Doctopic: Condensed matter ARTICLE IN PRESS PLA:25129

Physics Letters A ••• (••••) •••–•••

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Ranking the spreading influence of nodes in complex networks: An extended weighted degree centrality based on a remaining minimum degree decomposition

Fan Yang [a,b,*], Xiangwei Li [a], Yanqiang Xu [a], Xinhui Liu [a], Jundi Wang [a], Yi Zhang [a], Ruisheng Zhang [b], Yabing Yao [b]

[a] *School of Software Engineering, Lanzhou Institute of Technology, Lanzhou 730050, China*
[b] *School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China*

ARTICLE INFO

Article history:
Received 1 December 2017
Received in revised form 18 May 2018
Accepted 18 May 2018
Available online xxxx
Communicated by C.R. Doering

Keywords:
Complex networks
Spreading influence of nodes
Remaining minimum degree (RMD) decomposition
Extended weighted degree (EWD) centrality
Susceptible-Infectious-Recovered (SIR) model

ABSTRACT

Ranking the spreading influence of nodes is crucial for developing strategies to control the spreading process on complex networks. In this letter, we define, for the first time, a remaining minimum degree (RMD) decomposition by removing the node(s) with the minimum degree iteratively. Based on the RMD decomposition, a weighted degree (WD) is presented by utilizing the RMD indices of the nearest neighbors of a node. WD assigns a weight to each degree of this node, which can distinguish the contribution of each degree to the spreading influence. Further, an extended weighted degree (EWD) centrality is proposed by extending the WD of the nearest neighbors of a node. Assuming that the spreading process on networks follows the Susceptible-Infectious-Recovered (SIR) model, we perform extensive experiments on a series of synthetic and real networks to comprehensively evaluate the performance of EWD and other eleven representative measures. The experimental results show that EWD is a relatively efficient measure in running efficiency, it exposes an advantage in accuracy in the networks with a relatively small degree heterogeneity, as well as exposes a competitive performance in resolution.

1. Introduction

Spreading phenomena in the real world can be represented as spreading process on complex networks [1], some typical examples include disease spreading [2], computer virus propagation [3], rumors diffusion [4] and so on. Understanding the spreading phenomena and further controlling the spreading process are of great theorical and practical significance in complex networks [5–7]. One of the fundamental methods to control the spreading process is ranking the spreading influence of nodes [8,9], which has gained great attentions in recent years [10].

Some classical ranking measures include degree centrality (DC), betweenness centrality (BC), closeness centrality (CC) [11] and eigenvector centrality (EVC) [12]. To optimize the performance, some improved measures [13–16] are proposed based on these classical measures. Meanwhile, some novel ideas that are different from the classical measures are developed. Decomposing a network is an effective idea to rank the spreading influence of nodes [9,17,18]. A representative method is K-Shell (KS) decomposition [17]. However, it tends to assign many nodes a same KS index but neglects these nodes may be different in their true spreading influence [19]. Following KS decomposition, many improved measures are proposed, such as coreness centrality (C_{nc+}) [20], gravity centrality (GC) [21] and extended local K-Shell sum [22]. However, the accuracy of [17,20–22] is unsatisfactory.

Combining multiple measures is another feasible idea. Zeng et al. [23] proposed a mixed degree decomposition (MDD). Gao et al. [24] proposed a local structural centrality. Ma et al. [25] proposed a hybrid degree centrality. Wang et al. [26] proposed a measure based on node position and neighborhood. The limitation of [23–26] is that their performance depend on different adjustable parameters. Fu et al. [27] proposed a two-step framework (IF). Liu et al. [28] proposed a multi-attribute ranking measure. Madotto et al. [29] proposed a meta-centrality. Bian et al. [30] proposed a measure based on multiple attribute decision making technique. Particularly, in this type of methods, several works combined the shortest distance and other measures. Liu et al. [31] proposed a Θ measure by combining the shortest distance and KS decomposition. Bao et al. [32] and Tian et al. [33] combined the shortest

* Corresponding author at: School of Software Engineering, Lanzhou Institute of Technology, Lanzhou 730050, China.
E-mail addresses: fanyang2014@lzu.edu.cn, yangfan82416@163.com (F. Yang).
URLs: http://www.lzit.edu.cn, http://www.lzu.edu.cn (F. Yang).

https://doi.org/10.1016/j.physleta.2018.05.032
0375-9601/

distance and degree centrality. Obviously, the performance of this type of methods will depend on the combined measures.

Besides, Lü et al. [34] suggested that the h-index (HI) can better quantify node influence than either DC or KS. Wang et al. [35] proposed a new efficiency centrality (EC). However, the accuracy of [34,35] is unsatisfactory. Some other works [36–39] proposed different frameworks to employ a certain existing measure to rank the influence of nodes. However, the performance of these frameworks [36–39] will be determined by the measure they employed.

