PROACTIVE WORM PREVENTION BASED ON P2P NETWORKS

KAI-GUI WU, YONG FENG

College of Computer Science and Technology, BJ Industry University, Beijing , 100011,China

Abstract:

According to the features of worm propagation, a novel model for proactive worm prevention based on peer-to-peer (P2P) networking technologies is proposed in this paper. At first, we present the backgrounds and security issues related to P2P networks. Then, a structured P2P overlay network designed for worm prevention is specified. Based on that, we present the methods for Internet worm forecasting which is assumed to be unknown. Furthermore, we study effective strategies to restrain and eliminate worms under the model. We find that our proposed P2P-based worm prevention model can efficiently reduce the impact of worms on the networks.

Keywords:

Worm prevention; P2P overlay network; Internet worm forecasting

1. Introduction

Many People believe that Internet worm is a puzzle which has disturbed us for so many years. Although IDS has been widely used and operating systems are upgraded continually, Internet worms burst out occasionally and have rather increased than shown any symptoms of abating. The term worm was first described as follows: A worm is a program that can run by itself and can propagate a fully working version of itself to other machines. It is derived from the word tapeworm, a parasitic organism that lives inside a host and saps its resources to maintain itself [1]. Worms was used to be considered as monkeyshines made by some fastuous programmers. In recent years, as the technologies of worms become more and more sophisticated and the damage caused by worms (example - Code red, Nimda and Blaster) is greatly enhanced, much attention is given to this problem. Technology advances that combat worms, such as firewalls, Intrusion detection systems, and honey pots, continue to appear [2], [3]. However, these traditional ways to detect worms are not as effective as expectation in forecasting and preventing unknown Internet worms.

Nowadays, Internet worms are much more intelligent than before. New worms attempt to breach an enclave's defenses by learning which internal IPs they can reach and what ports and services are available for attack. They scan not only randomly selected external addresses on the Internet, but also internal address space, virtually ensuring that if there is one infection within a local network there will soon be multiple ones. Additionally, today's worms have achieved considerable success by doing something that previous generations of these types of malware generally did not do—they deceive users into thinking that messages generated and sent by these worms are from someone they know [4]. It's almost certain that ordinary users' systems will be infected by some kinds of worms if they are connected to the network in spite of firewalls or virus walls. Thereby, we should not regard the worms as common viruses that can be easily controlled and perished.

According to the features of worm propagation, we will present a new model for worm forecasting and preventing based on peer-to-peer networking technologies in this paper. P2P networking technologies have gained popularity as a mechanism for users to share files without the need for centralized servers. A P2P network provides a scalable and fault-tolerant mechanism to locate nodes anywhere on a network without maintaining a large amount of routing state. This allows for a variety of applications beyond simple file sharing. In this worm forecasting and preventing model, a new application of P2P networking technologies is presented. As worms in an infectious computer tend toward infecting adjacent nodes, the peers whose original addresses is contiguous, especially those in the same LAN, will constitute a peer group. A core node is in charge of the analysis of suspicious information exchanged in the group and all these core nodes will be integrated to form a high-level P2P network. In this way, the distributed peers are joined together to obtain a global overview of the network status.

The rest of paper is organized as follows: In Section 2, we present the backgrounds and security issues related to P2P networks. In Section 3, a structured P2P network designed for worm prevention is specified. Based on that, we present the methods to forecast the Internet worm which is assumed to be unknown. In Section 4, we study effective strategies to restrain and eliminate worms under the model. Finally, the performance results are given.

2. P2P networks and security issues

2.1. Backgrounds

Peer-to-peer networks begin with Napster, Gnutella, and several other related systems, and has become immensely popular in the past few years. Researchers have defined structured P2P overlays such as CAN [5], Chord [6], Pastry [7] and Tapestry [8] which give a self-organizing substrate for large-scale P2P applications. Rather than being designed specifically for the purpose of file sharing, these systems serves as a powerful platform for the construction of a variety of decentralized services, including network storage, content distribution, web caching, searching and indexing, and application-level multicast.

The reason why we should use P2P networking technologies to withstand the attack of Internet worms is due to the features of worm propagation. Typical worm propagation can be described by epidemic models, such as Simple Epidemic Model [9], Kermack-Mckendrick Model [10], SIS (Susceptible Infectious Susceptible) Model [11], Two-Factor Model [12]. Although these models analyze worms' behavior from different points of view, at least one consentaneous conclusion can be drawn—in the initial propagation phase of a worm, exponential increase in infections is observed. Unfortunately, most impact and losses are caused by worms in this period of propagation. Traditional ways based on firewalls, IDS, honey pots and so on, are limited to one host or a small group of hosts, while worms' propagation can cover a wide range of networks in short time. Therefore, a distributed P2P network is required in preventing Internet worms proactively.

