



A Stochastic Approach for Malware Detection in Mobile Network

¹M. Durgadevi , ²S. Dhanalakshmi S

^{1,2}Department of Computer Science & Engineering

Sri Muthukumaran Institute Of Technology, Mangadu, Chennai-69.

¹ksmdurgadevi@gmail.com, ²dhanalakshmisnr@gmail.com

ABSTRACT

Wireless mobile devices have turned out to be the integral part of all human communication. As a result, the computer malware is now drifting from computers to mobile phones. The problem of optimal distribution of the content-based signatures of malware helps to detect the corresponding malware and disable further propagation, in order to minimize the number of infected nodes. But in some cases, the malicious nodes may inject some dummy signatures targeting no malware into the network and induce denial-of-service attacks to the defence system. Enhancement of the system is done by developing an attack detection system that uses Multivariate Correlation Analysis (MCA) for accurate network traffic characterization by extracting the geometrical correlations between network and traffic features. MCA-based attack detection system employs the principle of anomaly-based detection in attack recognition. This makes the solution capable of detecting known and unknown attacks effectively by learning the patterns of legitimate network traffic only. Furthermore, a triangle-area-based technique is proposed to enhance and to speed up the process of MCA. The effectiveness of proposed detection system is evaluated, and the influences of both non-normalized data and normalized data on the performance of the proposed detection system are examined.

Index Terms- MCA, Triangle Area Based, Correlation, Traffic Features.

I. Introduction

The evolution of wireless mobile devices such as mobile phones, personal digital assistants, net books, and tablet PCs have been witnessed in last decades[1]. With the appearance and development of intelligent operating system, mobile devices are getting smarter and more functional. For example, they can connect to the Internet, receive and send emails and SMS /MMS, and connect to other devices for exchanging information and activating various applications. Mobile phones and personal digital assistants are becoming increasingly important in our daily life since they enable us to access a large variety of ubiquitous services.

Mobile networks, formed by the connection of mobile devices following some relationships among

mobile users, provide good platforms for spread of malware. The ideal targets of mobile malware are such devices because they are popular, programmable and general of purpose, and highly dependent on common software platforms such as Android, Symbian, Windows Mobile, and Linux.

Malware known as Malicious software, is specifically designed to damage or disrupt the devices such as a tablet or smartphone. Most mobile malware is designed to disable or interrupt a mobile device, allow a malicious user to remotely control the device or to steal personal information stored on the device. Variety of forms of hostile or intrusive software is referred by the term malware. Computer viruses, worms, Trojan horses, ransomware,

spyware, adware, scareware and other malicious programs are generally included in malware.

Malware attacks constitute a serious security risk that threatens to slow down the large scale proliferation of wireless applications. We need to quantify the maximum damage inflicted on the system owing to such outbreaks and identify the most vicious attacks as a initial step against this security threat. An epidemic model in which the worm can dynamically control the rate at which it kills the infected node and also the transmission range and/or the media scanning rate is represented in a battery constrained mobile network with propagation of the malware.

II. System Description

Mobile computing is used for creating an information management platform, which is free from constraints such as spatial and temporal. The users can access and process desired information from anywhere in the space due to independency of constraints. The capability of managing information is not affected by the state of the user, static or mobile conditions of the mobile platform.

As malware attacks become more frequently in mobile networks, to prevent serious spreading and outbreaks deploying an efficient defense system to protect against infection and to help the infected nodes to recover is necessary. The problem of optimal distribution of content based signatures of malware is investigated and number of infected nodes is minimized. Encounter based distribution algorithm is used to develop the optimized system welfare utility through the signature location. Through extensive simulations with both synthetic and real mobility traces. It achieves the optimal solution.

Based on communication media, mobile malware has 2 types:

- BT based malware
- SMS based malware.

