

LETTER

## Identifying effective multiple spreaders by coloring complex networks

To cite this article: Xiang-Yu Zhao *et al* 2014 *EPL* **108** 68005

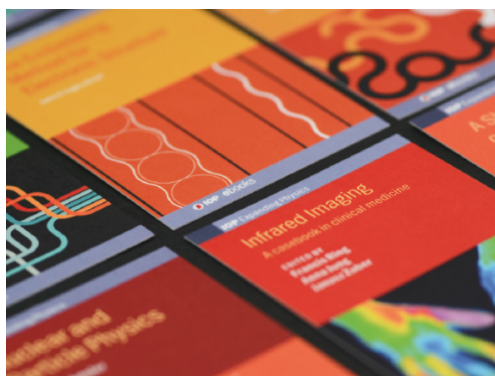
View the [article online](#) for updates and enhancements.

### Related content

- [Identifying multiple influential spreaders via local structural similarity](#)  
J.-G. Liu, Z.-Y. Wang, Q. Guo *et al.*
- [Rapid identifying high-influence nodes in complex networks](#)  
Song Bo, Jiang Guo-Ping, Song Yu-Rong *et al.*
- [Iterative resource allocation for ranking spreaders in complex networks](#)  
Zhuo-Ming Ren, An Zeng, Duan-Bing Chen *et al.*

### Recent citations

- [Chinenye Ezech \*et al\*](#)
- [Identifying influential spreaders in reversible process](#)  
Junyi Qu *et al*
- [Instantaneous success and influence promotion in cyberspace - how do they occur?](#)  
Ya Chen *et al*



**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# Identifying effective multiple spreaders by coloring complex networks

XIANG-YU ZHAO<sup>1,2</sup>, BIN HUANG<sup>3</sup>, MING TANG<sup>1(a)</sup>, HAI-FENG ZHANG<sup>4,5(b)</sup> and DUAN-BING CHEN<sup>1</sup>

<sup>1</sup> Web Sciences Center, University of Electronic Science and Technology of China - Chengdu 611731, PRC

<sup>2</sup> School of Computer Science and Technology, University of Science and Technology of China - Hefei 230027, PRC

<sup>3</sup> School of Applied Mathematics, Chengdu University of Information Technology - Chengdu 610225, PRC

<sup>4</sup> School of Mathematical Science, Anhui University - Hefei 230601, PRC

<sup>5</sup> Department of Communication Engineering, North University of China - Taiyuan, Shan'xi 030051, PRC

received 7 October 2014; accepted in final form 5 December 2014

published online 7 January 2015

PACS 89.75.Hc – Networks and genealogical trees

PACS 89.75.Fb – Structures and organization in complex systems

PACS 87.23.Ge – Dynamics of social systems

**Abstract** – How to identify influential nodes in social networks is of theoretical significance, which relates to how to prevent epidemic spreading or cascading failure, how to accelerate information diffusion, and so on. In this letter, we make an attempt to find *effective multiple spreaders* in complex networks by generalizing the idea of the coloring problem in graph theory to complex networks. In our method, each node in a network is colored by one kind of color and nodes with the same color are sorted into an independent set. Then, for a given centrality descriptor, the nodes with the highest centrality in an independent set are chosen as multiple spreaders. Comparing this approach with the traditional method, in which nodes with the highest centrality from the *entire* network perspective are chosen, we find that our method is more effective in accelerating the spreading process and maximizing the spreading coverage than the traditional method, no matter in network models or in real social networks. Moreover, the low computational complexity of the coloring algorithm guarantees the potential applications of our method.

Copyright © EPLA, 2014

**Introduction.** – Spreading phenomenon is ubiquitous in nature, and describes many important activities in the society [1]. Examples include the propagation of infectious diseases, the dissemination of information (*e.g.*, ideas, rumors, opinions, behaviors), and the diffusion of new technological innovations. With the advancement of complex network theory, spreading dynamics on complex networks have been intensively studied in the past decades. Many studies have revealed that the spreading process is strongly influenced by the network topologies [2,3].

