This is the accepted manuscript made available via CHORUS. The article has been published as:

Traffic-driven epidemic spreading in correlated networks
Han-Xin Yang, Ming Tang, and Ying-Cheng Lai
Phys. Rev. E 91, 062817 — Published 29 June 2015
DOI: 10.1103/PhysRevE.91.062817
Traffic-driven epidemic spreading in correlated networks

Han-Xin Yang,¹,∗ Ming Tang,²,† and Ying-Cheng Lai³
¹Department of Physics, Fuzhou University, Fuzhou 350108, China
²Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 610051, China
³School of Electrical, Computer and Energy Engineering, Arizona State University, AZ 85287, USA

In spite of the extensive previous efforts on traffic dynamics and epidemic spreading in complex networks, the problem of traffic-driven epidemic spreading on correlated networks has not been addressed. Interestingly, we find that the epidemic threshold, a fundamental quantity underlying the spreading dynamics, exhibits a non-monotonic behavior in that it can be minimized for some critical value of the assortativity coefficient, a parameter characterizing the network correlation. To understand this phenomenon, we use the degree-based mean-field theory to calculate the traffic-driven epidemic threshold for correlated networks. The theory predicts that the threshold is inversely proportional to the packet-generation rate and the largest eigenvalue of the betweenness matrix. We obtain consistency between theory and numerics. Our results may provide insights into the important problem of controlling/harnessing real-world epidemic spreading dynamics driven by traffic flows.

PACS numbers: 89.75.Hc, 05.70.Ln, 05.60.-k

I. INTRODUCTION

Both epidemic spreading [1–13] and traffic transportation [14–20] are two types of fundamental dynamical processes on complex networks. The past fifteen years have witnessed a great deal of effort in exploring and understanding these processes. Intuitively, spreading should be greatly affected by the transportation dynamics as the traffic flows determine the paths along which a virus or information can propagate. Nevertheless, for a long time, the two types of dynamical processes were studied separately. The first series of works to incorporate traffic dynamics into epidemic spreading were done using the metapopulation model [21–27]. In particular, consider a spatially extended system of a large number of individuals. In a metapopulation model, all the individuals are grouped into a number of spatially structured subpopulations interacting with each other in a manner that can be described as a network. The links in the network represent the traveling paths of individuals across different subpopulations. An infected individual can infect other individuals in the same subpopulation. The metapopulation model has been used to simulate the spread of human and animal diseases (such as SARS and H1N1) among different cities. In 2009, Meloni et al. proposed another traffic-driven epidemic spreading model [28], where each node corresponds to a router (e.g., in a computer network) and the virus can spread among nodes through the transmission of packets. A susceptible node will be infected with certain probability every time it receives a packet from an infected neighboring node. This model is quite suitable for studying computer virus propagation.

The model of Meloni et al. [28] has become a prototypical framework to address a variety of issues in traffic-driven epidemic spreading. For traffic flow dynamics on complex networks, an essential ingredient is the routing strategy. In the model, one can demonstrate that epidemic spreading can be modulated or controlled through a proper choice of the local routing strategy [29] or global routing protocol [30]. For a fixed routing strategy, the network structure can also affect the spreading dynamics. For example, by increasing the average network connectivity, one can effectively delay epidemic outbreaks [31]. Furthermore, the epidemic threshold can be enhanced by deliberately removing a subset of links associated with the large-degree nodes or the edges with the largest algorithmic betweenness [32].

In previous studies of traffic-driven epidemic spreading, the networks were assumed to be uncorrelated. In such a network, the average degree of the neighboring nodes is uniform across nodes. However, many real-world networks display various degrees of mixing patterns [33]. Qualitatively, a network is assortatively (disassortatively) mixing if high-degree nodes tend to connect with high-degree (low-degree) nodes. Quantitatively, degree mixing can be characterized by the assortativity coefficient [33]:

\[ c = \frac{M^{-1} \sum_{i,j,k} j_i k_i - [M^{-1} \sum_{i,j} (\frac{1}{2} j_i + k_i)]^2}{M^{-1} \sum_i \frac{1}{2} (j_i^2 + k_i^2) - [M^{-1} \sum_{i,j} (\frac{1}{2} j_i + k_i)]^2}, \]

where \( j_i, k_i \) are the degrees of the nodes at the ends of the \( i \)th edge, \( M \) is the number of edges in the network, and \( i = 1, \ldots, M \). For standard network models such as the Erdős-Rényi random graphs [34] and the Barabási-Albert scale-free networks [35], the assortativity coefficient \( c \) is zero, indicating complete lack of degree correlation. For many social networks, the assortativity coefficient \( c \) is positive. However, technological and biological networks tend to be disassortative with negative values of \( c \).

