Path diversity improves the identification of influential spreaders

Duan-Bing Chen¹,², Rui Xiao², An Zeng²(α) and Yi-Cheng Zhang¹,²

¹ Web Sciences Center, School of Computer Science and Engineering, University of Electronic Science and Technology of China - Xiyuan Avenue 2006, Chengdu 611731, PRC
² Department of Physics, University of Fribourg - Chemin du Musée 3, CH-1700 Fribourg, Switzerland

received 31 May 2013; accepted in final form 17 December 2013
published online 24 January 2014

PACS 89.75.-k – Complex systems
PACS 89.20.Ff – Computer science and technology
PACS 89.65.Ef – Social organizations; anthropology

Abstract – Identifying influential spreaders in complex networks is a crucial problem which relates to wide applications. Many methods based on the global information such as K-shell and PageRank have been applied to rank spreaders. However, most of the related previous works overwhelmingly focus on the number of paths for propagation, while whether the paths are diverse enough is usually overlooked. Generally, the spreading ability of a node might not be strong if its propagation depends on one or two paths while the other paths are dead ends. In this letter, we introduced the concept of path diversity and find that it can largely improve the ranking accuracy. We further propose a local method combining the information of path number and path diversity to identify influential nodes in complex networks. This method is shown to outperform many well-known methods in both undirected and directed networks. Moreover, the efficiency of our method makes it possible to apply it to very large systems.

Copyright © EPLA, 2013

Introduction. – How to identify influential nodes in complex networks is a crucial issue since it is highly related to the information spreading and epidemic controlling [1–4]. So far, a number of centrality indices have been proposed to address this problem such as degree, betweenness [5,6], closeness [7,8] and eigenvector centralities [9]. Among these indices, degree centrality is a very straightforward and efficient one. However, the performance of degree centralities is not satisfying enough. Recently, Kitsak et al. [1] claimed that the location of a node in the network actually plays a more important role than its degree. They accordingly proposed a coarse-grained method by using K-shell decomposition to quantify a node’s influence based on the assumption that nodes in the same shell have similar influence and nodes in higher shells are likely to infect more nodes. After this, some methods are proposed to further improve the ranking performance of this network decomposition process [10,11]. In directed networks, the ranking methods are mainly based on the iterative process. The representative methods include the well-known HITs [12] and PageRank [13], as well as some recently proposed algorithms like LeaderRank [14] and TwitterRank [15]. It has been demonstrated that these methods outperform out-degree centrality in terms of ranking effectiveness.

With big-data era coming, the design of ranking algorithms on very large-scale social networks is becoming a big challenge nowadays [16]. The online social systems can have millions of users or even more. The spreader ranking algorithms will be very time-consuming if they are based on the global information of the network. Therefore, the spreader ranking algorithm should be not only effective but also efficient. To solve this problem, it is better to design the ranking algorithm based on the local information of the network. For example, a semi-local index by considering the second nearest neighbors is devised [17]. This index is shown to well identify influential nodes and obtain a good trade-off on effectiveness and efficiency compared with global indices.

Most of the previous ranking methods are designed based on the number of paths for propagation. Actually, the diversity of paths for spreading is also very important. The spreading ability of a node will be significantly lowered if many of its propagation paths overlap. In this case, if the virus/information fails to go through the overlapped path, the following spreading of many paths will be terminated. However, this factor has not been taken

(α) E-mail: an.zeng@unifr.ch
into account in designing the spreading ranking algorithm so far, to the best of our knowledge.

Accordingly, we introduced in this letter the concept of path diversity [18] which is mathematically characterized by the information entropy [19,20]. If the spreading paths of a node are not highly overlapped in a small number of links, we consider its spreading paths to be diversified. Actually, the concept of diversity has been considered in ranking spreaders before [21,22]. The diversity is defined as the dispersion of the distribution of neighboring nodes in the network. More precisely, they assume the spreading ability of a node to be high if its neighboring nodes belong to different communities. Even though a node has high diversity in this aspect, its spreading path can still highly overlap in several links (i.e. with low path diversity).

After applying the path diversity to a very simple spreader ranking method, we find that the performance of this ranking method can be further enhanced. Combining the information of path number and diversity, we further propose a local but effective ranking method (called KED method) to identify influential nodes in large-scale social networks. We make use of the SIR spreading model [23] with tunable infectivity [24–26] to test the effectiveness of our method in four real social networks, including two undirected networks, Youtube and Orkut [27], and two directed networks, EmailEU [28] and Digg [29]. Experimental results show that KED performs much better than K-shell, Duan-Bing Chen et al., the simplest degree centrality, PageRank and LeaderRank, mental results show that KED performs much better than directed networks, EmailEU [28] and Digg [29]. Experimental results show that KED performs much better than K-shell, Duan-Bing Chen et al., the simplest degree centrality, PageRank and LeaderRank, mental results show that KED performs much better than directed networks, EmailEU [28] and Digg [29].

