International Journal of Innovative Computing, Information and Control
Volume **11**, Number **5**, October **2015**

ICIC International ©2015 ISSN 1349-4198
pp. 1765–1777

IDENTIFYING INFLUENTIAL NODES IN WEIGHTED NETWORK BASED ON EVIDENCE THEORY AND LOCAL STRUCTURE

Jiadong Ren[1,2], Chunyan Wang[1,2,*], Hongdou He[1,2] and Jun Dong[1,2]

[1]College of Information Science and Engineering
Yanshan University
[2]The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province
No. 438, Hebei Ave., Qinhuangdao 066004, P. R. China
{ jdren; dongjun }@ysu.edu.cn; 717431731@qq.com
*Corresponding author: yan7616494321@sina.com

Received April 2015; revised August 2015

Abstract. *It is a fundamental and important issue to identify influential nodes in complex network. In the existing evidential semi-local centrality, it modified the evidential centrality according to the actual degree distribution, but the topological connections among the neighbors of a node in weighted network are not taken into account. In this paper, a novel measure called evidential local structure centrality is proposed to identify influential nodes. Firstly, the value of modified evidential centrality is calculated by taking actual degree distribution. Secondly, local structure centrality combined with modified evidential centrality is extended to be applied in weighted networks. Then, in order to evaluate the performance of the proposed method, we use the susceptible-infected-recovered (SIR) model and susceptible-infected (SI) model to simulate the spreading process on real networks. Experiment results show that our method is effective and efficient to identify influential nodes.*
Keywords: Complex network, Influential node, Weighted network, Dempster-Shafer theory of evidence, Local structure

1. **Introduction.** How to identify influential nodes in complex network has become a hot topic in many fields such as effectively controlling disease [1,2] and computer viruses spreading, rumors diffusion [3], as well as promoting new products, looking for the leaders [4], and ranking scientists and publications [5,6].

In a complex network analysis, a variety of centrality indices were proposed to identify influential nodes in weighted networks. In [7], degree centrality (DC), betweenness centrality (BC), and closeness centrality (CC) were extended to be widely applied in weighted networks. DC is simple and less relevant, since a few highly influential neighbors may be more influential than a node with a larger number of less influential neighbors. Global metrics such as BC and CC can well identify influential nodes by ranking nodes, but their computational complexity is considerable, so they are not feasible for large-scale networks. Simultaneously, another constraint of CC is lack of applicability to networks with disconnected components, two nodes that belong to different components but do not have a finite distance between them. After that, Gao et al. [8] proposed a local structure centrality (LSC) measure which considers both the number and the topological connections of the neighbors of a node. For nodes with the same number of neighbors, the one with denser connected neighbors is supposed to be more influential since denser connected neighbors get more chance to influence each other. However, it is incapable of being applied in weighted networks. In the past two years, several centrality measures are also proposed to identify influential nodes, such as Weighted LeaderRank [9], neighborhood coreness centrality [10], and Weighted k-shell decomposition [11].

Dempster-Shafer evidence theory (D-S evidence theory for short) was first proposed by Dempster [12], formed by the further expansion of Shafer [13]. In this theory, belief function and plausibility function of proposition A are respectively represented by lower bound and upper bound of evidence interval. Furthermore, the D-S evidence theory is qualified to combine a pair of evidence or belief functions to obtain a new evidence or belief function. Based on the D-S theory, evidential centrality (EVC) [14] measure is raised as a tradeoff between degree and strength of each node to derive node importance in weighted network. In the literature [15], evidential semi-local centrality (ESC) measure is proposed by a combination of the modified evidential centrality which considers degree distribution of real network and the extension of semi-local centrality in weighted network. The values of centrality measure for each node are obtained by both these centrality measures, respectively. Then, we can obtain the orders of the nodes by comparing these values. It can be seen that the higher the value is, the more influential the node is.

However, the evidential centrality measure is similar with DC – simple but of little relevance, since it does not take account of the global structure information of the network. Hence, in order to rank nodes effectively, it is better to design the ranking algorithms based on the local information of the network. For example, the local structure centrality considers not only the number of node's neighbors, but also the topological connections among its neighbors. Inspired by both of the ideas, in this paper, combining the modified evidential centrality with taking degree distribution into account and the extension of local structure centrality in weighted network, a new centrality measure is proposed to identify influential nodes. The value of the new centrality measure for each node is ranked in descending order. The higher the value of centrality measure is, the more influential the node is. In order to validate the performance of the proposed method, the susceptible-infected-recovered (SIR) model is used to examine spreading influence of nodes ranked by different centrality measures in real network.