An important issue in analyzing complex networks is to identify the most influential nodes in a spreading process, which is crucial for developing efficient strategies to control epidemic spreading, or accelerate information diffusion. For this reason, more and more attentions have been paid to identify the most influential nodes in networks [4–11]. Many centrality descriptors have been proposed, such as,

degree centrality (defined as the degree of a node) [12], betweenness centrality (quantified by the number of times that all shortest paths travel through the node) [13], eigenvector centrality (defined as the dominant eigenvector of the adjacency matrix) [14], neighborhood centrality (defined as the average connectivity of all neighbors) [15] and closeness centrality (reciprocal of the sum of the lengths of the geodesic distance to every other node) [16]. Recently Kitsak *et al.* proposed a *k*-core decomposition to identify the most influential spreaders, which is found to be better than the degree centrality descriptor in many real networks [6]. However, most of these methods quantify the influence of each node from the viewpoint of entire network, which may be particularly suitable to the case in which single spreader of information is considered (*i.e.*, only one node is selected as the initial spreader) [17,18]. Many times, the spreading processes of rumors, ideas, opinions, or advertisements may initiate from different spreaders. In this case, the traditional methods that only select the nodes in the top of one certain ranking (*e.g.*, the

<sup>(a)</sup>E-mail: tangminghuang521@hotmail.com

<sup>(b)</sup>E-mail: haifengzhang1978@gmail.com

ranking obtained by ordering the nodes according to the degree centrality) may be not the optimal strategy since these chosen nodes cannot be dispersively distributed [19]. Thus, how to properly choose the multiple initial spreaders is an important and challenging problem [4,5,20,21]. To this end, apart from the consideration for the influence of each node, we need to make the chosen spreaders being sufficiently dispersive to ensure that the information can quickly diffuse.

Motivated by the above reasons, in this letter, we propose a method to detect the *effective multiple spreaders* which can enhance the spreading processes. For this method, the independent sets with disjointed nodes are obtained by coloring a network at first, and then the nodes with the highest centrality in an independent set are chosen as the initial spreaders. By implementing extensively simulations on network models as well as the real networks, we find that our method can effectively enhance the spreading process. In addition, the computational complexity of our method is  $O(N^2)$  when the size of network is  $N$ , and is independent of other parameters (*e.g.*, number of edges and number of seeds to be selected), which shows better applicability than the existing methods such as the original greedy algorithm [4] and its improvement [5].

The remainder of this paper is organized as follows. In the next section, we first describe the details of our method. Then we present the main results in the third section. Finally, we summarize the conclusions in the last section.

**Methods.** – The four-color theorem is the most famous theorem in the graph coloring problem, which states that, given any a plane graph, no more than four colors are required to color the regions of the plane graph so that no two adjacent regions have the same color [22–24]. In other words, the number of colors for all vertexes in a plane graph is not greater than four. The graph coloring problem has a wide range of applications, including the problem of the wireless channel allocation [25], the problem of the scheduling [26,27], and so on. Here, we generalize the idea of the graph vertex coloring problem to complex networks to obtain the *effective multiple spreaders*. The main steps are as follows. We first color a given network  $G = (V, E)$  ( $V$  denotes the set of nodes and  $E$  denotes the set of edges) with the Welsh-Powell algorithm [27] [see below] and each node in set  $V$  is colored by one kind of color. Secondly, sorting the nodes with the same color into the same subset  $V_i, i = 1, 2, \dots, K$  (each subset is called an independent set,  $K$  denotes the number of colors being used to color the network), which ensures that  $V = V_1 \cup V_2 \dots \cup V_K$  and  $V_i \cap V_j = \emptyset, \forall i \neq j$ , where  $\emptyset$  is an empty set. As nodes with the same color are not directly connected, the distance between any two nodes in an independent set will not be smaller than two. This concept is significantly different from the *community structure* based on link density, which means the internal members are more connected than those members from different

communities. In order to make multiple spreaders dispersively distributed, we choose the nodes with highest centrality descriptor in an independent set as the initial spreaders. To ensure that there are sufficient nodes from which to choose, we give priority to those large independent sets with more nodes, especially *the largest independent set* with the maximum nodes in this letter.

Though there are many graph coloring algorithms, an ideal algorithm should have the qualities: the time complexity is low and the number of colors to color the network is few since many real networks have huge sizes. In view of this, we choose the Welsh-Powell algorithm, whose time complexity is  $O(N^2)$  [28]. For a given network  $G = (V, E)$  with  $V = \{v_1, v_2, \dots, v_N\}$ , we let the color function be  $\pi$  and the color set be  $C = \{1, 2, \dots, \Delta + 1\}$ , where  $\Delta$  be the maximum degree of network  $G$ . The details of the Welsh-Powell algorithm are the following [27].

