International Journal of Advancements in Computer Networks and Its Security– IJCNS
Volume 4 : Issue 4 [ISSN 2250-3757]
Publication Date : 27 December,2014

Infection dectection in complex networks with community structures

Yi Yu, Gaoxi Xiao

***Abstract*—Infection detection is of significant importance as it allows early reaction and proper measures for infection control. Existing studies typically propose algorithms for finding the best locations for a given number of monitors in order to achieve most effective early detection. In this work, we examine the influences of community structures on infection detection. Specifically, a few different cases are tested where monitors are deployed in two community networks, namely community random network and community scale-free network, respectively. By comparing the average/maximum infection sizes in different networks with different community strengths, we show that the existence of community structures, in most cases, helps significantly reduce the infection size. We also test the case where each monitor has a certain probability failing to detect the infection. Simulation results show that in the community networks, similar to that in random networks without community structures, even a low probability of monitor failure may significantly increase the infection size.**

***Keywords*—complex network, infection detection, community structure, system reliability**

I. Introduction

Complex networks have been studied for more than a decade. A lot of the real-life systems, including the Internet autonomous systems [1], computer networks [2], World Wide Web [3], human sexual contacts [4] etc., can be modelled as complex networks. Various network models [5, 6, 7] have been proposed to catch the statistical characters of the networks and extensive results have shown that network structures can significantly influence the dynamics and properties of the systems. A typical example is the scale-free network of which the nodal degrees follow a power-law distribution [7]. It is shown that having such a structure can strongly influence the dynamics of the systems, e.g., leading to zero epidemic threshold in infinite networks [8] yet allowing effective immunization [9, 10, 11], etc.

One topic with very limited existing results, however, is the detection of infection spreading in complex networks. Typically it is assumed that a certain number of monitors are deployed in the network and the infection, once reaches a node installed with a monitor, will trigger an alarm and the consequent measures for infection control. Existing results have focused on finding the best locations for a given number of monitors in order to minimize the average/maximum infection size, respectively [12, 13].

We consider a different problem in this paper, namely the effects of community structures on the efficiency of infection detection in complex networks. Community structures have attracted extensive research interests as such structures widely exist in most real-life systems [14]. Existing studies however have been mainly focusing on detecting [15, 16, 17] and modelling [18, 19] community structures, while rather limited work has been done for understanding why community structures exist so widely and how such structures influence the properties of the complex systems.

In this paper, we study the effects of community structures on the efficiency of infection detection by conducting extensive simulations on two community network models: random network with community structure [18] and scale-free network with community structure [19]. We apply the monitor deployment algorithms in [12, 13] for minimizing the maximum/average infection size and compare the infection size of networks with different numbers of communities. We show that the existence of community structures generally speaking helps significantly reduce the infection size. Further, to examine how community structures influence the reliability of the detection scheme, we test the case where each monitor has a certain probability of failing to detect the infection. Simulation results show that even a low probability of monitor failure may significantly increase the infection size.

The rest of the paper is organized as follows. The community network models, epidemic models and infection detection models will be briefly described in Section 2. Numerical simulations for testing maximum/average infection sizes in different community networks will be reported in Section 3. In Section 4, we test on the reliability of the detection schemes in community networks. Finally, Section 5 concludes the paper.

II. Models and algorithms

A. *Random networks with community structures*

We adopt the algorithm in [18] to generate random networks with community structures. The basic idea of the algorithm is to initially assign the nodes into different communities and connect the nodes within the same community with a much higher probability than that between nodes belonging to different communities. Assume that each pair of nodes in the same community are connected at probability of p and each pair of nodes in different communities are connected at probability of q, and denote $\alpha = p/q$. For a network with N nodes and M communities

---

Yi Yu, Gaoxi Xiao
School of Electrical and Electronics Engineering
Nanyang Technological University, Singapore

with n_i nodes in the i-th community, the total number of links can be calculated as:

$$V=\sum_{i=1}^{M}\frac{1}{2}n_i(n_i-1)p+\sum_{i<j}^{M}n_in_jq$$

For a network with equal community size the expected number of intra-community links is:

$$V_{intra}=\frac{N^2p}{2M}-\frac{Np}{2}$$

The expected number of inter-community link is:

$$V_{inter}=\frac{N^2(M-1)p}{2M}$$

We shall use these two equations to control the number of links in the rest part of the paper.

