Simulating Email Worm Propagation Based on Social Network and User Behavior

Kexin Yin[1], Wanlong Li[1], Ming Hu[3], and Jianqi Zhu[2]
[1]School of Computer Science and Engineering, Changchun University of Technology, China
[2]School of Computer Science and Technology, Jilin University, China
[3]School of Computer Technology and Engineering, Changchun Institute of Technology, China

Abstract: *Email worms pose a significant security threat to organizations and computer users today. Because they propagate over a logical network, the traditional epidemic model is unsuitable for modeling their propagation over the internet. However, it is no doubt that accurate modeling the propagation of email worms is helpful to contain th9eir attacks in advance. This paper presents a novel email worms' propagation model, which is based on a directed and weighted social network. Moreover, the effects of user's behavior are also considered in this model. To the author's knowledge, there is little information available considering the effects of them in modeling their propagation. A simulation algorithm is designed for verifying the effectiveness of the presented model. The results show that the presented model can describe the propagation of email worms accurately. Through simulating different containing strategies, we demonstrate that the infected key nodes in email social community can speed up the worm propagating. Last, a new General Susceptible Infectious Susceptible (G-SIS) email worm model is presented, which can predict the propagation scale of email worms accurately.*

Keywords: *Network security, Email worm propagation, social network, user Behavior, G-SIS.*

Received April 2, 2016; accepted November 27, 2017

1. Introduction

Email is one of the most convenient and indispensable communication mediums in our life [18]. However, email worms quickly evolved the ability to spread through the Internet by various means and constitute one of the major Internet security problems. Email worms are defined as a piece of malicious code that spreads through email by including a copy of itself in the email attachment [1, 17, 20]. Some famous email worms such as Melissa in 1999, love letters in 2000, W3/Sircam in 2001, SoBig in 2003 spread throughout the Internet and cause millions or even billions of dollars in damage [6, 19]. "Random scanning worms (Code Red or Slammer)" find vulnerabilities in computers by scanning IP addresses and compromise computers. Such email worms as "MyDoom" use social engineering techniques to lure email users to execute worm attachment [2]. Due to the characteristics of slow start and exponential propagation exhibited by email worms, it is challenging to model their propagation accurately and detect it at the early stage.

Email worms typically exploit the social network to propagate from one computer to another. Recent studies reveal that Internet exhibits a common property of community structure [12, 14]. Generally speaking, community structure is the natural division of the network into groups of vertices with denser connections within each group and fewer connections between groups, where vertices and connections represent network users and their social relationships, respectively. Members in each community usually share some common interests and thus, tend to socialize with other members more frequently than with ones from outside communities. Moreover, due to less interactions between communities, email worm propagates within a single community would be much faster than it does between communities on a social network. That is, the propagation of email worm is affected by the social network topology. Specially, a computer with high-degree (active node) in a community group, either in-degree or out-degree, should get higher priority for being monitored because it is more likely to be infected or to infect others. Since email worms propagate over a logical network that makes the traditional epidemic model unsuitable for modeling their propagation over the internet. The epidermic models describe the viral infections based on the birth and death rates of a virus. These models usually assume that each node in Internet has the same chance of being infected. [5, 9, 10, 11, 24] investigated the topological effects on the propagation of email worms, however, these researches focus mainly on the random scanning worms, which are not suitable for modeling the Internet email worms. Recently, Gang *et al.* [7] uses Barrat Barthelemy Vespignani (BBV) [3] weighted scale-free network model to simulate a power law network topology, and explores the virus propagation in this model. Yang *et al.* [23] studies the propagation of "Rose" in different scenarios through analytic model and discusses the impact of immune factors. Wang *et al.* [22] proposes a topology aware worm propagation model. Hayashi *et*

al. [8] studies the propagation of email virus through SIR model and obtained the optimal time to control the virus propagation. The above literature studied the topological effects on modeling the propagation of email worms, but, for one thing, they failed to consider the user's behavior impacts (such as the probability of opening attachment). Speaking of the effects of user's behavior, [26] presents an email worm model by considering the user's email checking time and the probability of opening attachments, however, the probability is seen as a random event and also they restricted treating the email network as an undirected graph, which cannot reflect the of relationships of computers in a community and their email exchanges. For another, they also neglected the capability of email worms have in exploiting the users' address book. Specifically, email worms can search a user's local address book and send emails to other users appears in the book. As a result, a victim computer receiving this email will most likely open it because he believes it comes from someone he knows or trusts.

Thus, we believe that an effective worm propagation model should take into account the depth of relationships between computers in network，which will be a key factor that can precisely describe the propagation of email worms. By figuring out the social interactions between computers, i.e., which computers are more likely to exchange emails with each other, we can predict the likeliest propagation path of such email worm. Except [26], to the author's knowledge, there is little information available in literature considering the depth of relationships between computers in modeling the email worms' propagation. Thus, the central point of this research lies in that we include the depth of relationships (represented by mutual trust degree in this paper) and the user's behavior in modeling the propagation of email worms, which helps us predict their propagation more accurately.

