



Viral conductance: Quantifying the robustness of networks with respect to spread of epidemics

Mina Youssef^{a,*}, Robert Kooij^{a,b,c}, Caterina Scoglio^a

^a K-State Epicenter, Department of Electrical and Computer Engineering, Kansas State University, Manhattan, KS 66506, USA

^b Faculty of Electrical Engineering, Mathematics, and Computer Science, Delft University of Technology, P.O. Box 5031, 2600 GA, Delft, The Netherlands

^c TNO Information and Communication Technology, Delft, The Netherlands

ARTICLE INFO

Article history:

Received 28 September 2010

Received in revised form 19 February 2011

Accepted 2 March 2011

Available online 2 April 2011

Keywords:

Spread of epidemics

Complex networks

Robustness of networks

ABSTRACT

In this paper, we propose a novel measure, viral conductance (VC), to assess the robustness of complex networks with respect to the spread of SIS epidemics. In contrast to classical measures that assess the robustness of networks based on the epidemic threshold above which an epidemic takes place, the new measure incorporates the fraction of infected nodes at steady state for all possible effective infection strengths. Through examples, we show that VC provides more insight about the robustness of networks than does the epidemic threshold. We also address the paradoxical robustness of Barabási–Albert preferential attachment networks. Even though this class of networks is characterized by a vanishing epidemic threshold, the epidemic requires high effective infection strength to cause a major outbreak. On the contrary, in homogeneous networks the effective infection strength does not need to be very much beyond the epidemic threshold to cause a major outbreak. To overcome computational complexities, we propose a heuristic to compute the VC for large networks with high accuracy. Simulations show that the heuristic gives an accurate approximation of the exact value of the VC. Moreover, we derive upper and lower bounds of the new measure. We also apply the new measure to assess the robustness of different types of network structures, i.e. Watts–Strogatz small world, Barabási–Albert, correlated preferential attachment, Internet AS-level, and social networks. The extensive simulations show that in Watts–Strogatz small world networks, the increase in probability of rewiring decreases the robustness of networks. Additionally, VC confirms that the irregularity in node degrees decreases the robustness of the network. Furthermore, the new measure reveals insights about design and mitigation strategies of infrastructure and social networks.

Published by Elsevier B.V.

1. Introduction

Our daily activities increasingly rely on complex networks, for example the power grid, the Internet, and transportation networks. In contrast to simple networks, such as regular or Erdős–Rényi (ER) random networks [2], complex networks are characterized by a large number of vertices (from hundreds of thousands to billions of nodes), a low density of links, clustering effects, and power-law node-degree distribution [3,4]. Being so large, complex networks are often controlled in a decentralized way and show properties of self-organization. However, even if decentralization and self-organization theoretically reduce the risk of failure, complex networks can experience disruptive and massive failure. While many technological networks are still vulnerable to attacks, their robustness can be analyzed through their topological metrics. For example, the robustness of overlay networks in peer-to-peer

networks and service overlay networks is assessed through their topological metrics [5]. Since our daily routines would cease if the technological information infrastructure were to disintegrate, maintaining the highest levels of robustness in complex networks is crucial. Therefore, as a first step, we need to be able to assess the robustness of networks, which obviously depends on the type of attack.

Among the various scenarios of attack that damage infrastructure, the spread of epidemics caused by large scale attacks has become a crucial issue affecting modern life. For example, in 2001 and 2004, respectively, the Code Red and Sasser computer viruses infected numerous computers, resulting in costly global damages. Furthermore, many attacks range from theft, modification, and destruction of data, to dismantling of entire networks. In addition to technological networks, epidemics also spread in social networks, much like human diseases, like influenza [9,10], widely spread among individuals in contact with each other. In this paper, we study the robustness of networks with respect to the spread of epidemics that follow the susceptible/infected/susceptible (SIS) model. The SIS model can describe how an infection spreads in

* Corresponding author.

E-mail address: mkamel@ksu.edu (M. Youssef).

networks, and its dynamic is described by any change of a node's state with time. Thus, an infected node infects any susceptible neighbor with infection rate β , which is the number of infection trials per unit time. Then, the infected node cures itself with cure rate δ , and it becomes susceptible to a new infection. The ratio between β and δ is called the effective spreading rate τ . The epidemic threshold τ_c is a function of network characteristics and a specific value of the effective spreading rate above which an epidemic outbreak takes place. When an epidemic outbreak takes place in the network, a persistent fraction of infected population exists at the steady state, and this fraction does not depend on the initial condition of the infection in the network, but it only depends on the effective infection rate and the epidemic threshold. Consequently, assessing the robustness of complex networks not only allows us to compare the robustness among different network structures, but also gives insights about mitigation strategies plans and the design of future network infrastructures. However, even given the network topology, assessing the robustness of networks is difficult. On one hand, one way to measure the impact of the epidemic is by the number of infected nodes at steady state, which is not explicitly related to the topological characteristics of the networks. On the other hand, the epidemic threshold τ_c , used as a measure of the robustness of the network in the literature [6,7], does not take into account the number of infected nodes. Ultimately, the epidemic threshold is not adequate for assessing the network robustness.

In this paper, we propose a novel metric viral conductance (VC) to assess the robustness of networks with respect to the spread of SIS epidemics. The new metric considers the value of epidemic threshold and the number of infected nodes at steady state above the threshold. Thus the higher the value of VC, the lower the robustness of the network. One very interesting result obtained using VC, concerns the robustness of Barabási–Albert (BA) preferential attachment (PA) networks. While previous work stressed the fact that the epidemic threshold is close to zero in those networks, VC can quantify the fraction of infected nodes for increasing value of the effective spreading rates above the threshold. As a matter of fact, infinite Barabási–Albert (BA) preferential attachment (PA) networks with vanishing epidemic thresholds can still require strong epidemics to have major outbreaks, while in homogeneous random networks, an epidemic does not need to be very much beyond the threshold to cause major outbreaks. Due to the computational complexity, we propose a heuristic to actually compute the VC for large networks. Moreover, we derive upper and lower bounds for VC. We performed extensive simulations to validate the new metric, the bounds and the heuristic, and to show how the new heuristic outperforms our previous heuristic presented as work-in-progress in [1]. The numerical evaluation focuses on assessing the robustness of different types of networks, e.g. correlated preferential attachment (PA) networks, Watts–Strogatz (WS) small world networks [4] and Barabási–Albert (BA) preferential attachment (PA) networks [8]. Our results show that assortative PA networks have lower VC than disassortative PA networks when only the average node degree is preserved. Qualitatively, we compare between the robustness of Watts–Strogatz (WS) networks and Barabási–Albert (BA) preferential attachment (PA) networks, and we show cases where Watts–Strogatz (WS) networks are less robust than Barabási–Albert (BA) preferential attachment (PA) networks and vice versa. In addition, for a given irregular network, we rewired the network to make it almost regular by decreasing the variance of the node degrees, and we computed the VC before and after the regularization.

We summarize the contributions of the paper as follows:

- Introducing a new robustness measure VC with respect to the spread of SIS epidemics.
- Comparing between VC and the epidemic threshold and showing cases where the epidemic threshold fails to assess the robustness of networks.
- Showing that Barabási–Albert (BA) preferential attachment (PA) networks require strong epidemics to have major outbreaks.
- Introducing a computational heuristic for VC and upper and lower bounds.
- Presenting a framework to compare the robustness of Watts–Strogatz (WS) networks and Barabási–Albert (BA) preferential attachment (PA) networks.
- Applying VC to a weighted social network to evaluate the efficiency of mitigation strategies.

The rest of this paper is organized as follows. In Section 2, we review the literature and related work. In Section 3, we review the mathematical models of epidemic spread and their applications to different types of networks, and we compare the epidemic threshold with average fraction of infection. We propose the new robustness metric with respect to epidemic spread, the viral conductance VC, we compare between the classical measure and VC, and we discuss the robustness of Barabási–Albert (BA) preferential attachment networks in Section 4. In Section 5, a summary of some useful properties of infected nodes at steady state are presented. In Section 6, we propose a heuristic to compare viral conductance, and we derive upper and lower bounds for viral conductance. We present the numerical results in Section 7, and the main conclusions and future work are summarized in Section 8.

2. Related work

In biology, the history of epidemic modeling dates back to the eighteenth century when Bernoulli proposed the first deterministic epidemic model for smallpox, which was followed by other deterministic and stochastic epidemic models during the last two centuries [33–36]. Those models introduced the basic reproductive number R_0 , which is the average number of secondary infections due to a single infected case in the population. If $R_0 > 1$, the epidemic spreads in the population, while if $R_0 < 1$, the epidemic dies out without causing an outbreak. In complex network theory, the work in [7,16] found the epidemic threshold, which is a function of R_0 , for SIS epidemics using the heterogeneous mean field approach. The authors of [7,16] conclude that the threshold is a function of the heterogeneity of the network represented by the second moment of the node degree distribution. Later, an exact expression of the epidemic threshold emerged in the framework of spectral graph theory in [14,32]. The work in [7] also concludes that epidemics can spread on any scale-free network regardless of its effective infection rate when the number of nodes in the network approaches infinity. In other words, all scale-free networks are vulnerable to epidemic spread. However, finite scale-free networks have a non-zero epidemic threshold. Therefore, they are not always vulnerable to epidemic attacks.

Recently, biological epidemic models have been adapted to study the spread of viruses in technological networks as shown in [11–14]. For example, in [11], the authors propose deterministic and probabilistic models to predict the size of the infected population in homogeneous networks. Unfortunately, the models do not consider the heterogeneous structure of networks, and hence, they cannot be used to measure the robustness of generic networks with respect to the spread of epidemics. In [13], the authors propose a discrete-time epidemic model to predict the infection size, finding that the epidemic threshold is the reciprocal of the largest eigenvalue of the network adjacency matrix. Meanwhile, the N-Intertwined model in [14] reproduced the epidemic threshold in [13] based on a continuous time Markov chain process. Also the

N-Intertwined model explicitly reveals the role of the network in spreading epidemics. However, it does not introduce a metric to assess the robustness of networks.