2.2. Secure the P2P networks

Since P2P system considers that network peers have certain intelligence and are able to contribute their local resource to the system, the worm can easily propagate to the system in a short time and degrade P2P network services significantly. The P2P system per se is liable to the worm-based attack in itself. Hence, before a P2P system is employed in worm prevention, we should take its security into account.

A novel mathematical model for worm propagation in the P2P system is proposed as formula (1) [13].

$$X(t) = \frac{X_0(1-\rho)}{X_0(1-\rho-X_0)e^{(U+P-S)Dt}} \tag{1}$$

Where X_0 denotes the initial infected hosts, U denotes the rate at which infection hosts are detected and eliminated without the software patching and P denotes the rate at which an infected or vulnerable host become invulnerable due to the software patching. D is the number of average network neighbors which the infected host can initiate the worm propagation directly, and X is the ratio of infected hosts in the whole system. It's clear that we should improve the defense capability $U + P$.

To secure networks against worm attack based on P2P technologies, approaches are presented as follows:

- Control and isolate the infected hosts. When the worm attack has been detected in the network, several measures will be taken to slow down the spreading velocity of the worm and the isolated regions will be constructed.
- Distribute the worm patching to throttle the worm. As an inherent superiority, the distribution of the worm patches can be easily implemented through P2P overlay networks once the patches are developed.

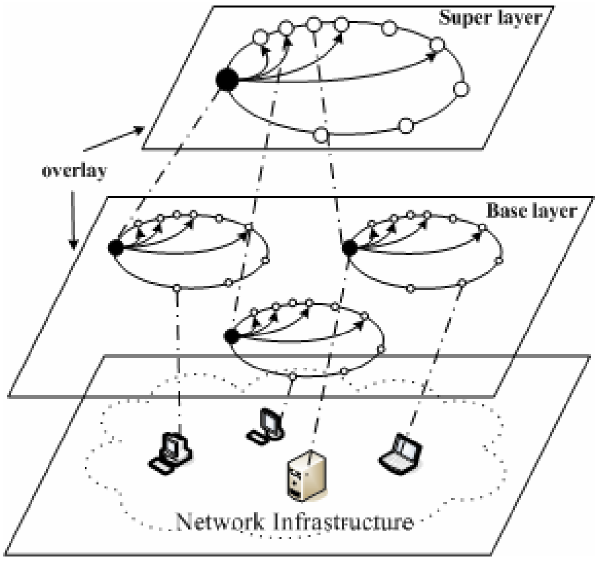
In the following sections, we are referring to each of approaches respectively.

3. Forecast worms

To minimize the impact and losses caused by unknown worms, we should sniff abnormal behaviors in the network and identify the worms rapidly at first. In this section, we will introduce a structured P2P overlay network, on which approaches to forecast and control the worms are described.

3.1. P2P overlay network for worm prevention

As worms in an infectious computer tend toward infecting adjacent nodes, the peers whose original addresses is contiguous, especially those in the same LAN, will constitute a peer group. A core node is in charge of the analysis of suspicious information exchanged in the group and all these core nodes will be integrated to form a high-level P2P network. In this way, the distributed peers are joined together to obtain a global overview of the network status. Figure 1 illustrates the architecture of the proposed P2P overlay network.



As we can see from Figure 1, the architecture is constructed by two levels: base level and super level. In the base level, the peers are grouped according to adjacency degree. To simplify the scheme, we use IP address as the measurement of adjacency degree and seed of peer's serial number. Each group can be regarded as a Chord except that the peer's serial numbers are not generated by SHA-1. Since we have to ensure that the nearest nodes are kept adjacent in the overlays, the peer's serial numbers are coded according to the sort order of its IP address. The most stabile and powerful peer in a group play the roll of super peer which detect the worm-like behaviors and exchange information with other groups. The super level is composed of these super peers with the same structure as groups in the base level.

3.2. Worm alarm

Worm early detection techniques have been researched for many years and several ingenious methods were presented by researchers. For example, Wu proposed a victim counter-based detection algorithm that tracks the increased rate of new infected hosts [14]. Worm alerts are output when anomaly events occur consecutively over a certain number of times. Kalman filter-based detection algorithm proposed by Zou in [15] detects the trend of illegitimate scans to a large unused IP space. Berk proposed to use ICMP "Destination Unreachable" messages collected at border routers to infer worm activities [16].