The contributions of this research are:

1. We construct a directed and weighted social network structure for email worm's propagation. This structure describes the social relationships between computers which are usually exploited by email worms for spreading and help us have a better understanding of their propagation in network. The simulation results verify that the presented network structure is conforming to the characteristics of social network.
2. We propose a new email worm simulation algorithm based on the behaviors of email users and email social network, on which we carry out extensive simulations that verify the critical roles of the combination of the user's behavior and the social network in modeling the propagation.
3. We discuss the key factors that affect the propagation of worms and compare the inhibitory effects of different measures. At last, based on the simulations, it demonstrates that the key node has the highest priority for being protection (the computer with the highest risk to be infected or infect others is the key node). At the end of the paper, we build an email worm analytic model General Susceptible Infectious Susceptible (G-SIS) to predict the propagation scale of email worm.

The rest of the paper is organized as follows: A directed and weighted email social network structure and simulations are introduced in section 2. In section 3 we present an analytic model for predicting the propagation of email worms. In Section 4 we analyze the protections against the propagation of email worms. Finally, section 5 concludes this paper with some discussions.

2. Email Worm Propagation Simulation

2.1. Construction of Email Social Network

In [26], email network was described as a simple undirected graph, which can't reflect the mutual relationships and email exchanges between users. Based on the concept of social network, this section presents an email directed and weighted social relationship graph based on Enron email dataset [13]. Enron email dataset is a public data collection of email communication in the real network and is widely used in the study of social network. The dataset collected 517,431 email communication records of Enron Corporation from October 1998 to July 2002, and included more than 80,000 network nodes. It does not include spam or virus emails, and reflect the users' interactions. The presented email social network is shown in Figure 1. Some notations marked in Figure 1 are illustrated in definition 1 and 2.

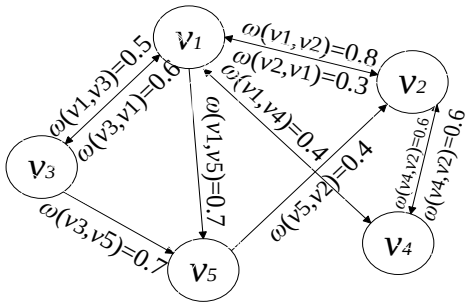


Figure 1. Email social network topology.

- *Definition* 1: $G =< V , E >$ is the directed and weighted network, $V=\{v_1,v_2,\ldots,v_n\}$ ($v_i \in V$), v denotes an email user. $|V|$ is the total number of email users. Degree of node v_i (denoted as $\lambda(v_i)$) is defined as the number of edges connected to v_i. $\bar{\lambda}$ is the average degree of email network. E is the edge set and $E=\{e(v_i, v_j)|v_i$ sends at least one email to v_j, $v_i, v_j \in V$, $1 \le i, j \le n$, $i \ne j\}$. $\omega(v_i, v_j)$ attaches to $e(v_i,v_j)$ represents the trust degree between v_i and v_j.
- *Definition* 2: The *key node* in email network is the node with high degree {v_i is the key node $|$ $\lambda(v_i) > M$

(M >0), M is the threshold}, which has a large amount of communications and with higher reliability. The communication frequency between users is derived by analyzing the Enron Email Dataset. In order to construct the topology of email communications, the weight matrix of edges between two users is constructed by the following steps.

- *Step* 1: Let weight $\omega(v_i, v_j)$ denote the frequency of sending emails from v_i to v_j. The more frequency of email exchanges between them the higher trust degree they have;
- *Step* 2: According to the node degree, adjust the weight with (1). The node with high degree is considered to be the key node;

$$\omega(v_i, v_j) = \omega(v_i, v_j)\sqrt{\frac{\lambda(v_j)}{\bar{\lambda}}} \quad (1)$$

- *Step* 3: Compute the average weight $\bar{\omega}$ of email network. The construction of email social network is achieved by the generation of weight matrix. This paper uses the Pajek [4] topology generator to produce Enron email network topology as shown in Figure 2. It is shown that the network topology shows a community structure in which a small number of key nodes are centers with higher degree.

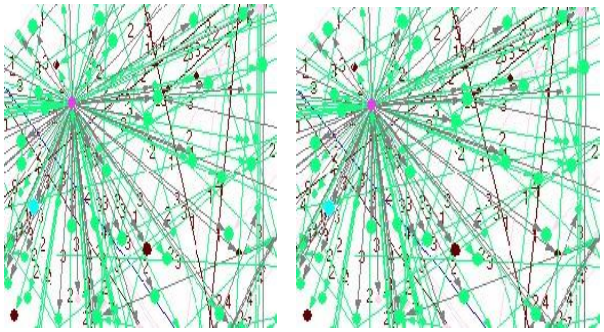


Figure 2. Enron email network (partial).