In summary, using only the epidemic threshold as a measure for robustness is common practice: the larger the epidemic threshold, the more robust a network is against the spread of an epidemic, as in [6]. However, in Section 4, we will show some examples of networks where the epidemic threshold fails to assess their robustness, properly. Moreover, none of the works in the literature combine both the effective spreading rate and the infected population to describe the robustness of the network with respect to epidemic spreading. Specifically, then, we show that the robustness of networks depends on the value of the epidemic threshold τ_c (or its reciprocal) as well as the fraction of the infected population above the threshold.

3. Review of epidemic models

In this section, we review basic results about the spread of susceptible–infected–susceptible (SIS) epidemics on networks. The SIS infection model, which arose in mathematical biology, is often used to model the spread of epidemics [11,15,16,23,33,38], epidemic algorithms for information dissemination in unreliable distributed systems like P2P and ad-hoc networks [17,18], and propagation of faults and failures in BGP networks [19]. The SIS model analytically reveals how a node's state changes between the two S and I states in complex networks. To clarify, during the spread of an epidemic, a node is in one of the two states. First, an infected node can infect susceptible neighbors with infection rate β . Also, an infected node can cure itself with a cure rate δ and become susceptible to re-infection. Additionally, the ratio between β and δ is called the effective infection rate, $\tau = \beta/\delta$. Moreover, the epidemic threshold τ_c can be defined as follows: for effective spreading rates below τ_c , the epidemic contamination in the network dies out; the mean epidemic lifetime is of order $\log n$, while for effective spreading rates above τ_c , the epidemic is prevalent, i.e. a persistent fraction of nodes remains infected with the mean epidemic lifetime [15] of the order $e^{\tau n}$, for τ sufficiently large. Thus, the epidemic threshold was found to be $\tau_c = 1/\rho$, i.e. the reciprocal of the spectral radius ρ of the network adjacency matrix [13,14]. Another basic epidemic model in the literature is the SIR [37], which differs from the SIS model in many aspects. In the SIR model, the susceptible node becomes infected and later recovers without being susceptible to further infection. Therefore, this model has no steady state infected population fraction since all infected nodes recover, while in the SIS model, a steady state infected population fraction exists, and it depends on the effective infection rate and the network structure.

Below, we review the important results of the spread of SIS epidemics on regular, bi-partite, and random networks, and we also discuss how the epidemic threshold is related to the average fraction of infection at steady state.

3.1. Epidemic spread on regular, complete bi-partite, and random networks

We compare the fraction of infected nodes in the SIS model for different example networks and show that the value of the effective spreading rate τ determines for which network this fraction is higher. The first and second example networks belong to the class of regular and complete bi-partite networks, respectively, while the last example addresses two different types of random network.

3.1.1. Epidemic spread on regular networks

This model is based on a classical result for SIS models by Kephart and White [11,20]. We consider a connected network on N nodes where every node has degree k . We denote the number of

infected nodes in the population at time t by $Y(t)$. If the population N is sufficiently large, we can convert $Y(t)$ to $y(t) = Y(t)/N$, yielding a continuous quantity representing the fraction of infected nodes. Now the rate at which the fraction of infected nodes changes is due to two processes: susceptible nodes become infected and infected nodes become susceptible. Obviously, the rate at which the fraction $y(t)$ grows is proportional to the fraction of susceptible nodes, i.e. $1 - y(t)$. Therefore, for every susceptible node, the rate of infection is the product of the infection rate per node (β), the degree of node (k), and the probability that on a given link the susceptible node connects to an infected node ($y(t)$). Therefore, we obtain the following differential equation describing the time evolution of $y(t)$:

$$\frac{dy(t)}{dt} = \beta ky(t)(1 - y(t)) - \delta y(t) \quad (1)$$

The steady state solution $y_\infty(\tau)$ of Eq. (1) satisfies

$$y_\infty(\tau) = \frac{\beta k - \delta}{\beta k} = 1 - \frac{1}{\tau k} \quad (2)$$

Because an epidemic state only exists if $y_\infty(\tau) > 0$, we conclude that the epidemic threshold satisfies

$$\tau_c = \frac{\beta}{\delta} = \frac{1}{k} \quad (3)$$

Because for k -regular networks, the spectral radius of the adjacency matrix is equal to k , see [21], Eq. (3) is in line with the result in [13].

3.1.2. Epidemic spread on complete bi-partite networks

A complete bi-partite network $K_{M,N}$ consists of two disjoint sets containing respectively M and N nodes, such that all nodes in one set are connected to all nodes in the other set, while within each set no connections occur. Notice that (core) telecommunication networks often can be modeled as a complete bi-partite topology. For instance, the so-called double-star topology (i.e. $K_{M,N}$ with $M=2$) is quite commonly used because it offers a high level of robustness against link failures. For example, the Amsterdam Internet Exchange (see www.ams-ix.net), one of the largest public Internet exchanges in the world, uses this topology to connect its four locations in Amsterdam to two high-density Ethernet switches. Sensor networks are also often designed as complete bi-partite networks.

In [22,14], a model for epidemic spreading on the complete bi-partite network $K_{M,N}$ is presented. Using differential equations and two-state Markov processes, the authors show that above the epidemic threshold $\tau_c = 1/\sqrt{MN}$, the fraction of infected nodes at the steady state for $K_{M,N}$ satisfies

$$y_\infty(\tau) = \frac{(MN\tau^2 - 1)(M + N)\tau + 2}{\tau(M + N)(M\tau + 1)(N\tau + 1)} \quad (4)$$

It is easy to verify that for the case $M=N$, Eq. (4) reduces to Eq. (2), with $k=N$.

3.1.3. Epidemic spread on random networks

Many classes of random networks are described by their statistical properties, for example Erdős–Rényi (ER), Watts–Strogatz (WS) and Barabási–Albert (BA) preferential attachment (PA) networks. Erdős–Rényi networks (ER) are homogeneous networks with binomial node degree distribution and average node degree k . In Erdős–Rényi networks (ER), every node selects its neighbors with probability p independently. For large number of nodes N and for $pN=k$, the node degree distribution becomes Poisson distribution with a tail that decays exponentially for large node degrees. Meanwhile, Barabási–Albert networks BA are built using the preferential attachment (PA) mechanism in which each node prefers

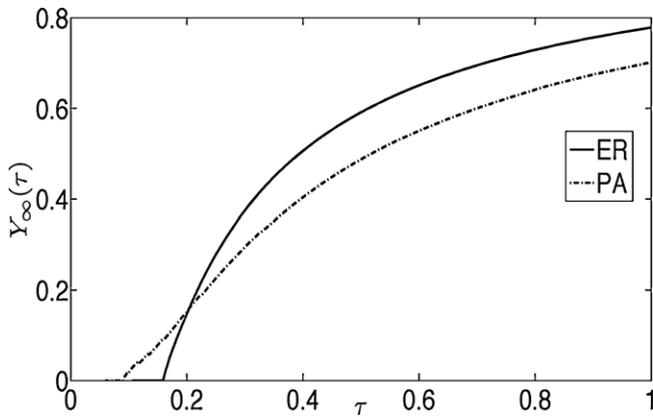


Fig. 1. Fraction of infected nodes for an ER network and a Barabási–Albert (BA) preferential attachment (PA) network with $N = 10^4$.

to connect with high node degree neighbors, and therefore the node degree distribution follows the power law distribution. We will use PA and BA equivalently to refer to Barabási–Albert preferential attachment networks. The literature shows that large BA networks are scale-free (SF) networks, and consequently are the most vulnerable networks to spread of epidemics. However, we study a counter example in which an Erdős–Rényi (ER) network can be more vulnerable to spread of epidemics than Barabási–Albert (BA) preferential attachment (PA) network given the same number of nodes. Fig. 1 shows the fraction of infected nodes at steady state $y_\infty(\tau)$ due to the spread of SIS epidemics for different effective infection rates $\tau = \beta/\delta$ on an ER network and a PA network with 10^4 nodes. We observed that the epidemic threshold of PA is smaller than the epidemic threshold of an ER network showing that PA network is more vulnerable than ER network. However, there is an inversion in $y_\infty(\tau)$ curves and after the inversion, the ER network has higher steady state infection fraction than the PA network. In this region, the ER network is more vulnerable than the PA network. Thus, the two networks have two opposing features and it is not trivial to measure their robustness.

For a given network, because the range of τ values for which the epidemic prevails is infinitely large, from now on, instead of considering the effective spreading rate τ , we look at the reciprocal of τ , that is the effective curing rate $s = \delta/\beta$. We are interested in $y_\infty(s)$, the fraction of infected nodes in steady state, as a function of the effective curing rate. Note that the behavior of $y_\infty(s)$ around $s = 0$ reflects the behavior of the original system for $\tau \rightarrow \infty$. Moreover, the value of $y_\infty(s)$ for $s = \rho$ (the reciprocal of the epidemic threshold) is 0. Such a conversion leads to a closed area under the $y_\infty(s)$ curve. We denote \bar{y}_∞ to be the average fraction of infection at steady state defined over $0 \leq s \leq \rho$ and it is given by $\bar{y}_\infty = (1/\rho) \int_0^\rho y_\infty(s) ds$.

3.2. Average infection fraction versus the epidemic threshold

Next, we hypothesize a case study that shows that the epidemic threshold is not capable to assess the robustness of networks having the same average node degree k within the same network class. Assume that the epidemic can have any effective cure rate $0 \leq s = (\delta/\beta) \leq \rho$, and every effective cure rate has a probability of infection at steady state $y_\infty(s)$. The average value of $y_\infty(s)$ over the defined range of s is \bar{y}_∞ infection fraction. Fig. 2 shows how the average infection fraction \bar{y}_∞ is inversely proportional to the maximum eigenvalue ρ (reciprocal of epidemic threshold) for networks that not only belong to the same class, but also have the same average node degree. Moreover, in Fig. 2(b), many networks with low average node degree and low maximum eigenvalue have a higher average infection fraction than other networks with high maximum

eigenvalues and high average node degree. However, the average infection fraction shows a general tendency to increase with the maximum eigenvalue over different average node degree values. This tendency reflects that as the number of links increases, the average node degree increases too, so the chance for an outbreak to take place becomes more relevant and causes a higher number of incidences at steady state.