Our work based on P2P overlay network is different from other related work in that we take advantage of the distributed feature of overlays to make a conjunct analysis of dubious network traffic. If the adjacent hosts show the same abnormal behaviors, there is great probability of worm propagation. We can take account of the similarity of the abnormal behaviors between adjacent hosts and set a threshold to judge whether a worm is propagating or not.

Since this model don't rely on a specific detection method, we can abstract the abnormal behaviors detected in a peer by using a vector $\left(B_{P_m}[1], B_{P_m}[2], B_{P_m}[3], \ldots B_{P_m}[n]\right)$, whose elements denote a certain parameter of the system, such as the increase rate of source addresses.

The similarity of the vectors can be gained by calculating the weighted Euclidean distance as formula (2).

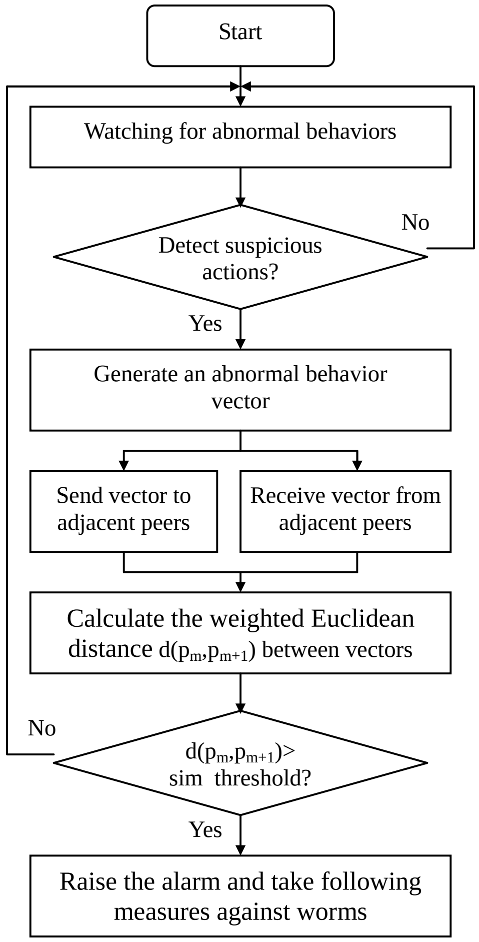
$$d(Pm, Pm+1) = \sqrt{w_1 \left|B_{P_m}[1] - B_{P_{m+1}}[1]\right|^2 + \cdots + w_n \left|B_{P_m}[n] - B_{P_{m+1}}[n]\right|^2} \quad (2)$$

Figure 2 shows the algorithm for worm detection in a peer. When an Internet worm is propagating in the overlay network, it will be quickly detected in this mechanism. Thereby, we can take actions to control the worm propagation as soon as possible.

Additionally, by the collaboration of all peers in the overlay, we can get a global view to the worm propagation.

4. Restrain and eliminate Internet worms

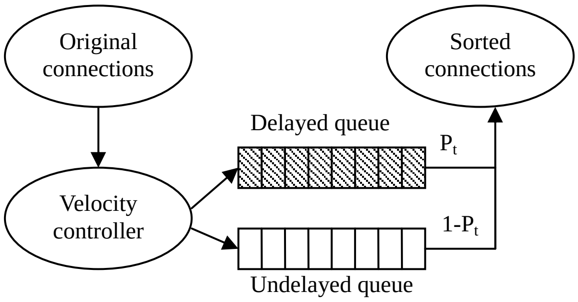
There are several traditional measures could be taken to withstand the worm attack, such as filtering and isolating. However, since the worm is cunning in disguising its true intentions and the deletion of peers would impact the structure of the P2P overlays, we have to find a suitable method for worm prevention in this architecture. Williamson describes a novel approach to the network security problem [17]. The situation can be significantly improved by using "benign" responses, those that slow but do not stop the virus. The main idea is to delay the worm by so long as to earn time for patching. Feedback control strategy is desirable in such systems because well-established techniques exist to handle and control such a system.


Start
Watching for abnormal behaviors
Detect suspicious actions?
No
Yes
Generate an abnormal behavior vector
Send vector to adjacent peers
Receive vector from adjacent peers
Calculate the weighted Euclidean distance d(pm,pm+1) between vectors
No
d(pm,pm+1)> sim threshold?
Yes
Raise the alarm and take following measures against worms

This technique is based on the fact that an infected machine tries to make connections at a faster rate than the machine that is not infected. The idea is to implement a controller, which restricts the rate at which a computer makes connection to other machines. The delay introduced by such an approach for normal traffic is very low (0.5 –1 Hz). This rate can severely restrict the spread of high-speed worm spreading at rates of at least 200 Hz.