2.2. Simulation and Results

Intuitively, the propagation of email worm depends on the time and frequency distribution of using mailboxes, that is, the user's behavior pattern, which is different from RedCode (vulnerability scanning). Thus, the large-scale and rapid propagation of worm happens in the period of high frequency of using mailbox. If there are few mailboxes being used, the speed of worm propagation will slow down. Figure 3 is the statistics of Nyxem email worm [15] and the virus spreading fluctuates over time, it is shown that the virus spreading is at the low point on public holidays. As shown in Figure 4-a), the peak use of mailbox concentrates on working hours (2-4, 8-10 and 14-16 o'clock). In these periods, the virus spreading is faster. After 18 o'clock, the frequency of use is lower and the spreading will slow down. Figure 4-b) shows the ratio of active users per week. Thus, from these two figures, it is clear that the user's behavior can affect the worm propagation.

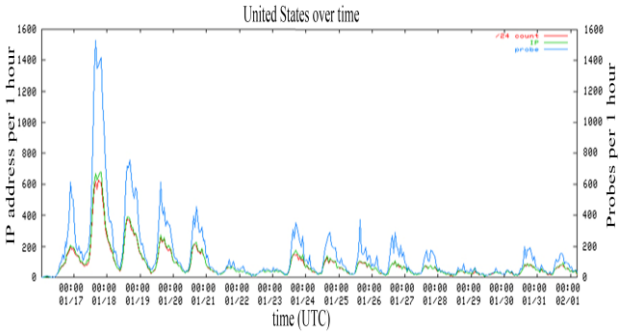
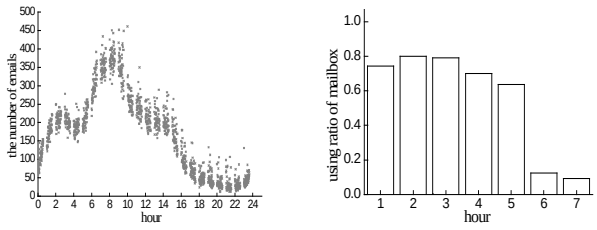


Figure 3. Propagation of nyxem worm (from CAIDA).



a) Statistics of user email each day. b) Ratio of active users per week.

Figure 4. Statistics of user's behaviors.

In order to know the effects of user's behaviors in modeling the propagation of email worms, this paper defines behavior pattern functions $\alpha(t)$ and $\mu(t)$, where $\alpha(t)$ denotes the statistics of user's emails each day and $\mu(t)$ denotes the ratio of active users per week. Let V denote the entire nodes (users), V_I denote the infected nodes, V_S denote the easily infected nodes and V_S^* denote the actively easily infected nodes, V_R denote the repaired nodes. $|V|$, $|V_I|$ and $|V_S|$ represent the size of V, V_I and V_S, respectively. Define $U(t)=\alpha(t)*\mu(t)$ to represent the ratio of active users at time t. In order to simulate the real propagation of email worm, the algorithm takes into account the user's behavior, the social network topology constructed in section 2.1. and the lower probability of reinfection of repaired nodes, etc., For continent, the variables used in this section are listed in Table 1.

Table 1. Variables definitions.

notation	implication
$\alpha(t)$	statistics of user's emails each day
$\mu(t)$	ratio of active users per week
N	the total number of users
V	the entire nodes, $\|V\| = N$
V_I	the infected nodes
V_S	the easily infected nodes
V_S^*	the actively easily infected nodes
$U(t)=\alpha(t)*\mu(t)$	the ratio of active users at time t
n_0	initially infected nodes
$N^*= U(t)N$	active nodes

2.2.1. Non-Reinfection Algorithm

- *Step* 1: Select n_0 initially infected nodes randomly, $|V_I| = n_0$, $|V_S| = N-n_0$ and $|V| = N$;
- *Step* 2: Select $N^*= U(t)N$ nodes as the active nodes randomly;

- *Step* 3: $\forall v \in V_S^*$, if $\exists u \in V_I$, $\exists e=(u,v) \in E$, it means that u and v have email exchange, moreover, v has received virus email from infected u. We assume that the infection probability of v is $p = \bar{p}\sigma\sqrt{\omega(e)/\bar{\omega}}$ ($\bar{p}$ is basic infection probability and σ is adjustment coefficient). Once infected, $V_S = V_S - \{v\}$; $V_I = V_I + \{v\}$.
- *Step* 4: Select $\gamma|V_I|$ infected nodes in V_I as the elements of V_R, execute the following operations: $V_S = V_S + V_R$; $V_I = V_I - V_R$; $\forall v \in V_R$, $\forall u \in V$ $(u \neq v)$, if $\exists e = (u,v) \in E$, $\omega(e) = \omega(e)/2$.
- *Step* 5: $t = t + 1$, go to Step2.