In order to compute the probability of the gray nodes getting infected, we consider the SIR model [23] with infection rate $\mu$ in this toy network. It has been pointed out that in real case an individual only contact a fraction of his/her friends each time [24–26]. For example, an active user with many friends in the online social system for sure will interact with more friends than the inactive user with much fewer friends. However, it is not likely that such active user will interact with all his/her friends at the same time. We assume that the contact fraction nonlinearly correlates with the nodes’ degree. More specifically, each node $i$ contacts $\sqrt{k_i}$ neighbors in each step (i.e. only $\sqrt{k_i}$ neighbors have the possibility $\mu$ to be infected by node $i$ in each step). Actually, as long as the contact number $N_c$ is growing slower than the degree (e.g. $N_c \sim k_i^{\alpha}$ and $\alpha < 1$), our following analysis is valid. In fig. 1, if we set the red nodes as the initial infected nodes, the expected number of infected nodes in fig. 1(a) and fig. 1(b) at the end will be $1 + \sqrt{5\mu} + \sqrt{15\mu^2}$ and $1 + \sqrt{5\mu} + \sqrt{3\mu^2}$, respectively. Apparently, the red node in fig. 1(a) can infect more nodes than that in fig. 1(b) simply due to the diverse paths.

In fact, the diversity of local paths can be represented by the degree evenness of the target nodes’ neighbors. One can easily see that the degree of the red node’s neighbors is more uneven in fig. 1(b) than that in fig. 1(a). To characterize such unevenness, we employ the information entropy [19] $h_i$ of each node $i$ in a network as

$$h_i = \sum_{j \in \Gamma_i} -p_j \log(p_j),$$

where $p_j = \frac{k_j}{\sum_{i \in \Gamma_i} k_i}$, $\Gamma_i$ is the set of neighbors of node $i$, and $k_j$ is the degree of node $j$. A high value of $h_i$ indicates that the degree of neighbors of the target node is even, which corresponds to diverse paths for the target node to propagate information. However, there is a shortcoming in eq. (1): the entropy of two nodes with different degree will not be the same, even though the degrees of their neighbors are entirely even. Therefore, the entropy defined in eq. (1) should be normalized according to the target

Fig. 1: (Color online) The example of two toy networks. Red nodes in (a) and (b) have the same degree and the same number of second neighbors, while the distributions of their neighbors’ degree are different.
node’s degree to overcome this drawback,
\[ H_i = \frac{\sum_{j \in G_i} -p_j \log(p_j)}{\sum_{j \in G_i} -p_j \log(p_j)} = \frac{\sum_{j \in G_i} -p_j \log(p_j)}{\log(k_i)}. \] (2)

After normalization, \( H_i \in [0, 1] \), independently of the target node’s degree. To calculate \( H_i \) in the directed network, one can simply replace \( k_i \) by \( k_i^\text{out} \) in the above formulae.

Four real social networks including Youtube [27], Orkut [27], EmailEU [28] and Digg [29] are used to empirically investigate the information entropy distribution of nodes. Among these networks, Youtube and Orkut are undirected networks, EmailEU and Digg are directed ones. Youtube is a video-sharing website that includes a social network. Orkut is a free online social network where users make friends with each other. EmailEU is generated by using email data from a large European research institution from October 2003 to May 2005. Given a set of email messages, each node corresponds to an email address and a directed edge indicates that one node sent at least one email to the other. Digg contains data about stories promoted to Digg’s front page over a period of a month in 2009. The data contains the voters’ friendship networks where a link means that one node is watching the activities of (is a fan of) the other node. Some basic statistical properties of these networks are listed in table 1.

For each network, we get the counterpart random networks according to the link swap method [30]. At each step, we randomly select a pair of edges \( A-B \) and \( C-D \). These two edges are then rewired to be \( A-D \) and \( B-C \). We make sure no multiple edges or self-loop is created in this rewiring process. The distribution of \( H \) in the original network and the counterpart random networks are compared, as shown in fig. 2. For simplicity, we only take into account the 1000 highest-degree nodes since they are more likely to have strong spreading ability than those with smaller degree. In fig. 2, one can see that the information entropy distribution of nodes in the original networks is much wider than that in the randomized ones, especially in the Youtube and EmailEU networks. For some nodes in real network, their neighbors’ degree can be very uneven. Therefore, the degree unevenness of the neighboring nodes should not be neglected when ranking the spreaders.

