IMA Journal of Complex Networks (2016) Page 1 of 17
doi:10.1093/comnet/xxx000

Infection spreading, detection and control in community networks

YI YU
School of Electronic and Electrical Engineering, Nanyang Technological University, Singapore

GAOXI XIAO*
School of Electronic and Electrical Engineering, Nanyang Technological University, Singapore
*Corresponding author: egxxiao@ntu.edu.sg

[Received on 31 May 2016]

Community structures widely exist in various complex networks. Extensive studies have been carried out on defining and quantifying community structures as well as developing algorithms for detecting them in extra-large complex systems. Despite all these efforts, however, our understanding of why community structures widely exist in so many real-life systems, or in other words, the benefits/drawbacks for real-life systems to have community structures, remains to be rather limited. In this work, we discuss on the effects of community structures on infection propagation, detection and control in complex networks. Specifically, we investigate (i) the effects of community structures on transmission speed and infection size; (ii) when monitors can be deployed in the network to detect the infection spreading, the effects of community structures on early-stage infection detection; and (iii) in adaptive networks with link rewiring for isolating the infected nodes, the effects of community structures on infection control. Our results show that the existence of community structures generally speaking helps slow down the infection spreading; whether it helps reduce the overall infection size when no control method is adopted however depends on the network topology. When infection detection and controlling methods such as link rewiring are adopted, the existence of community structures steadily helps improve the efficiency of infection detection and control, though having too many communities may not necessarily bring along additional benefits.

Keywords: complex network, community structure, link rewiring

2000 Math Subject Classification: 05C80, 05C82, 90B15, 91D30

1. Introduction

Community structures widely exist in all kinds of networks, such as social and biological systems [7], information networks [37], musical compositions [8], and brain networks [34], etc. There have to be some good reasons why community structures widely exist in so many real-life systems: intuitively, we may expect that the existence of community structures may in a certain way help make the systems function more efficiently.

Majority of the existing studies have been focusing on defining, measuring and detecting community structures in complex systems. Pioneering works conducted by Newman et al. [14, 19–21] proposed to measure community structures by measuring their modularities. Follow-up works include the extensions to weighted networks [24], definition and detection of overlapping communities [9, 10, 23, 32], methods for tuning the modularity of the networks [41], as well as some debates on the limitations of

the modularity-based approaches [13, 25, 40], etc. Since there is no general statistic model capable of describing all kinds of community structures, various models have been proposed, each of which typically trying to reveal a certain aspect of the community structures. For example, a random network model with community structures and adjustable modularity was proposed in [17]; while in [12, 16], a network model was proposed to have both the nodal degrees of inter-community and intra-community links follow power-law distributions. Studies have also been carried out to reveal the community structures of real-life networks [6, 37] and to propose algorithms for efficiently detecting community structures in extra-large systems [3, 29, 39], etc.

Despite these extensive studies, our understanding of the importance of the community structures and how their existence benefits the systems remains to be rather limited. In this paper, we study the effects of community structures on infection spreading, detection, and control in complex networks. As accurate theoretical analysis on such problems remains largely an open issue, extensive numerical simulations have to be carried out. It is found that community structures help slow down the infection spreading; whether they help significantly reduce the overall infection size when no detection/controlling approaches are adopted, however, depends on detailed network topology. When detection/controlling approaches are adopted, the existence of the community structures helps make the early detection of infection and isolation of the infected nodes much easier and more efficient, though having too many communities may not necessarily lead to significant additional benefits. Since infection spreading and control is of critical importance in many complex systems, especially human social systems, such observations may to a certain extent help explain the wide existence of community structures in complex systems: when the infection spreading is welcome (say, in spreading of a new idea), the existence of the community structures may not lower the impacts of the spreading at the end of the day though it may slightly slow down the spreading process. When the agent/idea being spread is dangerous, however, the existence of the community structures gives the systems a better chance to react and put the infection under control.

The rest of the paper is organized as follows. A brief survey to the related existing results on infection spreading, detection and control in complex networks is presented in Section 2. In Section 3, we describe the network models, epidemic models and algorithms adopted in this paper. Extensive numerical simulations results and discussions are presented in Section 4. Finally, Section 5 concludes the paper.

2. Related Existing Results: A Very Brief Survey

Epidemic spreading and immunization is among the most extensively studied topics in complex networks (e.g., [4, 5, 18, 22, 26–28, 35, 36]). For epidemic spreading in community networks, the typical existing works include: in [17], the author studied on a simple random network model with adjustable modularity level. It was shown that when the modularity score increases, which denotes a larger number of intra-community links and a smaller number of inter-community links, the epidemic threshold becomes lower. It however does not discuss the effects of community structures on the infection size in an epidemic breakout. In [12], the author adopted a scale-free network model with adjustable community structures [16]. It is shown that when the Susceptible-Infected (SI) model is adopted, which means that the infected nodes are never recovered [1], the epidemic spreading is relatively slower in networks with stronger community structures. In [33], the author considered a simple two-community network model. The SIS model is adopted to study the epidemic spreading in networks with and without adaptive rewiring respectively. It is showed that the steady state infection sizes may be significantly different in two communities in a static network. In an adaptive network with rewiring operation, however, the

two communities would have similar infection sizes. It is also showed that the time for the infection to penetrate into the second community is inversely proportional to the density of inter-community links. The results are interesting yet cannot be easily generalized to other cases.

Studies on the detection of infection in complex networks typically adopt the main idea to deploy a certain number of monitors in the networks as detection agents to reduce the expected detection time and/or infection size as much as possible. The first work in the area was reported in [15], which developed an algorithm termed *cost effective lazy-forward selection* method to deploy monitors for minimizing the average damage of an infection. The algorithm was evaluated on water supply network and blog space network for multiple penalty reductions. A follow-up work in [38] developed a heuristic algorithm called *minimizing maximum infection size* (*MMI*) to minimize the worst-case infection size in early spreading of a strong infection.

Infection control in complex network has also been well studied. The most well-known studies are probably those on immunization of network nodes [4, 5, 28, 35, 36]. Such works showed that random immunization is not effective in scale-free networks while targeted immunization, which immunizes the hubs, significantly reduces the infection size [28]. In [5], it was further revealed that another effective method is to cure the hubs with a higher probability. Methods for separating the networks into a certain number of clusters in order to control the infection spreading to be within a single cluster have also been proposed (e.g., [4]), typically at a rather high cost. A more popular approach is the link rewiring, where healthy nodes may break their connections with infected nodes and rewire to other healthy ones [11, 31]. It is shown that fast-enough rewiring operation may quickly eliminate an infection spreading.

3. Models and algorithms

3.1 *Network models with community structures*

3.1.1 *Generation of homogeneous random networks with communities* We adopt the algorithm proposed in [17] to generate random networks with community structures. The main idea is to generate inter-community links with a lower probability than that for intra-community links. The algorithm can be briefly described as follows [17]:

- Randomly divide N nodes into M communities with n_i nodes in i^{th} community such we have $\sum_{i=1}^{M} n_i = N$.
- Within each community, connect every pair of nodes with a probability p.
- For each pair of nodes belonging to different communities, connect them with a lower probability q. Denote $\sigma = p/q$.