The primary contributions of this paper can be summarized as follows.

- A new centrality measure called evidential local structure centrality (ELSC) is proposed to identify influential nodes by combining the modified evidential centrality with taking degree distribution into account and the extension of local structure centrality in weighted network.
- The proposed method is raised as a tradeoff between degree and strength of each node; meanwhile, it also considers both the number and the topological connections of node's neighbors.
- The local structure centrality is extended to be applied in weighted network very well.

The rest parts are organized as follows. In Section 2, we give an overview of centrality measures in brief and introduce evidence theory. A new method for identifying the influential nodes is proposed in Section 3. In Section 4, we present data sets and apply the SIR and SI models to evaluate the performance of the proposed method. Finally, some conclusions are summarized in the last section.

2. Preliminaries.

2.1. Definition.

Assuming that a weighted and undirected network $G = (V, E, W)$ is composed of $|V| = N$ nodes and $|E| = M$ edges. W is the weight set of E, i.e., the edge E_{uv} from node u to node v has a weight $\omega_{uv} \in W$.

In the literature [8], local structure centrality of node v, denoted as $C_{LS}(v)$, is defined as follows.

$$\begin{aligned} C_{LS}(v) &= \sum_{u \in \Gamma_1(v)} Q(u) \\ &= \sum_{u \in \Gamma_1(v)} \left(\alpha N(u) + (1-\alpha) \sum_{w \in \Gamma_2(u)} c_w \right) \end{aligned} \tag{1}$$

Here, local clustering coefficient [16] of node w quantifies how close its neighbors connect each other, which is defined as follows.

$$c_w = \frac{2\left|\{e_{ij} : i, j \in \Gamma_1(w), e_{ij} \in E\}\right|}{C_d(w)(C_d(w) - 1)} \tag{2}$$

where $C_d(w)$ is the degree of node, and $\Gamma_h(v)$ denotes the set of neighbors within h-hops from node v. $N(v)$ is the amount of the nearest and the next nearest neighbors of node v, i.e., $N(v) = |\Gamma_2(v)|$.

For each neighbor node u of node v, the $Q(u)$ is seen as node u's contribution to the final local structure centrality value of node v. We think about its nearest and next nearest neighbor set $\Gamma_2(u)$ to calculate $Q(u)$ for each node u. As we mentioned before, the local structure centrality measure considers both the number of the neighbor and the topological connections among the neighbors. Specially, For each node $w \in \Gamma_2(u)$, the former contribution of node w to $Q(u)$ is simply 1, namely each node in $\Gamma_2(u)$ is equally treated and only counted once in the calculation of $Q(u)$. The latter contribution of node w to $Q(u)$ is its local clustering coefficient c_w. We set a parameter α to balance both of the contributions. The total contribution of node w to $Q(u)$ is $\alpha * 1 + (1-\alpha) * c_w$. By summing contributions of all the nodes in $\Gamma_2(u)$, the $Q(u)$ can be obtained. Then, by summing all the $Q(u)$ for each neighbor node u of node v, we can get the local structure centrality $C_{LS}(v)$ as defined in Formula (1).

Local structure centrality considers not only the number of neighbors of a node, but also the topological connections among the neighbors. And it can be used to analyze a large-scale network. Nevertheless, it just can be applied in unweighted network, so in this paper, it is extended to be applied in weighted network. A new centrality measure called evidential local structure centrality (ELSC) is given.

Definition 2.1. *(ELSC in a Weighted Network) ELSC value of node v is denoted by $elsc(v)$, which satisfies*

$$\begin{aligned} elsc(v) &= \sum_{u \in \Gamma_1(v)} Q^{\omega}(u) \\ &= \sum_{u \in \Gamma_1(v)} \left(\alpha N^{\omega}(u) + (1-\alpha) \sum_{w \in \Gamma_2(u)} c_w^{\omega} \right) \end{aligned} \tag{3}$$

where $\Gamma_h(v)$ is the set of neighbors within h-hops from node v. $N^{\omega}(u)$ denotes the sum of the nearest and the next nearest neighbors' mec of node u. α $(0 \le \alpha \le 1)$ is balance parameter. In weighted network, c_w^{ω} is the local clustering coefficient of node w [17], and its definition is as follows.

$$c_w^{\omega} = \frac{1}{s_w(k_w - 1)} \sum_{j,k} \frac{\omega_{wj} + \omega_{wk}}{2} a_{wj} a_{jk} a_{kw} \tag{4}$$

where j, k are any two nodes connected to the node w, s_w is the sum of weights of all edges connecting with node w, and k_w means the degree of node w. The weight between node w and node j is denoted by ω_{wj}, and the value of a_{wj} is 1 if node w is linked to node j, otherwise 0. It is clear that the closer connection among nodes is, the higher the local clustering coefficient is.

