A Preliminary Study on the Effects of Fear Factors in Disease Propagation

Yubo Wang[1], Jie Hu[1], Gaoxi Xiao[1], Limsoon Wong[2],
Stefan Ma[3], and Tee Hiang Cheng[1]

[1] School of Electrical and Electronic Engineering, Nanyang Technology University,
Singapore 639798
`egxxiao@ntu.edu.sg`
[2] School of Computing and School of Medicine, National University of Singapore,
Singapore 117543
[3] Epidemiology & Disease Control Division, Ministry of Health, Singapore 169854

Abstract. Upon an outbreak of a dangerous infectious disease, people generally tend to reduce their contacts with others in fear of getting infected. Such typical actions apparently help to reduce the outbreak size. Thanks to today's broad media coverage, the fear factor may also contribute to preventing an outbreak from happening at all. We are motivated to conduct a careful study on modeling and evaluating such effects with a complex network approach. As a first step of this study, we consider the relatively simple case where involved individuals randomly remove a certain fraction of links between them. Analytical and simulation results show that such an action cannot effectively prevent an epidemic outbreak from happening. However, it may significantly reduce the fraction of all the people ever getting infected when an outbreak does happen.

Keywords: complex networks, scale-free networks, fear factor, epidemic threshold, average outbreak size, prevalence size.

1 Introduction

When a dangerous infectious disease is detected and known by public, especially revealed by the media, people generally tend to reduce their contacts with others in fear of getting infected. Such reactions can be easily observed in past experiences, e.g., during the outbreak of Severe Acute Respiratory Syndrome (SARS) [1]. It can be expected that such a fear factor helps to reduce the number of infected or even prevent an epidemic outbreak from happening. Detailed evaluations on its effects, however, have never been conducted to the best of our knowledge.

We are motivated to conduct a careful study on the effects of fear factor by adopting a complex network approach, where individuals are modeled as nodes (vertexes) and the contacts between them for possible disease propagation are modeled as links (edges) [2]-[4]. An infectious disease can spread between nodes via links. As a first step of this study, we consider a relatively simple case where a certain fraction of the links is randomly removed. It can be viewed as a benchmark for the more "realistic" cases which we will investigate in our future studies.

Epidemic spreading in networks is strongly influenced by the pattern of connectivity in the underlying contact networks [5]-[7]. Statistical studies on real-life complex networks, including human society and sexual contact network etc., show that many of them share some nontrivial features. One of the most noticeable features is that they are usually scale-free networks of which the nodal degrees follow a power-law distribution [8]. Specifically, the probability that a node is connected to k other nodes

$$P(k) \sim k^{-r}, \tag{1}$$

where the exponent r usually ranges between 2 and 3 [9]-[10]. In such networks, a small number of nodes tend to connect to a virtually infinite number of other nodes. This feature significantly affects epidemic propagation in the scale-free networks.

Existing research shows that epidemic spreading in infinitely large scale-free networks with an exponent $r \leq 3$ does not possess any epidemic threshold, below which the infection cannot produce a major epidemic outbreak or an endemic state [5]-[7], [11]. In other words, a disease can easily survive and cause an outbreak in scale-free networks even if its spreading capability is very weak. Further studies on the finite-size scale-free networks show that the epidemic threshold remains to be low and decreases with an increasing network size [12]. In such studies, it has been assumed that the network topologies remain static throughout the disease propagation procedure.

In this paper, we investigate the disease propagation in a scale-free network with a fraction of its links being randomly removed. As a case study, we mainly evaluate the Susceptible-Infected-Remove (SIR) [13]-[14] model on top of the well-known Barabasi-Albert (BA) model [8]. Simple as it is, the special case nevertheless provides some insights into the effects of the fear factor. Specifically, analytical and simulation results show that the random removals of links cannot efficiently increase the epidemic threshold but may significantly reduce the number of nodes ever getting infected throughout the outbreak.

The rest part of this paper is organized as follows: In Section 2, we introduce an analysis of network topology after random link removals. Section 3 studies the SIR model in scale-free networks after link removals. Considering that some nodes may become orphan nodes due to link removals, we propose different models where the orphan nodes are taken or not taken into account in calculating the outbreak size, respectively. Finally, Section 4 concludes the paper.