We conclude that neither the fraction of infection at steady nor the epidemic threshold can solely describe the robustness of networks with respect to spread of SIS epidemics in networks. Therefore, we need a new metric to quantify the robustness of networks considering both the fraction of infection at steady state and the epidemic threshold.

4. Viral conductance

Based on the above conclusions and the mathematical background, in this section, we propose a new metric to quantify the robustness of networks with respect to spread of SIS epidemics.

4.1. Definition of viral conductance VC

Because we are considering the effective curing rate s as an independent variable, the area under the curve $y_\infty(s)$ is bounded. A logical way to consider the range of s values as well as the fraction of infected nodes is to evaluate the area under the $y_\infty(s)$ curve. We can now introduce a new robustness measure with respect to epidemic spread, viral conductance VC, that takes into account all values of s , and hence τ .

Definition 4.1. Viral conductance VC is a robustness measure of a given network G with respect to the spread of epidemics. It represents the average fraction of infected nodes for all types of epidemic attacks that are capable of producing outbreaks in the network. For non-negative continuous variable, $y_\infty(s) = 1/N \sum_{i \in N} v_\infty^i(s)$ where $v_\infty^i(s) = (\sum_{j \in \text{neighbors}(i)} v_\infty^j(s)) / (s + \sum_{j \in \text{neighbors}(i)} v_\infty^j(s))$ is the probability that node i is infected at the steady state [14]; mathematically, VC is defined as follows:

$$VC(G) = \int_0^\rho y_\infty(s) ds = \rho \bar{y}_\infty \tag{5}$$

where ρ is the spectral radius (i.e. maximum eigenvalue) of the adjacency matrix of network G and \bar{y}_∞ is the average value of the fraction of infected nodes for all $0 \leq s \leq \rho$.

We will now state two theorems for the viral conductance VC, of a network G .

Theorem 4.2. For regular networks G_k , where every node has k neighbors, $VC(G_k) = k/2$.

Proof. This follows directly from Eq. (2). ■

Theorem 4.3. For complete bi-partite networks $K_{M,N}$, $VC(K_{M,N})$ is as follows:

$$VC(K_{M,N}) = \frac{(M+N)\sqrt{MN} - MN}{M+N} + \frac{(M-N)(N \ln(M + \sqrt{MN}))}{M+N} - \frac{(M-N)(M \ln(N + \sqrt{MN}))}{M+N} + \frac{(M-N)(M \ln M - N \ln N)}{M+N} \tag{6}$$

Proof. This follows from applying Eq. (5) to Eq. (4). ■

The viral conductance VC can also be applied to random networks. For example in Fig. 1, the values of VC for Erdős–Rényi (ER)

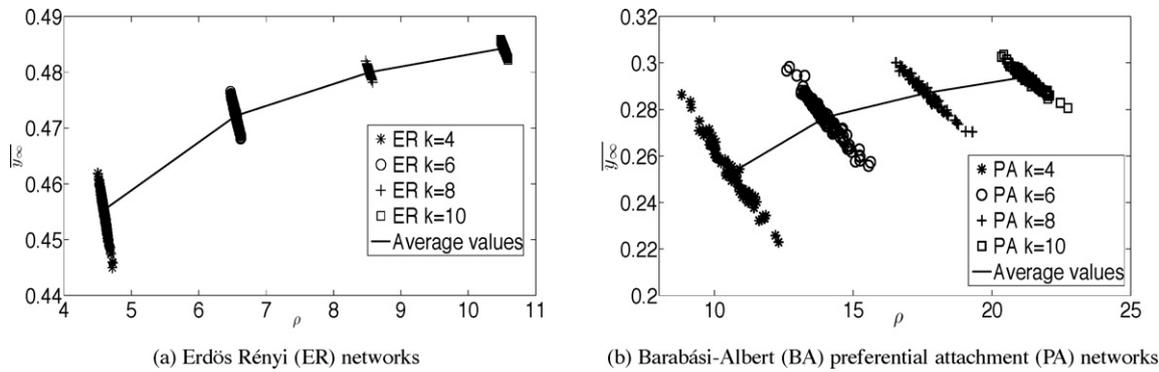


Fig. 2. The relationship between ρ and \bar{y}_∞ for Erdős-Rényi (ER) and Barabási-Albert (BA) preferential attachment (PA) networks with $N=10^4$ and $k=4, 6, 8$ and 10 . For network class and every average node degree k , there are 100 samples of networks. The solid line represents the average values of ρ and the average values of \bar{y}_∞ .

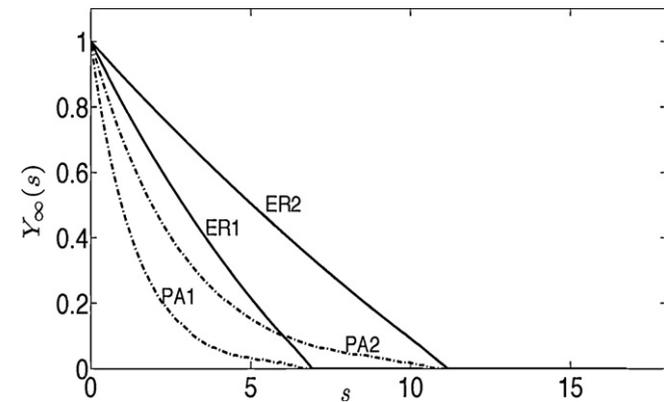


Fig. 3. Two examples of ER and PA networks with $N=10^4$ and with the same ρ (i.e. the same epidemic threshold).

and Barabási-Albert (BA) preferential attachment (PA) networks are 5.103 and 2.887, respectively. In this case, PA is more robust than ER.

4.2. Viral conductance versus the epidemic threshold

Traditionally, the epidemic threshold has been used to evaluate the robustness of networks with respect to spread of epidemics [6,7]. However, we present a case-study in which the epidemic threshold fails to assess the robustness of networks. Fig. 3 shows examples of networks with $N=10^4$ that almost have the same maximum eigenvalues (i.e. same epidemic thresholds), in which the epidemic threshold thus fails to accurately assess their robustness. Note that the difference between the maximum eigenvalue values in every pair of networks, {ER1, PA1} and {ER2, PA2}, is very small and in the order of 10^{-4} . Fig. 3 represents two ER networks and two PA networks. ER1 and ER2 networks have the same epidemic threshold as PA1 and PA2 networks, respectively.

Table 1 shows the numerical values of ρ and VC of the four networks discussed in Fig. 3. We notice that the value of ρ for both ER1 and PA1 networks is the same, while the VC value of ER1 network is higher than the VC value of PA1 network. The same observation is applied on ER2 and PA2 networks. The difference between VC and ρ is that VC represents the area under the $y_\infty(s)$ curve.

Table 1
The values of VC of the networks in Fig. 3 with the same ρ values.

Metric	ER1	PA1	ER2	PA2
ρ	7.16	7.16	10.98	10.98
VC	3.093	1.652	5.097	2.672

4.3. Paradoxical robustness of Barabási-Albert preferential attachment networks

We addressed the robustness of Erdős-Rényi (ER) networks and Barabási-Albert (BA) preferential attachment (PA) networks in Fig. 1 and we showed that there are two opposing features relate to the epidemic threshold and the fraction of infection at steady state. We summarize the opposing features as follows: On one hand, the PA network has a lower epidemic threshold than the ER network showing that the PA network is more vulnerable than ER network. On the other hand, the ER network has a higher fraction of infection at steady state than does the BA network. Therefore, looking independently look at the epidemic threshold or the steady state infection fraction to measure the robustness of networks is not sufficient. Additionally, to address the paradoxical robustness of PA networks, let us study the steady state infection curves for finite ER and PA networks as shown in Fig. 4. Because of the shrinking tail behavior of the fraction of infection at steady state, PA networks with very high ρ (very low epidemic threshold) would still have tiny fractions of infected population within the region beyond and away from the maximum eigenvalue (reciprocal of the threshold). Infinite PA networks with large maximum eigenvalues (vanishing epidemic threshold) can still require strong epidemics to have major outbreaks, while in ER networks, an epidemic does not need to be very much beyond the reciprocal of the epidemic threshold to cause a major outbreak.

5. Properties of steady state infected population fraction

Here, we summarize basic properties of the steady state infected population $y_\infty(s)$ presented in the literature, which are very useful

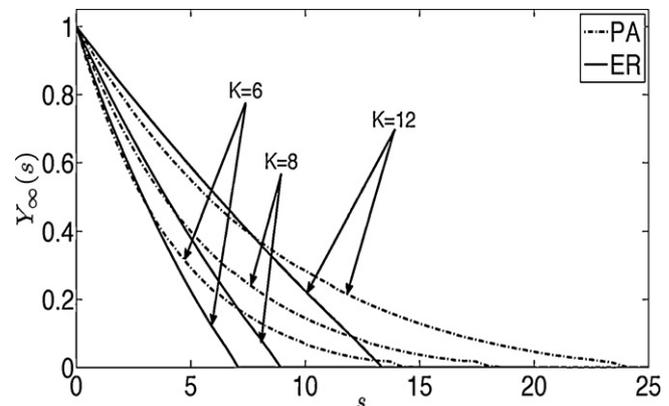


Fig. 4. Examples of the fraction of infection at steady state $y_\infty(s)$ given different values of $s=\delta/\beta$ for both ER and PA networks with $N=10^4$ and $k=6, 8$ and 12 .

