© 2015 IJSRST | Volume 1 | Issue 3 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology

Providing Security to User Confidentiality Data in Large Scale Networks

Dharavath Champla, Yasaram Ganesh

Assistant Professor, Computer Science and Engineering, Christu Jyoti Institute of Technology & Science, Jangaon, Telangana, India

ABSTRACT

Malware is pervasive in networks, and poses a critical threat to network security. However, we have very limited understanding of malware behavior in networks to date. In this paper, we investigate how malware propagates in networks from a global perspective. We formulate the problem, and establish a rigorous two layer epidemic model for malware propagation from network to network. Based on the proposed model, our analysis indicates that the distribution of a given malware follows exponential distribution, power law distribution with a short exponential tail, and power law distribution at its early, late and final stages, respectively. Extensive experiments have been performed through two real-world global scale malware data sets, and the results confirm our theoretical findings.
Keywords : Arduino, Wi-Fi (ESP 8266), Load cell, Database System

I. INTRODUCTION

Malware are malicious software programs deployed by cyber attackers to compromise computer. These Malwares are being created at an alarming rate in order to gain political and financial rewards. These malwares are sent to infect the whole network and gain confidential information. The systems that are affected by these Malwares are called as bots. The action against these malwares can be taken only when the propagation pattern, the behaviour pattern of the malwares are studied.

To date, we do not have a solid understanding about the size and distribution of malware or botnets. Researchers have employed various methods to measure the size of botnets, such as botnet infiltration, DNS redirection, external information. These efforts indicate that the size of botnets varies from millions to a few thousand. There are no dominant principles to explain these variations. As a result, researchers desperately desire effective models and explanations for the chaos. Dagon, Zou and Lee revealed that time zone has an obvious impact on the number of available bots. Mieghem et al. indicated that network topology has an important impact on malware spreading through their rigorous mathematical analysis. Recently, the emergence of mobile malware, such as Cabir, Ikee , and Brador, further increases the difficulty level of our understanding on how they propagate. More details about mobile malware can be found at a recent survey paper. To the best of our knowledge, the best finding about malware distribution in large-scale networks comes from Chen and Ji the distribution is non-uniform. All this indicates that the research in this field is in its early stage.

The epidemic theory plays a leading role in malware propagation modelling. The current models for malware spread fall in two cate.

Gories: the epidemiology model and the control theoretic model. The control system theory based models try to detect and contain the spread of malware. The epidemiology models are more focused on the number of compromised hosts and their distributions, and they have been explored extensively in the computer science community. Zou et al. used a susceptible-infected (SI) model to predict the growth of Internet worms at the early stage. Gao and Liu recently employed a susceptible-infected-recovered (SIR) model to describe mobile virus propagation. One critical condition for the epidemic models is a large vulnerable

population because their principle is based on differential equations. More details of epidemic modelling can be find in. As pointed by Willinger et al., the findings, which we extract from a set of observed data, usually reflect parts of the studied objects. It is more reliable to extract theoretical results from appropriate models with confirmation from sufficient real world data set experiments. We practice this principle in this project.

In this project, we study the distribution of malware in terms of networks (e.g., autonomous systems, ISP domains, abstract networks of smartphones who share the same vulnerabilities) at large scales. In this kind of setting, we have a sufficient volume of data at a large enough scale to meet the requirements of the SI model. Different from the traditional epidemic models, we break our model into two layers. First of all, for a given time since the breakout of a malware, we calculate how many networks have been compromised based on the SI model. Secondly, for a compromised network, we calculate how many hosts have been compromised since the time that the network was compromised. With this two layer model in place, we can determine the total number of compromised hosts and their distribution in terms of networks. Through our rigorous analysis, we find that the distribution of a given malware follows an exponential distribution at its early stage, and obeys a power law distribution with a short exponential tail at its late stage, and finally converges to a power law distribution. We examine our the oretical findings through two large-scale realworld data sets: the Android based malware and the Conficker. The experimental results strongly support our theoretical claims. To the best of our knowledge, the proposed two layer epidemic model and the findings are the first work in the field. Our contributions are summarized as follows.

We propose a two layer malware propagation model to describe the development of a given malware at the Internet level. Compared with the existing single layer epidemic models, the proposed model represents malware propagation better in largescale networks.