The overall number of links in the network can be estimated as:

$$V = \sum_{i=1}^{M} \frac{1}{2} n_i(n_i - 1)q + \sum_{i<j}^{M} n_i n_j q \tag{3.1}$$

In the rest of this paper, the number of edges is calculated using this equation. We control the strength of community structures and the average nodal degree of the network by adjusting p and q.

3.1.2 *Generation of scale-free networks with communities* In scale-free networks, the nodal degrees follow the power-law distribution, i.e.,

$$p(k) \sim k^{-\gamma} \tag{3.2}$$

where $p(k)$ denotes the probability that a randomly selected node has a degree k, and γ is the exponent which typically lies between 2 to 3 in many real-life systems [2]. The well-known BA model [2], which generates a scale-free network by growth and preferential attachment, has an exponent value of $\gamma = 3$. A method for generating scale-free networks with community structures, still adopting the main idea of growth and preferential attachment but with a different approach, was proposed in [12]. Specifically, when a new node joins the network, it is randomly assigned to a community and brings with itself a fixed number (denoted as m) of intra-community links. It also has a certain probability to have a fixed number (denoted as n) of inter-community links. The intra-community links are attached to nodes in the same community, while the inter-community links are connected to nodes in other communities, both by preferential attachment (i.e., the probability that a node is selected is linearly proportional to its current degree).

The construction process can be briefly described as follows [12]:

- Initialization: Build an initial network with the given number of communities, where the nodes within each community are connected into a complete graph (i.e., there is a link between every pair of nodes in the same community). For every pair of communities, randomly choose one node in each of them and connect these two nodes with an inter-community link. Keep separate records of each node's intra- and inter-community degrees respectively.
- Evolving: A newly added node will be randomly assigned to a community. Connect this new node with m existing nodes in the same community with intra-community preferential attachment, i.e., the new node connects to an existing node in the same community with a probability proportional to the intra-community degree of that existing node. Each new node, with a probability α, has n inter-community links. Specifically, with the probability α, the new node connects to n existing nodes in other communities. The chance that it is connected to a node in another community is proportional to the inter-community degree of that existing node. This above procedure is repeated until the network grows to the expected size.

The network generated in this way roughly follows a power-law degree distribution where [12]:

$$p(k) = \frac{2(m+\alpha n)^2 t}{M m_0 + t} k^{-3} \tag{3.3}$$

considering its combined intra- and inter-community degrees.

3.2 *Epidemic spreading models*

In this work, we adopt the well-known Susceptible-Infected-Recovered (SIR) model [1] in evaluating the infection spreading speed and infection size. Specifically, we assume that time is slotted and in each time slot, any susceptible node adjacent to an infected node has a probability λ of being infected. Meanwhile, each infected node has a probability μ to recover. For the SIR model, recovered nodes become permanently immunized to the infection. We also assume that the epidemic starts from a single infection source. In the rest of the paper, unless otherwise specified, we set $\lambda = 0.05$ and $\mu = 0.1$ for all the simulations adopting the SIR model.

While evaluating the maximum/average infection sizes in networks equipped with monitors for infection detection, as that in existing studies, the SI model is adopted. For the SI model, it differs from the SIR model by not allowing any node to recover. In early stage of a strong infection when recovery hardly starts yet, the SI model closely resembles the real-life cases [38]. In Section 4.2 where the SI model is adopted in simulations, we set $\lambda = 1$ following the assumptions in [38].

3.3 *Monitor deployment algorithm for infection detection*

For infection detection, we adopt the algorithms surveyed in Section 2 for minimizing the average/maximum infection size. Specifically, for minimizing the average infection size, we adopt the cost effective lazy-forward selection algorithm [15], also referred to as MAI algorithm [38]. The simple greedy algorithm adds monitors one by one, each time trying to find the best location for the newly added monitor to minimize the average infection size. For minimizing the maximum infection size, a heuristic algorithm MMI was proposed [38]. The algorithm adopts a simple iterative approach which guarantees to converge to a local optimal solution. Repeating the calculations for a large enough times will ensure reaching a suboptimal solution of the problem.

These two algorithms will be adopted in the numerical simulations in Section 4.2

3.4 *Adaptive rewiring model*

For infection control, we adopt the adaptive model [11] surveyed in Section 2. Specifically, in each time step, there is a certain probability that a susceptible node will break the connection with each of its infected neighbors and establish a new connection either to another susceptible node (S) or immunized node (R) in the same community. For each link that connects a susceptible node A and an infected node B, rewiring happens with a probability of p_{intra} if A and B are from the same community and with a probability of p_{inter} if the two nodes are in different communities. When rewiring happens, the susceptible node A shall rewire the link to connect to a susceptible or immunized node C which is in the same community as node B. Note that in Section 4.3, we will compare two different cases where C is a susceptible nodes (S) and an immunized nodes (R) respectively.

4. Simulation results and discussions

In this section, extensive simulations are carried out to evaluate the effects of community structures on infection spreading, detection and control in complex networks. Specifically, three topics will be discussed: i) we investigate the infection spreading in community networks, revealing the influences of communities on dynamics of infection size and spreading speed; ii) for the case where monitors are deployed in community networks for infection detection, we evaluate the influences of community structures on the maximum/average infection size at the moment when infection spreading is detected; and iii) for the case where link rewiring is adopted to isolate infected nodes for infection control, we investigate the effects of community structures when a few different link rewiring schemes are adopted respectively.

4.1 *Infection spreading in community networks*

In this subsection, we evaluate how community structures influence the infection size and infection spreading speed. Specifically, as defined in Section 3.1, we generate single-community, 5-community and 10-community random and scale-free networks, with specific values of σ and α, respectively. For the epidemic spreading, we adopt the SIR model. Ten networks are generated for each given number of communities, and for each network, we let each node serve as the infection source for 10 times and average the results of all these $100N$ realizations where N is the network size.

The method for generating random networks with 1, 5, and 10 communities respectively is as follows. Each network has 500 nodes and 2500 edges. For these three different cases, we choose $\sigma = p/q = \infty$ (single community), 500 (5 communities), 250 (10 communities), respectively. For the

