J. Basic. Appl. Sci. Res., 2(10)10130-10136, 2012
© 2012, TextRoad Publication
ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

A Novel Dynamic Modeling Method of Passive Worms Propagation in Peer to Peer Networks

Nasser Khanjari[1], Ehsan Sharifi[2], M.A Azgomi[3]

[1]Islamic Azad University, Zanjan Branch, Department of Computer Engineering, Zanjan, Iran
[2]Islamic Azad University, South Tehran Branch, Department of Computer Engineering, Tehran, Iran
[3]Computer engineering department, Science and Technology University, Tehran, Iran

ABSTRACT

Propagation of peer to peer (P2P) Network's worms on the Internet Due to its complexity has become a serious challenge for network security researches. Considering the complexities of this issue, none of the existing works has been able to solve the problem of propagation modeling of P2P Network's worms. In addition, there is no comprehensive and integrated software for the worms' propagation modeling .software's such as Peersim, NS2, etc. are data-oriented, i.e. data is input and output of them. Here it is decided to present an improved dynamic model of passive worms' propagation. In this paper method of propagation and modeling of passive worms in P2P networks is studied .Three new models called DSI, DSIS and DSIR are presented based on three basic models of SI, SIS, and SIR, which are accurate and closer to the actual operating environment .

KEY WORDS: Network security, peer to peer (P2P) networks, passive worms, worms propagation modeling.

INTRODUCTION

Importance of network security have made researchers to work on Worms and their species for years [1].Worms are autonomous programs that spread automatically through computer networks by Searching, invading and infecting the remote computer [2].worm primary task is Scanning [3].P2P network worms cause common vulnerabilities on the network's member hosts and spread topologically in them .Potentially, there are very useful strategies for random scanning in order to find victims' place [4] .Peer to Peer network worms were widely introduced in IPT Ps2005 international conference and their great threats to network security were discussed in this assembly [4],[5].Passive worms are one of the main types of peer to peer worms attaching themselves to the shared folders and spread through downloading and running these files by other peers . Passive worms' modeling is efficient in understanding how various factors affect their propagation [5].Several models of passive worms' propagation are introduced in previous works [1],[3],[4],[6],[7],[8],[9],[10],[11].But major defect of them is lack of dynamism. paper is structured based on the sections that follow , in Section 2 Related works are reviewed. New models are presented in Section 3.In Section 4, simulation results with different parameters are expressed and they have been evaluated .conclusions are presented in Section 4.

Related Works

According to authors, peer to peer worms can spread secretly in the whole network and disable defense mechanism of most of them using scanner worms .Generally, mathematics' epidemic appearance is attributed to McKendrick [12]. Chain et al. considered a simulation framework based on work density to describe three types of non-scanner worms (passive, reactive and hyperactive worms), and identified parameters influencing their propagation. They proved that the sort of the worm spreading in such networks is not detectable with conventional methods [13]. Kalafut et al. represented that 28% of executable files transferred in a one-month period include passive worms [14].Xia et al. presented epidemic models of P2P worms in three structured and P2P networks in which they investigated the rapid spread capability of worms and the negative effect of network topology on the worms' propagation [9].Thommas et al. used an analytical model to estimate the effect of identification solution for propagation of P2P worms [8]. Chaosheng Feng and Zhiguang Qin presented three basic propagation models of SI, SIS and SIR [15]but did not consider the effect of bandwidth and the number of infected files which causes the simulation results not to be as same as actual operating environment. The proposed model is derived from three basic models of SI, SIS, and SIR which considers the effect of bandwidth, infected and non-infected files. In addition, its most important feature is dynamism and the parameters can be varied during propagation using Peersim simulator software.

***Corresponding Author:** Nasser Khanjari, Computer engineering department, Azad university, Zanjan , Iran

1. Passive Worm's Propagation Models

Precisely investigating SI, SIS and SIR models, this is concluded that however these models have good efficiency and well modeling of propagation, but some of assumptions included in them do not correspond to reality [15]. They also have deficiencies, which are expressed in the following.

Fixed network size and constant files number: network size has been considered to be fixed in these models and all of previous ones, i.e. the numbers of online peers are changeless. In this case, no peer is added or removed and no new file is added.

Neglecting the file size: File size will have much effect on the increasing or decreasing of propagation rate. Since it is clear that as it becomes larger, download speed and consequently propagation rate reduces. Currently, the average file size on network is 4 MB [16].

