The Impact of the Network Topology on the Viral Prevalence: A Node-Based Approach

Lu-Xing Yang¹, Moez Draief², Xiaofan Yang³*

¹ College of Computer Science, Chongqing University, Chongqing, China, ² Department of Electrical and Electronic Engineering, Imperial College London, South Kensington Campus, London, United Kingdom, ³ School of Software Engineering, Chongqing University, Chongqing, China

* xfyang1964@gmail.com

Abstract

This paper addresses the impact of the structure of the viral propagation network on the viral prevalence. For that purpose, a new epidemic model of computer virus, known as the node-based SLBS model, is proposed. Our analysis shows that the maximum eigenvalue of the underlying network is a key factor determining the viral prevalence. Specifically, the value range of the maximum eigenvalue is partitioned into three subintervals: viruses tend to extinction very quickly or approach extinction or persist depending on into which subinterval the maximum eigenvalue of the propagation network falls. Consequently, computer virus can be contained by adjusting the propagation network so that its maximum eigenvalue falls into the desired subinterval.

1 Introduction

The rapidly popularized Internet has brought us lots of benefits. On the flip side of the coin, computer viruses can propagate their replicates through the Internet much more rapidly than ever before, resulting in great disruptions. Although antivirus software is recognized as the major means of defending against electronic viruses, there is a marked lag from the appearance of a new virus to the availability of its vaccine.

As an important supplement to antivirus techniques, the epidemic dynamics of computer viruses aims to understand the laws governing the spread of malware on networks and, thereby, to work out proper strategies to contain the viral prevalence. Since Kephart and White’s seminal work on the compartment modeling of computer viruses in the early 1990s [1, 2], a multitude of compartment-based computer virus propagation models, ranging from the SIR models [3] and the SIRS models [4, 5] to the SEIRS models [6], have been suggested. Most of these models are suited to infectious diseases and computer viruses equally well. In reality, however, some computer viruses have peculiarities most infectious diseases do not have. As we know, for most infectious diseases, there is a non-ignorable interval from the time when an individual gets infected to the time when it can infect other individuals. As opposed to this, for most computer viruses, one computer can infect other computers as soon as it gets infected. To capture
this common feature of most computer viruses, a series of epidemic models of computer virus, named as the SLBS models, were proposed [7, 8].

The network through which computers communicate with one another is frequently used to propagate viruses, and it has been recognized that the structure of the network has significant impact on the prevalence of virus [9]. In the early 2000s, it was empirically found that many real-world networks, ranging from the Internet and the World Wide Web to some email networks, are highly structured [10–12]. Later, a wave of research on virus epidemic dynamics was initiated, with focus on the propagation of virus on scale-free networks [13–19].

One common defect of all compartment-based epidemic models is that only partial knowledge on the network topology (the degree distribution or the degree correlation, say) can be used when establishing such models. In sharp contrast to this, when establishing a node-based epidemic model, one can make the best of the complete knowledge on the network topology [20]. As a result, some interesting properties concerning the viral spread, ranging from the mean propagation time and the expected number of infected nodes to the most probable network state, have been found [21–25].

With the aid of a node-based epidemic model, Wang et al. [26] found that whether viruses approach extinction depends heavily on the spectral radius of the underlying network. Next, by studying the N-interwined SIS model, Mieghem et al. [27] found that whether viruses decline toward extinction depends on the maximum eigenvalue of the network. Later, by examining a node-based SIR model, Youssef and Scoglio [28] indicated that the maximum number of infected nodes is closely related to the spectrum of the network. For more information on this topic, see Refs. [29–34].

This paper addresses the impact of the network topology on the viral prevalence, provided that a computer can infect other computers as soon as it gets infected. For that purpose, a node-based virus epidemic model, known as the node-based SLBS model, is proposed. After exhaustive research, it is found that the maximum eigenvalue of the underlying network is a key factor determining the viral prevalence. Specifically, the value range of the maximum eigenvalue is partitioned into three subintervals: viruses tend to extinction very quickly or approach extinction or persist depending on into which subinterval the maximum eigenvalue of the network falls. Consequently, computer virus can be contained by adjusting the network topology so that its maximum eigenvalue falls into the desired subinterval. Numerical examples support our results.

The rest of this paper is organized as follows: Preliminary knowledge is presented in Section 2, and the compartment-based SLBS models are briefly reviewed in Section 3. Section 4 describes the node-based SLBS model, Section 5 conducts a comprehensive analysis of this model, Section 6 gives some numerical examples, and Section 7 discusses the potential applications of the proposed model. Finally, Section 8 summarizes this work and presents some topics that are worthy of study.

2 Preliminaries

In this paper, the underlying network through which viruses propagate is denoted by a simple graph \( G = (V, E) \) on \( N \) non-isolated nodes numbered 1 through \( N \), where nodes stand for terminal devices of the network, and edges stand for network links through which viruses can propagate. Let \( A = [a_{ij}]_{N \times N} \) denote the adjacency matrix of graph \( G \), let \( \{d_k, 1 \leq k \leq N\} \) denote the degree sequence of \( G \), and let \( \{\lambda_k, 1 \leq k \leq N\} \) denote the spectrum of \( A \). As \( A \) is real and symmetric, we may assume \( \lambda_{\max} = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \).

For the purpose of analyzing the new computer virus epidemic model introduced in the next section, we need the following two lemmas.
Lemma 1 [35] Consider a smooth dynamical system \( \frac{dx(t)}{dt} = f(x(t)) \) defined at least in a compact set \( C \). \( C \) is positively invariant if for any smooth point \( y \) of \( \partial C \), \( f(y) \) is pointing into \( C \).

