Immunization strategy based on the critical node in percolation transition

Yang Liu[a], Bo Wei[b], Zhen Wang[c], Yong Deng[a,d,*]

[a]*School of Computer and Information Science, Southwest University, Chongqing 400715, China*
[b]*Institute of Intelligent Control and systems, Harbin Institute of Technology, Harbin 150080, China*
[c]*Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Kasuga-koen, Kasuga-shi, Fukuoka 816-8580, Japan*
[d]*School of Automation, Northwestern Polytechnical University, Xian, Shaanxi, 710072, China*

Abstract

The problem of finding a better immunization strategy for controlling the spreading of the epidemic with limited resources has attracted much attention since its great theoretical significance and wide application. In this paper, we propose a novel and successful targeted immunization strategy based on percolation transition. Our strategy immunizes the fraction of critical nodes which lead to the emergence of giant connected component. To test the effectiveness of the proposed method, we conduct the experiments on several artificial networks and real-world networks. The results show that the proposed method outperforms the existing well-known methods with 18% to 50% fewer immunized nodes for same degree of immunization.

Keywords: Immunization strategy; Percolation transition; Emergence; Critical nodes.

*Corresponding author: Yong Deng, School of Computer and Information Science, Southwest University, Chongqing, 400715, China. Email address: prof.deng@hotmail.com; yongdeng@nwpu.edu.cn

1. Introduction

Epidemic spreading and controlling in complex networks has attracted much attention from various research fields [1, 2, 3, 4, 5, 6]. In the real world, it is ubiquitous that there are only a few infection sources in the early stage of the spreading process. However, it is possible that these infected sources spread the epidemic to large area in the world within a few days if the epidemic is ineffectively controlled [7, 8], such as the outbreak of severe acute respiratory syndrome (SARS) and swine flu [9]. Crucial method of controlling the epidemic spreading is to apply immunization measure, which refers to assigning efficient protecting strategy to the most key individuals individuals with limited resources. A feasibility and effective immunization strategy can save thousands of human lives with low cost. Over the past decades, many immunization strategies have been extensively investigated [10, 11, 12]. So far, it is still an open issue.

In the complex networks, numerous strategies have been proposed to immunize nodes, such as random immunization strategy [13], target immunization strategy [14, 15], acquaintance immunization strategy [16, 17], active immunization strategy [18] and other related scenarios [16, 19, 20, 21]. It is widely accepted that the most efficient immunization strategies are basted on targeted strategies [22, 23, 24]. The target immunization is first to identify the potential constrain the spreading ability of each node and then removes the nodes from the highest ability to the lowest until the network reaches the certain immunization fraction.

In this paper, motivated by Refs. [25, 26], we proposed a novel immunization strategy. Based on percolation transition [26], the main idea of the proposed method is to find the critical nodes, which lead to the emergence of the giant connected component, for immunizing. To test the performance of our proposed strategy, we conduct the experiments on Erdős-Rényi (ER) networks [27], scale-free (SF) networks [28] and also several real-world networks. The results show that the proposed method has 18% to 50% improvement than the degree centrality strategy, the betweenness centrality strategy and the adaptive degree centrality strategy (for the networks studied here).

The rest of the paper is organized as follows: in Section 2, the proposed

strategy is described; in Section 3, related experiments are conducted and experimental results are evaluated and conclusions are presented in Section 4.

2. Method

Consider a network $G(N, E)$ where N and E are the set of nodes and edges, respectively. The notation d_i denotes the degree of node i. For a given degree threshold d_t, we classify the nodes into two categories: visible nodes if $d_i \leq d_t$ and invisible nodes if $d_i > d_t$,

$$x_i = \begin{cases} 1 & d_i \leq d_t; \\ 0 & d_i > d_t. \end{cases} \tag{1}$$

$$e_{ij} = \begin{cases} 1 & x_i = 1 \quad and \quad x_j = 1, (i,j) \in E; \\ 0 & x_i = 0 \quad or \quad x_j = 0, (i,j) \in E. \end{cases} \tag{2}$$

where $x_i, e_{ij} \in \{0, 1\}$ are binary variables. x_i equal to 1 if node i is visible and 0 otherwise. Let $G^*(d_t)$ be the subnetwork of G, which is constructed by the visible nodes and edges of G under a given threshold d_t. Let $C(d_t)$ denote the number of nodes of the connected components and $C_{max}(d_t), C_{max-1}(d_t), ...$ correspond to the largest, the second-largest connected components and so forth. In this way, the function $G^*(d_t)$ of d_t is constructed, and the critical node can be determined based on the emergence of giant connected components as the value of d_t increases (Fig. 1 and Fig. 2).