In this letter, we define, for the first time, a novel remaining minimum degree (RMD) decomposition by removing the node(s) with the minimum degree iteratively. Meanwhile, the corresponding algorithm is proposed. Inspired by an idea that the influence of a node largely depends on its neighbors [12,36,38,40], we present a weighted degree (WD) by utilizing the RMD indices of the nearest neighbors of a node. WD assigns a weight to each degree of this node, which can distinguish the contribution of each degree to the spreading influence. Further, motivated by an idea that utilizing the neighbor information within a more extensive range is beneficial to rank the influence of nodes accurately [14,20,22], an extended weighted degree (EWD) centrality is proposed by extending the WD of the nearest neighbors of a node. Among these measures, including [12,14,20,22,36,38,40] and EWD, a similar feature is that they rank the spreading influence of nodes by utilizing their neighbors. However, EWD is different from all these measures due to it is based on RMD decomposition. Finally, assuming that the spreading process on networks follows the Susceptible-Infectious-Recovered (SIR) model, we comprehensively evaluate the performance of EWD, WD and existing eleven representative measures on a series of synthetic and real networks. The results show that EWD is a relatively efficient measure in running efficiency, it exposes an advantage in accuracy in the networks with a relatively small degree heterogeneity, as well as exposes a competitive performance in resolution.

2. Extended weighted degree (EWD) centrality

2.1. Remaining minimum degree (RMD) decomposition

Denote $G(V,E)$ as an undirected and unweighted network, where V and E are the nodes set and edges set, respectively. RMD decomposition is as follows,

Definition 1. RMD decomposition: given a G, find the minimum degree (md) in current G, remove all nodes with degree=md and generate a new G, the removed nodes are assigned with a RMD index 1. In the similar way, find the md in current G, remove all nodes with degree=md, generate a new G, the removed nodes are assigned with a RMD index 2. Repeat the process and assign a corresponding RMD index to the removed nodes in each iteration. Finish the process when G is empty.

Table 1 shows the algorithm to implement the RMD decomposition. A complete RMD decomposition process is demonstrated in Fig. 1.

Algorithm analysis. The variable $CurrentIter$ in line(02) is to record the current iterative time in whole iterative process. The array $Iter$ in line(03) is to record the iterative time of each node when it is removed, which also represents the RMD indices of nodes. Line(05) calculate the degree of V in current G, the time complexity of this step is $O(m)$. The variable md in line(06) records the minimum degree in current G. The variable $count$ in line(07) is to count the number of removed nodes in each iteration. Lines(08–14) assign a RMD index to the node with degree$=md$, meanwhile, record this node in a temporary array $DelArray$, the time complexity of this

Table 1
The algorithm for RMD decomposition.

Input:	a network G.
01	$n =$ the size of V;
02	$CurrentIter = 1$;
03	$Iter$ is an array with size of n;
04	while($n > 0$)
05	calculate the degree of G and store them in an array deg;
06	$md = \min(deg)$;
07	$count = 0$;
08	for each v in G
09	if ($deg[v] == md$)
10	$Iter[v] = CurrentIter$;
11	record v in a temporary array $DelArray$;
12	$count = count + 1$;
13	end if
14	end for
15	delete the node(s) in $DelArray$ from G and generate a new graph G';
16	$G = G'$;
17	$n = n - count$;
18	$CurrentIter = CurrentIter + 1$;
19	end while
20	$MaxIter = \max(Iter)$;
Output:	the array $Iter$ and the variable $MaxIter$.

step is $O(n)$. Lines(15–16) remove the node(s) in array $DelArray$ and generate a new network G. In lines(17–18), $CurrentIter$ and n are updated after node(s) being removed in each iteration. Finally, taking the loop *while* into account, the time complexity of RMD decomposition is $O(mn)$ in the worst case. In fact, with node(s) being removed in each iteration, the size of V in G is decreased continuously. Thus, the actual running time of RMD is far less than $O(mn)$.

2.2. Extended weighted degree (EWD) centrality

Based on RMD decomposition, a weighted degree (WD) is defined as follows,

$$\mathrm{WD(v)} = \sum_{u \in neighbors(v)} \frac{Iter(u)}{MaxIter} \tag{1}$$

where, the array $Iter$ and variable $MaxIter$ can be obtained from the RMD decomposition algorithm in Table 1. In equation (1), we can observe that each item in $\sum$ corresponds to assign a weight to each degree of a node. Therefore, the contribution of each degree to the spreading influence of this node is distinguished. On one hand, implementing RMD decomposition must utilize the global topological structure of a network. Thus, RMD index reflects the global position of a node in this network. On the other hand, the calculation of WD is based on the nearest neighbors of a node, which utilizes the local topological structure of a network. Overall, WD considers both of the global and the local structure simultaneously. Further, we propose an extended weighted degree (EWD) centrality by extending the WD of the nearest neighbors of a node,

$$\mathrm{EWD(v)} = \sum_{u \in neighbors(v)} \mathrm{WD(u)} \tag{2}$$

Take the node 15 in Fig. 1(a) for example, $\mathrm{WD}(15) = \frac{Iter(2)}{7} + \frac{Iter(3)}{7} + \frac{Iter(6)}{7} + \frac{Iter(10)}{7} + \frac{Iter(13)}{7} \approx 4.143$, $\mathrm{EWD}(15) = \mathrm{WD}(2) + \mathrm{WD}(3) + \mathrm{WD}(6) + \mathrm{WD}(10) + \mathrm{WD}(13) \approx 11.429$ (the array $Iter$ and variable $MaxIter$ are calculated in Fig. 1).

Time complexity analysis. From Section 2.1, we know the time complexity of RMD is $O(mn)$ in the worst case. With the result of RMD decomposition, by equation (1), we know that obtaining the WD values of nodes requires $O(m)$. By equation (2), we know that

Please cite this article in press as: F. Yang et al., Ranking the spreading influence of nodes in complex networks: An extended weighted degree centrality based on a remaining minimum degree decomposition, Phys. Lett. A (2018), https://doi.org/10.1016/j.physleta.2018.05.032