Lemma 2 [36] Consider an \( n \)-dimensional dynamical system
\[
\frac{dx(t)}{dt} = Bx(t) + G(x(t)),
\]
where \( x(0) \in \Omega \), a positively invariant compact convex set containing the origin, \( G(x) \in C^1(\Omega) \), \( \lim_{x \to 0} \frac{|G(x)|}{|x|} = 0 \). Suppose matrix \( B^T \) has a real eigenvector \( z \) such that

(C1) \( \inf_{x \in \Omega - \{0\}} \frac{x \cdot z}{|z|} > 0 \),

(C2) \( \sup_{x \in \Omega} (\mathbf{G}(x), z) \leq 0 \), and

(C3) the origin forms the largest positively invariant set included in the set \( \{x \in \Omega | (\mathbf{G}(x), z) = 0 \} \).

Let \( s(B) \) denote the maximum real part of all eigenvalues of \( B \). Then, we have

(a) \( s(B) < 0 \) implies the global stability of the origin, and

(b) \( s(B) > 0 \) implies that \( x(0) \neq 0 \Rightarrow \liminf_{t \to \infty} \|x(t)\| > 0 \).

Lemma 3 [37] For a graph \( G \) with \( \{d_k, 1 \leq k \leq N\} \) as the degree sequence, its largest eigenvalue \( \lambda_{\max} \) has the following bounds.
\[
\sqrt{\frac{1}{N} \sum_i d_i^2} \leq \lambda_{\max} \leq \min \left\{ \sqrt{\frac{N-1}{N} \sum_i d_i}, \max_i d_i \right\}
\]

3 A brief review of the compartment-based SLBS models

This section gives a brief review of the previously proposed SLBS models.

Under an SLBS model, every node in a network is assumed to be in one of three possible states: susceptible, i.e. uninfected, latent, i.e., infected and with all viruses in the node being in their latent phase, and exploding, i.e., infected and with at least one virus in the node being in its exploding phase. For a compartment-based SLBS model, all nodes in a network are grouped into three classes (i.e., compartments) according to their states, and the change in the fraction of each compartment is the focus of study.

The original compartment-based SLBS models were established based on the homogeneously mixed assumption of the propagation network [7, 8]. However, most real-world networks, including the world-wide-web and the Internet, have been empirically found to be highly structured rather than simply homogeneously [11]. Therefore, a new compartment-based SLBS model was later suggested based on the assumption that the propagation network admits a prescribed degree distribution [18].

All of the above mentioned SLBS models suffer from a common defect that it is not possible to make full use of the knowledge concerning the structure of the propagation network. 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.

4 The new computer virus epidemic model

As with the traditional compartment-based SLBS models [8, 18], at any time, each and every node in the network is in one of three possible states: susceptible, latent, and exploding. Let \( X_i(t) = 0 \) (respectively, 1, 2) stands for that node \( i \) is susceptible (respectively, latent, exploding) at
time $t$. Then the state of the whole network at time $t$ can be represented by the vector

$$X(t) = [X_1(t), X_2(t), \ldots, X_N(t)].$$

Let $s_i(t)$ (respectively, $l_i(t), b_i(t)$) denote the probability of the event that node $i$ is susceptible (respectively, latent, exploding) at time $t$,

$$s_i(t) = \Pr(X_i(t) = 0), \quad l_i(t) = \Pr(X_i(t) = 1), \quad b_i(t) = \Pr(X_i(t) = 2).$$

Now, let us impose a set of statistical assumptions on the state transitions of a node.

(H1) A susceptible node is infected by a latent (respectively, exploding) neighbor with probability per unit time $\beta_1$ (respectively, $\beta_2$). As a result, when the number of infected nodes is small, a susceptible node $i$ gets infected approximately with average probability per unit time $\beta_1 \sum_j a_{ij} l_j(t) + \beta_2 \sum_j a_{ij} b_j(t)$. As the mission of all the viruses staying in a latent node is to infect other nodes, whereas the mission of all the exploding viruses staying in an exploding node is to destruct the system, we assume $\beta_1 > \beta_2$.

(H2) Some virus in a latent node breaks out with probability per unit time $\alpha$.

(H3) A latent (respectively, exploding) node gets cured with probability per unit time $\gamma_1$ (respectively $\gamma_2$). As an exploding node has more chance to be cured than a latent node, we assume $\gamma_2 > \gamma_1$.

**Fig 1** shows these assumptions schematically.

Let $\Delta t$ be a very small time interval. By the total probability formula, we have the following relations:

$$s_i(t + \Delta t) = s_i(t) \Pr(X_i(t + \Delta t) = 0 \mid X_i(t) = 0) + l_i(t) \Pr(X_i(t + \Delta t) = 0 \mid X_i(t) = 1)$$

$$+ b_i(t) \Pr(X_i(t + \Delta t) = 0 \mid X_i(t) = 2),$$

$$l_i(t + \Delta t) = s_i(t) \Pr(X_i(t + \Delta t) = 1 \mid X_i(t) = 0) + l_i(t) \Pr(X_i(t + \Delta t) = 1 \mid X_i(t) = 1)$$

$$+ b_i(t) \Pr(X_i(t + \Delta t) = 1 \mid X_i(t) = 2),$$

$$b_i(t + \Delta t) = s_i(t) \Pr(X_i(t + \Delta t) = 2 \mid X_i(t) = 0) + l_i(t) \Pr(X_i(t + \Delta t) = 2 \mid X_i(t) = 1)$$

$$+ b_i(t) \Pr(X_i(t + \Delta t) = 2 \mid X_i(t) = 2).$$
Assumptions (H1)-(H3) imply the following equations:

\[
\Pr(X_i(t + \Delta t) = 0 | X_i(t) = 0) = 1 - \left[ \beta_i \sum_j a_{ij}(t) + \beta_j \sum_j a_{ji}(t) \right] \Delta t + o(\Delta t),
\]

\[
\Pr(X_i(t + \Delta t) = 1 | X_i(t) = 0) = \left[ \beta_i \sum_j a_{ij}(t) + \beta_j \sum_j a_{ji}(t) \right] \Delta t + o(\Delta t),
\]

