

# Identifying Influential nodes in complex network based on local k-shell iteration factor

DiHua Wu<sup>1</sup>, Liming Zheng<sup>1</sup>

<sup>1</sup>(College of information science and technology, Jinan University, Guangzhou, China)

\*\*\*\*\*

## Abstract:

Identifying vital nodes in complex network is significant to promote or control the Information dissemination process. Centrality is the measure of evaluating the influence of various nodes and current research find that neighbourhood centrality outperforms the basic centrality. Inspired by the idea of closeness centrality and k-shell iteration factor, we proposed a new method to estimate the spreading ability of node based on the sum of k-shell iteration factor of neighbour nodes and the shortest distance between two nodes. To evaluate the effectiveness of the proposed method, Susceptible-Infected-Recovered (SIR) epidemic model is utilized to simulate the spreading process on six real world networks and the experimental results exhibit our method is more accurate than the existing measures.

**Keywords** —Complex networks, Influential nodes, k-shell iteration factor, Neighbour influence

\*\*\*\*\*

## I. INTRODUCTION

Many networks in different fields can be described and analysed by complex network model, such as power grid [1],[2], social network [3][4] and biological network [5],[6], etc. The network topology actually reflects the connected relation among different entities and the nodes in the network show different spreading capacity because of the diversities of the physical positions and connectivity. The very fact that the number of key nodes which have a great impact on the whole network is generally smaller than the common ones [7],[8]. Therefore, the correct identifying of the set of influential nodes has the important practical significance, which can effectively help people control spreading process like cascading failures [9], viral marketing [4],[10], news or rumours diffusion [11], pandemics [12] and so on.

Centrality is the measure of evaluating the influence of various nodes in a network. Nodes with higher centrality index are viewed as more important under the same assessment measure. A variety of centrality measures have been constructed in the past decades. According to the scope of the topology, all the methods can be classified into local metrics, global metrics and

hybrid metrics. Local metrics are focused on the local information from target nodes as well as their neighbour nodes, such as degree centrality (DC) [13] and H-index [14]. However, the local metrics generate more inaccurate ranking because of the limitation of local information. To solve this problem, global metrics such as betweenness centrality (BC) [15] and closeness centrality (CC) [16] are proposed and the results shows the better performance than the traditional local metric, but these metrics suffer from high computational complexity, which can not apply to large scale networks. Recently, Kitsak et al. [17] applied the k-shell decomposition method [18]-[20] to identifying the influential nodes, demonstrating the proposed method outperforms the degree or the betweenness centrality in most cases. They argued that the most efficient spreaders depend on their location in the network and the nodes with the largest k-shell index are deemed core of the network. However, recent research found that after k-shell decomposition, nodes with different spreading capacity may be assigned in the same k-shell value [21]-[23]. To overcome this shortcoming, Zeng et al. [21] put forward a mixed degree decomposition (MDD) procedure balance between the residual degree and the exhausted degree. However, an adjustment parameter need to be set in advance to fit the networks with various structures. Benefit from

the low computational complexity of k-shell method, more and more hybrid metrics on the foundation of this approach are designed in 3 to further improve the accuracy of node importance ranking. Bae et al. [24] presented neighbourhood coreness centrality ( $C_{nc}$ ) by taking into consideration the sum of k-shell value of the nearest neighbour nodes. Wang et al. [23] noted that the iteration information is generated but neglected while in k-shell decomposition. They constructed a hybrid method called k-shell iteration factor (KS-IF) on account of the degree and k-shell iteration information refer to both the node and their nearest neighbour nodes. Yang et al [25] also found the effect of iteration information and proposed weighted degree centrality (WD) and extended weighted degree centrality (EWD). All of these methods only allow for the properties of the node itself and the nearest neighbour nodes. Liu et al. [26] further investigated the spreading influence from not only the nearest neighbours but also indirect neighbours. Based on degree and k-shell index, experiments for neighbour nodes from different step lengths are conducted. The results indicate nodes within two steps can achieve a great balance between accuracy and time cost. However, an adjustable parameter are set manually as well and therefore not robust to diverse network structures.