2.2. **Dempster-Shafer theory of evidence.** In Dempster-Shafer evidence theory, problem domain $\Theta = \{a_1, a_2, \cdots, a_n\}$ is a nonempty set which consists of a finite number of mutually exclusive and exhaustive hypotheses, called the frame of discernment.

Suppose Θ is the frame of discernment, a mass function is mapping m: $2^\Theta \to [0, 1]$, (2^Θ is the power set of the Θ), basic probability assignment (BPA) is defined as follows.

$$m(\Phi) = 0 \text{ and } \sum_{A \subset 2^\Theta} m(A) = 1 \tag{5}$$

where Φ is the empty set and A is any element of 2^Θ, and mass $m(A)$ represents how strongly the evidence supports A.

Assuming that masses m_1 and m_2 are both basic probability assignments of Θ, orthogonal sum $m(A)$ is calculated from the two sets of masses m_1 and m_2 in Dempster's rule of combination.

$$m(A) = \frac{1}{1-K} \sum_{B \bigcap C = A} m_1(B) m_2(C) \tag{6}$$

with

$$K = \sum_{B \bigcap C = \Phi} m_1(B) m_2(C) \tag{7}$$

where A, B and C are elements of 2^Θ.

3. **Influential Node Identifying by Evidential Local Structure Centrality.** According to the definition of evidential centrality, it seems that evidential centrality is defined as a tradeoff between degree and strength of each node to derive node importance in weighted network. Nevertheless, just like the degree centrality, evidential centrality only captures the characteristics of single node, rather than the local structure feature of the network. Here, we extend the local structure centrality in weighted network to identify the local structure feature of the network. Thus, in this paper, the influence of the node is identified by a new centrality measure, called evidential local structure centrality.

In ELSC measure, the influences of degree and strength of each node are remarked by basic probability assignments (BPAs). The BPA obtained from degree of a node is based on the real degree distribution. Then for these BPAs of each node as to degree and strength, the influence value of each node is obtained by Dempster-Shafer theory of evidence. Further, both the neighbors information of a node and the topological connections among neighbors are taken into consideration. The sum of the nearest and next nearest neighbors influence value of each node is calculated and the topological connection among neighbors is measured by the local clustering coefficient. Finally, the influence of node is identified by ranking the value of ELSC. In a word, the ELSC combines the modified evidential centrality with taking degree distribution into account and the extension of local structure centrality in weighted network. The algorithm called Identify Influential Nodes-ELSC for identifying influential node by ELSC is performed as follows.

In the algorithm, there are two evaluation indices which are *high* or *low* for the influence of degree and strength of nodes in weighted network. Hence, in Step 1, a frame of discernment θ is denoted as $\theta = (high, low)$. Then in Step 2, in order to modify the

Algorithm 1 Identify Influential Nodes-ELSC

Input: Adjacent matrix and adjacent list corresponding to weighted networks.
Output: The ranked list and the corresponding value of ELSC with different balance parameter α ($0 \leq \alpha \leq 1$).
Step 1. Construct a frame of discernment Θ.
Step 2. Ascertain a corrected parameter to modify the BPA of degree.
Step 3. Calculate BPAs of each node with respect to the degree and strength.
Step 4. Achieve the BPA of influence value of the ith node by the Dempster's rule of combination.
Step 5. Let $m_i(\theta)$ allocate to $m_i(h)$ and $m_i(l)$ normally.
Step 6. Calculate the modified evidential centrality $mec(i)$.
Step 7. Ensure $mec(i)$ is a positive number.
Step 8. Calculate the sum of the nearest and the next nearest neighbors' mec of each node.
Step 9. Obtain the local clustering coefficient of each node according to the definition of evidential local structure centrality.
Step 10. Calculate the values of ELSC for each node.
Step 11. Rank the values of ELSC for each node in descending order.