2.2.2. Reinfection Algorithm

- *Step* 1: Select n_0 initially infected nodes randomly, $|V_I| = n_0$, $|V_S| = N\text{-}n_0$ and $|V| = N$;
- *Step* 2: Select $N^* = U(t)N$ nodes as the active nodes randomly;
- *Step* 3: $\forall v \in V_S^*$, if $\exists u \in V_I$, $\exists e=(u,v) \in E$, we assume that the infection probability of v is $p = \bar{p}\sigma\sqrt{\omega(e)/\bar{\omega}}$, once infected, $V_I = V_I + \{v\}$;
- *Step* 4: Select $\gamma|V_I|$ *infected* nodes in V_I as the elements of V_R, execute the following operations: $V_I = V_I - V_R$; $\forall v \in V_R$, $\forall u \in V (u \neq v)$, if $\exists e = (u,v) \in E$, $\omega(e) = \omega(e)/2$;
- *Step* 5: $t = t + 1$, go to Step2.

Based on Enron Email Dataset, this paper builds the email directed and weighted network simulation environment, which includes 87106 nodes, 359817 directed and weighted edges. The average weight of directed edges is $\bar{\omega}$ =5 and the average node degree is $\bar{\lambda}$ = 8. The node degree and the edge weight are power law exponentially distributed with the exponents of 1.95 and 1.86, respectively. We select initially infected nodes randomly with normal distribution function $X \sim N$ (1,800) and simulate non-reinfection and reinfection cases, respectively. Figure 5 shows the propagation curves after 150 simulation runs. It is clear that the simulation results are basically consistent with the real propagation law of email worms (shown in Figure 3) and the algorithm further illustrates the critical roles of the user's behavior and the social network in modeling the propagation of email worms. Figure 6 shows the numbers of active users and active infected users at each time step. As shown in Figure 7, possible due to the old version of antivirus software or the weak security awareness of users, the worm propagation exhibits a rapid growth at the start of propagation. With the gradually increasing number of infected nodes, regular updating of software and increasing security awareness, the worm spread approaches to the repair rate. At the approximate 500 time step, it reaches to the maximum and after which levels off.

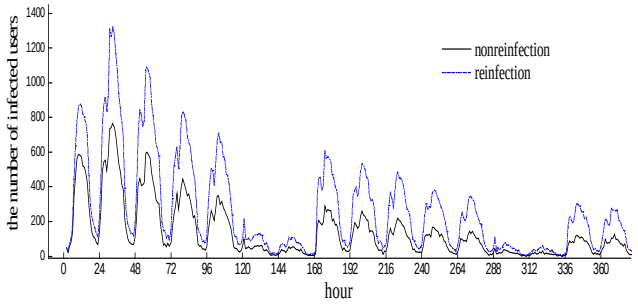


Figure 5. Behaviors of email worm propagation on simulation network.

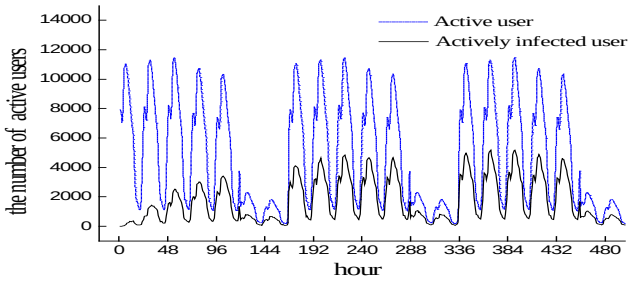


Figure 6. Statistics of active users.

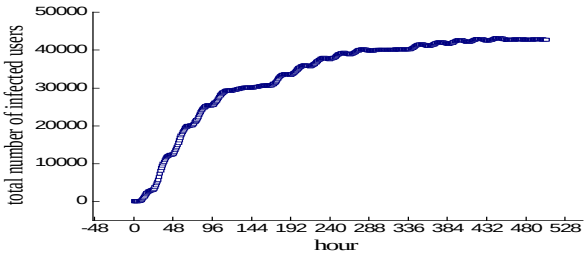


Figure 7. Statistics of email worm. propagation on simulation network.

3. Propagation Analytic Model

Many email worms' propagation models such as K-M, Susceptible Infectious Susceptible (SIS) [21] and two factor worm model [25] do not take into account the effects of user's behavior so that they can't describe the propagation accurately. Based on the classic SIS model, considering the user's behavior and the social network topology, we derive a general Internet worm propagation model called G-SIS.