B. *Scale-free networks with community structures*

Scale-free networks have a power-law degree distribution where

$$p(k)\sim k^{-\gamma}.$$

In most real-life cases, γ is between 2 and 3. The traditional BA model generates a scale-free network with $\gamma=3$[7].

In this paper, we will use the algorithm proposed in [19] to generate scale-free networks with adjustable community structures. Initially, we build a network with M communities with m_0 nodes in each community. Within each community, the nodes are connected into an complete graph. For each pair of communities, a random link is built to connect them. For each node, the number of inter-community links and intra-community links are recorded separately. At each time step of the evolving process, a single node is added into the network. It randomly chooses a community to join and brings m intra-community links. It connects to a node of its own community with a probability proportional to that node's intra-community nodal degree. Meanwhile, it has a probability of σ to create n inter-community links. It connects to a node of a different community with a probability proportional to that node's inter-community nodal degree. The process repeats until the network grows to the expected size. It can be calculated that the network created has a degree distribution of:

$$P(k)=\frac{2(m+\sigma n)^2t}{Mm_0+t}k^{-3}$$

C. *Infection and infection detection models*

Follow the work in [12], we adopt the SI model [20] in our simulations. Specifically, we assume that the infection starts from a single node. At each time step, each susceptible node adjacent to any infected node has a certain probability p of getting infected. An infected node is never recovered. SI model closely resembles the early stage of an outbreak of infection. In the rest of the paper, following the assumption in [12], we consider $p=1$ which indicates an extreme (benchmark) case of strong infection.

For detecting the infection, a given number of monitors are deployed in the network. We assume that the monitors are able to detect the infection and proper measures can be taken to stop the infection immediately. Among all the different cases with different infection sources, the worst one leading to the biggest infection size when the infection is detected generates the so-called maximum infection size; while the average infection size of all the different cases is termed the average infection size..

