PRS model. By this stage, the lower the offline rate of an online node is, the more online nodes in state S there are, the more potential targets of attack for active worms are and the faster the infection ratio rises. However after infection ratio reaches its early maximum, more and more nodes in state S have been transited to state L or state I, the number of potential targets that can be attacked by worm nodes gradually dwindles away, meanwhile more and more worm nodes have been detected and quarantined by monitoring software due to spread of worm file fragments, causing drop of infection ratio. Then more and more various offline nodes are transited to state S since they have reinstalled OS and been back online. The tendency will even reinforce more with increase of offline rate on online nodes, providing more new potential targets for active worms. Hence, the higher the offline rate of an online node is, the faster the infection ratio rises at the middle and later stage of worm propagation.

Fig. 5 shows influence of an online node's different levels of immune response on propagation speed of active worms in PRS model. Judging from the figure given, the higher the level of immune response of an online node is, the lower the infection ratio is. Likewise, the infection ratio caused by active worm propagation falls from the early peak and approximates successively to an equilibrium point at the middle and later stage.

It could be explained as the higher the level of immune response of online nodes is, the larger the probability of security software discovering various kinds of online nodes with security vulnerability is. Therefore the lower the infection ratio is. When the number of incremental infected nodes equals to the subtractive infected nodes' in a single attack round, infection ratio achieves the balance and tends to maintain it. Obviously, the level of immune response of online nodes has great influence on active worm propagation.

B. Numerical Analysis Results of HLS Model

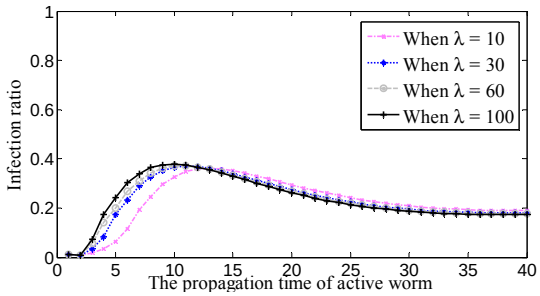

Figure 6. The scanning rate of a worm node in HLS model affects active worm propagation

Fig. 6 shows that the scanning rate of a worm node based both on random scanning strategy and Hit-list scanning strategy produces the same effect on the propagation speed of active worms. The higher the scanning rate of a worm node is, the earlier the peak of infection ratio reaches and the greater the infection ratio is in early stages of worm propagation; while at the middle and later stage of worm propagation, the lower scanning rate of a worm nodes is, the greater the infection ratio is, (the reason of this result is shown in the analysis of figure 2) but the scanning rate of a worm node based on Hit-list scanning strategy has less impact on the infection ratio than that of a worm node based on random scanning strategy since all online active worms based on HLS only select targets from the same Hit-list, there are only small number of non-scanned targets left in the Hit-list at the middle or later stage of worm propagation. Hence the number of remaining worm nodes influenced by scanning rate has little effect on infection ratio.

Fig. 7 shows the influence of the detection rate of monitoring software on propagation speed of active worms in HLS model. As the figure shown, the lower the detection rate of monitoring software is, the greater the infection ratio is. Likewise, the infection ratio caused by active worm propagation falls from the early peak and approximates successively to an equilibrium point at the middle and later stage. This figure can be analyzed

referring to figure 3, we won't dwell on it due to space limitation.

Fig. 8 shows influence of the offline rate applied to an online node on the propagation speed of active worms

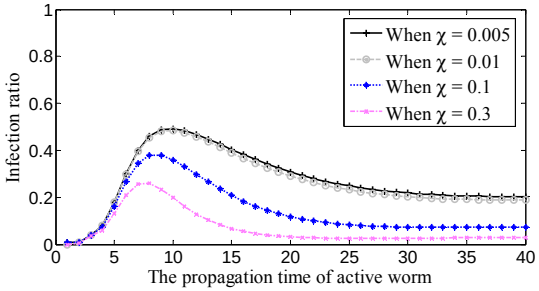

Figure 7. The detection rate of monitoring software in HLS model affects active worm propagation

