Impact of Global Edge-Removal on the Average Path Length

Lock Yue Chew[1], Ning Ning Chung[2], Jie Zhou[2], and Choy Heng Lai[3,4]

[1] Nanyang Technological University, Division of Physics and Applied Physics, 21 Nanyang Link, Singapore 637371
lockyue@ntu.edu.sg
http://www.ntu.edu.sg/home/lockyue/
[2] National University of Singapore, Temasek Laboratories, Singapore 117508
[3] National University of Singapore, Beijing-Hong Kong-Singapore Joint Centre for Nonlinear and Complex Systems (Singapore), Kent Ridge 119260, Singapore
[4] National University of Singapore, Department of Physics, Singapore 117542

Abstract. In this paper, we further explore into the impact of link removal from a global point of view. While diseases spread more efficiently through the best spreaders and removal of local links attached to them can have great impact, it is also important to have a method to estimate the cost of edge-removal from a global point of view since the removal of a link may also affect certain global properties of the network. We discuss global strategies on link removal and study their effectiveness in controlling the propagation of infectious diseases based on the spreading control characteristics (SCC). The SCC framework opens up a comprehensive way for researchers to assess and compare the efficacy of their strategies against the potential cost of their implementation from a global perspective.

Keywords: complex network, epidemic control, epidemic spreading, social and economic cost.

Finding an efficient way to slow down the propagation of infectious diseases within a society has always been an important subject in network sciences. Over the past decades, with the availability of large-scale, comprehensive data sets that trace the movement and interaction of the population, there have been increasing number of investigation on the control of the spreading of transmittable diseases [1,2,3,4,5,6,7,8,9,10,11,12,13]. Since the spreading of epidemics affects the society as a whole, the study of its control should not be separated from the associated economic and social dimensions. For this reason, we have recently examined into the efficacy of local control strategies on the propagation of infectious diseases by removing local connections attached to the best spreaders with the associated economic and social costs taken into account [14]. When local links attached to the best spreader are removed, we found an increase in the spreading time while the centrality betweenness of the best spreader decreases as a result of the reduced connectivity of the network topology. Nevertheless, our studies reveal that it is possible to trade minimal reduction in connectivity

K. Glass et al. (Eds.): COMPLEX 2012, LNICST 126, pp. 52–57, 2013.

of an important hub with efficiencies in epidemic control by removing the less busy connections. In other words, we have uncovered the surprising results that removing less busy connections can be far more cost-effective in hindering the spread of the disease than removing the more popular connections.

In this article, we study the impact of edge-removal on the global property of the network. Specifically, when a link is removed, instead of investigating the local cost paid by the best spreader, we are interested to know the impact of the removal on the average path length (APL) of the network. Most real-world networks have very short average path lengths such that communication within the network can be carried out efficiently. In the context of a small world, this means that any two nodes in the network are well connected through a very short path. As edges are removed, connectivity of a network decreases. Hence, if an edge-removal strategy is employed for epidemic control, a cost has to be paid by all of the nodes in the network, i.e. a decrease in the efficiency of information or mass transport in the network. Our aim is to look for an optimal edge-removal strategy that allows us to trade a minimal reduction in the connectivity of a network with efficiencies in epidemic control.

For this, we quantify the relative effectiveness (E_{ij}) of the removal of an edge ij in increasing the epidemic threshold by the decrease in the extreme eigenvalue ($\Delta\lambda_{ij}$) of the network adjacency matrix, normalized by the maximum eigenvalue of the original network, i.e.

$$E_{ij} = \frac{\Delta\lambda_{ij}}{\lambda_m} \,. \tag{1}$$

Note that E_{ij} captures how removal of link ij increases the difficulty for the epidemic outbreak to take place. On the other hand, the global cost (C_{ij}) of an edge removal is quantified by the decrease in the network average inverse path length (AIPL) [15], ΔL_{ij}, normalized by the average inverse path length of the original network, L_0:

$$C_{ij} = \frac{\Delta L_{ij}}{L_0} \,. \tag{2}$$

Here, AIPL is used instead of APL since edge-removal may fragment the network into disconnected components, causing APL of the network to diverge, but AIPL of the network remains finite.

Then, we simulate the global impact of the removal of a single edge in three real-world complex networks from different fields, i.e.: 1) the US air transportation network [16], 2) the collaboration network in computational geometry [17] and 3) the Gnutella peer-to-peer internet network [18,19]. As shown in Fig. 1, E_{ij} and C_{ij} relate differently to edge betweenness (B_E) [20] of the removed link. The most effective link-removal does not necessary comes with the largest cost to pay. This implies that it is possible to have a optimized global gain in the control of epidemic spreading if the link to be removed is picked properly. For global edge-removal, we define the gain in spreading control for removing edge ij as:

$$G_{ij} = \frac{\Delta\lambda_{ij}}{\lambda_m}\frac{L_0}{\Delta L_{ij}} \,. \tag{3}$$