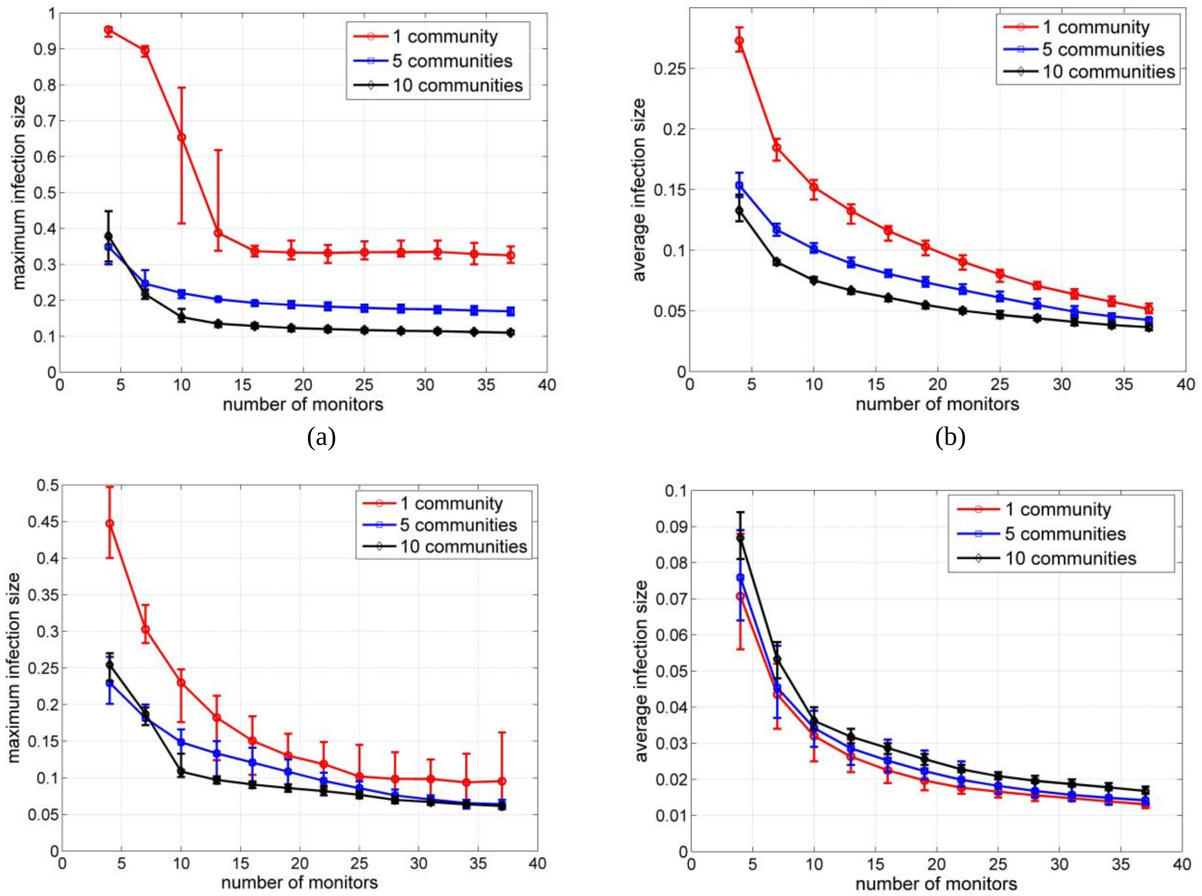
D. *Monitor deployment algorithms*

The algorithms in [12] and [13] respectively try to find the best locations for a given number of monitors in order to achieve the minimum maximum/average infection size, and both of them demonstrate good performances. Specifically, [12] developed a heuristic algorithm called MMI to minimize the maximum infection size. The basic idea of MMI algorithm is to start with a random deployment of monitors and then iteratively improve the detection performance until a local minimum is reached. It is shown that by repeating the algorithm with a large enough times, each time with a different initial deployment, satisfactory performance can be achieved. In [13], a greedy algorithm approach is developed for minimizing the average infection size. This algorithm was conveniently termed as MAI algorithm in [12]. In this paper, both MMI and MAI algorithms will be adopted for performance evaluation.

III. Detection of strong infection in community networks

In this section, we will apply the two existing monitor placement algorithms in the community networks to see how the maximum/average infection sizes are influenced by the community structures.

We carry out simulations on both random and scale-free networks with community structures. For the random network models, we generate them with 1 community (i.e., without community structure), 5 communities, and 10 communities, respectively. Each of these networks has 500 nodes and 2500 edges. Let $\alpha=p/q=\infty,100,200$ in the three networks respectively, which makes the inter-community links count for about 5% of all the network links in the latter two community networks. For the scale-free network model, we also generate them with 1 community, 5 communities, 10 communities respectively. Each of them has 1000 nodes, 3000 edges.Let $\sigma=0.02$ in the latter two community networks. For each random network, we let each of the 500 nodes serve as the infection source for 10 times and average the results of these 5000 realizations. For each scale-free network, we let each of the 1000 nodes serve as the infection source for 10 times and average the results of these 10000 realizations. In both network models, the SI model is adopted and the



(a) (b)

(c) (d)

Figure 1. (a) Maximum infection size versus the number of monitors deployed in random networks with community structures; (b) Average infection size versus the number of monitors deployed in random networks with community structures; (c) Maximum infection size versus the number of monitors deployed in scale-free networks with community structures; and (d) Average infection size versus the number of monitors deployed in scale-free networks with community structures.