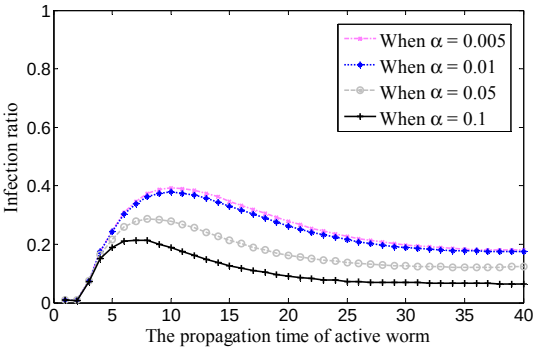

Figure 8. The offline rate of an online node in HLS model affects active worm propagation

under HLS model. As the figure shown, the higher the offline rate of an online node is, the lower the infection ratio is. The infection ratio caused by active worm propagation falls from the early peak and approximates successively to an equilibrium point at the middle and later stage.

As you can see in Fig. 4 and Fig. 8, the offline rate of an online node has more significant effect on active worm propagation based on PRS than that of HLS. The reason is that all active worms based on HLS only select targets from the Hit-list, the higher the offline rate of an online node is, the shorter the valid period of Hit-list is, the lower probability of active worms selecting effective online targets is and the lower the infection ratio caused by active worm propagation based on HLS is. Most of the nodes in Hit-list are still online in early stages of worm propagation, making them effective targets of active worms based on HLS and thus the infection ratio rises rapidly; after infection ratio reaches its early maximum, the number of effective targets that can be selected by active worms left in Hit-list gradually dwindles down, causing the infection ratio drop down; when the incremental infected nodes equal to the subtractive infected nodes in a single attack round, the infection ratio achieves the balance and maintains it.

Fig. 9 shows the influence of an online node's different levels of immune response on propagation speed of active worms in HLS model. As the figure shown, the higher the level of immune response of an online node is, the lower the infection ratio is. Likewise, the infection ratio caused by active worm propagation falls from the early peak and approximates successively to an equilibrium point at the middle and later stage. This figure can be analyzed referring to figure 5, we won't dwell on it due to space limitation.

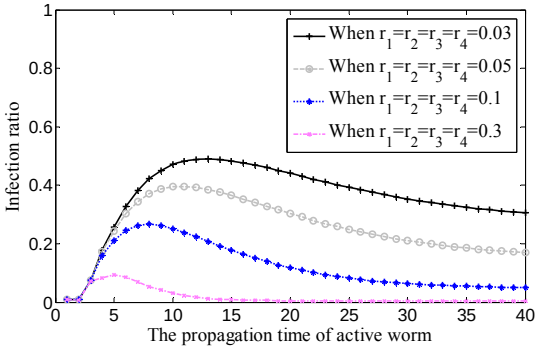

Figure 9. The level of immune response of an online node in HLS model affects active worm propagation n

C. Numerical Analysis Results of TPS Model

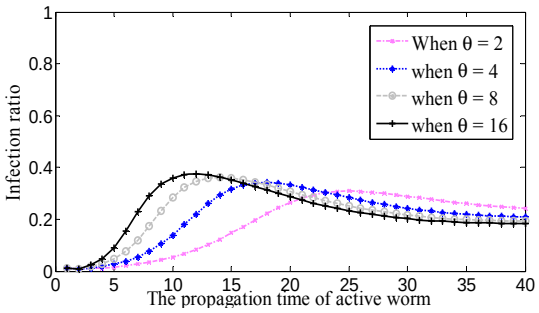

Figure 10. Degree of nodes in TPS model affects active worm

Fig. 10 shows the influence of degree of a node in structured P2P networks on the propagation speed of active worms. As the figure shown, the higher the degree of a nodes is, the greater the infection ratio is and the earlier the peak of infection ratio reaches in early stages of worm propagation; while at the middle and later stage of worm propagation, the higher the degree of a node is, the lower the infection ratio is; when the degree of a node in structured P2P networks reaches a fixed value, a small increase in the degree of a node may not address significant impact on the propagation speed of active worms. Experimental results of the propagation model of active worms based on TPS model in unstructured P2P networks can be found in [14], which we do not offer specifics due to space limitation.

