RESEARCH ARTICLE OPEN ACCESS

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

DiHua Wu[1], Liming Zheng[1]

[1] (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 theinfluence of various nodes in a network. Nodes withhigher centrality index are viewed as moreimportant under the same assessment measure. Avariety of centrality measures have beenconstructed 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 identifyingthe influential nodes, demonstrating the proposedmethod outperforms the degree or the betweennesscentrality 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 fromare deemed core of the network. However, recentresearch found that after k-shell decomposition,nodes with different spreading capacity may beassigned in the same k-shell value [21-23]. Toovercome this shortcoming, Zeng et al. [21] putforward a mixed degree decomposition (MDD)procedure balance between the residual degree andthe exhausted degree. However, an adjustmentparameter need to be set in advance to fit thenetworks with various structures. Benefit from the low computational complexity of k-shell method, more and more hybrid metrics on thefoundation 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 howto utilize the information from neighbour nodes hasbecome one of the key point of identifying theinfluential nodes. In this letter, we continue tofurther investigate the performance of k-shellmethod based on neighbour nodes. Inspired by theidea of closeness centrality, we consider a furtherrange of neighbour nodes whose influence refer tothe shortest distance from central node. The initialinfluence of neighbour nodes with the samedistance is denoted as the sum of their value ofiteration factor. Further, the epidemic thresholdvalue β_c in Susceptible-Infected-Recovered (SIR)epidemic model [27] is regarded as the one-stepweighted parameter of influence in order toevaluate the difficulty of diffusion in a givennetwork. Finally, assuming the neighbour influenceof nodes is mainly relevant to the shortest distance,the weighted factor of each neighbour nodedecrease exponentially with increasing distancefrom the central node. Given these, we proposed anovel measure, k-shell iteration factor closenesscentrality (IFC), to identify influential nodes fordifferent types of network. To evaluate theeffectiveness of the proposed method, SIR epidemicmodel is utilized to simulate the spreading processon real world network datasets and the experimentalresults exhibit our method is more accurate than theexisting 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 neighbourinformation plays an important role in improvingthe performance of evaluating the spreadingcapacity of nodes. That is to say, the influentialnode would be the ones which lie close toinfluential neighbours. But most measures onlyallow for the node and the nearest neighbour information. Therefore, we further investigate theimpact from further neighbour nodes. Based on theassumption that the neighbour nodes with largerdistance are less likely to be infected, the weightvalues should be assigned to neighbour nodes withdifferent distances.

Commonly, the epidemic threshold in SIRmodel is used to characterize whether the localdisease is able to develop into an epidemic in thenetwork. Furthermore, if a constant transmissionprobability is given, when the threshold of networkis large, it indicates that it is more difficult towidely spread in the given network for the disease,while the disease can be easily outbreak over thenetwork that with small threshold. Therefore,epidemic threshold can be seen as the measure ofthe difficulty of disease spreading in differentnetworks. In view of the above analysis, weintroduce the epidemic threshold as the weightvalue of one-step propagation in a given network,and the weight values decreases exponentially withthe increase of distance. Besides, k-shell iterationfactor δ[23] is employed to distinguish the initialinfluence of neighbour nodes, where m is themaximum number of iterations in the step and n isthe current iteration when node i is removed:

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

At last, a new measure, the iteration factorcloseness centrality (IFC), based on both iterationand neighbour information is proposed. The node$I$ inthe network is defined as:

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

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

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

Take the node2 in Fig.1 as the example, the epidemic threshold $\beta_c\approx0.43$. If given the radius r=2, IFC(2)=$0.43^1\cdot[\delta(1)+\delta(3)+\delta(4)+\delta(9)+\delta(16)]+0.43^2\cdot[\delta(5)+\delta(6)+\delta(10)+\delta(17)+\delta(18)+\delta(19)+\delta(20)]\approx13.01$, EIFC(2)=IFC(1)+IFC(3)+IFC(4)+IFC(9)+IFC(16)≈ 52.71.

