

An Efficient Control Strategy for Traffic-Driven Epidemic Spreading by Deleting Edges

Fei Shao^{1,2,3}, Guo-Ping Jiang²

¹Jiangsu Information Analysis Engineering Laboratory, Nanjing 211169, China

²Center for Control and Intelligence Technology, Nanjing University of Posts & Tele., Nanjing 210003, China

³School of Information Technology, Jinling Institute of Technology, Nanjing 211169, China

Received: 17 Oct. 2012, Revised: 25 Jan. 2013, Accepted: 28 Jan. 2013

Published online: 1 Jun. 2013

Abstract: The vast majority of traditional control strategies aim at the scenario that the propagation is driven by reaction processes from nodes to all neighbours. A novel and effective control strategy for restraining epidemic spreading in the situation that epidemic pathways are defined and driven by traffic flows is proposed in this paper. It is performed by deleting edges in proper order according to their weights. The advantages of our control strategy are of facility and practical importance that such a strategy can better retain the integrity of networks. Simulation results show that the control strategy aiming at deleting edges according to the product of the degrees of two nodes of the edge is proved to be more effective. Moreover, we analyze the curve of critical epidemic threshold on deleting more and more edges and find that there is a sharp transition at the critical point after kicking out a certain fraction of edges.

Keywords: Control Strategy, Epidemic Threshold, Traffic Flow, SIS Model.

1 Introduction

Since the seminal work on the small-world phenomenon by Watts and Strogatz [1] appeared in Nature in 1998 and on scale-free networks by Barabási and Albert [2] one year later in Science, the studies of complex networks have attracted more and more interest in recent years. The important issue is to understand how the network effects influence dynamical processes taking place upon it. Many models have been proposed to characterize the epidemics spread among human, animal, and plant [3,4,5,6]. In those most extensively studied models, an individual is represented by a node which can be classified in three states: susceptible (which will not infect others but may be infected), infected (which is infective) and recovered (which has recovered from the disease and has immunity). On the basis of those former studies, how to control the epidemic spreading is one of the hot topics of recent research taking place in complex networks [7,8,9,10,11,12].

A simple model often used to study the generic features of epidemic spreading is the SIS model, which is often used for these in which the infected nodes will return to the susceptible state again and does not take into

account the possibility of nodes removal due to death or acquired immunization, and thus nodes run stochastically through the cycle susceptible \rightarrow infected \rightarrow susceptible. It is generally used to study epidemics leading to endemic states with a stationary average density of infected nodes. Assume that a susceptible node will be infected by a certain infected one during one time step with probability ν , and the recovering rate of infected ones is δ . Then the effective spreading rate λ is defined as $\lambda = \nu/\delta$ for the disease. Without lack of generality, we set can set $\delta = 1$, since it only affects at the definition of the time scale of the disease propagation.

2 Models

In the previous epidemic spreading model which assumes that each node's infectivity is strictly equal to its degree, each infected node will contact every neighbor once within onetime step and infect them with a certain probability. However, lots of networks are found that the nodes only interact with some of their neighbors at intervals. For instance, Internet connection at a given time depends on the specific traffic and routing protocols. In

* Corresponding author e-mail: shaofei@jit.edu.cn

this paper, we also investigate the propagation process driven by traffic flows along the shortest path. The traffic-driven epidemic spreading SIS model can be described as follows: all the nodes can create packets with addresses of destination, receive packets from other nodes, and route packets to their destinations; at each time step, an information packet is generated at every node with probability β , with randomly chosen sources and destinations and all the packets are forwarded one step toward their destinations through the shortest path; each node has unbounded packet delivery capability for simplicity which means congestion cannot arise in the model; a packet, upon reaching its destination, is removed from the system; a susceptible node has an effective spreading rate λ of becoming infected every time it receives a packet from an infected neighbor while the recovering rate is fixed to 1 for simplicity.

In the familiar immunization strategies such as random immunization [4] and targeted immunization [9], they focus on how to immunize nodes to realize the control of epidemic spreading on complex networks which often isolate the related nodes from networks and break the integrity. In this paper, we propose a new control strategy which can maintain the integrity by deleting some edges.

The edges have close relationship with the nodes they linked. The degree k_i of node i which has been widely used to describe the importance of a node is usually defined to be the total number of its links. However, it is just a local measure of a node. Another important measure that takes advantage of the global information of a network is betweenness. The most widely used is that of Freeman [13,14], called shortest-path betweenness of node i , which is defined as:

$$b_{sh}(i) = \sum_{s,t} \frac{\sigma(s,i,t)}{\sigma(s,t)} \quad (1)$$

where $\sigma(s,i,t)$ is the number of shortest paths between nodes s and t that pass through node i and $\sigma(s,t)$ is the total number of shortest paths between node s and t , and the sum is over all pairs s,t of all distinct nodes. In the same way, to find which edges in a network are most between other pairs of nodes, we generalize Freeman's betweenness centrality to edges and define the edge betweenness of an edge linking nodes i and j , e_{ij} , as the number of shortest paths between pairs of nodes that run along it.