\[
\Pr(X_i(t + \Delta t) = 2 | X_i(t) = 0) = o(\Delta t),
\]

\[
\Pr(X_i(t + \Delta t) = 1 | X_i(t) = 1) = \gamma_i \Delta t + o(\Delta t),
\]

\[
\Pr(X_i(t + \Delta t) = 2 | X_i(t) = 1) = \gamma_j \Delta t + o(\Delta t),
\]

\[
\Pr(X_i(t + \Delta t) = 0 | X_i(t) = 1) = \gamma_i \Delta t - \alpha \Delta t + o(\Delta t),
\]

\[
\Pr(X_i(t + \Delta t) = 1 | X_i(t) = 2) = o(\Delta t),
\]

\[
\Pr(X_i(t + \Delta t) = 2 | X_i(t) = 2) = -\gamma_j \Delta t + o(\Delta t).
\]

Substituting these equations into the above relations and letting \(\Delta t \to 0\), we get the following 3N-dimensional differential dynamical system:

\[
\begin{align*}
\frac{ds_i(t)}{dt} &= \gamma_i l_i(t) + \gamma_j b_i(t) - \left[ \beta_i \sum_j a_{ij}(t) + \beta_j \sum_j a_{ji}(t) \right] s_i(t), \quad i = 1, 2, \ldots, N, \\
\frac{dl_i(t)}{dt} &= -\left( \gamma_i + \alpha \right) l_i(t) + \left[ \beta_i \sum_j a_{ij}(t) + \beta_j \sum_j a_{ji}(t) \right] s_i(t), \quad i = 1, 2, \ldots, N, \\
\frac{db_i(t)}{dt} &= \alpha l_i(t) - \gamma_j b_i(t), \quad i = 1, 2, \ldots, N.
\end{align*}
\]

As \(s_i(t) + l_i(t) + b_i(t) \equiv 1\), this system is equal to the following 2N-dimensional system:

\[
\begin{align*}
\frac{dl_i(t)}{dt} &= -\left( \gamma_i + \alpha \right) l_i(t) + \left( 1 - l_i(t) - b_i(t) \right) \left[ \beta_i \sum_j a_{ij}(t) + \beta_j \sum_j a_{ji}(t) \right], \\
\quad i = 1, 2, \ldots, N, \\
\frac{db_i(t)}{dt} &= \alpha l_i(t) - \gamma_j b_i(t), \quad i = 1, 2, \ldots, N.
\end{align*}
\]

We shall refer to system (1) (equivalently, system (2)) as the node-based SLBS model.

**Remark 1** When \(\beta_2 = \gamma_2 = \alpha = 0\), the node-based SLBS model degrades into the N-interwined SIS model [27]. When \(\beta_1 = \beta_2\) and \(\gamma_1 = \gamma_2\), our model again degenerates into the N-interwined SIS model.

The major task in the subsequent sections is to study the dynamical properties of system (2) (equivalently, system (1)).

### 5 Model analysis

Obviously, system (2) always has the origin as an equilibrium. This trivial equilibrium stands for that all viruses in the network die out almost surely. This section is focused on the stability properties of the trivial equilibrium.
First, consider the asymptotic stability of the trivial equilibrium of system (2). For that purpose, let
\[ \Omega = \{ x = (x_1, x_2, \ldots, x_{2N})^T \in \mathbb{R}^{2N} \mid x_i + x_{i+N} \leq 1, i = 1, 2, \ldots, N \}. \]

Let \( x(t) = (l_1(t), \ldots, l_N(t), b_1(t), \ldots, b_N(t))^T \), and rewrite system (2) in matrix notation as
\[ \frac{dx(t)}{dt} = Bx(t) + G(x(t)) \quad (3) \]
with initial condition \( x(0) \in \Omega \), where
\[
B = \begin{pmatrix}
\beta_1 A - (\alpha + \gamma_1) I & -\beta_2 A \\
\alpha I & -\gamma_2 I
\end{pmatrix},
\]
\[
G(x(t)) = -((l_1(t) + b_1(t)) \left( \beta_1 \sum_j a_{ij} l_j(t) + \beta_2 \sum_j a_{ij} b_j(t) \right), \ldots, \\
(l_N(t) + b_N(t)) \left( \beta_1 \sum_j a_{ij} l_j(t) + \beta_2 \sum_j a_{ij} b_j(t) \right), 0, \ldots, 0)^T.
\]

Finally, let
\[ R_0 = \frac{(\alpha + \gamma_1) \gamma_2}{\beta_1 \gamma_2 + \beta_2 \alpha}. \quad (4) \]

We are ready to present a criterion for the asymptotic stability of the trivial equilibrium.

**Theorem 1** Consider system (2).

(a) The trivial equilibrium is asymptotically stable if \( \lambda_{\text{max}} < R_0 \).

(b) The trivial equilibrium is unstable if \( \lambda_{\text{max}} > R_0 \).

**Proof** The characteristic equation for the Jacobian of system (3) evaluated at the trivial equilibrium is
\[
\det (\eta I - B) = \det \begin{pmatrix}
(\eta + \alpha + \gamma_1) I - \beta_1 A & -\beta_2 A \\
-\beta_1 A & (\eta + \gamma_2) I
\end{pmatrix} = \det ((\eta + \alpha + \gamma_1) (\eta + \gamma_2) I - (\beta_1 (\eta + \gamma_2) + \beta_2 \alpha) A) = 0 \quad (5)
\]

We distinguish between two possibilities.

