

# The Emerging Threat of Peer-to-Peer Worms

Nassima Khiat<sup>(1,2)</sup>, Yannick Carlinet<sup>(1)</sup> and Nazim Agoulmine<sup>(2)</sup>

<sup>(1)</sup> France Telecom R&D

{nassima.khiat, yannick.carlinet}@orange-ft.com

<sup>(2)</sup> LRSM, University of Evry, France

nazim.agoulmine@iup.univ-evry.fr

**Abstract- Security of Peer-to-Peer (P2P) systems is a subject of strategic importance for ISPs (Internet Service Providers). Worms making use of P2P overlay networks to propagate present a new threat potentially more harmful than non-P2P worms. This is essentially due to their way of propagation, which can be faster and more furtive than other existing worms. In this paper, we highlight some aspects related to the P2P worms problem: their propagation, modeling, detection and containment. We also suggest areas of investigation in the design of a detection and mitigation system, from a network operator's point of view.**

## I. INTRODUCTION

Internet worms are one of the most crucial threats for the ISPs' customers (and ISPs themselves), as shown in [16]. Detection and protection systems (listed in [12] and [15]) were designed to counter them; however we must now face a new trend in worm design: peer-to-peer worms. A P2P worm is a worm that makes use of a P2P system to spread from one machine to another. Peer-to-peer systems have become one of the "killer-applications" of the Internet and, as such, they are major drivers for the large growth of the broadband user population. While P2P applications are being used in a questionable manner currently, they will become more and more popular for distributing content legally. Moreover analyses of the traffic in the network have shown that a large majority of the traffic volume is generated by peer-to-peer applications. Therefore, peer-to-peer applications have an important strategic value for network operators. Consequently, the challenge of securing peer-to-peer systems is important for two reasons, first to prevent attacks on a valuable application, and also to prevent the propagation of worms in general. Indeed, worms are a threat to all the ISP customers as well as to the network infrastructure due to the harmful actions they can carry out (such as credit card theft, denial-of-service attacks, information theft and so on).

Peer-to-peer worms can potentially be faster and more furtive than other types of worms. This is because of three main factors. Firstly, the users of a P2P application must run the peer program on their terminal. While they can use different versions/implementations of the software, practical experience shows that most of the users run the same software. So, a vulnerability in this software could potentially allow a worm to infect millions of hosts (for instance there are more than 10 millions users connected simultaneously on the eMule network). Secondly, by making use of the overlay topology of

the P2P system, P2P worms do not waste time probing unused IP addresses (like a random-scanning worm would do, for example). Thirdly, they do not generate high rates of failed connections and they can blend into the normal traffic patterns of the P2P network (as explained in [1]). Göldi and Hiestand [15] have shown that detection systems based on the analysis of worm scans cannot differentiate the normal network activity of a P2P client from a worm.

The aim of this paper is to make out the problem of P2P worms, first by giving an overview of the state-of-art research on the subject, and then by suggesting some possible areas of investigation in the design of detection and mitigation systems, from a network operator view-point. In section II, we present different ways by which the worm can spread in the P2P system. In section III, we show models for the propagation of P2P worms and in section IV, we mention some solutions for the detection of P2P worms. Section V gives an overview of possible leads to tackle the issue of P2P worms. Finally, section VI concludes the paper.

## II. PROPAGATION OF P2P WORMS

According to its propagation strategy, a P2P worm can belong to one of the following two classes:

*Topological scan based P2P worms:* in this case, the worm uses information about victim's neighbors in the P2P system to spread through it. This can significantly increase the speed of the worm, because the worm does not waste time in searching for victims. A hit-list can be used to improve the efficiency of the topological scan of this class of P2P worms. In that case the addresses of the peers in the P2P system are collected before the worm is launched, which will permit the worm to gain more time at the beginning of the attack. Once a host is infected, it will participate in infecting the remainder hosts of the hit-list, until there will be no more machines to hit in this list. To the best of our knowledge, no topological scan based worm has been seen in the Internet yet.

*Passive P2P worms:* they correspond to the existing P2P worm class. In this case, the worm does not search for targets, but waits for them. In fact, the worm resides in the shared folder of the infected machine, 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. This is the case of the Benjamin [3] worm, which was launched on May 18<sup>th</sup>, 2002, and spread over the KaZaA network. Another P2P worm which has already been seen in the Internet in February 2001 is Gnuman [4], which behaves a little bit differently. It simulates a Gnutella peer, and answers positively to all the requests coming from other clients, by changing the name of the corrupted file by the words used in the request. The requesting host will then download the corrupted file thinking that it is the file he sought out, which will permit the propagation of the worm. This kind of worms cannot operate in centralized P2P system, due to the fact that in this case, the requests are not sent to all the peers, but only to some of them, named super-peers. So, the worm on the other peers cannot answer to all requests. A list of known passive P2P worms can be found in [10].