In this paper, we study quantitatively how degree mixing, the most pronounced feature of correlated networks, affects traffic-driven epidemic spreading. Interestingly, we find a non-monotonic behavior in that the threshold can be minimized for some critical value of the assortativity coefficient. To understand this phenomenon, we employ a degree-based mean-field theory, which allows us to calculate the traffic-driven epidemic threshold in correlated networks. We find
that the threshold is inversely proportional to the packet-generation rate - a key parameter characterizing the traffic dynamics, as well as the largest eigenvalue of the betweenness matrix. Moreover, the epidemic threshold is negatively correlated with the maximum algorithmic betweenness. Our theory predicts the existence of a critical value of the assortativity coefficient at which the epidemic threshold reaches its minimum, which agrees with the results from direct numerical simulations.

II. CORRELATED SCALE-FREE NETWORK AND TRAFFIC-DRIVEN EPIDEMIC SPREADING MODELS

Construction of scale-free networks with tunable assortativity coefficient. We first generate an ensemble of scale-free networks according to the uncorrelated configuration model [36]. The steps are: (i) assign to each node $i$, in a set of $N$ initially disconnected nodes, a number $k_i$ of stubs, where $k_i$ is drawn from the probability distribution $P(k) \sim k^{-\gamma}$ and subject to the constraints $k_{\min} \leq k_i \leq \sqrt{N}$ (here we set $k_{\min} = 2$ and $N = 4900$) and $\sum_k k_i$ being even, and (ii) generate a network by randomly choosing stubs and connecting them to form edges, taking into account the preassigned degrees and avoiding multiple and self-connections. We then use the algorithm in Ref. [37] to obtain networks with desired degree mixing patterns. Specifically, in order to obtain an assortative network, at each step we randomly choose two different edges with four different ends, and then purposefully swap the two edges by linking the vertices with higher degrees and lower degrees, respectively. Repeating this procedure while forbidding multiple connections and disconnected components, we generate a network with certain degree assortativity without altering the degree distribution of the original network. Through the opposite operation that one edge links the highest and the lowest nodes and the other edge connects the two remaining nodes, we can obtain networks with disassortative mixing. When the assortativity coefficient of a network reaches some expected value, we stop the edge-swapping process.

Traffic-driven epidemic spreading model. Following the pioneering work of Meloni et al. [28], we incorporate the traffic dynamics into the standard susceptible-infected-susceptible epidemic spreading model [38], as follows. At each time step, $R$ new packets are generated with randomly chosen source and destination nodes, and each node $i$ can deliver at most $D_i$ packets toward their destinations. (For simplicity, we set $D_i$ to be infinite.) Packets are forwarded according to a given routing algorithm (we use the shortest-path routing scheme). The queue length of each node is assumed to be unlimited, to which the first-in-first-out principle is applied. Each newly generated packet is placed at the end of the queue of its source node. Once a packet reaches its destination, it is removed from the system. A node can be in two discrete states: susceptible or infected. After a transient time, the total number of delivered packets at each time step will reach a steady value. Subsequently, an initial fraction of nodes $\rho_0$ is set to be infected (we choose $\rho_0 = 0.1$ in numerical experiments). Differing from the conventional spreading dynamics where there is a certain probability that a node can be infected if one or more of its neighboring nodes are infected, here the infection spreads in the network through packet exchanges. In particular, all packets queuing at an infected node are infected, while all packets at a susceptible node are uninfected. A susceptible node has the probability $1 - \beta$ of being infected every time it receives an infected packet from any infected neighboring nodes. With probability $1 - \beta$, the virus in an infected packet will be cleaned by some anti-virus software (say, in a computer network) in the susceptible node. The infected nodes are recovered at the rate $\mu$ (we set $\mu = 1$).