Case 1: \( \beta_1 (\gamma_2 - \gamma_1) = \alpha(\beta_1 - \beta_2) \). Then, \( R_0 = \frac{\alpha}{\beta_1} \), and Eq (5) degrades into
\[
\beta_1^N (\eta + \alpha + \gamma_1)^N \det \left( \frac{\eta + \gamma_2}{\beta_1} I - A \right) = 0
\]
This equation has \( -(\alpha + \gamma_1) \) as a root with multiplicity \( N \), and has \( \beta_1 \lambda_k - \gamma_2, 1 \leq k \leq N \) as the remaining \( N \) roots. If \( \lambda_{\text{max}} < R_0 \), then \( \beta_1 \lambda_k - \gamma_2 \leq \beta_1 \lambda_{\text{max}} - \gamma_2 < 0 \) for all \( k \). So, the roots of Eq (5) are all negative. Hence, the trivial equilibrium of system (2) is asymptotically stable [38]. Otherwise, if \( \lambda_{\text{max}} > R_0 \), then \( \beta_1 \lambda_{\text{max}} - \gamma_2 > 0 \). So, Eq (5) has a positive equilibrium. Thus, the trivial equilibrium is unstable [38].
Case 2: \( \beta_1(\gamma_2 - \gamma_1) \neq \alpha(\beta_1 - \beta_2) \). Then, \(-\gamma_2 - \frac{\alpha x}{\beta_1} \) is not a root of Eq (5). Thus,

\[
\det \left( \frac{(\eta + x)(\eta + \gamma_2)}{\beta_1(\eta + \gamma_2) + \beta_2 x} I - A \right) = 0.
\]

This implies that \( \eta \) is a root of Eq (5) if and only if for some \( k \) (\( 1 \leq k \leq N \)), \( \eta \) is a root of equation

\[
\eta^2 + a_k \eta + b_k = 0, \quad (6)
\]

where

\[
a_k = \alpha + \gamma_1 + \gamma_2 - \beta_1 \lambda_k,
\]

\[
b_k = (\alpha + \gamma_1)\gamma_2 - \lambda_k(\beta_1 \gamma_2 + \beta_2 x).
\]

If \( \lambda_{\max} < R_0 \), we have \( a_k > 0 \) and \( b_k > 0 \). So, it follows from the Hurwitz criterion [38] that the two roots of Eq (6) have negative real parts. As a result, all roots of Eq (5) have negative real parts. Hence, the trivial equilibrium is asymptotically stable [38]. Otherwise, if \( \lambda_{\max} > R_0 \), the equation

\[
\eta^2 + a_k \eta + b_k = 0
\]

has a root with positive real part. As a result, Eq (5) has a root with positive real part. Hence, the trivial equilibrium is unstable [38]. The proof is complete.

**Remark 2** This theorem can also be formulated as (a) \( \lambda_{\max} < R_0 \Rightarrow s(B) < 0 \), and (b) \( \lambda_{\max} > R_0 \Rightarrow s(B) > 0 \).

Second, consider the global stability of the trivial equilibrium of system (2). For that purpose, the following lemma is indispensable.

**Lemma 4** The set \( \Omega \) is positively invariant for system (2). That is, \( x(0) \in \Omega \) implies \( x(t) \in \Omega \) for all \( t > 0 \).

**Proof** \( \partial \Omega \) consists of the following \( 3N \) hyperplanes:

\[
S_i = \{ x \in \Omega \mid x_i = 0 \}, \quad 1 \leq i \leq N,
\]

\[
T_i = \{ x \in \Omega \mid x_{i+N} = 0 \}, \quad 1 \leq i \leq N,
\]

\[
U_i = \{ x \in \Omega \mid x_i + x_{i+N} = 1 \}, \quad 1 \leq i \leq N.
\]

For \( 1 \leq i \leq N, S_i, T_i, \) and \( U_i \) have

\[
n_i = \{0, \cdots, 0, \underbrace{1}_{i}, 0, \cdots, 0\},
\]

\[
n_{N+i} = \{0, \cdots, 0, \underbrace{1}_{N+i}, 0, \cdots, 0\},
\]

and

\[
n_{2N+i} = \{0, \cdots, 0, \underbrace{1}_{i}, 0, \cdots, 0, \underbrace{1}_{N+i}, 0, \cdots, 0\}
\]

as their respective outer normal vectors. Let \( x \) be a smooth point of \( \partial \Omega \). We distinguish among three possibilities.

---

Network Topology and Viral Prevalence

PLOS ONE | DOI:10.1371/journal.pone.0134507 July 29, 2015 7/15
Case 1: \( x_i = 0 \) for some \( 1 \leq i \leq N \). Then, \( x_{N+i} < 1 \), and \( x_j > 0 \) for all \( j \neq i \). As graph \( G \) has no isolated node, we have

\[
(Bx + G(x), n_i) = -(1 - x_{N+i}) \left( \beta_1 \sum_j a_{ij} x_j + \beta_2 \sum_j a_{ij} x_{N+j} \right) < 0.
\]

Case 2: \( x_{N+i} = 0 \) for some \( 1 \leq i \leq N \). Then, \( x_i > 0 \). Thus,

\[
(Bx + G(x), n_{N+i}) = -x_i < 0.
\]

Case 3: \( x_i + x_{N+i} = 1 \) for some \( 1 \leq i \leq N \). Then,

\[
(Bx + G(x), n_{2N+i}) = -\gamma_1 x_i - \gamma_2 (1 - x_i) < 0.
\]

Combining the above discussions, we get that \( Bx + G(x) \) is pointing into \( \partial \Omega \). The claimed result then follows from Lemma 1. The proof is complete.

We are ready to present a criterion for the global stability of the trivial equilibrium.

**Theorem 2** The trivial equilibrium of system (2) is globally asymptotically stable if \( \lambda_{\text{max}} < R_0 \).