This is because there are a large number of uninfected nodes with security vulnerability in early stages of worm propagation. The higher the degree of a nodes in structured P2P networks is, the more effective targets active worms will be found in a single scanning round and the faster the infection ratio grows; however more and more nodes in state S have been transited to state L or state I with rapid growth of worm node numbers in

P2P networks; after infection ratio reaches its maximum, the number of the effective targets that can be found by worm nodes gradually dwindles away in each scanning round, which leads to the decline of infection ratio at the middle and later stage of worm propagation. When the number of incremental infected nodes equals to the subtractive infected nodes' in a single round, infection ratio achieves the balance and tends to maintain it. Meanwhile, the lower the degree of a node is in structured P2P networks, the fewer the number of nodes can be simultaneously scanned by active worms based on TPS in each scanning round, and the lower the probability of a worm node based on TPS detected and quarantined by monitoring software is. Therefore, the lower the degree of a node is, the greater the infection ratio is at the middle and later stage of worm propagation.

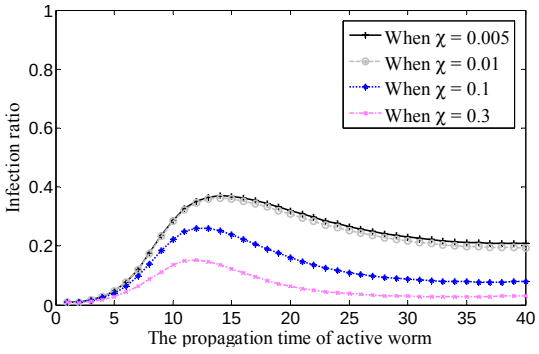

Figure 11. The detection rate of monitoring software in TPS model affects active worm

Fig. 11 shows the influence of detection rate of monitoring software on propagation speed of active worms in TPS model. As the figure shown, the lower the detection rate of monitoring software is, the greater the infection ratio is. Likewise, the infection ratio caused by active worm propagation falls from the early peak and approximates successively to an equilibrium point at the middle and later stage. This figure can be analyzed referring to figure 3, therefore we won't dwell on it due to space limitation.

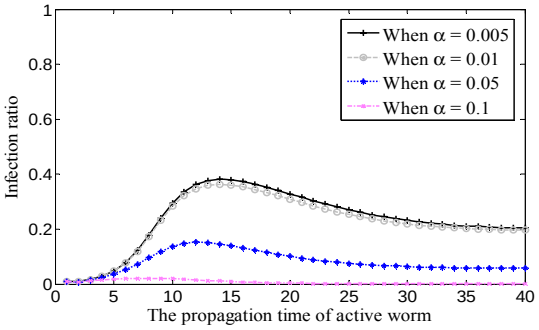

Figure 12. The offline rate of an online node in TPS model affects active worm

Fig. 12 shows the influence of an online node's offline rate on propagation speed of active worms in TPS model. As the figure shown, the higher the offline rate of an online node is, the lower the infection ratio will be and the infection ratio caused by active worm propagation falls from the early peak and approximates successively to an equilibrium point at the middle and later stage of worm propagation.

The reason is that active worms based on TPS select targets according to network topology; the higher the offline rate of an online node is, the lower probability of active worms selecting effective online targets is; the higher the offline rate of an online node is, the lower the infection ratio caused by active worm propagation based on TPS is. Most of the nodes in P2P networks are still online in early stages of worm propagation, making them effective targets of active worms based on TPS, and the infection ratio rises gradually; after infection ratio reaches its early maximum, the number of online targets that can be selected by active worms left in P2P networks dwindles away and causes gradual drop of infection ratio; when the number of incremental infected nodes equals to subtractive infected nodes' in a single attack round, infection ratio achieves the balance and tends to maintain it. As you can see in figure 8 and figure 12, the offline rate of an online node has more significant effect on active worm propagation based on TPS than that of HLS.