2   Contact Network with Fear Factor

Consider a random scale-free network with a given nodal-degree distribution $P(i)$. Assume that a fraction $(1-\rho)$ of all the links are randomly removed from this network due to the fear factor. Then each link in the network is included in the subnet for disease propagation with a probability ρ and left out with a probability $(1-\rho)$ (as shown in Fig. 1). The procedure can be viewed as a random sampling of links in the original network. Hereafter we term ρ as the *sampling probability* and the subnet as the *sampled subnet*. Apparently, the topology of the subnet is substantially different from that of the original network.
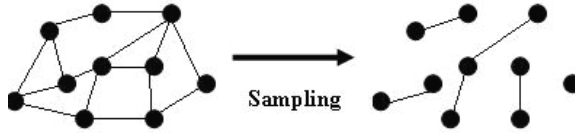
Fig. 1. Sampling on the network: each link is picked randomly with a probability ρ to be included in the subnet

To study the topological changes of a given random network after sampling, the probability-generating function (PGF) [15] is employed as a compact and conventional tool. Specifically, a random network with a degree distribution $P(i)$ can be defined in terms of its probability-generating function $G_0(x)$ [5] as follows

$$G_0(x) = \sum_{i=0}^{\infty} P(i)x^i,\tag{2}$$

where $P(i)$ can be derived from $G_0(x)$ as $P(i) = \frac{1}{i!} \frac{d^i G_0}{dx^i}\big|_{x=0}$. Note that for networks without any un-connected nodes, $P(0)$ always equals to zero.

Let the degree distribution of the sampled subnet be $P^*(k)$ (Hereafter, we use the "*" to indicate functions for the sampled subnet.). The probability that a node of degree i in the original network is connected to k other nodes $(k \le i)$ in the sampled subnet is given by the binomial formula,

$$P^*(k) = \sum_{i=k}^{\infty} P(i)\binom{i}{k}\rho^k (1-\rho)^{i-k}.\tag{3}$$

Therefore, PGF of the subnet is

$$\begin{aligned}
G_0^*(x) &= \sum_{k=0}^{\infty} P*(k)x^k \\
&= \sum_{k=0}^{\infty}\sum_{i=k}^{\infty} P(i)\binom{i}{k}\rho^k (1-\rho)^{i-k} x^k \\
&= \sum_{i=0}^{\infty} P(i)\sum_{k=0}^{i}\binom{i}{k}(\rho x)^k (1-\rho)^{i-k} \\
&= \sum_{i=0}^{\infty} P(i)(1-\rho+\rho x)^i \\
&= G_0(1+\rho(x-1)).
\end{aligned}\tag{4}$$

Equation (4) describes the degree distribution of the sampled subnet, including those orphan nodes (those nodes with degree zero). If orphan nodes are to be excluded from the subnet, PGF of the subnet would be

$$G_0^*(x) = \frac{G_0(1+\rho(x-1)) - G_0(1-\rho)}{1 - G_0(1-\rho)}. \tag{5}$$

It makes sense to remove orphan nodes since they will neither be infected by others nor infect any other nodes during the epidemic outbreak. Certainly in real life, seldom anyone will go total isolation even when in panics. Most people may choose to keep at least a few links (e.g., with their family members) despite of the risk of getting infected. Such a case however requests different analyses, and hence is out of the scope of this paper.

It is easy to see that the sampled subnet of a scale-free network does not remain as a scale-free network [16]. Similar conclusion has been drawn on random node sampling in [17]. Therefore disease propagation in the sampled subnet reveals some different features from those in the original networks, as we shall discuss below.

3   Epidemic Spreading in the Sampled Subnet

3.1   Analysis of the Classic SIR Model

Diseases spread through human populations by contacts between infective individuals (those carrying the disease) and susceptible individuals (those who are healthy yet vulnerable to the disease). Meanwhile, the infected individuals will be finally removed (either immunized or dead). Such a procedure can be concisely described by the well-known SIR model. For the case where people can be re-infected, SIS model [13]-[14] may have to be adopted which will be discussed in a separate report.

As that in [5], the generalized SIR model can be described as follows: denote the probability that the disease propagates from an infective individual i to a susceptible individual j in a unit time as r_{ij}, which can be drawn from an arbitrary distribution $P(r)$. Meanwhile the time τ for which individuals remain infected can be drawn from an arbitrary distribution $Q(\tau)$. The SIR epidemic process on a network is equivalent to the bond percolation on the same network with a uniform bond occupation probability [5], [18] - [21]

$$T = 1 - \int_0^\infty P(r)Q(\tau)e^{-r\tau}drd\tau. \tag{6}$$

T is called the transmissibility of disease, which reflects the transmission capability of the disease. The value of T ranges between 0 and 1. For different cases where T is lower or higher than the epidemic threshold (denoted as T_c) respectively, the size of infections is measured differently. Specifically, for $T < T_c$, we define the *average outbreak size* [5] as the average number of nodes ever getting infected where the disease starts with a randomly chosen initial carrier; while for $T \geq T_c$, the average relative size of the infected giant component versus the whole network is measured, termed as the *prevalence size* [5].