The aboved-mentioned studies show that how to utilize the information from neighbour nodes has become one of the key point of identifying the influential nodes. In this letter, we continue to further investigate the performance of k-shell method based on neighbour nodes. Inspired by the idea of closeness centrality, we consider a further range of neighbour nodes whose influence refer to the shortest distance from central node. The initial influence of neighbour nodes with the same distance is denoted as the sum of their value of iteration factor. Further, the epidemic threshold value  $\beta_c$  in Susceptible-Infected-Recovered (SIR) epidemic model [27] is regarded as the one-step weighted parameter of influence in order to evaluate the difficulty of diffusion in a given network. Finally, assuming the neighbour influence of nodes is mainly relevant to the shortest distance, the weighted factor of each neighbour node decrease exponentially with increasing distance from the central node. Given these, we proposed a novel measure, k-shell iteration factor closeness centrality (IFC), to identify influential nodes for different types of network. To evaluate the effectiveness of the proposed method, SIR epidemic model is utilized to simulate the spreading process on real world network datasets and the experimental

results exhibit our method is more accurate than the existing measures.

The remainder of this letter is structured as follows: In Section 2, motivation is described and the new method is proposed. In Section 3, SIR simulation and experiments on real world network datasets are set up to determine the optimal distance to other nodes and evaluate the performance of the proposed method in comparison with other centrality measures. Finally, we summarize the conclusion of this letter in Section 4.

## II. PROPOSED METHOD

The existing studies show that the neighbour information plays an important role in improving the performance of evaluating the spreading capacity of nodes. That is to say, the influential node would be the ones which lie close to influential neighbours. But most measures only allow for the node and the nearest neighbour information. Therefore, we further investigate the impact from further neighbour nodes. Based on the assumption that the neighbour nodes with larger distance are less likely to be infected, the weight values should be assigned to neighbour nodes with different distances.

Commonly, the epidemic threshold in SIR model is used to characterize whether the local disease is able to develop into an epidemic in the network. Furthermore, if a constant transmission probability is given, when the threshold of network is large, it indicates that it is more difficult to widely spread in the given network for the disease, while the disease can be easily outbreak over the network that with small threshold. Therefore, epidemic threshold can be seen as the measure of the difficulty of disease spreading in different networks. In view of the above analysis, we introduce the epidemic threshold as the weight value of one-step propagation in a given network, and the weight values decreases exponentially with the increase of distance. Besides, k-shell iteration factor  $\delta$ [23] is employed to distinguish the initial influence of neighbour nodes, where  $m$  is the maximum number of iterations in the step and  $n$  is the current iteration when node  $i$  is removed:

$$\delta(i) = ks(1 + \frac{n}{m}) \quad (1)$$

At last, a new measure, the iteration factor closeness centrality (IFC), based on both iteration and neighbour information is proposed. The node  $i$  in the network is defined as:

$$IFC(i) = \sum_{j \in N(i)} \beta_c^{d_{ij}} \cdot \delta(j) \quad (2)$$

where  $\beta_c$  is equal to  $\langle k \rangle / (\langle k^2 \rangle - \langle k \rangle)$  [27] and  $\langle k \rangle$ ,  $\langle k^2 \rangle$  denote the average degree and the mean square of degree.  $N(i)$  represents the set of neighbour nodes whose distance less or equal to a given radius  $r$ .  $r=2$  represents only the nearest neighbours and next nearest neighbours are considered. Further, the extended iteration factor closeness centrality (EIFC) by considering the IFC of the nearest neighbour nodes, which is calculated by:

$$EIFC(i) = \sum_{j \in \Gamma(i)} IFC(j) \quad (3) \quad \text{Take the node}$$

2 in Fig.1 as the example, the epidemic threshold  $\beta_c \approx 0.43$ . If given the radius  $r=2$ ,  $IFC(2) = 0.43^1 \cdot [\delta(1) + \delta(3) + \delta(4) + \delta(9) + \delta(16)] + 0.43^2 \cdot [\delta(5) + \delta(6) + \delta(10) + \delta(17) + \delta(18) + \delta(19) + \delta(20)] \approx 13.01$ ,  $EIFC(2) = IFC(1) + IFC(3) + IFC(4) + IFC(9) + IFC(16) \approx 52.71$ .