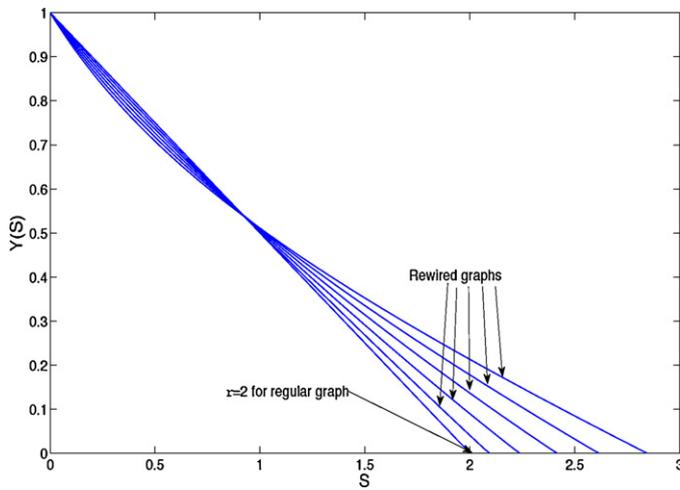


Fig. 5. The steady infected populations of a ring network and its rewired versions.

for the computational heuristic for VC in Section 6. We also show new results and ideas based on the basic properties of an infected population at steady state. First, the role of the epidemic threshold was studied in the literature and found that for any connected network G , denote ρ is the maximum eigenvalue of the adjacency matrix, such that $y_\infty(\rho) = 0$. Secondly, the fraction of infected population at the steady state $y_\infty(s)$ can be written as $y_\infty(s) = 1 - \sigma s + O(s^2)$, where $\sigma = (1/N) \sum_{i=1}^N 1/d_i$ and d_i is the degree of node i . Lastly, given a regular network with node degree $d = k$, the fraction of infected nodes $y_\infty(s)$ for $s = k/2$ equals $1/2$. This result directly follows from Eq. (2). For further reading and proofs, we refer the reader to [14]. Below, we show new results about $y_\infty(s)$ for $s = k/2$ (i.e. $\tau = 2/k$) for any network with average node degree k .

Lemma 5.1. For any complete bi-partite network and for $s = k/2$, the fraction of infected nodes $y_\infty(k/2)$ is bounded by 0.5147.

Proof. Let the total number of nodes in any complete bi-partite network be $T = M + N$. Then, substituting $M = T - N$ in Eq. (4) for $s = k/2 = (MN/M + N)$, we obtain the following equation:

$$y_\infty\left(\frac{k}{2}\right) = \frac{T^4 + T^3N - 3T^2N^2 + 4TN^3 - 2N^4}{2T^4 + NT^3 - T^2N^2} \quad (7)$$

Note that Eq. (7) is symmetric at $N = T/2$. By differentiating $y_\infty(k/2)$ with respect to N , we find that the N values at which the function obtains a maximum are $0.5T$, $2.5908T$, $-1.5908T$, $0.8587T$ and $0.1413T$. Note that we only consider solutions that satisfy the condition $0 \leq N \leq T$. Due to the symmetry of Eq. (7), the maximum value of $y_\infty(k/2)$ is 0.5147 for $N = 0.8587T$ and $N = 0.1413T$. ■

To address the effect of the network structure on the steady state probability of infection when $s = k/2$, we show a simple example on a ring network structure ($\rho = k = 2$), where we rewired every link towards a common node. Fig. 5 represents the $y_\infty(s)$ curve of every rewiring step. The figure shows that all the steady state infected population curves pass close to the point $((k/2), 0.5)$. In addition, we

performed extensive simulations to validate our assumption on different network classes with network sizes ranging from 100 nodes up to 3×10^5 nodes with different connectivities. All the simulations are averaged over 10^3 runs. We randomly rewired a 100-node regular network with $k = 10$ and a 100-node Barabási–Albert (BA) preferential attachment (PA) network with $k = 7.2$ each with 10^3 rewiring steps, and we show that $y_\infty(k/2)$ is very close to 0.5 as Table 2 shows. Moreover, we created 10^4 -node Barabási–Albert (BA) preferential attachment (PA) networks with $k = 8, 12, 16$ and 20. For every type of network, we computed the average and the variance values of $y_\infty(k/2)$ as in Table 2. The results show that the average of samples of $y_\infty(k/2)$ is very close to 0.5. Furthermore, we evaluated $y_\infty(s = k/2)$ for Watts–Strogatz (WS) networks with network size $N = 10^4$, average node degree $k = 4$ and 6, and rewiring probability $0 \leq p \leq 1$ as shown in Fig. 6(a). We noticed that Watts–Strogatz (WS) networks with $k = 4$ have a larger deviation from the value 0.5 than the networks with $k = 6$. Additionally, the deviation converges with the rewiring probability p reaches 1. Thus, the results show that $y_\infty(k/2)$ can be approximated to 0.5 with small deviation. To study the effect of the finite network size on the value of $y_\infty(k/2)$, we created Barabási–Albert (BA) preferential attachment (PA) networks with different sizes and different average node degrees, and we evaluated $y_\infty(s = k/2)$ as shown in Fig. 6(b). Again, the results validate our assumption. We believe that for any connected network $G(N, L)$ with average node degree k , for $s = k/2$, the fraction of infected nodes at steady state is $0.5 + O(\varepsilon)$ where $|\varepsilon| \ll 1$. Therefore, in Section 6, where we propose a heuristic for the new robustness metric with respect to spread of epidemics, we neglect the parameter ε and assume that $y_\infty(k/2) = 0.5$.

6. Computation of VC and bounds

For general networks with heterogeneous structure, we cannot analytically compute the fraction of infected nodes $y_\infty(s)$, and hence the viral conductance, explicitly. Moreover, computing the solution of $y_\infty(s)$ for $0 \leq s \leq \rho$ numerically is not feasible for large scale networks. Therefore, in this section, we propose a heuristic for computing the viral conductance for general networks. We will use the lemmas and theorems of the previous section to construct a heuristic to compute the new robustness metric, VC. Our objective is to make the heuristic as simple as possible to avoid computation complexity.

6.1. A heuristic for VC

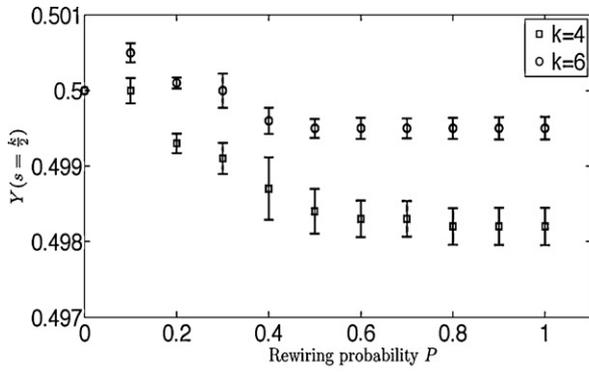
The heuristic mainly depends on fitting linear and non-linear functions passing through three main points on the steady state infected population curve $y_\infty(s)$. The three points $(s, y_\infty(s))$ are as follows:

- Point (0,1) where s equals 0 (i.e. $\delta = 0$) and hence the whole network is infected at steady state (i.e. $y_\infty(0) = 1$).
- Point $((k/2), 0.5)$ as discussed in Section 5.
- Point $(\rho, 0)$, which is the reciprocal of the epidemic threshold where the network is cured.

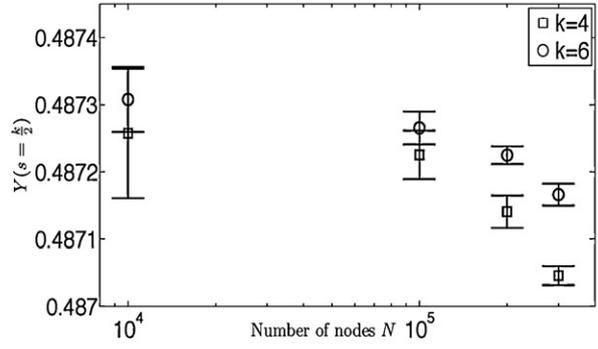
Table 2

The average and variance values of $y_\infty(s)$ are evaluated at $s = k/2$ for a randomly rewired regular network and different Barabási–Albert (BA) preferential attachment (PA) networks.

Network	N	k	Average	Variance
Randomly rewired regular	10^2	10	0.499	2.7843×10^{-7}
Preferential attachment	10^2	7.2	0.4849	2.59×10^{-5}
Preferential attachment	10^4	8	0.4885	2.34×10^{-5}
Preferential attachment	10^4	12	0.4883	2.01×10^{-5}
Preferential attachment	10^4	16	0.4881	1.806×10^{-5}
Preferential attachment	10^4	20	0.4881	1.8013×10^{-5}



(a) The value of $y_{\infty}(s = \frac{k}{2})$ for Watts-Strogatz (WS) networks given $N = 10^4$, $k = 4$ and 6 , and rewiring probability $0 \leq p \leq 1$.



(b) The value of $y_{\infty}(s = \frac{k}{2})$ for Barabási-Albert (BA) preferential attachment (PA) given $k = 4$ and 6 , and different network size N .

Fig. 6. Evaluation of the steady state infected population $y_{\infty}(s)$ for $s = k/2$ on Watts-Strogatz (WS) and Barabási-Albert (BA) preferential attachment (PA) networks given different network sizes N and different average node degrees k .

