Modelling and Analyzing Passive Worms over Unstructured Peer-to-Peer Networks

Fangwei Wang[1,2], Yunkai Zhang[1], and Jianfeng Ma[2]
(Corresponding author: Fangwei Wang)

Network Center, Hebei Normal University[1]
No.113, Yu Hua Rd., Chang An District, Shijiazhuang, 050016, China (Email: fw_wang@hebtu.edu.cn)
Key Lab of Computer Networks and Information Security of Ministry of Education, Xidian University[2]
No.2, Tai Bai Rd., Xi'an, 710071, China

(Received Sept. 10, 2007; revised and accepted Apr. 23, 2008)

Abstract

Passive worm have posed serious security threats to the functioning of unstructured P2P networks. A delayed SEIRS epidemic model with death, off line and online rate is constructed based on the actual situation of P2P users. The basic reproduction number that governs whether a passive worm is extinct or not is obtained. In this model, time delay consists of latent and temporary immunity periods. The impact of different parameters on this model is studied with simulation results, especially the effect of time delay, which can provide an important guideline in the control of unstructured P2P networks as well as passive worm defense.

Keywords: Basic reproduction number, equilibrium, propagation model, passive worms, Peer-to-peer networks

1 Introduction

Unstructured Peer-to-Peer (P2P) overlay networks are distributed systems in nature, without any hierarchical organization or centralized control [11, 18, 19]. Each peer is a client as well as a server. All peers play the same role in logic instead of in the function. Unstructured P2P networks provide P2P worms some facilities to propagate, although they can decrease reliability and search capabilities, and increases network traffic.

P2P worm is formally proposed in the International Conference IPTPS 2005 [20] and their tremendous harm to network security have aroused widespread concerns in the community. Passive worms, one of the major forms of P2P worms, attach themselves to shared files and propagate as these files are downloaded and executed on other hosts. The worm resides in the shared folder of the infected hosts, under several names. When another peer downloads one of those files, the worm spreads to this host, and when the user runs the file, the worm duplicates itself under several attractive names in the shared folder of the new victim, and waits for other victims, and so on.

Modelling passive worms is useful to understand how particular factors can affect their propagation. Some epidemiological models have been proposed to study the propagation of scanning strategy-based worms and local subnet scanning worms [2, 4]; however, they can not be used to model passive P2P worms, because of the particular characteristics of passive P2P worms' propagation. Several passive worm propagation models have been investigated in earlier work [1, 3, 8, 10, 14, 15, 16, 21, 17, 20]. An important omission of the above models is the effect of network throughput. Modelling the propagation of passive worms, the authors assume that a vulnerable peer can be infected in a unit time. This assumption is certainly not accordant to the reality since the average file is 4MB [7], whereas, average bandwidth per peer is 132KBps [9], which neglects the diversity of file size in the shared folders. As a result, the downloading time is not a negligible factor, which has a significant effect on the simulating the passive worms propagation. Thus, inspired by the reference [1], we propose a new delayed SEIRS epidemic mathematically-based model accounting for the effect of network throughput based on the actual situation of P2P end users.

The rest of this paper is organized as follows: Section 2 gives a brief overview of the propagation models of passive worms. Section 3 proposes a new unstructured P2P-based passive worm attack model with death, off line and online rate based on the actual situation of P2P users, and analyze its steady-state behavior. Section 4 simulates the propagation of passive worms with different parameters. Section 5 concludes our paper.

2 Related Work

The issue of worms in peer-to-peer is addressed wherein the authors perform a simulation study of dangers posed by P2P worms and proceed to outline possible mitigation mechanisms [20]. The authors point out that P2P worm can stealthily propagate through the overlay topology and invalidate numerous defending mechanisms aiming at scanning worms. Moreover, they obtain the upper bound of the number of infected host. However, Zhou et al. do not take into account the effect of the node position on the worm propagation.