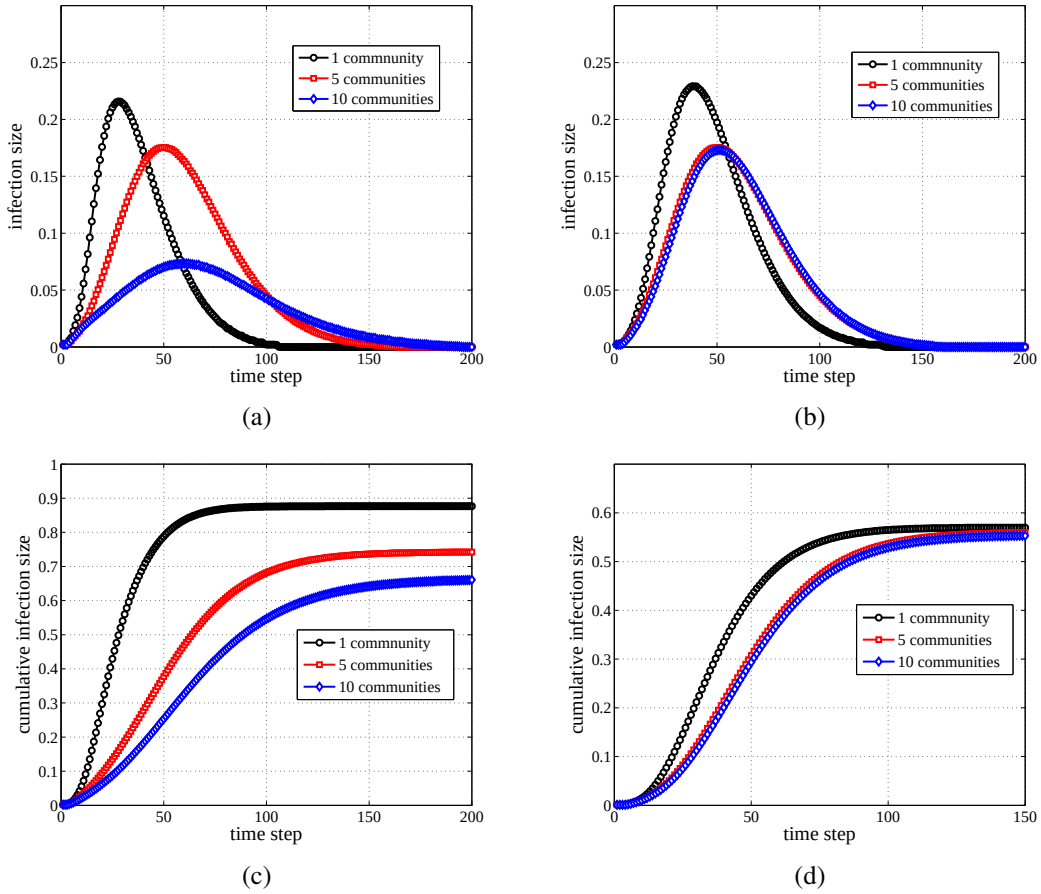

FIG. 1. Evolution of infection size along time in (a) random networks with community structures and (b) scale-free networks with community structures; cumulative infection sizes in (c) random networks with community structures and (d) scale-free networks with community structures.

networks with 5 and 10 communities, the values of σ we chose let the inter-communities links account for about 1% of all the network links. For each case we generate 10 random networks; and in each network, we let each of the 500 nodes serve as the infection source for 10 times and average the results of all these 50,000 realizations. For the scale-free networks, we also generate them with 1 community, 5 communities, 10 communities respectively. Each network has 1000 nodes and 3000 edges. For networks with 5 and 10 communities, we let $\alpha = 0.02$. The inter-community links thus make up about 2% of all the network links. Again, we generate 10 random networks for each case, and for each network we let each of the 1000 nodes serve as the infection source for 10 times and average the results of these 100,000 realizations. Unless otherwise specified, the networks generated for simulation all adopt the parameters as described above.

Figure 1 demonstrates the evolution of infection size (Figures 1(a) and 1(b)) and the overall cumu-

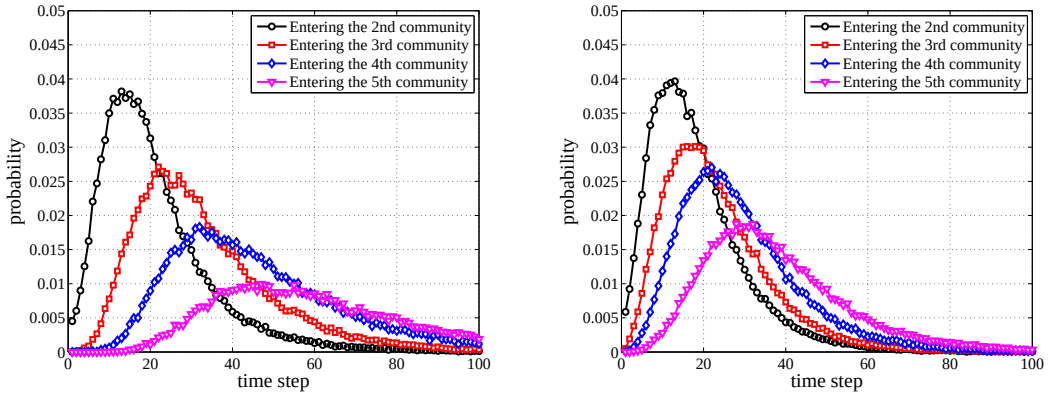

FIG. 2. Distribution of the time for the infection to enter into new communities in (a) random network with community structures; and (b) scale-free network with community structures.

lative infection size (Figures 1(c) and 1(d)) under SIR model. We find that the existence of community structures not only delays the time to reach peak infection size, but also reduces the peak infection size itself. In random networks, community structures even help reduce the overall cumulative infection size. In scale-free networks, however, reduction in cumulative infection size is insignificant. The reason lies in the existence of high-degree hubs in scale-free networks: in random networks, having more communities slows down the infection propagation. Since the infected nodes recover at a fixed rate, having a slower propagation speed helps reduce the cumulative infection size. In scale-free networks, however, the hubs with high inter-community degrees make the inter-community infection spreading much easier and faster. Thus the existence of community structures does not make significant difference to the cumulative infection size.

Also note that the 5-community and 10-community scale-free networks have nearly the same infection size over time: when having community structures helps slow down infection spreading in scale-free networks, having more communities does not necessarily lead to more benefits. The effects of the hub nodes as discussed above quickly exhaust the benefits of having more communities: penetrating into a large number of small communities is not significantly more difficult than penetrating a small number of large communities when hub nodes with high intercommunity degrees exist.

To have a closer look at how the infection propagates into different communities, we count the time when infection for the first time enters into each community, respectively. Specifically, we test on the networks with 5 communities. We still generate 10 networks and for each network let each node serve as the infection source for 10 times. For each realization, we count the time that the infection for the first time enters into the $2nd$, $3rd$, $4th$ and $5th$ communities, respectively. From all the $100N$ realizations (N is the size of the network), we can obtain the distribution of the time that the infection for the first time enters into the $2nd$, $3rd$, $4th$ and $5th$ communities respectively.