**Proof** Look at system (3). As matrix \( B^T \) is irreducible and its off-diagonal entries are all non-negative, it follows from [36] that \( B^T \) has a positive eigenvector \( z = (z_1, z_2, \cdots, z_{2N}) \) belonging to its eigenvalue \( s(B^T) \). Let \( r = \min_i z_i (> 0) \). Then, for all \( x \in \Omega \), we have

\[
\langle x, z \rangle \geq r \sum_i x_i = r \left\| x \right\|_1,
\]

\[
\langle G(x), z \rangle = -\sum_i z_i (x_i + x_{N+i}) (\beta_1 \sum_j a_{ij} x_j + \beta_2 \sum_j a_{ij} x_{N+j}) \leq 0
\]

Moreover, \( \langle G(x), z \rangle = 0 \) implies that \( x = 0 \). In view of Theorem 1 and Lemma 3, the claimed result follows from Lemma 2. The proof is complete.

**Remark 3** The global stability of the trivial equilibrium of system (2) implies that, almost surely, the viruses in the network decline toward extinction.

Next, consider the global exponential stability of the trivial equilibrium of system. For that purpose, let

\[
R_i = \frac{\gamma_i}{\beta_i}.
\]

Now, let us give a criterion for the global exponential stability of the trivial equilibrium.

**Theorem 3** The trivial equilibrium of system (2) is globally exponentially stable if \( \lambda_{\text{max}} < R_i \).

**Proof** We have

\[
\frac{d(l_i(t) + b_i(t))}{dt} = \beta_i (1 - l_i(t) - b_i(t)) \sum_j a_{ij} (l_j(t) + b_j(t)) - \gamma_i (l_i(t) + b_i(t)) - (\beta_i - \beta_j) (1 - l_i(t) - b_i(t)) \sum_j a_{ij} b_j(t) - (\gamma_i - \gamma_j) b_i(t)
\]

\[
\leq \beta_i (1 - l_i(t) - b_i(t)) \sum_j a_{ij} (l_j(t) + b_j(t)) - \gamma_i (l_i(t) + b_i(t)) - \gamma_j (l_i(t) + b_i(t))
\]

\[
\leq \beta_i \sum_j a_{ij} (l_j(t) + b_j(t)) - \gamma_i (l_i(t) + b_i(t)), \quad i = 1, \cdots, N.
\]
Consider the comparison system

\[
\frac{dp_i(t)}{dt} = \beta_i \sum_j a_{ij} p_j(t) - \gamma_i p_i(t), \quad i = 1, \ldots, N,
\]

with initial condition \( p_i(0) = l_i(0) + b_i(0), \quad i = 1, \ldots, N \). Let \( p(t) = [p_1(t), \ldots, p_N(t)]^T \) and rewrite the comparison system in matrix notation as

\[
\frac{dp(t)}{dt} = (\beta_1 A - \gamma_1 I)p(t).
\]

The solution to this comparison system is

\[
p(t) = e^{(\beta_1 A - \gamma_1 I)t}p(0).
\]

Let \( U \) denote the \( N \times N \) matrix having an eigenvector belonging to the eigenvalue \( \lambda_i \) of matrix \( A \) as the \( i \)-th column. Then, we have the following spectral decomposition of matrix \( A \):

\[
A = U \times \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N) \times U^T.
\]

As a result, we have

\[
p(t) = U \times \text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots, e^{\lambda_N t}) \times U^T \times p(0)
\]

Hence,

\[
\|p(t)\|_1 \leq \|U\|_1 \times \|\text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots, e^{\lambda_N t})\|_1 \times \|U^T\|_1 \times \|p(0)\|_1.
\]

Therefore, \( p(t) \to 0 \) at an exponential speed if \( \lambda_{\text{max}} < R_1 \). It follows from the comparison theorem [38] that if \( \lambda_{\text{max}} < R_1 \), then for each \( i \), \( l_i(t) + b_i(t) \to 0 \) at an exponential speed. The proof is complete.

Finally, let us consider what happens if \( \lambda_{\text{max}} > R_0 \). By applying Lemma 2 to system (2) and in view of Theorem 1, we get the following result.

**Theorem 4** Consider system (2). If \( \lambda_{\text{max}} > R_0 \), then

\[
\sum_i (l_i(0) + b_i(0)) > 0 \Rightarrow \lim_{t \to \infty} \inf \sum_i (l_i(t) + b_i(t)) > 0.
\]

**Remark 4** This theorem shows that if \( \lambda_{\text{max}} > R_0 \), then, almost surely, viruses in the network persist.

**Remark 5** As the largest eigenvalue of a network is an indicator of the structure of the network, Theorems 1–4 clearly reveal the impact of the network topology on the viral prevalence; a network with smaller largest eigenvalue is inclined to contain viruses.

It follows from Theorems 1–4 that it is proper to partition the value range \( (0, \infty) \) of \( \lambda_{\text{max}} \) into three subintervals: \( I_1 = (0, R_1) \), \( I_2 = (R_1, R_0) \), and \( I_3 = (R_0, \infty) \). When \( \lambda_{\text{max}} \in I_1 \), viruses in the network tend to extinction almost surely, at an exponential speed. When \( \lambda_{\text{max}} \in I_2 \), viruses in the network declines toward annihilation almost surely. When \( \lambda_{\text{max}} \in I_3 \), viruses in the network persist.

**6 Numerical examples**

In this section, we shall verify the results obtained in the previous section using numerical examples. For that purpose, let \( p(t) \) denote the percentage of infected nodes in all nodes at time \( t \), \( p(t) = \frac{1}{N} \sum_i (l_i(t) + b_i(t)) \).
Example 1 Consider the node-based SLBS model, and take a complete graph on 100 nodes as the viral propagation network. Then, $\lambda_{\text{max}} = 99$.

