J Syst Sci Complex (2014) 27: 650–665

IDENTIFYING INFLUENTIAL SPREADERS IN ARTIFICIAL COMPLEX NETWORKS*

WANG Pei · TIAN Chengeng · LU Jun-an

DOI: 10.1007/s11424-014-2236-4
Received: 5 November 2012 / Revised: 22 March 2013
©The Editorial Office of JSSC & Springer-Verlag Berlin Heidelberg 2014

Abstract A long-term common belief in complex networks is that, the most connected nodes are the most efficient spreaders. However, recent investigations on real-world complex networks show that the most influential spreaders are those with the highest k-shell values. It is well-known that, many real-world complex networks have scale free (SF), small world (SW) properties, therefore, identification of influential spreaders in general artificial SF, SW as well as random networks will be more appealing. This research finds that, for artificial ER and SW networks, degree is more reliable than k-shell in predicting the outcome of spreading. However, for artificial SF networks, k-shell is remarkably reliable than degree and betweeness, which indicate that the four recently investigated real-world networks [Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E, Makse H A, Identification of influential spreaders in complex networks, Nat. Phys., 2010, 6: 888–893.] are more similar to scale free ones. Moreover, the investigations also indicate us an optimal dissemination strategy in networks with scale free property. That is, starting from moderate-degree-nodes will be ok and even more economical, since one can derive roughly similar outcome with starting from hubs.

Keywords Complex network, influential spreader, k-shell, scale free, small world.

1 Introduction

With the development of complex network science, lots of real-world complex systems could be described by complex networks[1−37]. And the dynamical disease models and spreading behaviors, such as susceptible-infectious-susceptible (SIS), susceptible-infectious-recovered (SIR) models and so on, have been extensively investigated in various complex networks[2−13]. Just

WANG Pei
School of Mathematics and Information Sciences, Henan University, Kaifeng 475004, China.
Email: wp0307@126.com.
TIAN Chengeng · LU Jun-an (Corresponding author)
School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China.
Email: zzzzqqqa@163.com; jalu@whu.edu.cn.
*This research is supported by the National Natural Science Foundation of China under Grant Nos. 11172215, 61304151, 61174028, China-Australia Health and HIV/AIDS Facility (FA36 EID101), and the Science Foundation of Henan University under Grant No. 2012YBZR007.
◇ *This paper was recommended for publication by Editor LÜ Jinhu.*

as an old Chinese saying "One who mixes with vermilion will turn red, one who touches pitch shall be defiled therewith". Obviously, whether a person will turn "red" or not as well as how fast of this transition depends on the properties of the original spreader in the networks. Therefore, identifying influential spreaders in complex networks have fundamental importance[14−16]. Traditionally, degree centrality, betweenness centrality are two indexes to intuitionally predict the outcome of spreading. And a long-term common belief in complex social network theory is that, the most highly connected or the most central people are the most efficient spreaders, because they have the most amount of neighbors or the most amount of shortest paths through them[13,17].

Recently, Kitsak and coworkers[14] investigated four real-world complex networks: a) The LiveJournal.com community; b) The email contacts network in the Computer Science Department of University College London; c) The contact network of inpatients collected from hospitals in Sweden; and d) the network of actors who have costarred in movies labeled by imdb.com as adult. They use degree, betweenness centrality as well as another index: k-shell[18], to predict the outcome of spreading, and they found that for networks with single initial spreader, k-shell can predict the outcome of spreading more reliably. While for multiple origins, a better spreading strategy is to choose either the highest degree or the highest k-shell nodes with the requirement that no two of the spreaders are directly linked to each other. Which is mainly because that k-shell can reflect the global location of a spreader in the overall network, while degree only reveals a node's local information[15]. Following, Chen and coauthors proposed a semi-local index, which has better performance than degree and betweenness on finding influential nodes in real-world complex networks[15].

However, Kitsak and coworkers' conclusions were derived only from four real-world complex networks[14], it is still not clear whether these conclusions can be extended to all kinds of real-world networks. It is well-known that many real-world complex networks have SF, SW properties[1,19−21], ranging from the Internet, the World Wide Web, citation networks to some other social networks[22]. Therefore, identification of effective spreaders in theoretical-based artificial SF, SW networks are more interesting yet important. Motivated by the above discussions, based on SIS and SIR models, one identifies influential spreaders in theoretical SW, SF, and Erdős-Rényi (ER) random networks, where k-shell, degree as well as betweenness centrality will be considered to act as indexes to predict the spreading outcome.

The rest paper is organized as follows. Section 2 constructs the ER, SF, and SW networks, and network statistical properties will be investigated simultaneously. One presents dynamical simulations and robustness analysis in Section 3. Conclusion remarks and discussions will be in the last Section 4.

2 Networks Construction

The random networks are based on the ER random network, each two nodes in the network has a probability $p(0 \leq p \leq 1)$ of gaining an edge[23]. The SF networks used in this paper are generated based on the BA* algorithm that proposed in [24], and the algorithm contains the classical BA algorithm as its special case. The main processes are that the network begins with m_0 vertices. At each time step, one vertex and m edges will be added to the network, a little revision from the classical BA algorithm, instead that every vertex has at least some baseline probability of gaining an edge, both endpoints of edges in the BA* algorithm are chosen according to a mixture of probability $\omega(\omega \in [0,1])$ for preferential attachment and $1-\omega$ for uniform attachment. That is,

$$P(e_i) = \omega \frac{e_i}{2mt} + (1-\omega)\frac{1}{m_0 + t}, \tag{1}$$

where e_i denotes the current number of edges incident on vertex i; $2mt$ represents total links at time t. m_0 is initial network size, $m_0 + t$ denotes total number of nodes at time t. Obviously, if $\omega = 1$, then the generated network will be the classical BA SF network, while for $\omega = 0$, the generated network is a uniform attachment random network. The new algorithm can well quantify the degree to which the rich nodes grow richer, and how new (or poorly connected) nodes can compete[24], and therefore, more realistic. Through ω, one can control the continuous transition of network properties from random networks to scale free ones. It is noted that, in order to enrich k-shell layers in the above algorithm, except the restriction of (1), one randomly adds the number of edges to the network at each step, but guarantees that the averagely added edges in all steps are around m.