Figure 2 illustrates the simulation results. For the 5-community random networks, the average time steps it takes for the infection to enter into the $2nd$, $3rd$, $4th$ and $5th$ communities are 20.17, 31.18, 43.29, and 60.17, respectively. For the 5-community scale-free networks, the corresponding values are 18.52, 23.32, 29.08, and 38.05, respectively. We find that averagely it takes significantly longer time for

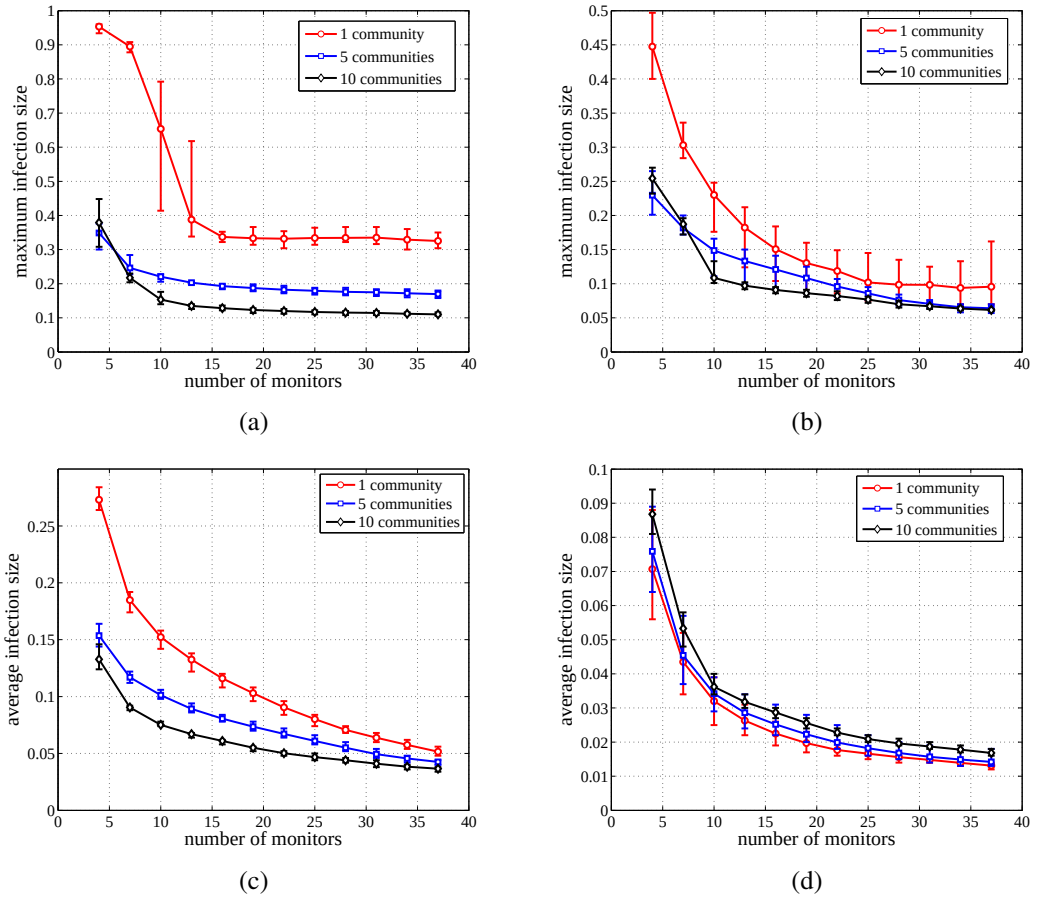

FIG. 3. Maximum infection size versus the number of monitors deployed in (a) random networks with community structures and (b) scale-free networks with community structures; average infection size versus the number of monitors deployed in (c) random networks with community structures and (d) in scale-free networks with community structures.

the infection to spread from the first community into the 2*nd* community than that for further penetrating into other communities. For example, in the 5-community scale-free networks, it takes an average of about 18 time steps for infection to spread into the 2*nd* community; and after that, it takes only about 5 to 9 time steps to penetrate into each of the following communities. This means that once the first inter-community link is breached, the infection will soon appear in multiple communities, especially in scale-free networks containing hub nodes with high inter-community degrees. Such observations may help explain why it becomes much more difficult to control disease spreading once the opportunity to act at a very early stage has been missed [30]. Some further discussions will be presented in Section 4.3.

4.2 *Monitoring the community network*

In this subsection, we evaluate infection detection in community networks. Specifically, a certain num-

ber of monitors are deployed in community networks, and we record the infection size at the moment when infection spreading is detected. As to the monitor allocation, we adopt the MMI and MAI algorithms [38] as described in Section 3.3 to minimize the maximum/average infection size, respectively. Specifically, simulations are again carried out on both the random and scale-free networks, with 1, 5 and 10 communities, respectively. For each given number of communities, we randomly generate 10 different networks. The MMI algorithm is performed 5000 times on each random network and 1000 times on each scale-free network (As pointed out in [38], detection results have a larger variation on a random network due to its uniform degree distribution. Hence more rounds of algorithm are needed.), and we choose the monitor location with the best performance. For MAI algorithm, we only need to run it once for each network. For the epidemic spreading, we adopt the SI mode as defined in Section 3.2 with $\lambda = 1$. Each node serves as the infection source in turn. The maximum/average infection size for all the different infection sources is the maximum/average infection size of the network. Averaged simulation results in 10 networks are presented in Fig. 3 where error bars indicate 95% confidence intervals.

Figures 3(a) and (c) show that the existence of community structures significantly reduces both the maximum and average infection sizes in random networks. In scale-free networks, however, only the maximum infection size is significantly reduced (Fig.3(b)), while the average infection size is actually slightly larger in networks with more communities (Fig.3(d)). To understand such an observation, we go back to Fig.3(b) where we can see that, with 4 and 7 monitors, the maximum infection size in the 10-community scale-free networks is slightly larger than that in 5-community ones. Notice that for these cases, we use fewer than 10 monitors in scale-free networks with 10 communities. For the worst case where infection starts from an un-monitored community, infection spreading can easily penetrate into other un-monitored communities (if any) through hub nodes with high inter-community degrees. Hence the worst-case infection size becomes larger in networks with a large number (larger than the number of monitors) of communities. However, when more than 10 monitors are deployed, the infection can hardly spread out of its source community; having a larger number of communities under such case hence helps reduce the maximum infection size. For the average infection size in scale-free network, the case is different: it is shown in [38] that scale-free networks are typically quite easy to be monitored and most of the monitors will be installed on hub nodes in an optimal/suboptimal solution. When monitors are installed on high-degree hubs, most of other nodes have a short distance to at least one of these monitors, making early detection of infection spreading a likely event. In community scale-free networks with a larger number of communities, the hub nodes on which monitors are deployed tend to have a smaller average nodal degree and other nodes tend to have a slightly larger distance to their nearest monitor. Thus the average infection size becomes slightly larger. For example, when 10 monitors are deployed in single community, 5-community and 10-community scale-free networks, the average nodal degrees of the 10 nodes on which monitors are deployed are 55, 46 and 25 respectively; and the average distances between each of the other nodes and its nearest monitor node are 1.58, 1.64 and 1.78 hops, respectively.