Many researches already show that the degree centrality is not enough to accurately identify influential nodes [1,17]. The ranking methods considering also the average degree of the neighboring nodes can effectively improve the pure degree method in ranking spreaders [17].

Table 1: Basic information of four real networks. \( N \) is the network size, \( k_{\text{max}} \) is the highest degree, \( \langle k \rangle \) is the average degree and \( \langle c \rangle \) is the clustering coefficient.

| Network | \( N \)  | \( k_{\text{max}} \)  | \( \langle k \rangle \)  | \( \langle c \rangle \) |
|---------|---------|----------------|----------------|----------------|
| Youtube | 1,134,890 | 28,754 | 5.2650 | 0.1723 |
| Orkut   | 3,072,441 | 33,313 | 76.2814 | 0.1698 |
| EmailEU | 265,214  | 929  | 1.5838 | 0.3003 |
| Digg    | 279,630  | 12,097 | 9.3623 | 0.0775 |

Fig. 2: (Color online) The entropy distribution in four social networks and their corresponding random ones.

After taking into account the neighboring nodes’ average degree, the score for each target node can be
\[ f_i = k_i + \lambda \frac{\sum_{j \in G_i} k_j}{k_i}, \] (3)
where \( \lambda \) is a tunable parameter.

We employ the Susceptible-Infected-Recovered (SIR) model [23] to simulate the spreading process on networks. The simulation runs in discrete time steps. In a social network, a user neither contact all of her neighbors nor only a single neighbor, but part of her neighbors [24–26]. Moreover, the node with larger degree may contact more neighbors. Therefore, at each time step in our simulation, every infected node \( i \) will select each of its neighbors (or followers) with probability \( p = \frac{1}{k_i} \) and then transmit the information to it with probability \( \mu \) if the selected neighbor (or follower) is a susceptible one. The recovery rate is set as 1 here. The simulation stops when there is no infected node anymore. Notice that this model is slightly different from the standard SIR model where all the followers of an infected node have the chance to be infected. The present mechanism is usually used to mimic the limited neighbor contact capability that is a positive correlation to the degree of individuals [24,25].

The number of nodes that are finally infected when the infection starts from a given node \( i \) is denoted as its spreading ability \( s_i^\mu \), where \( \mu \) is the infection rate in the SIR model. Here, we select a relatively small value of \( \mu \) so that the infected percentage of the nodes is not so large (\( \mu = 0.04 \) in the Orkut network and \( \mu = 0.05 \) in the other three networks). We select 1000 highest-degree nodes from each network and calculate Kendall’s tau correlation coefficient (\( \tau \)) between \( f_i \) and \( s_i^\mu \). The results are reported in fig. 3. Clearly, \( \tau \) increases with \( \lambda \) at first, and then decreases after reaching a maximum. This indicates that the average degree of the neighboring nodes are beneficial for improving the identification of the most influential nodes.
results are averaged over 100 independent simulations.

spreaders. When $\lambda$ gets very large, the node with small degree but with large average degree of neighbors will have high ranking, leading to a small $\tau$ value.

We further move to improve the ranking accuracy by the path diversity. Specifically, we will make use of the entropy $H$ to modify eq. (3) as

$$f_i^* = k_i + H_i \lambda \sum_{j \in \Gamma_i} \frac{k_j}{k_i}.$$  

Kendall’s tau correlation coefficient between $s_i^\mu$ and $f_i^*$ is also shown in fig. 3. One can see that the optimal $\tau$ can be even higher after $H_i$ is introduced.

So far, we have shown that the local path diversity is very valuable for identifying influential nodes. For the local path number, the nodes’ degree and the average degree of neighboring nodes are two key factors. Therefore, we combine these two pieces of information (local path diversity and number) and propose a local spreader ranking method named KED in this letter. It can be described as

$$f_i = k_i E_i D_i,$$  

where $E_i = H_i$, $D_i$ is a function of the degree of the neighbors of node $i$ with $D_i = \frac{e^{\sum_{j \in \Gamma_i} \rho_{ij}}}{k_i}$. The power in $D$ is the sum of the neighbors’ degree normalized by its maximum possible value among all nodes, so that the value of $D_i$ is bounded in $[1, e]$. In this ranking metric, the node with larger $f_i$ is supposed to have higher spreading ability. As an example, the $f_i$ for the red node is 13.5914 in fig. 1(a) and 6.5672 in fig. 1(b), which indicates that the red node in fig. 1(a) has stronger spreading ability than that in fig. 1(b). It is also noted that the model described in eq. (5) is not only used to rank nodes in undirected networks, but also in directed networks.