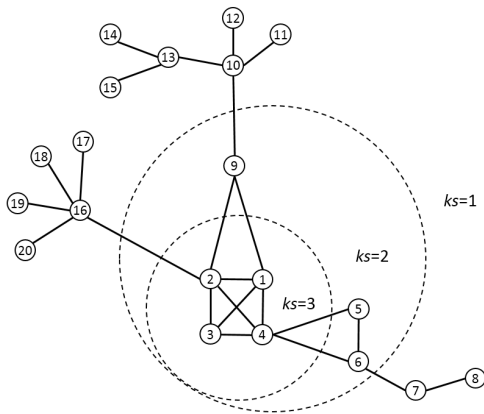


Fig. 1. Schematic network

TABLE I

Some statistical properties of datasets: node number (n), edge number(m), average degree ( $\langle k \rangle$ ), maximum degree ( $k_{max}$ ), epidemic threshold ( $\beta_c$ ), average clustering coefficient (C), assortativity (A), diameter (D).

Network	n	m	$\langle k \rangle$	$k_{max}$	$\beta_c$	C	A	D
Karate	34	78	4.589	17	0.148	0.571	0.476	5
Dolphins	62	159	5.129	12	0.172	0.259	0.044	8
Netscience	379	914	4.823	34	0.142	0.741	0.082	17
Email	1133	5451	9.622	71	0.057	0.220	0.078	8
Power	4941	6594	2.669	19	0.348	0.080	0.003	46
PGP	10680	24316	4.554	205	0.056	0.266	0.238	24

### III. EXPERIMENTAL RESULTS

#### A. The SIR epidemic model

In this section, the susceptible-infected-recovered (SIR) is implemented to evaluate the spreading influence of nodes on real world networks. In each time step, there is only one state for a node including susceptible, infected, recovered. Before the infection, a given node  $i$  is preset as infected while all other nodes are in the state of susceptible. Next, the infectious nodes will infect each of their susceptible nearest neighbour nodes with probability  $\beta$  and then recover with probability  $\gamma$ . The recovered nodes cannot be infected again. The propagation process continues until no more infected ones left in the network. Without loss of generality, we set  $\gamma=1$ . At last, the independent simulation experiment is run for 1000 times in which the average number of recovered number is regarded as the spreading influence of node  $i$ .

#### B. Datasets

For the purposed of validating the performance of proposed methods, a schematic network and the real networks are used, including:

- Karate: social network of a karate club [28]
- Dolphins: dolphins social network [29]
- Netscience: collaboration network of scientists [30]
- Email: E-mail network of the University RoviraiVirgili **Error! Reference source not found.**
- Power: Western States Power Grid [32]
- PGP: an encrypted communication network [33]

Some statistical properties of all the networks are listed in Table II.

#### C. Pretreatment

In Ref. [26], the authors found that there exists a saturation effect that is negative for ranking performance, which is considered as the result of the radius of neighbour nodes set too large. Moreover, the computational complexity of the measure will increase quickly when a more extensive range of neighbour information is utilized. Therefore, to balance the accuracy and time cost, the parameter  $r$  of the proposed method need to be determined previously.

The Kendall's tau coefficient [34] is widely used to measure the rank correlation between two ordered lists. Suppose two equal-sized ranking lists X and Y are given, and the joint set is XY. For any two selected elements  $(x_i, y_i)$  and  $(x_j, y_j)$ , if  $x_i > x_j$ ,