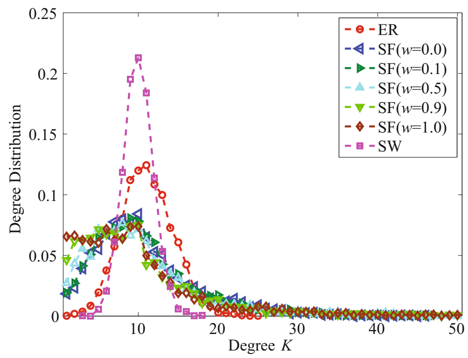
Table 1 Network properties

Network	N	$\langle K \rangle$	$K_{\max}$	$K_{\min}$	DH	CC	PLE
ER	3208	11.1135	25	1	1.0862	0.0030	−−
SF($\omega = 0.0$)	3208	10.8541	46	1	1.3613	0.0055	2.593
SF($\omega = 0.1$)	3208	10.7999	47	1	1.3968	0.0058	2.598
SF($\omega = 0.5$)	3208	10.9308	82	1	1.6235	0.0080	2.595
SF($\omega = 0.9$)	3208	11.0879	185	1	2.3037	0.0166	2.762
SF($\omega = 1.0$)	3208	11.1272	171	1	2.6423	0.0240	2.951
SW	3208	10.0000	18	3	1.0355	0.3557	−−

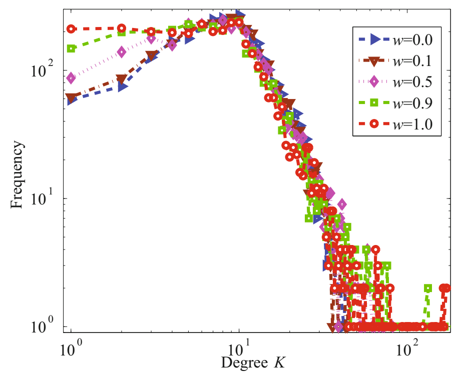
In Table 1, $\langle K \rangle$ denotes average degree. $K_{\max}, K_{\min}$ represents maximum degree and minimum degree, respectively. DH is degree heterogeneity, defined as DH=$\langle K^2 \rangle / \langle K \rangle^2$[15,27]. CC denotes clustering coefficient[21]. PLE denotes power-law exponent for SF network.

The SW networks are based on the randomly rewiring of the corresponding regular networks[25]. In the following, N denotes total nodes in a network. K and K_s denote node degree and k-shell value, respectively. B_c denotes betweenness centrality. In this paper, ER random networks and SW networks are generated by Netlogo[26], SF networks are constructed in Matlab, the spread simulations are performed by Netlogo.

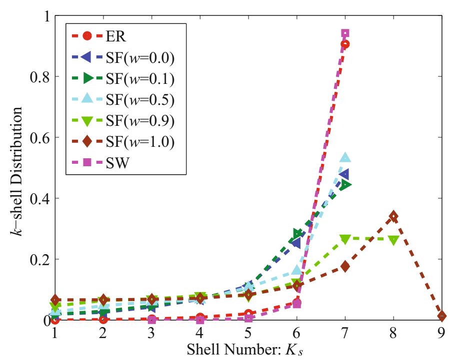
It is noted that, for different networks, one guarantees the same number of nodes and roughly similar average degrees. One pays extensive attention to the case with 3208 nodes and average degree around 11. For other cases, robustness of the related investigation will be discussed in detail in Section 3. Network properties of different kinds of networks are summarized in Table 1, where, one computes average degree, maximum and minimum degree, degree heterogeneity[15,27], clustering coefficient[21] for each network. Additionally, power-law exponents are computed for SF networks. From Table 1, one can easily see that the SF network has a higher degree heterogeneity, and with the increasing of ω, the degree heterogeneity becomes higher and higher. Therefore, one can guess, hierarchy structures in SF networks can then become richer and richer with the increasing of ω. The SW network has the largest clustering coefficient. Power-law exponents for SF networks range from around 2.6 to 2.95.



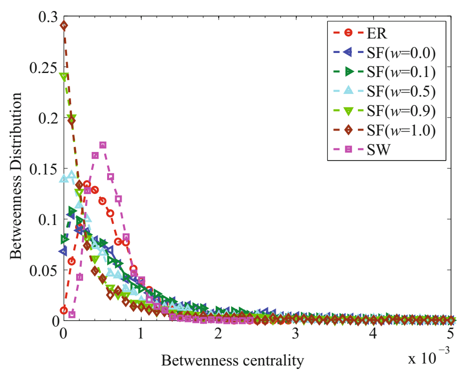
(a) Degree distribution for all networks



(b) Degree distribution for the SF networks in log-log coordinate



(c) k-shell distribution



(d) Betweenness centrality distribution

Figure 1 Degree, k-shell, betweenness centrality distribution for the three kinds of networks in Table 1. In (a) and (d), for clarity, the long tail for each curve is cut off

Figure 1 shows degree, k-shell as well as betweenness centrality distribution of the networks. From Figure 1(a), one finds that the ER and SW networks have the similar degree distributions, but the distribution of the SW network is a little narrower. The degree distributions of the SF networks power exponent decades for $K \geq 10$, there are long tails with high degree nodes, and this distribution appropriates to power-law in log-log coordinate system (see Figure 1(b)). Figure 1(c) shows k-shell K_s distributions of the three kinds of networks. The K_s distributions for the SF networks become more and more uniform with the increasing of ω. More importantly, this panel also reveals that the SF networks have more distinct hierarchical structures than the ER and SW networks. The biggest shell number for the ER and SW are both 7, more than 80% nodes locate at the 7th layer for the ER and SW networks, therefore, hierarchical structures are not so obvious in these two networks. The distribution of the ER is a little wider than the SW, there are only four shell numbers ($K_s = 4, 5, 6, 7$) for the SW network.

Betweenness describes the number of geodesic paths that pass through a node, or the number of "times" that any node needs a given node to reach any node by the shortest path[16,17,28,29]. In this paper, B_c for node i is computed as[16]:

$$B_c(i) = \frac{2}{(N-1)(N-2)} \sum_{j=0}^{N} \sum_{k=0,k\neq i}^{N} \frac{\psi_{jk}(i)}{\psi_{jk}}. \tag{2}$$

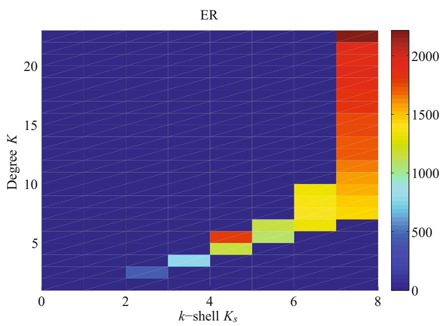
Here, $\psi_{jk}(i)$ represents the number of the shortest paths between nodes j and k running through node i, and ψ_{jk} denotes the total number of the shortest paths between nodes j and k. From Figure 1(d), betweenness centrality distributions of the three classes of networks show similar appearances as the corresponding degree distributions. With the increasing of ω, the SF networks have more and more nodes with lower B_c, and a small fraction of nodes with very high B_c, which is because that most of the shortest paths through hub nodes, however, the hub nodes only take up a small fraction.