As indicated before, $U(t)=\alpha(t)*\mu(t)$ denotes the ratio of active users at time t, in which $\alpha(t)$ denotes the *active degree* of user at different times of a day, $\mu(t)$ is used to adjust the user's *activity ratio* in a week, then we describe the effect of cyclical fluctuation of user's behavior on the worm propagation. Let $I(t)$ denote the number of infected users at time t, $S(t)$ denote the number of easily infected users at time t, N denote the total number of users. Then $I^*(t)=U(t)I(t)$ denotes the number of actively infected users, that is, $I^*(t)$ users will continue to spread virus at time t. $S^*(t)=U(t)S(t)$ denotes the number of actively easily infected users at time t.

$N^*(t)=U(t)N$ denotes the total number of active users at time t.

In fact, the infected nodes may be repaired by updating the system when he feels the system anomaly or by updating the antivirus software, et al. However, the email worm infection depends on the user's trust degree instead of the system vulnerability so that the repaired node may be reinfected. Thus, we use SIS epidemic model [16] to describe the user's state, and $R(t)$ represents the number of repaired nodes at time t. $\frac{dR(t)}{dt}=r\mu(t)I(t)$, r is the repair rate of infected users, $\mu(t)I(t)$ denotes the number of active infected users. The above variables are listed in Table 2.

Table 2. Variables definitions.

notation	implication
I(t)	the number of infected users at time t
S(t)	the number of easily infected users at time t
I*(t)=U(t)I(t)	the number of actively infected users
S*(t)=U(t)S(t)	the number of actively easily infected users at time t
N*(t)=U(t)N	the total number of active users at time t
R(t)	the number of repaired nodes at time t
μ(t)I(t)	the number of active infected users

Then, the change in the number of easily infected users after Δt period of time is as following:

$$S(t+\Delta t)-S(t)=-\beta(t)S^*(t)I^*(t)\Delta t+[R(t+\Delta t)-R(t)]$$

So,

$$\frac{dS(t)}{dt}=-\beta(t)S^*(t)I^*(t)+\frac{dS(t)}{dt} \quad (2)$$

The differential equation of the number of infected users at time t is:

$$\frac{dI(t)}{dt}=\beta(t)S^*(t)I^*(t)-\frac{dR(t)}{dt} \quad (3)$$

Where, $S^*(t)=N^*(t)-I^*(t)$, so,

$$\frac{dI(t)}{dt}=\beta(t)[N^*(t)-I^*(t)]-\frac{dR(t)}{dt} \quad (4)$$

Based on above,

$$\frac{dI(t)}{dt}=\beta(t)\alpha^2(t)\mu^2(t)[N-I(t)]I(t)-r\mu(t)I(t) \quad (5)$$

In [25], Zou simulated the effect of network congestion caused by worm scanning on the worm propagation, and the infection rate $\beta(t)$ is the function of time t. Note that the continuous random scanning does not work on email worm environment, thus, it is not easy to appear the fluctuation of infection rate caused by large-scale abnormal netflow. In email network, the infection rate $\beta(t)$ reflects the ability to spread virus at time t, that is, the degrees of infected nodes are correlated at time t. The greater the power law exponent is the more average distribution of the node degree is. This paper uses the power law exponent of scale-free network to measure the effect of social network on the infection rate. Define $\beta(t)=\beta(1-S(t)/N)^\gamma/t$ (γ is the power law exponent). We study the email worms' propagation and the parameters N=87106=$S(0)$, γ=0.002, $I(0)$=31, $\beta(t)=\beta(1-S(t)/N)^\gamma/t, (\beta=\bar{\lambda}*I(0)/N\approx 0.00285)$.

Figure 8 is the prediction of worm propagation scale by the integration of $\frac{dI(t)}{dt}(I(t)=\int_0^t\frac{dI(t)}{dt})$, which has a high degree of agreement with the results of simulation algorithm, thus, the G-SIS model can accurately predict the email worm propagation.

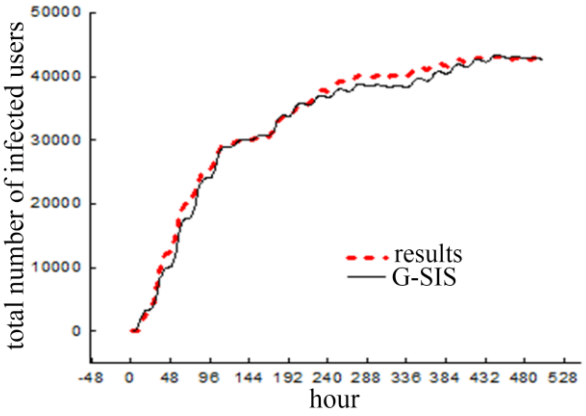


Figure 8. Comparisons of results between model prediction and simulation.

