Applied Mathematical Modelling 63 (2018) 491–507

Contents lists available at ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

Modeling and analyzing the dynamic spreading of epidemic malware by a network eigenvalue method

Wanping Liu[a,b,*], Shouming Zhong[b]

[a] College of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China
[b] School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China

ARTICLE INFO

Article history:
Received 5 October 2017
Revised 13 June 2018
Accepted 4 July 2018
Available online 12 July 2018

Keywords:
Malware spreading
Stability analysis
Network eigenvalue
Propagation threshold

ABSTRACT

This paper mainly focuses on studying the influence of network characteristics on malware spreading. Firstly, a generalized model with weakly-protected and strongly-protected susceptible nodes is developed by considering the possibility of an intruded node converting back to a weakly-protected susceptible one. The dynamics of the generalized compartmental model is intensively discussed and analyzed, deriving several sufficient conditions for its global stability. Following this work, a novel node-based model is newly proposed to describe malware propagation over an arbitrary connected network including synthesized and real networks. From a microscopic perspective, we establish the novel model by introducing several different variables for each node which describe the probabilities of a node locating at respective states. Our theoretical analysis shows that the largest eigenvalue of the propagating network is a key factor determining malware prevalence. Specifically, the range of the leading eigenvalue can be split into three subintervals in which malware approaches extinction very quickly, or tends to extinction, or persists, depending on into which subinterval the largest eigenvalue of the propagating network falls. Theoretically, the trivial equilibrium of our new node-based model is clearly proved to be exponentially globally stable when the maximum eigenvalue is less than a threshold. We also illustrate the predictive effectiveness of our model by designing some numerical simulations on some regular and scale-free networks. Consequently, we conclude that malware prevalence can be effectively prevented by properly adjusting the spreading network, e.g., reducing the number of nodes and deleting some edges, so that its maximum eigenvalue falls into the appropriate subinterval.

1. Introduction

Although a number of new defense techniques are constantly developed to detect and prevent malware attacks since the appearance of computer viruses, the world of malicious code does not fade away, but is dramatically evolving in terms of diversity and scale. Besides computer viruses and Internet worms, a variety of new types of malicious threats spring up in the past few years, such as spam, phishing and Trojans. For instance, more than 430 million new unique pieces of malware were discovered by Symantec in 2015, which is an increase of 36 percent from the year before [1]. Especially, a remarkable

* Corresponding author at: College of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China.
E-mail address: lwphe@163.com (W. Liu).

https://doi.org/10.1016/j.apm.2018.07.010
0307-904X/

phenomenon is that people may be no longer surprised by these statistical numbers, since cybercrime has already become a part of our daily lives.

As an important addition to traditional anti-malware technology, the research field of malware dynamics aims to understand the propagating mechanisms of malicious threats over networks so as to suggest some proper strategies for controlling malware diffusion [2–5]. Motivated by the epidemic models in the field of infectious diseases, Kephart and White [6] originally applied the mean-field theory to study the modeling of computer virus spreading in the early 1990s. Following their work, a number of compartmental models for computer virus propagation have been developed, such as the Susceptible-Infected-Recovered (SIR) models and the Susceptible-Exposed-Infected-Recovered-Susceptible (SEIRS) models, most of which are applicable to both epidemic diseases and malicious software [7,8]. However, some features of malware are actually different from most infectious diseases. Thus, in recent years, some models considering the unique characteristics of epidemic malware are proposed [9–13]. For instance, the malware transition between different states is very similar to the fault detection process, correction process and introduction process during software testing [14]. In order to model and compute the service reliability for the distributed computing (DC) system under virus epidemics, Li and Peng [15] studied the spreading of viruses in DC systems by using differential equations. Especially, Liu et al. [16] recently considered that each node over the network can be in a state of weakly-protected susceptible, strongly-protected susceptible, or infected, and then proposed a compartment-based model (referred as WSIS model in the sequel) by incorporating the heterogenous immunization of terminal devices. In this model, all nodes of the propagating network are categorized into three groups according to their states. Generally, compartment-based models, which mainly focus on the evolution of the number or fraction of nodes in each compartment, are developed based on a fundamental assumption of the spreading network being homogeneously hybrid. However, most real networks from nature and society, e.g., the Internet and social networks, have been empirically found to be highly structured rather than simply homogeneously [17,18]. Hence, the WSIS model was further discussed based on an assumption that the spreading network admits a prescribed scale-free degree distribution, deriving a new network-based WSIS model for epidemic malware propagation [19].

Note that the compartmental models mentioned above suffer from a common defect that the structure knowledge of the propagation network is not fully considered. As a result, it is extremely difficult to deeply understand the impact of the network topology on the viral prevalence by solely studying such compartment-based models. Wang et al. [20] proposed a discrete-time model, of which the major merit is the incorporation of an arbitrary network characterized by an adjacency matrix denoted by $\mathbf{A}$. This model extends the homogeneous Kephart and White model [6], where the only network characteristic was the average degree. By an approximate analysis, they conclude a major and intriguing result that the epidemic threshold of the effective infection rate is specified by $1/\lambda_{\max}(\mathbf{A})$, where $\lambda_{\max}(\mathbf{A})$ represents the largest eigenvalue of the adjacency matrix. As far as we know, this result first relates the malware spreading to a specific characteristic, i.e., the spectral radius of the propagating network. Later, by applying the mean field theory, Van Mieghem et al. [21] strictly addresses the existence of a well-defined threshold, i.e., the inverse of largest eigenvalue of the adjacent matrix. This result identifies the belief that scale-free networks like the Internet or World Wide Web possess an extremely small epidemic threshold and, thus, are vulnerable to malware. This finding has promoted massive research on immunization strategies for scale-free (complex) networks [22–24].

Motivated by previous studies, both diverse immunization against malware and the network structure are found to have significant impacts on viral spreading, respectively. In order to consider their combined influence on malware prevalence, a node-based dynamical model is newly developed and analyzed in this paper. Specifically, our major motivation is to understand the influence of network characteristics on malware spreading in the sense of the WSIS model. It is found by theoretical analysis that the dynamics of the model, especially the global stability of trivial equilibrium, considerably depend on the maximum eigenvalue of the adjacent matrix of the spreading network. Some numerical simulations are also designed to illustrate the main results, implying that the largest network eigenvalue plays a key role in controlling malware spreading.

2. Motivation and compartments

In reality, the security awareness of users has significant influence on the immunization of terminal devices against malicious threats. However, due to the uncertainty of user's behavior, different users actually possess distinct levels of security consciousness. Specifically, compared with users with lower security awareness, vigilant users who have stronger security awareness will take certain measures to enhance the protection level of their computers, such as newly installing effective anti-malware or regularly updating security products. Thereby, computers of users with high security consciousness will exhibit greater ability to resist malware invasion. On the contrary, the devices used by people with lower security awareness have weaker defences against malicious attacks. Based on this consideration, susceptible nodes are further divided into two groups: weakly-protected and strongly-protected. For clarity, we define herein that

(1) A susceptible node (computer) is *weakly-protected* if the node is installed with expired anti-malware or even naked in terms of security protection (not protected by any security products).
(2) A susceptible node (computer) is *strongly-protected* if the node is equipped with effective real-time security products.
(3) A node (computer) is *intruded* if the node has already suffered from a malware intrusion and has derived the ability of infecting other neighboring susceptible nodes.