1. Suppose $\beta_1 = 0.002$, $\beta_2 = 0.001$, $\gamma_1 = 0.2$, $\gamma_2 = 0.3$, and $\alpha = 0.1$. As $\lambda_{\text{max}} \in I_1$, Theorem 3 predicts that $p(t) \to 0$ at an exponential speed. Fig 2(1) shows the trend of $p(t)$ provided (a) there are initially 90 susceptible nodes and 10 latent nodes, or (b) there are initially 90 susceptible nodes and 10 exploding nodes. It can be seen that viruses tend to extinction very quickly, in consistency with the prediction.

2. Suppose $\beta_1 = 0.0025$, $\beta_2 = 0.001$, $\gamma_1 = 0.2$, $\gamma_2 = 0.3$, and $\alpha = 0.1$. As $\lambda_{\text{max}} \in I_2$, Theorem 2 predicts that $p(t) \to 0$. Fig 2(2) shows the trend of $p(t)$ provided (a) there are initially 90 susceptible nodes and 10 latent nodes, or (b) there are initially 90 susceptible nodes and 10 exploding nodes. It can be seen that viruses tend toward annihilation, in agreement with the prediction.

3. Suppose $\beta_1 = 0.01$, $\beta_2 = 0.005$, $\gamma_1 = 0.2$, $\gamma_2 = 0.3$, and $\alpha = 0.1$. As $\lambda_{\text{max}} \in I_3$, Theorem 4 predicts that $p(t) \not\to 0$. Fig 2(3) shows the trend of $p(t)$ provided (a) there are initially 90 susceptible nodes and 10 latent nodes, or (b) there are initially 90 susceptible nodes and 10 exploding nodes. It can be seen that viruses persist, agreeing with the prediction.
Example 2 Consider the node-based SLBS model, and take a star-shaped graph on 100 nodes as the viral propagation network. Then,
\[ \lambda_{\text{max}} = \sqrt{99} = 9.95. \]

1. Suppose \( \beta_1 = 0.01, \beta_2 = 0.006, \gamma_1 = 0.2, \gamma_2 = 0.3, \) and \( \alpha = 0.1. \) As \( \lambda_{\text{max}} \in I_1, \) Theorem 3 predicts that \( p(t) \to 0 \) at an exponential speed. Fig 3(1) shows the trend of \( p(t) \) provided (a) the hub is initially latent, and the remaining 99 nodes are initially susceptible, or (b) one leaf node is initially latent, and the remaining 99 nodes are initially susceptible. It can be seen that viruses tend to extinction very quickly, coinciding with the prediction.

2. Suppose \( \beta_1 = 0.035, \beta_2 = 0.01, \gamma_1 = 0.2, \gamma_2 = 0.4, \) and \( \alpha = 0.2. \) As \( \lambda_{\text{max}} \in I_2, \) Theorem 2 predicts that \( p(t) \to 0. \) Fig 3(2) shows the trend of \( p(t) \) provided (a) the hub is initially latent, and the remaining 99 nodes are initially susceptible, or (b) one leaf node is initially latent, and the remaining 99 nodes are initially susceptible. It can be seen that viruses tend to extinction, in agreement with the prediction.

3. Suppose \( \beta_1 = 0.05, \beta_2 = 0.02, \gamma_1 = 0.2, \gamma_2 = 0.3, \) and \( \alpha = 0.1. \) As \( \lambda_{\text{max}} \in I_3, \) Theorem 4 predicts that \( p(t) \not\to 0. \) Fig 3(3) shows the trend of \( p(t) \) provided (a) the hub is initially latent, and the remaining 99 nodes are initially susceptible, or (b) one leaf node is initially latent, and the remaining 99 nodes are initially susceptible. It can be seen that viruses persist, consistent with the prediction.
Example 3 Consider the node-based SLBS model, and take an Erdos-Renyi graph on 50 nodes, which is produced randomly with connection probability 0.2, as the viral propagation network. Numerical calculation gives $\lambda_{\text{max}} = 10.19$.

1. Suppose $\beta_1 = 0.015, \beta_2 = 0.01, \gamma_1 = 0.2, \gamma_2 = 0.4, \text{and } \alpha = 0.2$. As $\lambda_{\text{max}} \in I_1$, Theorem 3 predicts that $p(t) \to 0$ at an exponential speed. Fig 4(1) shows the trend of $p(t)$ provided there are initially 10 latent nodes and 40 susceptible nodes. It can be seen that viruses go to extinction very quickly, in consistency with the prediction.

2. Suppose $\beta_1 = 0.03, \beta_2 = 0.01, \gamma_1 = 0.2, \gamma_2 = 0.4, \text{and } \alpha = 0.2$. As $\lambda_{\text{max}} \in I_2$, Theorem 2 predicts that $p(t) \to 0$. Fig 4(2) shows the trend of $p(t)$ provided there are 10 latent nodes and 40 susceptible nodes. It can be seen that viruses tend toward annihilation, in coherence with the prediction.

3. Suppose $\beta_1 = 0.05, \beta_2 = 0.03, \gamma_1 = 0.2, \gamma_2 = 0.4, \text{and } \alpha = 0.2$. As $\lambda_{\text{max}} \in I_3$, Theorem 4 predicts that $p(t) \not\to 0$. Fig 4(3) shows the trend of $p(t)$ provided there are initially 10 latent nodes and 40 susceptible nodes. It can be seen that viruses persist, in agreement with the prediction.
In a word, the above given numerical examples are all in perfect agreement with the theoretical results.

7 Further discussions

It can be seen from the main results in Section 5 that an effective approach to the containment of electronic virus is to adjust the system parameters so that $R_0$ or $R_1$ is large enough. Simple calculations yield

$$
\frac{\partial R_0}{\partial \beta_1} < 0, \quad \frac{\partial R_0}{\partial \beta_2} < 0, \quad \frac{\partial R_0}{\partial \gamma_1} > 0, \quad \frac{\partial R_0}{\partial \gamma_2} > 0, \quad \frac{\partial R_1}{\partial \alpha} > 0,
$$

$$
\frac{\partial R_1}{\partial \beta_1} < 0, \quad \frac{\partial R_0}{\partial g_1} > 0, \quad \frac{\partial R_0}{\partial g_2} > 0,
$$

As a result, the following practical measures are strongly recommended.