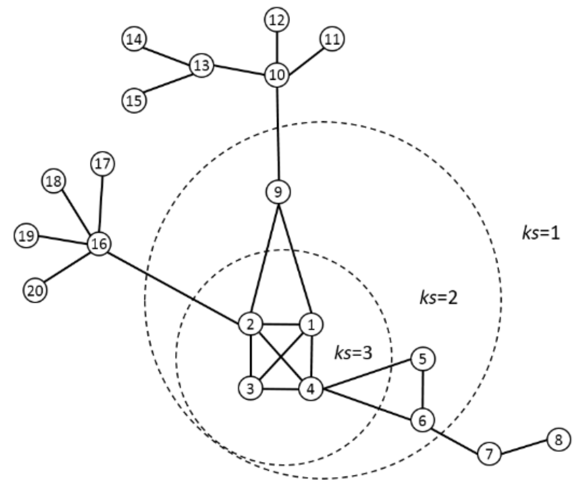

Fig 1. Schematic network

III. EXPERIMENT RESULT

A. The SIR epidemic model

In this section, the susceptible-infected-recovered (SIR) is implemented to evaluate thespreading influence of nodes on real worldnetworks. In each time step, there is only one statefor a node including susceptible, infected, recovered. Before the infection, a given node i ispreset as infected while all other nodes are in thestate of susceptible. Next, the infectious nodes willinfect each of their susceptible nearest neighbournodes with probability β and then recover withprobability γ. The recovered nodes cannot beinfected again. The propagation process continuesuntil no more infected ones left in the network.Without loss of generality, we set γ=1. At last, theindependent simulation experiment is run for 1000times in which the average number of recoverednumber is regarded as the spreading influence ofnode i.

B. Datasets

For the purposed of validating the performanceof proposed methods, a schematic network and thereal 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 UniversityRoviraiVirgil [31]

Power: Western States Power Grid [32] PGP: an encrypted communication network [33]Some statistical properties of all the networks are listed in Table 1.

Table 1. Some statistical properties of datasets: node number (n), edge number(m),average degree ($\langle k \rangle$), maximum degree (kmax), epidemic threshold (β_c),average clustering coefficient (C), assortativity (A), diameter (D).

Network	|V|	|E|	$\langle k \rangle$	k_{max}	β_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

C. Pretreatment

In Ref. [26], the authors found that there existsa saturation effect that is negative for ranking performance, which is considered as the result ofthe radius of neighbour nodes set too large.Moreover, the computational complexity of themeasure will increase quickly when a moreextensive range of neighbour information is utilized.Therefore, to balance the accuracy and time cost,the parameter r of the proposed method need to bedetermined previously.The Kendall's tau coefficient [34] is widelyused to measure the rank correlation between twoordered lists. Suppose two equal-sized ranking listsX and Y are given, and the joint set is XY. For anytwo selected elements (x_i, y_i) and (x_j, y_j), if x_i>$x_j$and y_i>$y_j$or x_i<x_j and y_i<y_j, the pairs are consideredconcordant while they are discordant if x_i>$x_j$andy_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 nordiscordant. The Kendall's tau coefficient is definedas:

$$\tau = \frac{2(n_c - n_d)}{n(n-1)} \qquad (4)$$

Where n_c and n_d are the number of concordantpairs and discordant pairs, respectively. n is thetotal number of pairs.

To study the optimal radius of IFC, weconduct experiments on four datasets with differentsizes by considering the shortest distance $r \le 5$ toneighbour nodes, which are denoted C1, C2, C3, C4and C5 as Eq.(2), respectively. Then Kendall's taucoefficient is introduced to calculate the correlationof ranking lists obtained from C1~C5 and SIRmodel by varying the infection probability β. Theinfection probability β is set nearby the value ofepidemic threshold β_c in order to easily distinguishvital nodes. The larger value of τ corresponds tomore accurate results of measure ranking nodes.The comparative results are shown in Fig. 2.

As is shown in Fig.2, when the propagationprobability is small, such as β<0.07 in Karate network and Netscience network and β<0.03 in PGP network, C1 performsbetter than C2~C5 as the spreading process isconfined to local areas and cannot spread outward. Then with the increase of transmission probability, when β>β_c(the dot line), disease can disseminate ona large scale with more neighbour nodes acting asthe transmitter of the disease, so C4, C5 have betteraccuracy by considering wider neighbourinformation. In addition, there also exists thesaturation effect that with the increase of nodedistance, the improvement of evaluation ability islimited. The reason is that on the one hand,especially in small-scale networks (Fig.2 (a), (b)),the network diameter is small and the distancebetween nodes is limited, so C3, C4, C5 willachieve the similar experimental results. On theother hand, as the definition of IFC shown byEq.(2), the weight values of neighbour nodes withr=1,2,3…n are respectively assigned β_c^1, β_c^2, $\beta_c^3 \ldots \beta_c^n$.The smaller the β_c of the network is, the faster theweight value of the remote node decreases. Thepurpose is that when the epidemic threshold issmall, more attention should be paid to the initialspreading region of nodes, while the weight ratio ofremote neighbours get higher in the network withlarger epidemic threshold. By the consideration ofaccuracy and time consumption, we set the value of$r$=3 in IFC and EIFC, and comparison experimentsamong all the measures are conducted in the laterpart.