The advent of mathematical Epidemiology is generally credited to McKendrick [12]. Chen et al. [3] provide a workload-driven simulation framework to characterize three types of non-scanning worms (e.g., passive worm, reactive worm, and proactive worm) and identify the parameters influencing their propagations, which states that the type of worm that would spread over such a network would not be detected by many of current methods. Kalafut et al. [8] point out the fact that 68% of the executable files contain passive worms through months of data. Xia et al. [16] present epidemic models of P2P worms in three typical structured P2P networks, outline the worms' rapid spreading capability, and reveal the negative influences of overlay topologies on the worms' propagation. Krishna et al. [10] give a model for Gnutella-type P2P systems by addressing a parameter (TTL), and consider the possible victims of an infected peer are limited to those which are TTL hops away from it and not the whole P2P network. The introduction of such a characteristic would avoid false positives. Based on the two-factor model, Zhou et al. [21] take into account the effect of the time delay, and present a propagation model for passive worms. Thommes et al. [15] use an analytical model to assess the impact of a detection solution (Credence) on the P2P worm propagation, and to determine approximately how widespread the Credence system must be so as to combat the worm efficiently. Richard et al. [14] propose an improved SEI (Susceptible- Exposed-Infected) model to simulate virus propagation. However, they do not show the length of latency and take into account the impact of anti-virus software. The model (SEIR) proposed by Yan et al. [17] assumes that recovery hosts have a permanent immunization period with a certain probability, which is not consistent with real situation. In order to overcome limitation, Bimal et al. [1] present a SEIRS model with latent and temporary immunity periods, which can reveal common worm propagation. Due to the fact that it does not take into account the real situation of end-users, the model is also not be used to simulate passive worms over unstructured P2P networks.

3 Propagation Model for Passive Worms

3.1 Model Assumptions

We make the following assumptions in order to concisely and accurately reflect the propagation behaviors of passive worms. (1) The total number of peer population $N(t)$ is a variable changing with time t. (2) The total population is partitioned into four compartments: the susceptible compartment (S), the exposed compartment (E), the infected compartment (I), and the recovered compartment (R). (3) The latency period ω is a variable related to the downloaded file size; moreover, the immunity period τ is a constant related to the anti virus software. (4) The waiting time of the exposed, infected and recovered compartment is an exponent distribution. (5) When a peer is removed from the infected class, it obtains temporary immunity with probability p and died from the attack of passive worms with probability $(1-p)$.

From the assumptions above, the standard incidence of the total variable population size can be expressed as

$$N(t) = S(t) + E(t) + I(t) + R(t). \qquad (1)$$

Table 1 lists parameters and notations needed in this paper.

Table 1: Parameters and notations in this paper

Parameters	Notations
$S(t)$	Number of susceptible peers at time t.
$E(t)$	Number of exposed peers at time t.
$I(t)$	Number of infected peers at time t.
$R(t)$	Number of recovered peers at time t.
$b, \mu, \varepsilon, \alpha$	Per peer online rate, off line rate, death rate, recovery rate, respectively.
ω, τ	The latency and temporary immunity, respectively.
γ_1	Rate at which peers terminate ongoing downloads.
γ_2	Rate at which peers renew interest in downloading a file after having deleted it.
λ	Rate at which a peer generates queries.
p	Probability of temporary immunity acquired when a peer is recovered from the infected.
TTL	Number of hops a query can reach.
bw, s, m	Average bandwidth of all peers, size and the number of chunks of sharing files, respectively.

3.2 Propagation Model for Passive Worms

In order to clearly understand the propagation process of passive worms, we first study the search mechanisms adopted by some P2P networks, such as Gnutella. Among search mechanisms available, flooding method is the most popular. Peer A sends out a query for an expected file to all its neighbors. Peer B receiving such a request first check its local shared folder, and responds this request if possessing the file and then check the hop count of the query. If the value is greater than zero, it will forward the query to its neighbors; otherwise, the query is discarded. Due to the fact that Gnutella, Kazaa networks follow a power-law degree distribution [13], we use the generating function [13] to quantify the number of peers available while searching for the file. Define the generating function for vertex degree probability distribution as

$$G_0(x) = \Sigma_{k=0}^{\infty} p_k x^k, \tag{2}$$

where p_k is the probability that a randomly selected vertex on the network has degree k, which is given by the following equation $p_k = P_r(N = k)Ck^{-\delta}$ (C and δ are constant). The query is used to reach a recursive definition for the k-hop neighbors of a node in the network. Because an edge is chosen at random, it is more likely that it leads to a node with a higher degree. The generating function for the probability distribution of reaching a k degree node by traversing a randomly chosen edge can then be obtained as

$$\frac{\Sigma_k k p_k x^k}{\Sigma_k k p_k} = \frac{xG_0^{'}(x)}{G_0^{'}(1)}. \tag{3}$$