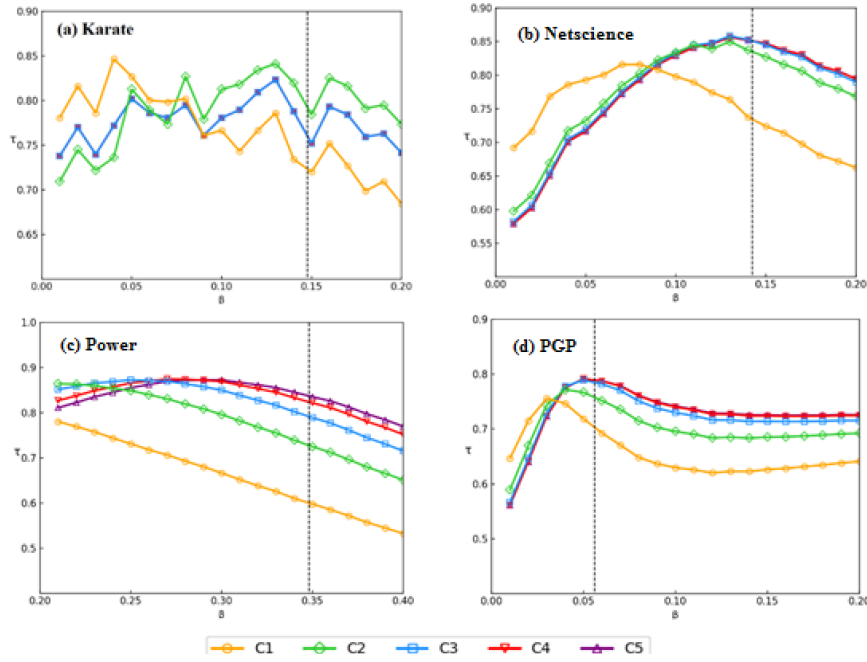


Fig.2. The effect on Kendall's tau when the value of parameter r of IFC is set from 1 to 5.

and  $y_i > y_j$  or  $x_i < x_j$  and  $y_i < y_j$ , the pairs are considered concordant while they are discordant if  $x_i > x_j$  and  $y_i > y_j$  or  $x_i < x_j$  and  $y_i < y_j$ . Moreover, if  $x_i = x_j$  and  $y_i = y_j$ , the pairs are considered neither concordant nor discordant. The Kendall's tau coefficient is defined as:

$$\tau = \frac{n_c - n_d}{0.5n(n-1)} \quad (4)$$

where  $n_c$  and  $n_d$  are the number of concordant pairs and discordant pairs, respectively.  $n$  is the total number of pairs.

To study the optimal radius of IFC, we conduct experiments on four datasets with different sizes by considering the shortest distance  $r \leq 5$  to neighbour nodes, which are denoted C1, C2, C3, C4 and C5 as Eq.(2), respectively. Then Kendall's tau coefficient is introduced to calculate the correlation of ranking lists obtained from C1~C5 and SIR model by varying the infection probability  $\beta$ . The infection probability  $\beta$  is set nearby the value of epidemic threshold  $\beta_c$  in order to easily distinguish vital nodes. The larger value of  $\tau$  corresponds to more accurate results of measure ranking nodes. The comparative results are shown in Fig. 2.

As is shown in Fig.2, when the propagation probability is small, such as  $\beta < 0.07$ , C1 performs better than C2~C5 as the spreading process is confined to local areas and cannot spread outward.

Then with the increase of transmission probability, when  $\beta > \beta_c$  (the dot line), disease can disseminate on a large scale with more neighbour nodes acting as the transmitter of the disease, so C4, C5 have better accuracy by considering wider neighbour information. In addition, there also exists the saturation effect that with the increase of node distance, the improvement of evaluation ability is limited. The reason is that on the one hand, especially in small-scale networks (Fig.2 (a), (b)), the network diameter is small and the distance between nodes is limited, so C3, C4, C5 will achieve the similar experimental results. On the other hand, as the definition of IFC shown by Eq.(2), the weight values of neighbour nodes with  $r=1,2,3\dots$  are respectively assigned  $\beta_c^1, \beta_c^2, \beta_c^3\dots$ . The smaller the  $\beta_c$  of the network is, the faster the weight value of the remote node decreases. The purpose is that when the epidemic threshold is small, more attention should be paid to the initial spreading region of nodes, while the weight ratio of remote neighbours get higher in the network with larger epidemic threshold. By the consideration of accuracy and time consumption, we set the value of  $r=3$  in IFC and EIFC, and comparison experiments among all the measures are conducted in the later part.

**D. Results**

To further evaluate our methods' performance of identifying influential nodes, IFC and EIFC have been compared to some other representative measures, including degree centrality (DC), closeness centrality (CC) [16], k-shell decomposition (KS) [17]-[20], Mixed degree centrality (MDD) [21], extended neighbourcoreness centrality ( $C_{nc+}$ ) [24], k-shell iteration factor centrality (KS-IF) [23], extended weighted degree centrality (EWD) [25], whose definitions are shown in section 2. Moreover, we view the iteration factor (IF), part of KS-IF method, as an independent measure so as to compare the accuracy with that of k-shell decomposition.

Similarly, we investigate the accuracy of each measure by computing the value of  $\tau$  with infection probability varied around  $\beta_c$ . The results obtained from this experiment on all six real datasets are shown in Fig.3. We can observe that both IFC and EIFC can get higher values of  $\tau$  in most cases, especially in the case of  $\beta > \beta_c$ . In comparison with k-shell method, IF offers higher correctness in all the datasets except for Email. The reason is that the average of degree in Email network is large enough that the iteration information plays a less important role. But overall, iteration information is also a plus for identifying vital nodes in the network.