Epidemic, however, more likely spreads around randomly, infecting who it encounters. To address these problems, Newman suggested a more sophisticated betweenness measure, usually called random-walk flow betweenness. The random-walk betweenness of a node i is equal to the number of times that a random walk starting at s and ending at t passes through i along the way, averaged over all s and t . In this paper, we use the

traditional random-walk betweenness introduced by Newman as follows [15]:

$$b_{rw}(i) = \frac{\sum_{s<t} I_i(st)}{n(n-1)/2} \quad (2)$$

We give our control strategies as follow: at first, we defined the weight of the edge linking two nodes i and j , w_{ij} , in different way according to different control strategies as shown in Table 1 which will not be changed; then we sort the edges by their weights from large to small to delete a fraction f_e of the edges ranked first. However, we should maintain the integrity of the network which means that, if deleting an edge will cause some nodes to be disconnected, we will not delete it, and go to deal with the edge ranked next.

Table 1 The Weight of Edge in Different Control Strategies

Weight of Edge	RBP	DP	SBP	EB
w_{ij}	$b_{rw}(i) * b_{rw}(j)$	$k_i * k_j$	$b_{sh}(i) * b_{sh}(j)$	e_{ij}

In a BA network with n nodes and BA parameter m , there are totally $n * m$ edges. And there should be at least $n - 1$ edges to maintain the integrity of the network. So we can kick out about $(n * m - n + 1) \approx n * (m - 1)$ at most, that is $(n * (m - 1)) / (n * m) = (m - 1) / m$ of the total amount. In BA networks, the average node degree $\langle k \rangle = 2 * m$, so the fraction of edges we can kick out is $(m - 1) / m = (\langle k \rangle - 2) / \langle k \rangle$. After kicking out some edges, there are fewer connections between the hub nodes which are always super spreader, a larger effective spreading rate is expected to make the epidemic become endemic. But after deleting $(\langle k \rangle - 2) / \langle k \rangle$ of the total edges, the network become a treelike topology and a smaller value of spreading rate can cause epidemic. So there must be a critical point where a phase transition occurs, if more edges are deleted, the effect of the control strategy is less obvious.

3 Simulation Results and Analysis

It's of first importance to obtain the critical epidemic threshold λ_c of networks where a phase transition occurs from absorbing phase (healthy state) to active phase (infected state). This critical threshold λ_c can best reflect the maximum capability of a certain network handling epidemic spreading. If the value of λ is above the critical threshold, $\lambda \geq \lambda_c$, the epidemic spreads and becomes endemic. On the contrary, the epidemic will die out gradually. We are interested in determining critical threshold λ_c in order to address which kind of control strategy is more susceptible to phase transition and therefore epidemic spreading. By using mean-field theory, Pastor-Satorras and Vespignani figured out the epidemic threshold of the traditional SIS dynamics model where

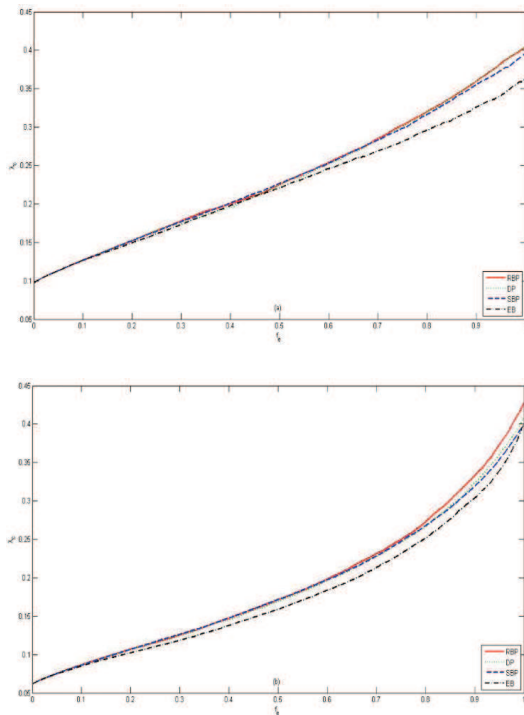


Fig. 1 λ_c VS f_e for different control strategies in the traditional SIS model in the BA networks with $n = 1000$ nodes. a) $m = 2$, b) $m = 4$.