4.3 *Infection control in adaptive networks with community structures*

In this subsection, we consider the link rewiring schemes for infection control, with a main focus on evaluating the effects of community structures when different link rewiring methods are adopted. Specifically, by implementing two different rewiring strategies, namely rewiring to susceptible nodes (S) and immunized nodes (R) respectively, and adjusting the speed and starting time of the rewiring operation,

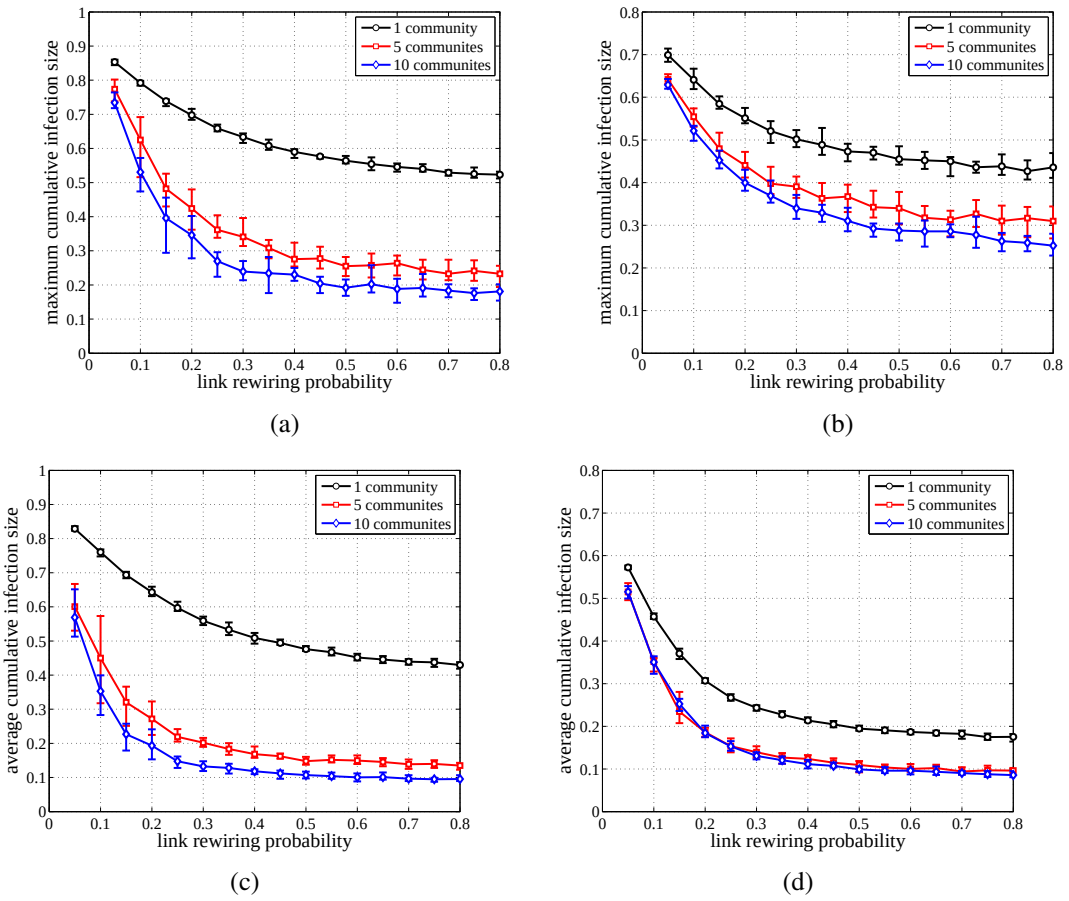

FIG. 4. The maximum expected cumulative infection size versus the probability of link rewiring in (a) random networks with community structures and (b) scale-free networks with community structures; average cumulative infection size versus the probability of link rewiring in (c) random networks with community structures and (d) scale-free networks with community structures. Susceptible nodes are rewired to immunized nodes.

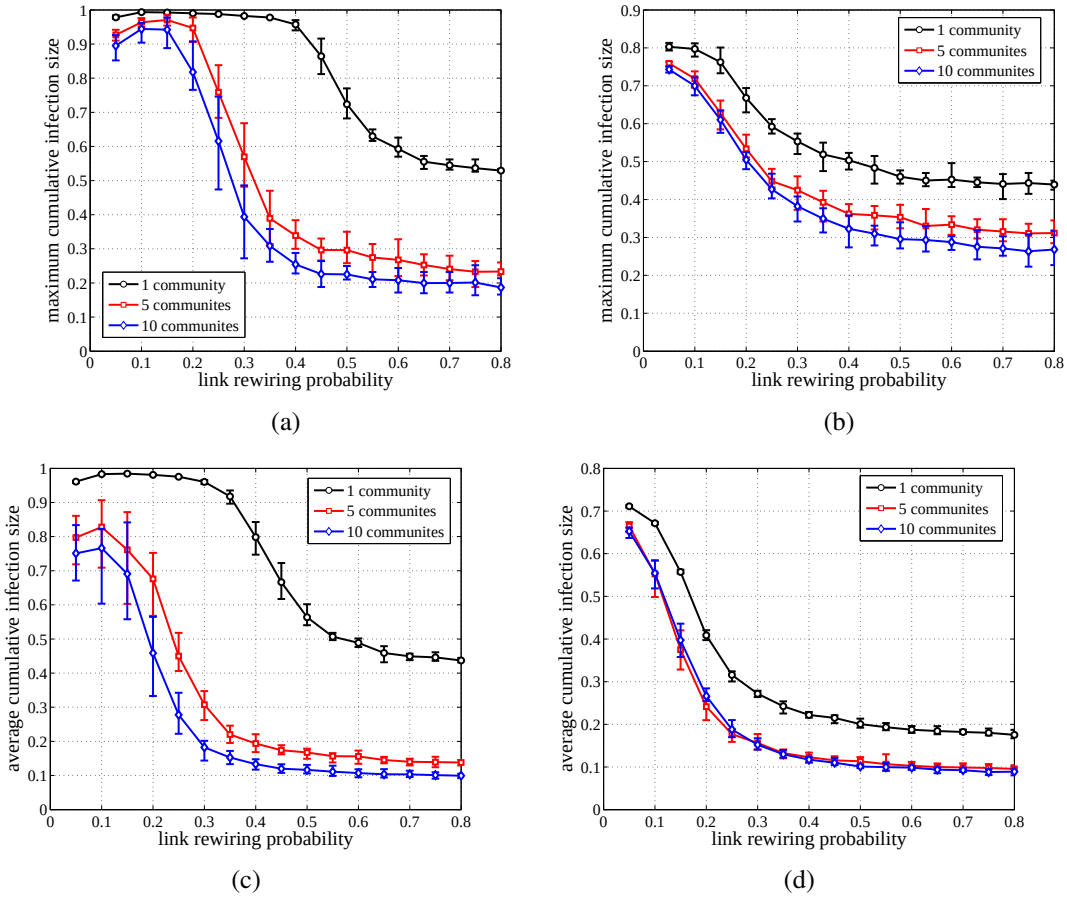

FIG. 5. The maximum expected cumulative infection size versus the probability of link rewiring in (a) random networks with community structures and (b) scale-free networks with community structures; average cumulative infection size versus the probability of link rewiring in (c) random networks with community structures and (d) scale-free networks with community structures. Susceptible nodes are rewired to susceptible nodes.

