Commun Nonlinear Sci Numer Simulat 19 (2014) 1935–1944

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

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

A new epidemic model of computer viruses

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

[a] School of Electronic and Information Engineering, Southwest University, Chongqing 400715, China
[b] College of Computer Science, Chongqing University, Chongqing 400044, China

ARTICLE INFO

Article history:
Received 22 September 2012
Received in revised form 5 May 2013
Accepted 30 September 2013
Available online 8 October 2013

Keywords:
Computer virus
Epidemic model
Endemic equilibrium
Global stability
Asymptotically autonomous system

ABSTRACT

This paper addresses the epidemiological modeling of computer viruses. By incorporating the effect of removable storage media, considering the possibility of connecting infected computers to the Internet, and removing the conservative restriction on the total number of computers connected to the Internet, a new epidemic model is proposed. Unlike most previous models, the proposed model has no virus-free equilibrium and has a unique endemic equilibrium. With the aid of the theory of asymptotically autonomous systems as well as the generalized Poincare–Bendixson theorem, the endemic equilibrium is shown to be globally asymptotically stable. By analyzing the influence of different system parameters on the steady number of infected computers, a collection of policies is recommended to prohibit the virus prevalence.

1. Introduction

As a class of malicious programs that can replicate themselves and spread among computers, computer viruses have formed a great threat to human society; during the breaking-out period, a computer virus can perform devastating operations, even destroying the whole computer system [1]. Antivirus software is currently the major means of defending against viruses. Yet, the development of new antivirus software always lags behind the emergence of new viruses and, hence, cannot help laying down long-term policies of containing viruses. As a promising supplement to the conventional antivirus technique, the epidemic dynamics of computer viruses aims to understand the way that computer viruses spread across a network and to work out global policies of inhibiting their prevalence. Inspired by the compelling analogies between computer viruses and their biological counterparts, Cohen [2] and Murray [3] inventively suggested to exploit the compartment modeling techniques developed in the epidemic dynamics of biologically infectious diseases to study the spread of computer viruses. Later, Kephart and White [4] borrowed a biological epidemic model (the SIS model) to investigate the virus spreading on the Internet. The relevant researches have since been conducted mainly along two distinct ways, which are reviewed below.

One decade ago, it was found that the Internet follows diverse power-law degree distributions [5–7]. This finding has stimulated the interest in virus spreading on complex networks, leading to the surprising finding that the epidemic threshold vanishes for scale-free networks with infinite size [10]. Previous work, however, was limited mainly to three simplistic epidemic models: the SI model [8,9], the SIS model [10–20], and the SIR model [18,20–23], and the global stability of the endemic equilibrium, if present, was examined mainly experimentally. Although Pastor-Satorras and Vespignani [10] pointed out that it is worthwhile to study more rational epidemic models, less work has been reported in the literature.

* Corresponding author. Tel./fax: +86 23 65342815.
E-mail addresses: ylx910920@gmail.com (L.-X. Yang), xfyang1964@gmail.com (X. Yang).

1007-5704/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cnsns.2013.09.038

Due to the desire to fully understand the spreading mechanism of computer viruses, on the other hand, multifarious virus spreading models, ranging from conventional models to unconventional models such as delayed models, impulsive models, and stochastic models, have been proposed. For ease in treatment, all of these models adopt the homogeneously mixed assumption, i.e., every computer in the Internet is equally likely to be accessed by any other computer in the network [24–45]. Previous work was focused mainly on the theoretical study of dynamical properties of these models, ranging from the global stability of equilibria to the occurrence of a variety of bifurcations.