4. Containing Email Worms

The propagation of email worm code has randomness and uncertainty. This section will study the features of worm propagation from a micro-perspective by considering the effects of social network, and give the propagation trend from a macro-perspective. In this section we study the roles of key node in email social network by which a network operator would arrest a worm's propagation before it causes complete network-scale infection. Patch propagation techniques have been developed for delivering worm signatures to contain worms. However, in such a bandwidth-constrained environment, patches cannot be propagated by a network operator to all computers at the same time. Moreover, the patch would have to compete with the bandwidth already being consumed by a propagating email worm. Taking into account the spreading capability of email worms by exploiting the users' email address book, we take an approach towards patching the key nodes within the community first. Intuitively, such nodes once infected are most likely to infect the entire social community and hence they must be patched first.

- *Definition* 3: Assume $v\in V_I$ and $u\in V_S$. If there is a path from v to u (denoted as $L=(e_1, e_2, \ldots\ldots, e_m)$, e_i is the edge) and the infection probability through e_i is P_i, the infection probability of u through path L is $P_L=\prod_{i=1}^{m}P_i$, L is called the infection path of u with the probability of P_L.

- *Definition* 4: The degree of node u is $d=\lambda(u)$. That is, there are at least d paths from V_I to u. Let $D(u)$ denote the number of infection path of u ($D(u) \geq d$).

Assume that the nodes a, $b \in V_I$ and $\lambda(a) >> \lambda(b)$. Due to the randomness of the initially infected nodes, the length of infection path to a (or to b) and the probability of worm propagation on the edge are independent distributed. The infection probability of each path is independent Gaussian variable, that is, $P_g \propto N(\mu_p, \sigma_p^2)$. The experiment shows that the greater degree of the node is the larger average weight of the edge connected to it, as shown in Figure 9. Because $\lambda(a) >> \lambda(b)$, $\overline{\omega}(a) > \overline{\omega}(b)$ (the average weights of a and b are $\overline{\omega}(a)=\frac{\sum_{i=1}^{\lambda(a)}\omega(e_i)}{\lambda(a)}$ and $\overline{\omega}(b)=\frac{\sum_{j=1}^{\lambda(b)}\omega(e_j)}{\lambda(b)}$).

The infection probabilities of a and b are:

$$P(a)=\sum_{i=1}^{\lambda(a)} P_{g(i)}\sqrt{\frac{\omega(a_i)}{\overline{\omega}}} \text{ and } P(b)=\sum_{i=1}^{\lambda(b)} P_{g(i)}\sqrt{\frac{\omega(b_i)}{\overline{\omega}}} \text{ (assume } \sigma = 1)$$

Then,

$E(P(a))=\lambda(a)\mu_p\overline{\omega}(a)$ and $E(P(b))=\lambda(b)\mu_p\overline{\omega}(b)$ (μ_p is the average value of P_g, P_g is the independent Gaussian variable).

$E(P(a)) > E(P(b))$ and a has a greater infection probability than b so that the greater degree of the node is the greater chance of being infected it will. Figure 10 is the statistics of the average degrees of the infected nodes at each time step. It shows that the average degree has power law characteristic. That is, the node with larger degree (key node) will be infected quickly in the early stage of worm propagation. Therefore, key nodes should be protected and patched first. In order to illustrate this point, we simulate the scenarios of worm propagation under four different containing strategies as shown in Figure 11. It is shown that the protection to the key node can slow down the worm spread effectively.

- *Strategy* 1: Set the protection strength *Strength*(v) for node v whose degree is larger than the average degree. Then the greater degree of a node is the higher degree of protection should be taken.

$$Strength(v)=\begin{cases}(\lambda(v)\sqrt{\lambda})*2\%, \lambda(v) < 50\overline{\lambda}; \\ 100\% \quad ,other\end{cases} \tag{6}$$

- *Strategy* 2: Set the protection strength *Strength*(v) for the key node (M>150, Num=800 and M is threshed of degree). Then the greater degree of a node is the higher degree of protection should be taken.
- *Strategy* 3: Only protect the key node (M>100, Num=1500, Strength=85%).
- *Strategy* 4: Only protect the key node (M>50, Num=3000, Strength=60%).

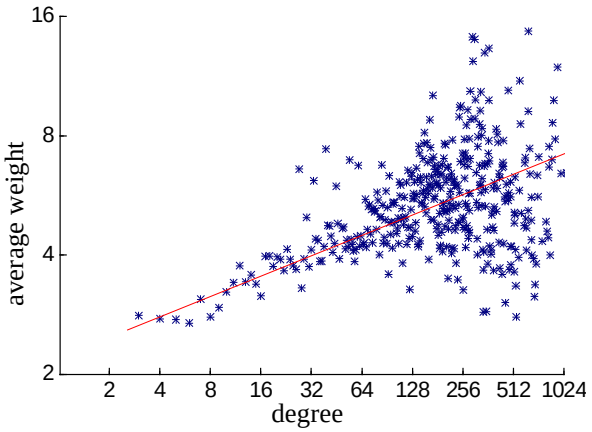


Figure 9. Relation between degree and average weight of a given node.