Neglecting the file type: whether the downloaded file is infected or not affects the propagation process.

Another influential factor is the network bandwidth being not considered in this models. Currently, the average bandwidth is 132 KB per second [17].

Now, according to above mentioned expressions, proposed model assumptions are as follow.

2.1 Proposed Model assumptions

Assume that each user puts the files in his sharing folders which can be downloaded by others and other users put their downloaded files in sharing folders. Online peers refer to pairs which are running.

The number of online peers is dynamic. In this case, arbitrary number of infected or non-infected peers can enter the network during the simulation.

The number of files is dynamic. In this case, arbitrary number of infected or non-infected files can enter the network during the simulation.

A file will be implemented immediately after downloading.

The time spent for searching, connecting, downloading and executing a file is constant called unit time. A unit time lasts when an uploaded pair is damaged or remained secure.

When a pair is infected, "c" infected files are established in shared folders and take "c" different names. All infected pairs share the same "c" infected files.

Network Bandwidth is dynamic, i.e. it can be changed during simulation.

To analyze passive worms' modeling, the parameters involved in propagation modeling are listed in Table I

TableI : Parameters of DSI, DSIS and DSIR Models

N(t)	Total number of peers in p2p networkas a function of twhich is consideredvariable.
S(t)	The number of susceptible peers in time t.
I(t)	The number of infected peers in time t.
R(t)	The number ofsecured or retrieved peers in time t.
K(T)	The number of infected files in time t.
M(t)	The number of non-infected files (healthy) in time t.
h(t)	The possibility of downloading infected files in time t. $\frac{k(t)}{M(t)+k(t)}$H(t) = α
S_i	The average size of infected files.
S_{ui}	The average size of non-infected files (healthy).
bw	The average bandwidth of all peers considered fixed here.
λ_{di}	The average download rate of infected file by any peer (this includesthe time spent for searching, creating,connecting to another peer and performing downlosd files). $\lambda_{di} = \frac{bw}{S_i}$
λ_{dui}	The average download rate of non-infected file by any peer (this includesthe time spent for searching, creating,connecting to another peer and performing downlosd files). $\lambda_{dui} = \frac{bw}{S_{ui}}$
λ_{ir}	The average rate of host in every time unit, in which the infected host returns to the susceptible one.
λ_{sr}	The average rate of removing in every time unit, in which the predisposed host becomes immune.
λ_{ir}	The average rate of removing in every time unit, in which theinfected host becomes immune

2.2 DSI propagation model

In this model, the peers' state in P2P network is classified into two categories .susceptible and infected ones .Susceptible pairs share no infected file but they are exposed to the risk of infected files downloading .When a pair downloads an infected file, it becomes infected immediately. Upon execution, the infected files will be put in the sharing folder of pairs. Advancement mode for all peers in this model is shown in Figure 1.

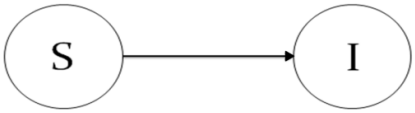

Figure 1 : Advancement mode in DSI model

Way of Calculation of h(t) is expressed in SI model {15}.As it is known , downloading rate of infected files depends on to the network bandwidth and average number of infected files, i.e. the smaller the file size and the larger the bandwidth, the more the download rate of file, and vice versa, so we can express this relationship as following equation :

$$\lambda_{di} = \frac{bw}{S_i} \qquad (1)$$

Where bw is each peer's average bandwidth and S_i is the average size of infected files; S and Irate will change because the peers' infecting rate depends on the downloading rate of infected files and also the probability of downloading infected file. So the infection probability of a susceptible peer as a function of time will be equal to $\lambda_{di} I_r(t)$.It is evident that the rate of I will be opposite to S since the number of susceptible peers will be reduced by the increasing of infected peers.

Above expressions is applicable to the non-infected files, i.e:

$$\lambda_{dui} = \frac{bw}{S_{ui}} \qquad (2)$$

Hence, equations of this model are as follows:

$$\frac{dS(t)}{dt} = -\lambda_{di} h(t) S(t) \qquad (3)$$

$$\frac{dI(t)}{dt} = \lambda_{di} h(t) S(t) \qquad (4)$$

$$\frac{dK(t)}{dt} = \lambda_{di} h(t) S(t) \times c \qquad (5)$$

$$\frac{dM(t)}{dt} = \lambda_{dui} N(1-h(t)) \qquad (6)$$

Where $N(t) = S(t) + I(t)$.