D. Results

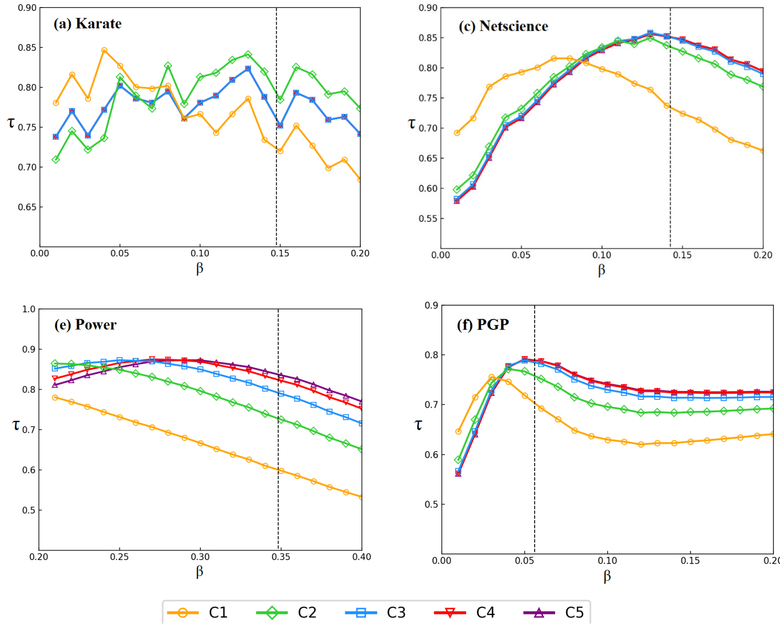

Fig.2. The effect on Kendall's tau when the value of parameter *r* of IFC is set from 1 to 5

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-shelldecomposition (KS) [17-20], Mixed degreecentrality (MDD) [21], extended neighbourcorenesscentrality (C_{nc+}) [24], k-shell iteration factorcentrality (KS-IF) [23], extended weighted degreecentrality (EWD) [25], whose definitions are shownin section 2. Moreover, we view the iteration factor(IF), part of KS-IF method, as an independentmeasure so as to compare the accuracy with that ofk-shell decomposition.

Similarly, we investigate the accuracy of eachmeasure by computing the value of τ with infectionprobability varied around β_c. The results obtainedfrom this experiment on all six real datasets areshown in Fig.3. We can observe that both IFC andEIFC can get higher values of τ in most cases,especially in the case of $\beta>\beta_c$. In comparison withk-shell method, IF offers higher correctness in allthe datasets except for Email. The reason is that theaverage of degree in Email network is large enoughthat the iteration information plays a less importantrole. But overall, iteration information is also a plusfor identifying vital nodes in the network.

In addition, nodes with different spreadingcapacity in the network may get the same scoremeasured by a given method, which is the case wedo not expect to see. Therefore, the capacity ofdifferentiating nodes is one of the criteria forevaluating the performance of centrality methods aswell. Here, the monotonicity *M* [24] is employed toevaluate the resolution of different rankingmeasures. The monotonicity is calculated by:

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

where R is the given ranking list, *n* is thenumber of ranks and nr is the number of the sameranks of R. Obviously, *M* ranges from 0 to 1, and*M* reaches the maximum value 1 only when the rankof an element vary from each other. The ranks ofnodes is viewed as the element of ranking lists anda larger value of *M* indicates the better capacity ofdiscriminating nodes of the metrics. Theexperimental results shown in Table II exhibitthat both IFC and EIFC are more suitable for thediscrimination in specification of node than theother methods. Moreover, IF method alsooutperforms k-shell method in terms ofmonotonicity. The high values of monotonicity ofproposed methods come from taking into accountmore extensive neighbour information as well asiteration information, which based on the idea ofcloseness centrality and iteration factor.