III. NUMERICAL RESULTS AND MEAN-FIELD THEORY

Numerical results of the dependence of the epidemic threshold on the assortativity coefficient. A useful quantity to characterize correlated networks is the average degree of the nearest neighbors, denoted as $\langle k_{nn}(k) \rangle$. Figure 1 shows $\langle k_{nn}(k) \rangle$ as a function of degree $k$ for different values of the assortativity coefficient $c$. From Fig. 1, one can see that $\langle k_{nn}(k) \rangle$ is almost independent of $k$ when $c = 0$. For assortative (disassortative) networks, $\langle k_{nn}(k) \rangle$ is an increasing (decreasing) function of $k$, indicating that large-degree nodes tend to connect with large (small)-degree nodes.

A fundamental quantity in SIS dynamics is the epidemic threshold $\beta^*$, below which the epidemic dies off asymptotically [39, 40]. Figure 2 shows $\beta^*$ versus the assortativity coefficient $c$ for different values of the scale-free degree exponent $\gamma$. We see a non-monotonic behavior. Especially, for each value of $\gamma$ simulated, there exists a critical value of $c$, de-
noted as $c_{\text{wor}}$, at which the epidemic threshold is minimized. The inset of Fig. 2 shows that $c_{\text{wor}}$ decreases from 0.3 to 0.15 as $\gamma$ increases from 2.1 to 4.

**Mean-field analysis of epidemic threshold for correlated networks.** We seek to understand the counterintuitive behavior exemplified in Fig. 2. The base of our analysis is the standard degree-based mean-field theory [41] for complex networks. First, we write the dynamical rate equations for the traffic-driven SIS model as

\[
\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \beta k R [1 - \rho_k(t)] \sum_{k'} b_{k' \rightarrow k} P(k'|k) \rho_{k'}(t),
\]

where $\rho_k$ is the density of the infected nodes of degree $k$, $P(k'|k)$ is the conditional probability that a node of degree $k$ is connected to a node of degree $k'$, and $b_{k' \rightarrow k}$ is the algorithmic betweenness [42] of a directed edge from a node of degree $k'$ to a node of degree $k$. The first term in Eq. (2) is the recovery rate of the infected nodes, and the second term takes into account the probability that a node of degree $k$ belongs to the susceptible class of fraction $1 - \rho_k(t)$ and gets infected via packets from some infected nodes. The algorithmic betweenness of a directed edge, $b_{i \rightarrow j}$, is the number of packets passing from node $i$ to its neighboring node $j$ each time if the packet-generation rate is $R = 1$. For an undirected network, we have $b_{i \rightarrow j} = b_{j \rightarrow i}$. For the shortest-path routing protocol, the algorithmic betweenness of a directed edge $b_{i \rightarrow j}$ can be calculated through

\[
b_{i \rightarrow j} = \frac{1}{N(N-1)} \sum_{m \neq n} \frac{\sigma_{m \rightarrow n}(i \rightarrow j)}{\sigma_{m \rightarrow n}},
\]

where $\sigma_{m \rightarrow n}$ is the total number of the shortest paths going from node $m$ to $n$, and $\sigma_{m \rightarrow n}(i \rightarrow j)$ is the number of shortest paths going from $m$ to $n$ and passing from node $i$ to its neighboring node $j$. The algorithmic betweenness of a node $i$ is given by $b_i = \sum_{j \in \Omega_i} b_{j \rightarrow i}$, where the summation runs over the nearest neighbor set $\Omega_i$ of node $i$.