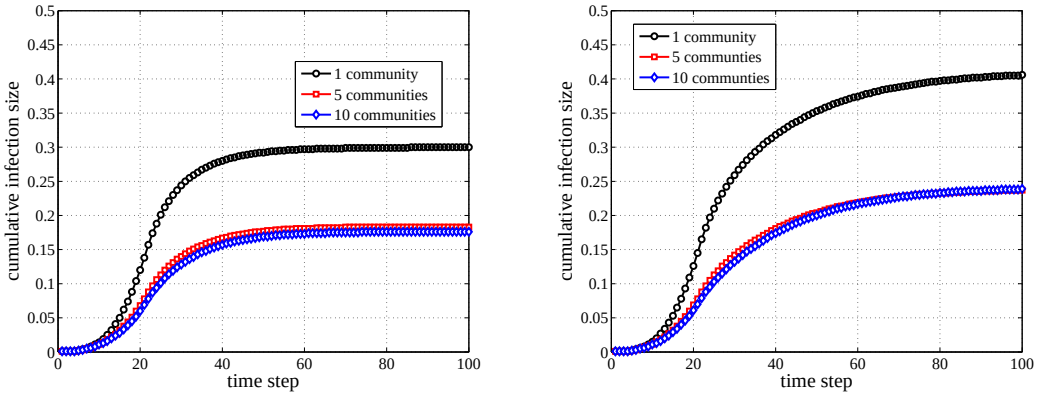

FIG. 6. Evolution of cumulative infection size in scale-free network with different community structures for (a) susceptible nodes are rewired to immunized nodes; and (b) susceptible nodes are rewired to susceptible nodes. The rewiring probability at each time step is $p = 0.2$. Reaction starts from $t = 20$.

we aim to evaluate (i) the effects of community structures on the overall infection size when different rewiring methods are adopted; and (ii) the effectiveness of rewiring inter-community links.

We adopt the adaptive link rewiring model described in Section 3.4, where in each time step, each susceptible node rewires each of its inter-community and intra-community links connected to an infected node, if any, at probabilities p_{inter} and p_{intra} respectively. Simulations are still conducted on random networks and scale-free networks with 1, 5 and 10 communities and SIR model is adopted for the epidemic spreading. For each given number of communities, we generate 10 networks, while in each network we let each node serve as the infection source for 10 times. The average cumulative infection size of the 10 realizations for each infection source is calculated. For all the infection sources, the maximum/average of their cumulative average infection sizes is the maximum/average expected cumulative infection size of the network.

Figures 4 and 5 illustrate the relationship between rewiring probability and cumulative infection size while using the two different rewiring strategies respectively. Here we set $p_{intra} = p_{inter} = p$ and let p increase from 0.05 to 0.8 with a step length of 0.05. The link rewiring starts at time step 20 to allow an initial spreading of the infection. We find that when we rewire susceptible nodes to immunized nodes, a moderate rewiring probability already depresses the infection size to a low level while further increasing the probability does not help too much (Figure 4). When we rewire susceptible nodes to susceptible nodes, however, a small rewiring probability hardly helps. Only when we increase the rewiring probability to a high enough level (about 0.3-0.4) , would the infection size be significantly reduced. After this sharp decrease, further increasing the rewiring probability does not help too much in reducing the cumulative infection size (Figure 5).

The above observations can be explained: when susceptible nodes are rewired to susceptible nodes, rewiring operation increases the nodal degree of susceptible nodes yet may not stop the infection from reaching these nodes. Once such relatively high-degree susceptible nodes are infected, the infection may soon go out of control. In such cases, rewiring only delays the infection but hardly reduce the infection size. Only when rewiring probability is high enough, would some susceptible nodes have a

reasonably good chance to be cut off from infected nodes altogether and therefore get protected from infection. When susceptible nodes are rewired to immunized nodes, on the other hand, the chance that a susceptible node gets infected is reduced, which helps lower the infection size.

A common observation in the above cases is that in networks with rewiring operations, the existence of community structures generally speaking helps lower both the maximum and average expected infection sizes; stronger community structures lead to a smaller infection sizes. The reason is not difficult to understand: in Section 4.1, it was shown that infection propagation is slowed down in networks with community structures. As a result, when link rewiring starts, community networks averagely have a smaller infection size. The slower infection propagation speed also allows the subsequent rewiring operation to be more effective. Therefore, for a given rewiring speed, the average/maximum infection size in community networks is smaller.

The effects of community structures however have their limit in reducing the average infection size in scale-free networks. As we can see in Figures 4(d) and 5(d), having 5 communities and 10 communities in scale-free networks lead to nearly the same average cumulative infection size regardless of the rewiring strategy we adopt. The reason still lies in the existence of hubs with large inter-community degree, which makes the inter-community spreading of infection much easier. Figure 6 shows the infection spreading process when rewiring starts at $t = 20$ adopting the strategies of rewiring to immunized nodes (Figure 6(a)) and susceptible nodes (Figure 6(b)) respectively. Similar to what we have observed in Figure 1(d), having a larger number of smaller communities may not help further slow down the infection spreading; and it remains to be the case when rewiring operation is adopted.

Figures 7 and 8 evaluate the effects of two factors: the starting time of the rewiring operation and rewiring of inter-community links. To make comparisons, we test on two different cases with different probabilities of rewiring each intra-community link (p_{intra}) and inter-community link (p_{inter}), respectively. Specifically, we test on two "extreme" cases: case 1 where $p_{intra} = 0.2$ and $p_{inter} = 1$, and case 2 where $p_{intra} = 0.2$ and $p_{inter} = 0$. For each case, we test on the average/maximum infection size when link rewiring starts at different time. Simulation results on 5-community networks are presented though the conclusions apply to all the other networks as well.

For both of the above rewiring strategies, we observe that it is critically important to start link rewiring in early stage of infection spreading. This can be easily understood since, as we can observe in Fig.1, infection size increases quickly with infection time. A too-late action therefore would not help avoid having a large infection size. Besides, it is observed that rewiring inter-community links also has to start very early in order to be of any significant help, even when the inter-community link rewiring probability is set to 1. The reason is obvious: as observed in Fig.2, infection penetrates into different communities rather quickly, especially in scale-free networks. Rewiring or removing inter-community links after infection has penetrated into most communities certainly would not be of any big help. Further, it is not a surprise to observe that, for either inter-community or intra-community links, link rewiring to immunized nodes is more effective than rewiring to susceptible nodes.