The basic heuristic equation is as follows:

$$y_{heuristic\infty}(s) = \begin{cases} \frac{y_{1\infty}(s) + y_{2\infty}(s)}{2} & 0 \leq s \leq \frac{k}{2} \\ \frac{y_{1\infty}(s) + y_{3\infty}(s)}{2} & \frac{k}{2} \leq s \leq \rho \end{cases} \quad (8)$$

The function $y_{heuristic\infty}(s)$ is defined over two intervals of the range of s . In each interval, $y_{heuristic\infty}(s)$ is the average of two fitting curves. For example, in the first interval where $0 \leq s \leq (k/2)$, $y_{1\infty}(s)$ is a decreasing exponential function, it continues to the second interval to the point $(\rho, 0)$ and therefore is constrained to pass by the three main points. The function $y_{1\infty}(s)$ is defined as follows:

$$y_{1\infty}(s) = \left(\frac{\rho - s}{\rho}\right) e^{-as} \quad (9)$$

To obtain the value of the exponent coefficient a , we solve $y_{1\infty}(s)$ at the point $((k/2), 0.5)$. Therefore, the value of a becomes as follows:

$$a = -\frac{2}{k} \ln\left(\frac{\rho}{2\rho - k}\right) \quad (10)$$

Next, the second function in the first interval $y_{2\infty}(s)$ is a linear decreasing function, proposed to equalize any underestimation due to the exponential function $y_{1\infty}(s)$. It is defined as follows:

$$y_{2\infty}(s) = 1 - \frac{s}{k} \quad (11)$$

In the second interval, the function $y_{3\infty}(s)$ is a decreasing exponential function that passes through the points $((k/2), 0.5)$ and $(\rho, 0)$, proposed to follow the tail of the $y_{\infty}(s)$ curve that depends on the irregularity of the network. Thus, the function $y_{3\infty}(s)$ is defined as follows:

$$y_{3\infty}(s) = b(\rho - s)e^{-cs} \quad (12)$$

The exponent coefficient c depends on the irregularity of the network since $\rho \geq k$ for all irregular networks [21]. Also, the total exponent of the exponential function of $y_{3\infty}(s)$ should be unitless; therefore, the exponent coefficient c has to hold a unit of $(\text{node degree})^{-1}$. To compute the value of c , we propose the following equation:

$$c = \frac{1}{2} \sqrt{\frac{-1 + \sum_{i=1}^{N-1} 1_{i \in ND}}{\rho k}} \quad (13)$$

$$1_{i \in ND} = \begin{cases} 1 & \text{if } i \in ND \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

where $1_{i \in ND}$ is a set function of the node degree, and ND is the set of node degrees that exists in the network, and therefore $\sum_{i=1}^{N-1} 1_{i \in ND}$ represents the irregularity of the network. Then, the constant b is computed as follows:

$$b = \frac{e^{c(k/2)}}{2\rho - k} \quad (15)$$

Note that for regular networks with degree $k(k = \rho)$, the functions $y_{1\infty}(s)$ and $y_{3\infty}(s)$ become linear and they are similar to $y_{2\infty}(s)$, which is the exact $y_{\infty}(s)$ curve for regular networks.

By integrating $y_{heuristic\infty}(s)$ in Eq. (8), we obtain the final mathematical formula for the $VC_{heuristic}$ as follows:

$$VC_{heuristic} = \frac{1}{2} \left[\frac{3k}{8} + \frac{1}{a} - \frac{1}{\rho a^2} - \frac{e^{-a\rho}}{a} - \frac{b\rho}{c} e^{-c\rho} + \frac{b\rho}{c} e^{-c(k/2)} + \frac{e^{-a\rho}}{\rho} \left(\frac{\rho}{a} + \frac{1}{a^2} \right) + b e^{-c\rho} \left(\frac{\rho}{c} + \frac{1}{c^2} \right) - b e^{-c(k/2)} \left(\frac{k}{2c} + \frac{1}{c^2} \right) \right] \quad (16)$$

6.2. Upper and lower bounds for VC

We formulated upper and lower bounds for VC, depending on the topological characteristics of the network, to avoid underestimating and overestimating the actual value of VC.

6.2.1. Upper bound

We know that the steady state infection population $y_{\infty}(s)$ is always a convex function [24] since connecting two points on the curve with a linear decreasing function renders the area under the linear function greater than the area under the actual curve (see Fig. 7). Therefore, we computed that area under the following linear functions: (1) a linear function that connects the points $(0,1)$ and $((k/2), 0.5)$, and (2) a linear function that relates the points $((k/2), 0.5)$ and $(\rho, 0)$. Thus, we formulated the upper bound as follows:

$$VC_{UB} = \int_0^{k/2} \left(1 - \frac{s}{k}\right) ds + \int_{k/2}^{\rho} \frac{1}{k - 2\rho} (s - \rho) ds \quad (17)$$

The VC upper bound VC_{UB} is as follows:

$$VC_{UB} = \frac{1}{4} (k + \rho). \quad (18)$$

6.2.2. Lower bound

Section 5 shows that the slope of the steady state $y_{\infty}(s)$ at $s = 0$ satisfies $-\sigma$, where $\sigma = (1/N) \sum_{i=1}^N 1/d_i$. Then using the line connecting the points $((k/2), 0.5)$ and $(k, 0)$ and the intersection point

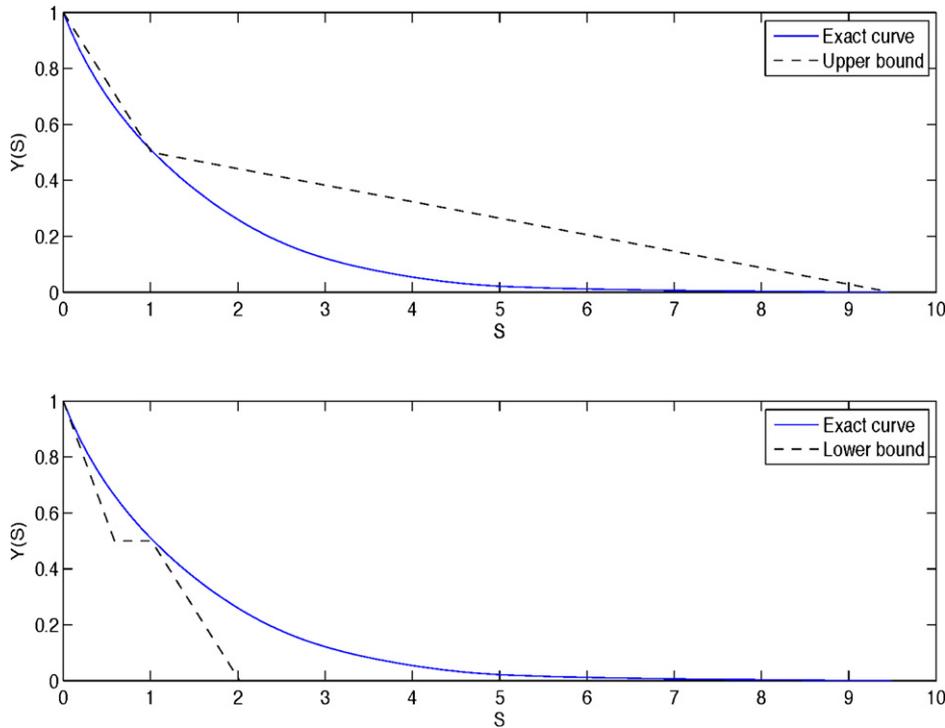


Fig. 7. An example of the upper and lower bounds of the viral conductance for the Abilene network with $N=886$, $k=2.022$ and $\rho=9.489$. The upper figure represents the upper bound VC_{UB} , while the lower figure shows the lower bound VC_{LB} .

between the tangent line $y_{\infty}(s) = 1 - \sigma s$ and the line $y_{\infty}(s) = 0.5$, we constructed a lower bound VC_{LB} as shown in Fig. 7. By applying Eq. (5), we found the following value for the lower bound:

$$VC_{LB} = \frac{1}{8} \left(\frac{1}{\sigma} + 3k \right) \quad (19)$$

For any regular network with node degree k , both the upper bound VC_{UB} and lower bound VC_{LB} lead to the same value of $VC = k/2$.

7. Numerical results

In this section, we numerically evaluate the robustness metric VC, the accuracy of the heuristic, and the bounds on different types of networks like synthetic networks with 10^4 nodes, real-world networks, and survey-based networks. The results are averaged over 10^3 runs.

7.1. Assortative and disassortative preferential attachment networks

In this subsection, we show how the new robustness measure VC can differentiate between assortative and disassortative networks in which node degree correlation is observed. Such correlation is important since correlated networks exist in the real world. For example, social networks are classified as assortative networks, while technological and biological networks are classified as disassortative networks [25]. In assortative networks, nodes with similar node degrees are connected together, while in disassortative networks, nodes with different nodes degrees are connected together. Accordingly, we generated assortative and disassortative preferential attachment (PA) networks using the algorithm in [26]. The algorithm starts with a connected network having $m_0 \ll N$ nodes. Every new node is connected to already existing nodes through two stages: In the first stage, a new node is connected to an existing node u with probability $\pi_u = (d_u / \sum_j d_j)$ where d_u is the degree of node u ; in the second stage, a new link between the new node

and one of the neighbors s of the chosen node u in the first stage is added with probability $p_s = d_s^{\alpha} / \sum_{v \in \Gamma_u} d_v^{\alpha}$, where α is an assortative tuning coefficient, and Γ_u is the set of neighbors of individual u chosen in the first stage. The generated assortative and disassortative preferential attachment (PA) networks have different node degree distributions. We will address the analytical and numerical studies of the robustness of correlated PA networks having the same node degree distribution in our future work.

To show how VC can distinguish among correlated networks and also that the heuristic is accurate and close to the exact VC, we evaluated VC and $VC_{heuristic}$ on correlated PA networks with 10^4 nodes given different average node degrees k as in Fig. 8. We noticed that $VC_{heuristic}$ increases with k , and it is close to the exact VC for both types of networks. In addition, both exact VC and $VC_{heuristic}$ of

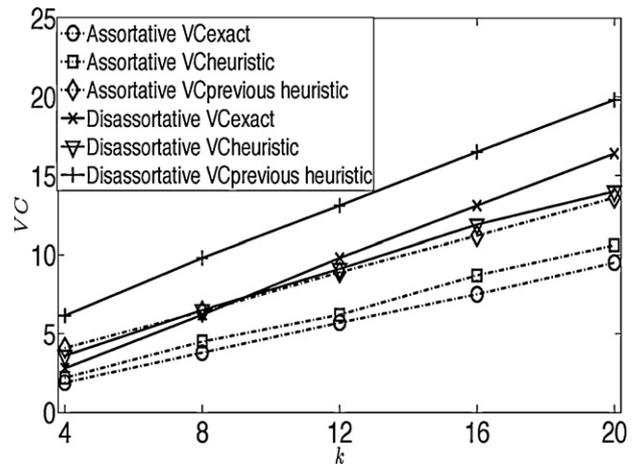


Fig. 8. The exact value of VC is compared with the heuristic approach $VC_{heuristic}$ in Eq. (16), and with the previous heuristic approach $VC_{previous heuristic}$ presented in [1] on assortative and disassortative preferential attachment (PA) networks with 10^4 nodes given different average node degrees $k = 4, 8, 12, 16$ and 20 .