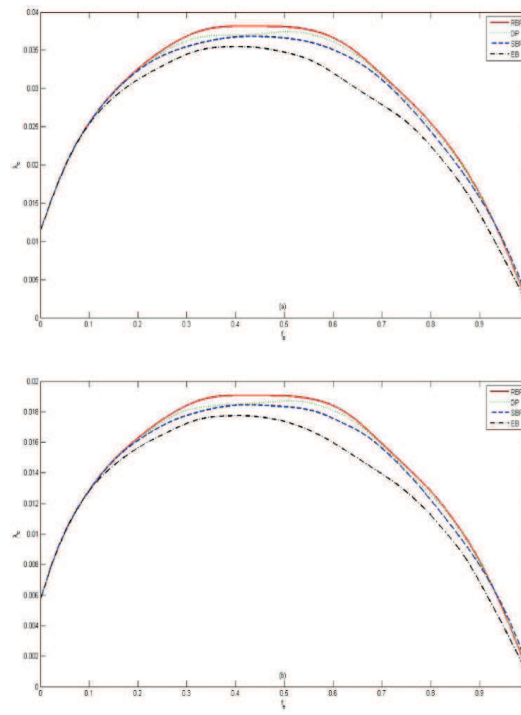


Fig. 2 λ_c VS f_e in the traffic-driven SIS model, $n = 1000$ and $m = 2$. a) $\beta = 1$, b) $\beta = 2$.

epidemic spreads to all the neighbors, the critical epidemic threshold λ_c is given as follow [16]:

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle} \tag{3}$$

where k denotes the degree, with $\langle \dots \rangle$ indicating average over all nodes.

While in the traffic-driven SIS model where the epidemic pathways are defined and driven by traffic flows, the critical epidemic threshold λ_c is [17]:

$$\lambda_c = \frac{\langle b \rangle}{\langle b^2 \rangle * \beta * n} \tag{4}$$

where β is the packet generation rate, n is the total nodes number, and b denotes the shortest-path betweenness of a certain node, with $\langle \dots \rangle$ indicating average over all nodes.

To find out which control strategy is more effective, numerical simulations will be performed on the BA network to get the critical epidemic threshold λ_c . BA networks with $n = 1000$ or 1500 nodes are generated and the parameter m is different for comparison. And f_e is the fraction of deleted edges of the $n * (m - 1)$ edges which can be deleted while maintain the integrity of the network. It means that when f_e is up to 1, there are still about $(n - 1)$ edges in the network.

Fig. 1 shows the critical threshold of different control strategies in the traditional SIS model (In all simulations, we generate 50 instances, and the result is the average of the 50 instances). One can see clearly from Fig. 1 that while edges are deleted continuously, the value of spreading threshold is increasing gradually which means deleting edges has great effect on restraining the epidemic spreading. And the different definition of the weight of edges has little effect on the control strategies.

While it changes to the other situation where the epidemic pathways are defined by traffic-driven, different curves of threshold are derived, as shown in Fig. 2. In both Fig. 2 (a) and Fig. 2 (b), no matter how much the packet generation rate is, the curves of epidemic threshold present an inverse U shape which always goes up when $f_e \leq 0.5$ and declines when more edges are deleted in all control strategies. Especially, when $0 < f_e \leq 0.1$ of edges are deleted, the value of epidemic threshold increases rapidly. It shows that all control strategies are much more efficient while kicking out a small fraction of edges with large weight, whereas the efficiency will be lower when more than half of edges which can be deleted are kicked out. And when $n * (m - 1)$ edges are deleted, which means $f_e = 1$ and the number of edges is almost equal to the number of nodes, the network will be modified to a treelike topology that only one shortest path between two nodes. The epidemic can become endemic with a very smaller spreading rate.

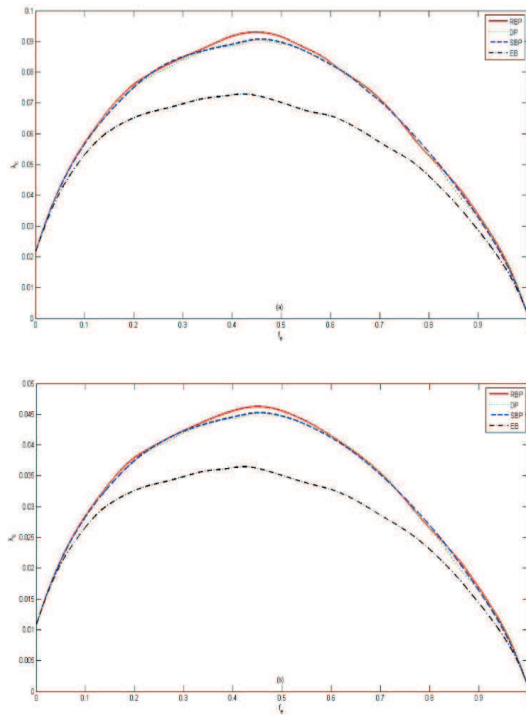


Fig. 3 λ_c VS f_e in the traffic-driven SIS model, $n = 1000$ and $m = 4$. a) $\beta = 1$, b) $\beta = 2$.