Table2.THE MONOTONICITY OF EACH MEASURES ON SIX REAL WORLD NETWORKS.

Network	M(DC)	M(CC)	M(KS)	M(IF)	M(MDD)	$M(C_{nc+})$	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	**0.9577**
Dolphins	0.8312	0.9737	0.3769	0.8781	0.9091	0.9873	0.9968	0.9958	**0.9979**	**0.9979**
Netscience	0.7642	0.9928	0.6421	0.7606	0.8226	0.9893	0.9948	0.9943	0.9951	**0.9955**
Email	0.8874	0.9988	0.8088	0.8977	0.9233	0.9991	0.9996	**0.9999**	**0.9999**	**0.9999**
Power	0.5927	0.9998	0.2460	0.6234	0.6940	0.9419	0.9848	0.9944	0.9998	**0.9999**
PGP	0.6193	0.9996	0.4806	0.6487	0.6679	0.9851	0.9912	0.9985	**0.9997**	**0.9997**

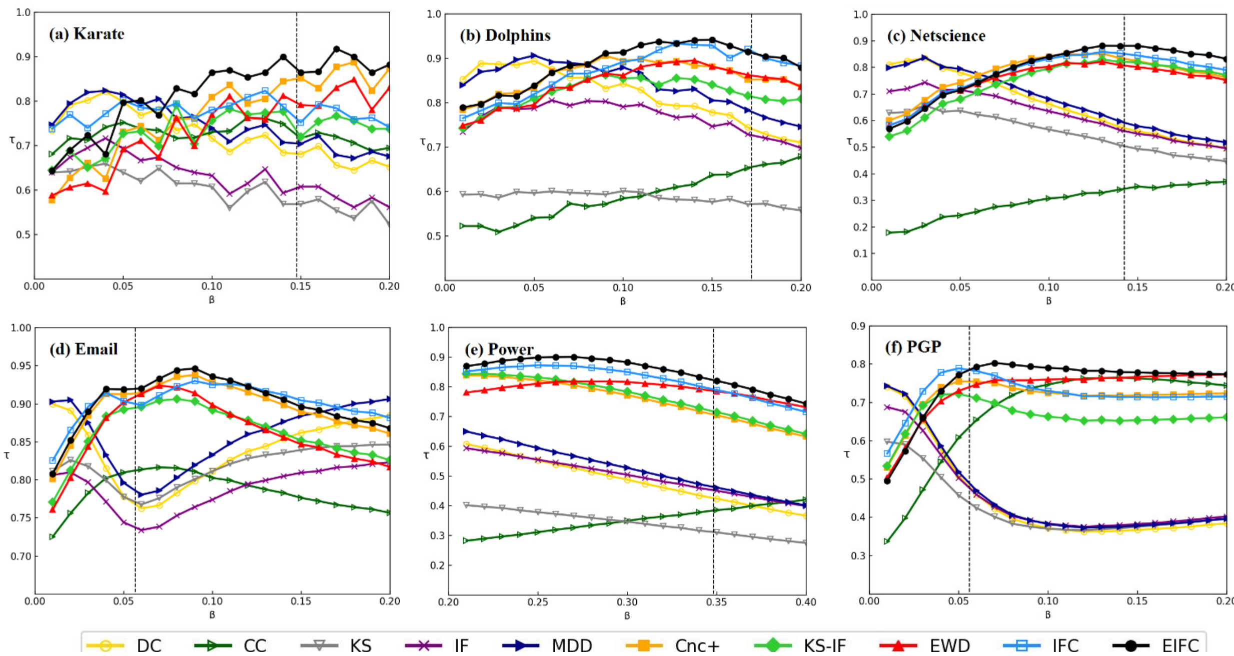

Fig. 3. The Kendall's tau between the ranking lists offered by ten measures and the ranking list generated by SIR simulation with the infected probabilityβ varied around β_c (dot line) in six real networks.