The distribution of the outgoing edges from the vertex chosen has one power of x lesser than the expression (3) and can thus be expressed as $G_1(x) = G_0^{'}(x)/G_0^{'}(1)$. In a similar fashion, the generating function for the number of two-hop neighbors is $\Sigma_k p_k[G_1(x)]^k = G_0(G_1(x))$. As a result, the recursive formulation for the distribution of the number of m-hop neighbors is expressed by the following equation $G_0(G_1(\cdots G_1(x)\cdots))$, which can be given as

$$G^{(m)}(x) = \begin{cases} G_0(x) & m = 1; \\ G^{(m-1)}(G_1(x)) & m \geq 2. \end{cases} \tag{4}$$

Through differentiating the generating function and substituting $x = 1$, we can obtain the average number of one and two hop neighbors of a peer which are given by $Z_1 = G_0^{'}(1) = \sum_k kp_k$ and $Z_2 = G_0^{''}(1)$, respectively. Similarly, the number of m-hop neighbors can be expressed as $Z_m = Z_1(Z_2/Z_1)^{m-1}$. Consequently, the average number of search neighborhood of a peer during TTL hops is obtained as

$$Z_{av} = \sum_{i=1}^{TTL} Z_i. \tag{5}$$

Due to the restriction of network bandwidth, the time taken a file be simultaneously downloaded by multi-peers is different to single peer. Now, we study the average time. Let the number of peers in a large-scale unstructured P2P network be fixed at N, and a limited number of peers, say 1, can serve them. Suppose the file has s bits and each peer has a limited download capacity, say bw bps. For simplicity, we assume that $N = Z_{av}^k$ users wish to obtain the expected file which is initially available at one peer. As a result, we can obtain the following Lemma 1 about the average delay for all peers in the whole transmitting process.

Lemma 1. *For the number of peers N, the average delay for peers is $\omega = \pi log_{Z_{av}}^N$ at least, where $\pi = s/bw$.*

Proof. From above hypothesis, we can obtain that Ns bits are required to be exchanged in order to serve N requests. It is clear that a good dissemination strategy is to first serve Z_{av} users at rate bw, at which point the service capacity grows to $bw(Z_{av} + 1) \approx bZ_{av}$, and then have Z_{av} peers serve remnant peers, until the N users are served. Under this idealized strategy, peers can complete service every $\pi = s/bw$ seconds, at which point there are an exponential growth of $Z_{av}^{t/\pi}$ in the number of peers available to serve the file. If the network follows these dynamics the N peers will be served by time $\pi log_{Z_{av}}^{N+1} = \pi k$. As a result, the average download delay experienced by peers can be computed as follows. Let ω_j denote the delay experienced by the jth peer to complete, and note that $Z_{av}^{(i-k)}N$ peers complete service at time $(i+1)\pi$, thus, an average delay for peers is

$$\begin{aligned} \omega &= \tfrac{1}{N}\textstyle\sum_{j=1}^{N} \omega_j \\ &= \textstyle\sum_{i=0}^{k-1} Z_{av}^{i-k}\pi(i+1) \\ &= k\pi - \tfrac{N-1}{N}\tau \\ &= \pi(log_{Z_{av}}^N - \tfrac{N-1}{N}) \\ &\approx \pi log_{Z_{av}}^N. \end{aligned} \tag{6}$$

□

In some networks (such as eDonkey), multi-part downloads strategy is utilized in order to improve the downloading speed. Suppose the file is divided into m chunks of identical size. In the following, we study its average delay for peers under this scenario.

Lemma 2. *For the number of peers N, the average delay for peers for downloading a big file (m chunks) is $\omega^{(m)} = (\pi/m)log_{Z_{av}}^N$ at least, where $\pi = s/bw$.*

Proof. When a peer has finished downloading a file chunk, it can start to server it. To illustrate this idea consider the following idealized strategy. We shall track service completions in time slots of size $(s/m)bw = \pi/m$. We suppose that the source of file sends Chunk 1 to a peer, Chunk 2 to another peer, and so on until it finishes disseminating the last Chunk m on slot m. Meanwhile each Chunk i is being duplicated in the network. Then at time k slot, the N peers can be partitioned into k sets A_i $(i = 1, 2, \cdots, k)$, with $|A_i| = Z_{av}^{k-i}$, and A_i corresponds to peers which have only received the ith Chunk. Now consider the $(k+1)$th time slot. Suppose the peers in A_1 send chunk 1 to the N/Z_{av} peers that have not yet received it. Meanwhile the peers in $A_i (i > 1)$, send Chunk i to a node in A_1 choosing a peer that has at this point only received Chunk 1. This process continues until all chunks are eventually delivered to all peers by time slot $k+m = (\pi/m)(log_{Z_{av}}^{(N-1)} + m)$. Because N/Z_{av} peers have received all chunks when Chunk $m-1$ completes duplication across all peers at time slot $k + m - 1$ and the rest ones will receive Chunk m during the last time slot $k + m$, thereafter the average delay experienced by peer can be expressed as follows.