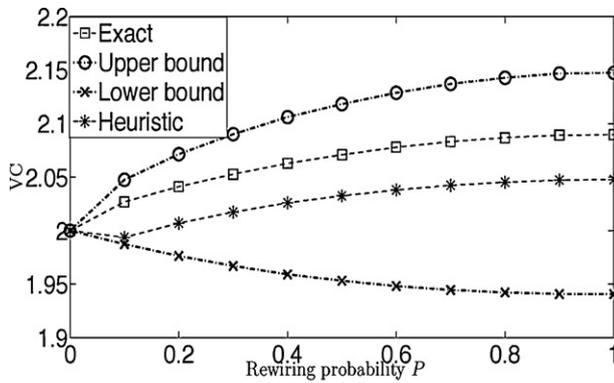
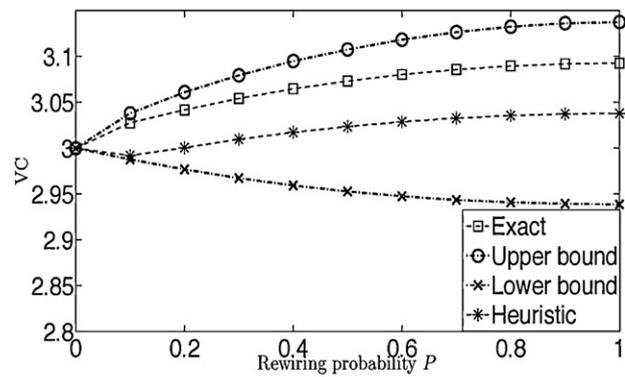
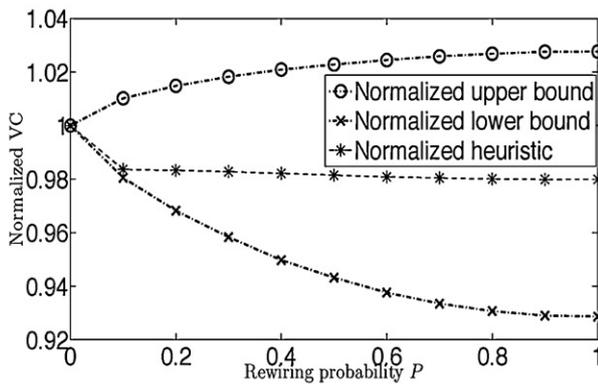
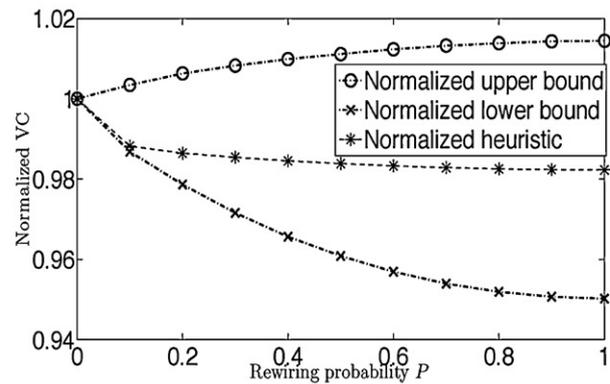
(a) Absolute values of VC given $k = 4$ (b) Absolute values of VC given $k = 6$ (c) Normalized values of VC given $k = 4$ (d) Normalized values of VC given $k = 6$

Fig. 9. Absolute values of exact VC , heuristic VC , and upper and lower bounds VC in Fig. 9(a) and (b); normalized values of heuristic VC , and upper and lower bounds VC with respect to exact value of VC in Fig. 9(c) and (d) for Watts–Strogatz (WS) networks given $N = 10^4$ and probability of rewiring $0 \leq p \leq 1$.

assortative networks are lower than their corresponding values for disassortative networks, showing that the new heuristic is capable of evaluating the robustness of networks. Moreover, we compared the new heuristic presented in this paper with our previous heuristic presented as work-in-progress in [1]. In Fig. 8, we observed that the new heuristic $VC_{heuristic}$ is closer to VC_{exact} than the previous heuristic $VC_{previous\ heuristic}$. Additionally, the values of $VC_{heuristic}$ of disassortative networks do not overlap with VC_{exact} and $VC_{heuristic}$ for assortative networks. Consequently, the new heuristic outperforms the previous heuristic, and it is capable of evaluating the robustness of large networks for which the computation of VC_{exact} may not be feasible.

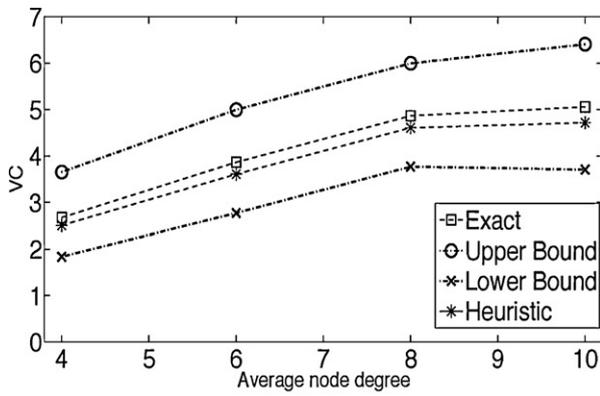
The robustness of assortative networks with respect to the spread of epidemics is also discussed in [25] showing that the giant component in assortative networks is smaller than in disassortative networks. On one hand, nodes with high degrees are connected together causing any epidemic to persist in the network. On the other hand, the epidemic survives in only a small portion of the network. Here, our results show that disassortative preferential attachment (PA) networks are more vulnerable than assortative preferential attachment (PA) networks. Notably, the algorithm used to create the correlated PA networks does not constrain the degree of the nodes in the networks. Therefore, few nodes with very large node degrees appear in the disassortative networks, and they are connected with low node degree nodes. Consequently they increase the network heterogeneity properties; once an epidemic reaches

any node with large node degree, a major outbreak takes place in the network.

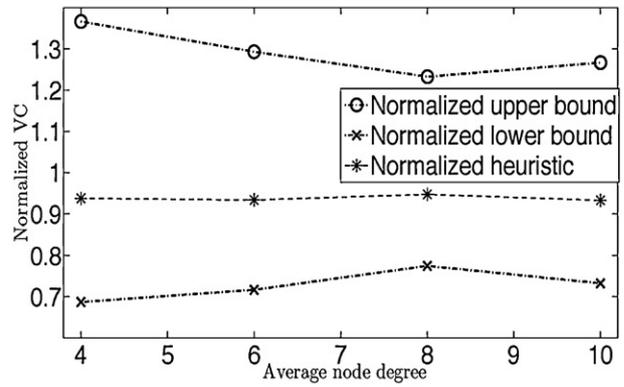
7.2. Watts–Strogatz small world model

We generated Watts–Strogatz (WS) small world networks [4] with a given number of nodes N and average node degree k . To create a WS small world network, we started with a k -regular network; each node has $k/2$ links that connect it to its nearest counterclockwise neighbors, while the other $k/2$ links connect the same node to its nearest clockwise neighbors. Given a rewiring probability p , we started rewiring the clockwise links for every node. We created WS networks given $N = 10^4$ nodes with average node degrees $k = 4, 6, 8,$ and 10 and different probability of rewiring p ranging from 0 to 1. Then, we used those networks to evaluate the heuristic value $VC_{heuristic}$ compared with the exact VC values, and to study the effect of rewiring the network links on the values of the VC .

Fig. 9(a) and (b) shows how the exact value of VC changes with a probability of rewiring p given $k = 4$ and 6 , respectively. For networks with given regular node degrees ($p = 0$), VC equals $k/2$. In addition, the exact VC value and $VC_{heuristic}$ non-linearly increase with the probability of rewiring p because the irregularity of the network increases with p . We also verified the validity of the upper and lower bounds VC_{UB} and VC_{LB} , noticing that for all networks with different average node degrees and different probability of rewiring, exact values of VC as well as heuristic values $VC_{heuristic}$ are



(a) Absolute values of VC given $k = 4, 6, 8$ and 10



(b) Normalized values of VC given $k = 4, 6, 8$ and 10

Fig. 10. Absolute values of exact VC , heuristic VC , and upper and lower bounds VC in Fig. 10(a); normalized values of heuristic VC , and upper and lower bounds VC with respect to exact value of VC in Fig. 10(b) for Barabási–Albert BA networks given $N = 10^4$ and different average node degrees.

bounded. To evaluate the deviation of $VC_{heuristic}$, the upper bound and the lower bound, we normalized the results with respect to the exact value of VC as shown in Fig. 9(c) and (d) for $k=4$ and 6 , respectively. We observed that the deviation of the heuristic value of VC is bounded with the increase of rewiring probability p . All the above analysis of VC was also applied to the WS small world networks with $k=8$ and 10 with the same observations as discussed in this subsection.

7.3. Barabási–Albert preferential attachment network model

We generated Barabási–Albert (BA) preferential attachment (PA) networks as follows: we started with a small number m_0 of disconnected nodes; next we connected a new node to an existing node u with probability $\pi_u = d_u / \sum_j d_j$, where d_u is the node degree of the existing node u . The generated BA network is characterized as an uncorrelated Barabási–Albert (BA) preferential attachment (PA) network. We evaluated VC on the generated BA networks with $N = 10^4$ and $k = 4, 6, 8$ and 10 as shown in Fig. 10(a). The value of VC increases as k increases, and also the heuristic value of VC is close to its corresponding VC value. Moreover, both exact and heuristic VC values are bounded. Also, we evaluated the deviation of the heuristic VC values, and the upper and lower bounds by computing their normalized values with respect to the exact values of VC as shown in Fig. 10(b). We found that the heuristic slightly deviates from its exact value, and therefore it perfectly estimates the exact value of VC .

