Applied Mathematical Modelling 43 (2017) 110–125

Contents lists available at ScienceDirect

Applied Mathematical Modelling

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

The impact of patch forwarding on the prevalence of computer virus: A theoretical assessment approach

Lu-Xing Yang[a,b,*], Xiaofan Yang[a], Yingbo Wu[a]

[a] School of Software Engineering, Chongqing University, Chongqing 400044, China
[b] Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, GA, Delft 2600, The Netherlands

ARTICLE INFO

Article history:
Received 14 January 2016
Revised 31 August 2016
Accepted 11 October 2016
Available online 2 November 2016

Keywords:
Computer virus
Virus patch
Node-level epidemic model
Equilibrium
Global stability
Spectral radius

ABSTRACT

Virus patches can be disseminated rapidly through computer networks and take effect as soon as they have been installed, which significantly enhances their virus-containing capability. This paper aims to theoretically assess the impact of patch forwarding on the prevalence of computer virus. For that purpose, a new malware epidemic model, which takes into full account the influence of patch forwarding, is proposed. The dynamics of the model is revealed. Specifically, besides the permanent susceptible equilibrium, this model may admit an infected or a patched or a mixed equilibrium. Criteria for the global stability of the four equilibria are given, respectively, accompanied with numerical examples. The obtained results show that the spectral radii of the patch-forwarding network and the virus-spreading network both have a marked impact on the prevalence of computer virus. The influence of some key factors on the prevalence of virus is also revealed. Based on these findings, some strategies of containing electronic virus are recommended.

1. Introduction

The malware epidemic dynamics has been recognized as an effective approach to the assessment of prevalence of computer virus [1]. Since Kephart and White's seminal work [2], a multitude of malware epidemic models, ranging from simple models [3–10] to advanced models such as delayed models [11,12], impulsive models [13–15] and stochastic models [16,17], have been proposed.

Patches targeting malware can be disseminated through technological networks (such as the Internet, the world-wide web, and online social networks) to a large fraction of network nodes at an extremely high velocity. Moreover, the patch dissemination can be carried out in a distributed way (every node that was newly patched tries to forward the patches to all of its neighbors), so as to reduce the overhead of network resources. For the purpose of theoretically assessing the effectiveness of the patch forwarding strategy, Yang and Yang [18,19] suggested two malware epidemic models taking into account the influence of patch forwarding. As the models assume that every node is either susceptible or infected or patched, they are referred to as the Susceptible-Infected-Patched models, abbreviated as the SIPS models. The distinction between the two models lies in that the first model ignores the influence of infected removable storage media, whereas the second model considers that influence. The two SIPS models are both compartmental, that is, all nodes are grouped into three compartments according to their current states, and the major concern is the change in the fraction of each

* Corresponding author at: School of Software Engineering, Chongqing University, Chongqing 400044, China.
E-mail addresses: ylx910920@gmail.com (L.-X. Yang), xfyang1964@gmail.com (X. Yang), wyb@cqu.edu.cn (Y. Wu).

http://dx.doi.org/10.1016/j.apm.2016.10.028
0307-904X/

compartment. As these SIPS models cannot accommodate the complete information concerning the network structures, in most situations the effectiveness of the patch-forwarding strategy cannot be accurately assessed by studying them. For related work on this topic, see Refs. [20–24].

What node-level epidemic models mean are epidemic models that accommodate the probability of every node being in a state. As node-level epidemic models can accommodate the full knowledge concerning the network topology, the impact of the network topology on the prevalence of virus can be revealed by studying such models. In 2009, Mieghem et al. [25] introduced the first node-level epidemic model by remoulding a traditional SIS model. Later, Xu et al. [26] proposed a node-level SIR model capturing the dynamics of multivirus. By introducing the added alert state, Sahneh and Scoglio [27] established a node-level SAIS model. It was found under these epidemic models that the spectral radius of the virus-spreading network plays a key role in determining the virus prevalence [25–27]. For more information on this topic, see Refs. [28–32]. In the context of patch forwarding, viruses propagate through the virus-spreading network, whereas patches are disseminated through the patch-forwarding network, and the two networks may be different. Consequently, it is of practical importance to study the combined impact of the patch-forwarding network and the virus-spreading network on the viral prevalence. To our knowledge, however, there is yet no report in literature on this topic.

