ISSN: 2366-1313

MALWARE IS SPREADING RAPIDLY THROUGHOUT EXTENSIVE NETWORKS

#1J.Raja Kala, Assistant Professor,
#2Dr. T.Veeranna, Associate Professor,
#3V. V.Siva Prasad, Assistant Professor,
Department of Computer Science and Engineering,
SAI SPURTHI INSTITUTE OF TECHNOLOGY, SATHUPALLY, KHAMMAM.

ABSTRACT:

Malicious malware that significantly exploits computer systems is readily available. Unfortunately, there is still a lack of complete knowledge on how harmful software operates on computers. Using a worldwide perspective, this research looks into the issue of harmful software spreading from system to system. We discovered the issue and developed a two-stage pandemic model to illustrate the propagation of malware from one device to another. According to our findings, a particular virus's spread may be represented as an exponential process with power law transmission, power law dispersion with a brief exponential drop, and power law circulation at the onset, apex, and nadir of the process, respectively. We checked our hypotheses by examining data from two large databases that collect intelligence about harmful software from throughout the world. Our research proves our claims to be true.

Keywords: Malware, Propagation, Modelling, Power Law.

1. INTRODUCTION

Cybercriminals employ malicious software to compromise computer systems by taking advantage of security flaws. Malicious software authors will attempt elaborate tactics to propagate their programs across various devices because doing so can bring major financial and political benefits. Simply described, a "bot" is a single infected computer, whereas a "botnet" is a network of infected devices all under the command of the same virus. Individuals' online safety is in serious jeopardy because botnets are the major tool utilized by hackers at the moment. Security specialists, in order to mount a successful defense against cybercriminals, must have a deep comprehension of the intricate workings of malware, including its techniques of spread, botnet growth, and bot dispersion. The software development process can benefit from thinking about a pandemic. Malware diffusion can be understood using either a model inspired by disease transmission or one inspired by control theory. Models based on hypotheses are used by the control system to detect malware and stop its propagation. The number of hosts and the behavior of these hosts during transmission are often the core focus of disease transmission models. The software engineering team has devoted a great deal of time to studying this problem. The researchers used a SIR model (weakly infected recovered) to show how viruses can transfer from one device to another. At first, a model called weakly infected (SI) was used to foretell how Internet viruses would spread. Recently, the SIR model has been used to show how an infection can quickly propagate throughout a network of devices. Plague models require a big population of vulnerable individuals because their recommendations are based on a wide range of factors. Having the observed data corroborate the predicted outcomes strengthens the case for further investigation. Audience trust is increased when hypotheses are tested using validated data sets. The general public has more faith in the results obtained through the use of fitting models. The inquiry is ongoing, however measures are being taken to guarantee complete observance of this regulation.

2. EXISTING AND PROPOSED ALGORITHM

Existing System

Models of malware spread strongly depend on epidemiological factors. Research into the spread of malware is now dominated by epidemiological and control-theoretic models. To detect and contain malware, models based on control system theory are used. The discipline of computer science has done much study on epidemiological models, particularly those that count infected hosts and map their distribution. New Internet bug outbreaks were predicted using the susceptible-infected (SI) model by Zou et al. Gao and Liu used the susceptible-infected-recovered (SIR) paradigm to depict the propagation of viruses through portable electronic devices.

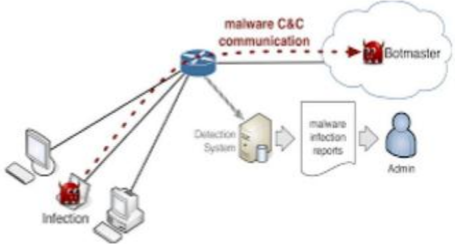


Fig.1. System Architecture of Proposed System.

Proposed Algorithm

Using real-world examples such autonomous systems, ISP domains, and abstract networks of Smartphones sharing an internet connection, this article investigates how viruses spread from one network to another.

3. PROBLEM STATEMENT

Cyberdefenders are actively seeking a way to counteract the massive propagation of viruses due to the lack of complete network security equipment on the market. Our epidemic model is unique in that it accounts for two different degrees of severity. The top layer handles information about broad networks like Internet domains, whereas the bottom layer handles information about the individual hosts that make up that network. This two-layer model improves upon the standard practice of utilizing single-layer models to simulate malware outbreaks. Furthermore, the suggested dual-layer architecture makes it easier for malware to spread through low-level networks. Our planned study will begin with an analysis of the completed level's playability. The extent of the power law distribution's exponential tail is a prime candidate for further study. Defenders may place more value on network security than on meeting the conditions necessary for implementing the two-layer approach. To stop the spread of malware within corporate networks, it may be important for individuals to interact with their ISPs.