7.4. VC versus ρ

Here, we numerically present the differences between VC and ρ as shown in Fig. 11, which comprises models with 100 samples of results for a given network class with a certain average node degree k . In general, the values of VC tends to increase with ρ , and simultaneously both measures increase with the average node degrees. Note that within a module of samples, VC does not always need to increase with ρ . We also observed that the slope for Watts–Strogatz (WS) networks results is higher than the slope in Barabási–Albert (BA) preferential attachment (PA) networks results. To relate the slope with VC and ρ , recall that $VC = \bar{y}_\infty \rho$, and therefore, the slope of the trend line is \bar{y}_∞ . To address the differences between VC and ρ , let us compare the robustness of a WS network with $k=8$ and a PA network with the same value of k as shown in the black box in Fig. 11. The value of ρ of the PA network is at least twice the value of ρ of WS network, and therefore the value of \bar{y}_∞ of PA net-

work is at most half its corresponding value of WS network. Such a trade-off was discussed earlier in the paper, so on one hand, we cannot measure the robustness of networks by considering only ρ or the average infection \bar{y}_∞ , yet on the other hand, VC combines both ρ and \bar{y}_∞ to account for them to measure the robustness of networks. Moreover, for any average node degree in PA networks, the samples widely scatter in the direction of ρ than VC in a given module, while samples from WS networks do not scatter largely. As a result, for a given value of ρ , it is not difficult to obtain distinct values of VC for different networks. For example, for a given value of ρ , WS networks are less robust than PA networks. Therefore, it is not necessary that PA networks are always more vulnerable than WS networks, and also ρ is not a unique robustness measure with respect to the spread of epidemics.

7.5. Internet AS-level networks

Next, we apply VC to measure networks like the Internet AS-level networks. Specifically, we employed three different networks, SKITTER [27] with 9204 nodes and $k=6.29$, BGP [28] with 17446 nodes and $k=4.68$, and WHOIS [29] with 7485 and $k=15.22$, to measure their robustness and to study the effect of rewiring the networks with the value of VC . The rewiring process aims to decrease the variance of node degree by regularizing the network using the algorithm in [30]. Next, we compared the robustness

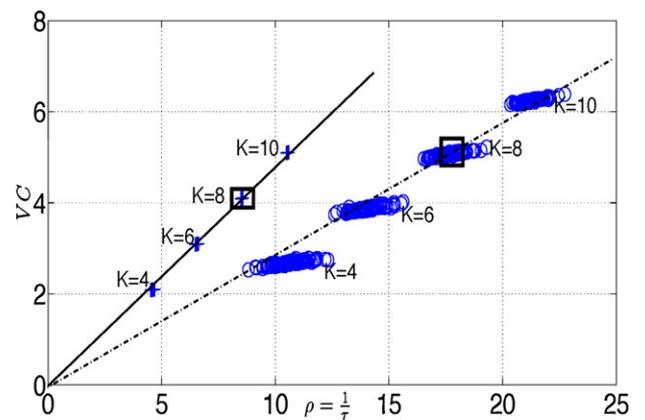


Fig. 11. Numerical relationship between VC and $\rho = 1/\tau$ (the reciprocal of the epidemic threshold) collected from Watts–Strogatz (WS) networks ‘+’ with rewiring probability $p=1$ and Barabási–Albert (BA) preferential attachment (PA) networks ‘o’ with $N = 10^4$ given different average node degrees.

Table 3
 $VC_{heuristic}$, VC_{UB} , VC_{LB} , ρ and κ of Internet AS-level networks and their regularized networks.

	VC_{LB}	$VC_{heuristic}$	VC_{UB}	ρ	Heterogeneity κ
SKITTER	2.6041	4.3661	21.4559	79.53	42.0340
Regularized SKITTER	3.1415	3.1428	3.1559	6.33	6.3282
WHOIS	6.0708	9.7964	41.5197	150.86	135.6138
Regularized WHOIS	7.6082	7.61	7.6130	15.23	15.224
BGP	1.9750	3.2877	19.4348	73.06	57.2345
Regularized BGP	2.3334	2.3385	2.3549	4.73	4.7221

of each network in two cases: the original network and its regularized network. Table 3 shows that regularized networks have a lower $VC_{heuristic}$ than original networks. In addition, we used the metric κ to assess the heterogeneity of the irregular and regularized networks. The metric κ was introduced in the literature as the ration between the second moment and the first moment of the node degree distribution. After regularizing a network, the heterogeneity of the network decreases leading to decrease the degrees of nodes with large degrees and to increase the degrees of nodes with low degrees. Therefore, all the node degrees are centered around the average degree value k with low variance in node degrees, while the maximum degree and the minimum degree are $\lceil k \rceil$ and $\lfloor k \rfloor$, respectively. Consequently, $y_{\infty}(s)$ curve of the regularized networks shrinks dramatically and approaches a linear decreasing function. Assessing the $VC_{heuristic}$ and κ , both metrics confirm that regular networks are more robust than irregular networks against the spread of epidemics. Moreover, the values of ρ for irregular and regularized networks in Table 3 show that both values of VC and ρ decrease when the networks are regularized. For every network, this behavior occurs due to the decrease in the maximum node degree, which upper-bounds the spectral radius of the network, resulting in a

largely decreased area under the curve, and therefore VC decreases too.

7.6. Survey-based network

In this subsection, we applied VC to a social network created through a survey to study the spread of epidemics in rural regions. The survey was conducted in Clay Center, the county seat of Clay County in the State of Kansas, and the network was created based on the responses of Clay Center residents. The survey included questions about three main places, $\{R, W, G\}$, that the residents visit, and questions about three levels of contact each respondent makes with other individuals. The three levels of contact were defined as follows: (1) *proximity contact*, which happens when another person is passing within five feet, (2) *direct-low contact*, which happens when a person is directly touching another person for short time period, and (3) *direct-high contact*, which happens when a person is directly touching another person for a long time period. Every respondent i provided the number of contacts $n_{x,i}$ for every contact level x . We used the survey responses to create a weighted contact network with 138 nodes (respondents) and 9222 links (contacts) as shown in Fig. 12. The weight between respondent i and respondent

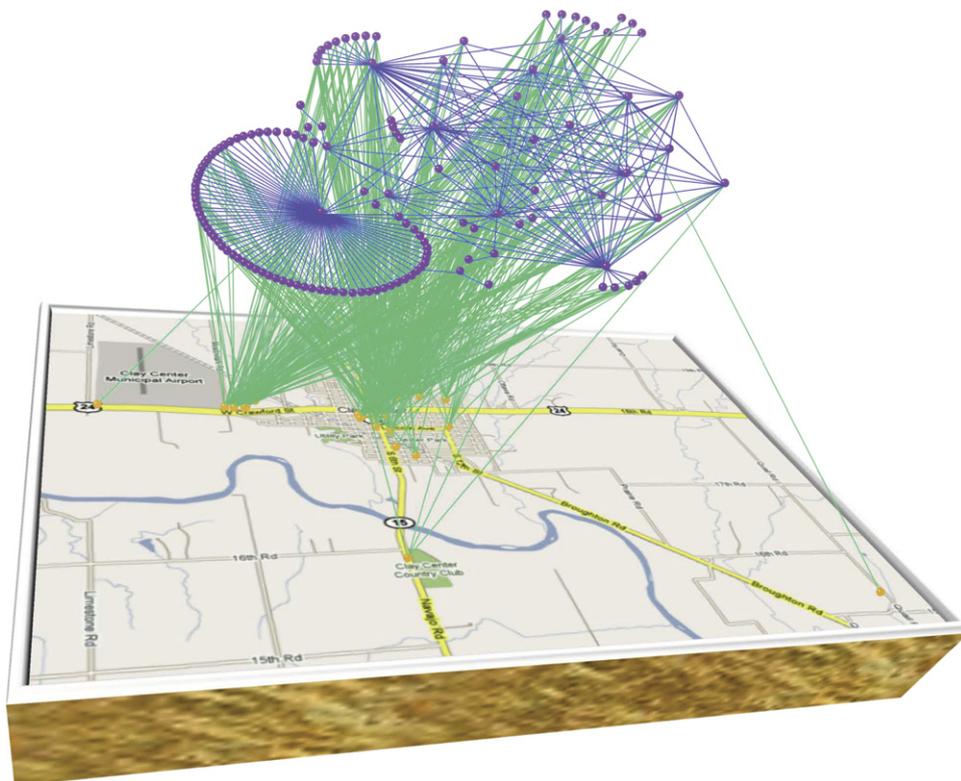


Fig. 12. The survey-based network of individuals and popular locations in Clay Center, Kansas, where the nodes (survey respondents) in the cloud network are connected via green edges to the locations in Clay Center according to the survey responses. The map is courtesy of Google.

Table 4

The value of VC for the survey-based social network in case of the absence of mitigation strategies and the most two effective mitigation strategies.

Mitigation strategy	VC	ρ	Immunized nodes
None (weighted network)	71.4985	31.91	0
High strength node immunization	49.1791	23.89	14
High strength node immunization in location R	49.4439	24.08	14

j is the average of the three contact levels. For contact level x , we proposed the following equation:

$$w_{x,i,j} = (1 - (1 - \mu_{i,j}\pi_x)^{n_{x,i}})(1 - (1 - \mu_{i,j}\pi_x)^{n_{x,j}}) \quad (20)$$

where π_x is a constant that depends on the level of contact x and $\mu_{i,j}$ quantifies the location responses for both respondents i and j as follows:

$$\mu_{i,j} = \frac{1 + l_{i,j}}{1 + d} \quad (21)$$

where d is the total number of locations, and $l_{i,j}$ represents number of common locations that respondents i and j used to visit. For more details about the survey questions and the link weights, we refer the reader to [31].

We applied the measure VC to quantify the robustness of the social network with respect to the spread of SIS epidemics. We also applied some mitigation strategies to the network and used VC to rank them. Also, we studied mitigation strategies where 10% of the nodes were immunized. From a network point of view, the immunized nodes were removed from the network. Immunized nodes were selected based on (1) node strength (the sum of link weights of a node), (2) random selection, (3) random selection from a specific location, and (4) highest strength nodes from a specific location. Table 4 shows the values of VC and ρ in case of the absence of mitigation strategies and the most two effective mitigation strategies. Clearly, we notice that VC values, when mitigation strategies are applied, are lower than the VC value in absence of mitigation. Also, the highest strength mitigation strategy outperforms other mitigation strategies since it has the lowest VC value because highest strength nodes play a major role in spreading any epidemic. Also, the highest strength mitigation strategy outperforms the highest strength mitigation strategies that are applied at different locations because the former considers all nodes with the highest strength in the network regardless of their locations. This result agrees with the effect of mitigation strategies presented in the literature (for example see [39]). Therefore, the highest strength mitigation strategy has the best effectiveness for reducing the spread of epidemics. Observing the values of robustness metrics in our example, we see that both VC and ρ rank the mitigation strategies similarly.

