IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

MALWARE PROPAGATION IN LARGE SCALE NETWORKS

[1]Mastanvali, [2]Ch.Sai Preetham Reddy, [3]K.Bhavana, [4]K.Santhoshi Kumari

[1]Assistant professor, [2]Student, [3]Student, [4]Student

Department of Information Technology, JB Institute of Engineering and Technology(JBIET), Hyderabad, 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 project, 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.

***Index Terms*: Malware, Propagation, Epidemic Model, Power Law, Exponential Distribution.**

1.INTRODUTION

Malware, short for malicious software, is any software used to disrupt computer operations, gather sensitive information, gain access to private computer systems, or display unwanted advertising. Malware is defined by its malicious intent, acting against the requirements of the computer user, and does not include software that causes unintentional harm due to some deficiency. The term badware is sometimes used, and applied to both true (malicious) malware and unintentionally harmful software.

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. This paper describes the distribution of malware in terms of networks (e.g., autonomous systems (AS), ISP domains, abstract networks of smart phones 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. Second, 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.

2.LITERATURE 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.

Self-propagating codes, called worms, such as Code Red, Nimda, and Slammer, have drawn significant attention due to their enormously adverse impact on the Internet. Thus, there is great interest in the research community in modeling the spread of worms and in providing adequate defense mechanisms against them. In this paper, we present a (stochastic) branching process model for characterizing the propagation of Internet worms. The model is developed for uniform scanning worms and then extended to preference scanning worms. This model leads to the development of an containment strategy that prevents the spread of a worm beyond its early stage. Specifically, for uniform scanning worms,

we are able to 1) provide a precise condition that determines whether the worm spread will eventually stop and 2) obtain the distribution of the total number of hosts that the worm infects. We then extend our results to contain preference scanning worms. Our strategy is based on limiting the number of scans to dark-address space. The limiting value is determined by our analysis. Our automatic worm containment schemes effectively contain both uniform scanning worms and local preference scanning worms, and it is validated through simulations and real trace data to be nonintrusive.

While multi-hop broadcast protocols, such as Trickle, Deluge and MNP, have gained tremendous popularity as a means for fast and convenient propagation of data/code in large scale wireless sensor networks, they can, unfortunately, serve as potential platforms for virus spreading if the security is breached. To understand the vulnerability of such protocols and design defense mechanisms against piggy- backed virus attacks, it is critical to investigate the propagation process of these protocols in terms of their speed and reachability. In this paper, we propose a general framework based on the principles of epidemic theory, for vulnerability analysis of current broadcast protocols in wireless sensor networks. In particular, we develop a common mathematical model for the propagation that incorporates important parameters derived from the communication patterns of the protocol under test. Based on this model, we analyze the propagation rate and the extent of spread of a malware over typical broadcast protocols proposed in the literature. The overall result is an approximate but convenient tool to characterize a broadcast protocol in terms of its vulnerability to malware propagation.

Conficker is the most recent widespread, well-known worm/bot. According to several reports, it has infected about 7 million to 15 million hosts and the victims are still increasing even now. In this paper, we analyze Conficker infections at a large scale, about 25 million victims, and study various interesting aspects about this state-of-the-art malware. By analyzing Conficker, we intend to understand current and new trends in malware propagation, which could be very helpful in predicting future malware trends and providing insights for future malware defense. We observe that Conficker has some very different victim distribution patterns compared to many previous generation worms/botnets, suggesting that new malware spreading models and defense strategies are likely needed. We measure the potential power of Conficker to estimate its effects on the networks/hosts when it performs malicious operations. Furthermore, we intend to determine how well a reputation-based blacklisting approach can perform when faced with new malware threats such as Conficker. We cross-check several DNS blacklists and IP/AS reputation data from Dshield and FIRE and our evaluation shows that unlike a previous study which shows that a blacklist-based approach can detect most bots, these reputation-based approaches did relatively poorly for Conficker. This raises a question of how we can improve and complement existing reputation-based techniques to prepare for future malware defense? Based on this, we look into some insights for defenders. We show that neighborhood watch is a surprisingly effective approach in the case of Conficker.

3.1 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.

3.2 DISADVANTAGES:

One critical condition for the epidemic models is a large vulnerable population because their principle is based on differential equations.