In addition, nodes with different spreading capacity in the network may get the same score measured by a given method, which is the case we do not expect to see. Therefore, the capacity of differentiating nodes is one of the criteria for evaluating the performance of centrality methods as well. Here, the monotonicity  $M$  [24] is employed to evaluate the resolution of different ranking measures. The monotonicity is calculated by:

$$M(R) = \left( 1 - \frac{\sum_{r \in R} n_r(n_r - 1)}{n(n-1)} \right)^2 \quad (5)$$

where  $R$  is the given ranking list,  $n$  is the number of ranks and  $n_r$  is the number of the same ranks of  $R$ . Obviously,  $M$  ranges from 0 to 1, and  $M$  reaches the maximum value 1 only when the rank

of an element vary from each other. The ranks of nodes is viewed as the element of ranking lists and a larger value of  $M$  indicates the better capacity of discriminating nodes of the metrics. The experimental results shown in Table IIIIV exhibit that both IFC and EIFC are more suitable for the discrimination in specification of node than the other methods. Moreover, IF method also outperforms k-shell method in terms of monotonicity. The high values of monotonicity of proposed methods come from taking into account more extensive neighbour information as well as iteration information, which based on the idea of closeness centrality and iteration factor.

To better understand the distribution of nodes in each rank, complementary cumulative distribution function (CCDF) of six real networks are plotted in Fig.3. CCDF can be calculated by:

$$CCDF(r) = \frac{n - \sum_{i=1}^r n_i}{n} \quad (6)$$

where  $n$ ,  $n_i$  denotes the total number of nodes and the number of nodes ranking at  $i$ , respectively.  $r$  denotes the given ranking number in the list. The faster the curve of CCDF falls to zero, indicating that the nodes are distributed in the smaller ranking interval. As observed in Fig. 4, the CCDF curve of IFC, EIFC, KS-IF and EWD trend are basically the same in small size networks. However, results of Power and PGP network show that networks with larger size can cause the discrimination of KS-IF and EWD to weaken, while our proposed methods still perform the best. Next, the time complexity of the proposed measures is analysed. The calculation procedure of IFC consists of three parts: first, to compute the epidemic threshold in each network we need time  $O(n)$ . Then, obtaining the iteration factor values of nodes requires  $O(m)$ , whose procedure is based on the algorithm of k-shell decomposition[35]. Last, summing over the iteration factor of neighbour nodes within a radius of 3 runs in  $O(\langle k \rangle^3 n)$ . Therefore, the total time complexity of IFC is  $O(n+m+\langle k \rangle^3 n) \approx O(\langle k \rangle^3 n)$ . EIFC is computed by the sum of IFC from nearest neighbour, which requires total time complexity  $O(n+m+\langle k \rangle^3 n+\langle k \rangle n) \approx O(\langle k \rangle^3 n)$ . At last, the time complexity of eight



TABLE VVI  
THE MONOTONICITY OF EACH MEASURES ON SIX REAL WORLD NETWORKS

Network	M(DC)	M(CC)	M(KS)	M(IF)	M(MDD)	M(C <sub>nc+</sub> )	M(KS-IF)	M(EWD)	M(IFC)	M(EIFC)
Karate	0.7079	0.8993	0.4958	0.6463	0.7536	0.9472	0.9542	0.9542	0.9542	<b>0.9577</b>
Dolphins	0.8312	0.9737	0.3769	0.8781	0.9091	0.9873	0.9968	0.9958	<b>0.9979</b>	<b>0.9979</b>
Netscience	0.7642	0.9928	0.6421	0.7606	0.8226	0.9893	0.9948	0.9943	0.9951	<b>0.9955</b>
Email	0.8874	0.9988	0.8088	0.8977	0.9233	0.9991	0.9996	<b>0.9999</b>	<b>0.9999</b>	<b>0.9999</b>
Power	0.5927	0.9998	0.2460	0.6234	0.6940	0.9419	0.9848	0.9944	0.9998	<b>0.9999</b>
PGP	0.6193	0.9996	0.4806	0.6487	0.6679	0.9851	0.9912	0.9985	<b>0.9997</b>	<b>0.9997</b>