**Step 1:** according to the degree centrality, re-rank the node set  $V$  in descending order, such that  $k(v_1) \geq k(v_2) \geq \dots \geq k(v_N)$ .

**Step 2:** let  $\pi(v_1) = 1, i = 1$ .

**Step 3:** if  $i = N$ , stop; otherwise, let  $C(v_{i+1}) = \{\pi(v_j) | j \leq i, \text{ and } v_j \text{ is connected by } v_{i+1}\}$ . Let  $m$  be the minimal positive integral of the subset  $C \setminus C(v_{i+1})$  (where  $C \setminus C(v_{i+1})$  is the complementary set of the subset  $C(v_{i+1})$  in the set  $C$ ). Then  $\pi(v_{i+1}) = m$ .

**Step 4:** let  $i = i + 1$ , and back to step 3.

In the above algorithm,  $k(v_i)$  denotes the degree of node  $v_i$ , and  $\pi(v_i) = m$  denotes that the node  $v_i$  is colored by a color labeled  $m$ .

When the multiple spreaders are selected, a spreading model should be used to check the effectiveness of the proposed method. In many literatures, the susceptible-infected-recovered (SIR) epidemic model is used to simulate the spreading process in networks, in which each node can be in one of three states: susceptible, infected, or recovered. A susceptible node is healthy and can catch the disease from each infected neighbor with transmission rate  $\beta$ , whereas an infected node becomes recovered with recovery rate  $\mu$  and is immune to the disease. In the classical SIR model, each infected node can contact *all* of its neighbors at per time step with transmission rate  $\beta$ . In reality, at a time step, one often can contact *one* neighbor at most, taking the sex activity and the telephone marketing activity as examples. Thus, in this letter, we use the SIR epidemic model based on a contact process to simulate the spreading process and evaluate the effectiveness of the proposed method [29–31]. Its worth noting that our method can also be applied to the classical SIR model. The epidemic spreading process ends when there is no infected node in the network. We define the effective

transmission rate  $\lambda = \beta/\mu$  by fixing the recovery rate  $\mu = 0.1$ .

**Results.** – One should note that our method is compared with the traditional method according to one given centrality descriptor. Taking the degree centrality descriptor as an example, for our method, the nodes in the *largest independent set* are ranked in a descending order according to the degree centrality descriptor, and then the nodes at the top of the ranking are selected as spreaders (labeled as IS method). For the traditional method, all nodes are ranked according to the degree centrality descriptor from *entire network* perspective, and the same amount nodes at the top of the ranking are selected as the spreaders (labeled as EN method).

To evaluate the effectiveness of the IS method, we define the relative difference of outbreak size  $\Delta r_R$  as  $\Delta r_R = (R_i - R_e)/R_e$ , where  $R_i$  and  $R_e$  are the final number of recovered nodes for the IS method and the EN method, respectively. Thus, the larger value of  $\Delta r_R$  is, the better effectiveness of the IS method is. All results are averaged over 500 independent realizations.

We first perform the Welsh-Powell algorithm to Barabási-Albert (BA) network with size  $N = 10000$  and average degree  $\langle k \rangle = 12$  [32]. From the inset of fig. 1(a), one can see that such a network can be successfully divided into different independent sets, where the number of color is  $K = 8$  and the node number in the largest independent set is over 2000. Figure 1 also compares the IS method with the EN method based on the degree centrality descriptor. In general,  $\Delta r_R$  is larger than 0 for the different initial spreaders  $n_0$  and the different values of transmission rate  $\lambda$ . This indicates that the IS method is better than the latter case in most situations. More importantly, from fig. 1(a) one can see that the advantage of the IS method is the most striking when  $\lambda$  is not too small or too large. As we know, when  $\lambda$  is very small, the information initiated from any node can only spread to a very small fraction of nodes. The influence regions of multiple spreaders for these two methods scarcely overlap each other. The outbreak size is approximately equal to the sum of multiple spreaders' spreading coverage (*i.e.*, the number of infected nodes). Thus, the difference of the two methods cannot be distinguished obviously at a small  $\lambda$ , which results in a small value of  $\Delta r_R$ . With the increase of transmission rate  $\lambda$ , single node can induce a greater spreading coverage. More dispersive locations of multiple spreaders for the IS method lead to less overlap of the influence regions, and the IS method thus performs better. When  $\lambda$  is very large, the information can diffuse to a very wide range even single node is selected as the initial spreader. In this case, there are too many overlap influence regions to play the role of the IS method. In fig. 1(b),  $\Delta r_R$  as a function of  $n_0$  displays distinct trends for different values of  $\lambda$ , which stems from the combined effects of both the dispersive locations of multiple spreaders and the intricate spreading processes at different  $\lambda$ .