### III. MODELING P2P WORMS

Modeling worms is useful to understand how particular elements can affect their spreading. It also helps in studying the effectiveness of some solutions of worm detection and mitigation, to find the optimal application of these solutions that permits to fight worms effectively. Understanding the factors affecting the spread of malware is useful to design network infrastructures resilient to such attacks [2]. Epidemiological models have been studied to represent random scan worms and local subnet scanning worms, like in [11], but they cannot be used to model P2P worms, because of the particular characteristics of P2P worms spreading. So, new studies have been carried out to model P2P worms propagation. In the following, we describe some of the proposed models.

Yu et al. [6] gave an analytical model of the propagation of topological scan based P2P worms. The simulation results show that: P2P based worms are more efficient compared to random scan based worms, and hit-list based P2P worms are more efficient than P2P worms which does not use a hit list. This is due to the fact that in the first case, the worm does not waste time in searching for victims, since it knows the structure of the P2P system before the attack is launched. Another observation is that for all attack strategies, the increase of the P2P system size leads to the improvement of the performance of the attack. The topology degree also has an impact on P2P worm propagation. A greater topology degree conducts the worm to be more effective, and the increase of P2P hosts vulnerability results in the increase of the efficiency of the attack.

Passive scan based P2P worms have also been represented with an analytical model in [7]. This model has been used to assess the impact of a detection solution, Credence, described in the next section, on the P2P worm propagation, and to determine approximately how widespread the Credence system must be so as to combat the worm efficiently.

In [2], authors give a model for Gnutella-type P2P systems, by addressing a shortcoming observed in [7]. In fact, the authors in [2] note that, in [7], the assumption that a

vulnerable peer can be infected by any of the infected ones in the network is not true. This is because, the possible victims of an infected peer are limited to those which are TTL hops away from it and not the whole P2P network; otherwise it would lead to an overestimation of the spreading. Thus, the introduction of such a characteristic would avoid false positives (alerts when there is no attack).

### IV. SOLUTIONS TO P2P WORMS

Most of the research and existing systems on worm detection (such as Intrusion Detection Systems (IDS) [12]) cannot be used to detect P2P worms, because they are based on characteristics that cannot be found in P2P worms. Göldi and Hiestand [15] found out that solutions based on anomalous behavior in the network traffic are inefficient, because P2P worm traffic blend into the normal traffic patterns of the P2P system. In this section, we present some solutions proposed specifically for P2P worms.

Zhou et al. [1] propose a self-defense infrastructure inside a P2P network, by defining some "Guardian nodes" among all the system peers. Those "Guardian nodes" have an automatic worm detection property. The authors suggest to use an approach based on the observation that a majority of worms modify the control flow of a vulnerable program to execute malicious code injected from the network or from the memory [1]. The role of a guardian node is to trigger an alarm when it detects an attack, to warn the other nodes of the system, which will take appropriate actions to become immune (e.g. stop the vulnerable application). If there are super-nodes in the P2P system, they can be used as guardian nodes to disseminate more quickly the alarm, due to their high connectivity. This solution has some limitations because the detection mechanism involves hardware modification and causes performance degradation.

Another solution, Peer Pressure, has been proposed by Keyani et al. [8], against P2P systems fragmentation attacks. This kind of attack aims to disconnect the most connected nodes of a decentralized P2P system, which becomes a set of isolated pieces. Such an attack targets the P2P network itself and could be performed by a worm.

The idea of the proposed solution is to build a virtual P2P system with exponential topology (homogeneous network where all nodes have roughly the same number of links [8]) from a P2P system which has inhomogeneous or scale-free topology. This can be done by building a list of virtual neighbors for each node in the system. The nodes in the virtual system will not be connected and will not send requests. When the system is under an attack, each lost neighbor of a node is replaced by a virtual node, until the end of the attack, where the list is rebuilt using other nodes. The function of detecting an attack is performed by all nodes of the system. This happens when the percentage of the second degree neighbors (two hops neighbors) lost is greater than the percentage of the first degree ones (direct neighbors), and greater than a threshold (to avoid false positives). This threshold is used to differentiate an attack from the random

system failure, where the percentage of dead neighbors of the first and second degree is the same. One limitation of this solution is that a large percentage of users have to adopt it to be effective [8] which is not easy.

Walsh and Sirer [9] have proposed a reputation system, named Credence, which is a decentralized system to evaluate file authenticity and to avoid corrupted file downloads by clients. This solution is generally used against pollution, where file content does not correspond to file description, or is corrupted. Each node votes by giving its level of trust for a file (i.e. if it considers that the file corresponds to its descriptor or not). When a client wants to download a file, it looks beforehand at the votes given by the other peers to this file. Other reputation systems exist, like [13], but they are based on node reputation. In contrast, Credence defines file reputation, which is more relevant to the context of worms, due to the fact that node reputation can change over time. This can happen if the host is patched or changes its identity in the network. Solutions based on object reputation can mitigate the propagation of a worm, because they can prevent peers from downloading corrupted files (including files containing a worm) but they do not detect if there is an attack in the system.