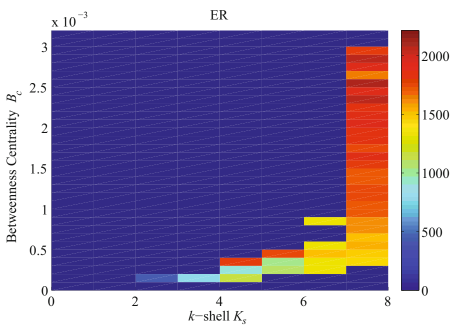
3 Main Results

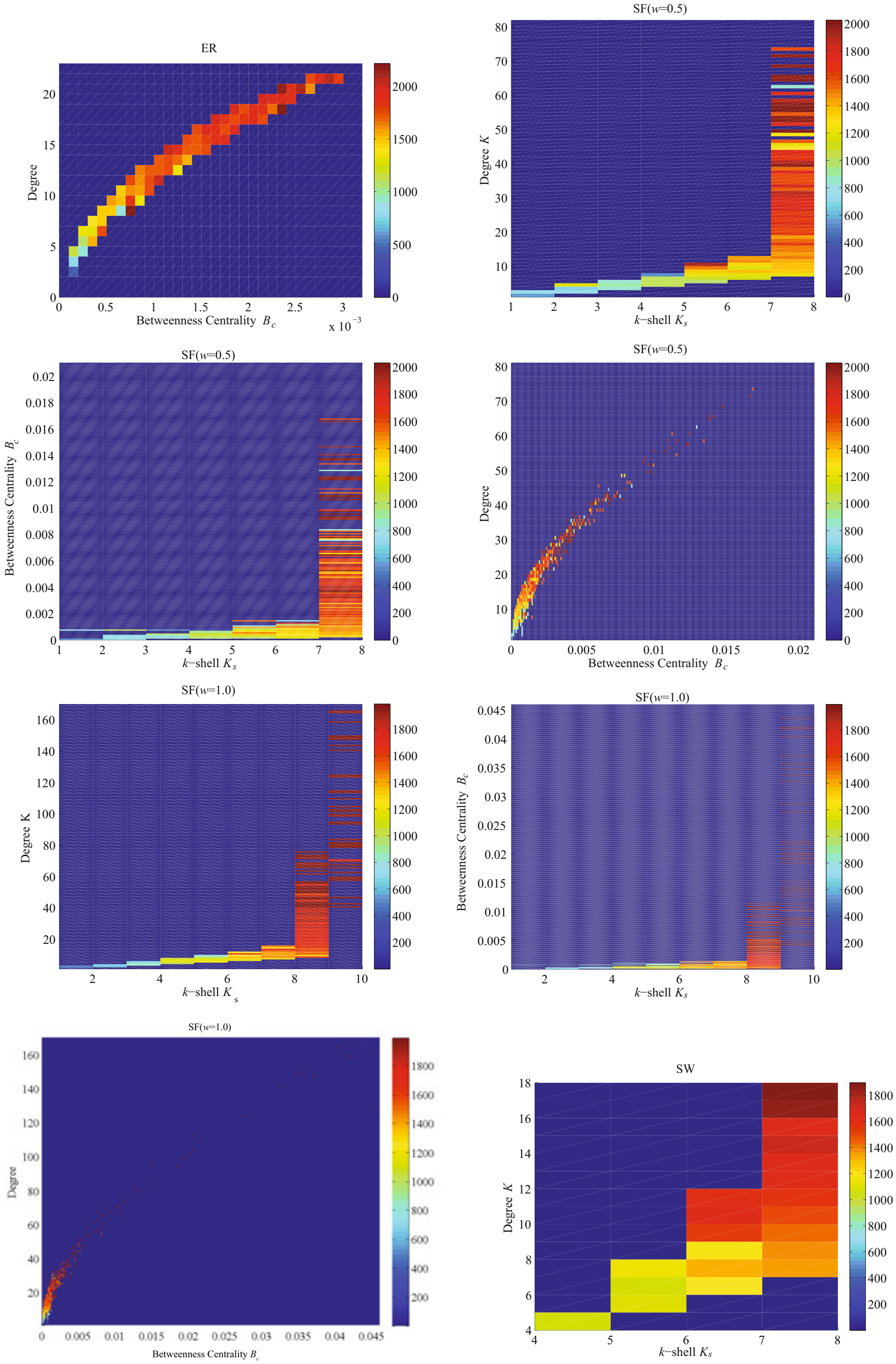
3.1 Reliable Indexes in Artificial Complex Networks

Just like in [14], one considers the spreading with single original spreader, and denotes the probability that an infectious node will infect a susceptible neighbor as α. For SIS model, the natural recovery rate is assumed to be γ, and the recovered ones become susceptible and can be infected again in the SIS model, while recovered ones can not be infected again in SIR model.

Following, one identifies which index can best predict the outcome of spreading for these three kinds of theoretically constructed networks. In the following simulation, one sets $\alpha = 3.5\%, \gamma = 15\%$, unless otherwise noted. The amount of infected nodes are recorded when the propagation processes reach their steady states, and for each node, 10 simulation runs are averaged over.