To illustrate how to determine the critical node, the network of coauthorships between scientists [29] is considered. In order to clearly understand, we obtain the experimental network (Fig. 1(a)) by extracting the largest component and deleting $\{51, 52\}$ nodes, and then removing the little isolated components. $C(d_t)$ increases as d_t increases (Fig. 1(b) to Fig. 1(c)). In Fig. 1(d), the critical node 223 becomes visible and connects the largest and the second-largest connected components when $d_t = 14$. The details of C_{max} and C_{max-1} of d_t are showed in Fig. 2 where the critical node is that connecting the largest connected component with the maximal second-largest connected component. We also obtain the other critical nodes until

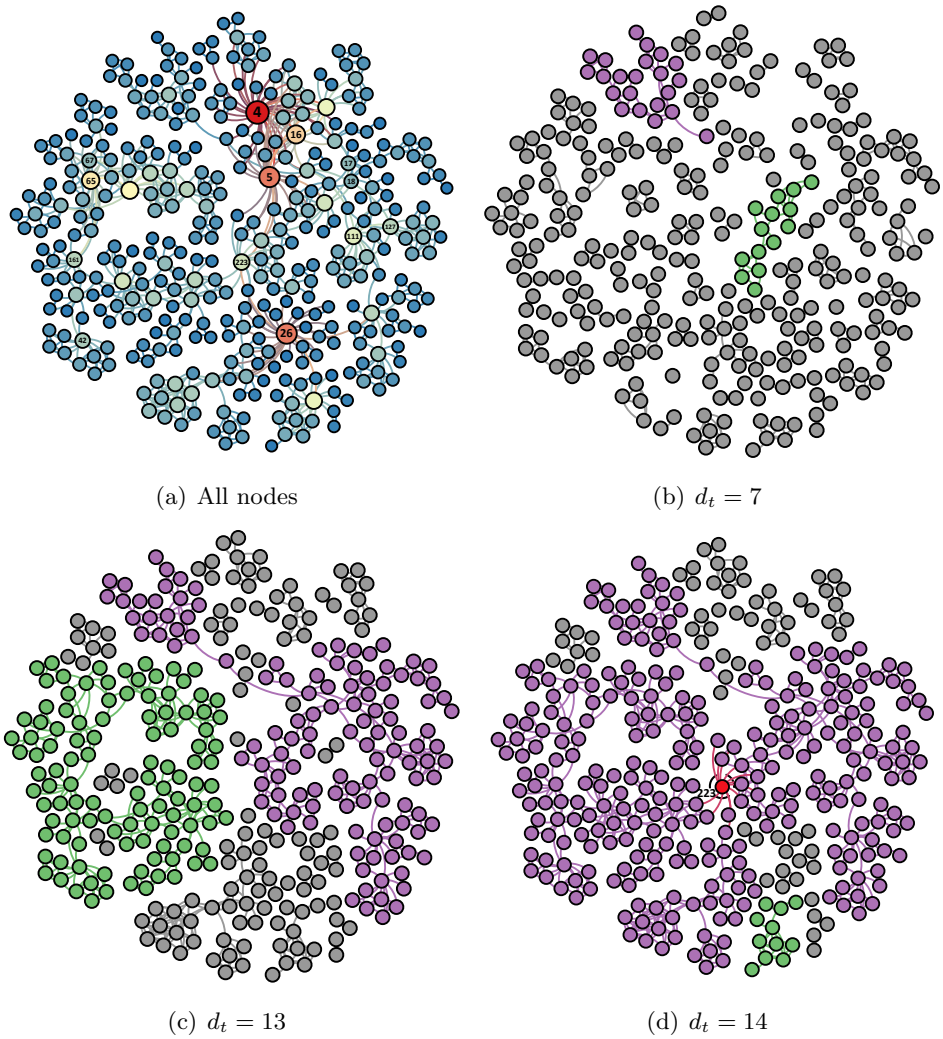
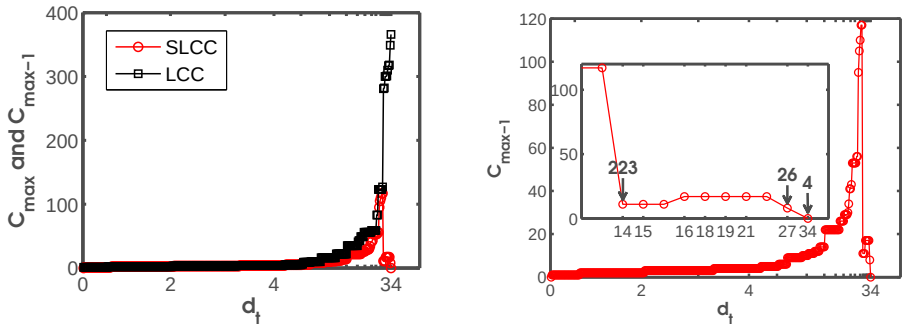