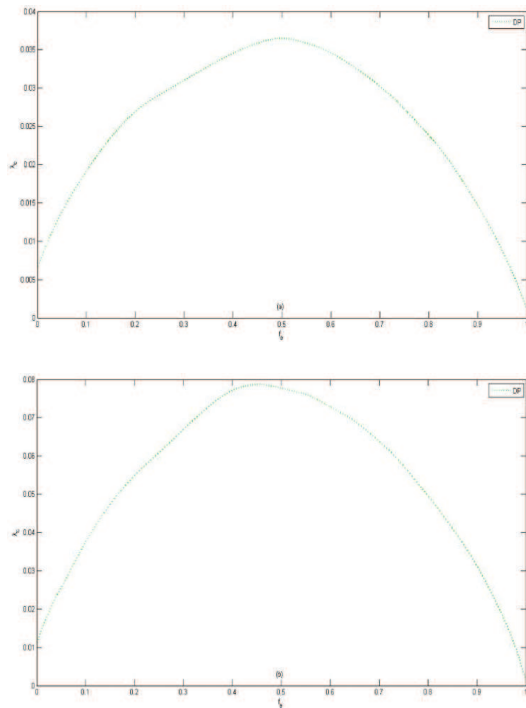


Fig. 4 λ_c VS f_e in the traffic-driven SIS model, $n = 1500$ and $\beta = 1$. a) $m = 2$, b) $m = 4$.

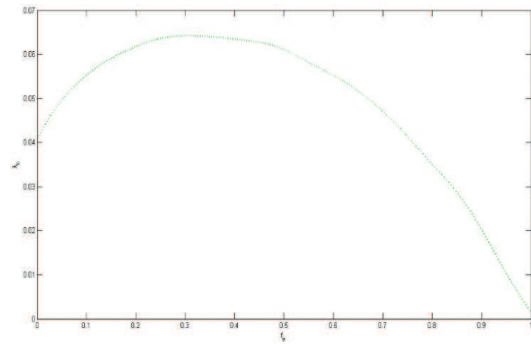


Fig. 5 λ_c VS f_e in the traffic-driven SIS model in the email network, $\beta = 1$.

If a more thorough study is carried out on the epidemic threshold of four strategies as shown in both Fig. 2 (a) and Fig. 2 (b), we can discover that there is only some slight difference between the four strategies. As a result, we can choose the control strategy DP, where the weight of edge is defined as the product of the degree of the node at the end of the edge, as the best one for constraining the epidemic spreading in complex networks. For one thing, to calculate the degree of a certain node is using only local information while the others need to know the topology of the whole network; for another, it takes the least time to calculate while the time complexity of calculating the shortest-path betweenness of node is up to $O(n^3)$.

Comparing Fig. 2 (a) and Fig. 2 (b), we also notice that while the packet generation rate is doubled, the epidemic threshold is reduced to half which agree well with formula (4). In traffic-driven epidemic spreading SIS model, epidemics only spread through links which transfer the packets, thus the number of nodes which can spread epidemic is based on the paths at every time step. At lower packet generation rate, there are only a few packets traveling throughout the network and a majority of nodes do not interact with others, so the epidemic simply dies out. When the packet generation rate increases, there are more links between nodes which spread the epidemic to a larger fraction of nodes in the network. So the epidemic spreads and becomes endemic even it has a smaller spreading rate itself.

To further demonstrate this proposition, we change m from 2 to 4. The results are shown in Fig. 3. All the curves show a common property which is the same as shown in the previous figures: the epidemic threshold increase along with the fraction of deleting edges in traditional SIS model while it has a sharp transition after deleting about half of edges which can be deleted in traffic-driven SIS model. And DP strategy is also proved more effective.

Then we check the validity of DP control strategy with different parameters. The results are shown in Fig. 4 which also presents the same feature.

Finally we employ a real-world network to test our test DP control strategy which is the E-mail network with 1133 nodes [18]. As shown in Fig. 5, the epidemic threshold is enhanced by 38.5% while kicking out only 10% of edges. It means our DP control strategy also makes effect in some real-world network.