3.3 PROPOSED SYSTEM:

In proposed system when input image is given we are extracting information from given image and training data in the way as information location is extracted from image and features are compared using API and accurate location information and land mark information is displayed.

3.4 ADVANTAGES:

Automatic extracting of location and land mark form image which helps in searching related images and related data.
Results are accurate for most of the images.

4. SYSTEM DESIGN

Fig 4.1 System Architecture(User)

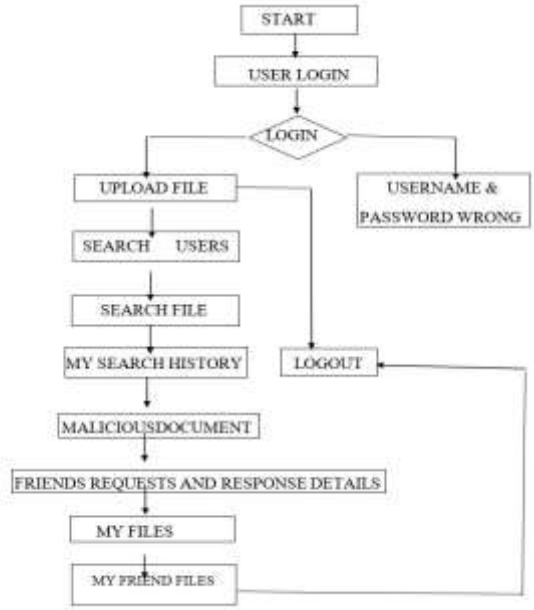

Fig 4.2 System Architecture(Admin)