2.3 DSIS propagation model

In this model, state variation of peers can be depicted as Fig.2.

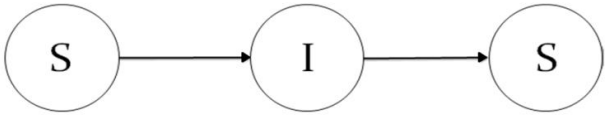

Figure 2 : Advancement mode in DSIS model

When all the infected files of a peer die naturally or be removed(cleaned) peer returns to its susceptible state. Assume that λ_{is} is the average return rate of infected peers to susceptible state .Thus $\lambda_{is} S(t)$ number of infected peers return to susceptible condition .According to the DSI model analyzation , relative equations of this model are as follows:

$$\frac{dS(t)}{dt} = -\lambda_{di} h(t) S(t) + \lambda_{is} I(t) \quad (7)$$

$$\frac{dI(t)}{dt} = \lambda_{di} h(t) S(t) - \lambda_{is} I(t) \quad (8)$$

$$\frac{dK(t)}{dt} = \lambda_{di} h(t) S(t) \times c - c \times \lambda_{is} I(t) \quad (9)$$

$$\frac{dM(t)}{dt} = \lambda_{dui} N(1 - h(t)) \quad (10)$$

Where $N(t) = S(t) + I(t)$.

2.4 DSIS propagation model

The peers' advancement of this model is shown in figure 3.

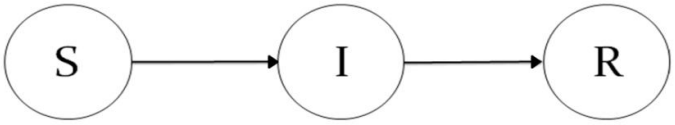

Figure 3 : Advancement mode in the DSIS model

Unlike DSIS model, In this model infected peers retrieve to the immune state rather than returning to the susceptible one. This means that all infected files on the peer are removed and the peer cannot be polluted no longer. Assuming that the susceptible and infected peers change with $\lambda_{sr}\ S(t) + \lambda_{ir}\ I(t)$ rate,the infected files will be reduced with the rate of $c\ \lambda_{ir}\ I(t)$ in the unit time;So new equations can be defined as the following:

$$\frac{dS(t)}{dt} = -\lambda_{di} h(t) S(t) - \lambda_{sr} S(t) \quad (11)$$

$$\frac{dI(t)}{dt} = \lambda_{di} h(t) S(t) - \lambda_{sr} I(t) \quad (12)$$

$$\frac{dR(t)}{dt} = \lambda_{sr} S(t) - \lambda_{ir} I(t) \quad (13)$$

$$\frac{dK(t)}{dt} = \lambda_{di} h(t) S(t) \times c - c \times \lambda_{ir} I(t) \quad (14)$$

$$\frac{dM(t)}{dt} = \lambda_{dui} N(1 - h(t)) \quad (15)$$

Where $N(t) = S(t) + I(t) + R(t)$.

Since number of secured peers is more than those coming back to the vulnerable state,the ratio of vulnerable peers won't be considered. Numerical values simulation of equations is performed In order to investigate the accuracy of them . Hence, Peersim Simulator software is used according to its precisely executing simulation and output data are plotted in MATLAB software.

SIMULATION AND RESULTS

In this section, it is investigated that how passive worms' propagation with different parameters under different conditionsoccurs .initial values of parameters are listed in table II:

TableII : initial values of some parameters used in simulation

S(0)	I(0)	S_i	S_{ui}	bw	λ_{sr}	λ_{ir}	λ_{is}	c
10000	10	100	1000	32	0.001	0.002	0.001	**10**

The average file size should be considered small to be able to spread more quickly in networks.

3.1 Models' Evaluation

Figure 4 shows the infected peers' number versus time unit with the default values of parameters.