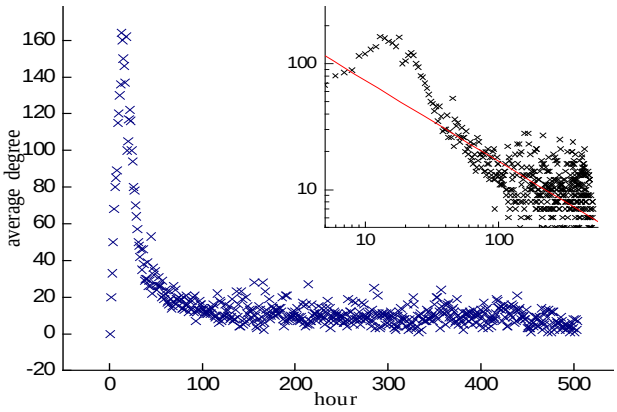


Figure 10. Average degree of infected nodes per hour.

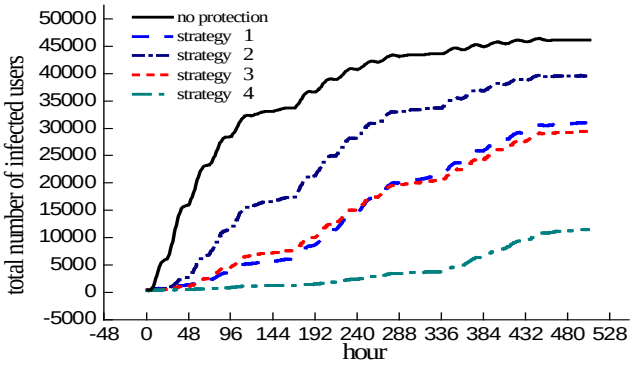


Figure 11. Impacts of different protection strategies on worm propagation.

5. Conclusions

Email worms have been posing a significant security threat to internet communities. The limitations of traditional epidemic models make it unsuitable for modeling their propagation.

In order to accurately modeling the propagation of email worms, this paper constructs an email social network with a directed and weighted topological structure and on which an email worm's propagation simulation algorithm is presented by considering the social network and the user's behavior. To the author's knowledge, there is little information available in literature considering the effects of them in modeling

their propagation. The simulation results show that the propagation simulation algorithm could reflect the propagation of email worms accurately, and also verifies the social character of email worms' propagation. Moreover, we analyze the critical roles of key nodes in social community and point out that preventing email worms from propagation can be achieved effectively by distributing patches to the most influential users (the key nodes in network). At the end of the paper, we give a G-SIS model for predicting the propagation scale of email worms. The simulation verifies the model's effectiveness.

Acknowledgments

This work was supported by the National Nature Science Foundation(61572229); CERNET Innovation Project(NGII20161004); Jilin Provincial International Cooperation Foundation (20150414004GH).

References

[1] Abdulla S., Ramadass S., Altaher A., and Al-Nassiri A., "Employing Machine Learning Algorithms to Detect Unknown Scanning and Email Worms," *The International Arab Journal of Information Technology*, vol. 11, no. 2, pp. 140-148, 2014.

[2] A Study of Mass-mailing Worms, http://www.cert.org/incident_notes/IN-2004-01.html, Last Visited, 2004.

[3] Barrat A., Barthelemy M., and Vespignani A., "Weighted Evolving Networks: Coupling Topology and Weighted Dynamics," *Physical Review Letters*, vol. 92, no. 22, pp. 22870-1-22870-4, 2004.

[4] Batagel V. and Mrvar A., "Pajek Program for Large Network Analysis," *Connections*, vol. 21, no. 2, pp. 47-57, 1998.

[5] Chen Z., Gao L., and Kwiat K., "Modeling the Spread of Active Worms," *in Proceedings of IEEE INFOCOM 22nd Annual Joint Conference of the IEEE Computer and Communications Societies*, San Francisco, pp. 1890-1900, 2003.

[6] Frequently Asked Questions about the Melissa Virus, http://www.cert.org/incident_notes/IN-2003-03.html, Last Visited, 2003.

[7] Gang Y., Tao Z., Jie W., Zhong-Qian F., and Bing-Hong W., "Epidemic Spread in Weighted Scale-Free Networks," *Chinese Physics Letters*, vol. 22, no. 2, pp. 510-513, 2005.

[8] Hayashi Y., Minoura M., and Matsukubo J., "Oscillatory Epidemic Prevalence in Growing Scale-Free Networks," *Physical Review E*, vol. 69, pp. 161121-161128, 2004.