- Install and timely update antivirus software on computers, so as to reduce the two cure rates of infected computers.
- Filter and block suspicious messages with firewall located at the gateway of a domain, so as to lower the two infecting rates of susceptible computers.

On the other hand, it benefits the inhibition of virus to adjust the structure of the propagation network so that its maximum eigenvalue is small enough. As there is no closed-form formula for the maximum eigenvalue of a general adjacency matrix, it is difficult to verify this condition. To circumvent this difficulty, let us present an easily verified condition for the final extinction of virus as follows.

**Theorem 5** All viruses in a network would tend to extinction if

$$
\min \left\{ \sqrt{\frac{N-1}{N} \sum_i d_i}, \max_i d_i \right\} < \max \{ R_0, R_1 \}.
$$

**Proof** The claim follows by combining Lemma 3 and Theorems 2–3.

This theorem suggests that simultaneously reducing the number of links and the maximum node degree in a network should contribute to the annihilation of virus.

8 Conclusions and remarks

To understand the way that the spread of virus on a network is affected by the structure of the network, a new epidemic model of computer virus has been proposed. The model analysis reveals that the maximum eigenvalue of the network is a key factor determining the viral prevalence; viruses tend to extinction very quickly or approach extinction or persist depending on where the maximum eigenvalue of the network lies. As a result, viruses can be contained by properly adjusting the structure of the propagation network.

Towards this direction, lots of work has yet to be done. For instance, our model assumes that all computers have the same infection rate, the same bursting rate, and the same curing rate. In reality, however, these rates vary from computer to computer. Hence, our model should be generalized so that different nodes have different infection rates, different bursting rates, and different curing rates. Additionally, that computers are likely to be infected by removable storage media may lead to the emergence of a non-trivial steady state. In this situation, it makes sense to suppress the fraction of the infected nodes. Third, the immunization strategy we adopt also has significant impact on the viral prevalence. To a certain extent, the static immunization problem reduces to that of assigning different curing rates to different nodes so
that the best virus containment effect is achieved, given that the sum of curing rates of all nodes is fixed [33, 40], while the dynamic immunization problem can be solved by use of the optimal control theory [41]. Last, but not least, the methodology developed in this work can be applied to the situation of infectious diseases [42–45].

**Acknowledgments**

The authors are grateful to the two anonymous reviewers and the editor for their valuable comments and suggestions. This work is supported by China Scholarship Council (Grant No. 201406050107).

**Author Contributions**

Conceived and designed the experiments: LY. Performed the experiments: LY. XY MD. Wrote the paper: LY XY MD.

**References**

1. Kephart JO, White SR. Directed-graph epidemiological models of computer viruses. in: Proc. 1991 IEEE Comput. Soc. Symp. Res. Secur. Privacy 1991; 343–359.
2. Kephart JO, White SR. Measuring and modeling computer virus prevalence. in: Proc. 1993 IEEE Comput. Soc. Symp. Res. Secur. Privacy 1993; 2–15.
3. Ren J, Yang X, Zhu Q, Yang LX, Zhang C. A novel computer virus model and its dynamics. Nonlinear Anal. RWA 2012; 13: 376–384. doi: 10.1016/j.nonrwa.2011.07.048
4. Gan C, Yang X, Liu W, Zhu Q. A propagation model of computer virus with nonlinear vaccination probability. Commun. Nonlinear Sci. Numer. Simulat. 2014; 19: 92–100. doi: 10.1016/j.cnsns.2013.06.018
5. Muroya Y, Enatsu Y, Li H. Global stability of a delayed SIRS computer virus propagation model. Int. J. Comput. Math. 2014; 91: 347–367. doi: 10.1080/00207160.2013.790534
6. Mishra BK, Pandey SK. Dynamic model of worms with vertical transmission in computer network. Appl. Math. Comput. 2011; 217: 8438–8446. doi: 10.1016/j.amc.2011.03.041
7. Yang X, Yang LX, Towards the epidemiological modeling of computer viruses. Discrete Dyn. Nat. Soc. 2012; 2012: Article ID 259671.
8. Yang LX, Yang X, Zhu Q, Wen L. A computer virus model with graded cure rates. Nonlinear Anal. RWA 2013; 14: 414–422. doi: 10.1016/j.nonrwa.2012.07.005
9. Balthrop J, Forrest S, Newman MEJ, Williamson MW. Technological networks and the spread of computer viruses. Science 2004; 304: 527–529. doi: 10.1126/science.1095845 PMID: 15105484
10. Barabasi AL, Albert R. Emergence of scaling in random networks. Science 1999; 286: 509–512. doi: 10.1126/science.286.5439.509 PMID: 10521342
11. Albert R, Barabasi AL. Statistical mechanics of complex networks. Rev. Mod. Phys. 2002; 74: 47–97. doi: 10.1103/RevModPhys.74.47
12. Ebel H, Mielsch LI, Bornholdt S. Scale-free topology of email networks. Phys. Rev. E 2002; 66: Article ID 036103. doi: 10.1103/PhysRevE.66.036103
13. Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. Phys. Rev. Lett. 2001; 86: 3200–3203. doi: 10.1103/PhysRevLett.86.3200 PMID: 11290142
14. Pastor-Satorras R, Vespignani A. Epidemic dynamics and endemic states in complex networks. Phys. Rev. E 2001; 63: Article ID 066117. doi: 10.1103/PhysRevE.63.066117
15. Castellano C, Pastor-Satorras R. Thresholds for epidemic spreading in networks. Phys. Rev. Lett. 2010; 105: Article ID 218701. doi: 10.1103/PhysRevLett.105.218701
16. Huang CY, Sun CT. Effects of resource limitations and cost influences on computer virus epidemic dynamics and tipping points. Discrete Dyn. Nat. Soc. 2012; 2012: Article ID 473136. doi: 10.1155/2012/473136
17. Huang CY, Lee CL, Wen TH, Sun CT. A computer virus spreading model based on resource limitations and interaction costs. J. System Software 2013; 86: 801–808. doi: 10.1016/j.jss.2012.11.027
18. Yang LX, Yang X, Liu J, Zhu Q, Gan C. Epidemics of computer viruses: A complex-network approach. Appl. Math. Comput. 2013; 219: 8705–8717. doi: 10.1016/j.amc.2013.02.031
19. Yang LX, Yang X. The spread of computer viruses over a reduced scale-free network. Physica A 2014; 396: 173–184. doi: 10.1016/j.physa.2013.11.026