Bluetooth based malware:

A BT based malware is able to infect its geographical neighbors with same Operating System. It has a wide range of applications, such as

wireless headsets, dial-up networking, and peer to peer file sharing. Bluetooth virus can infect all Bluetooth activated phones within a distance from 10 to 30m, resulting in a spatially localized spreading pattern[2].

When a phone is infected with Bluetooth based malware, it automatically turns on the Bluetooth service of itself. Then, the infected phone arbitrarily picks out a susceptible phone as its target. Susceptible phones are the mobile devices in the vicinity of infected mobile phone having its Bluetooth on. The mobile malware spreads out from the infected device to such susceptible phones present in its Bluetooth range. The market for Bluetooth devices has been growing tremendously in recent years.

SMS based malware:

SMS based mobile malware propagation does not have a geographical boundary restrictions[3]. The SMS virus can send a copy of itself to all mobile phones whose numbers are found in the infected phone's address book, a long range spreading pattern. A SMS based malware can propagate from one device to other devices which are millions of kilometers farther. So this may create havoc. Operational behavior of users plays an important role in malware propagation.

Once a phone becomes infected with an SMS virus, after certain time it sends a copy of itself to each mobile phone number found in the handset's phone book[5]. It is done with the list of numbers the handset's user communicated with during a long observational period such as months. SMS virus can reach only a small fraction of users due to the fragmentation of the call graph.

The mobile phones that are within the range of Bluetooth from infected phone gets affected with the Bluetooth virus. The mobility patterns of the mobile user is determined based on which the spread is established. An MMS virus can infect all susceptible phones whose number is found in the infected phone's phonebook, resulting in a long-range spreading pattern that is independent of the infected phone's physical location.

It is found that about 275 new threat families that run on Android and 1 new threat family run each on iPhone and Symbian families based on similarities in the code.

A mobile phone, smart phone or personal digital assistant infected by mobile malware can be a huge inconvenience to a mobile phone user. A compromised phone can cause its user service interruption, financial loss, privacy and confidentiality loss, slowdown of processing speed, unnecessarily huge consumption of memory and loss of battery power.

It is observed that the spreading rate of the mobile malware depends on the market share of the different operating systems. Cybercriminals will, hence, see mobile users as highly profitable targets and will be driven to develop new ways to compromise user data, and potentially breach privacy by tracking individual locations. Mobile malware solutions are in their infancy so their capabilities to protect users and networks are very limited.

SMS-based viruses are more hazardous than BT-based viruses when dealing with the propagation speed and severity. Defending against proximity malware is particularly challenging since it is difficult to piece together global dynamics from just pair-wise device interactions. Traditional network defenses depend upon observing aggregated network. With proximity malware, however, observations are inherently local since they do not involve network infrastructure. Thus the malware detection must begin at the device.

III. Problem Formulation

To design a defense system for both MMS and proximity malware by deploying an efficient defense system to help infected nodes to recover and prevent healthy nodes from further infection. Multivariate correlation analysis, an attack detection system for accurate network traffic characterization is used. It extracts the geometrical correlations between network and traffic features. It employs the principle of anomaly based detection in attack recognition[12] [14]. Triangle area based technique is proposed to enhance and to speed up the process of MCA. The effectiveness of proposed detection system is evaluated, and the influences of both non normalized data and normalized data on the performance of the proposed detection system are examined.

A) Feature generation:

Basic features are generated from ingress network traffic to the internal network where protected servers reside in and are used to form traffic records for a well-defined time interval. Monitoring and analysing at the destination network reduce the overhead of detecting malicious activities by concentrating only on relevant inbound traffic. This also enables our detector to provide protection which is the best fit for the targeted internal network because legitimate traffic profiles used by the detectors are developed for a smaller number of network services.