BPA of degree, the real degree distribution is considered as a parameter. Assume node i with degree k_i follows a degree distribution $P(k_i)$. Thus, the parameter is defined as $\lambda_i = \sum_{j \leq k_i} P(j)$, where j is a set of degree of nodes which is lower than k_i. Next, in Step 3, BPAs of *high* or *low* influence for the degree of ith node are represented for $m_{id}(h)$ or $m_{id}(l)$ ($i = 1, 2, \ldots, N$), separately; Likewise, BPAs of *high* or *low* influence for the strength of ith node are represented for $m_{i\omega}(h)$ or $m_{i\omega}(l)$ ($i = 1, 2, \ldots, N$), respectively. They are expressed as follows.

$$m_{id}(h) = \lambda_i \frac{|k_i - k_m|}{\sigma} \tag{8}$$

$$m_{id}(l) = (1 - \lambda_i) \frac{|k_i - k_M|}{\sigma} \tag{9}$$

$$m_{i\omega}(h) = \frac{|\omega_i - \omega_m|}{\delta} \tag{10}$$

$$m_{i\omega}(l) = \frac{|\omega_i - \omega_M|}{\delta} \tag{11}$$

where σ and δ are given as

$$\sigma = k_M + \mu - (k_m - \mu) = k_M - k_m + 2\mu \tag{12}$$

$$\delta = \omega_M + \varepsilon - (\omega_m - \varepsilon) = \omega_M - \omega_m + 2\varepsilon \tag{13}$$

where $0 < \mu < 1$, $0 < \varepsilon < 1$, paper [14] demonstrated that the values of μ and ε have no impact on the ranking orders of nodes in weighted network. k_M and k_m are the maximum and minimum values of degree, and ω_M and ω_m correspond to the maximum and minimum values of weight, respectively. According to the above statement, the BPAs of degree and strength of ith node are obtained, respectively, as

$$M_d(i) = (m_{id}(h), m_{id}(l), m_{id}(\theta)) \tag{14}$$

$$M_\omega(i) = (m_{i\omega}(h), m_{i\omega}(l), m_{i\omega}(\theta)) \tag{15}$$

where

$$m_{id}(\theta) = 1 - (m_{id}(h) + m_{id}(l)) \tag{16}$$

$$m_{i\omega}(\theta) = 1 - (m_{i\omega}(h) + m_{i\omega}(l)) \tag{17}$$

For the above BPAs of the ith node as to degree and strength, the BPA of influence value of ith node is achieved by the Dempster's rule of combination in Step 4, and is listed by

$$M(i) = (m_i(h), m_i(l), m_i(\theta)) \tag{18}$$

where $\theta = (high, low)$. In Equation (18), $m_i(\theta)$ means the probability of *high* or *low* of the ith node. In Step 5, letting $m_i(\theta)$ allocate to $m_i(h)$ and $m_i(l)$, then the probabilities of *high* or *low* influence of the ith node are given by

$$M_i(h) = m_i(h) + \frac{1}{2m_i(\theta)} \tag{19}$$

$$M_i(l) = m_i(l) + \frac{1}{2m_i(\theta)} \tag{20}$$

Apparently, the higher the value of $M_i(h)$ is, the more influential the node is. On the contrary, the lower the value of $M_i(l)$ is, the more influential the node is. So the modified evidential centrality $mec(i)$ of the ith node is defined as

$$mec(i) = M_i(h) - M_i(l) = m_i(h) - m_i(l) \tag{21}$$

In Equation (21), the value of $mec(i)$ is a positive or negative number. Thus, in Step 7, to ensure $mec(i)$ is a positive number, the numerical treatment and normalization are denoted as below.

$$mec(i) = \frac{|\min(mec)| + mec(i)}{\sum_{i=1}^{N}\{|\min(mec)| + mec(i)\}} \tag{22}$$

where $|\min(mec)|$ is the absolute minimum value of mec. In Step 10, for the results from the Step 8 and Step 9, the value of ELSC for each node is achieved according to the definition of ELSC. In the end, the value of ELSC for each node is ranked in descending order. To sum up, the higher the value of ELSC is, the more influential the node is.

4. **Experimental Analysis.** SIR model [18] is a widely used tool to examine the spreading influence of nodes in weighted networks, and there are three states, namely $Susceptible(S)$, $Infected(I)$ and $Recovered(R)$. At the initial time, only one node is in infected state. At each step, each node in the infected state randomly selects their susceptible neighbors with probability P and enters the recovered state with probability equal to 1. The spreading process terminates when there is no node which is infected. In weighted networks, node j is infected by node i with probability

$$P = \left(\frac{\omega_{ij}}{\omega_M + 1}\right)^{\beta}, \quad \beta > 0$$

[19], where ω_{ij} is the weight of edge E_{ij} and ω_M is maximum of weights. Notice that this model is slightly different from the standard SIR model where all the neighbors of an infected node have the chance to be infected. The present mechanism is usually used to mimic the limited spreading capability of individuals [20,21]. The spreading capacity of node v is defined as the number of nodes that are finally infected at the end of spreading process which originates from node v. Except for the standard SIR model, the standard SI model [20] is applied to examine the spreading influence of top-ranked nodes, and it has two compartments, namely $Susceptible(S)$ and $Infected(I)$. The spreading process stops when all nodes in the network become infected. At time t, the number of infected node is denoted by $F(t)$, and it is treated as an indicator to estimate the influence of

initially infected node. At the steady state, the number of infected nodes is equal to the total number of nodes in network. In our case, we set $t = 5$ for further investigation, because the spreading in early stage is more important in fact. The results are obtained by averaging over 100 independent realizations.