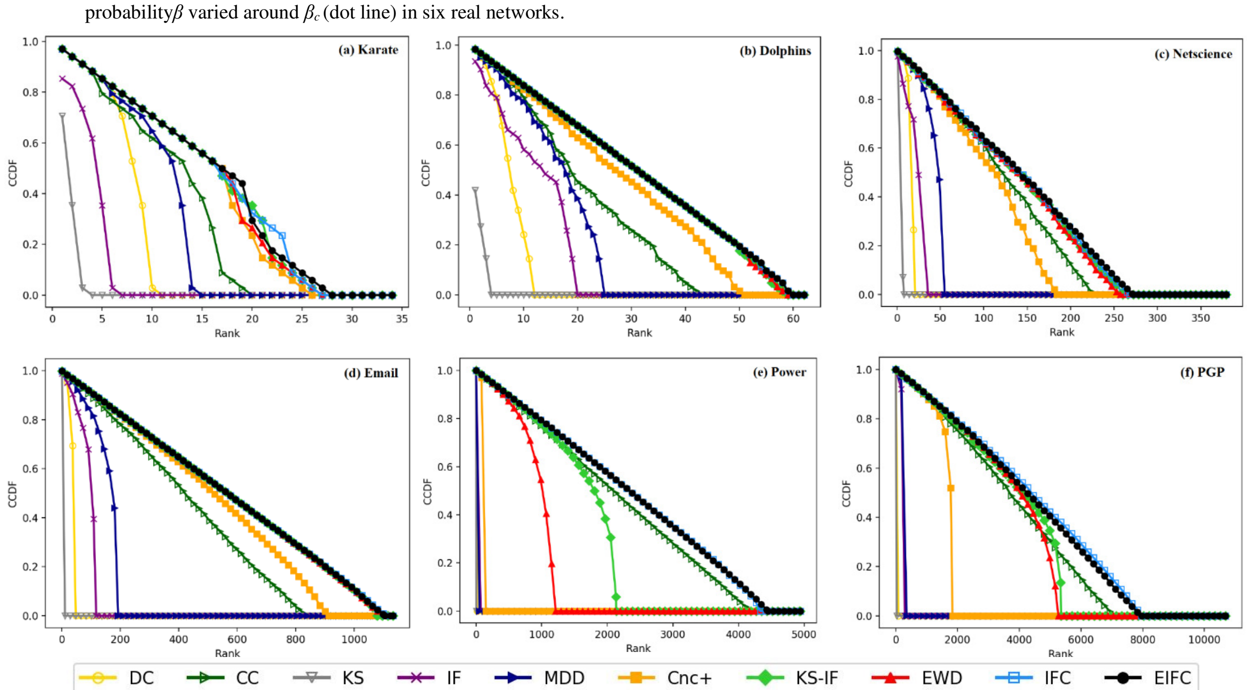

Fig. 4. The complementary cumulative distribution function (CCDF) plots for six real networks.

To better understand the distribution of nodesin each rank, complementary cumulative distribution function (CCDF) of six real networksare 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 andthe number of nodes ranking at i, respectively. Rdenotes the given ranking number in the list. Thefaster the curve of CCDF falls to zero, indicatingthat the nodes are distributed in the smaller rankinginterval. As observed in Fig. 4, the CCDF curve ofIFC, EIFC, KS-IF and EWD trend are basically thesame in small size networks. However, results ofPower and PGP network show that networks withlarger size can cause the discrimination of KS-IFand EWD to weaken, while our proposed methodsstill perform the best.

TABLE 3. THE COMPUTATIONAL COMPLEXITY OF TEN METHODS.

Ranking measure	Computation complexity	Category
DC	O(n)	Local measures
CC	$O(n^3)$ or $O(n^2 \log n+nm)$	Global measure
KS	O(m)	Global measure
IF	O(m)	Global measure
MDD	O(m)	Hybrid measure
C_{nc+}	$O(\langle k\rangle n)$	Hybrid measure
Ks-if	$O(\langle k\rangle n)$	Hybrid measure
EWD	$O(\langle k\rangle)$	Hybrid measure
IFC	$O(\langle k\rangle^3 n)$	Hybrid measure
EIFC	$O(\langle k\rangle^3 n)$	Hybrid measure

Next, the time complexity ofthe proposed measures is analysed. The calculationprocedure of IFC consists of three parts: first, tocompute the epidemic threshold in each network weneed time O(n). Then, obtaining the iteration factorvalues of nodes requires O(m), whose procedure isbased on the algorithm of k-shelldecomposition[35]. Last, summing over theiteration factor of neighbour nodes within a radiusof 3 runs in $O(\langle k\rangle^3 n)$.Therefore, the total timecomplexity 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 nearestneighbour, 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 timecomplexity of tenmethods is shown in Table 3. Due toconsideration of more neighbour information, bothIFC and EIFCcost slightly more computational complexity thanother hybrid measures. In addition, in comparison with CC, themethod of taking account into the neighbour nodeswith shortest distance as well, our methods greatlyincrease the efficiency by considering limitedneighbour influence. It can be concluded that both IFC and EIFC are sufficiently competitive in computing time consumption.