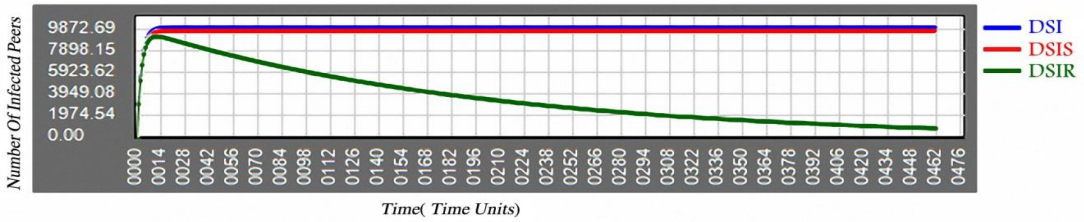

Figure 4 : infected peers' number in terms of time unit

As shown in the Fig.5, propagation rate is very high in DSI model and all peers are infected in time unit 40.This result is not comparable with the same model SI. This is due to sudden rate change, the effect of the average file size and bandwidth. when the average size of infected files are small, They can be downloaded quickly and consequently spreaded rapidly. In addition, there is an interesting point in DSIS model simulation, that is 9988 peers are infected in time unit 25 and other peers are in the susceptible state thereafter. on the other hand , there are no more infected peers to enter to the susceptible state, and this is due to effect of λ_{is}parameter. DSIR model has the best performance such that in time unit 460, the number of infected peers are reduced to 1000 showing relatively better performance compared to the SIR model. Other parameters of this model are investigated in following parts.

3.2 The average size of the infected file parameter(S_i)

According toFig.4, as S_i gets smaller, propagation speed becomes faster .because according to equation (1), λ_{di} increases while$S_i$decreases, thus infection rate increases according to equation (4).Now effect of changing this parameter to 4000 is investigated. It is expected that the propagation rate decrease. Figure 5illustrates the effect of this parameter.

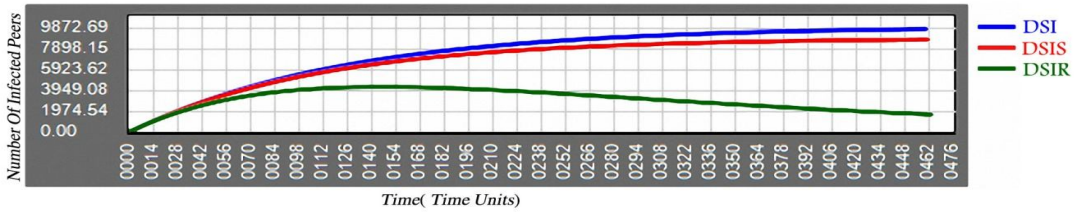

Figure 5 : The effect of the infected files' size parameteron number of infected peers

As expected, the propagation rate decreases by $S_i$increasing .Simulation depicted that 9761 peers became infected in time unit 460using DSI model. This rate was 8766 and 1682 when using DSIS and DSIR models, respectively. It is notable that in SIR model, the number of infected peers in the highest case was 4282 in time unit 145 representing better performance compared to the same DSI model.

3.3 Bandwidth Parameter (bw)

The effect of this parameter is investigated in two modes. At first, bw set to 16 and it was 256 at second mode. figure6 indicates the results with each bandwidth .In addition, the value of $S_i$was set to 1000to balance establishing of both modes .

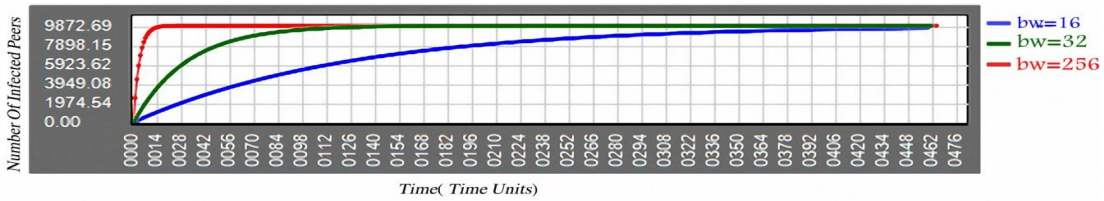

Figure 6 : The effect of bw parameter on number of infected peers

As Fig.6depicts , the speed of the infection propagation increases with the increase of bandwidth. Simulations indicated that all peers of the network became infected with the bandwidth of 256 in time unit 50, which is too

close to the effect of the parameter of the average size of the infected file. Also , with the bandwidth of 16 in time unit 460, the number of infected peers reached to 10004 representing the reduction of the propagation rate .In the normal state , i.e. the bandwidth of 32 in time unit 460, 10009 peers became infected.