Figure 1: Critical node of the network. (a) A representation of all nodes. The size of a node is proportional to its degree: node 4 holds the largest degree. (b) A representation of visible nodes when $d_t = 7$. The largest and second-largest connected component are clarified by purple and green, respectively. (c) A representation of visible nodes when $d_t = 13$. (d) The critical node (red) is obtained when $d_t = 14$.



(a) (b)

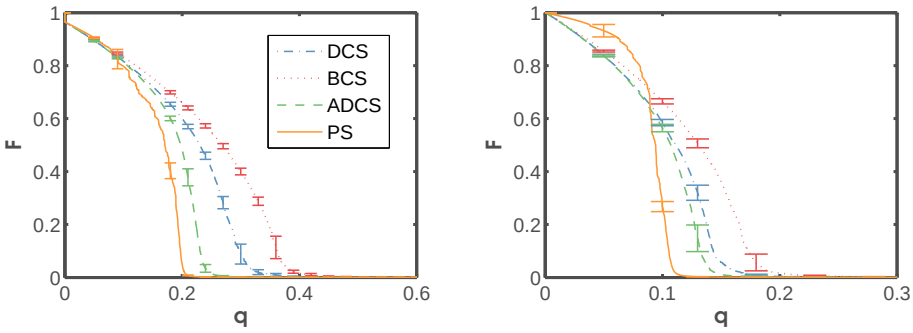
Figure 2: The number of nodes of the largest (LCC) and second-largest (SLCC) connected component C_{max} (black square) and C_{max-1} (red circle) versus the degree threshold d_t. Each circle or square represents that a node becomes visible. The critical node (node 223 but not node 26 and node 4) is the node which connects LCC with the maximal SLCC.

to reach the predetermine by repeating this process on the largest connected component of the network after the critical node removal.

3. Results

To test the effectiveness of the proposed strategy (PS), we plot F (the fraction of the size of the largest cluster that can be infected) versus q (the fraction of immunized nodes) for several artificial networks (including Erdős-Rényi (ER) networks [27] and scale-free(SF) networks [28] in Fig. 3) and empirical networks from different fields in Fig. 4: the electrical power grid of the western United States (Power Grid) [30, 31], a trust network of mutual signing based on the pretty good privacy algorithm (PGP network)[32, 33], collaboration network of arxiv condensed matter (ca-CondMat) [34, 35], and email communication network from Enron (email-Enron) [36, 37, 38]. The detailed information of the networks is showed in Table 1. Note that we here remove 58 self-loop edges from the source data of ca-CondMat.

We compare the efficiency of the PS to the degree centrality strategy (DCS), the betweenness centrality strategy (BCS) and the adaptive degree centrality strategy (ADCS). In the ADCS we immunize a node by recalculating the importance of every node after the immunized node removal.