To validate consistence of the ranked list generated by a centrality measure and one by SIR, we use Kendall's tau coefficient (τ) [22]. It considers a set of joint observations from two random variables X and Y (in this paper, X is the values of a certain centrality measure and Y can be the simulation results for all nodes). Any pair of observations (x_i, y_i) and (x_j, y_j) are said to be concordant if the ranks for both elements agree: if both $x_i > x_j$ and $y_i > y_j$ or if both $x_i < x_j$ and $y_i < y_j$. They are said to be discordant if $x_i > x_j$ and $y_i < y_j$ or if $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$, the pair is neither concordant nor discordant. The Kendall's tau coefficient τ is given a definition as

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

where n_c and n_d mean the number of concordant and discordant pairs respectively. The higher the value of τ is, the more accurate the ranked list generated by a centrality measure is.

4.1. **Experimental data.** In this paper, three real weighted networks are applied. (i) Zachary's Karate Club Network [23], the undirected and weighted network consists of 34 nodes. The data is collected from the members of a university karate club by Wayne Zachary. A node represents the member of the Club, every link means they have a friendship outside the Club activities, and the weight of edge signifies how closely the members associate each other. (ii) Les Miserable Network [24], the network is a weighted network with 77 characters, and a character is denoted by a node, each edge represents the two characters appearing in the same chapter of book, and the weight indicates how often such a co-appearance occurred. (iii) Netscience network, the network of co-authorships between scientists who are themselves publishing on topic of network. There are in total 1589 scientists in this collaboration network [25]. Here, we consider the largest component with 379 scientists. In Table 1, the basic topological properties of these three networks are shown. n and m are the total number of nodes and links respectively. $< k >$ and $k_{\max}$ denote the average and maximum degree. $< \omega >$ and $\omega_{\max}$ denote the average and maximum weight. C is the clustering coefficient.

TABLE 1. The basic topological features of the three real networks

Network	n	m	$< k >$	$k_{\max}$	$< \omega >$	$\omega_{\max}$	C
Zachary	34	78	4.5882	17	2.9615	7	0.5817
Les miserable	77	254	6.5974	36	3.2283	31	0.6057
Netscience	379	914	4.8232	34	0.5356	4.75	0.7610

4.2. **Experimental results.** In Club network, firstly, by setting the balance parameter α of ELSC to be 0.8, the τ value by ELSC is compared with the ones by ESC and EVC under different β ($1 \leq \beta \leq 2$) values corresponding to different spreading probabilities and the results are shown in Figure 1. As seen in Figure 1, ELSC can achieve better performance on a wide range of β value. The Kendall's tau coefficient τ can only estimate the consistency of the ranked list generated by a certain centrality measure and the ranked list generated by SIR model, while the real spreading ability of top-ranked nodes is incapable of being evaluated. Thus, Figure 2 shows the average number of infected

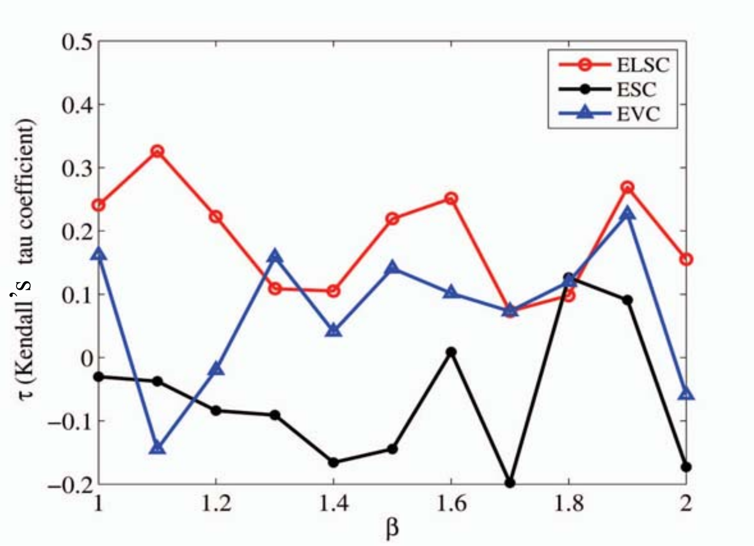


FIGURE 1. The Kendall's tau τ values corresponding to three centrality measures in Club network