Fig. 3: (Color online) Kendall’s tau correlation coefficient $\tau$ between real spreading ability $s_i$ and the ranking score $f_i$ of the 1000 highest-degree nodes on four real networks when the simple ranking methods in eq. (3) and eq. (4) are applied. The results are averaged over 100 independent simulations.

For comparison, we consider two ranking algorithms on directed networks: PageRank [13] and LeaderRank [14], and one algorithm on undirected networks: K-shell [1,31]. PageRank is depicted as a random walk on hyperlinked networks. In the algorithm, the random walker has a certain probability to jump from a web page to a random web page and the probability is set as 0.15 in our simulation. LeaderRank is also a random-walk–based ranking algorithm. Different from PageRank, LeaderRank introduces a ground node $g$ connecting to every node $i$ in the network. The K-shell method is a ranking algorithm based on network decomposition. It starts by removing all nodes with one connection until no more of such nodes remain, and assigns them to the 1-shell. After that, all nodes with residual degree 2 are recursively removed and the 2-shell is created. This procedure continues as the residual degree increases until all nodes in the networks have been assigned to one of the shells.

Fig. 4: (Color online) The ratio $\rho$ of the average final coverage $\langle s^\mu \rangle$ of top-$L$ nodes obtained by KED to that by degree, K-shell, PageRank or LeaderRank on four real networks. The spreading starts from each single node separately in the top-$L$ list. The results are averaged over 100 independent simulations.

Experiments and results. – To investigate the influence of a node in information spreading, we initially set this node to be infected. As discussed above, the final coverage of this node $s_i^\mu$ is used to represent the spreading ability of $i$. The average final coverage $\langle s^\mu \rangle$ of top-$L$ ranked nodes obtained by each ranking algorithm is used to investigate the performance of these algorithms. We define $\rho$ as a ratio of $\langle s^\mu \rangle$ of KED to $\langle s^\mu \rangle$ of other methods. The results in fig. 4 show that $\rho$ is larger than 1 in four networks. The result indicates that the information can spread more widely from the top ranked nodes obtained by KED than that obtained by degree, K-shell, PageRank and LeaderRank.

Kendall’s tau correlation coefficient between the ranking of nodes and the real spreading ability is also calculated. We do not compare KED with K-shell since many nodes
Path diversity improves the identification of influential spreaders

Fig. 5: (Color online) Kendall’s tau correlation coefficient $\tau$ between real spreading ability $s_i$ and the ranking score $f_i$ of the 1000 highest-degree nodes on four real networks. The results are averaged over 100 independent simulations.

are with the same K-shell value, which makes the $\tau$ value of the K-shell method very low. Like before, we take into account 1000 highest-degree nodes. Kendall’s tau correlation coefficient between the ranking of top-$L$ ($L \leq 1000$) nodes from different methods and their real spreading ability is shown in fig. 5. One can see that generally $\tau$ of KED is the largest among all the ranking methods. That is to say, the KED ranking algorithm can generate a more accurate ranking of spreading ability than degree centrality, PageRank and LeaderRank.

In fact, $D$ and $E$ can be regarded as the correction factors to the existing spreader ranking algorithms. Specifically, $D$ is the correction factor of path number and $E$ is the correction factor of path diversity. To make our idea more general, we also apply the correction factors $D$ and $E$ to the K-shell, PageRank and LeaderRank algorithms. We show Kendall’s tau between the rank of top-1000 nodes and their spreading coverage in table 2. The accuracy of ranking the spreaders is considerably increased due to the correction of $E$ and $D$. In the K-shell, PageRank and LeaderRank methods, the information of path number is implicitly embedded. Interestingly, we find that even adding $D$ alone can improve these methods. In the above analysis, the infection rate is the same as the values in figs. 3, 4 and 5 ($\mu = 0.04$ in the Orkut network and $\mu = 0.05$ in the other three networks). We actually try different infection rates in our simulation, and the results show that adding $D$ and $E$ can always improve the $\tau$ value of degree, PageRank and LeaderRank methods. However, there is some parameter range in which $D$ and $E$ may lower the $\tau$ value of the K-shell method.