(a) ER (b) BA

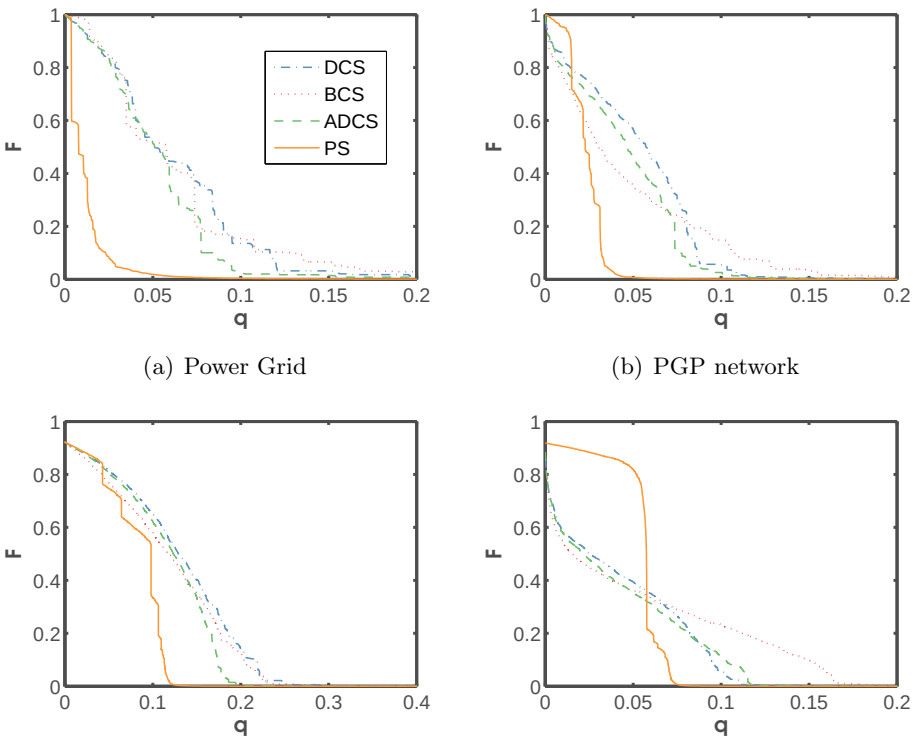
Figure 3: The fraction F of the size of the largest cluster that can be infected versus the fraction of immunized nodes q for DCS, BCS, ADCS and PS for (a) ER network with $N = 10^4$ and $< k >= 3.50$, (b) BA network with $N = 10^4$, degree exponent $\alpha = 2.6$ and $< k >= 4.0$. We also show the error bars in F, which are derived from simulating realization.

Table 1: Some statistical properties of the real-world networks: network size (n), edge number (m), average degree $< k >$, maximum degree k_{max}, degree heterogeneity ($H =< k^2 > / < k >^2$), degree assortativity (r), clustering coefficient ($< C >$).

Network	n	m	$< k >$	k_{max}	H	r	$< C >$
Power Grid	4941	6594	2.6691	19	1.4504	0.0035	0.0801
PGP	10680	24316	4.5536	205	4.1465	0.2382	0.2659
ca-CondMat	23133	93439	8.0784	279	2.7305	0.1340	0.6334
email-Enron	36692	183831	10.0202	1383	13.9796	-0.1108	0.4970

For both of ER and BA networks, the PS exhibits notable advantage of less nodes to be immunized compared to the other strategies. Regarding to threshold point q_c, where F approaches 0 ($F < 0.0005$), the PS shows over 50% improvement than the DCS and BCS, 18% in ER network and 26% in BA network than ADCS. For the real-world networks, the PS also shows a larger improvement, over 50%, against both DCS and BCS, and gives an advantage of over 30% compared to ADCS.