[9] Kephart J. and White S., "Directed-Graph Epidemiological Models of Computer Viruses," *in Proceedings of IEEE Computer Society Symposium on Research in Security and Privacy*, Oakland, pp. 343-359, 1991.

[10] Kephart J., White S., and Chess D., "Computers and Epidemiology," *IEEE Spectrum*, vol. 30, no. 5, pp. 20-26, 1993.

[11] Kesidis G., Hamadeh I., and Jiwasurat S., "Coupled Kermackmckendrick Models for Randomly Scanning and Bandwidth-Saturating Internet Worms," *in Proceedings of the 3rd International Conference on Quality of Service in Multiservice IP Networks*, Catania, pp. 101-109, 2005.

[12] Kim K., Lee H., Hong J., Cho M., Fava M., Mischoulon D., Kim D., and Jeon H., "Poor Sleep Quality and Suicide Attempt Among Adults with Internet Addiction: A Nationwide Community Sample of Korea," *PlOS One*, vol. 12, no. 4, pp. 1-13, 2017.

[13] Klimt B. and Yang Y., "Introducing the Enron Corpus," *in Proceedings of the CEAS*, Mountain View, pp. 1-2, 2004.

[14] Massa F., "Guardians of the Internet: Building and Sustaining the Anonymous Online Community," *Organization Studies*, vol. 38, no. 7, pp. 959-988, 2017.

[15] Moore D. and Shannon C., "CAIDA: The nyxem email virus: analysis and inferences [EB/OL]," http://www.caida.org, Last Visited, 2004.

[16] Newman M., "The Structure and Function of Complex Networks," *SIAM Review*, vol. 45, no. 2, pp.167-256, 2003.

[17] Sheng W., Wei Z., Jun Z., Yang X., Wanlei Z., Weijia J., and Cliff Z., "Modeling and Analysis on the Propagation Dynamics of Modern Email Malware," *IEEE Transactions on Dependable and Secure Computing*, vo. 11, no. 4, pp. 361-374, 2013.

[18] Sneha S. and Swapna P., "Analyze and Prevent Modern Email Malware Propagation Using Sell Model," *IIOAB Journal*, vol. 7, no. 9, pp. 696-702, 2016.

[19] U.S. Department of Homeland Security Announces Partnership with Carnegie Mellon's CERT Coordination Center. http://securityresponse.symantec.com/avcenter/venc/data/w32.beagle.f@mm.html, Last Visited, 2004.

[20] Wang Y., Wen S., Xiang Y., and Zhou W., "Model the Propagation of Worms in Networks: A Survey," *IEEE Communications Survey and Tutorials*, vol. 16, no. 2, pp. 942-960, 2014.

[21] Wang Y. and Wang C., "Modeling the Effects of Timing Parameters on Virus Propagation," *in Proceedings of ACM Workshop on Rapid Malcode*, Washington, pp. 61-66, 2003.

[22] Wang Y., Ji-Wu J., Ji X., and Qi L., "Topology Aware Worm Simulation and Analysis," *Journal*

of Software, vol. 19, no. 6, pp. 1508-1518, 2008.
[23] Yang S., Jin H., Liao X., and Liu S., "Modeling Modern Social-Network-Based Epidemics: A Case Study of Rose," *in Proceedings of International Conference on Autonomic and Trusted Computing*, Oslo, pp. 302-315, 2008.
[24] Zou C., Gao L., Gong W., and Towsley D., "Monitoring and Early Warning for Internet Worms," *in Proceedings of the 10th ACM Conference on Computer and Communications Security*, Washington, pp. 190-199, 2003.
[25] Zou C., Gong W., and Towsley D., "Code Red Worm Propagation Modeling and Analysis," *in Proceedings of the 9th ACM Conference on Computer and Communications Security*, Washington, pp. 138-147, 2002.
[26] Zou C., Towsley D., and Gong W., "Modeling and Simulation Study of the Propagation and Defense of Internet Email Worm," *IEEE Transactions on Dependable and Secure Computing*, vol. 4, no. 2, pp. 105-118, 2007.

Kexin Yin received Ph.D. degree from Changchun University of Science and Technology in 2008. She is professor of computer science and engineering in Changchun University of Technology. Her areas of interests are machine learning, image processing and network security.

Wanglong Li received M.S. degree from Jilin University in 2002. He is professor of computer science and engineering in Changchun University of Technology. His areas of interests are computer network.

Ming Hu received Ph.D. degree from Jilin University in 2005. He is professor and present of Changchun Institute of Technology. His areas of interests are artificial intelligence and data mining.

Jianqi Zhu received Ph.D. degree from Jilin University in 2009. He is professor of computer science and technology of Jilin University. His areas of interests are network security, machine learning and data mining.