Algorithm:

```

Node a encounters node b;
For each signature in i's buffer do
    Examine {F1,...FX,...FY};
    if there exists X that satisfies cj(X) > FX then
        if node j never receives this signature
        before then
            Node a duplicates and delegates the
            signature to b;
        end if
        for each device Y do
            Set FY = max{cb(Y), original FY} on nodes
            a and b;
        end for
    end if
end for

```

B) Multivariate Correlation Analysis:

Multivariate correlation analysis, in which the “triangle area map generation” technique is applied to extract the correlations between two distinct features within each traffic record coming from the first step or the traffic record normalized by the “feature normalization” module in this step. The occurrence of network intrusions cause changes to these correlations so that the changes can use as indicators to identify the intrusive activities. All the extracted correlations, namely, triangle areas stored in triangle area maps (TAMs), are then used to replace the original basic features or the normalized features to represent the traffic records. This

provides higher discriminative information to differentiate between legitimate and illegitimate traffic records. The behavioural based detection mechanism is adopted in decision making

C) Attack detection:

The behavioural model extracts important behaviours of normal profiles and stores in normal behaviour database. These normal behaviour profiles are then mapped with tested profiles. This technique further increases the accuracy of attack detection.

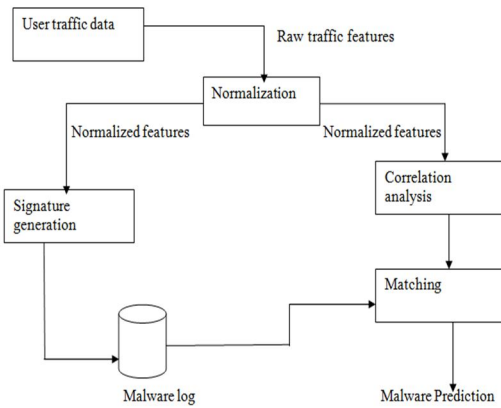


Fig 1: Architecture

D) Decision Making:

Decision making involves two phases

- ✓ ☐ Training phase
- ✓ Test phase

Normal profile generation is work in “Training phase” to generate a profile for individual traffic record and the generated normal profile are stored in a database. In test phase “tested profile generation” are used to built profiles for individual observed traffic records. Then at last the tested profiles are handed over to “Attack Detection” it compares tested profile with stored normal profiles. This module distinguishes the Dos attack from legitimate traffic.

IV. Performance Evaluation

The technique proposed DoS attack detection system with the capability of distinguishing both known and unknown DoS attacks from legitimate network traffic with high accuracy. The network features have been used for

the separation of the legitimate and the attack traffic. The influence of original (non- normalised) and normalised data has been studied and the results have revealed that utilising correlation technique eliminates the bias from the data and boosts up the detection accuracy in terms of detection accuracy and computational complexity.

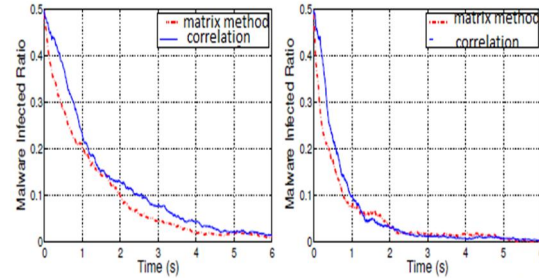


Fig 2: System performance under different methods

For performance evaluation and modeling of mobile malware spreading, the epidemic model, based on the matrix calculation [1] traditionally used in networks, has been extensively used. Actually, the system performance of the epidemic model can be approximated by the Matrix covariance method [9] being widely used to model the denial of service. These works show that when the number of nodes in a network is large, the deterministic models can successfully represent the dynamics of malware spreading, which is demonstrated by simulations and matching with actual data.

V. Conclusion

In this paper, the problem of optimal signature distribution to defend mobile networks against the propagation of both proximity and MMS-based malware is investigated. A distributed algorithm that closely approaches the optimal system performance of a centralized solution is implemented. Through both theoretical analysis and simulations, the efficiency of our defence scheme in reducing the amount of infected nodes in the system is demonstrated. The security and authentication mechanisms of the mobile network are considered. Finally, the defense system with the consideration of denial of service deployed.