$$\begin{aligned} \omega^{(m)} &= \tfrac{1}{N}\textstyle\sum_{j=1}^{N}\omega_j^{(m)} \\ &= \tfrac{1}{2}((k+m-1)+(k+m))\tfrac{\pi}{m} \\ &= \tfrac{\pi}{m}(log_{Z_{av}}^{N}+\tfrac{2m-1}{2}) \\ &\approx \tfrac{\pi}{m}log_{Z_{av}}^{N}. \end{aligned} \qquad (7)$$

□

From Lemma 2, we can obtain that the larger m is, the smaller the average delay for the downloading process is.

On the other hand, anti virus software can provide a temporary immunity with time τ instead of permanent immunity. τ is governed by the version and virus database of anti virus software. The temporary immunity period τ may get long if the anti virus software is often updating.

From the model assumptions in Section 3.1, we can obtain that the number of infected peers at time t is the number of exposed ones ($\frac{\lambda Z_{av}S(t-\omega)I(t-\omega)}{N(t-\omega)}$) at time $t-\omega$. Due to the off line of some peers, the probability of infected peers that are still online is $e^{-\mu\omega}$ from time $t-\omega$ to t. Using the same procedure and assumption, we can obtain that the number of peers from the recovered class to the susceptible class is $\alpha I(t-\tau)$, and the probability is $e^{-\mu\tau}$ from time $t-\tau$ to t.

As a result, the population transfer among compartments is schematically depicted in the transfer diagram in Figure 1.

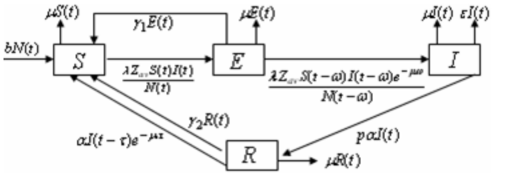
$\mu S(t)$
$\gamma_1 E(t)$
$\mu E(t)$
$\mu I(t)$
$\varepsilon I(t)$
$bN(t)$
S
E
I
$\frac{\lambda Z_{av}S(t)I(t)}{N(t)}$
$\frac{\lambda Z_{av}S(t-\omega)I(t-\omega)e^{-\mu\omega}}{N(t-\omega)}$
$\gamma_2 R(t)$
$\alpha I(t-\tau)e^{-\mu\tau}$
$p\alpha I(t)$
R
$\mu R(t)$

Figure 1: The structure of compartments

We can obtain the following SEIRS model according to the modelling idea of the epidemic dynamic compartments.

$$\begin{cases} \frac{dS(t)}{dt} = bN(t) - \mu S(t) - \frac{\lambda Z_{av}S(t)I(t)}{N(t)} + \gamma_2 R(t) \\ \qquad -\alpha I(t-\tau)e^{-\mu\tau} + \gamma_1 E(t) \\ \frac{dE(t)}{dt} = \frac{\lambda Z_{av}S(t)I(t)}{N(t)} - \frac{\lambda Z_{av}S(t-\omega)I(t-\omega)e^{-\mu\omega}}{N(t-\omega)} \\ \qquad -\mu E(t) - \gamma_1 E(t) \\ \frac{dI(t)}{dt} = \frac{\lambda Z_{av}S(t-\omega)I(t-\omega)e^{-\mu\omega}}{N(t-\omega)} - \mu I(t) - \varepsilon I(t) - \alpha I(t) \\ \frac{dR(t)}{dt} = p\alpha I(t) - \alpha I(t-\tau)e^{-\mu\tau} - \gamma_2 R(t) - \mu R(t). \end{cases} \qquad (8)$$