In a random network with a degree distribution $P(k)$, following a randomly chosen edge, the probability of reaching a node with a degree k is proportional to $kP(k)$, corresponding to the generating function

$$G_1(x) = \frac{\sum_k kP(k)x^{k-1}}{\sum_k kP(k)} = \frac{G_0'(x)}{G_0'(1)}.\tag{7}$$

Following the bond percolation theory for epidemic spreading in networks [5], [18]-[21], we can get the expressions for calculating the epidemic threshold T_c and the average outbreak size $<s>$ that

$$T_c = \frac{1}{G_1'(1)} = \frac{G_0'(1)}{G_0''(1)},\tag{8}$$

$$\langle s \rangle = 1 + \frac{TG_0'(1)}{1 - TG_1'(1)},\tag{9}$$

The prevalence size for $T \geq T_c$ is then given as

$$S = 1 - G_0(u), \quad \text{where} \quad u = 1 - T + TG_1(u).\tag{10}$$

Since it is generally not easy to get a close-form expression of the parameter u, S is normally calculated using numerical methods.

3.2 Epidemic Spreading in the Sampled Subnet with Orphan Nodes

The PGF for the sampled subnet has been shown in Eq. (4). Therefore the epidemic threshold is governed by

$$T_c^* = \frac{1}{G_1^{*'}(1)} = \frac{G_0^{*'}(1)}{G_0^{*''}(1)}.\tag{11}$$

From Eq. (4), we can easily derive that $G_0^{*'}(1) = \rho G_0'(1)$ and $G_0^{*''}(1) = \rho^2 G_0''(1)$. Hence

$$T_c^* = \frac{1}{\rho}T_c,\tag{12}$$

where T_c is the epidemic threshold of the original network. Equation (12) indicates that link removals can only slowly increase the epidemic threshold. For example, a 50% link removals merely double the (originally very low) epidemic threshold. Such observations are also revealed in numerical simulations presented in Fig. 2. Therefore, we draw the conclusion that link removals cannot easily prevent an epidemic outbreak from happening.

Consider a particular scale-free network model – an N-node BA model with a degree distribution $P(k) = 2m^2 / k^3$, an average degree $< k >= 2m$, and a maximum nodal degree $k_m = mN^{1/2}$ [22]. Applying the continuum theory, we have

$$T_c^* = \frac{1}{\rho} \frac{G_0'(1)}{G_0''(1)} = \frac{1}{\rho} \frac{\sum_k kP(k)}{\sum_k k(k-1)P(k)} = \frac{1}{\rho} \frac{\int_m^{k_m} \frac{2m^2}{k^2} dk}{\int_m^{k_m} (k-1) \frac{2m^2}{k^2} dk} \tag{13}$$

$$= \frac{1}{\rho} \frac{1}{m \ln N}$$

We compare the analytical results in Eq. (13) with numerical results. As shown in Fig. 2, there is a good match between them. A linear relationship between the inverse epidemic threshold and the sampling probability can also be clearly observed.
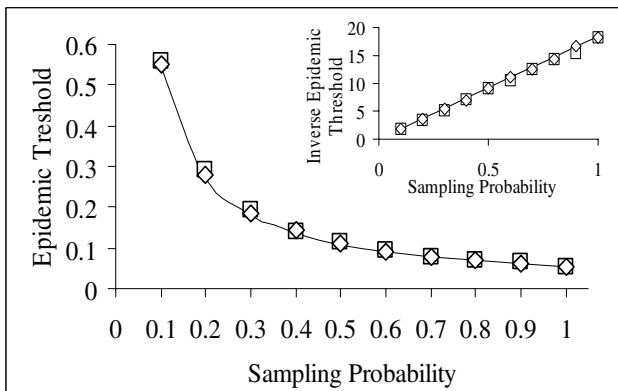

Fig. 2. Dependence of epidemic threshold T_c^* on the sampling probability ρ. The rectangle boxes represent the numerical simulation with orphan nodes while the diamond boxes indicate the results without orphan nodes. Analytical results are shown by the solid lines. The embedded sub-figure shows the linear relationship between the inverse epidemic threshold $(T_c^*)^{-1}$ and sampling probability ρ. The simulation results are based on a BA model with 10,000 nodes and an average nodal degree of $< k >= 4$.