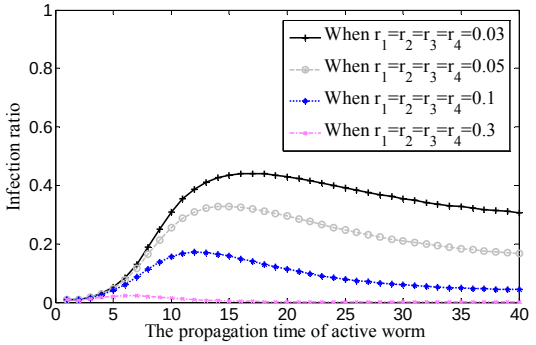

Figure 13. The level of immune response of an online node in TPS model affects active worm

Fig. 13 shows the influence of different levels of immune response of an online node on propagation speed of active worms in TPS model. As the figure shown, the higher the level of immune response of an online node is, the lower the infection ratio is. Likewise, the infection ratio caused by active worm propagation falls from the early peak and approximates successively to an equilibrium point at the middle and later stage. As you can see in figure 8 and figure 12, the level of immune response of an online node has more significant effect on active worm propagation based on TPS than that of active HLS. The analysis can be referred to figure 5, we won't dwell on it due to space limitation.

In conclusion, it can effectively suppress propagation of active worms by increasing detection rate of monitoring software and improving the level of immune response of online nodes, no matter which attack strategy active worms are based on. Besides, increasing offline rate has better inhibition to propagation of active worms based on HLS or TPS.

V. Conclusion And Future Work

In this paper, we studied three major attack strategies of active worms in P2P networks, provided states transition process of nodes as active worms spread in accordance with these strategies respectively. Then three propagation models based on these attack strategies were developed with fair consideration of dynamic features within P2P nodes. Next a mathematical analysis towards propagation models proposed was conducted and testified with various parameters affecting active worm propagation. Finally we deduced a number of key parameters affecting propagation speed of active worms in P2P networks.

For future work, we plan to further the study of improving detection rate of monitoring software based on dynamic characteristics of active worms, as well as building an efficient defense system for preventing active worms from these work.

Acknowledgment

This work is supported by China Postdoctoral Science Foundation(Grant No. 2014M552308), NCET, NSFC (Grant No. 61272400，Grant No. 61309031, 61309032, 61272400), Chongqing Innovative Team Fund for College Development Project (Grant No. KJTD201310), Program for Innovation Team Building at Institutions of Higher Education in Chongqing (Grant No. KJTD201310), Natural Science Foundation of Chongqing, (Grant No. cstc2013jcyjA40026), Scientific and Technological Research Program of Chongqing Municipal Education Commission (Grant No. KJ130525, KJ130523), and CQUPT Research Fund for Young Scholars (Grant No. A2012-79).

References

[1] Z. S. Chen, L. X. Gao, K. Kwiat, “Modeling the Spread of Active Worms,” *Proc. IEEE INFOCOM 2003*, vol. 3, pp. 1890-1900, March, 2003.

[2] W. Yu, C. Boyer, S. Chellappan, and D. Xuan, “Peer-to-Peer System-based Active Worm Attacks: Modeling and Analysis,” *Proc IEEE International Conference on Communications (ICC05)*, vol 1, pp. 295-300, May, 2005.

[3] W. Yu, “Analyzing the performance of Internet worm attack approaches,” *Proc 13th International Conference on Computer communications and Networks*, *Chicago*, vol. 1, pp. 501-506, Oct, 2004.

[4] C. F. Jia, X. Liu, Z. C. Hu, G. Y. Liu, and Z. Wang, “Defending P2P Networks against Malicious Worms Based on Benign Worms,” *Advances in Electric and Electronics*, vol. 155, pp. 653-660, 2012.

[5] X. S. Wang, J. L. Zhu, H. Z. Lin, X. M. Su and Y. Q. Jiang, “Modeling Propagation of Active P2P Worm in Chord Network,” *Advances in Intelligent and Soft Computing*, vol. 133, pp. 389-396, 2012.