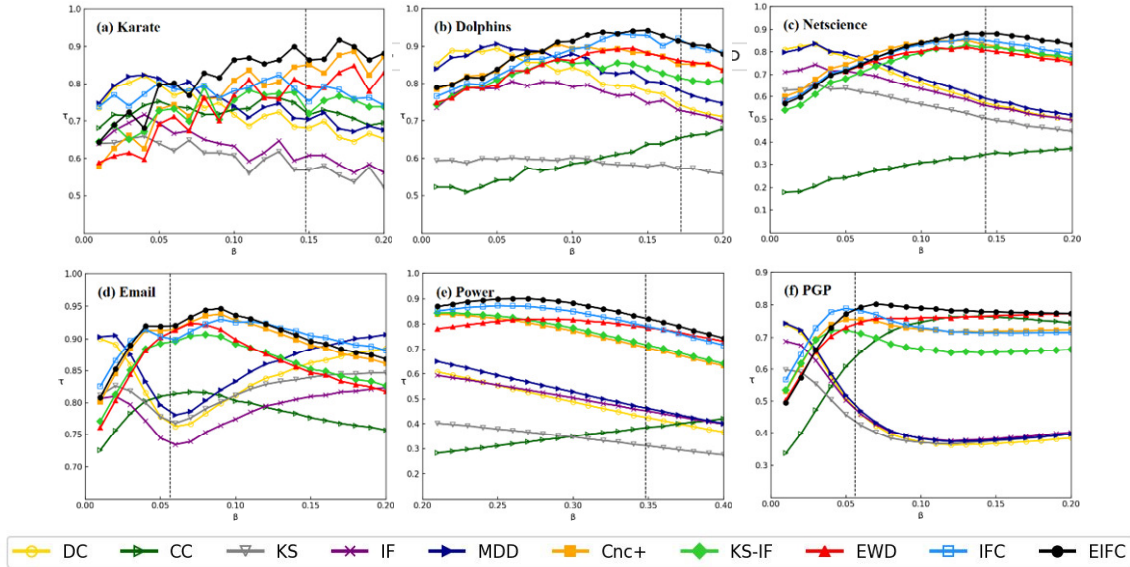


Fig. 3. The Kendall's tau between the ranking lists offered by eight measures and the ranking list generated by SIR simulation with the infected probability  $\beta$  varied around  $\beta_c$  (dot line) in six real networks.

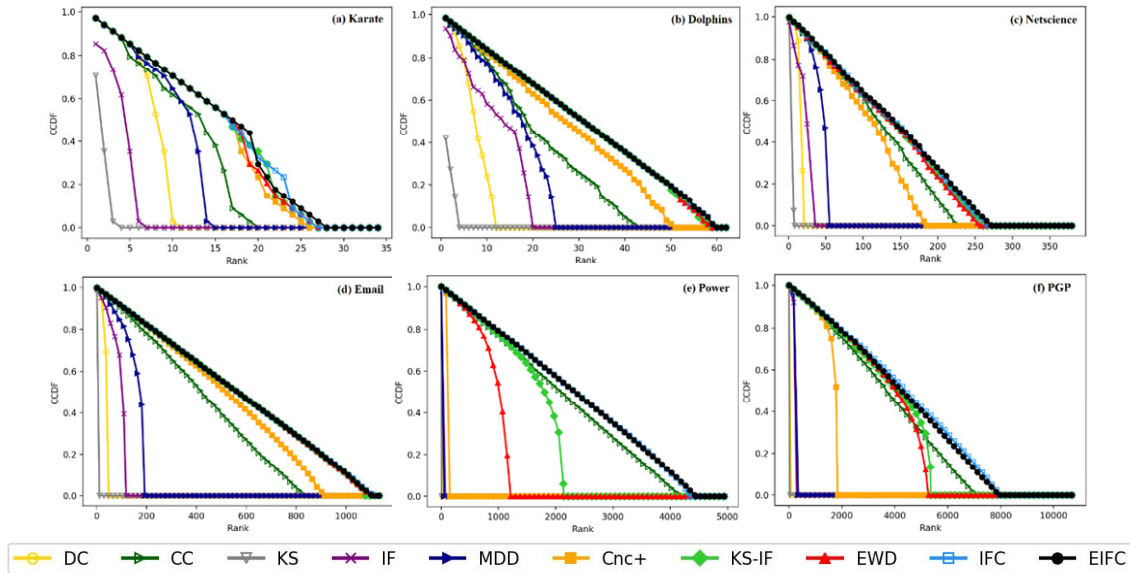


Fig. 4. The complementary cumulative distribution function (CCDF) plots for six real networks.

methods is shown in Table VIIVIIIIX. Due to consideration of more neighbour information, both IFC and EIFC cost slightly more computational complexity than  $C_{nc+}$ , KS-IF and EWD when the average of degree is small. In addition, in comparison with CC, the method of taking account into the neighbour nodes with shortest distance as well, our methods greatly increase the efficiency by considering limited neighbour influence.