In order to further understand the effect of the factor $E$ and $D$ in the KED method, we add the power $\alpha$ and $\beta$ to $E$ and $D$, respectively. The modified KED method reads

$$f_i = k_i E_i^\alpha D_i^\beta.$$  \hspace{1cm} (6)

Table 2: Kendall’s tau coefficient when the correction factors $D$ and $E$ are applied to the degree (K), K-shell (KS), PageRank (PR) and LeaderRank (LR) algorithms

| Method | Youtube | Orkut | EmailEU | Digg |
|--------|---------|-------|---------|------|
| K      | 0.5416  | 0.4811| 0.6516  | 0.5850|
| K+D    | 0.5581  | 0.4816| 0.6906  | 0.6519|
| K+E    | 0.5810  | 0.5467| 0.7509  | 0.6115|
| K+E+D  | 0.6064  | 0.5559| 0.7687  | 0.6746|
| KS     | 0.4033  | 0.4031|         |      |
| KS+D   | 0.4970  | 0.6250|         |      |
| KS+E   | 0.3914  | 0.6360|         |      |
| KS+E+D | 0.4418  | 0.7032|         |      |
| PR     |         |       | 0.4725  | 0.4389|
| PR+D   |         | 0.5051| 0.4492  |      |
| PR+E   |         | 0.5518| 0.4583  |      |
| PR+E+D |         | 0.5821| 0.4684  |      |
| LR     |         |       | 0.4964  | 0.5837|
| LR+D   |         | 0.5325| 0.6218  |      |
| LR+E   |         | 0.5868| 0.6083  |      |
| LR+E+D |         | 0.6184| 0.6437  |      |

Fig. 6: (Color online) The dependence of $\tau$ on $\alpha$ and $\beta$ in the modified KED method. The infection rate is set as $\mu = 0.12$ in Youtube, emailEU and Digg network and $\mu = 0.06$ in Orkut since it is a very dense network. The results are averaged over 100 independent simulations.

The dependence of $\tau$ on the $\alpha$ and $\beta$ in the modified KED method is reported in fig. 6. It shows that $\beta$ can indeed enhance $\tau$ even without $E$ (i.e., $\alpha = 0$). With both $\alpha$ and $\beta$, $\tau$ can be further improved to a much higher value. For example, the improvement in $\tau$ solely from $\beta$ is 69.61% in Youtube, 82.82% in Orkut, 88.81% in emailEU and 42.39% in Digg, respectively. However, with both $\alpha$ and $\beta$, the improvement in $\tau$ can reach 133.97% in Youtube, 232.17% in Orkut, 124.27% in emailEU and 49.58% in Digg. The results indicate that the path diversity is indeed an important factor in ranking spreaders. Interestingly, there is an optimal $\beta$ in each subfigure and it depends on the infection rate $\mu$. We remark that a similar heat map
can be observed when applying $E$ and $D$ to other ranking algorithms such as K-shell, PageRank and LeaderRank.

**Conclusion.** – In this letter, we introduced the concept of path diversity in ranking spreaders. Measured by information entropy, it is introduced to a simple spreader ranking algorithm. The results show that not only the number of paths that determines the spreading ability of a node, but also the path diversity is also a significant factor. We further proposed a local spreader ranking algorithm named KED, in which the information of path number and that of path diversity are combined. We use the SIR model to simulate the spreading process on four real social networks. The results show that the performance of KED is better than that of degree, PageRank and LeaderRank, and in most cases, better than that of K-shell. With two parameters to adjust the contribution of path number and path diversity in the KED method, we find that the ranking accuracy can be further improved. From the practical point of view, our method can be easily applied to very large real systems since it is based on only local information. Finally, we find that the concept of combining path number and path diversity can be applied to improve the K-shell, PageRank and LeaderRank algorithms as well.

Many extensions can be made based on this work. Most existing ranking algorithms are designed to rank the spreading ability of individual nodes. In our simulation, we notice that their performance is almost the same when they are applied to pick up a group of nodes for spreading (i.e. instead of originating from an individual node, the spreading will start from a group of the top ranking nodes). Therefore, the method for this task will require further investigation in the future. Moreover, how to identify influential nodes in temporal networks and bipartite networks is also a long-term challenge problem. Some progress has been made in this direction [32,33], but systematic analysis is still lacking. We remark here that in these systems, the combination of path number and path diversity may also lead to an improvement in ranking spreaders.