VI. References

1. Yong Li, Pan Hui, Member, IEEE, Depeng Jin, Li Su, and Lieguang Zeng, “Optimal Distributed

- Malware Defense in Mobile Networks with Heterogeneous Devices” Member, IEEE, 2014
2. M. Khouzani, S. Sarkar, and E. Altman, “Maximum Damage Malware Attack in Mobile Wireless Networks,” Proc. IEEE Infocom, 2010. 390 Ieee Transactions On Mobile Computing, Vol. 13, No. 2, February 2014
3. Shanghai Jiao Tong Univ., Traffic Information Grid Team, Grid Computing Center, “Shanghai Taxi Trace Data,” <http://wirelesslab.sjtu.edu.cn/>, 2013.
4. L. Hu, J. Boudec, and M. Vojnovic, “Optimal Channel Choice for Collaborative Ad-Hoc Dissemination,” Proc. IEEE INFOCOM, 2010.
5. S. Ioannidis, L. Massoulie’, and A. Chaintreau, “Distributed Caching over Heterogeneous Mobile Networks,” Proc. ACM Int’l Conf. Measurement and Modeling of Computer Systems (SIGMETRICS ’10), 2010.
6. T. Ning, Z. Yang, and H. Wu, “Counting in Delay-Tolerant Mobile Networks,” Proc. IEEE Int’l Conf. Comm. (ICC), pp. 1-5, 2010.
7. E. Altman, G. Neglia, F. De Pellegrini, and D. Miorandi, “Decentralized Stochastic Control of Delay Tolerant Networks,” Proc. IEEE INFOCOM, 2009.
8. A. Keranen, J. Ott, and T. Karkkainen, “The ONE Simulator for DTN Protocol Evaluation,” Proc. Second Int’l Conf. Simulation Tools and techniques, 2009
9. G. Lawton, “On the Trail of the Conficker Worm,” Computer, vol. 42, no. 6, pp. 19-22, June 2009.
10. F. Li, Y. Yang, and J. Wu, “CPMC: An Efficient Proximity Malware Coping Scheme in Smartphone-Based Mobile Networks,” Proc. IEEE INFOCOM, 2009.
11. A. Mei and J. Stefa, “SWIM: A Simple Model to Generate Small Mobile Worlds,” Proc. IEEE INFOCOM, pp. 2106-2113, 2009.
12. P. Wang, M. Gonzalez, C. Hidalgo, and A. Barabasi, “Understanding the Spreading Patterns of Mobile Phone Viruses,” Science, vol. 324, no. 5930, pp. 1071-1076, 2009.
13. Z. Zhu, G. Cao, S. Zhu, S. Ranjan, and A. Nucci, “A Social Network Based Patching Scheme for Worm containment in Cellular Networks,” Proc. IEEE INFOCOM, 2009.
14. G. Zyba, G. Voelker, M. Liljenstam, A. Mehes, and P. Johansson, “Defending Mobile Phones from Proximity Malware,” Proc. IEEE INFOCOM, 2009.
15. P. Hui, J. Crowcroft, and E. Yoneki, “Bubble Rap: Social-Based Forwarding in Delay Tolerant Networks,” Proc. ACM MobiHoc, 2008.
16. M. Hypponen, “Mobile Malwar,” Proc. 16th USENIX Security Symp., 2007.
17. J. Kumpula, J. Onnela, J. Sarama’ki, K. Kaski, and J. Kerte’sz, “Emergence of Communities in Weighted Networks,” Physical Rev. Letters, vol. 99, no. 22, p. 228701, 2007.
18. M. Grossglauser and D. Tse, “Mobility Increases The Capacity of Ad-Hoc Wireless Networks,” Proc. IEEE INFOCOM, pp. 1360- 1369, 2001.
19. R. May and A. Lloyd, “Infection Dynamics on Scale-Free Networks,” Physical Rev. E, vol. 64, no. 6, p. 066112, 2001.
20. P. Brémaud, Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues. Springer Verlag, 1999.