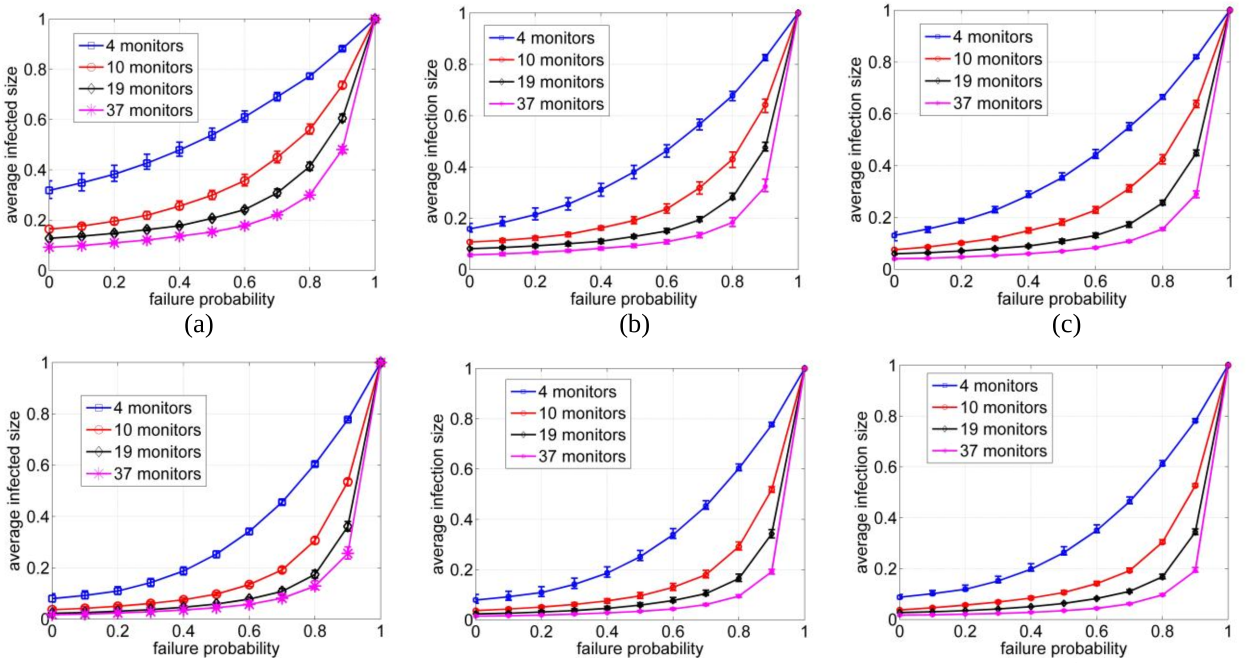
spreading probability is $p = 1$. For deploying monitors minimizing the maximum infection size, we run the MMI algorithm proposed in [12] for 5000 and 1000 times respectively and select the best results of these realizations. For deploying monitors minimizing the average infection size, we adopt the MAI algorithm [13].

Figure 1 shows both the maximum and average infection size in random networks and scale-free networks with community structures. In random networks, both the maximum and average infection size are significantly reduced when there exist community structures. In scale-free networks with community structures, however, only the maximum infection size is reduced while the average infect size is not significantly changed. In fact, having more communities leads to slightly bigger average infection size. This is not hard to understand: in scale-free networks, the inter-community links are more likely to be connected to the nodes with higher inter-community degree. Once such nodes are infected, the infection easily spread to all the communities.

IV. Reliability of infection detection schemes

In this section, we further consider the case where each monitor has a certain chance failing to detect the infection. This is a quite common situation in real-world cases and we want to examine the tolerance of system when community structures are introduced.

Following the argument in [12], we only consider the average infection size as there always exists a certain chance, though typically extremely small, that all the monitors fail to trigger an alarm when getting infected and consequently, the infection takes over the whole network. To calculate average infection size, for each network, we use the MAI algorithm to find the monitor locations; then we let each node be the infection source and run the simulation for 10 times. During the process, each monitor has a probability θ of failing to detect the infection. The average of these 10 realizations is the expected infection size for infection sourced from this node. Let all nodes serve as the infection source and the average of them is the average infection size of the network.