From Eqs. (4) and (9), the average outbreak size is given by

$$\langle s^* \rangle = 1 + \frac{TG_0^{*'}(1)}{1 - TG_1^{*'}(1)} = 1 + \frac{T\rho G_0'(1)}{1 - T\rho G_0''(1) / G_0'(1)} \tag{14}$$

Hence for a N-node BA model, we have

$$< s^* >= 1 + \frac{2T\rho m}{1 - T\rho(\frac{1}{2} m \ln N - 1)} \tag{15}$$
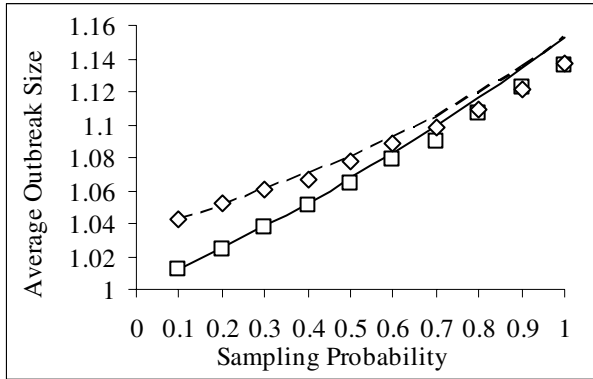
Fig. 3. The average outbreak sizes corresponding to different sampling probabilities on top of a BA model with 10,000 nodes and an average degree $<k>= 4$. $T = 0.06$. The rectangle and diamond boxes represent the numerical results while including and excluding orphan nodes, respectively. The corresponding analytical results are respectively presented by the solid and dashed lines.
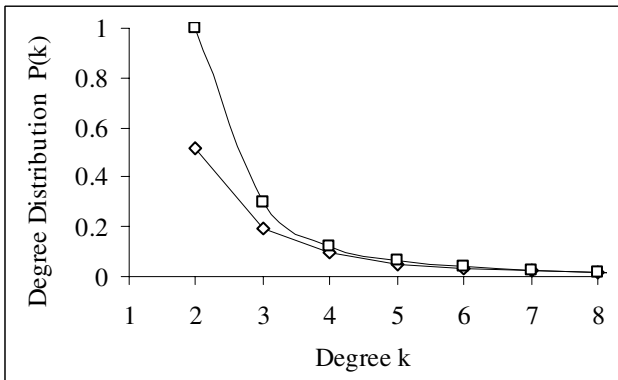
Fig. 4. The mismatches between the approximately calculated (represented by rectangle boxes) and measured $P(k)$ (represented by diamond boxes). The measured results come from a BA model with 10,000 nodes and an average degree $<k>= 4$.

Analytical and simulation results are presented in Fig. 3. We see that the average outbreak size is slightly decreased, which is not a surprise since, without an epidemic outbreak, the average number of nodes getting infected is quite low even in the original network.

The results in Fig. 3 do not show a very good match between analytical and simulation results especially under high sampling probabilities. We believe that the mismatch is mainly due to two different reasons: (i) The finite-size BA model is not strictly a random network model [23]; and more significantly, (ii) the equation $P(k) = 2m^2/k^3$ is a good approximation only when k is of a large enough value. For

example, where $m = 2$ and $k = 2$, from the above equation we shall have $P(2) = 1$, while the real value of $P(2)$ measured from the network is about 0.5194. For more details, Fig. 4 shows the mismatches between the measured and approximately calculated $P(k)$ in a 10,000-node BA model for different values of k.

When transmissibility is higher than the threshold T_c^*, Eq. (13) can be re-written as follows for calculating the prevalence size