From Equations (1) and (8), we can get

$$\frac{dN(t)}{dt} = (b-\mu)N(t) - [\varepsilon + (1-p)\alpha]I(t). \qquad (9)$$

For the continuity of the precondition, we require,

$$E(0) = \int_{-\omega}^{0} \frac{\lambda Z_{av}S(u)I(u)}{N(u)}e^{u\mu}du, \qquad (10)$$

and

$$R(0) = \int_{-\tau}^{0} p\alpha I(u)e^{u\mu}du. \qquad (11)$$

3.3 Model Analysis

We do some transforms because of the denominators containing variables. Define

$$s(t) = \frac{S(t)}{N(t)}, e(t) = \frac{E(t)}{N(t)}, i(t) = \frac{I(t)}{N(t)}, r(t) = \frac{R(t)}{N(t)},$$

then the Equation (8) can be expressed as Equation (12)

$$\begin{cases} \frac{ds(t)}{dt} = b - m(t)s(t) - \lambda Z_{av}s(t)i(t) + \gamma_1 e(t) \\ \qquad -\alpha i(t-\tau)exp(-\int_{t-\tau}^{t} m(q)dq) \\ \frac{de(t)}{dt} = \lambda Z_{av}s(t)i(t) - m(t)e(t) - \gamma_1 e(t) \\ \qquad -\lambda Z_{av}s(t-\omega)i(t-\omega)exp(-\int_{t-\omega}^{t} m(q)dq) \\ \frac{di(t)}{dt} = \lambda Z_{av}s(t-\omega)i(t-\omega)exp(-\int_{t-\omega}^{t} m(q)dq) \\ \qquad -m(t)i(t) - \varepsilon i(t) - \alpha i(t) \\ \frac{dr(t)}{dt} = p\alpha i(t) - p\alpha i(t-\tau)exp(-\int_{t-\tau}^{t} m(q)dq) \\ \qquad -\gamma_2(t) - m(t)r(t), \end{cases} \qquad (12)$$

where $m(t) = b - (b + (1-p)\alpha)\varepsilon i(t)$, and $1 = s(t) + e(t) + i(t) + r(t)$.

Let $S(t)$, $E(t)$, $I(t)$, $R(t)$ be the solution of Equation (8). The $s(t)$, $e(t)$, $i(t)$, $r(t)$ is the solution of (12) with

$$e(0) = \int_{-\omega}^{0} \lambda Z_{av}s(u)i(u)exp(-\int_{u}^{0} m(q)dq)du$$

and

$$r(0) = \int_{-\tau}^{0} p\alpha i(u)exp(-\int_{u}^{0} m(q)dq)du.$$

If $s(t)$ and $i(t)$ are positive on the initial interval, then $s(t)$ and $i(t)$ are positive for all finite $t \geq 0$. (Corollary 2.1 [5].)

It is easy to check that the region $Y = \{(s(t), e(t), i(t), r(t)) | s(t), e(t), i(t), r(t) \geq 0, s(t) + e(t) + i(t) + r(t) = 1\}$ is a positive invariant set of Equation (12). We consider the passive worm free equilibrium. When the infected fraction $i = 0$, then $e = r = 0$, and $s = 1$. This is the only equilibrium on the boundary of Y. According to reference [1], we have the following threshold for the existence of the interior equilibrium: $R_0 = (\lambda Z_{av}e^{-\mu\omega})/(b + \alpha + \varepsilon + \gamma_1 + \gamma_2)$. The threshold R0 is also called the basic reproduction number. When R_0 is smaller than 1, the equilibrium is globally stable, and the worm gradually disappears. When R_0 is larger than 1, the worm will exponentially propagate. The quantity $1/(b + \alpha + \varepsilon + \gamma_1 + \gamma_2)$ is the average waiting time in the infective class. For $(\lambda Z_{av}) < (b + \alpha + \varepsilon + \gamma_1 + \gamma_2)$, the solutions of Equation (12) approach the passive worm free equilibrium as $(t \longrightarrow \infty)$ [5].

4 Performance Evaluation

4.1 Simulation Model

The network used for simulations consists of 30,0000 peers. The network is growing using the methodology proposed by Holme [6]. This ensures that the peer degree follows a power law distribution and the network has a high clustering coefficient. The average peer degree and the exponent of power law of the network are 4.5 and 3.4, respectively, which are close to the real values of Gnutella network as measured in [3, 9].

1) Performance metrics: the system attack performance is defined as follows: the time taken t (X axis) to infected peer numbers (Y axis).

2) Evaluation systems: a tuple: $< TTL, s, m, p, \tau >$ is used to represent the configuration parameters. As we focus mainly on selected important parameters that are sensitive to the propagation of passive worms, the following parameters are set as constant values ($I(0) = 1$, $s = 4,000KB$,$bw = 132KBps$, $\varepsilon = 0.01$, $b = 0.04$, $\mu = 0.03$, $\alpha = 0.2$, $\lambda = 0.001$, $\gamma_1 = 0.005$, $\gamma_2 = 0.001$) in all simulations. The size of average files s and average bandwidth available bw is the same to the real P2P networks.