To further evaluate the performance of the proposed strategy, the susceptible-



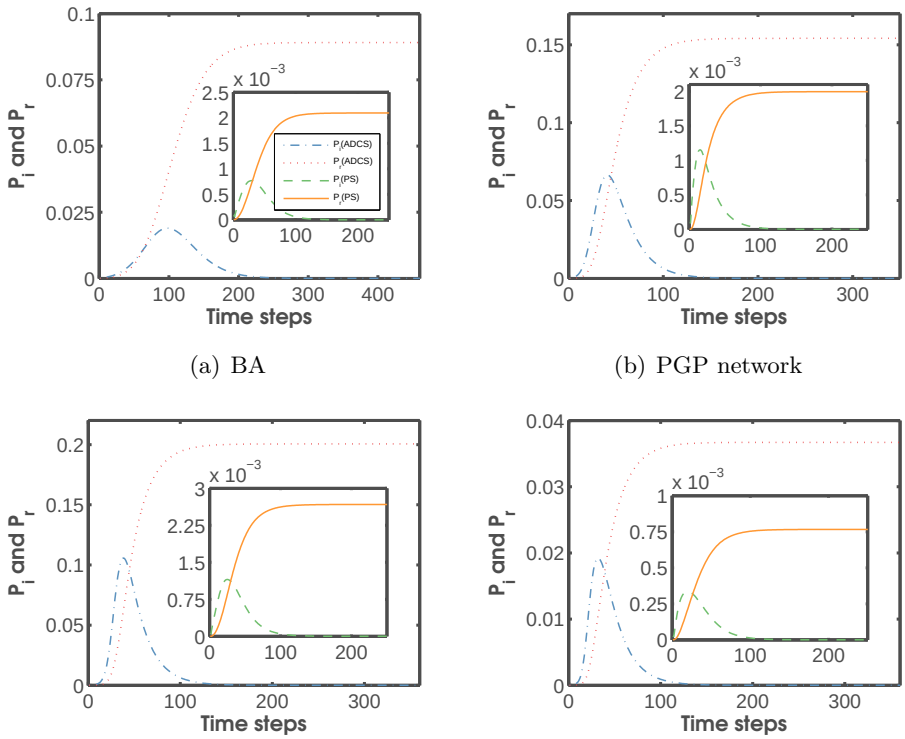
(a) Power Grid

(b) PGP network

(c) ca-CondMat

(d) email-Enron

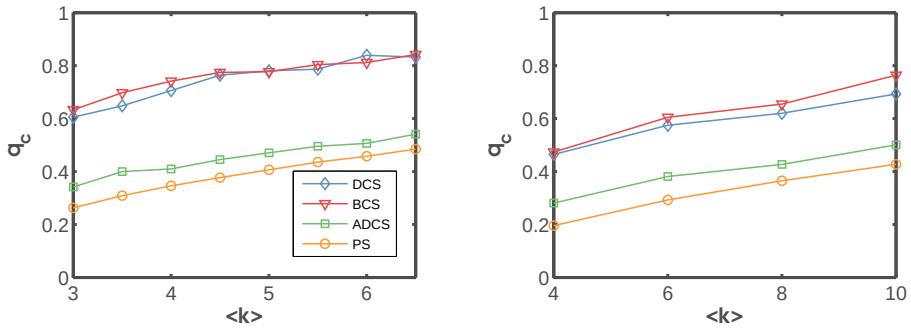
Figure 4: The fraction F of the size of the largest cluster versus the fraction of immunized nodes q for DCS, BCS, ADCS and PS for (a) the Power Grid network, (b) the PGP network, (c) the ca-CondMat network, (d) the email-Enron network.



(a) BA (b) PGP network

(c) ca-CondMat (d) email-Enron

Figure 5: Infected fraction P_i and Recovered fraction P_r of time steps for the SIR simulation. Comparison between the PS and ADCS for (a) immunizing a fraction $q = 0.11$ of the nodes in the BA network with $\alpha = 2.6$, $N = 10^4$ and $< k >= 4$, (b) immunizing a fraction $q = 0.045$ of the nodes in the PGP network, (c) immunizing a fraction $q = 0.12$ of the nodes in the ca-CondMat network, (d) immunizing a fraction $q = 0.075$ of the nodes in the email-Enron network. 10^4 independent simulation for each network.



(a) ER (b) BA

Figure 6: The threshold q_c of average mode degree $< k >$ for the DCS, BCS, ADCS and PS for (a) ER networks with $N = 10^4$, (b) BA networks with $N = 10^4$ and $\alpha = 2.6$.

infectious-recovery (SIR) epidemic spreading model [39, 40, 41, 42] is used to evaluate the immunization effectiveness on the BA network, the PGP network, the ca-CondMat network and the email-Enron network. In the SIR model, each node belongs to one state of the susceptible state, the infected state and the recovered state. At the initial stage, the immunized nodes are removed from the network, including the incident links. We set one node that is randomly selected from the remaining network or networks (may be some disconnected subnetworks after immunized) to be infected to investigate the transmissibility of this node, and the others to be susceptible. At each time step, the infected nodes infect its susceptible neighbors with infection probability λ, and they recover with probability η. The recovered nodes are removed from the network. This process is repeated until there is no infected node in the network. The simulation results are shown in Fig. 5. 10^4 independent simulation and $\lambda = 0.4$ and $\eta = 0.05$ are used for each network. For all of the networks studied here, the recovered fraction is significantly (42 to 77 times) lower when using the PS compared to the ADCS with the same fraction of immunization doses.