7.7. Results summary

We summarize the above results and analysis of robustness with respect to spread of epidemics in the following conclusions.

1. *Viral conductance is a better measure than the epidemic threshold for robustness of networks:* VC incorporates the fraction of infected nodes at steady state for all possible infection strengths.
2. *Increasing the probability of rewiring decreases the robustness of Watts–Strogatz (WS) networks:* The initial regular network in the Watts–Strogatz (WS) model has the lowest value of VC, and therefore it is the most robust of any other obtained network given the probability of rewiring $0 < p \leq 1$.
3. *VC_{heuristic} is close to the exact value of VC:* The proposed heuristic satisfies the basic requirements of simplicity and high accuracy in addressing solutions for any expensive computation quantity.

4. *VC_{UB} and VC_{LB} effectively bound VC and VC_{heuristic} from above and from below, respectively:* Bounds give the feasible region in which the value of VC is predictable.

5. *Our numerical results show that the regular structure of a network has a minimum VC value compared to any other structure:* Given N

nodes and L links, we can obtain $\binom{N}{2}^L$ different network

structures. We believe that the regular structure of a network is the most robust to any spread of epidemic.

8. Conclusions and future work

This paper aims to introduce a new measure, viral conductance VC, to assess the robustness of complex networks with respect to spread of epidemics. Viral conductance integrates the fraction of infected population at the steady state for all possible effective infection strengths an epidemic may possess. We compared the epidemic threshold with both the average fraction of infection at steady state and the viral conductance, and we show that the epidemic threshold is not an adequate measure to assess the robustness of networks. We also numerically found that when the effective cure rate equals half the average node degree, the fraction of infection at steady state is almost half the population. We verified our finding through extensive simulations on different networks with different sizes. Based on the previous property, we introduce a computational heuristic for the viral conductance, which is a function of the infected population and network characteristics. The heuristic aims to reduce the computational complexity of calculating the infected population at the steady state. In addition, we derive upper and lower bounds for the new measure, and we derive the analytical expression for viral conductance of both regular networks and bi-partite networks. We applied the new robustness measure to different types of network structures, and found that in Watts–Strogatz networks, the increase in probability of rewiring decreases the robustness of networks. Additionally, we found that the irregularity in node degrees decreases the robustness of the network. Moreover, within the considered set of correlated preferential attachment networks, the assortative structure of preferential attachment networks is more robust than the disassortative structure of preferential attachment networks. For all tested networks, the heuristic value perfectly approximates the exact value of the viral conductance and both are bounded using the proposed upper and lower bounds. Furthermore, the new robustness measure shows the effectiveness of different mitigation strategies on social networks.

Future work will explicitly focus on studying the robustness of correlated preferential attachment (PA) networks having the same node degree distribution. Additionally, given the number of nodes and links, VC can be used to design networks that maximize the robustness of the network with respect to the spread of epidemics. Moreover, the concept of viral conductance can be extended to measure the robustness with respect to the spread of epidemics that follow the susceptible/infected/recovered (SIR) epidemic model. We will also address the finite size effect on viral conductance, and we will analytically prove that the fraction of infection at steady state in the SIS model is almost half, when the effective curing rate equals one half of the average node degree.

Acknowledgments

This research was partially supported by the National Agricultural Biosecurity Center at Kansas State University, the Netherlands Organization for Scientific Research (NWO) under project number 643.000.503, and the National Science Foundation (NSF) under award number 0841112. The authors would like to thank Jasmina Omic and Phillip Schumm for the helpful discussion and Ali Sydney for creating the survey-based Clay Center network map.

References

- [1] R.E. Kooij, P. Schumm, C. Scoglio, M. Youssef, A new metric for robustness with respect to virus spread, in: L. Fratta, et al. (Eds.), Proceedings of the 8th International IFIP-TC 6 Networking Conference, NETWORKING 2009, LNCS 5550, 2009, pp. 562–572.
- [2] P. Erdős, A. Rényi, On random graphs, *Publicationes Mathematicae* 6 (1959) 290–297.
- [3] R. Albert, H. Jeong, A. Barabási, Error and attack tolerance of complex networks, *Nature* 406 (2000) 278–282.
- [4] D. Watts, S. Strogatz, Collective dynamics of small-world networks, *Nature* 393 (1998) 440–442.
- [5] M. Youssef, C. Scoglio, On graph-based characteristics of optimal overlay topologies, *The International Journal of Computer and Telecommunications Networking, COMNET* 53 (2009) 913–925.
- [6] A. Jamakovic, R.E. Kooij, P. Van Mieghem, E. van Dam, Robustness of networks against the spread of viruses: the role of the spectral radius, in: The 13th Annual Symposium of the IEEE/CVT Benelux, Liège, Belgium, 2006.
- [7] R. Pastor-Satorras, A. Vespignani, Epidemic dynamics and endemic states in complex networks, *Physical Review E* 63 (2001) 066117.
- [8] A.-L. Barabási, A. Réka, Emergence of scaling in random networks, *Science* 286 (1999) 509–512.
- [9] S. Towers, Z. Feng, Pandemic H1N1 influenza: predicting the course of a pandemic and assessing the efficacy of the planned vaccination programme in the United States, *Euro Surveillance* 14 (41) (2009).
- [10] V. Colizza, et al., Estimate of novel influenza A/H1N1 cases in Mexico at the early stage of the pandemic with a spatially structured epidemic model, *PLoS Currents: Influenza* (November) (2009).
- [11] J.O. Kephart, S.R. White, Direct-graph epidemiological models of computer viruses, in: IEEE Computer Society Symposium on Research in Security and Privacy, 1991, pp. 343–359.
- [12] Z. Chen, C. Ji, Spatial-temporal modeling of malware propagation in networks, *IEEE Transactions on Neural Networks* 16 (2005) 1291–1303.
- [13] Y. Wang, D. Chakrabarti, C. Wang, C. Faloutsos, Epidemic spreading in real networks: an eigenvalue viewpoint, in: 22nd Symposium in Reliable Distributed Computing, Florence, Italy, 2003.
- [14] P. Van Mieghem, J. Omic, R.E. Kooij, Virus spread in networks, *IEEE/ACM Transactions on Networking* 17 (2009) 1–14.
- [15] A. Ganesh, L. Massoulié, D. Towsley, The effect of network topology on the spread of epidemic, in: Proceedings of the IEEE INFOCOM.05, Miami, FL, 2005.
- [16] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, *Review Letters* 86 (2001) 3200–3203.
- [17] D. Chakrabarti, J. Leskovec, C. Faloutsos, S. Madden, C. Guestrin, M. Faloutsos, Information survival threshold in sensor and P2P networks, in: Proceedings of the IEEE INFOCOM.07, 2007.
- [18] P. Eugster, R. Guerraoui, A. Kermarrec, L. Massoulié, From epidemics to distributed computing, *IEEE Computer* 37 (2004) 60–67.
- [19] F. Coffman, Z. Ge, V. Misra, D. Towsley, Network resilience: exploring cascading failures within BGP, in: Proceedings of the 40th Annual Allerton Conference on Communications, Computing and Control, 2002.
- [20] J.O. Kephart, G.B. Sorkin, M. Swimmer, An immune system for cyberspace, in: Proceedings of the Conference on Systems, Man, and Cybernetics IEEE International: Computational Cybernetics and Simulation, vol. 1, 1997.
- [21] D. Cvetkovic, M. Doob, H. Sachs, Spectra of Graphs, Theory and Applications, 3rd ed., Johan Ambrosius Barth Verlag, Heidelberg, 1995.
- [22] M. Newman, A. Barabási, D. Watts, The Structure and Dynamics of Networks, Princeton University Press, Princeton, NJ, 2006.
- [23] P. Schumm, C. Scoglio, T. Easton, D. Gruenbacher, Epidemic spreading on weighted contact networks, in: Proceedings of the BIONETICS'07, Budapest, Hungary, 2007.
- [24] P. Van Mieghem, J. Omic, In-homogeneous Virus Spread in Networks, Delft University of Technology, Report 2008081, 2008.
- [25] M.E.J. Newman, Assortative mixing in networks, *Physical Review Letter* 89 (2002) 208701.
- [26] Q. Guo, T. Zhou, J.-G. Liu, W.-J. Bai, B.-H. Wang, M. Zhao, Growing scale-free small-world networks with tunable assortative coefficient, *Physica A* 371 (2006) 814–822.
- [27] K.C. Claffy, T.E. Monk, D. McRobb, Internet tomography, *Nature*, 1999, <http://www.caida.org/tools/measurement/skitter/>.
- [28] University of oregon routeviews project, <http://www.routeviews.org/>.
- [29] Internet routing registries, <http://www.irr.net/>.
- [30] V. Nikiforov, Eigenvalues and degree deviation in graphs, in: Elsevier Linear Algebra and its Applications, 2006, p. 414.
- [31] C. Scoglio, et al., Efficient mitigation strategies for epidemics in rural regions, *PLoS One* 5 (7) (2010) e11569.
- [32] R. Durrett, Some features of the spread of epidemics and information on a random graph, *PNAS* 107 (2010) 10.
- [33] R. Anderson, R. May, Infectious Diseases of Humans: Dynamics and Control, Oxford Press, 1991.
- [34] G. Macdonald, The analysis of equilibrium in malaria, *Tropical Diseases Bulletin* (1952) 49.
- [35] R. May, R. Anderson, Population biology of infectious diseases. Part II, *Nature* (1979) 280.
- [36] R. Anderson, Population ecology of infectious disease agents, in: Theoretical Ecology, 1981.
- [37] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society of London A* 115 (1927) 700–721.
- [38] J.D. Murray, *Mathematical Biology: An Introduction*, Springer, 2002.
- [39] R. Pastor-Satorras, A. Vespignani, Immunization of complex networks, *Physical Review E* (2002) 65.