II. RESEARCH

a. Modeling Botnet Propagation Using Time Zones (2006).

Time zones play an important and unexplored role in malware epidemics. To understand how time and location affect malware spread dynamics, we studied botnets, or large coordinated collections of victim machines (zombies) controlled by attackers. Over a six month period we observed dozens of botnets representing millions of victims. We noted diurnal properties in botnet activity, which we suspect occurs because victims turn their computers off at night. Through binary analysis, we also confirmed that some botnets demonstrated a bias in infecting regional populations. Clearly, computers that are offline are not infectious, and any regional bias in infections will affect the overall growth of the botnet. We therefore created a diurnal propagation model. The model uses diurnal shaping functions to capture regional variations in online vulnerable populations. The diurnal model also lets one compare propagation rates for different botnets, and prioritize response. Because of variations in release times and diurnal shaping functions particular to an infection, botnets released later in time may actually surpass other botnets that have an advanced start. Since response times for malware outbreaks is now measured in hours, being able to predict short-term propagation dynamics lets us allocate resources more intelligently. We used empirical data from botnets to evaluate the analytical model.

b. My Botnet is Bigger than Yours (Maybe, Better than Yours): why size estimates remain challenging

As if fueled by its own fire, curiosity and speculation regarding botnet sizes abounds. Among researchers, in the press, and in the classroom—the questions regarding the widespread effect of botnets seem never-ending: what are they? how many are there? what are they used for? Yet, time and time again, one lingering question remains: how big are today's botnets? We hear widely diverging answers. In fact, some may argue, contradictory. The root cause for this confusion is that the term botnet size is currently poorly defined. We elucidate this issue by presenting different metrics for counting botnet membership and show that they lead to

widely different size estimates for a large number of botnets we tracked. In particular, we show how several issues, including cloning, temporary migration, and hidden structures significantly increase the difficulty of determining botnet size with any accuracy. Taken as a whole, this paper calls into question speculations about botnet size, and more so, questions whether size really matters.

c. Smartphone Malware and Its Propagation Modeling: A Survey

Smartphones are pervasively used in society, and have been both the target and victim of malware writers. Motivated by the significant threat that presents to legitimate users, we survey the current smartphone malware status and their propagation models. The content of this paper is presented in two parts. In the first part, we review the short history of mobile malware evolution since 2004, and then list the classes of mobile malware and their infection vectors. At the end of the first part, we enumerate the possible damage caused by smartphone malware. In the second part, we focus on smartphone malware propagation modeling. In order to understand the propagation behavior of smartphone malware, we recall generic epidemic models as a foundation for further exploration. We then extensively survey the smartphone malware propagation models. At the end of this paper, we highlight issues of the current smartphone malware propagation models and discuss possible future trends based on our understanding of this topic.

EXISTING SYSTEM:

- The epidemic theory plays a leading role in malware propagation modelling. The current models for malware spread fall in two categories: the epidemiology model and the control theoretic model.
- The control system theory based models try to detect and contain the spread of malware. The epidemiology models are more focused on the number of compromised hosts and their distributions, and they have been explored extensively in the computer science community.
- Zou et al. used a susceptible-infected (SI) model to predict the growth of Internet worms at the early stage.
- Gao and Liu recently employed a susceptible-infected-recovered (SIR) model to describe mobile virus propagation.

DISADVANTAGES OF EXISTING SYSTEM:

- One critical condition for the epidemic models is a large vulnerable population because their principle is based on differential equations.
- As pointed by Willinger et al. the findings, which we extract from a set of observed data, usually reflect parts of the studied objects. It is more reliable to extract the-oretical results from appropriate models with confirmation from sufficient real world data set experiments.

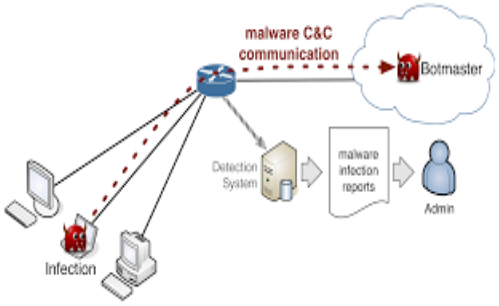
PROPOSED SYSTEM:

- In this paper, we study the distribution of malware in terms of networks (e.g., autonomous systems, ISP domains, and abstract networks of smartphones who share the same vulnerabilities) at large scales.
- In this kind of setting, we have a sufficient volume of data at a large enough scale to meet the requirements of the SI model. Different from the traditional epidemic models, we break our model into two layers.
- First of all, for a given time since the breakout of a malware, we calculate how many networks have been compromised based on the SI model.
- Secondly, for a compromised network, we calculate how many hosts have been compromised since the time that the network was compromised.