This paper addresses the theoretical assessment of the patch forwarding strategy. For that purpose, a node-level SIPS model, which takes into full account the influence of the patch forwarding, is proposed. The dynamics of the model is revealed. Specifically, besides the permanent susceptible equilibrium, this model may admit an infected or a patched or a mixed equilibrium. Criteria for the global stability of the four equilibria are given, respectively, accompanied with numerical examples. The obtained results show that the spectral radii of the patch-forwarding network and the virus-spreading network both have a marked impact on the prevalence of computer virus. The influence of some key factors on the prevalence of virus is also revealed. Based on these findings, some virus-containing policies are recommended.

The remaining materials of this paper are organized in this pattern: Section 2 formulates the new malware epidemic model. Section 3 theoretically analyzes the proposed model, and Section 4 illustrates the obtained results. Section 5 examines the impact of some factors on the virus prevalence and thereby draws some new insights on containing virus spreading. Finally, Section 6 summarizes this work and points out some directions of research.

2. Formation of a node-level SIPS model

Consider a networked system of N nodes labeled $1, 2, \ldots, N$. Let $V = \{1, 2, \ldots, N\}$. Let $G_v = (V, E_v)$ denote the virus-propagating network, where two nodes are adjacent if and only if computer viruses can propagate directly from one of them to the other. Let $\mathbf{A} = [a_{ij}]_{N\times N}$ denote the adjacency matrix of G_v, and let $\rho(\mathbf{A})$ denote the spectral radius of $\mathbf{A}$, which equals its maximum eigenvalue. Let $G_p = (V, E_p)$ denote the patch-forwarding network, where two nodes are adjacent if and only if patches can be forwarded directly from one of them to the other. Let $\mathbf{B} = [b_{ij}]_{N\times N}$ denote the adjacency matrix of G_p, and let $\rho(\mathbf{B})$ denote the spectral radius of $\mathbf{B}$, which equals its maximum eigenvalue. In what follows, it is always assumed that the virus-spreading network and the patch-forwarding network are both connected.

As with the compartmental SIPS models [18,19], it is assumed in the new model that at any time, each and every node in the system is in one of three states: *susceptible, infected*, and *patched*. A node is susceptible if it is uninfected and with no newest patches. Hence, a susceptible node can be infected by an infected G_v-neighbor or patched by a patched G_p-neighbor. In contrast, a node is patched if it is not only uninfected but with newest patches. As a result, a patched node cannot be infected by an infected G_v-neighbor. Finally, an infected node is with no newest patches and hence can be patched by a patched G_p-neighbor.

Let $S_i(t)$, $I_i(t)$, and $P_i(t)$ denote the probability that at time t, node i is susceptible, infected, and patched, respectively. Clearly, the vector

$$\widetilde{\mathbf{x}}(t) = (S_1(t), \ldots, S_N(t), I_1(t), \ldots, I_N(t), P_1(t), \ldots, P_N(t))^T,$$

probabilistically captures the t-time state of the system. Let

$$\widetilde{\Omega} = \{(S_1, \ldots, S_N, I_1, \ldots, I_N, P_1, \ldots, P_N)^T \in \mathbb{R}_+^{3N} \mid S_i + I_i + P_i = 1, i = 1, 2, \ldots, N\}.$$

Then $\widetilde{\mathbf{x}}(t) \in \widetilde{\Omega}$ for all $t \geq 0$.

As $S_i(t) + I_i(t) + P_i(t) \equiv 1, 1 \leq i \leq N$, the vector

$$\mathbf{x}(t) = (I_1(t), \ldots, I_N(t), P_1(t), \ldots, P_N(t))^T,$$

also probabilistically captures the t-time state of the system. Let

$$\Omega = \{(I_1, \ldots, I_N, P_1, \ldots, P_N)^T \in \mathbb{R}_+^{2N} \mid I_i + P_i \leq 1, i = 1, 2, \ldots, N\}.$$

Then $\mathbf{x}(t) \in \Omega$ for all $t \geq 0$. In what follows, let $\mathring{\Omega}$ and $\partial\Omega$ denote the interior and boundary of Ω, respectively.

Now, let us impose a set of statistical assumptions on the state transition of a node as follows.

(H1) Due to the propagation of viruses, a susceptible node is infected by an infected G_v-neighbor at a constant rate $\beta > 0$. At the early stage of invasion of viruses, a susceptible node i gets infected at time t approximately at rate $\beta \Sigma_j a_{ij} I_j(t)$.