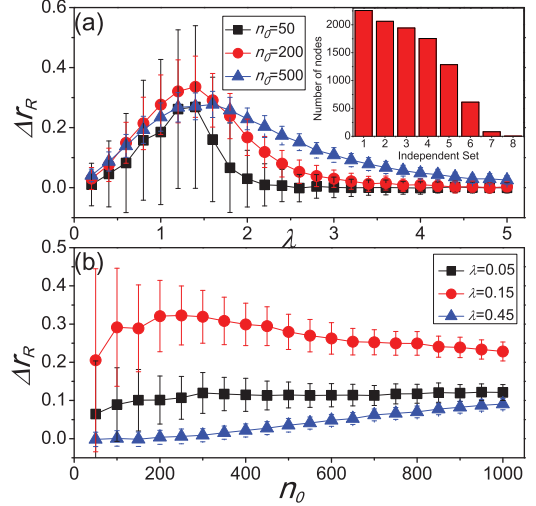


Fig. 1: (Color online) For the degree centrality descriptor, the EN method and the IS method are compared in the BA network. (a) The relative difference of outbreak size  $\Delta r_R$  as a function of the effective transmission rate  $\lambda$  for a different number of initial spreaders  $n_0$ ; (b)  $\Delta r_R$  as a function of  $n_0$  for different values of  $\lambda$ . The inset of subfigure (a) is the size distribution of independent sets. The error bars are given by the standard deviation.

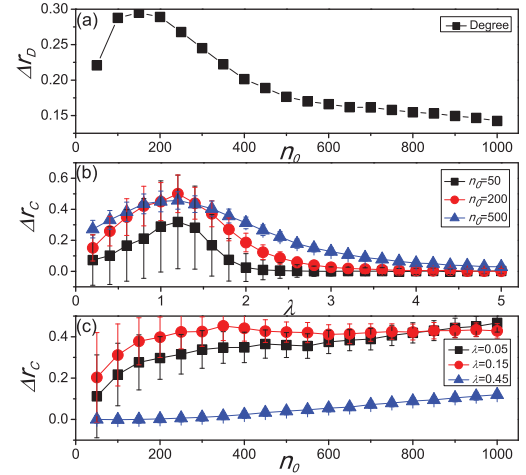


Fig. 2: (Color online) For the degree centrality descriptor, the relative difference of average distance  $\Delta r_D$  and the relative difference of effective contacts  $\Delta r_C$  are plotted to explain the phenomena in fig. 1. (a)  $\Delta r_D$  vs.  $n_0$ , (b)  $\Delta r_C$  vs.  $\lambda$  for different values of  $n_0$ , (c)  $\Delta r_C$  vs.  $n_0$  for different values of  $\lambda$ . The detail definitions of the two descriptors are given in the main text.

These distinct trends will be verified by the relative difference of effective contacts in fig. 2(c) later. Note that given the contacting randomness in the CP model [33,34], there is a large fluctuation of  $\Delta r_R$ , especially near the critical point [35].