For small values of $\rho_k$, we can linearize Eq. (2) to obtain

\[
\frac{d\rho_k(t)}{dt} \approx \sum_{k'} L_{kk'} \rho_{k'}(t),
\]

where the Jacobian matrix element is

\[
L_{kk'} = -\delta_{kk'} + \beta k R P(k'|k) b_{k' \rightarrow k},
\]

and $\delta_{kk'}$ is the Kronecker delta function. The solution $\rho_k = 0$ will be unstable if there exists at least one positive eigenvalue of the Jacobian matrix $L$. The endemic phase ($\rho_k > 0$) will thus take place for $-1 + \beta R \Lambda_m > 0$, where $\Lambda_m$ is the largest eigenvalue of the betweenness matrix $C$ with its elements given by

\[
C_{kk'} = k P(k'|k) b_{k' \rightarrow k}.
\]

The element $C_{kk'}$ gives the number of packets that a node of degree $k$ receives from all neighboring nodes of degree $k'$ at
each time step when the packet-generation rate $R = 1$. Since $C$ is non-negative and irreducible, the Perron-Frobenius theorem [43] stipulates that its largest eigenvalue is real and positive. The endemic state then occurs for

$$\beta > \beta_c = \frac{1}{RA_m}. \quad (7)$$

For uncorrelated networks, we can use two approximations: (i) $P(k’|k) \propto k’ P(k’)$, and (ii) $b_{k’ \rightarrow k} \approx b_k b_{k’}/kk’$, where the latter allows us to identify the key parameters of the underlying dynamics. Under these approximations, the element of the betweenness matrix can be reduced to

$$C_{kk’} \approx b_k b_{k’} P(k’)/b,$$

which has a unique non-zero eigenvector with the eigenvalue $\Lambda_{max} = \langle b^2 \rangle/b$. The epidemic threshold for uncorrelated networks can then be obtained as

$$\beta_{uc} = \frac{\langle b \rangle}{R \langle b^2 \rangle}. \quad (8)$$

where $b$ represents the algorithmic betweenness of a node. This result was obtained previously by a slightly different version of the mean-field theory (heterogeneous mean-field theory) [28].

To validate our theoretical results [Eq. (7)], we compare them with numerically calculated dependence of $\beta_c$ on the assortativity coefficients $c$. As shown in Fig. 2, the theoretical and numerical results are consistent.

How degree mixing affects the algorithmic betweenness and the infection probability for nodes of different degrees? The nodal algorithmic betweenness within the degree class $k$ is defined as $b_k = \sum b_i/N_k$, where $N_k$ is the number of nodes within each degree class $k$ and the summation runs over all nodes of degree $k$. Figure 3(a) shows the dependence of $b_k$ on degree $k$. We see that for dissortative (e.g., $c = -0.2$), uncorrelated ($c = 0$), and weakly assortative (e.g., $c = 0.25$) networks, $b_k$ increases with $k$, and the scaling of $b_k$ with $k$ follows a power law. However, for strongly assortative (e.g., $c = 0.6$) networks, the scaling is not as straightforward in that the values of $b_k$ for large-degree nodes are not necessarily high but instead spread in a relatively wide range. (Similar phenomena were observed previously in some real-world networks [44].) The density of the infected nodes within the degree class $k$ is $\rho_k = I_k/N_k$, where $N_k$ and $I_k$ are the numbers of nodes and of infected nodes within the degree class, respectively. From Fig. 3(b), we see that for dissortative (e.g., $c = -0.2$), uncorrelated ($c = 0$), and weakly assortative (e.g., $c = 0.25$) networks, $\rho_k$ increases with $k$. However, for strongly assortative (e.g., $c = 0.6$) networks, the value of $\rho_k$ for large-degree nodes can assume relatively small values.

Combining Figs. 3(a) and 3(b), we deduce that the algorithmic betweenness of a node is positively correlated with the probability of its being infected. We define the infection probability of a node $i$ as $p_i = T_i/T$, where $T$ is the total time lapse and $T_i$ is the number of time steps at which node $i$ is infected (here we set $T = 1000$). In Fig. 4, we plot the dependence of $p_i$ on $b_i$ for different values of $c$. We see that, as a general trend, $p_i$ increases with $b_i$ and the largest $b_i$ corresponds to the highest value of $p_i$, regardless of the value of $c$.