3) Evaluation method: we use numerical analysis of the differential equations by using Matlab Simulink to obtain performance data.

The basic reproduction number R_0 is 0.4436 ($TTL = 2$, $s = 4,000KB$) through the calculation. The passive worm will gradually disappear from the theory. Next we will validate the conclusion by the use of some experiments.

4.2 Performance Results

In this subsection, we report the performance results of propagation model along with observations.

Figure 2 shows the data on the performance sensitivity to different TTL. The general system is configured as $< *, 4000, 1, 0.4, 50 >$ and $TTL \in \{1, 2, 3\}$. From figure 2, we make the following observations: the number of hops (TTL) plays an important role in the propagation of passive worm. The larger TTL is, the more rapid the passive worm propagates. It is easy to understand that a peer may locate more peers during the search process, and the probability of finding an infected peer becomes larger with the increase of TTL. Thereafter, once a peer is infected, the passive worm will propagate rapidly. The tendency of the passive worm propagation in figure 2 is depressive, which is consistent with the theory analysis.

Figure 3 shows the data on the performance sensitivity to different file size. The general system is configured as $< 2, *, 1, 0.4, 50 >$ and $s \in \{2000, 4000, 8000\}$. From figure 3, we make the following observations: the sharing file size s has a larger impact on the passive worm. Along with the increase of s, the time of reaching its peak decreases. However, the number of infected peers is obviously small. As a result, sharing some big files (without chunks) as soon as possible is a simple and efficient method.

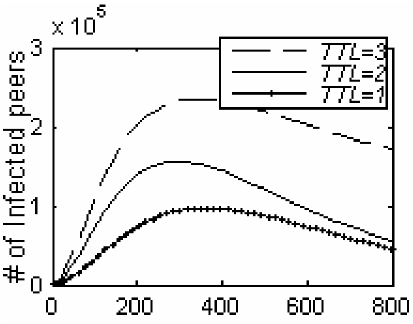

Figure 2: Effect of number of hops

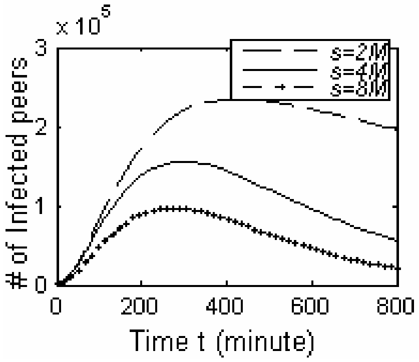

Figure 3: Effect of file size

Figure 4 shows the data on the performance sensitivity to different chunks m. The general system is configured as $< 2, *, *, 0.4, 50 >$, $s = 80,000KB$, and $m \in \{1, 5, 10, 20, 25\}$. From figure 4, we make the following observations: although the time of reaching the peak decreases, the peak has no obvious change ($m = 5, 10, 20, 25$). Unfortunately, when $m = 1$, the passive worm infects much larger peers than other values. Under the same assumptions, we can draw the conclusions from figure 2-4: a larger TTL, a smaller sharing file embedded passive worms and a larger number of chunks result in the increase of propagation speed.

Figure 5 shows the data on the performance sensitivity to temporary immunity probability p. The general system is configured as $< 2, 4000, 1, *, 50 >$, and $p \in \{0.1, 0.5, 0.8, 0.9\}$. Figure 5 shows the following observations: a large p results in the decrease of infected peers and much larger time to eradicate passive worms. This is mainly due to the fact that the number of recovered class becomes large with increase of p.