20. Anderson RM, May RM. Infectious Diseases of Humans. Oxford University Press; 1992.

21. Allen EJ. Stochastic differential equations and persistence time for two interacting populations. Dyn. Continuous Discrete Impulsive System. 1999; 5: 271–281.

22. Allen LJ, Allen EJ. Comparison of three different stochastic population models with regard to persistence time. Theor. Pop. Biol. 2003; 64: 439–449. doi: 10.1016/S0040-5809(03)00104-7

23. Ganesh A, Massoulie L, Towsley D. The effect of network topology on the spread of epidemics. INFOCOM 2005. 24th Annual Joint Conf. IEEE Computer Communications Societies. 2005; 2: 1455–1466.

24. Banerjee S, Chatterjee A, Shakkottai S. Epidemic thresholds with external agents; Available: arXiv:1312.0938v1. Accessed 3 Dec 2013.

25. Draief M, Ganesh A, Massoulie L. Thresholds for virus spread on networks. Ann. Appl. Probab. 2008; 18: 439–449. doi: 10.1214/07-AAP470

26. Wang Y, Chakrabarti D, Wang C, Faloutsos C. Epidemic spreading in real networks: an eigenvalue point. in 22nd Int. Symp. Reliable Distrib. System 2003; 25–34.

27. Mieghem PV, Omic J, Kooij R. Virus spread in networks. IEEE/ACM Trans. Networking 2009; 17: 1–14. doi: 10.1109/TNET.2008.925623

28. Yousef M, Scoglio C. An individual-based approach to SIR epidemics in contact networks. J. Theo. Biol. 2011; 283: 136–144. doi: 10.1016/j.jtbi.2011.05.029

29. Sahneh FD, Scoglio C. Epidemic spread in human networks. 50th IEEE Conf. Decision Control and European Control Conf. 2011.

30. Sahneh FD, Scoglio C. Optimal information dissemination in epidemic networks. 50th IEEE Conf. Decision Control, 2012.

31. Khanaf a A, Basar T, Gharesifard B. Stability properties of infected networks with low curing rates. American Control Conf. 2014.

32. Lin Y, Lui JC. Modelling multi-state diffusion process in complex networks: theory and applications. J. Comp. Netw. 2014; 2: 431–459. doi: 10.1093/comnet/cnu027

33. Preciado VM, Zargham M, Enyioha C, Jadbabaie A, Pappas GJ. Optimal resource allocation for network protection against spreading processes. IEEE Trans. Control Netw. System 2014; 1: 99–108. doi: 10.1109/TCNS.2014.2310911

34. Pastor-Satorras R, Castellano C, Mieghem PV, Vespignan A. Epidemic processes in complex networks; Available: arXiv:1408.2701v1. Accessed 12 Aug 2014.

35. Yorke JA. Invariance for ordinary differential equations. Theory Comput. Syst. 1967; 1: 353–372.

36. Lajmanovich A, Yorke JA. A deterministic model for gonorrhea in anhomogenous population. Math. Biosci. 1976; 28: 221–236. doi: 10.1016/0025-5564(76)90125-5

37. Mieghem PV. Graph Spectra for Complex Networks. Cambridge University Press; 2012.

38. Robinson RC. An Introduction to Dynamical Systems: Continuous and Discrete. Prentice Hall; 2004.

39. Wang Y, Yang X. The spread of computer viruses under the influence of removable devices. Appl. Math. Comput. 2012; 219: 3914–3922. doi: 10.1016/j.amc.2012.10.027

40. Shakeri H, Sahneh FD, Scoglio C. Optimal information dissemination strategy to promote preventive behaviours in multilayer epidemic networks. Math. Biosc. Eng. 2015; 12: 609–623. doi: 10.3934/mbe.2015.12.609

41. Eshghi S, Khouzani M, Sarkar S, Venkatesh S. Optimal patching in clustered malware epidemics, IEEE/ACM Trans. Networking doi: 10.1109/TNET.2014.2364034

42. Sun GQ, Liu QX, Jin Z. Chakraborty A, Li BL. Influence of infection rate and migration on extinction of disease in spatial epidemics. J. Theo. Biol. 2010; 264: 95–103. doi: 10.1016/j.jtbi.2010.01.006

43. Wang Y, Jin Z, Yang Z, Zhang ZK, Zhou T, Sun GQ. Global analysis of an SIS model with an infective vector on complex networks. Nonlinear Anal. RWA 2012; 13: 543–557. doi: 10.1016/j.nonrwa.2011.07.033

44. Zhang X, Sun GQ, Zhu YX, Ma J, Jin Z. Epidemic dynamics on semi-directed complex networks. Math. Biosci. 2013; 246: 242–251. doi: 10.1016/j.mbs.2013.10.001 PMID: 24140877

45. Jin Z, Sun GQ, Zhu HP. Epidemic models for complex networks with demographics. Math. Biosc. Eng. 2014; 11: 1295–1317. doi: 10.3934/mbe.2014.11.1295