Finally, in Fig. 6, we investigate the threshold q_c of the network parameters of ER and BA networks for different immunization strategies. The threshold q_c is defined as the fraction of nodes immunized or removed for

which F approaches 0 ($F < 0.0005$). In all networks tested here, the proposed strategy exhibits obviously effectiveness of lower q_c compared to the other immunization strategies. Although q_c increases with increasing the average degree $< k >$, the curves of PS are far below curves of DCS, BCS and ADCS.

4. Conclusion

In this paper, we have developed and applied a novel method as an efficient network immunization strategy based on the percolation transition. The proposed strategy chooses the immunization fraction of nodes by repeatedly looking for the critical node, which leads to the emergence of the giant connected component as the degree threshold increases, in the largest connected network. To test the performance of the proposed method, we conduct the experiments on several artificial and real-world networks.

The results show that the proposed method is more effective in assortatively mixed networks where high degree nodes tend to connect to other high degree nodes, which is a common feature for many real networks [16, 43]. In the disassortative networks, high degree nodes mostly have neighbors with a small number of connections. Our strategy identifies the critical node only based on the largest and the second-largest connected components. It is expected that the proposed strategy will immunize or remove some low degree nodes, which is the critical node of two components, at the early stage. However, those components are also connected by some high degree nodes. The immunization threshold q_c is fast reached after those low degree nodes removal. To summarize, our strategy holds the advantage of 18% to 50% compared to the degree centrality strategy, the betweenness centrality strategy and the adaptive degree centrality strategy.

Acknowledgments

The work is partially supported by National Natural Science Foundation of China (Grant No. 61174022), Specialized Research Fund for the Doctoral Program of Higher Education (Grant No. 20131102130002), R&D Program

of China (2012BAH07B01), National High Technology Research and Development Program of China (863 Program) (Grant No. 2013AA013801), the open funding project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No.BUAA-VR-14KF-02).

References

[1] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Physical review letters 86 (2001) 3200.

[2] A. Saumell-Mendiola, M. Á. Serrano, M. Boguñá, Epidemic spreading on interconnected networks, Physical Review E 86 (2012) 026106.

[3] J. Zhou, G. Xiao, S. A. Cheong, X. Fu, L. Wong, S. Ma, T. H. Cheng, Epidemic reemergence in adaptive complex networks, Physical Review E 85 (2012) 036107.

[4] Z. Wang, A. Szolnoki, M. Perc, Interdependent network reciprocity in evolutionary games, Scientific reports 3 (2013).

[5] Z. Wang, A. Szolnoki, M. Perc, Optimal interdependence between networks for the evolution of cooperation, Scientific reports 3 (2013).

[6] C. Buono, L. G. Alvarez-Zuzek, P. A. Macri, L. A. Braunstein, Epidemics in partially overlapped multiplex networks, PloS one 9 (2014) e92200.

[7] V. Colizza, A. Barrat, M. Barthelemy, A.-J. Valleron, A. Vespignani, Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions, PLoS medicine 4 (2007) e13.

[8] N. Perra, B. Gonçalves, R. Pastor-Satorras, A. Vespignani, Activity driven modeling of time varying networks, Scientific reports 2 (2012).

[9] C. M. Schneider, T. Mihaljev, S. Havlin, H. J. Herrmann, Suppressing epidemics with a limited amount of immunization units, Physical Review E 84 (2011) 061911.

[10] C. T. Bauch, D. J. Earn, Vaccination and the theory of games, Proceedings of the National Academy of Sciences of the United States of America 101 (2004) 13391–13394.

[11] L. Hébert-Dufresne, A. Allard, J.-G. Young, L. J. Dubé, Global efficiency of local immunization on complex networks, Scientific reports 3 (2013).

[12] A. Vespignani, Modelling dynamical processes in complex socio-technical systems, Nature Physics 8 (2012) 32–39.

[13] R. Pastor-Satorras, A. Vespignani, Immunization of complex networks, Physical Review E 65 (2002) 036104.

[14] C. M. Schneider, T. Mihaljev, H. J. Herrmann, Inverse targetingan effective immunization strategy, EPL (Europhysics Letters) 98 (2012) 46002.

[15] B. Mirzasoleiman, M. Babaei, M. Jalili, Immunizing complex networks with limited budget, EPL (Europhysics Letters) 98 (2012) 38004.

