Malware models for network and service management

Jérôme François, Radu State, Olivier Festor

Malware models for network and service management

Jérôme François, Radu State and Olivier Festor

MADYNES - INRIA Lorraine, CNRS, Nancy, France
{jerome.francois,radu.state,olivier.festor}@loria.fr

Abstract. Different kinds of malware like the botnets and the worms are a main threat on Internet for the current and future. Their efficiency to control systems is proved and we are investigating the malware mechanism that can be adapted to get an efficient and scalable management plane. Our work consists in modelling malware based network management and assessing its performance.

1 Introduction

The network and service management is a major component in order to provide value added services. Due to the multiplication of services and the third-party management delegation, the management boundaries are not clear and the management operations are faced with several problems: more and more devices to manage, an hostile environment due to security appliances (firewalls or intrusion detection systems), the distance between the manager and the devices to manage... However the creators of malware faced the same problems and today the efficiency of malware is well known. Our work is motivated by their results and we propose a malware based network management plane.

The section 2 will introduce the botnets and the worms. The section 3 presents the motivation of a management framework based on malware. A formal model will be presented in 4. A brief overview of related works is given in 5. Finally, we conclude and plan future works.

2 Malware efficiency

A botnet is a network of a compromised machines which are usually called zombies or bots, these machines wait for instructions from the master (the hacker) and perform operations on behalf of the latter. IRC networks are a well-known way to send a command on a channel to all the bots. We are very interesting in the botnets using IRC networks for three main reasons:

- the exchange of commands is simple and multiple operations are possible: to do denial of service attacks or retrieve personal data for example. So an application can be easily adapted to use an IRC network.
- IRC networks are very resilient [2]

– the botnets have already proven their efficiency. In [4] some statistics about botnets are presented showing that it is possible to control 400 000 bots.

The worms are a main threat for Internet due to the propagation speed [7] and their ability to bypass security equipments. The authors in [7] have shown that little worm can infect a large population 1 000 000 hosts in 2 generations. Moreover the worms are not limited to only propagate themselves, but can be used to send spam emails or retrieve private informations. A hacker can lead computers thanks to the P2P network built by the slapper worm [1].

3 Malware benefits

In [8], we propose a novel malware based management framework. We will focus on two kinds of malware introduced in the previous section. Our research is motivated by two main facts:

– a huge number of hosts could be managed without specific and different configurations
– the management applications are various because the command or propagation mechanisms are totally unlinked with the applications

However, even though there are a lot of studies about the power of botnets or worms we have to define a precise model to evaluate the efficiency of a malware management based. There are some important questions to answer: what is the probability to reach 90% of the hosts ? What is the probability to reach 50% of the hosts in 1 minute ? How much time is needed to have 99% of reached hosts with a probability of 0.99 ? Another aim is to optimize the malware communication scheme to have the best performances.

4 Botnet model

In order to establish a mathematical botnet model, we have to model an IRC network. The servers where the clients are connected are organized as a spanning tree. In [6], the author considers a recursive random tree i.e. each node has a number and the nodes are linked successively to a previous node. In fact the node j has the probability $\frac{1}{j-1}$ to be connected to a given node among nodes $1, 2, ..., j-1$. Indeed, the probability that the distance between node i and node j is d with $i < j$ is :

$$P(i, j, d) = \frac{1}{j-1}[P(1, i, d-1) + P(2, i, d-1) + ... + P(j-1, i, d-1)]$$

We propose to extend this model by finding a formula to determine the probability to have a distance less or equal to d between a given origin node o and a set N of other nodes between a total of n nodes of the tree: $P(o, d, N, n)$. By summing the probabilities for the different sets, we have the global probability independent from N.

The formula is recursive. At the beginning we have to determine how many nodes of N we want to connect to the origin node directly (distance = 1). There are $|N|$ possibilities: $i \in [1, |N|]$. Then we have to choose the combination of these i nodes which is $CHOICE =< c_1, ..., c_i > \subseteq N$. $REMAINDER = N - CHOICE =< r_1, ..., r_p >$ represents the nodes to be connected directly or not at the next hop nodes previously chosen. So we define a permutation with repetitions of the elements of $SUITE < s1, .., sp >$ with $\forall k,\ s_k \in CHOICE$. Indeed, we obtain the followind formula :

$$P(o, d, N, n) = \sum_{i=1..n \ et \ <c_1,...,c_i> \subseteq N} [\prod_{c_j=[c_1,c_m]} P_1(o, c_j) \times P(c_j, d-1, SET, n)]$$