ADVANTAGES OF PROPOSED SYSTEM

- Our rigorous analysis, we find that the distribution of a given malware follows an exponential distribution at its early stage, and obeys a power law distribution with a short exponential tail at its late stage, and finally converges to a power law distribution.

III. SYSTEM ARCHITECTURE



IV. IMPLIMENTATION

a. Admin Module:

Admin needs to login to perform his actions. He can view all user information. He should identify the malware behavior and register it in the malware inventory to protect the data from that malware. He can block the malicious user by monitoring the users.

b. User Module:

The User should register and login to access his account. He can follow the people in this media. He can share the message and his feeling with the followers. Followers can see his posts and they can send tweets or comment to that post

c. Malware Propagation to identify the malware in tweets:

This module can identify the malware type of tweets before sending it to the users post. If the tweet is a malware then it will block that tweet and it doesn't allow that comment on specific post.

V. CONCLUSION

In this paper, we thoroughly explore the problem of malware distribution at large-scale networks. The solution to this problem is desperately desired by cyber defenders as the network security community does not yet have solid answers. Different from previous modeling methods, we propose a two layer epidemic model: the upper layer focuses on networks of a large scale networks, for example, domains of the Internet; the lower layer focuses on the hosts of a given network. This two layer model improves the accuracy compared with the available single layer epidemic models in malware modeling. Moreover, the proposed two layer model offers us the distribution of malware in terms of the low layer networks. We perform a restricted analysis based on the proposed model, and obtain three conclusions: The distribution for a given malware in terms of networks follows exponential distribution, power law distribution with a short exponential tail, and power law distribution, at its early, late, and final stage, respectively. In order to examine our theoretical findings, we have conducted extensive experiments based on two real-world large-scale malware, and the results confirm our theoretical claims.

VI. REFERENCES

[1]. B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna, "Your botnet is my botnet: Analysis of a botnet takeover," in CCS '09: Proceedings of the 2009 ACM conference on computer communication security, 2009.

[2]. D. Dagon, C. Zou, andW. Lee, "Modeling botnet propagation using time zones," in Proceedings of the 13 th Network and Distributed System Security Symposium NDSS, 2006.

[3]. M. A. Rajab, J. Zarfoss, F. Monrose, and A. Terzis, "My botnet is bigger than yours (maybe, better than yours): why size estimates remain challenging," in Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets, 2007.

[4]. D. Dagon, C. C. Zou, and W. Lee, "Modeling botnet propagation using time zones," in NDSS, 2006.

[5]. P. V. Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," IEEE/ACM Transactions on Networking, vol. 17, no. 1, pp. 1-14, 2009.

[6]. Cabir, http://www.f-secure.com/en/web/labs global/2004- threat-summary.

[7]. Ikee, http://www.f-secure.com/vdescs/ worm iphoneos ikee b.shtml.

[8]. Brador, http://www.f-secure.com/v-descs/brador.shtml.

[9]. S. Peng, S. Yu, and A. Yang, "Smartphone malware and its propagation modeling: A survey," IEEE Communications Surveys and Tutorials, in press, 2014.

[10]. Z. Chen and C. Ji, "An information-theoretic view of network-aware malware attacks," IEEE Transactions

on Information Forensics and Security, vol. 4, no. 3, pp. 530- 541, 2009.
[11]. A. M. Jeffrey, xiaohua Xia, and I. K. Craig, "When to initiate hiv therapy: A control theoretic approach," IEEE Transactions on Biomedical Engineering, vol. 50, no. 11, pp. 1213-1220, 2003.
[12]. R. Dantu, J.W. Cangussu, and S. Patwardhan, "Fast worm containment using feedback control," IEEE Transactions on Dependable and Secure Computing, vol. 4, no. 2, pp. 119-136, 2007.