[6] V. Ravikumar and M. Rajani, “Peer-to-Peer Networks for Modeling Malware Propagation,” *International Journal of Advanced Research in Computer Science and Software Engineering*, *India*, vol. 2, pp. 1-4, 2012.

[7] X. Fan, Y. Xiang, “Modeling the Propagation of Peer-to-Peer Worms under Quarantine,” *Proc. Network Operations and Management Symposium (NOMS)*, *IEEE*, pp. 942-945, April 2010.

[8] X. Fan, Y. Xiang, “Modeling the Propagation Process of Topology-Aware Worms: An Innovative Logic Matrix Formulation,” *Proc. Network and Parallel Computing, 2009. NPC '09*. Sixth IFIP International Conference, Oct, 2009, pp.182-189.

[9] Y. J. Zhang, Z. T. Li, Z. B. Hu, Q. F. Huang, C. W. Lu, “Evolutionary Proactive P2P Worm: Propagation Modeling and Simulation,” *Proc. Genetic and Evolutionary Computing, WGEC '08. Second International Conference, Hebei*, pp. 261-264, September 2008.

[10] W. Yang, M. G. Chang, Y. Yao, X. M. Shen “Stability Analysis of P2P Worm Propagation Model with Dynamic Quarantine Defense,” *Journal of Networks*, *Finland*, vol. 6, No.1, pp. 153-162, 2011.

[11] T. Chen, X. S. Zhang, H. Li, X. D. Li, and Y. Wu, “Fast quarantining of proactive worms in unstructured P2P networks,” *Journal of Network and Computer Applications(JNCA)*, vol. 34, No. 5, pp. 1648-1659, 2011.

[12] C. S. Feng, Z. G. Qin, L. Cuthbet, and L. Tokarchuk, “Propagation Model of Active Worms in P2P Networks,” *Proc. Young Computer Scientists, The 9th International Conference*, pp. 1908-1912, 2008.

[13] H. Li, Z. Qin, X. H. Pan, and X. S. Zhang, “Propagation Model of Non-scanning Active Worm in Unstructured P2P Network,” *Proc. of Multimedia Information Networking and Security, 2009. MINES '09. International Conference*, pp. 378-381, 2009.

[14] K. Suto, H. Nishiyama, X. M. Chen, and N. Kato. “Designing P2P Networks Tolerant to Attacks and Faults Based on Bimodal Degree Distribution”, *Journal of Communications*, *Finland*, vol. 7, pp.587-595, August 2012.

[15] Z. Z. Zhou, Y. L. Luo, L M. Guo, L. P. Sun, “Assessment of P2P Trust Model Based on Fuzzy Comprehensive Evaluation,” *Journal of Software,* vol. 8, No. 11, pp. 2711–2714, November 2013.

[16] L. M. Guo, Y. L. Luo, Z. Z. Zhou, M. J. Ji, “A Recommendation Trust Method Based on Fuzzy Clustering in P2P Networks,” *Journal of Software,* vol. 8, No. 2, pp. 357–360, February 2013.

[17] F. B. Meng, L. Ding, S. Peng, G. X. Yue, “A P2P Network Model Based on Hierarchical Interest Clustering Algorithm,” *Journal of Software,* vol. 8, No. 5, pp. 1262–1267, May 2013.

[18] P. Silvey, L. Hurwitz, "Adapting Peer-to-Peer Topologies to Improve System Performance," *Proc. the Hawaii International Conference on System Sciences*, pp.199-208, 2004.

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 Co., Ltd. He has published more than 11 refereed journal/conference papers. His current research interests are in network security, P2P applications, cloud computing.

Yukui Lu received B.S. degree in computer science from Hebei University of Science and Technology, China, in 2003**.** Now he is the director of the Propaganda Department, Shijiazhuang YiLing Pharmaceutical Co., LTD. He has published more than 3 refereed journal/conference papers. His current research interests are in network security, network applications, network maintenance.

Shitong Zhou is 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 and 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 scholar at Global Information and Telecommunication Institute, Waseda University, Tokyo from Sep. 2010 to Sep. 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, Cloud computing and etc.