(a) (b) (c)

(d) (e) (f)

Figure 2 Average infection size versus the detection failure probability in (a) random network with 1 community; (b) random network with 5 communities; (c) random network with 10 communities; (d) scale-free network with 1 community; (e) scale-free network with 5 communities and (f) scale-free network with 10 communities.

Figure 2 shows the relationship between detection failure probability and the average infection size in random networks with 1, 5, 10 communities and scale-free networks with 1, 5, 10 communities, respectively. Similar to that reported in [12] for non-community random networks, we can observe that in all the cases, the infection size increases dramatically with the detection failure probability.

V. Conclusion

In this paper, we examined how community structures influence infection detection in complex networks. Extensive simulation results show that the existence of community structures in most case helps significantly reduce the maximum/average infection size in random networks. In scale-free networks, however, community structures help reduce the maximum infection size but not the average infection size. When every node has a certain chance of failing to report the infection spreading, the existence of community structures does not help significantly improve the reliability of detection schemes in both networks. The infection size increases dramatically with the detection failure probability.

Acknowledgment

This work is partially supported by Ministry of Education (MOE), Singapore, under research grant RG 69/12.

References

[1] C. Qian, C. Hyunseok, R. Govindan, and S. Jamin, "The origin of power laws in Internet topologies revisited," Proc. IEEE Infocom, vol. 2 pp. 608-617, 2002.

[2] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graphs over time: densification laws, shrinking diameters and possible explanations," Proc. ACM SIGKDD, 2005.

[3] A.-L. Barabási, R. Albert, and H. Jeong, "Scale-free characteristics of random networks: the topology of the world-wide web," Physica A, vol. 281, pp. 69-77, 2000.

[4] F. Liljeros, C. R. Edling, L. A. Amaral, H. E. Stanley, and Y. Aberg, "The web of human sexual contacts," Nature, vol. 411, pp. 907-908, 2001.

[5] Erdos P and Renyi A, "On random graphs: I," Publ. Math. 6 290–7, 1959.

[6] Moore C and Newman M E J, "Epidemics and percolation in small-world networks," Phys. Rev. E 61 5678–82, 2000.

[7] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," Science, vol. 286, pp. 509-512, October 15, 1999.

[8] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-Free networks," Physical Review Letters, vol. 86, pp. 3200-3203, 2001.

[9] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin, "Resilience of the Internet to random breakdowns," Phys. Rev. Lett., vol. 85, pp. 4626–4628, 2000.

[10] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, "Breakdown of the Internet under intentional attack," Phys. Rev. Lett. vol. 86, pp. 3682–3685, 2001.

[11] S. Xiao, G. Xiao, and T. H. Cheng, "Tolerance of intentional attacks in complex communications networks," IEEE Commun. Mag., vol. 46, no. 1, pp. 146-152, Jan. 2008.

[12] Y. Yu and G. Xiao. "On early detection of strong infections in complex networks," Journal of Physics A: Mathematical and Theoretical, 47(6):065101, 2014.

[13] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, "Cost-effective outbreak detection in networks," Proc. ACM SIGKDD, 2007.

[14] M. E. J. Newman, "Detecting community structure in networks," European Physical Journal B, vol. 38, pp. 321-330, 2004.

[15] M. E. J. Newman, "Fast algorithm for detecting community structure in networks," Physical Review E, vol. 69, p. 066133, 2004.
[16] M. J. Barber, "Modularity and community detection in bipartite networks," Physical Review E, vol. 76, p. 066102, 2007.
[17] G. Palla, I. Derenyi, I. Farkas, and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," Nature, vol. 435, pp. 814-818, 2005.
[18] Z. Liu and H. Bambi, "Epidemic spreading in community networks," EPL, vol. 72, p. 315, 2005.
[19] W. Huang and C. Li, "Epidemic spreading in scale-free networks with community structure," Journal of Statistical Mechanics, vol. 2007, p. P01014, 2007.
[20] N. T. J., Bailey The mathematical theory of infectious diseases and its applications, volume 66, London, UK: Griffin, 1975.