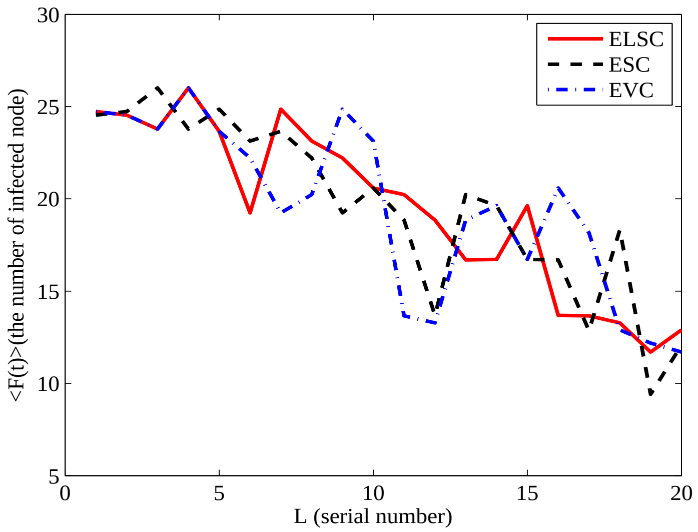


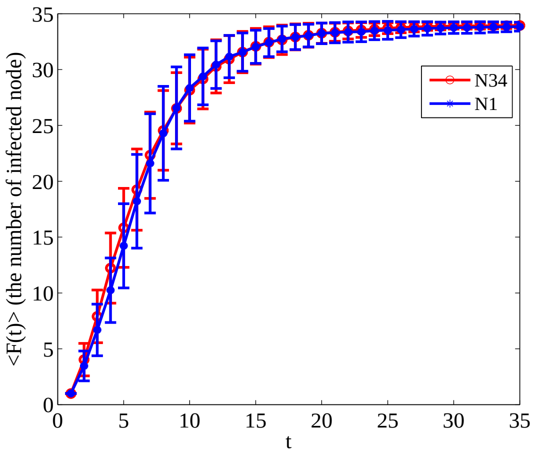
FIGURE 2. The average number of $F(t)$ $(t = 5)$ of top-L nodes as ranked by the three centrality measures

nodes (i.e., $F(t)(t = 5)$) by the top-L nodes as ranked by three centrality measures. Here, $\beta = 1.1$. Obviously, the curve for our proposed centrality measure is downward sloping more gently than ESC and EVC, namely the average spreading ability of top-L nodes obtained by our ELSC decreases more steady with the increasing of L.

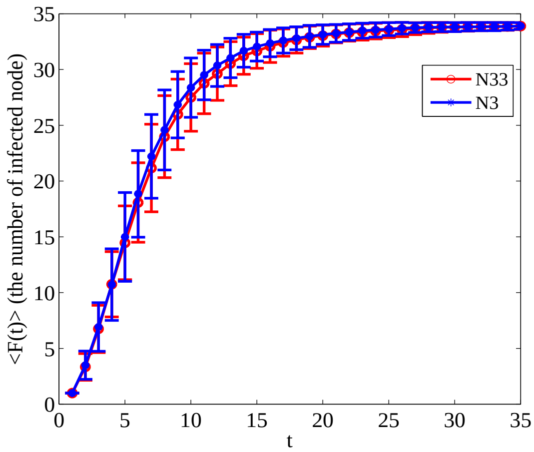
Meanwhile, the top-5 nodes ranked by the proposed method, ESC and EVC are listed in Table 2. Apparently, the spreading abilities between node 1 and node 34 as well as between node 3 and node 33 need to be distinguished to validate the efficiency of the proposed method. In Figure 3(b), the spreading speed and stability of node 3 and node 33 are almost the same. Here, $\beta = 1.8$. Besides, in Figure 3(a) there is subtle difference between node 1 and node 34 in the aspect of spreading ability, and in early stage, the number of nodes infected by regarding node 34 as the initial node is a little higher than the number of nodes infected by regarding node 1 as the initial node. Therefore, the proposed method can well identify the influential nodes in the Club network.