Figure 6 shows the data on the performance sensitivity to different temporary immunity period τ (minute). The general system is configured as $< 2, 4000, 1, 0.4, * >$ and

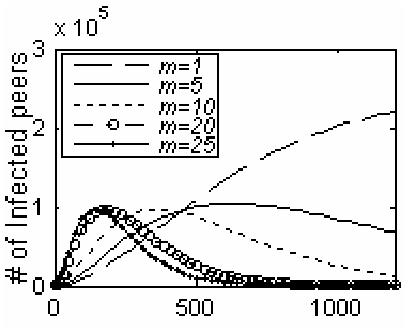

Figure 4: Impact of number of Chunks

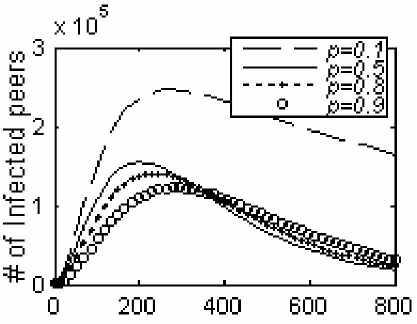

Figure 5: Impact of temporary immunity probability

$\tau \in \{20, 50, 80, 100\}$. From figure 6, we make the following observations: the τ plays an important role in both the number of infected peers and propagation speed. The larger τ is, the smaller the peak will be. This can win valuable time to defend against passive worms. For users, the method of improving temporary immunity period is to update the anti virus software in time. If all users can do this, the damages caused by passive worms will minimize.

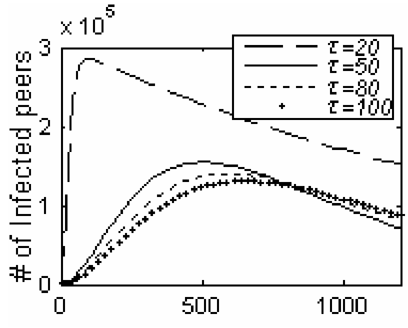

Figure 6: Impact of temporary immunity period

5 Conclusions

This paper proposes a delayed SEIRS model with online rate, off line rate and death rate by the use of the epidemic dynamics. Using the delayed SEIRS model, we simulate the propagation of passive worms, and obtain some valuable conclusions:

1) A larger (TTL), a smaller sharing file s embedded passive worms, and a larger number of chunks m result in the increase of propagation speed.

2) The temporary immunity probability p plays an important role in the propagation of passive worm. A large p results in the decrease of infected peers and much larger time to eradicate passive worms.

3) The temporary immunity period τ plays an important role in both the number of infected peers and propagation speed. The larger τ is, the smaller the peak will be.

The simulation results show the propagation of passive worms being mainly governed by the number of hops, the size and the number of chunks of sharing files, the temporary immunity probability, and the temporary immunity period. As a result, in order to effectively defend against passive worms, we must restrict the hops in configuring P2P systems, share large files as soon as possible, and update anti virus software in time. This can provide an important guideline in the control of unstructured P2P networks as well as passive worm defense.

In future work we will validate the model with simulations obtained by NS2 (Network Simulator version 2); develop a common platform of simulating passive worms in order to improve the simulating efficiency.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under No.60633020, the Natural Science Foundation of Hebei Province of China under No.F2006000177, Science Foundation of Ministry of Education of China under No.208012.

References

[1] K. M. Bimal, and K.S. Dinesh. "SEIRS epidemic model with delay for transmission of malicious objects in computer network," *Applied Mathematics and Computation*, vol. 188, no. 2, pp. 1476-1482, 2007.

[2] Y. W, C. Boyer, S. Chellappen, and D. Xuan. "Peer-to-peer system-based active worm attacks: modelling and analysis," *Proceedings of IEEE International Conference on Communications*, pp. 295-300, IEEE Press, Seoul Korea, 2005.

[3] G. L. Chen, S. G. Robert "Simulating non-scanning worms on peer-to-peer networks," *Proceedings of the 1st International Conference on Scalable Information Systems*, pp.29-42, ACM Press, Hong Kong, 2006.

[4] Z. S. Chen, L. Gao, and K. Kwiat. "Modelling the spread of active worms," *Proceeding of Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*, pp. 1890-1900, IEEE Press, San Franciso, CA, USA, 2003.

[5] K. Cooke, P. Driessche. "Analysis of an SEIRS epidemic model with two delays," *Journal of Mathematical Biology*, vol. 35, no. 2, pp. 240-260, 1996.

[6] P. Holme, B. J. Kim. "Growing scale-free networks with tunable clustering," *Physical Review E (Statistical Nonlinear, and Soft Matter Physics)*, vol. 64, no. 2, pp.1-4, 2002.

[7] C. Jacky, L. Kevin, and N. L. Brian. "Availability and popularity measurements of peer-to-peer file systems,". *Technical Report 04-36*, 2004.

[8] A. Kalafut, A. Acharya, and M. Gupta. "A study of malware in peer-to-peer networks," *Proceedings of the Sixth ACM SIGCOMM on Internet Measurement*, pp. 327-332, ACM Press, Rio de Janeriro, Brazil, 2006.