To maximize the spreading coverage, we hope these spreaders not only have high centrality, but also are dispersive enough so that a susceptible node has only one or few infected neighbor instead of surrounded by many



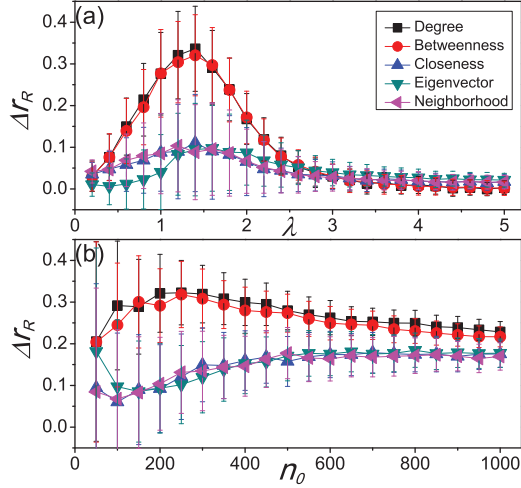


Fig. 3: (Color online) Comparing the IS method with the EN method based on five descriptors (*i.e.*, degree, betweenness, closeness, eigenvector, and neighborhood) in the BA network. (a)  $\Delta r_R$  vs.  $\lambda$  at  $n_0 = 200$ , (b)  $\Delta r_R$  vs.  $n_0$  at  $\lambda = 1.5$ .

infected nodes to reduce the overlap effect of the spreaders. As a result, if we can verify that the initial spreaders in the IS method have a larger average distance among them and produce less overlap of spreading influence (*i.e.*, more effective contacts between infected nodes and their susceptible neighbors) than that in the EN method, the phenomena in fig. 1 will be naturally explained. For this reason, we define two descriptors. One is the relative difference of average distance,  $\Delta r_D = (D_i - D_e)/D_e$ , where  $D_i$  and  $D_e$  are the average distance among the initial spreaders based on the IS method and the EN method, respectively. The other is the relative difference of effective contacts,  $\Delta r_C = (C_i - C_e)/C_e$ , where  $C_i$  and  $C_e$  are the number of effective contacts between infected nodes and their susceptible neighbors based on the IS method and the EN method, respectively. In each time step of the transmission processes, an infected node randomly chooses one of its neighbors to transmit the information with probability  $\beta$ . If the chosen neighbor is susceptible, this contact is defined as an effective contact; otherwise, it is not effective. A greater number of effective contacts denotes the less overlap of spreading regions initiated from multiple spreaders. In fig. 2, one can see that  $\Delta r_D$  is always larger than 0 (see fig. 2(a)). The reason can be explained as follows: the distance between any two nodes in an independent set is greater than or equal to 2, while the nodes with highest centrality are connected more easily in the BA network. In figs. 2(b) and (c),  $\Delta r_C$  is generally larger than 0, too. Moreover, comparing figs. 2(b) and (c) with figs. 1(a) and (b), respectively, one can observe that the effects of  $\lambda$  and  $n_0$  on  $\Delta r_D$  and  $\Delta r_C$  are similar to that on  $\Delta r_R$ . Thus, the reason for the advantage of the IS method is well explained. To be specific, the greater  $\Delta r_D$  induces the greater  $\Delta r_C$  and then results in the greater  $\Delta r_R$ .

In fig. 3, we further compare the IS method with the EN method based on five commonly used descriptors—degree

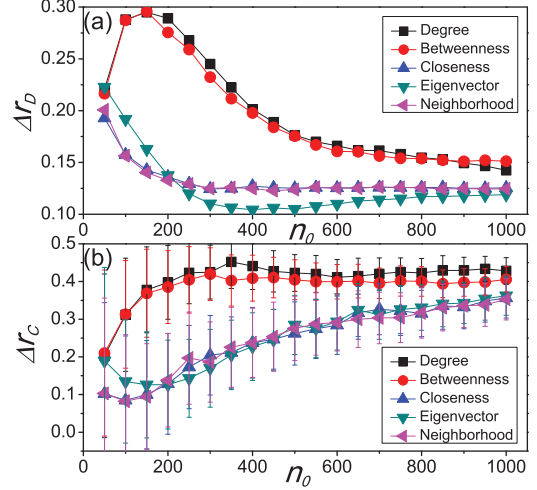


Fig. 4: (Color online)  $\Delta r_D$  and  $\Delta r_C$  as functions of  $n_0$  are given to explain the phenomena in fig. 3. (a)  $\Delta r_D$  vs.  $n_0$ , (b)  $\Delta r_C$  vs.  $n_0$ . The parameter is chosen as  $\lambda = 1.5$ .