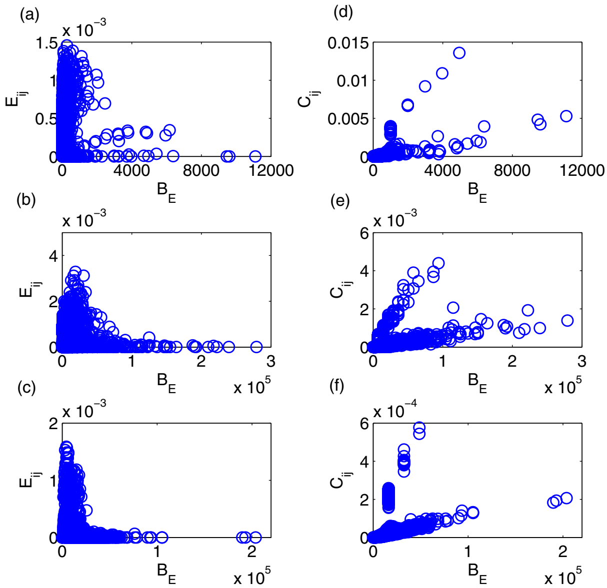

Fig. 1. Dependence of E_{ij} and C_{ij} on B_E of the removed link for (i) the US air transportation ((a) and (d)), (ii) the collaboration((b) and (e)) and (iii) the Gnutella ((c) and (f)) networks

Next, we look at four different global edge-removing strategies which involve removal of more than one link, namely, removing the edges following (1) the decreasing order of the edge betweenness, (2) the increasing order of the edge betweenness, (3) the decreasing order of the product of the k_s [21,22] values, and (4) the decreasing order of the gain. Note that the product of the k_s values of an edge ij is defined as the product of the k_s values of nodes i and j. For each strategies, edges are removed one by one according to their original ranking without further updating of the ranking during the process of removal. We measure the cost of removing q links (following a specified order according to the adopted strategy) on the network by the decrease in the network average inverse path length (AIPL), $\Delta_q L$, normalized by the average inverse path length of the original network, L_0:

$$C_q = \frac{\Delta_q L}{L_0} \,. \tag{4}$$

Meanwhile, the effectiveness in epidemic control of removing the q links are measured by

$$E_q = \frac{\Delta_q \lambda}{\lambda_m} \,, \tag{5}$$

where λ_m is the maximum eigenvalue before edge removal. By studying the global spreading control characteristics (GSCC) curves, i.e. the $E_q - C_q$ curve, one can analyze the difference between various global edge-removing strategies.

In Fig. 2, we show the GSCC curves for the three networks when edges are removed according to the four global edge removal strategies. Strategy 2 performs better than strategy 1 in both the US airline and the Gnutella networks. In these cases, removal of the less busy connections shows higher cost-efficiency than removal of the busy connections. However, in the collaboration network, strategy 2 performs better than strategy 1 only at the earlier stage of edge-removal. For strategy 4, we observe that the gradient near the end of the curves are larger than the earlier parts of the curves. This implies that the edge-removal strategy is not optimized here to yield the largest gain at each step. In particular, as shown in Fig. 2 (b) for the collaboration network, strategy 4 performs better only at the early stage of the edge-removal. After which, the curve is flat with almost no increase in E_q as C_q increases to 0.4. The gradient of the curve increases after that and drops quickly to zero again. In contrast, the performance of the third strategy is consistently better in all the three networks.

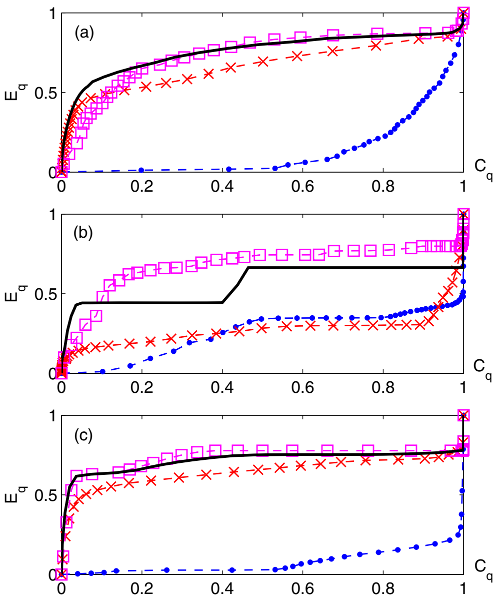

Fig. 2. Global spreading control characteristics for (a) the US air transportation, (b) the collaboration and (c) the Gnutella networks. Note that SCC are plotted as dashed curves with dots, dashed curves with crosses, dashed curves with squares and solid curves for the first, second, third and fourth strategies respectively.