ER
Degree
Betweenness Centrality B_c
x 10^{-3}
SF(w=0.5)
Degree K
k−shell K_s
SF(w=0.5)
Betweenness Centrality B_c
k−shell K_s
SF(w=0.5)
Degree
Betweenness Centrality B_c
SF(w=1.0)
Degree K
k−shell K_s
SF(w=1.0)
Betweenness Centrality B_c
k−shell K_s
SF(w=1.0)
Degree
Betweenness Centrality B_c
SW
Degree K
k−shell K_s

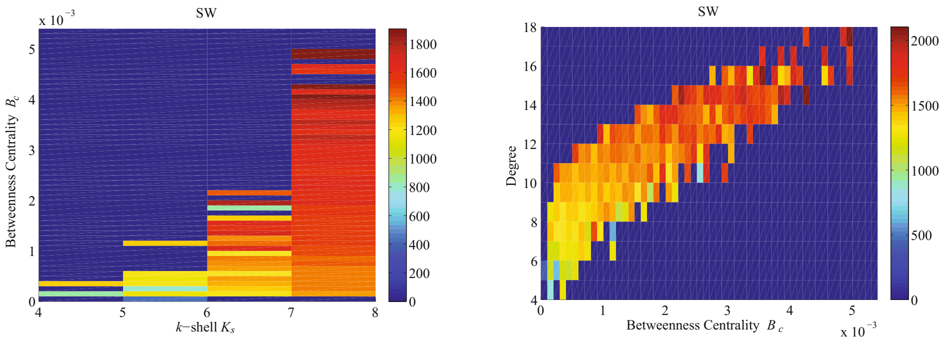
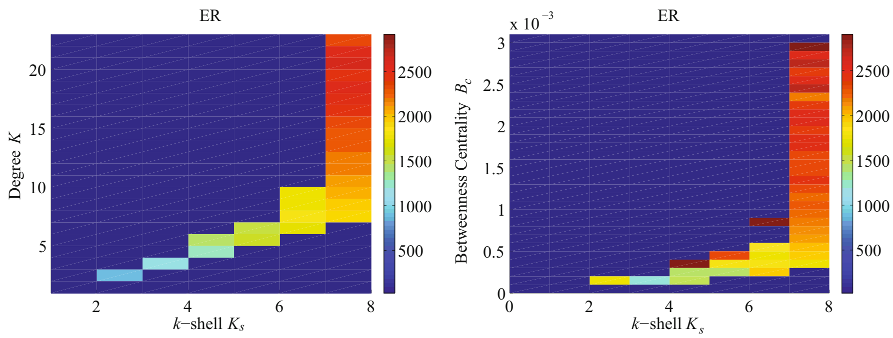
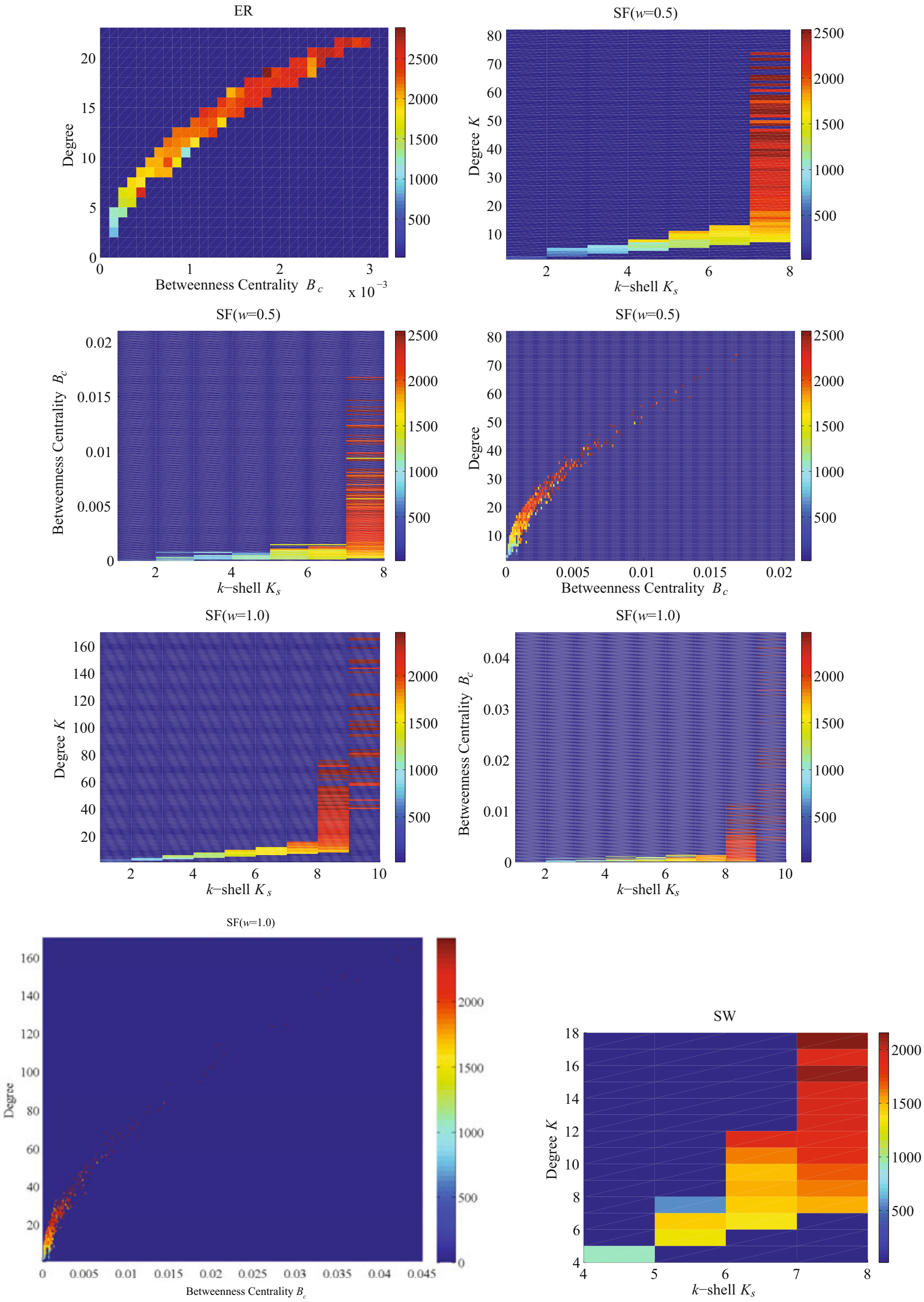


Figure 2 Reliable indexes in different artificial complex networks. Here SIS model is adopted, with infected ration $\alpha = 3.5\%$, and recovery ration $\gamma = 15\%$, the color bars represent average infected numbers

Figure 2 shows K_s, K versus infected population, K_s, B_c versus infected population as well as B_c, K versus infected population, where the first three subfigures show the cases for ER network. The fourth to ninth subfigures show the cases for the SF networks with $\omega = 0.5$ and $\omega = 1.0$, respectively. The last three figures correspond to the SW network.