3.4 Comparison of SI and DSI models

In this section, existing SI model and the proposed DSI model are compared .Figure 7 shows an example of two models' execution with default values .

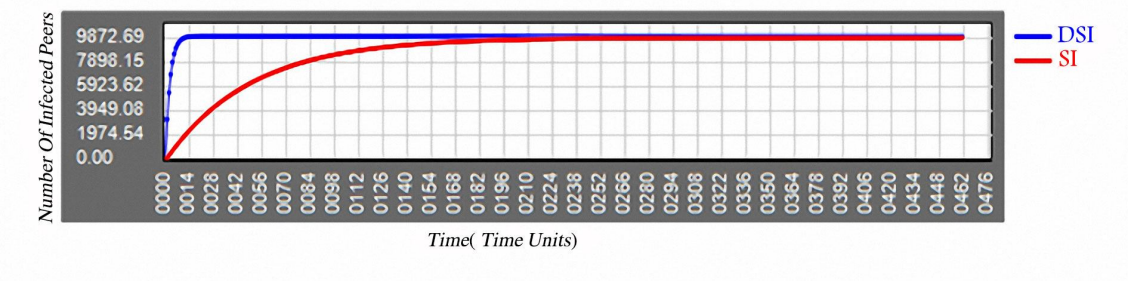

Figure7 : Comparison of SI and DSImodel based on numbers of infected peers

According to Fig.7 accurate simulation close to real cannot be performed due to not considering infected file size and bandwidth in SI model . beside it, in time unit 460 all of the peers (10010) are infected . While in DSI model, it is occurred in time unit 40 representing higher precision and realism of this model.

3.5 Comparison of SIS and DSIS models

In this section, existing SIS model and the proposed DSIS model are compared .Figure 8 shows an example of two models' execution with default values.

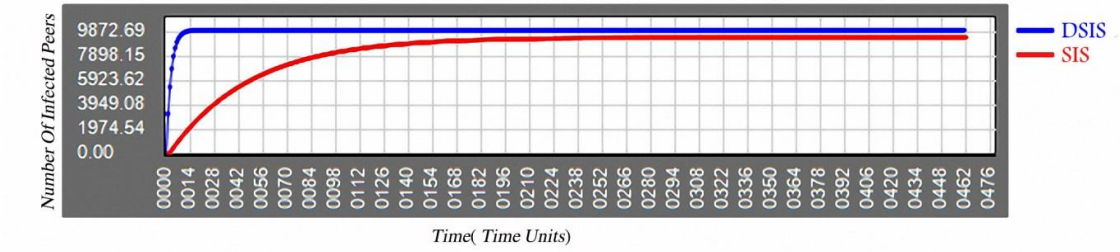

Figure8 : comparison of SIS and DSIS model based on numbers of infected peers

Simulation results showed that 9988 peers have become infected in DSIS model in time unit 25, and afterwards there is no infected peer to be converted to susceptible state. The number of infected peers reached 9542 in SIS model in time unit 476 and remained on this value. This means that there were no more infected peers to be converted to susceptible state. This discrepancy is due to the separation of the infected file from non-infected one in proposed model.

3.6 Comparison of SIR and DSIR models

In this part, SIR model and the proposed DSIR model are compared .Figure 9 shows an example of two models' execution with default values.

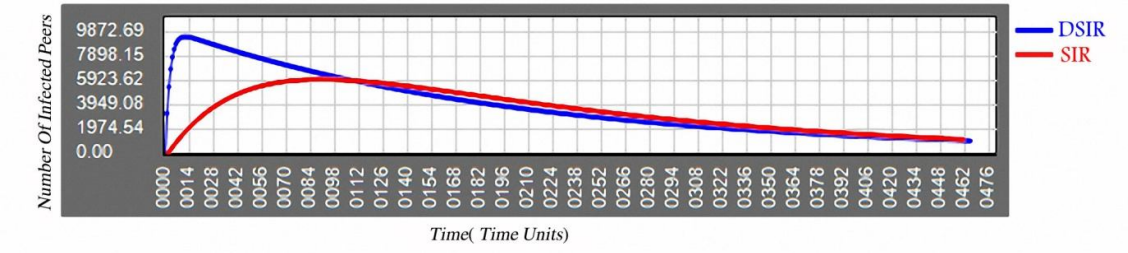

Figure 9 : comparison of SIR and DSIR model based on numbers of infected peers

Simulation results expressed that the number of infected peers reached 1246 in SIR model in time unit 460.also, in SIR model, the highest infection is related to the time unit 89 in which 6117 peers have become

infected. The number of infected peers reduced to 1000 in time unit 460in DSIR model, and the highest infection is related to the time unit 11 in which 9449 peers are infected.