[9] P. G. Krishna, J. D. Richard, and S. Saroiu, S. D. Gribble, H. M. Levy, and J. Zahorjan. "Measurement, modelling, and analysis of a peer-to-peer file-sharing workload," *Proceedings of the 19th ACM Symposium on Operating System Principles*, pp. 314-329, ACM Press, Bolton Landing, NY, 2003.

[10] R. Krishna and S. Biplab. "Modelling malware propagation in Gnutella type peer-to-peer networks," *The Third International Workshop on Hot Topics in Peer-to-Peer Systems*, pp. 8-15, IEEE Press, Rhodes Island, Greece, 2006.

[11] E. K. Lua, J. Crowcroft, M. Pias, R. Sharma, and S. Lim. "A survey and comparison of peer-to-peer overlay network schemes," *Journal of. IEEE Communications Survey and Tutorial*, vol. 7, no.2, pp. 72-93, 2005.

[12] A. G. McKendrick. "Applications of mathematics to medical problems," *Proceedings of the Edinburgh Mathematical Society*, vol. 44, No.1, pp.98-130, 1926.

[13] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. "Random graphs with arbitrary degree distribution and their applications," *Phys.Rev.E. 64: 026118*, 2001.

[14] W. T. Richard, and J. C. Mark. "Modelling virus propagation in peer-to-peer networks," *IEEE International Conference on Information, Communications and Signal Processing (ICICS 2005)*, pp. 981-985, IEEE Press, Beijing, China, 2005.

[15] R. Thommes, and M. Coates. "Epidemiological modelling of peer-to-peer viruses and pollution," *The Twenty-fifth Annual IEEE Conference on Computer Communications (INFOCOM'06)*, pp. 15-26, IEEE Press, Barcelona, Spain, 2006.

[16] C. H. Xia , Y. P. Shi, and X. J. Li. "Research on epidemic models of P2P worm in structured peer-to-peer networks," *Chinese Journal of Computers*, vol. 29, no. 6, pp. 952-959, 2006.

[17] P. Yan, S. Q. Liu. "SEIR epidemic model with delay," *Journal of the Australian Mathematical Society, Series B - Applied Mathematics*, vol. 48, no.1, pp. 119-134, 2006.

[18] Y. Zhang, L. Lin, and J. Huai, "Balancing trust and incentive in peer-to-peer collaborative system," *International Journal of Network Security*, vol. 5, no. 1, pp. 73-81, 2007.

[19] Q. Zhang and K. L. Calvert, "A peer-based recovery scheme for group rekeying in secure multicast," *International Journal of Network Security*, vol. 6, no. 1, pp. 15-25, 2008.

[20] L. Zhou, L. Zhang, F. Mcsherry, N. Immorlica, and S. Chien. "A first look at peer-to-peer worms: threats and defenses," *Proc. of 4th Int. Workshop on Peer-to-Peer Systems (IPTPS'05)*, pp. 24-35, Springer press, Ithaca, NY, 2005.

[21] H. X. Zhou, Y. Y. Wen, and H. Zhao. "Passive worm propagation modelling and analysis," *International Multi-Conference on Computing in the Global Information Technology (ICCGI'07)*, pp. 32-35, IEEE Press, Guadeloupe, French Caribbean, 2007.

Fangwei Wang received his B.S. degree in 2000 from College of Mathematics & Information Sciences, Hebei Normal University, his M.S. degree in 2003 from College of Computer Science and Software, Hebei University of Technologies. Currently he is a Ph.D. candidate in the College of Computer at Xidian University. His research interests include: network and information security, sensor networks.

Yunkai Zhang received his B.S. degree in 1986 from Department of Electronic and Information Engineering, Hebei University, his M.S. degree in 1997 from Department of Telecommunication Engineering, Beijing University of Posts and Telecommunications, and his Ph.D degree in 2005 from College of Computer at Xidian University. Currently he is a professor at Hebei Normal University, Shijiazhuang, China. His research interests include network and information security.

Jianfeng Ma received his B.S. degree in mathematics from Shaanxi Normal University, China in 1985, and obtained his M.E. and Ph.D. degrees in computer software and communications engineering from Xidian University, China in 1988 and 1995, respectively. Currently he is a Professor at Xidian University, Xi'an, China. His research interests include distributed systems, wireless and mobile computing systems, computer networks, and information and network security.