Recently, some defects of previous epidemic models were reported [46]. First, those models with the exposed compartment overlook the fact that a computer can infect other computers immediately after it gets infected. Second, those models having all infected computers in a single compartment neglect the marked difference between latent computers and breaking-out computers. Third, those models having a permanent immunity compartment fail to consider the possibility that a computer having been cured previously is prone to infection by new variants of old viruses or new virus strands. Fourth, the majority of these models ignore the effect of removable storage media. Finally, all of these models assumes that a computer is uninfected when connected to the Internet, which is inconsistent to the actual conditions. To surmount these defects, very recently a series of virus spreading models were proposed, all of which distinguish latent computers from breaking-out computers by introducing the L and B compartments [46–52]. Furthermore, the model presented in [49] assumes that an infected computer is likely to be connected to the Internet, whereas the model given in [50] takes into consideration the effect of removable media. Although each of these models can overcome part of the five demerits mentioned above, neither of them can conquer all of these flaws simultaneously. Besides, all of these models assume that the number of computers connected to the Internet keeps constant, in disagreement with the reality that there is a significant fluctuation in this number.

This paper is intended to develop an epidemic model of computer viruses that can get over all of the above-mentioned flaws of previous models. By incorporating the effect of removable media, considering the possibility of connecting infected computers to the Internet, and removing the conservative condition on the total number of computers connected to the Internet, simultaneously, a novel epidemic model is presented. The proposed model is proved to have no virus-free equilibrium and have a unique endemic equilibrium. With the aid of the theory of asymptotically autonomous systems as well as the generalized Poincaré–Bendixson theorem, the endemic equilibrium of the proposed model is shown to be globally asymptotically stable. By analyzing the influence of different system parameters on the steady number of infected computers, some policies are suggested to place a curb on the virus prevalence.

The remaining materials of this paper are organized in this fashion: Section 2 elaborates the new model. The global stability of the endemic equilibrium for the new model is proved in Section 3. Section 4 analyzes the influence of system parameters. This work is summarized in Section 5.

2. Model formulation

For brevity, computers are called as *nodes*. A node is *internal* or *external* depending on whether it is currently connected to the Internet or not. A node is *infected* or *uninfected* depending on whether it contains viruses or not. An infected node is referred to as *latent* or *breaking-out* depending on whether the viruses in it are all latent or at least one virus in it is breaking out. All internal nodes worldwide are classified as uninfected internal nodes (*S-nodes*), latent internal nodes (*L-nodes*), and breaking-out internal nodes (*B-nodes*). Likewise, all external nodes worldwide are classified as uninfected external nodes (*S∗-nodes*), latent external nodes (*L∗-nodes*), and breaking-out external nodes (*B∗-nodes*). Let $S(t), L(t)$, and $B(t)$ denote the numbers of S-, L-, and B-nodes at time t, respectively. Without loss of ambiguity, they will be abbreviated as S, L, and B, respectively.

Our model is based on the following collection of hypotheses.

(H1) Every external node is either virus-free or latent when it is being connected to the Internet. Moreover, S∗-nodes are connected to the Internet at constant rate $\mu_1 > 0$, and L∗-nodes are connected to the Internet at constant rate $\mu_2 > 0$. Let $\mu = \mu_1 + \mu_2$.
(H2) Every internal node is disconnected from the Internet with constant probability $\delta > 0$.
(H3) Due to the influence of infected removable storage media, every S-node is infected with constant probability $\theta > 0$.
(H4) At time t, every S-node is infected by L-nodes with probability $\beta_1 L(t)$, while every S-node is infected by B-nodes with probability $\beta_2 B(t)$, where β_1 and β_2 are positive constants.
(H5) Every L-node breaks out with constant probability $\alpha > 0$.
(H6) Every L-node is cured with constant probability $\gamma_1 > 0$, while every B-node is cured with constant probability $\gamma_2 > 0$.

The mean-field rate equations based on these hypotheses are formulated as

$$\begin{cases} \dot{S} = \mu_1 - (\beta_1 L + \beta_2 B)S + \gamma_1 L + \gamma_2 B - (\delta + \theta)S, \\ \dot{L} = \mu_2 + (\beta_1 L + \beta_2 B)S - (\gamma_1 + \alpha + \delta)L + \theta S, \\ \dot{B} = \alpha L - (\gamma_2 + \delta)B, \end{cases} \tag{1}$$

with initial conditions $(S(0), L(0), B(0)) \in R^3_+$. It is easily verified that R^3_+ is positively invariant for the system.