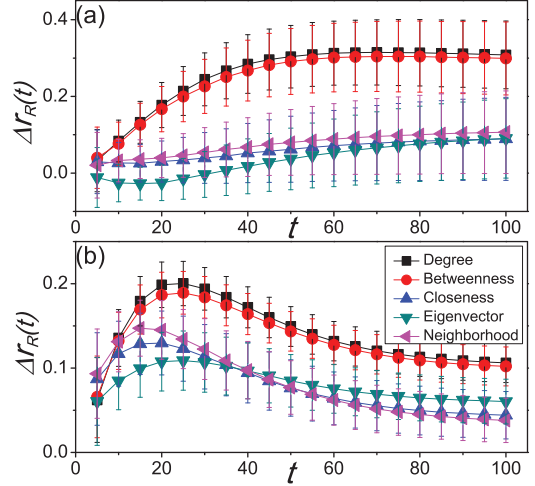


Fig. 5: (Color online) For five descriptors, the time evolutions of  $\Delta r_R(t)$  in the BA network are plotted for (a)  $\lambda = 1.5$  and  $n_0 = 200$ , (b)  $\lambda = 3.0$  and  $n_0 = 500$ .

centrality, betweenness centrality, closeness centrality, eigenvector centrality, and neighborhood centrality. Since the  $k$ -core decomposition can not quantify the relative influence of nodes in the BA network, this descriptor is not considered here [10]. As in fig. 1, the results in fig. 3 indicate that the IS method is more effective than the EN method for all cases. In particular, the advantage of the IS method is the most remarkable for the betweenness centrality descriptor. The positive  $\Delta r_D$  and  $\Delta r_C$  for different cases shown in fig. 4 can explain that the IS method is superior to the EN method for different centrality descriptors. Meanwhile, figs. 4(a) and (b) show that the methods based on the degree and betweenness centrality descriptors can generate the largest values of  $\Delta r_D$  and  $\Delta r_C$ , leading to the highest efficiency in enhancing the spreading coverage in fig. 3.

Table 1: Basic structural parameters of Blogs network and Email network.  $N$  is the total number of nodes,  $\langle k \rangle$  denotes the average degree, and  $H$  is the degree heterogeneity, defined as  $H = \langle k^2 \rangle / \langle k \rangle^2$ .  $D$  is the average shortest distance, and  $C$  and  $r$  are the clustering coefficient and assortative coefficient, respectively.

Network	$N$	$\langle k \rangle$	$H$	$D$	$C$	$r$
Blogs	3982	3.42	4.038	6.227	0.146	0.133
Email	1133	9.62	1.942	3.716	0.110	0.078

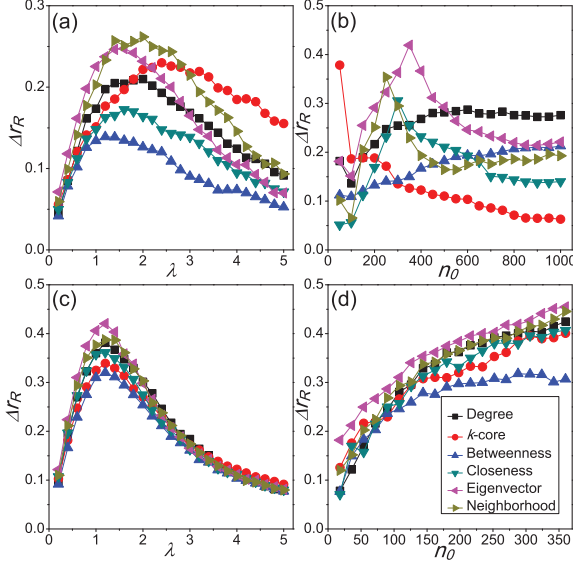


Fig. 6: (Color online) For different descriptors,  $\Delta r_R$  in Blogs and Email networks are given as the functions of  $\lambda$  and  $n_0$ , respectively.  $\Delta r_R$  as a function of the values  $\lambda$  (a) and  $n_0$  (b), respectively in Blogs network;  $\Delta r_R$  as a function of the values  $\lambda$  (c) and  $n_0$  (d), respectively in Email network. In panels (a) and (c), the parameter is chosen as  $n_0 = 200$ ; and in panels (b) and (d), the parameter is chosen as  $\lambda = 1.5$ . Detail network information is summarized in table 1.