TABLE 2. The top-5 ranked nodes by ELSC with $\alpha = 0.8$, ESC and EVC

L	ELSC	ESC	EVC
1	34	1	34
2	1	34	1
3	33	3	33
4	3	33	3
5	2	9	2



(a) Spreading ability between node 1 and node 34



(b) Spreading ability between node 33 and node 3

FIGURE 3. The number of infected nodes by initially infected nodes in the top-5 list

In Les miserable network, balance parameter α is equal to 0.5 and parameter β in spreading ability P ranges from 0 to 1. From Figure 4, we survey that the Kendall's tau calculated by ELSC is higher than the ones calculated by ESC and EVC on a large scale. That is to say, the ranked list generated by ELSC is much closer to the ranked list generated by the real spreading process.

Moreover, we compare the spreading ability of the nodes that either appear in the top-10 list by ELSC or other two centrality measures including ESC and EVC (not appearing in both lists). Note that without considering the effects of common nodes in both ranking lists, the differences of these methods can be well distinguished. The top-10 nodes generated by these three centrality measures are displayed in Table 3. Figure 5(a) and Figure 5(b) show the simulations on the cumulative infected nodes, namely $F(t)$, as a function of time for Les miserable network. The number of cumulative infected nodes increases with time and ultimately reaches the steady value. As shown in Figure 5(a), the average number of infected nodes by the proposed method in each step is a bit larger than that by ESC, that is, the result for our proposed method is slightly better than the result for ESC, and the new method is almost similar to the EVC in Figure 5(b). Hence, the proposed method has the better performance than other centrality measures in the Les miserable network.

Furthermore, when considering the Netscience network, we set the balance element α to be 0.4. The results for the comparison of the Kendall's tau value with respect to ELSC, ESC and EVC are shown in Figure 6. The value of parameter β ranges from 0 to 1. It is

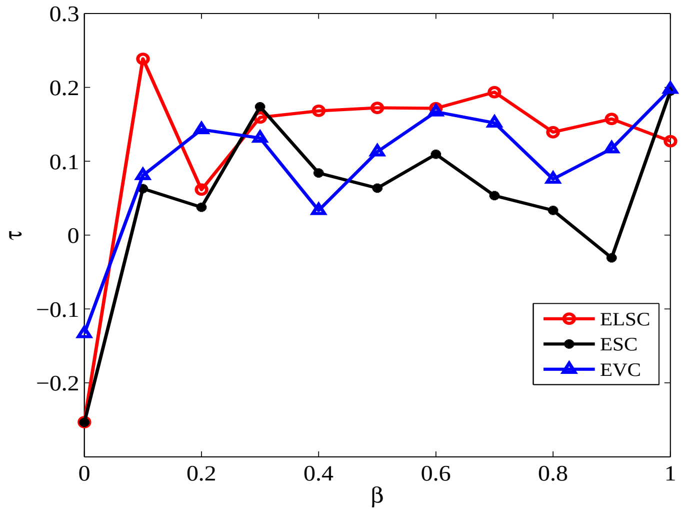
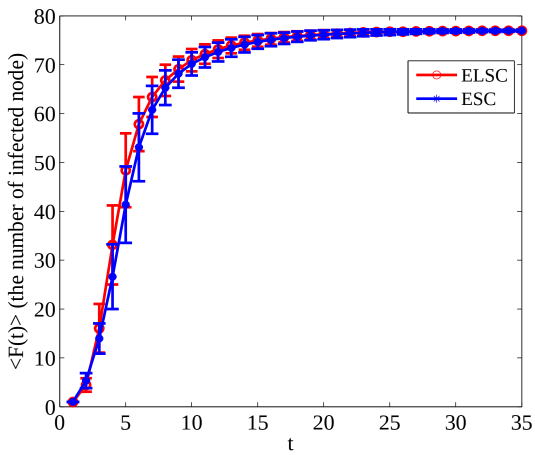


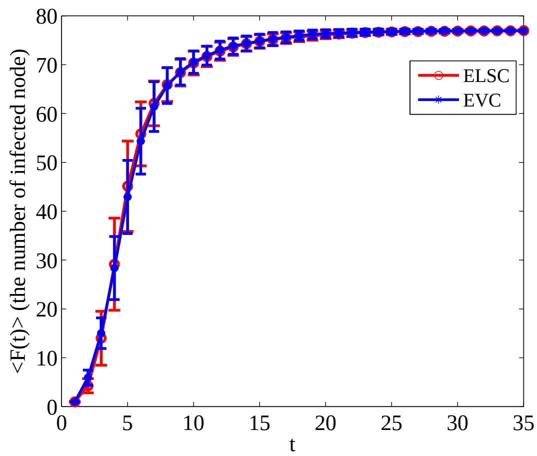
FIGURE 4. The Kendall's tau τ values corresponding to three centrality measures in Les miserable network