5. Conclusion and future work

In this paper, we studied the effects of community structures on propagation, detection and control of infection in complex networks. Specifically, we investigated the effects of community structures on transmission speed, infection size, effectiveness of early-stage infection detection and effectiveness of infection control by link rewiring, respectively. Our results on random and scale-free networks basically

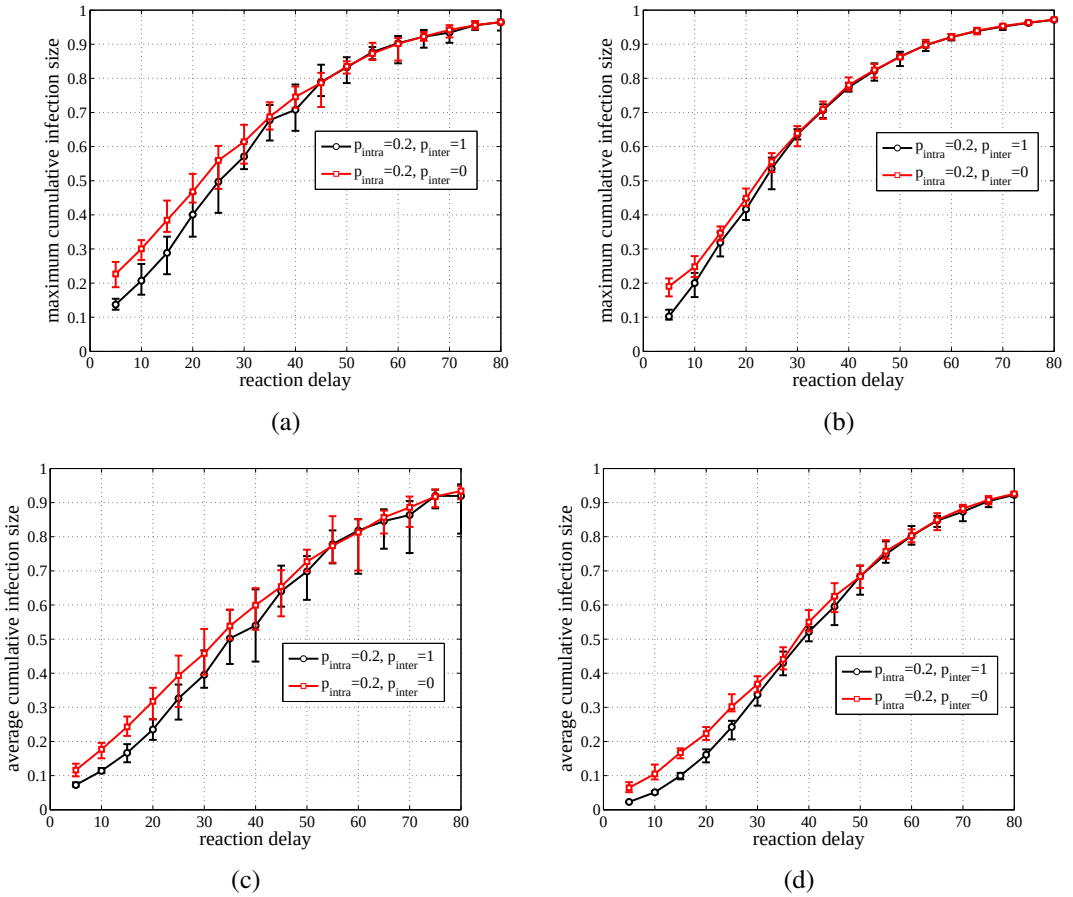

FIG. 7. The maximum expected infection size versus the starting time of rewiring operation in (a) random networks with community structures and (b) scale-free networks with community structures; average infection size versus the starting time of rewiring operation in (c) random networks with community structures and (d) scale-free networks with community structures. Susceptible nodes are rewired to immunized nodes.

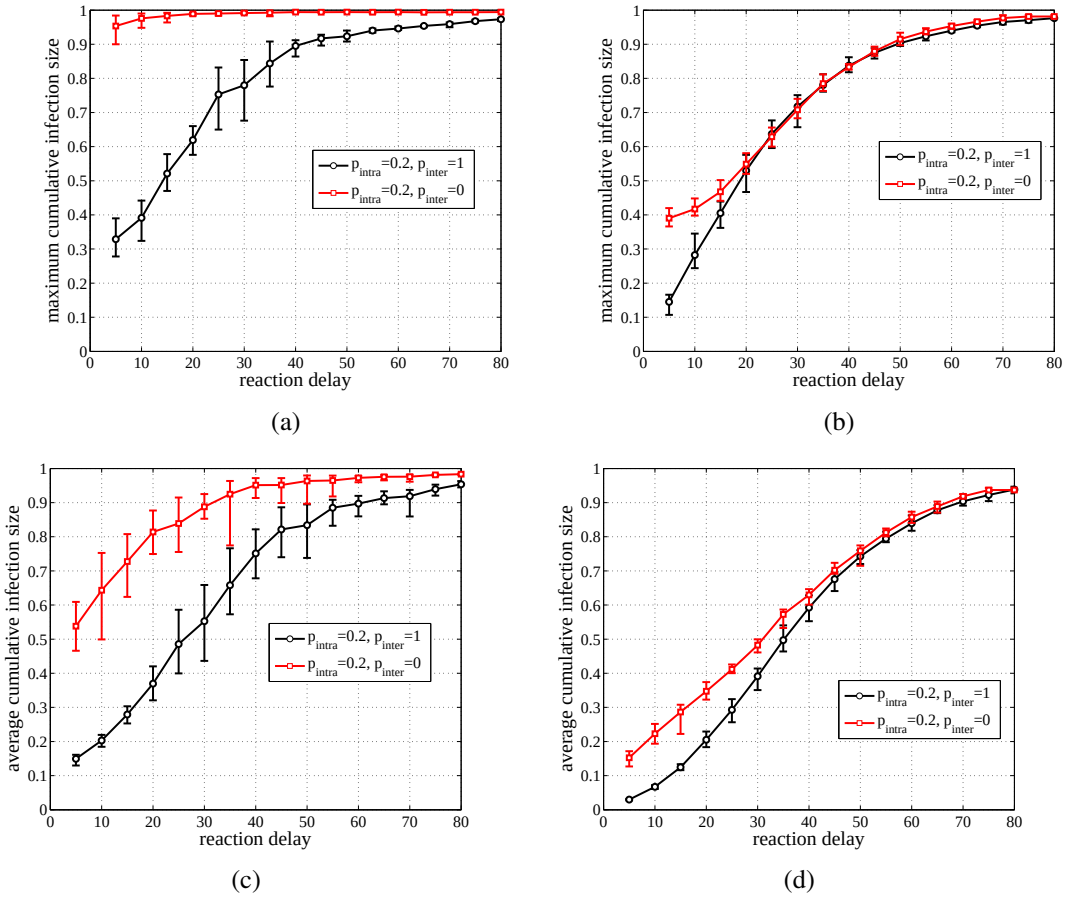

FIG. 8. The maximum expected infection size versus the reaction start time in (a) random networks with community structures and (b) scale-free networks with community structures; average infection size versus the reaction start time in (c) random networks with community structures and (d) scale-free networks with community structures. Susceptible nodes are rewired to susceptible nodes.

show that the existence of community structures helps slow down the infection spreading; whether it helps reduce the overall infection size when no control method is adopted depends on network topology. When infection detection and link rewiring are adopted, the existence of community structures helps improve the efficiency of infection detection and control, though having too many communities may not bring along additional benefits, especially in scale-free networks. The benefits of having community structures generally become more significant when a large enough number of monitors are installed in the network or when link rewiring could start at an early stage of infection spreading.

It is not fully understood whether infection detection and control help the emergence of the community structures, which may be of our future research interest.

Acknowledgment

The work is partially supported by Ministry of Education, Singapore, under the contracts RG28/14 and MOE2013-T2-2-006.