4 Conclusions

The research purpose of control strategy is to restrain epidemic spreading to the greatest possible extent. Four control strategies are proposed in this paper. By deleting some edges with larger weight, simulations have indicated that the critical epidemic threshold is enhanced greatly which means the proposed strategies can effectively control the epidemic. The strategy which deletes the edges according to the product of the degree of the nodes at the end of the edge is proved to be the most feasible. And there is found a transition at the critical point after kicking out about half of edges. Though our control strategy may be less cost effective than the traditional strategies, it can maintain the integrity of the whole network which makes it is worthwhile in some particular practical situations. Thus, our control strategy may be of great significance for controlling the epidemic spreading in real systems.

Acknowledgements

This work is supported in part by the National Natural Science Foundation of China (Grant No. 60874091), the Six Projects Sponsoring Talent Summits of Jiangsu Province (Grant No. SJ209006), the Specialized Research Fund for the Doctoral Program of Higher Education of China (SRFDP) (Grant No. 20103223110003), the Natural Science Foundation of Jiangsu province (Grant No. BK2010526, BK2012082), the Scientific Innovation Program for University Research Students in Jiangsu Province (Grant No. CXZZ11.0403 and No. CXLX11.0417) and sponsored by Qing Lan Project.

References

- [1] D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world' networks, *Nature*, 393(4) (1998) 440-442.
- [2] R. Albert, A.L. Barabási, Emergence of Scaling in Random Networks, *Science*, 286 (1999) 509-512.
- [3] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and Its Applications*. New York: Hafner Press (1975).
- [4] R. M. Anderson, R. M. May, *Infectious Diseases of Humans*. Oxford: Oxford University Press (1992).
- [5] H. W. Hethcote, The mathematics of infectious diseases. *SIAM Review*, 42 (2000) 599-653.

- [6] O. Diekmann, J. A. P. Heesterbeek, *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. JohnWiley & Sons, New York (2000).
- [7] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86 (2001) 3200-3203.
- [8] Z. Dezsö and A. L. Barabási, Halting viruses in scale-free networks. *Phys. Rev. E* 65 (2002) 055103.
- [9] R. Pastor-Satorras, A. Vespignani, Immunization of complex networks. *Phys. Rev. E*, 65 (2002) 036104.
- [10] R. Cohen, S. Havlin, D. Ben-Avraham, Efficient Immunization Strategies for Computer Networks and Populations. *Phys. Rev. Lett.*, 91 (2003) 247901.
- [11] R. Pastor-Satorras, A. Vespignani, Epidemics and immunization in scale-free networks. In Bornholdt S, Schuster H G, *Handbook of graph and networks*. Berlin: Wiley-VCH (2003).
- [12] C. Castellano, and R. Pastor-Satorras, Thresholds for epidemic spreading in networks. *Phys. Rev. Lett.*, 105 (2010) 218701.
- [13] L. C. Freeman, A set of measures of centrality based upon betweenness. *Sociometry* 40 (1977) 35-41.
- [14] L. C. Freeman, Centrality in social networks: Conceptual clarification. *Social Networks* 1 (1979) 215-239.
- [15] M. E. J. Newman, A measure of betweenness centrality based on random walks. *Social Networks*, 27 (2005) 39-54.
- [16] R. Pastor-Satorras, A. Vespignani, Epidemic dynamics and endemic states in complex networks. *Phys Rev E*, 63 (2001) 066117.
- [17] S. Meloni, A. Arenas, and Y. Moreno, Traffic-Driven Epidemic Spreading in Finite-Size Scale-Free Networks, *Proc. Natl. Acad. Sci.*, 106 (2009) 16897-16902.
- [18] R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt, A. Arenas, Self-similar community structure in a network of human interactions, *Phys. Rev. E*, 68 (2003) 065103.



Fei Shao was born in Jiangsu, China, in 1978. He received the Bachelor Degree from Jiangsu University of Science and Technology in 1998, Master degree in Computer Application Technology from Nanjing Uni-versity of Technology, Jiangsu, China in 2003. Since 2008, he has been pursuing

his Dr. Degree in the School of Computer Science & Technology at Nanjing University of Posts & Telecommunications. From 2010 to now, he is an associate professor at the School of Information Technology, Jinling Institute of Technology and working for Jiangsu Information Analysis Engineering Laboratory. His current research interests include information security and complex dynamical networks.



Guo-ping

Jiang Professor and Ph.D supervisor of Nanjing University of Posts and Telecommunications. His research interests include chaos synchronization and control, chaos-based communication, and complex dynamical networks.