We introduce this technique into our P2P overlays and make it more suitable for worm controlling. It is assumed that, if the number of requests increases too quickly, a portion of them will be sent to a delay queue to be served later. Parameters related to the size of the delay queue and the number of dropped connections are used to control the total number of connections resulting in a slow down of spreading worm.

The goal of our work is to slow down the spreading velocity of a worm by controlling the paroxysmal connections detected by the host. To achieve the goal, we propose a connection controlling model, which is shown on Figure 3.



In this model, original connections are monitored by a velocity controller which separates the connections into two queues—delayed queue and undelayed queue. Velocity controller identifies the abnormal connections by watching distinctly increasing connections, especially those made by the same procedure. Thus, no matter how quickly worms make connections, all procedures in the system will have the same probability to establish connections at the cost of slow-down of the whole system. The two queues are selected to be served in different probability. A parameter P_t is used here to denote the probability for delayed queue, while $(1-P_t)$ denote the probability for undelayed queue. We can adjust P_t to gain a good result in applications.

When the worm alarm is raised, the restrain process starts up. It's unavoidable that the whole system is degraded. However, this process will have more impact on worm than normal procedures in the system. In the following section, we will prove that in a mathematic way.

At first, we should calculate the whole connection requests in the system during a period of time, which is denoted by N_{all}.

$$N_{all} = (P_C \cdot R_p + W_C \cdot R_W) \cdot t \qquad (3)$$

Where P_c is the number of normal procedures, R_p is the average rate of connections made by normal procedures, W_c is the number of worms, R_w is the average rate of connections made by worms, and time is denoted by t. In general, R_w is far more larger than R_p. After the work of velocity controller, the whole connection requests N_{all} are separated into two portions. The portion in delayed queue is denoted by $N_{delayed}$, the other is denoted by $N_{undelayed}$. They are calculated respectively in formula (4) and (5).

Where δ_e denote the error factor in the process of separation. Since the connection requests are inserted into two queues, they will be served in different probability. As we have mentioned, P_t is the probability for delayed queue to be served. Finally, we can give the average rate of connections made by normal procedures after controlling which is denoted by $R_p{}'$. Also, the average rate of connections made by worm after controlling is calculated and denoted by $R_w{}'$.

$$R_p{}' = \frac{P_t \cdot N_{delayed} \big|_{P_C} + (1-P_t) \cdot N_{undelayed} \big|_{P_C}}{P_c \cdot t} \qquad (6)$$
$$= R_p \cdot (1 - \delta_e - P_t + 2 \cdot \delta_e \cdot P_t)$$

$$R_w{}' = \frac{P_t \cdot N_{delayed} \big|_{W_C} + (1-P_t) \cdot N_{undelayed} \big|_{W_C}}{W_c \cdot t} \qquad (7)$$
$$= R_w \cdot (\delta_e + P_t - 2 \cdot \delta_e \cdot P_t)$$

Generally speaking, δ_e and P_t are usually very small parameters. As we can see from formula (6) and (7), $R_p{}'$ is reduced for a little, while $R_W{}'$ is restrained deeply.

Although the worms are restrained, the ultimate solution is patching the peers. Fortunately, the propagation of worms is slowed down so that we have time to make patches available before worms have penetrated the entire network. There are many efficient ways to distribute the patches. In our proposed P2P overlay network, the distribution of patches will be implemented based on the file systems over P2P structures which have been studied by a good many researchers and have many credible schemes. The patch time will be limited to $O(\log kN)$, where N is the number of clients, and K is the fan-out factor of the peer.

5. Simulations

To study the propagation of worm through the proposed P2P overlay network, we developed a single machine simulation on worm propagation where the application of our scheme against worms could be studied. The main advantages of such an approach over using a real network are ease of configuration and low cost. The simulation is carefully designed in order to accurately model real-world behavior.

In addition, a worm model which has the general features of modern worms is introduced into the simulation. Our purpose is to validate our system and the results of our analytical models in an Internet-like setting.

The overall simulation works as follows. The network consists of a fixed set of fully interconnected nodes (any routers that may connect nodes are ignored). Each node has several procedures which request connections randomly and is modeled as a peer in the proposed P2P overlay network.

The performance of our proposed model in preventing the worm propagation is shown in Figure 4.