TABLE XXIXII  
The computational complexity of eight methods

Ranking measure	Computation complexity	Category
CC	$O(n^3)$ or $O(n^2 \log n+nm)$	Global measure
ks	$O(m)$	Global measure
IF	$O(m)$	Global measure
$C_{nc+}$	$O((k)n)$	Hybrid measure
KS-IF	$O((k)n)$	Hybrid measure
EWD	$O((k)n)$	Hybrid measure
IFC	$O((k)^3n)$	Hybrid measure
EIFC	$O((k)^3n)$	Hybrid measure

#### IV. CONCLUSION

How to identify the vital nodes in complex networks effectively has become the research hotspots in recent years. In this letter, some related centrality measures are introduced which includes classical methods like degree centrality, betweenness centrality, closeness centrality, and some hybrid measures based on k-shell decomposition method. The existing methods have confirmed that neighbour information has a great effect on identifying vital nodes. In closeness centrality, the shortest distance between nodes is considered to be an affected factor in propagation process. In this letter, more extensive neighbour nodes are further investigated in consideration of the negative effect on increasing distance to the central node. Then, the epidemic threshold is introduced as the attenuation factor in spreading process in order to adapt to propagation attenuation in different networks. Last, by combining distance and iteration information of neighbour nodes, a new hybrid measures named iteration factor closeness centrality (IFC) and its expansion forms are presented. Firstly, experiments are carried out on four real datasets with different scales to verify that the optimal neighbourhood radius is 3. Furthermore, the proposed methods are compared with the existing method on six real world networks in term of three aspects, including accuracy, discrimination and computational complexity. The experimental results show that our methods outperform other

existing methods and verify the positive effect of iteration information from k-shell decomposition on identifying key nodes. Finally, we analyse the computational complexity, the proposed methods improve the performance of CC and remain competitive in the existing methods. How to apply the proposed approach to specifying and ranking the spreading capacity of nodes in weighted and directed networks will be investigated in the future works.

#### References

- [1] Kinney R, Crucitti P, Albert R, et al. Modeling cascading failures in the North American power grid [J]. The European Physical Journal B - Condensed Matter and Complex Systems, 2005, 46(1):101-107.
- [2] Buldyrev S V, Parshani R, Paul G, et al. Catastrophic cascade of failures in interdependent networks [J]. Nature, 2010, 464(7291):1025-1028.
- [3] Albert R, Jeong H, Barabási A-L. Error and attack tolerance of complex networks [J], Nature, 2000 406(6794):378–382.
- [4] Hinz O, Schulze C, Takac C. New product adoption in social networks: Why direction matters [J]. Journal of Business Research, 2014, 67(1): 2836-2844.
- [5] Cui X, Li X, Li J, et al. inFRank: a ranking-based identification of influential genes in biological networks [J]. Oncotarget, 2017, 8(27): 43810.
- [6] Zemanová, Lucia, Zhou C, Kurths, Jürgen. Structural and functional clusters of complex brain networks [J]. Physica D Nonlinear Phenomena, 2006, 224(1):202-212.
- [7] Pastor-Satorras R, Vespignani A. Immunization of complex networks [J]. Physical Review E, 2002, 65(3): 036104.
- [8] Cohen R, Erez K, Ben-Avraham D, et al. Breakdown of the Internet under intentional attack [J]. Physical review letters, 2001, 86(16): 3682.
- [9] Motter A E, Lai Y C. Cascade-based attacks on complex networks [J]. Physical Review E, 2002, 66(6):065102.
- [10] Yang J, Yao C, Ma W, et al. A study of the spreading scheme for viral marketing based on a complex network model [J]. Physica A Statistical Mechanics & Its Applications, 2010, 389(4):859-870.
- [11] Doer B, Fouz M, Friedrich T. Why rumors spread so quickly in social networks [J]. Communications of the ACM, 2012, 55(6):70.
- [12] Pastor-Satorras R, Vázquez, Alexei, Vespignani A. Dynamical and Correlation Properties of the Internet [J]. Physical Review Letters, 2001, 87(25):258701.
- [13] Freeman L C. Centrality in Social Networks' Conceptual Clarification [J]. Social Networks, 1979, 1(3):215-239.
- [14] Lü, Linyuan, Zhou T, Zhang Q M, et al. The H-index of a network node and its relation to degree and coreness [J]. Nature Communications, 2016, 7:10168.
- [15] Freeman L C. Centrality in Social Networks' Conceptual Clarification [J]. Social Networks, 1979, 1(3):215-239.
- [16] Sabidussi G. The Centrality Index of a Graph [J]. Psychometrika, 1966, 31(4):581-603.
- [17] Kitsak M, Gallos L K, Havlin S, et al. Identification of influential spreaders in complex networks [J]. Nature Physics, 2010, 6(11): 888-893.