A. Implementation of Modules

In Malware propagation in large scale networks we have themodules such as discussed below.

- ➢ Malware,
- ➢ Propagation.
- ➢ Power law

Malware:

Cybercriminals probe for security holes in networks using malicious software. Malicious software engineers spend a lot of time and energy developing sophisticated new methods for penetrating distributed computer networks. Infected computers are called "bots," and they are utilized for criminal activity. The bot has joined the ranks of the bots. When launching assaults, cybercriminals often make use of botnets, which makes it more difficult to prevent unauthorized access to systems. In order to counteract cybercriminals, defenders need a deep understanding of malware's inner workings, including how it spreads, who it recruits, how large its botnet is, and where it is spread.

Propagation: Propagation takes place in three stages such as given below,

Early stage:

The infection rate is extremely high and expanding exponentially, even though the malware has only affected a small fraction of the possible targets now under attack.

Final stage:

The virus will have reached its maximal level of dissemination once it has successfully exploited all network vulnerabilities.

Late stage: A late stage means the time interval between the early stage and the final stage.

Power Law Distribution:

The research suggests that the number of hosts in complex networks follows a power law distribution with respect to time. Countless real-world data points, from city populations to national incomes, have been found to follow a power law distribution. For example, the population of a country can be represented by a power law. The distribution of file sizes, for example, was found by researchers to follow a power law. The power law has just recently been applied to the calculation of network sizes. Power law distributions can be seen in both the Pareto and Zipf distributions. These distributions are illustrated by examples below. All of these techniques can show objects that follow the same power law distribution. The degree of regularity in Zipf distributions is higher than that of Pareto distributions. In this analysis, Zipf distributions are used to probe the idea of a power law. It is crucial to examine when and why malware growth changed from an exponential to a power law distribution. When can we say with absolute certainty that an event has moved from its introductory to its final stage?

4. PERFORMANCE EVALUATION

Here, we put our theoretical conclusions to practical use, focusing on two well-documented forms of harmful software. Conficker and Android are just two examples of malware. Mobile device malware is a new phenomenon that is growing quickly. Malware is a common term when discussing threats to Android devices. Conficker is an advanced botnet that has spread throughout the web; it's not like the normal adware that infects Android smartphones. Both formats of information are generally recognized as valid. In order to give a thorough examination of the development of dangerous software from August 2010 to October 2011, we want to compile data on Android malware. There are 1260 malicious Android apps altogether, and they may be broken down into 49 different types. Malicious Android apps aren't very dangerous because they can't do much to the operating system. Due to this flaw, Android malware can spread to more smartphones than ever before. It's safe to assume that each of the 49 networks in the dataset has a sizable population size. The availability of such data allows for this to happen. In Figure 2, we see the different types of malicious software ranked from largest to smallest based on their relative prevalence within the sample. There is a linear relationship that goes toward convergence, as shown by the scatter plot, and the subclasses are shown in log format. In a nutshell, the power law characterizes the distribution of Android apps on different devices.

The I (t) pattern, which shows how the number of Android hosts infected with malicious software grows over time, is currently under investigation. Tabulated in Table 1 below are the results of these computations on the dataset. Figure 3 is a graphical representation of the data. The data shows that in the previous 15 months, a large number of people have joined Android malware groups. We've concluded that this information is unlawful under the power law and have no plans to disclose it.

Malware Propagation in Large-Scale Networks

Due to constraints on space, they are unable to take part. The "fewest squares" method (seen in Figure 3) is used to find linear relationships in data. Based on this information, we can estimate that I(0), where I is the number of seeds, has a value of about 0.2349. We can say with some certainty that the spread of dangerous Android malware was in its infancy, based on the facts at hand. Given how pervasive each Android network is, it's safe to assume that people are the primary vector for virus dissemination. We have amassed a large amount of Conficker information from many different places. Only a small portion of our theoretical study can be shown here due to time constraints. The Internet and other similar networks are often the first thing people think of when hearing the word "Autonomous Systems" (AS). Autonomous systems (ASs) make up the bulk of the internet.