Compared to a local edge-removal strategy, a global edge-removal strategy involves not only the edges attached to the best spreader but all the edges in the network. Hence, the number of removed edges can be a lot larger when a global edge-removal strategy is considered. After the removal of a large number of edges, the topological structure of the network can be very different and the ranking of the gain of the rest of the edges can change drastically. When this happens, one cannot obtain the optimal tradeoff between spreading control effectiveness and the cost to pay following the original order of the gain. Therefore, for global link-removal, removing edges following the highest gain does not consistently give optimal epidemic control efficiency as it does for local link removal. While one can update the ranking of the gain repeatedly each time after a link is removed to obtain an optimized result, the operation requires a huge amount of computational power. Specifically, it takes a time of $O(MN)$ to calculate ΔL of all the M edges in a network with N nodes. Therefore, we suggest to adopt strategy 3 if more than 50% of the effectiveness is required.

In conclusion, from both the local and global point of view, removal of the most popular connections is not the most cost-effective solution in slowing down the spreading of either viruses or rumors. In our earlier study on the local edge removal strategies, we show that it is possible to maximize the effectiveness in epidemic control with minimal reduction in the centrality of the network's hub by removing the less busy connections. Here, we further demonstrated that the decrease in the network global connectivity can be minimized with better performance in epidemic control by removing properly chosen edges.

Acknowledgements. This work is supported by the Defense Science and Technology Agency of Singapore under project agreement of POD0613356.

References

1. Hufnagel, L., Brockman, D., Geisel, T.: Forecast and control of epidemics in a globalized world. Proc. Natl. Acad. Sci. USA 101, 15124–15129 (2004)
2. Brownstein, J.S., Wolfe, C.J., Mandl, K.D.: Empirical evidence for the effect of airline travel on inter-regional influenza spread in the United States. Plos. Medicine 3, 1826–1835 (2006)
3. Holme, P., Kim, B.J., Yoon, C.N., Han, S.K.: Attack vulnerability of complex networks. Phys. Rev. E 65, 056109 (2002)
4. Pastor-Satorras, R., Vespignani, A.: Immunization of complex networks. Phys. Rev. E 65, 036104 (2002)
5. Marcelino, J., Kaiser, M.: Reducing influenza spreading over the airline network. PLOS Currents, RRN1005 (2009)
6. Kitsak, M., Gallos, L.K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H.E., Makse, H.A.: Identification of influential spreaders in complex networks. Nature Phys. 6, 888–893 (2010)
7. Zhou, J., Liu, Z.: Epidemic spreading in communities with mobile agents. Physica A 388, 1228–1236 (2009)
8. Chen, Y., Paul, G., Havlin, S., Liljeros, F., Stanley, H.E.: Finding a better immunization strategies. Phys. Rev. Lett. 101, 058701 (2008)
9. Colizza, V., Barrat, A., Barthélemy, M., Vespignani, A.: The role of the airline transportation network in the prediction and predictability of global epidemics. Proc. Natl. Acad. Sci. USA 103, 2015–2020 (2006)
10. Newman, M.E.J.: Spread of epidemic disease on networks. Phys. Rev. E 66, 016128 (2002)

11. Pastor-Satorras, R., Vespignani, A.: Epidemic spreading in scale-free networks. Phys. Rev. Lett. 86, 3200–3203 (2001)
12. Lessler, J., Kaufman, J.H., Ford, D.A., Douglas, J.V.: The cost of simplifying air travel when modeling disease spread. Plos One 4, e4403 (2009)
13. Crépey, P., Barthélemy, M.: Detecting robust patterns in the spread of epidemics: A case study of influenza in the United States and France. Am. J. Epidemiol. 166, 1244–1251 (2007)
14. Chung, N.N., Chew, L.Y., Zhou, J., Lai, C.H.: Impact of edge removal on the centrality of the best spreaders. Europhys. Lett. 98, 58004 (2012)
15. Beygelzimer, A., Grinstein, G., Linsker, R., Rish, I.: Improving network robustness by edge modification. Physica A 357, 593–612 (2005)
16. Colizza, V., Pastor-Satorras, R., Vespignani, A.: Reaction-diffusion processes and metapopulation models in heterogeneous networks. Nature Physics 3, 276–282 (2007)
17. Jones, B.: Computational Geometry Database (February 2002), `http://compgeom.cs.edu/~jeffe/compgeom/biblios.html`
18. Leskovec, J., Kleinberg, J., Faloutsos, C.: Graph evolution: densification and shrinking diameters. ACM Transactions on Knowledge Discovery from Data (ACM TKDD) 1(1) (2007)
19. Ripeanu, M., Foster, I., Iamnitchi, A.: Mapping the Gnutella Network: Properties of Large-Scale Peer-to-Peer Systems and Implications for System Design. IEEE Internet Computing Journal (2002)
20. Newman, M.E.J.: Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. Phys. Rev. E 64, 016132 (2001)
21. Seidman, S.B.: Network structure and minimum degree. Social Networks 5, 269–287 (1983)
22. Carmi, S., Havlin, S., Kirkpatrick, S., Shavitt, Y., Shir, E.A.: A model of internet topology using k-shell decomposition. Proc. Natl. Acad. Sci. USA 104, 11150–11154 (2007)