The summary of deduced results can be expressed as table III .

TableIII : The results obtained from passive worms' propagation modelling

Propagation process	operation
Increasing propagation rate	Adding infected peer during propagation
decreasing propagation rate	Adding non-infected peer during propagation
-	Adding infected file during propagation
decreasing propagation speed	Adding non-infected file during propagation
The smaller the size of the infected file, the more the speed of propagation	The size of the infected file
The larger the network bandwith, the more the speed of propagation	Network bandwidth
The more The number of non-infected files, the less the speed of propagation	The number of non-infected files

2. Conclusion

In this paper, we have proposed three new propagation models called DSI, DSIS and DSIR. the performance of each one was closer to the actual operating ,because the important parameters such as network bandwidth and the average size of the infected and non-infected files are considered. important point in of simulations is that the size of infected files is considered 100 KB which is close to the same value in the real world and this enhanced the propagation rate .

REFERENCES

[1] X. Fan and Y. Xiang, "Propagation modeling of peer-to-peer worms," 2010, pp. 1128-1135.

[2] C. C. Zou, W. Gong, and D. Towsley, "Worm propagation modeling and analysis under dynamic quarantine defense," 2003, pp. 51-60.

[3] Y. Yu, L. Xingrui, G. Fuxiang, and Y. Ge, "A potential approach of internet worm propagation based on P2P," *Wuhan University Journal of Natural Sciences*, vol. 11, pp. 1711-1714, 2006.

[4] L. Zhou, L. Zhang, F. McSherry, N. Immorlica, M. Costa, and S. Chien, "A first look at peer-to-peer worms: Threats and defenses," *Peer-to-Peer Systems IV*, pp. 24-35, 2005.

[5] F. Wang, Y. Zhang, and J. Ma, "Modelling and Analyzing Passive Worms over Unstructured Peer-to-Peer Networks," *International Journal of Network Security*, vol. 11, pp. 39-45, 2010.

[6] Z. Chen, L. Gao, and K. Kwiat, "Modeling the spread of active worms," 2003, pp. 1890-1900 vol. 3.

[7] R. Thommes and M. Coates, "Modeling virus propagation in peer-to-peer networks," 2005, pp. 981-985.

[8] R. Thommes and M. Coates, "Epidemiological modeling of peer-to-peer viruses and pollution," 2006, pp. 181-192.

[9] X. I. A. C. H. S. H. I. Yun-Ping and L. Xiao-Jian, "Research on Epidemic Models of P2P Worm in Structured Peer-to-Peer Networks [J]," *Chinese Journal of Computers*, vol. 6, 2006.

[10] Z. Hanxun, W. Yingyou, and Z. Hong, "Passive Worm Propagation Modeling and Analysis," in *Computing in the Global Information Technology, 2007. ICCGI 2007. International Multi-Conference on*, 2007, pp. 32-32.

[11] P. Yan and S. Liu, "SEIR epidemic model with delay," *Anziam Journal*, vol. 48, pp. 119-134, 2006.

[12] A. M'Kendrick, "Applications of mathematics to medical problems," *Proceedings of the Edinburgh Mathematical Society*, vol. 44, pp. 98-130, 1925.

[13] G. Chen and R. S. Gray, "Simulating non-scanning worms on peer-to-peer networks," 2006, p. 29.

[14] A. Kalafut, A. Acharya, and M. Gupta, "A study of malware in peer-to-peer networks," 2006, pp. 327-332.

[15] C. Feng, Z. Qin, L. Cuthbet, and L. Tokarchuk, "Propagation modeling of passive worms in P2P networks," 2008, pp. 1027-1031.

[16] J. C. Chu, K. S. Labonte, and B. N. Levine, "Availability and locality measurements of peer-to-peer file systems," 2002, p. 310.

[17] K. P. G. e. al., "Measurement, Modeling, and Analysis of a Peer-to-Peer File Sharing Workload," in *SOSP*, Bolton Landing, New York, USA, 2003, pp. 19–22.