From the simulation results for the ER and SW in Figure 2, one sees that degree can predict the outcome of spreading more reliably than K_s, since in the left panel for these two networks, one can see that, at the same k-shell layer, infected population increases with the increasing of node degree. Moreover, one can also see that B_c is better than k-shell. Degree and B_c have the best consistence. Whereas, for the SF network with $\omega = 1.0$, it seems that K_s is the most reliable index to predict the outcome of spreading. But for $\omega = 0.5$, this conclusion becomes not so obvious. Furthermore, for the SF networks with $\omega = 1.0$, degree and B_c seem to have the similar performance. In fact, the ER can be seen as the SF network with $\omega = 0.0$, therefore, the first to the ninth subfigures of Figure 2, ω increases from zero to one. As a consequence, hierarchy structures of networks are strengthened, and advantage of the k-shell becomes more and more obvious.




ER
Degree
Betweenness Centrality B_c
x 10^{-3}
SF(w=0.5)
Degree K
k–shell K_s
SF(w=0.5)
Betweenness Centrality B_c
k–shell K_s
SF(w=0.5)
Degree
Betweenness Centrality B_c
SF(w=1.0)
Degree K
k–shell K_s
SF(w=1.0)
Betweenness Centrality B_c
k–shell K_s
SF(w=1.0)
Degree
Betweenness Centrality B_c
SW
Degree K
k–shell K_s

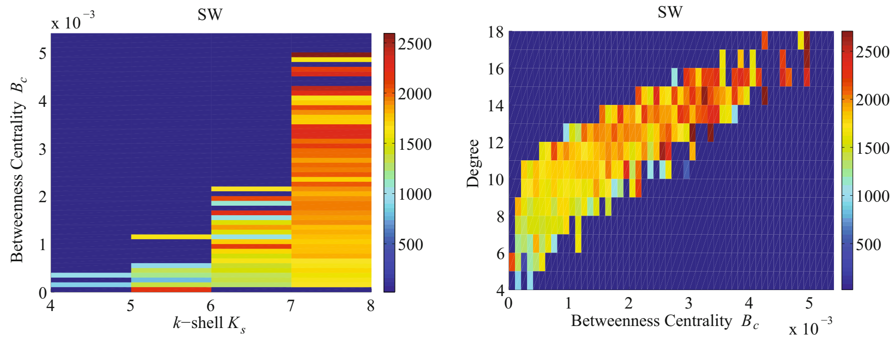
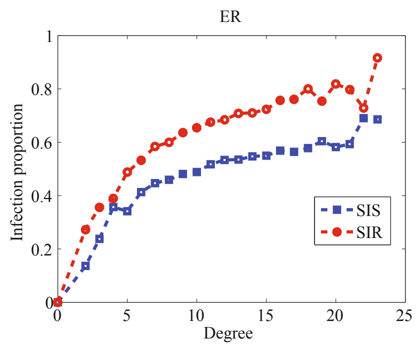


Figure 3 Reliable indexes in different artificial complex networks. Here SIR model is adopted, with infected ration $\alpha = 3.5\%$, and recovery ration $\gamma = 15\%$, the color bars represent average infected numbers

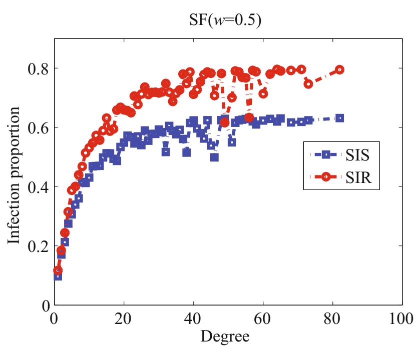
Figure 3 shows the cases under SIR model, where one supposes immune ration as $I = 100\%$, and the immune ones can not be infected anymore. Other parameters are similarly chosen as that in Figure 2. From Figure 3, one can also see that k-shell is a more reliable index than degree for the SF ($\omega = 1.0$) network. And degree is a little better than k-shell in the ER and SW networks. In [14], the authors found that K_s in real networks can well predict the outcome of spreading. From the conclusions on theoretical complex networks, one can conclude that the real-world complex networks investigated in [14] are more similar to the SF networks. It is intriguing to find what causes the differences between the SW, ER, and SF networks.

3.2 Roles of Degree and k-shell

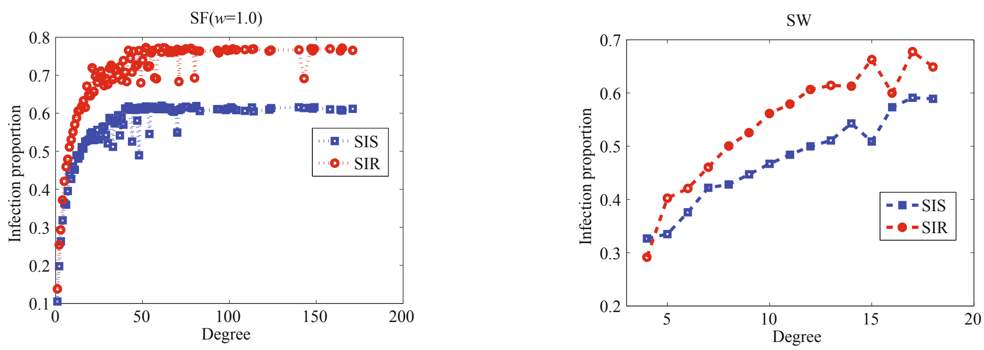
In order to find what causes the differences for different kinds of networks, one considers the influence of degree and k-shell separately.