***

We thank the anonymous referees for helpful comments to improve the paper. This work was partially supported by the National Natural Science Foundation of China under Grant Nos. 61103109 and 61003231. D-BC acknowledges the Huawei University-Enterprise Cooperation project YBCB2011057.

REFERENCES

[1] Kitsak M., Gallos L. K., Havlin S., Liljeros F., Muchnik L., Stanley H. E. and Makse H. A., Nat. Phys., 6 (2010) 888.
[2] Aral S. and Walker D., Science, 337 (2012) 337.
[3] Bond R. M., Fariss C. J., Jones J. J., Kramer A. I., Marlow C., Settle J. E. and Fowler J. H., Nature, 489 (2012) 295.
[4] Bauer F. and Lizier J. T., EPL, 99 (2012) 68007.
[5] Brands U., J. Math. Sociol., 25 (2001) 163.
[6] Freeman L. C., Sociometry, 40 (1977) 35.
[7] Noh J. D. and Rieger H., Phys. Rev. E, 92 (2004) 118701.
[8] Opsahl T., Agneessens F. and Skvoretz J., Soc. Netw., 32 (2010) 245.
[9] Katz L., Psychometrika, 18 (1953) 39.
[10] Zeng A. and Zhang C.-J., Phys. Lett. A, 377 (2013) 1031.
[11] Garas A., Schweitzer F. and Havlin S., New J. Phys., 14 (2012) 083030.
[12] Kleinberg J., J. ACM, 46 (1999) 604.
[13] Brin S. and Page L., Comput. Netw. ISDN Syst., 30 (1998) 107.
[14] Lü L., Zhang Y.-C., Wang B. H., Ren J., Bai W. J., Shi Z. W., Lü L., Zhang Y.-C., and Makse H. A., Nat. Phys., 8 (2012) 083030.
[15] Weng J., Lim E. P., Jiang J. and He Q., Proceedings of the 3rd ACM International Conference WSDM (ACM Press, New York, USA) 2010.
[16] Saito K., Kimura M., Ohara K. and Motoda H., Knowl. Inf. Syst., 30 (2012) 613.
[17] Chen D.-B., Lü L., Shang M.-S., Zhang Y.-C. and Zhou T., Physica A, 391 (2012) 1777.
[18] Teixeira R., Marzullo K., Savage S. and Voelker G. M., Proceedings of the 3rd ACM SIGCOMM Conference on Internet Measurement (ACM Press, New York, USA) 2010.
[19] Shannon S. E., Bell Syst. Tech. J., 27 (1948) 379.
[20] VanSiclen C. D., Phys. Rev. E, 56 (1997) 5211.
[21] Wang S., Wang F., Li Z. and Zhang X., Proceedings of the 2012 Second International Conference on Cloud and Green Computing (IEEE Computer Society, Washington DC, USA) 2012, p. 586.
[22] Banos R. A., Borge-Holthoefer J. and Moreno Y., EPJ Data Sci., 2 (2013) 6.
[23] Anderson R. M., May R. M. and Anderson B., Infectious Diseases of Humans: Dynamics and Control (Oxford University Press, London) 1992.
[24] Zhou T., Liu J. G., Bai W. J., Chen G. R. and Wang B. H., Phys. Rev. E, 74 (2006) 056109.
[25] Yang R., Wang B. H., Ren J., Bai W. J., Shi Z. W., Wang W. X. and Zhou T., Phys. Lett. A, 364 (2007) 189.
[26] Gómez S., Arenas A., Borge-Holthoefer J., Meloni S. and Moreno Y., EPL, 89 (2010) 38009.
[27] Yang J. and Leskovec J., Proceedings of the ACM SIGKDD on MDS (ACM Press, New York, USA) 2012.
[28] Leskovec J., Kleinberg J. and Faloutsos C., ACM Trans. Knowl. Discov. Data, 1 (2007) 2.
[29] Lerman K. and Ghosh R., Proceedings of the 4th International Conference on Weblogs and Social Media (AAAI Press, Menlo Park, Cal.) 2010.
[30] Maslov S. and Sneppen K., Science, 296 (2002) 910.
[31] Carmi S., Havlin S., Kirkpatrick S., Shavit Y. and Shiri E., Proc. Natl. Acad. Sci. U.S.A., 104 (2007) 11150.
[32] Zhou Y. B., Lü L. and Li M., New J. Phys., 14 (2012) 033033.
[33] Wei D., Deng X., Zhang X., Deng Y. and Mahadevan S., Physica A, 392 (2013) 2564.