REFERENCES

1. BAILEY, N. T. J. (1975) *The mathematical theory of infectious diseases and its applications*, volume 66. London, U.K,:Griffin.
2. Barabási, A.-L. & Albert, R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.
3. Barber, M. J. (2007) Modularity and community detection in bipartite networks. *Phys. Rev. E*, **76**, 066102.
4. Chen, Y., Paul, G., Havlin, S., Liljeros, F. & Stanley, H. E. (2008) Finding a Better Immunization Strategy. *Phys. Rev. Lett.*, **101**, 058701.
5. Dezső, Z. & Barabási, A.-L. (2002) Halting viruses in scale-free networks. *Phys. Rev. E*, **65**, 055103.
6. Flake, G., Lawrence, S., Giles, C. & Coetzee, F. (2002) Self-organization and identification of Web communities. *Computer*, **35**(3), 66 –70.
7. Girvan, M. & Newman, M. E. J. (2002) Community structure in social and biological networks. **99**(12), 7821–7826.
8. Gleiser, P. M. & Danon, L. (2003) Community structure in Jazz. *Advances in Complex Systems*, **06**(04), 565–573.
9. Gregory, S. An Algorithm to Find Overlapping Community Structure in Networks. In Kok, J., Koronacki, J., Lopez de Mantaras, R., Matwin, S., Mladeni, D. & Skowron, A., editors, *Knowledge Discovery in Databases: PKDD 2007*, volume 4702, pages 91–102. Springer Berlin Heidelberg.
10. Gregory, S. (2010) Finding overlapping communities in networks by label propagation. *New Journal of Physics*, **12**(10), 103018.
11. Gross, T., D'Lima, C. J. D. & Blasius, B. (2006) Epidemic Dynamics on an Adaptive Network. *Phys. Rev. Lett.*, **96**, 208701.
12. Huang, W. & Li, C. (2007) Epidemic spreading in scale-free networks with community structure. *Journal of Statistical Mechanics: Theory and Experiment*, **2007**(01), P01014.
13. Lancichinetti, A. & Fortunato, S. (2011) Limits of modularity maximization in community detection. *Phys. Rev. E*, **84**, 066122.
14. Leicht, E. A. & Newman, M. E. J. (2008) Community Structure in Directed Networks. *Phys. Rev. Lett.*, **100**, 118703.
15. Leskovec, J., Krause, A., Guestrin, C., Faloutsos, C., VanBriesen, J. & Glance, N. (2007) Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '07, pages 420–429.
16. Li, C. & Maini, P. K. (2005) An evolving network model with community structure. *Journal of Physics A: Mathematical and General*, **38**(45), 9741.
17. Liu, Z. & Hu, B. (2005) Epidemic spreading in community networks. *EPL (Europhysics Letters)*, **72**(2), 315.

18. Moore, C. & Newman, M. E. J. (2000) Epidemics and percolation in small-world networks. *Phys. Rev. E*, **61**, 5678–5682.

19. Newman, M. (2004a) Detecting community structure in networks. *The European Physical Journal B - Condensed Matter and Complex Systems*, **38**(2), 321–330.

20. Newman, M. E. J. (2004b) Fast algorithm for detecting community structure in networks. *Phys. Rev. E*, **69**, 066133.

21. Newman, M. E. J. & Girvan, M. (2004) Finding and evaluating community structure in networks. *Phys. Rev. E*, **69**, 026113.

22. Newman, M. E. J. & Watts, D. J. (1999) Scaling and percolation in the small-world network model. *Phys. Rev. E*, **60**, 7332–7342.

23. Nicosia, V., Mangioni, G., Carchiolo, V. & Malgeri, M. (2009) Extending the definition of modularity to directed graphs with overlapping communities. *Journal of Statistical Mechanics: Theory and Experiment*, **2009**(03), P03024.

24. Opsahl, T. & Panzarasa, P. (2009) Clustering in weighted networks. *Social Networks*, **31**(2), 155–163.

25. Palla, G., Dernyi, I., Farkas, I. & Vicsek, T. (2005) Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, **435**(7043), 814Ł818.

26. Pastor-Satorras, R. & Vespignani, A. (2001) Epidemic Spreading in Scale-Free Networks. *Phys. Rev. Lett.*, **86**, 3200–3203.

27. Pastor-Satorras, R. & Vespignani, A. (2002a) Epidemic dynamics in finite size scale-free networks. *Phys. Rev. E*, **65**, 035108.

28. Pastor-Satorras, R. & Vespignani, A. (2002b) Immunization of complex networks. *Phys. Rev. E*, **65**, 036104.

29. Psorakis, I., Roberts, S., Ebden, M. & Sheldon, B. (2011) Overlapping community detection using Bayesian non-negative matrix factorization. *Phys. Rev. E*, **83**, 066114.

30. Riley, S., Fraser, C., Donnelly, C. A., Ghani, A. C., Abu-Raddad, L. J., Hedley, A. J., Leung, G. M., Ho, L.-M., Lam, T.-H., Thach, T. Q., Chau, P., Chan, K.-P., Lo, S.-V., Leung, P.-Y., Tsang, T., Ho, W., Lee, K.-H., Lau, E. M. C., Ferguson, N. M. & Anderson, R. M. (2003) Transmission Dynamics of the Etiological Agent of SARS in Hong Kong: Impact of Public Health Interventions. *Science*, **300**(5627), 1961–1966.

31. Shaw, L. B. & Schwartz, I. B. (2008) Fluctuating epidemics on adaptive networks. *Phys. Rev. E*, **77**, 066101.

32. Shen, H., Cheng, X., Cai, K. & Hu, M.-B. (2009) Detect overlapping and hierarchical community structure in networks. *Physica A: Statistical Mechanics and its Applications*, **388**(8), 1706 – 1712.

33. Tunc, I. & Shaw, L. B. (2014) Effects of community structure on epidemic spread in an adaptive network. *Phys. Rev. E*, **90**, 022801.

34. Valencia, M., Pastor, M. A., Fernandez-Seara, M. A., Artieda, J., Martinerie, J. & Chavez, M. (2009) Complex modular structure of large-scale brain networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **19**(2), 023119.

35. Wang, Y. & Xiao, G. (2012) Epidemics spreading in interconnected complex networks. *Physics Letters A*, **376**(42-43), 2689–2696.

36. Wang, Y., Xiao, G., Hu, J., Cheng, T. H. & Wang, L. (2009) Imperfect targeted immunization in scale-free networks. *Physica A: Statistical Mechanics and its Applications*, **388**(12), 2535–2546.

37. Yang, J. & Leskovec, J. (2012) Defining and Evaluating Network Communities Based on Ground-truth. In *Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics*, MDS '12, pages 3:1–3:8, New York, NY, USA. ACM.

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

39. Zhang, S. & Zhao, H. (2012) Community identification in networks with unbalanced structure. *Phys. Rev. E*, **85**, 066114.

40. Zhang, S., Zhao, J. & Zhang, X.-S. (2012) Common community structure in time-varying networks. *Phys. Rev. E*, **85**, 056110.

41. Zhou, J., Xiao, G., Wong, L., Fu, X., Ma, S. & Cheng, T. H. (2010) Generation of arbitrary two-point correlated directed networks with given modularity. *Physics Letters A*, **374**(3132), 3129 – 3135.