(a) ER

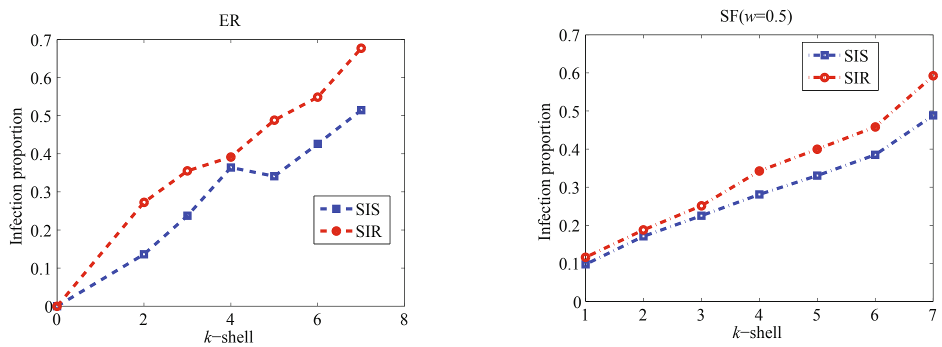


(b) SF($\omega = 0.5$)

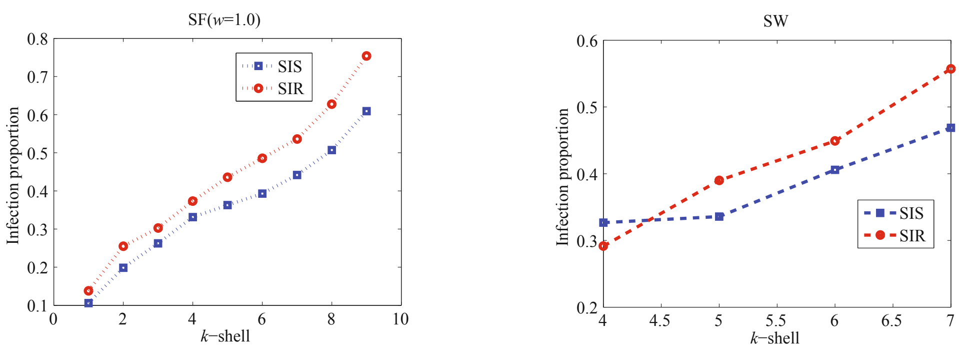


(c) SF($\omega = 1.0$) (d) SW

Figure 4 Average infected proportion versus degree, both SIS and SIR model are considered. Parameters are the same as that in Figures 2 and 3



(a) ER (b) SF($\omega = 0.5$)



(c) SF($\omega = 1.0$) (d) SW

Figure 5 Average infected proportion versus k-shell, both SIS and SIR model are considered. Parameters are the same as that in Figures 2 and 3

Figures 4 and 5 show average infected proportion versus degree and k-shell, respectively. Where infected proportion from each experiment is also recorded at steady state. Cases under SIS and SIR model are both drawn in Figures 4 and 5, where parameters are the same as that in Figures 2 and 3. Obviously, similar trends under two different models can be observed. From

Figures 4 and 1(a), one can derive the following conclusion: Since the SW have the narrowest degree distributions, the infected proportions in this network can increase with the increment of node degrees, but absolute numbers are not seriously promoted as the other networks. While for the SF with $\omega = 1.0$, its degree distribution is the widest, the increment of degree for the initial spreader can increase the infected proportions distinctly, especially for nodes with degrees lower than 50, but for degree above certain value, for example, 50, the infected proportions will have no obvious increment, and similar phenomenon can be found for $\omega = 0.5$. That is, there exists a degree sensitive threshold for the SF network. This figure also indicates us that, for the SF networks, excessive emphasis on hub nodes to spread a behavior is unnecessary, one should only control these nodes with degree more than 50, which will gain the roughly same end compared with controlling these hub nodes. Indubitability, this can lower the cost to a great extent.

From Figure 5 and combined with Figure 1(c), one can derive the following conclusions: For the SW network, since it has the narrowest k-shell distribution, infected proportions can be increased with the increment of k-shell values. However, absolute increments are also not so distinct as the other networks. While for the SF with $\omega = 1.0$, there does not exist such k-shell sensitive threshold as degree, the infected ratio increases almost linearly with k-shell values, and the increments of absolute infected numbers are very significant. Compared between Figures 4 and 5, one sees that curves of average infected proportion versus degree and k-shell for the SW and ER have great relevance. Therefore, it is difficult to distinct between these two indexes, but from Figures 2 and 3, degree seems to be better than k-shell in performance. The difference between Figures 4 and 5 is one direct reason that why k-shell is so reliable for the SF with $\omega = 1.0$, and not so reliable for the other networks.

Another conclusion that can be derived from Figures 4 and 5 is that, compared the result between SIS and SIR model, more people are infected under SIR models, which is because that, in SIS model, infected people can be recovered and re-infected, and these recovered people at each step have not been recorded. While in SIR, both "I" (infected) and "R" (recovered and acquired immunity) ones are recorded.

3.3 Robustness of the Findings with Respect to Parameters

In order to see whether the above conclusions are parameter-dependent, one chooses different model parameters, network sizes and so on, and investigate whether the reliable index can be obviously changed or not.

To measure whether an index is reliable or not under different parameters, one will firstly introduce an indicator. To start with, one defines averaged infected numbers for node with degree K_i, k-shell K_{sj} as $\Gamma_{ij}(K_i, K_{sj})$, where nodes with the same degree and k-shell values are equally treated. One further defines some functions that measure which indexes can best predict the outcome of spreading. The Shannon entropy[30] has many advantages, the entropy can achieve its maximum if the probability for each stochastic event is equally distributed. Based on this concept, the Shannon entropy-based function that compared between degree and

k-shell can be defined as:

$$H(K, K_s) = -\frac{1}{N_s} \sum_{j=1}^{K_{sm}} \sum_{i=1}^{K_m} p_{ij} \log_2 p_{ij}, \tag{3}$$

where K_m is the largest node degree, K_{sm} is the largest k-shell value, N_s denotes number of layers. p_{ij} is defined as:

$$p_{ij} = \frac{\Gamma_{ij}(K_i, K_{sj})}{\sum_{v=1}^{K_m} \Gamma_{vj}(K_v, K_{sj})}. \tag{4}$$

In fact, p_{ij} satisfies $\sum_{i=1}^{K_m} p_{ij} = 1$ for each $j \in [1, 2, \cdots, K_{sm}]$. It is noted that during computation, these combinations with $p_{ij} = 0$ are omitted. It is also noted that, the bigger $H(K, K_s)$, the better k-shell as an index to predict the outcome of spreading. Intuitively, if k-shell was a good index, then $p_{1j}, p_{2j}, \cdots, p_{K_m j}$ will be more uniform. Therefore, $H(K, K_s)$ will be relatively bigger. The function that compare between k-shell and betweenness centrality, degree between betweenness centrality can similarly be defined, one omits the details.