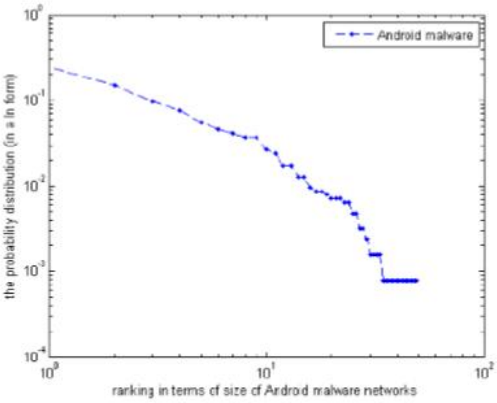
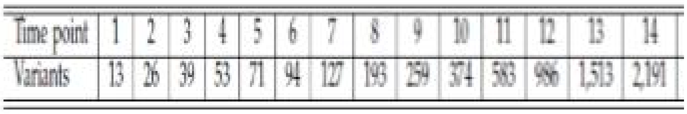


Fig.2. The probability distribution of Android malware in terms of networks.

TABLE 1. The Number of Different Android MalwareAgainst Time (Months) in 2010-2011



Time point	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Variants	13	26	39	53	71	94	127	193	259	374	583	986	1,513	2,191

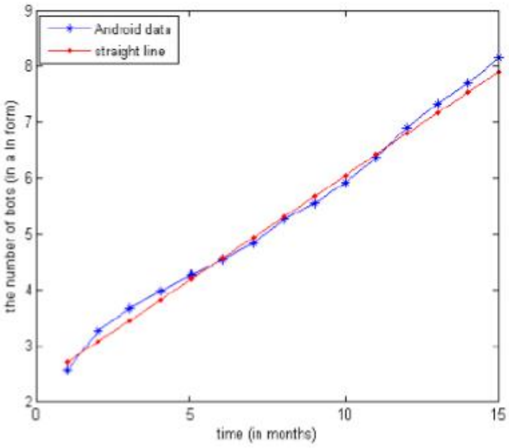
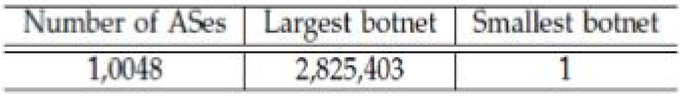


Fig.3. The growth of total compromised hosts by Android malware against time from August 2010 to October 2011.

TABLE 2. Statistics for Conficker Distribution inTerms of Ass



Number of ASes	Largest botnet	Smallest botnet
1,0048	2,825,403	1

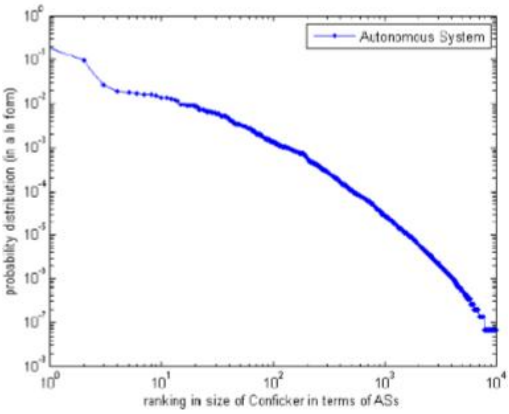
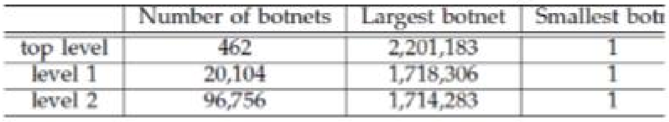


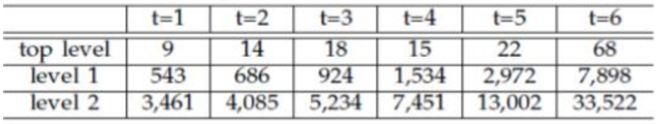
Fig.4. Power law distribution of Conficker in terms ofautonomous networks.

TABLE 3. Statistics for Conficker Distribution in Terms of Domain Names at the Three Top Levels



	Number of botnets	Largest botnet	Smallest bot
top level	462	2,201,183	1
level 1	20,104	1,718,306	1
level 2	96,756	1,714,283	1

TABLE 4. The Last Six Elements of Conficker Botnet from The Top Three Domain Name Levels



	t=1	t=2	t=3	t=4	t=5	t=6
top level	9	14	18	15	22	68
level 1	543	686	924	1,534	2,972	7,898
level 2	3,461	4,085	5,234	7,451	13,002	33,522