[16] R. Cohen, S. Havlin, D. Ben-Avraham, Efficient immunization strategies for computer networks and populations, Physical Review Letters 91 (2003) 247901.

[17] L. K. Gallos, F. Liljeros, P. Argyrakis, A. Bunde, S. Havlin, Improving immunization strategies, Physical Review E 75 (2007) 045104.

[18] X. Fu, M. Small, D. M. Walker, H. Zhang, Epidemic dynamics on scale-free networks with piecewise linear infectivity and immunization, Physical Review E 77 (2008) 036113.

[19] L. Lü, Y.-C. Zhang, C. H. Yeung, T. Zhou, Leaders in social networks, the delicious case, PloS one 6 (2011) e21202.

[20] K. Gong, M. Tang, P. M. Hui, H. F. Zhang, D. Younghae, Y.-C. Lai, An efficient immunization strategy for community networks, PloS one 8 (2013) e83489.

[21] W. Wang, M. Tang, H.-F. Zhang, H. Gao, Y. Do, Z.-H. Liu, Epidemic spreading on complex networks with general degree and weight distributions, Physical Review E 90 (2014) 042803.

[22] P. Holme, B. J. Kim, C. N. Yoon, S. K. Han, Attack vulnerability of complex networks, Physical Review E 65 (2002) 056109.

[23] L. K. Gallos, R. Cohen, P. Argyrakis, A. Bunde, S. Havlin, Stability and topology of scale-free networks under attack and defense strategies, Physical Review Letters 94 (2005) 188701.

[24] Y. Chen, G. Paul, S. Havlin, F. Liljeros, H. E. Stanley, Finding a better immunization strategy, Physical Review Letters 101 (2008) 058701.

[25] P. Erdős, A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hungar. Acad. Sci 5 (1960) 17–61.

[26] D. Li, B. Fu, Y. Wang, G. Lu, Y. Berezin, H. E. Stanley, S. Havlin, Percolation transition in dynamical traffic network with evolving critical bottlenecks, Proceedings of the National Academy of Sciences 112 (2015) 669–672.

[27] P. Erdős, A. Rényi, On random graphs i., Publ. Math. Debrecen 6 (1959) 290–297.

[28] A.-L. Barabási, R. Albert, H. Jeong, Mean-field theory for scale-free random networks, Physica A: Statistical Mechanics and its Applications 272 (1999) 173–187.

[29] M. E. Newman, Finding community structure in networks using the eigenvectors of matrices, Physical Review E 74 (2006) 036104.

[30] D. J. Watts, S. H. Strogatz, Collective dynamics of small-worldnetworks, nature 393 (1998) 440–442.

[31] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, science 286 (1999) 509–512.

[32] J. Duch, A. Arenas, Community detection in complex networks using extremal optimization, Physical Review E 72 (2005) 027104.

[33] Z. Lü, W. Huang, Iterated tabu search for identifying community structure in complex networks, Physical Review E 80 (2009) 026130.

[34] J. Leskovec, J. Kleinberg, C. Faloutsos, Graph evolution: Densification and shrinking diameters, ACM Transactions on Knowledge Discovery from Data (TKDD) 1 (2007) 2.

[35] `http://snap.stanford.edu/data/ca-CondMat.html`, 2015.

[36] B. Klimt, Y. Yang, Introducing the enron corpus., in: CEAS.

[37] J. Leskovec, K. J. Lang, A. Dasgupta, M. W. Mahoney, Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters, Internet Mathematics 6 (2009) 29–123.

[38] `http://snap.stanford.edu/data/email-Enron.html`, 2015.

[39] R. Albert, H. Jeong, A.-L. Barabási, Error and attack tolerance of complex networks, Nature 406 (2000) 378–382.

[40] M. E. Newman, Spread of epidemic disease on networks, Physical Review E 66 (2002) 016128.

[41] C. Castellano, R. Pastor-Satorras, Thresholds for epidemic spreading in networks, Physical Review Letters 105 (2010) 218701.

[42] C. Xia, L. Wang, S. Sun, J. Wang, An sir model with infection delay and propagation vector in complex networks, Nonlinear Dynamics 69 (2012) 927–934.

[43] D. V. Foster, J. G. Foster, P. Grassberger, M. Paczuski, Clustering drives assortativity and community structure in ensembles of networks, Physical Review E 84 (2011) 066117.