## V. OPEN ISSUES AND CHALLENGES

Solutions need to be found so as to detect this new kind of worms and mitigate their potentially devastating impact on ISP's networks and customer hosts. This offers many challenges:

Since P2P worm network activity is very similar to the network activity of a legitimate peer, we may need to monitor the traffic generated by peers at the application level, so as to detect abnormal behavior. For instance, some worms such as Gnutella generate traffic that is perfectly normal if it is observed in the network layer, but it is bogus at the level of the Gnutella protocol. A Gnutella peer can be detected by analyzing the requests/responses exchanges between the peers and if one of them answers positively to all the requests, then we can have some confidence that this peer behaves anomalously.

It would also be interesting to investigate the use of a high-interaction honeypot that acts as a peer in the P2P network and collects data about the network, so as to find suspicious behavior of other peers. It may also get information from a reputation system implemented in the P2P network.

Finally, after the detection has happened we need to have a containment system to prevent further propagation of the worms. We believe that this system should be located in the ISP network because end-host users often do not have the technical knowledge or the will to implement security measures on their PC. The Witty worm propagation (cf. [14]) has shown that even if the target hosts belong to users with the best security practice that can be expected, a worm can spread very widely and very quickly. Indeed the Witty worm was exploiting a vulnerability in several firewall software

products. Moreover the containment should be sufficiently distributed so as to be efficient.

## VI. CONCLUSION

In this paper, we have presented an overview of the research on an emerging threat on the Internet: worms spreading through peer-to-peer networks. We have first given a description of how they function and how they spread. We have presented research activities that model the spreading of topological and passive scan worms. Then we have described some detection and mitigation systems that can be used to cope with P2P worms. Few research activities have focused on P2P worms; this is why we have presented works that were not focused specifically on P2P worms, but that can be applied to detect them. Finally, we have suggested areas of investigation in the design of a detection and mitigation system, from a network operator's point of view.

## REFERENCES

- [1] L. Zhou, L. Zhang, F. McSherry, N. Immorlica, M. Costa and S. Chien, "A First Look at Peer-to-Peer Worms: Threats and Defenses," In Proceedings of Peer-to-Peer Systems IV, 4th International Workshop (IPTPS), pages 24-35, February 2005.
- [2] K. Ramachandran and B. Sikdar, "Modeling Malware Propagation in Gnutella Type Peer-to-Peer Networks," The Third International Workshop on Hot Topics in Peer-to-Peer Systems (Hot-P2P), Rhodes Island, Greece, April 2006.
- [3] [http://siliconvalley.internet.com/news/article.php/3531\\_1141841](http://siliconvalley.internet.com/news/article.php/3531_1141841)
- [4] [http://www.symantec.com/avcenter/venc/data/w32\\_gnuman.worm.html](http://www.symantec.com/avcenter/venc/data/w32_gnuman.worm.html)
- [5] N. Weaver, V. Paxson, S. Staniford, and R. Cunningham, "A taxonomy of computer worms," In The First ACM Workshop on Rapid Malcode (WORM), 2003.
- [6] W. Yu, C. Boyer, S. Chellappan and D. Xuan, "Peer-to-Peer System-based Active Worm Attacks: Modeling and Analysis," Communications, 2005. ICC 2005. 2005 IEEE International Conference on Volume 1, 16-20 May 2005 Page(s):295 - 300 Vol. 1.
- [7] R. Thommes and M. Coates, "Epidemiological Modeling of Peer-to-Peer Viruses and Pollution," In Proceedings of IEEE INFOCOM, April 2006.
- [8] P. Keyani, B. Larson and M. Senthil, "Peer Pressure: Distributed Recovery from Attacks in Peer-to-Peer Systems," In Intl. Workshop on Peer-to-Peer computing, pages 306-320, 2002.
- [9] K. Walsh and E. G. Sirer, "Thwarting P2P Pollution Using Object Reputation," Tech. Rep. Computer Science Department Technical Report TR2005-1980, Cornell University, February 2005.
- [10] <http://www.viruslist.com/en/virusesdescribed?chapter=153311928>
- [11] Z. Chen, L. Gao and K. Kwiat, "Modeling the spread of active worms," in Proceedings of IEEE INFOCOM 2003.
- [12] H. Debar, M. Dacier and A. Wespi, "Towards a taxonomy of intrusion-detection systems," Computer Networks, vol. 31, pp. 805-822, 1999.
- [13] F. Cornelli, E. Damiani, S. De Capitani di Vimercati, S. Paraboschi and P. Samarati, "Choosing Reputable Servents in a P2P Network". In International World Wide Web Conference, May 2002.
- [14] <http://www.caida.org/analysis/security/witty/>
- [15] C. Göldi and R. Hiestand, "Scan Detection Based Identification of Worm Infected Hosts", Thesis Report, Swiss Federal Institute of Technology, ETHZ, Zurich, 18 April 2005.
- [16] S. Staniford, V. Paxson, and N. Weaver, "How to Own the Internet in your spare time," In Proceedings of the 11th USENIX Security Symposium, August 2002.