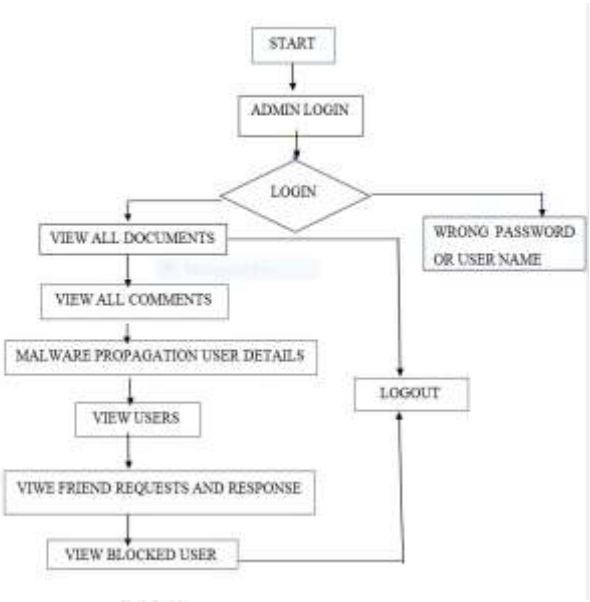

Fig 4.3 Class Diagram

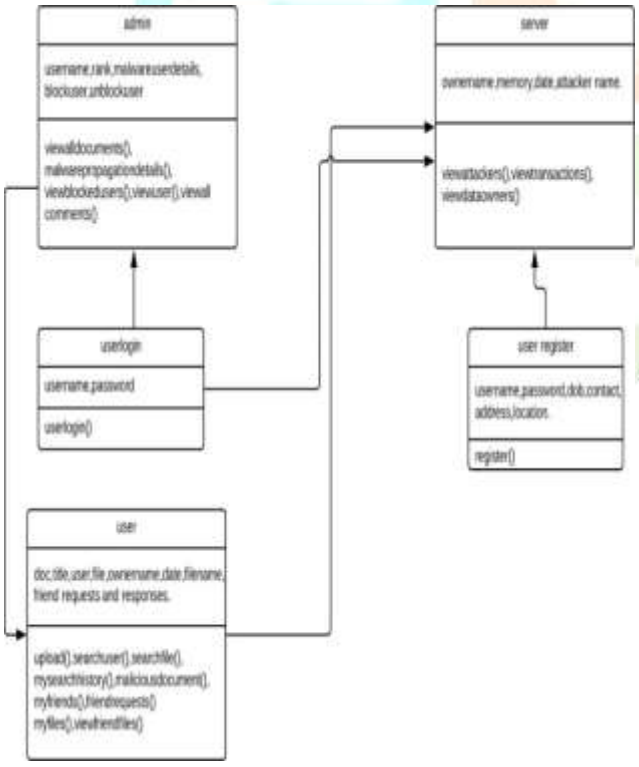

Fig 4.4 Activity Diagram

5.SAMPLE CODE:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<%@ include file="connect.jsp" %>
<%@page import="java.util.*,java.security.Key,java.util.Random,javax.crypto.Cipher,javax.crypto.spec.SecretKeySpec,org.bouncycastle.util.encoders.Base64"%>
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title>Malware Propagation in Large-Scale Networks</title>
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
<link href="css/style.css" rel="stylesheet" type="text/css" />
<link rel="stylesheet" type="text/css" href="css/coin-slider.css" />
<script type="text/javascript" src="js/cufon-yui.js"></script>
<script type="text/javascript" src="js/cufon-titillium-400.js"></script>
<script type="text/javascript" src="js/jquery-1.4.2.min.js"></script>
<script type="text/javascript" src="js/script.js"></script>
<script type="text/javascript" src="js/coin-slider.min.js"></script>
<style type="text/css">
<!--
.style4 {color: #0066FF}
.style5 {color: #FF0000}
.style6 {color: #CC0000}
.style7 {
color: #FF0033; font-weight: bold;
}
.style8 {font-size: 36px}
-->

</style>
</head>
<body>
<div class="main">
<div class="header">
<div class="header_resize">
<div class="logo">
<h1><a href="index.html">Malware Propagation in Large-Scale Networks<span></span></a></h1>
</div>
```

6.MODULES

i. **User module:** In this module, the user browses the required file and uploads to the Social network to share with their friends. The data provider also perform the following operations such as show in the below.

ii. **Add Document:** In this module, the user can add the document. If user wants to add the new document, then he will enter document name, enter a document title, and so on then submit and that data will stored in data base.

iii. **View documents:** In this module, the user can view the document details i.e, document name, document title, document image and document content, related images

iv. **Search users:** In this module, the user can search for the other user he is looking for.

v. **Search files:** In this module, the user can search for a particular file he is looking for, if the file is found in the network, it will display the file name and description.

vi. **View my search history:** In this module, the user can view his search history i.e. it shows what the user has searched for recently.

vii. **View my friends:** In this module the user can view all his friend list.

viii. **View friend requests:** In this module the user can see all his friend requests till date.

ix. **View my files:** In this module the user can see all the files uploaded by him up to the time of checking.

x. **View my friends files:** In this module the user can look at all the files of his friends.

xi. **Admin:** The Admin is responsible for performing some operations like to analyzing documents and contents to check whether the document contains malware. If documents are malware related then those documents will be scanned and Finds malicious users those who propagate malware in the social networks and block from the social networks and they will be keeping in the block list of the social networks.

xii. **View all documents:** In this module, the admin can see all the documents or files which have been uploaded by all the users of the network.

xiii. **View all comments:** In this module, the admin can see all the comments on the files done by the users of the network.

xiv. **View malicious user details :** In this module, the admin can see all the users who are responsible for uploading malicious files in to the network.

xv. **View friend requests/responses:** In this module, the admin can see all the details about which user has sent a friend request to whom? And who has accepted the friend requests.

xvi. **View blocked user:** In this module, the admin can see all the details about the blocked users of the network.

7. CONCLUSION

In this project, 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 modelling 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 modelling. Moreover, the proposed two layer model offers us the distribution of malware in terms of the low layer networks.

In regards to future work, we will first further investigate the dynamics of the late stage. More details of the findings are expected to be further studied, such as the length of the exponential tail of a power law distribution at the late stage. Second, defenders may care more about their own network, e.g., the distribution of a given malware at their ISP domains, where the conditions for the two layer model may not hold. We need to seek appropriate models to address this problem. Finally, we are interested in studying the distribution of multiple malware on large-scale networks as we only focus on one malware in this paper. We believe it is not a simple linear relationship in the multiple malware case compared to the single malware one.

8. 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 Proc. ACM Conf. Comput. Commun. Security, 2009, pp. 635–647.

[2] 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 Proc. 1st Conf. 1st Workshop Hot Topics Understanding Botnets, 2007, p. 5.

[3] D. Dagon, C. Zou, and W. Lee, "Modeling botnet propagation using time zones," in Proc. 13th Netw. Distrib. Syst. Security Symp., 2006.