$$S^* = 1 - G_0^*(u), \quad \text{where} \quad u = 1 - T + TG_1^*(u). \tag{16}$$

Similar to that for Eq. (10), numerical methods normally have to be adopted to solve the above equation. In the BA model, however, the approximated $P(k)$ does not always allow a feasible solution of u. Therefore only the simulation results have been presented in Fig. 5. We see that though even a relatively low sampling probability
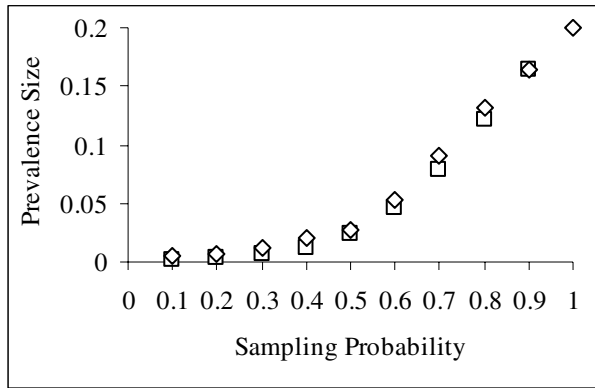
(a)
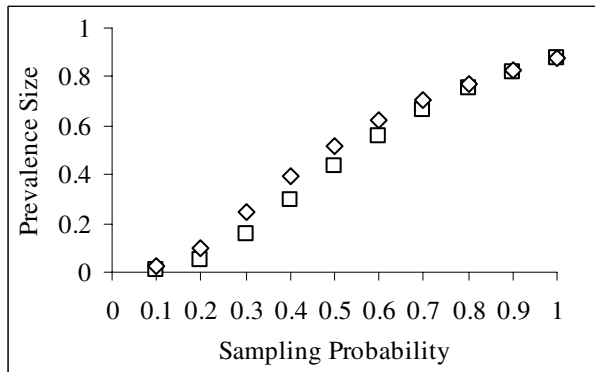
(b)

Fig. 5. The prevalence sizes corresponding to different sampling probabilities on top of a BA model with 10,000 nodes and an average degree $< k > = 4$. $T = 0.2$ in sub-figure (a) and 0.6 in sub-figure (b). The rectangle and diamond boxes represent the numerical results with and without orphan nodes, respectively.

(e.g., $\rho = 0.6$) does not lead to a significantly enhanced epidemic threshold, it does help to decrease the prevalence size significantly (by more than 3/4 from 0.20186 to 0.04639) under moderate transmissibility (when $T = 0.2$). For the case with strong transmissibility (e.g., $T = 0.6$), the effects in reducing prevalence size becomes less significant as shown in Fig. 5(b). This can be understood: under high transmissibility, reducing the number of contacts cannot easily significantly reduce the probability of being infected.

3.3 Epidemic Spreading in the Sampled Subnet Excluding the Orphan Nodes

For the sampled subnet excluding the orphan nodes, with the normalized PGF in Eq. (5) and following the same procedure as that in Section 3.2, we have the threshold as

$$T_c^* = \frac{1}{G_1^{*'}(1)} = \frac{G_0^{*'}(1)}{G_0^{*''}(1)} = \frac{1}{\rho}T_c,$$ (17)

which equals to that of the case with orphan nodes. This matches the straightforward observation that the epidemic threshold is not affected by whether we take orphan nodes into account or not.

Considering the average outbreak size for $T < T_c$, from Eq. (9), we have

$$\left\langle s^* \right\rangle = 1 + \frac{TG_0^{*'}(1)}{1 - TG_1^{*'}(1)} = 1 + \frac{T\rho G_0^{'}(1)/(1 - G_0(1 - \rho))}{1 - T\rho G_0^{''}(1)/G_0^{'}(1)}.$$ (18)

Comparing with Eq. (14), we can see Eq. (18) has one more term $1/(1 - G_0(1 - \rho))$, which suggests that by excluding the orphan nodes, the average outbreak size becomes larger. This can be easily explained: those (usually) unrealistic cases infecting only an orphan node (with an outbreak size 1) have been excluded while calculating the average outbreak size.

For a N-node BA model, the average outbreak size can be approximately expressed as

$$<s^*> = 1 + \frac{2mT\rho}{[1 - T\rho(\frac{1}{2}m\ln N - 1)](1 - \sum_{k=m}^{mN^{1/2}} \frac{2m^2}{k^3}(1 - \rho)^k)}.$$ (19)

Analytical and simulation results are presented in Fig. 3.

For $T \geq T_c^*$, Eq. (10) can be revised as follows for calculating the prevalence size

$$S^* = 1 - G_0^*(u), \quad \text{where} \quad u = 1 - T + TG_1^*(u).$$ (20)

Once again only the simulation results are presented in Fig. 5 since u is not always solvable based on the approximated $P(k)$. We see that the conclusions on the effects of link removals in reducing prevalence size remain unchanged. Moreover, by excluding the orphan nodes, the relative size of the infected giant component, i.e., the prevalence size, becomes larger.