- [18] Haviland J, Thomason A. Pseudo-Random Hypergraphs [J]. *Annals of Discrete Mathematics*, 1989, 75(1):255-278.
- [19] Seidman S B. Network structure and minimum degree [J]. *Social Networks*, 1983, 5(3):269.
- [20] Carmi S, Havlin S, Kirkpatrick S, et al. A model of Internet topology using k-shell decomposition [J]. *Proceedings of the National Academy of Sciences*, 2007, 104(27): 11150-11154.
- [21] Zeng A, Zhang C J. Ranking spreaders by decomposing complex networks [J]. *Physics Letters A*, 2013, 377(14):1031-1035.
- [22] Liu J G, Ren Z M, Guo Q. Ranking the spreading influence in complex networks [J]. *Physica A: Statistical Mechanics and its Applications*, 2013, 392(18):4154-4159.
- [23] Wang Z, Zhao Y, Xi J, et al. Fast ranking influential nodes in complex networks using a k-shell iteration factor [J]. *Physica A Statistical Mechanics & Its Applications*, 2016, 461:171-181.
- [24] Bae J, Kim S. Identifying and ranking influential spreaders in complex networks by neighbourhoodcoreness [J]. *Physica A: Statistical Mechanics and its Applications*, 2014, 395:549-559.
- [25] Fan Y, Li X, Xu Y, et al. Ranking the spreading influence of nodes in complex networks: An extended weighted degree centrality based on a remaining minimum degree decomposition [J]. *Physics Letters A*, 2018:S0375960118305632-.
- [26] Liu Y, Tang M, Zhou T, et al. Identify influential spreaders in complex networks, the role of neighbourhood [J]. *Physica A: Statistical Mechanics and its Applications*, 2016, 452: 289-298.
- [27] Moreno Y, Pastor-Satorras R, Vespignani A. Epidemic outbreaks in complex heterogeneous networks [J]. *The European Physical Journal B-Condensed Matter and Complex Systems*, 2002, 26(4): 521-529.
- [28] Zachary W W. An information flow model for conflict and fission in small groups [J]. *Journal of anthropological research*, 1977, 33(4): 452-473.
- [29] Lusseau D, Schneider K, Boisseau O J, et al. The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations [J]. *Behavioral Ecology and Sociobiology*, 2003, 54(4): 396-405.
- [30] Newman M E J. Finding community structure in networks using the eigenvectors of matrices [J]. *Physical review E*, 2006, 74(3): 036104.
- [31] Guimerà R, Danon L , DíazGuilera A, et al. Self-similar community structure in a network of human interactions [J]. *Phys Rev E Stat Nonlin Soft Matter Phys*, 2003, 68(6 Pt 2):065103.
- [32] Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks [J]. *Nature*, 1998, 393(6684): 440.
- [33] Boguñá M, Pastor-Satorras R, Díaz-Guilera A, et al. Models of social networks based on social distance attachment [J]. *Physical review E*, 2004, 70(5): 056122.
- [34] Kendall M G. A new measure of rank correlation [J]. *Biometrika*, 1938, 30(1/2): 81-93.
- [35] Batagelj V, Zaversnik M. An O(m) algorithm for cores decomposition of networks [J]. *arXiv preprint cs/0310049*, 2003.