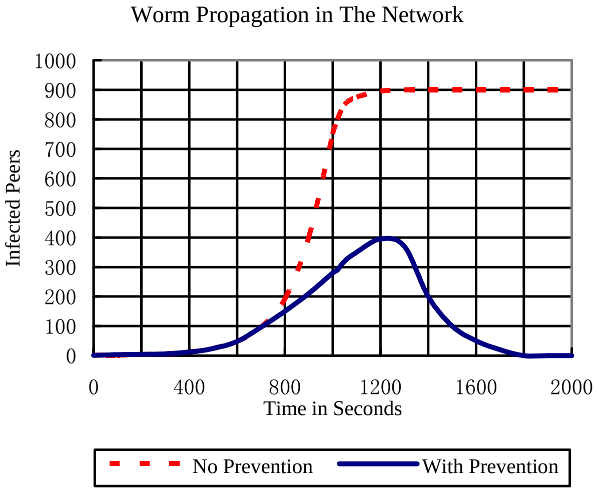


Figure 4. The Performance of Prevention Model

6. Conclusions

In the paper, we introduced a systematic approach for worm detection based on P2P overlay network. The results presented in this paper demonstrate that the approach has a good performance in forecasting, restraining and eliminating the worms on P2P overlay network. Further more, since the peers in the overlay is assumed to be a common Internet node, the prevention model can also be used to deal with worm on the Internet. Therefore, we believe our technique is complementary to other existing worm detection algorithms.

Acknowledgements

This paper is supported by the National Natural Science Foundation of China (Grant No. 30400446) and The Research Fund for the Doctoral Program of Higher Education of China (Grant No. 2004061102).

References

[1] E.H. Spafford, “The Internet Worm Program: An Analysis,”1988;www.cerias.purdue.edu/homes/spaf/tech-reps/823.pdf.

[2] R.V. Dantu, “An Architecture of Security Engineering”, ACSA Workshop on Application of Engineering Principles for Security System Design, November, 2002

[3] D. Moore, C. Shannon, G. Voelker, and S. Savage, "Internet quarantine: Requirements for containing self-propagating code." in IEEE INFOCOM, 2002.

[4] Dr Eugene Schultz, "Worms and viruses: are we losing control?", Computers & Security (2004) 23, 179-180

[5] Ratnasamy, S., Francis, P., Handley, M., Karp, R., Shenker, S.: A scalable content addressable network. In: Proc. ACM SIGCOMM’01, San Diego, California (2001)

[6] Stoica, I., Morris, R., Karger, D., Kaashoek, M.F., Balakrishnan, H.: Chord: A scalable peer-to-peer lookup service for Internet applications. In: Proc. ACM SIGCOMM’01, San Diego, California (2001)

[7] Rowstron, A., Druschel, P.: Pastry: Scalable, distributed object location and routing for large scale peer-to-peer systems. In: Proc. IFIP/ACM Middleware 2001, Heidelberg, Germany (2001)

[8] Zhao, B.Y., Kubiatowicz, J.D., Joseph, A.D.: Tapestry: An infrastructure for fault-resilient wide-area location and routing. Technical Report UCB//CSD-01-1141, U. C. Berkeley (2001)

[9] Streftaris G, Gibson GJ. Statistical inference for stochastic epidemic models. In: Proc. of the 17th Int’l Workshop on Statistical Modelling. Chania, 2002. 609~616.

[10] Frauenthal JC. Mathematical Modeling in Epidemiology. New York: Springer-Verlag, 1980.

[11] Wang Y, Wang CX. Modeling the effects of timing parameters on virus propagation. In: Staniford S, ed. Proc. of the ACM CCS Workshop on Rapid Malcode (WORM 2003). Washington, 2003.

[12] Zou CC, Gong W, Towsley D. Code Red worm propagation modeling and analysis. In: Proc. of the 9th

ACM Symp. on Computer and Communication Security. Washington, 2002. 138~147.

[13] Wei Yu, Analyze the Worm-based Attack in Large Scale P2P Networks, Proceedings of the Eighth IEEE International Symposium on High Assurance Systems Engineering (HASE'04)

[14] J. Wu, S. Vangala, L. Gao, and K. Kwiat. An efficient architecture and algorithm for detecting worms with various scan techniques. In Proceedings of the 11th Annual Network and Distributed System Security Symposium (NDSS'04), February 2004. to appear.

[15] C. C. Zou, L. Gao, W. Gong, and D. Towsley. Monitoring and early warning for internet worms. In Proceedings of 10th ACM Conference on Computer and Communications Security (CCS'03), October 2003.

[16] V.H. Berk, R.S. Gray, and G. Bakos. Using sensor networks and data fusion for early detection of active worms. In Proceedings of the SPIE AeroSense, 2003.

[17] Williamson, M.M, "Throttling Viruses: Restricting propagation to defeat malicious mobile code", Computer Security Applications Conference, 2002. Proceedings. 18th Annual, 9-13 Dec. 2002 Page(s): 61 –68