IV. CONCLUSION

How to identify the vital nodes in complexnetworks effectively has become the researchhotspots in recent years. In this letter, some relatedcentrality measures are introduced which includesclassical methods like degree centrality,betweenness centrality, closeness centrality, andsome hybrid measures based on k-shelldecomposition method. The existing methods haveconfirmed that neighbour information has a greateffect on identifying vital nodes. In closenesscentrality, the shortest distance between nodes isconsidered to be an affected factor in propagationprocess. In this letter, more extensive neighbournodes are further investigated in consideration ofthe negative effect on increasing distance to thecentral node. Then, the epidemic threshold isintroduced as the attenuation factor in spreadingprocess in order to adapt to propagation attenuationin different networks. Last, by combining distanceand iteration information of neighbour nodes, a newhybrid measures named iteration factor closenesscentrality (IFC) and its expansion forms arepresented. Firstly, experiments are carried out onfour real datasets with different scales to verify thatthe optimal neighbourhood radius is 3. Furthermore,the proposed methods are compared with theexisting method on six real world networks in termof three aspects, including accuracy, discriminationand computational complexity. The experimentalresults show that our methods outperform otherexisting methods and verify the positive effect ofiteration information from k-shell decomposition onidentifying key nodes. Finally, we analyse thecomputational complexity, the proposed methodsimprove the performance of CC and remaincompetitive in the existing methods. How to applythe proposed approach to specifying and rankingthe spreading capacity of nodes in weighted anddirected networks will be investigated in the futurework.

References

[1] Kinney R, Crucitti P, Albert R, et al. Modeling cascading failures in theNorth 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 offailures in interdependent networks [J]. Nature, 2010, 464(7291):1025-1028.

[3] Albert R, Jeong H, Barabási A-L. Error and attack tolerance of complexnetworks [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 ofinfluential genes in biological networks [J]. Oncotarget, 2017, 8(27):43810.

[6] Zemanová, Lucia, Zhou C, Kurths, Jürgen. Structural and functionalclusters of complex brain networks [J]. Physica D NonlinearPhenomena, 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 Internetunder intentional attack [J]. Physical review letters, 2001, 86(16): 3682.

[9] MotterA 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 viralmarketing based on a complex network model [J]. Physica A StatisticalMechanics & Its Applications, 2010, 389(4):859-870.

[11] Doer B, Fouz M, Friedrich T. Why rumors spread so quickly in socialnetworks [J]. Communications of the ACM, 2012, 55(6):70.

[12] Pastor-Satorras R, Vázquez, Alexei, Vespignani A. Dynamical andCorrelation 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 nodeand 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 influentialspreaders 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]. SocialNetworks, 1983, 5(3):269.

[20] Carmi S, Havlin S, Kirkpatrick S, et al. A model of Internet topologyusing k-shell decomposition [J]. Proceedings of the National Academyof Sciences, 2007, 104(27): 1150-11154.

[21] Zeng A, Zhang C J. Ranking spreaders by decomposing complexnetworks [J]. Physics Letters A, 2013, 377(14):1031-1035.

[22] Liu J G, Ren Z M, Guo Q. Ranking the spreading influence in complexnetworks [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 complexnetworks 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 complexnetworks 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 incomplex networks: An extended weighted degree centrality based on aremaining minimum degree decomposition [J]. Physics Letters A, 2018:S0375960118305632-.

[26] Liu Y, Tang M, Zhou T, et al. Identify influential spreaders in complexnetworks, 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 incomplex 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 insmall groups [J]. Journal of anthropological research, 1977, 33(4): 452-473.

[29] Lusseau D, Schneider K, Boisseau O J, et al. The bottlenose dolphincommunity of Doubtful Sound features a large proportion of longlasting associations [J]. Behavioral Ecology and Sociobiology, 2003, 54(4): 396-405.

[30] Newman M E J. Finding community structure in networks using theeigenvectors of matrices [J]. Physical review E, 2006, 74(3): 036104.

[31] Guimerà R, Danon L, DíazGuilera A, et al. Self-similar communitystructure 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 socialnetworks 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 decompositionof networks [J]. arXiv preprint cs/0310049, 2003.