The time evolutions of the relative difference of outbreak size  $\Delta r_R(t) = [R_i(t) - R_e(t)]/R_e(t)$  for different descriptors are also shown in fig. 5. As shown in fig. 5(a) ( $n_0 = 200$  and  $\lambda = 0.15$ ) and fig. 5(b) ( $n_0 = 500$  and  $\lambda = 0.3$ ), the value of  $\Delta r_R(t)$  is generally larger than 0. It means that, compared with the EN method, the IS method can not only extend the spreading coverage but also *speed up* the spreading process. Moreover, fig. 5(a) illustrates that  $\Delta r_R(t)$  monotonously increases with time step  $t$  when the values of  $n_0$  and  $\lambda$  are small. Nonetheless, when the values of  $n_0$  and  $\lambda$  are large,  $\Delta r_R(t)$  increases with  $t$  at first and then decreases to a stable level (see fig. 5(b)). For the latter case, the information begins to diffuse from the initial spreaders in the early stages, the IS method ensures these multiple spreaders are more dispersive, leading to the better effectiveness of the IS method. With the further increase of  $t$ , the information will diffuse to a wide range of network and the influence regions of the

multiple spreaders are more likely to overlap each other, the advantage of the IS method will thus be weakened.

Finally, we use two real networks —Blogs network [36] and Email network [37]. To further confirm the effectiveness of our method, where some basic structural features of the two networks are given in table 1. In fig. 6,  $\Delta r_R$  as a function of  $\lambda$  and  $n_0$  for six centrality descriptors are presented. Note that  $k$ -core descriptor is also considered, besides the five descriptors shown in fig. 3. No matter the Blogs network (figs. 6(a) and (b)) or the Email network (figs. 6(c) and (d)), the IS method is more effective in enhancing the spreading process than the EN method. In fig. 6(b), there is an sudden decrease of  $\Delta r_R$  at  $n_0 = 100$  for almost all the centrality descriptors, which is closely related to the decrease of  $\Delta r_D$  and  $\Delta r_C$ . In order to further verify our method, other independent sets such as the second largest one are also investigated. As expected, all simulations reveal the same conclusion.

**Summary.** — Even though great improvement has been made in the research of identifying influential spreaders, there are still many problems needed to be solved, among which how to find multiple effective spreaders is an important question. It is commonly believed that, to effectively speed up the spreading process, the selected multiple spreaders should be as dispersive as possible to reduce the overlap of the spreading regions initiated from multiple spreaders. But, how to design an effective method to achieve this goal is almost missing. In this letter, we have proposed an effective method by selecting the nodes with the highest centrality from an independent set as the initial spreaders rather than from the entire network. By testing such a method on BA network and two real networks, we found that our method can greatly enhance the average distance among these initial spreaders and the effective contacts between the susceptible nodes and the infected nodes. Therefore, our method can ensure that the information diffuses much wider and faster. Meanwhile, the computational complexity of the coloring algorithm used in the paper is  $O(N^2)$  and is independent of other structural and dynamical parameters, which shows better applicability than those existing methods. Although the efficiency of our method was studied from the perspective of the spreading process, it is immediately related to many other aspects [38], including network resilience to attacks, immunization of epidemics, commercial product promotions in markets and other aspects, which implies the potential applications of our method.

\*\*\*

X-YZ would like to thank KAI QI for stimulating discussions. This work was partially supported by the National Natural Science Foundation of China (Grant Nos. 11105025, 61473001, 11331009, 61433014) and China Postdoctoral Science Special Foundation (Grant No. 2012T50711).