3.3.1 Robustness with Respect to Model Parameters

Spreading rate[31–34] is an important concept in epidemic models, which is defined as the ratio between infected rate and recovery rate, that is:

$$\lambda = \frac{\alpha}{\gamma}. \tag{5}$$

Spreading rate is a characteristic quantity, which can comprehensively measure the disease dissemination behavior. Through dynamical mean field theory, References [31–34] have found that the ER and SW networks have non-zeros critical values λ_c of λ, if $\lambda < \lambda_c$, the disease can not be spread out, while for $\lambda \geq \lambda_c$, the infection spreads and becomes persistent. Theoretical threshold of spreading rate for ER network approximates to $\lambda_c^{ER} = \langle K \rangle^{-1}$, while for infinite size scale free network, $\lambda_c^{SF} = \langle K \rangle / \langle K^2 \rangle$. Combined with Table 1, $\lambda_c^{ER} = 1/11.1135 \approx 0.0900$, $\lambda_c^{SF} = \langle K \rangle / \langle K^2 \rangle \in [0.03401, 0.06768]$. In Section 3.1, λ is fixed as $3.5/15 \approx 0.2333$, which indicates that spreading behaviors can become persistent. Following, one chooses $\lambda = 0.02, 0.05, 0.1, 0.2, 0.3$, and also takes the constructed networks in Section 2 as examples, one performs robustness analysis with respect to λ.

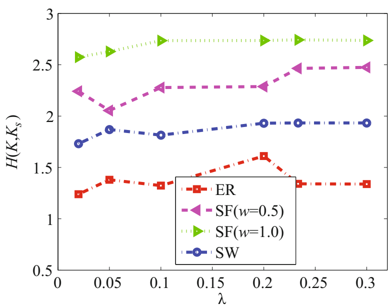


Figure 6 Robustness of reliable indexes with respect to spreading rate λ

Figure 6 (Here SIS model is adopted) shows $H(K, K_s)$ versus λ under SIS model. From Figure 6, one can see that the SF networks have higher $H(K, K_s)$ values than the ER and SW, which further demonstrates that k-shell is more reliable in the SF than in the ER and SW networks. Furthermore, the SF($\omega = 1.0$) has higher $H(K, K_s)$ value than the SF($\omega = 0.5$). And the $H(K, K_s)$ value for the SW is higher than the ER. For $\lambda \geq 0.1$, the curves $H(K, K_s)$ versus α for the SF($\omega = 1.0$) and the SW almost parallel with the λ axis, which indicates that, when $\lambda > \lambda_c$, reliable indexes in these two networks are not so sensitive to spreading rate. For the SF($\omega = 0.5$) and ER, the corresponding curves deviate from the stable lines, but fluctuations in all of these networks are not so high, the variances of $H(K, K_s)$ under these different λ values from the ER to SW networks are $0.0159, 0.0244, 0.0052, 0.0068$, respectively. Therefore, the conclusions derived in Section 3.1 are robust to spreading rate.

3.3.2 Robustness with Respect to Network Sizes

The above main conclusions are derived under fixed network size $N = 3208$, to see whether network sizes can influence the conclusions, one randomly generates ER and SW network with $N = 989, 1897, 3208, 4795$ nodes, SF($\omega = 1.0$) networks with $N = 1000, 2000, 3208, 5000$ nodes, and average degree for each network is all around 11. One investigates the evolutions of $H(K, K_s)$ with respect to network size N. Figure 7 shows $H(K, K_s)$ versus network sizes N for the three kinds of networks under SIS model.

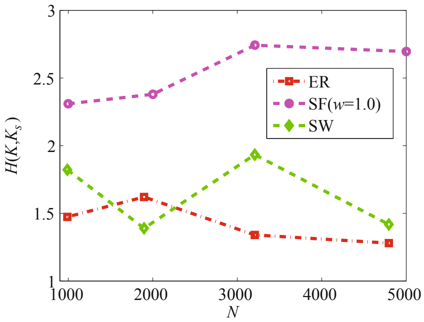


Figure 7 Robustness of reliable indexes with respect to network size N

From Figure 7 (Here, one also considers the SIS model), one can derive the conclusion that with the increasing of network sizes, $H(K, K_s)$ becomes larger and larger for the SF network. This indicates that k-shell becomes more and more reliable in predicting the outcome of virus spreading. Whereas, $H(K, K_s)$ decreases for the ER and SW, which means that k-shell can still not become so reliable as that in the SF networks, even under very large network sizes. Figure 7 illustrates that the conclusion derived in this paper is robust to network sizes. Moreover, the conclusion can also stand water for SIR model, due to space limitations, one omits detailed discussions.

3.3.3 Robustness with Respect to Different SF Networks

In the above analysis, when one considers the SF networks, one particularly investigates the cases with $\omega = 0.5, 1.0$, to see to what degree different w can affect the conclusions, one takes $\omega = 0.0, 0.1, 0.5, 0.9, 1.0$, and generates SF networks with $N = 2000, 3208, 5000$ nodes and average degrees around 11. Figure 8 shows $H(K, K_s)$ versus ω for different SF networks under SIS model.

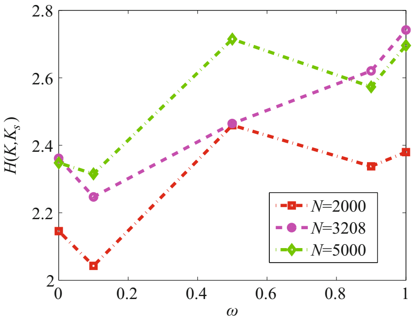


Figure 8 Robustness of reliable indexes with respect to parameter ω in BA* algorithm

From Figure 8, one sees that, with the increasing of ω, roughly speaking, $H(K, K_s)$ becomes larger and larger, which indicates that k-shell becomes better and better in predicting spreading outcomes. Therefore, one can say that the conclusions are also robust to parameter ω.

In Table 1, the PLE for SF networks with different ω ranges from around 2.6 to 2.95, and this PLE roughly increases with ω, combined with the result in Figure 8, one can also say that, the increment of PLE may make the advantage of k-shell much obvious, theoretical analysis of this point will be performed in our future work.

4 Conclusions

Virus as well as behavior spreading in complex networks are interesting topics in social networks. A fundamental yet important question is that which index can well predict the outcome of virus or behavior spreading, since it has wide real-world applications in the design of optimal dissemination strategies, disease control, and so on. Based on a recent work, this research tries to identify influential spreaders in theoretically constructed artificial complex networks, and one finds that for the ER and SW networks, degree seems to be better than k-shell in predicting the spreading outcomes, while for the SF network, k-shell is a very reliable index. The associated conclusions are very robust to parameters in different spreading models and network sizes.