In Figure 4, the data is displayed logarithmically, making it easy to see that the distribution is power-law shaped. The power law stands out since it doesn't require any sort of measurement device to be put to use. By dividing infected hosts' domain names into three tiers—top, intermediate, and bottom—we can calculate the total number of machines affected. Table 3 contains data from the aforementioned research project. Figure 5 (a), (b), and (c) redisplay the data, this time on a logarithmic scale. The statistics show without a doubt that straight lines make up the vast majority of the three scale measures. They have no choice but to accept the authority's rulings. The figure has a smooth, level top, although this flatness can be explained. The number 5 is the result of a Zipf-Mandelbrot distribution. The reliability of the second theorem can now be evaluated. Selecting the six data points with the lowest values from each of the three tails allows us to test the hypothesis that the tails follow an exponential distribution. Based on the data in Table 4, we can see that these networks have been breached six times in the past (where t = 1 is the sixth-to-last time point and t = 6 is the most recent time point). The exponential distribution of the data in Table 4 is depicted in a scatter diagram (Figure 6). This holds true not only for generic top-level domains, but also for second- and third-level domains. Our third theorem holds water in this particular case.

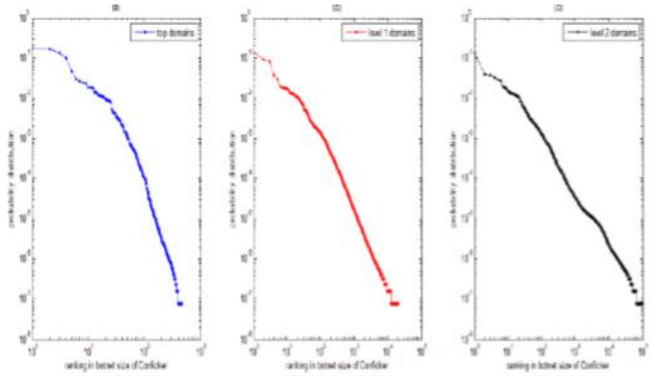


Fig.5. Power law distribution of Conficker botnet in thetop three levels of domain names.

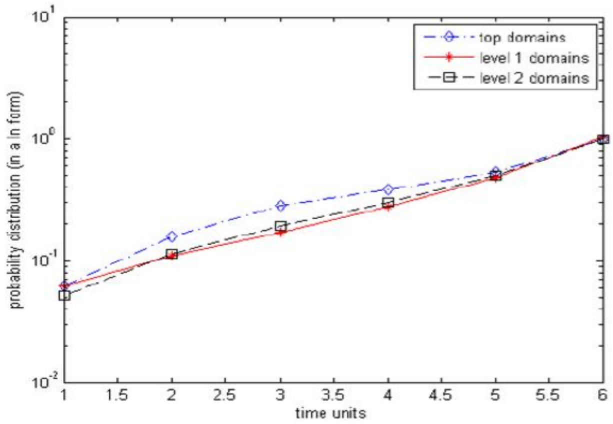


Fig.6. The three tails from the three domain name levels fitexponential distributions.

5. CONCLUSION

In this study, we look into how often malware is on enterprise-level LANs. Since the system security group hasn't come up with a solution, it's up to the digital defense group to provide one. In contrast to conventional ways, the display system we propose makes use of two tiers. On the highest level, we have the Internet and other platforms and technologies built to support infinitely many users. The most fundamental levels of operation for each host are unique. When compared to the more prevalent one-layer models, our two-layer model of malware provides more accurate results. Malware can also enter systems at a deeper level using the aforementioned dual-layer approach. The following three findings emerged from our early analysis of the proposed program: There are three separate stages to the spread of harmful software in computer systems, each with its own set of characteristics: an early phase of rapid development, an intermediate phase with steady decline and a brief period of strong growth, and a late phase of progressive dispersion. Our findings were validated when we examined two separate examples of widespread malicious software in great detail to test their validity.

REFERENCES

1. Shui Yu, Senior Member, IEEE, Guofei Gu, Member,IEEE, Ahmed Barnawi, Member, IEEE, Song Guo, Senior Member, IEEE, and Ivan Stojmenovic, Fellow, IEEE, "Malware

Propagation in Large-Scale Networks", IEEE 2015.

2. 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.
3. 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.
4. 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.
5. D. Dagon, C. C. Zou, and W. Lee, "Modeling botnet propagation using time zones," in NDSS, 2006.
6. P. V. Mieghem, J. Omic, and R. Kooij, "Virus spread innetworks," IEEE/ACM Transactions on Networking, vol. 17, no. 1, pp. 1–14, 2009.
7. Cabir,http://www.f-secure.com/en/web/labsglobal/2004- threat-summary.
8. Ikee, http://www.f-secure.com/vdescs/worm iphone os ikee b.shtml.
9. Brador, http://www.f-secure.com/v-descs/brador.shtml.
10. S. Peng, S. Yu, and A. Yang, "Smartphone malware and its propagation modeling: A survey," IEEE Communications Surveys and Tutorials, in press, 2014.