with $SET = \{r_k\} \setminus s_k = c_i$ and $P_1(o, c_j)$ which is the probability to have the node c_j linked to o.
In our case, we assume that a node is connected to a previous at random and thus:

$$P_1(o, c_j) = distribution\_function(o, c_i) = \frac{1}{max(o, c_i) - 1}$$

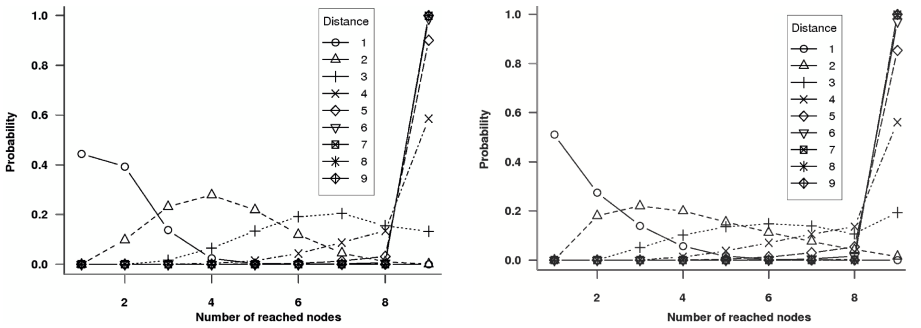
This description gives a clear idea of how we can compute the probabilities but the formula we used for our experiments is more complex and integrates the following facts:

- the recursive tree implies an order in the node. For example it is impossible to have node 1 connected to node 3 and node 3 connected to node 2.
- because we want precise probabilities, we need to have the probability for an exact set N and not a set containing N

Figure 1(b) shows that we are able to determine the probability to reach a given number of nodes at a maximal given distance. For instance, we can see that the probability to reach three nodes at a distance less or equal to two is about 0.2. The node are the servers of the IRC network, so we can deduce from this value the probability to reach a certain number b if there are randomly connected to server by multiplying this value by $\frac{b}{n}$. If we want to optimize the number of reached node, we can compare the probability with the different origin node as in 1(a).

5 Related works

The classical management frameworks have shown their limits related to the scalability and several authors proposed solutions to deal with this problem. In [3] a decentralized management is proposed where a query is transformed into different subqueries and where the results are aggregated of each one. The idea to use malware for management was proposed in [5] where a worm patrols on different hosts to detect malfunctioning. However the framework is not modeled and the experiments were very limited. Multiple articles or books about what botnets or worms can do and how they are efficient can be found in [2] and [7].



(a) from the origin node 5 (b) average over all origin nodes

Fig. 1. The probability to reach a certain number of nodes at a defined distance in tree with a total number of 10 nodes

6 Conclusion

The efficiency of botnets and worms have been proven in the past. We are convinced that a benign usage can be obtained for addressing scalability in the network management plane. It implies to know exactly what we are able to do with them. Thus the model aims to prove the efficiency of this new framework. We attend to extend the model with more parameters and metrics related to the performance of the management plane. A second activity will consist in optimizing the management plane with respect to the previously mentioned metrics. Finally a realistic testbed implementation will be performed.

References

1. Ivan Arce and Elias Levy. An analysis of the slapper worm. *IEEE Security and Privacy*, 1(1):82–87, 2003.
2. Evan Cooke, Farnam Jahanian, and Danny Mcpherson. The zombie roundup: Understanding, detecting, and disrupting botnets. pages 39–44, June 2005.
3. K.-S. Lim and R. Stadler. Real-time views of network traffic using decentralized management. In *Integrated Network Management, 2005. 9th IFIP/IEEE International Symposium on*, 2005.
4. Laurianne McLaughlin. Bot software spreads, causes new worries. *IEEE Distributed Systems Online*, 5(6), 2004.
5. Hiroyuki Ohno and Akihiro Shimizu. Improved network management using nmw (network management worm) system. In *Proceedings of INET'95: Honolulu, Hawai'i, June 27-30, 1995.*, 1995.
6. Vladimir N. Sachkov. *Probabilistic methods in combinatorial analysis*, chapter 6 - Random Graphs and Random Mappings. Cambridge University Press, 1997.
7. Stuart Staniford, David Moore, Vern Paxson, and Nicholas Weaver. The top speed of flash worms. In *WORM '04: Proceedings of the 2004 ACM workshop on Rapid malcode*, pages 33–42, New York, NY, USA, 2004. ACM Press.
8. Radu State and Olivier Festor. Malware: a future framework for device, network and service management, 2006.