## REFERENCES

- [1] PASTOR-SATORRAS R., CASTELLANO C., VAN MIEGHEM P. and VESPIGNANI A., arXiv:1408.2701 (2014).
- [2] NEWMAN M. E. J., *Networks - An Introduction* (Oxford University Press, New York) 2010.
- [3] PASTOR-SATORRAS R. and VESPIGNANI A., *Phys. Rev. E*, **63** (2001) 066117.
- [4] KEMPE D., KLEINBERG J. and TARDOS É., in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM) 2003, pp. 137–146.
- [5] LESKOVEC J., KRAUSE A., GUESTRIN C., FALOUTSOS C., VANBRIESEN J. and GLANCE N., in *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM) 2007, pp. 420–429.
- [6] KITSACK M., GALLOS L. K., HAVLIN S., LILJEROS F., MUCHNIK L., STANLEY H. E. and MAKSE H. A., *Nat. Phys.*, **6** (2010) 888.
- [7] ZHANG J., XU X.-K., LI P., ZHANG K. and SMALL M., *Chaos*, **21** (2011) 016107.
- [8] BORGE-HOLTHOEFER J., RIVERO A. and MORENO Y., *Phys. Rev. E*, **85** (2012) 066123.
- [9] CHEN D., LÜ L., SHANG M.-S., ZHANG Y.-C. and ZHOU T., *Physica A*, **391** (2012) 1777.
- [10] ZENG A. and ZHANG C.-J., *Phys. Lett. A*, **377** (2013) 1031.
- [11] DE ARRUDA G. F., BARBIERI A. L., RODRÍGUEZ P. M., RODRIGUES F. A., MORENO Y. and COSTA L. DA F., *Phys. Rev. E*, **90** (2014) 032812.
- [12] NEWMAN M. E. J., *SIAM Rev.*, **45** (2003) 167.
- [13] FREEMAN L. C., *Sociometry*, **40** (1977) 35.
- [14] ESTRADA E. and RODRIGUEZ-VELAZQUEZ J. A., *Phys. Rev. E*, **71** (2005) 056103.
- [15] MASLOV S. and SNEPPEN K., *Science*, **296** (2002) 910.
- [16] SABIDUSSI G., *Psychometrika*, **31** (1966) 581.
- [17] PEI S., MUCHNIK L., ANDRADE J. S. jr., ZHENG Z. and MAKSE H. A., *Sci. Rep.*, **4** (2014) 5547.
- [18] LIU Y., TANG M., ZHOU T. and DO Y., arXiv:1409.5187 (2014).
- [19] HU Z.-L., LIU J.-G., YANG G.-Y. and REN Z.-M., *EPL*, **106** (2014) 18002.
- [20] CHEN W., WANG Y. and YANG S., in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM) 2009, pp. 199–208.
- [21] CHEN W., WANG C. and WANG Y., in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM) 2010, pp. 1029–1038.
- [22] BOLLOBÁS B., *Modern Graph Theory, Graduate Texts in Mathematics*, Vol. **184** (Springer) 1998.
- [23] APPEL K., HAKEN W. *et al.*, *Illinois J. Math.*, **21** (1977) 429.
- [24] APPEL K., HAKEN W., KOCH J. *et al.*, *Illinois J. Math.*, **21** (1977) 491.
- [25] RIIHIJÄRVI J., PETROVA M. and MÄHÖNEN P., *Frequency allocation for WLANs using graph colouring techniques*, in *Proceedings of WONS*, Vol. **5**, *Second Annual Conference on IEEE 2005*, pp. 216–222.
- [26] LEIGHTON F. T., *Nat. Bur. Standard*, **84** (1979) 489.
- [27] WELSH D. J. and POWELL M. B., *Comput. J.*, **10** (1967) 85.
- [28] KLOTZ W., *Math. Rep.*, **5** (2002) 1.
- [29] CASTELLANO C. and PASTOR-SATORRAS R., *Phys. Rev. Lett.*, **96** (2010) 038701.
- [30] YANG R., HUANG L. and LAI Y.-C., *Phys. Rev. E*, **78** (2008) 026111.
- [31] BORGE-HOLTHOEFER J. and MORENO Y., *Phys. Rev. E*, **85** (2012) 026116.
- [32] BARABÁSI A.-L. and ALBERT R., *Science*, **286** (1999) 509.
- [33] GONG K., TANG M., YANG H. and SHANG M., *Chaos*, **21** (2011) 043130.
- [34] SHU P., TANG M., GONG K. and LIU Y., *Chaos*, **22** (2012) 043124.
- [35] SHU P., WANG W., TANG M. and DO Y., arXiv:1410.0459 (2014).
- [36] XIE N., *Social network analysis of blogs*, PhD Thesis, MSc Dissertation, University of Bristol (2006).
- [37] GUIMERA R., DANON L., DIAZ-GUILERA A., GIRALT F. and ARENAS A., *Phys. Rev. E*, **68** (2003) 065103.
- [38] HUANG B., ZHAO X.-Y., QI K., TANG M. and DO Y., *Acta Phys. Sin.*, **62** (2013) 218902.