The related works can be seen as an extension and perfection of the existing works [14]. Through the conclusions of this investigation, one can predict that these real-world complex networks are more similar to the SF ones, and these networks have the property that "winners don't take all". In fact, many real-world networks have SF property, and the degree distribution always has fat-tailed distribution, which actually corresponds to the SF networks generated by BA* algorithm. The findings also indicate that if the statistical properties of a network is clear, then one can determine which index can well predict the outcome of virus or behavior spreading at once.

References

[1] Bullmore E and Sporns O, Complex brain networks: Graph theoretical analysis of structural and functional systems, *Nat. Rev.*, 2009, **10**: 186–198.

[2] May R M and Lloyd A L, Infection dynamics on scale-free networks, *Phys. Rev. E*, 2001, **64**: art.no. 066112.

[3] Lloyd A L and May R M, How virus spread among computers and people, *Science*, 2001, **292**: 1316–1317.

[4] Xu X, Peng H, Wang X, and Wang Y, Epidemic spreading with time delay in complex networks, *Physica A*, 2006, **367**: 525–530.

[5] Nekovee M, Moreno Y, Bianconi G, and Marsili M, Theory of rumour spreading in complex social networks, *Physica A*, 2007, **374**: 457–470.

[6] Chen J, Zhang H, Guan Z, and Li T, Epidemic spreading on networks with overlapping community structure, *Physica A*, 2012, **391**: 1848–1854.

[7] Ni S, Weng W, and Zhang H, Modeling the effects of social impact on epidemic spreading in complex networks, *Physica A*, 2011, **390**: 4528–4534.

[8] Anderson R M, May R M, and Anderson B, *Infectious Diseases of Humans: Dynamics and Control*, Oxford Science Publications, 1992.

[9] Hethcote H W, The mathematics of infectious diseases, *SIAM Rev.*, 2000, **42**: 599–653.

[10] Fan W and Yeung K H, Online social networks-Paradise of computer viruses, *Physica A*, 2011, **390**: 189–197.

[11] Christley R M, Pinchbeck G L, Bowers R G, Clancy D, French N P, Bennett R, and Turner J, Infection in social networks: Using network analysis to identify high-risk individuals, *Am. J. Epidemiol.*, 2005, **162**: 1024–1031.

[12] Liu J, Wu J, and Yang Z R, The spread of infectious disease on complex networks with household-structure, *Physica A*, 2004, **341**: 273–280.

[13] Pastor-Satorras R and Vespignani A, Epidemic spreading in scale-free networks, *Phys. Rev. Lett.*, 2001, **86**: 3200–3203.

[14] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E, and Makse H A, Identification of influential spreaders in complex networks, *Nat. Phys.*, 2010, **6**: 888–893.

[15] Chen D, Lü L, Shang M, Zhang Y, and Zhou T, Identifying influential nodes in complex networks, *Physica A*, 2012, **391**: 1777–1787.

[16] Kuhnert M, Geier C, Elger C E, and Lehnertz K, Identifying important nodes in weighted functional brain networks: A comparison of different centrality approaches, *Chaos*, 2012, **22**: art. no. 023142.

[17] Freeman L C, Centrality in social networks: Conceptual clarification, *Social Networks*, 1979, **1**: 215–239.

[18] Carmi S, Havlin S, Kirkpatrick S, Shavitt Y, and Shir E, A model of Internet topology using k-shell decomposition, *Proc. Natl. Acad. Sci. USA*, 2007, **104**: 11150–11154.

[19] Centola D, The spread of behavior in an online social network experiment, *Science*, 2010, **329**: 1194–1197.

[20] Barabási A and Albert R, Emergence of scaling in random networks, *Science*, 1999, **286**: 509–512.

[21] Watts D J and Strogatz S H, Collective dynamics of 'small world' networks, *Nature*, 1998, **393**: 440–442.

[22] Fan Z, Chen G, and Zhang Y, A comprehensive multi-local-world model for complex networks, *Phys. Lett. A*, 2009, **373**: 1601–1605.

[23] Bollobás B, *Random Graphs*, 2nd Edition, Cambridge University Press, 2001.

[24] Pennock D M, Flake G W, Lawrence S, Glover E J, and Giles C L, Winners don't take all: Characterizing the competition for links on the web, *Proc. Natl. Acad. Sci. USA*, 2002, **99**: 5207–5211.

[25] Wang X, Li X, and Chen G, *Complex Networks Theory and Its Applications*, Tsinghua University Press, 2006 (in chinese).

[26] Wilensky U, NetLogo, http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, 1999.

[27] Hu H and Wang X, Unified index to quantifying heterogeneity of complex networks, *Physica A*, 2008, **387**: 3769–3780.

[28] http://www.analytictech.com/networks/centrali.htm.

[29] Wasserman S and Faust K, *Social Network Analysis: Methods and Applications*, Cambridge University Press, 1994.

[30] Shannon C E, Prediction and entropy of printed English, *Bell Syst. Tech. J.*, 1951, **30**: 50–64.

[31] Pastor-Satorras R and Vespignani A, Epidemics and immunization in scale-free networks, *Handbook of Graphs and Networks* (eds. by Bornholdt S and Schuster H G), Wileyvch Publisher, Berlin, 2003.

[32] Wang X and Chen G, Complex networks: Small-world, scale-free and beyond, *IEEE Circuits and Systems Magazine*, 2003, **3**: 6–20.

[33] Pastor-Satorras R and Vespignani A, Epidemic spreading in scalefree networks, *Phys. Rev. Lett.*, 2001, **86**: 3200–3203.

[34] Pastor-Satorras R and Vespignani A, Epidemic dynamics and endemic states in complex networks, *Phys. Rev. E*, 2001, **63**: art.no. 066117.

[35] Lü J and Chen G, A time-varying complex dynamical network model and its controlled synchronization criteria, *IEEE Trans. Automat. Control*, 2005, **50**: 841–846.

[36] Liu H, Lu J, Lü J, and Hill D J, Structure identification of uncertain general complex dynamical networks with time delay, *Automatica*, 2009, **45**: 1799–1807.

[37] Wang P, Lü J, and Ogorzalek M J, Global relative parameter sensitivities of the feed-forward loops in genetic networks, *Neurocomput.*, 2012, **78**: 155–165.