Since the node having the maximum algorithmic betweenness $b_{max}$ is most likely to be infected, it plays a crucial role in determining the traffic-driven epidemic threshold. In particular, only when the spreading rate is large enough to successfully infect the node with $b_{max}$ will the epidemic spreading be able to sustain. To confirm this, we calculate the dependence of $b_{max}$ on the assortativity coefficient $c$ for networks with different values of the degree exponent $\gamma$, as shown in

FIG. 4: For scale-free networks of degree exponent $\gamma = 3$, dependence of the infection probability $p_i$ on the algorithmic betweenness $b_i$ for different nodes. The values of the assortativity coefficients $c$ are -0.2, 0 and 0.6 for the left, middle and right panels respectively. Other parameters are: packet-generation rate $R = 1500$ and density of infected nodes $\rho \approx 0.04$. 

FIG. 5: (Color online) For scale-free networks of different values of the degree exponent $\gamma$, the maximum nodal algorithmic betweenness $b_{\text{max}}$ versus the assortativity coefficient $c$. Each data point is result of averaging over 100 random network realizations.

FIG. 6: (Color online) For scale-free networks of degree exponent $\gamma = 3$ with different values of the assortativity coefficient $c$, epidemic threshold $\beta_c$ as a function of the packet-generation rate $R$. The slope of the fitted line is about -1. Each data point results from an average over 100 network realizations.

FIG. 7: (Color online) In the case of classical epidemic spreading, the epidemic threshold $\beta_c$ as a function of the assortativity coefficient $c$ for different values of degree exponent $\gamma$. Each data point results from an average over 100 network realizations.

Fig. 5. We see that, for each value of $\gamma$, $b_{\text{max}}$ peaks at some moderate value of $c$. Comparing with the results in Fig. 2, we see that the $b_{\text{max}}$ versus $c$ behavior is opposite to that of $\beta_c$ versus $c$, implying that increasing $b_{\text{max}}$ causes $\beta_c$ to decrease. The critical value of $c$ leading to the maximum $b_{\text{max}}$ is in fact indistinguishable from the value that minimizes $\beta_c$. As a by-product, the phenomenon shown in Fig. 5 suggests a relationship between the maximum algorithmic betweenness and the largest eigenvalue of the betweenness matrix, for which a mathematical understanding is not yet available.

Next, we study the epidemic threshold $\beta_c$ as a function of the packet-generation rate $R$ for different values of the assortativity coefficient $c$. The results are shown in Fig. 6. We see that $\beta_c$ scales inversely with $R$ for each value of $c$, as predicted by our mean-field theory [Eq. (7)]. This is intuitively correct, as an increase in the traffic flow generally facilitates epidemic outbreak.

Finally, we compare results between traffic-driven and classical epidemic spreading in the same type of correlated networks. From Fig. 7, we see that in the case of classical epidemic spreading [39], the epidemic threshold $\beta_c$ decreases as the assortativity coefficient $c$ increases. The similar phenomenon was also observed in Ref. [45]. Combining Figs. 2 and 7, one can find that traffic dynamics plays an important role in epidemic spreading.

IV. CONCLUSIONS

To summarize, our investigation of traffic-driven epidemic spreading in correlated scale-free networks reveals the existence of a critical level of network correlation for which the outbreak of epidemic is maximally promoted. We use an extended degree-based mean-field theory, taking into account the traffic flow dynamics, to account for the phenomenon. In addition, we find that nodes with larger algorithmic betweenness are more likely to be infected. We note that in traditional epidemic spreading where infections are transmitted as a reactive process from nodes to all neighbors, the epidemic threshold is largely determined by the nodes with the largest degree [46]. Our results reveal, however, that for traffic-driven epidemic spreading, the threshold mainly depends on the node with the maximum algorithmic betweenness. In particular, the threshold tends to decrease with the value of the maximum algorithmic betweenness. While for uncorrelated or disassortative networks, the largest-degree node typically has the maximum algorithmic betweenness, this is not the case for assortative networks. Our results indicate that in traffic-driven spreading dynamics, the structural properties of the network can have a significant effect on epidemic spreading.
and outbreak, providing a potential mechanism to control various spreading dynamics in real-world situations.

Acknowledgments

This work was supported by the National Science Foundation of China (Grants No. 61403083, No. 91324002 and No. 71301028), and the Natural Science Foundation of Fujian Province, China (Grant No. 2013J05007). YCL was supported by the Army Research Office (ARO) under Grant No. W911NF-14-1-0504.