4   Conclusions

Under the threat of a dangerous infectious disease, people tend to reduce their contacts with others in fear of getting infected. We investigated the effects of such common reactions. Our preliminary study showed that such behaviors may not efficiently prevent an outbreak from happening. However, they could be helpful in reducing the size of infected.

Studies in this paper are based on the simple assumption of random link removals. In our future research, more realistic cases will be investigated. Specifically, we will evaluate the effects of fair factors on epidemic behaviors where links connected to high-degree nodes have a higher chance to be removed than those connected to low-degree nodes: while high-degree nodes usually have a large number of dispensable links, low-degree nodes may cherish the last a few links they have, typically connecting them with their families and close friends etc., even when there is an outbreak of a dangerous epidemic disease. An (arguably) even more realistic and more difficult case may be that link removals are based on *local* information of epidemic spreading. For example, a node may tend to remove more of its links when there are more confirmed infected nodes in its neighborhood area. Such a case will also be studied in our future research.

References

1. World Heath Organization, `http://www.who.int/csr/sars/en/`
2. Gupta, S., Anderson, R.M., May, R.M.: Networks of Sexual Contacts: Implications for the Pattern of Spread of HIV. AIDS 3, 807–817 (1989)
3. Morris, M.: Sexual Networks and HIV. AIDS 97: Year in Review 11, 209–216 (1997)
4. Liljeros, F., Edling, C.R., Amaral, L.A.N., Stanley, H.E., Áberg, Y.: The Web of Human Sexual Contacts. Nature 411, 907–908 (2001)
5. Newman, M.E.J.: Spread of Epidemic Disease on Networks. Phys. Rev. E 66, 016128 (2002)
6. Moreno, Y., Pastor-Satorras, R., Vespignani, A.: Epidemic Outbreaks in Complex Heterogeneous Networks. Eur. Phys. J. B 26, 521–529 (2002)
7. Boguñá, M., Pastor-Satorras, R.: Epidemic Spreading in Correlated Complex Networks. Phy. Rev. E 66, 047104 (2002)
8. Barabási, A.-L., Albert, R.: Emergence of Scaling in Random Networks. Science 286, 509–512 (1999)
9. Bornholdt, S., Schuster, H.G. (eds.): Handbook of Graphs and Networks. Wiley-VCH, Berlin (2003)
10. Barabási, A.-L., Albert, R.: Statistical Mechanics of Complex Networks. Rev. Mod. Phys. 74, 47–98 (2002)
11. Pastor-Satorras, R., Vespignani, A.: Epidemic Spreading in Scale-free Networks. Phys. Rev. Lett. 86, 3200 (2001)
12. Pastor-Satorras, R., Vespignani, A.: Epidemic Dynamics in Finite Size Scale-free Networks. Phys. Rev. E 65, 035108 (2002)

13. Anderson, R.M., May, R.M.: Infectious Diseases of Humans. Oxford University Press, Oxford (1991)
14. Hethcote, H.W.: The Mathematics of Infectious Diseases. SIAM Review 42, 599–653 (2000)
15. Herbert, S.W.: Generatingfunctionology. Academic Press, Harcourt Brace Jovanovich (1991)
16. Martin, S., Carr, R.D., Faulon, J.-L.: Random Removal of Edges from Scale-free Graphs. Physica A 371(2), 870–876 (2006)
17. Stumpf, M.P.H., Wiuf, C., May, R.M.: Subnets of Scale-free Networks Are Not Scale-free: Sampling Properties of Networks. PNAS 102(12) (2005)
18. Newman, M.E.J., Strogatz, S.H., Watts, D.J.: Random Graphs with Arbitrary Degree Distributions and Their Applications. Physical Review E 64 (2001)
19. Warren, C.P., Sander, L.M., Sokolov, I.M.: Firewalls, Disorder, and Percolation in Epidemics. cond-mat/0106450 v1 (2001)
20. Sander, L.M., Warren, C.P., Sokolov, I., Simon, C., Koopman, J.: Percolation on Disordered Networks as a Model for Epidemics. Math. Biosci. 180, 293–305 (2002)
21. Moore, C., Newman, M.E.J.: Exact Solution of Site and Bond Percolation on Small-World Networks. Phy. Rev. E 62, 7059–7064 (2000)
22. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of Networks. Adv. Phys. 51, 1079 (2002)
23. Newman, M.E.J.: Assortative Mixing in Networks. Phys. Rev. Lett. 89, 208701 (2002)