TABLE 3. The top-10 ranked nodes by ELSC with $\alpha = 0.5$, ESC and EVC

L	ELSC	ESC	EVC
1	12	12	12
2	49	49	56
3	28	56	59
4	56	59	49
5	26	26	63
6	59	28	26
7	24	65	65
8	27	63	28
9	65	64	27
10	71	66	60



(a) Comparison by ELSC or ESC



(b) Comparison by ELSC or EVC

FIGURE 5. The number of infected nodes by initially infected nodes in the top-10 list under $\beta = 0.5$

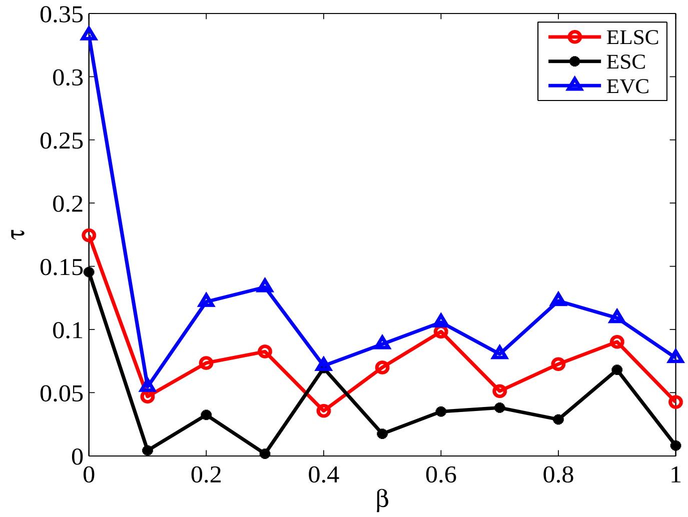


FIGURE 6. The Kendall's tau τ values corresponding to three centrality measures in Netscience network

TABLE 4. The top-10 ranked nodes by ELSC with $\alpha = 0.4$, ESC and EVC

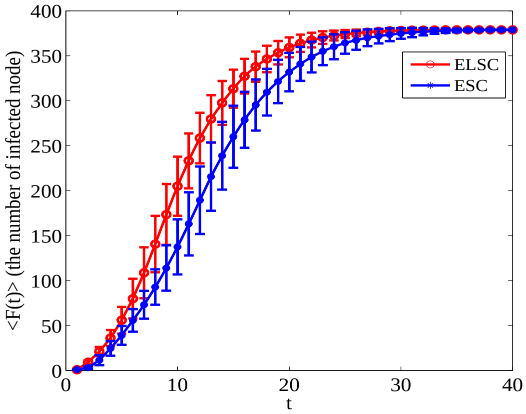
L	ELSC	ESC	EVC
1	4	5	4
2	5	4	26
3	16	16	5
4	26	15	51
5	15	45	95
6	45	46	67
7	70	47	16
8	231	176	52
9	67	177	169
10	51	1	70

clear that the curve for ELSC is located between the curves for ESC and EVC, in other words, ELSC can achieve better performance than ESC, but EVC performs slightly better than ELSC.

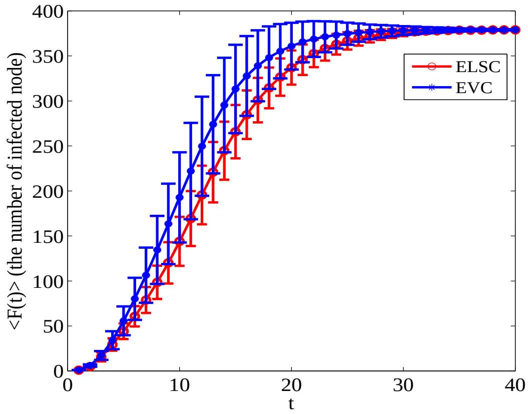
Table 4 shows the top-10 nodes generated by ELSC, ESC and EVC, respectively. We compare the spreading ability of different nodes in the top-L (top 5 and top 10) by the proposed method or each of the two centrality measures, and the results are shown in Figure 7 and Figure 8. From the error bar of Figure 7 and Figure 8, it is observed that no matter whether the value of L is 5 or 10, both reveal that the proposed method performs a quicker spreading than ESC, but spreading speed of the top-L nodes ranked by EVC is a little faster than that by the proposed method. Thus, to some extent, our method is effective as well for identifying the influential nodes in Netscience network.

5. **Conclusions.**

In this paper, we propose a new approach to identify influential nodes in weighted network called evidential local structure centrality which is based on the Dempster-Shafer theory of evidence. The proposed centrality measure considers not only the degree and strength of a node, but also the topological connections among the neighbors in weighted network. Firstly, the value of modified evidential centrality is calculated

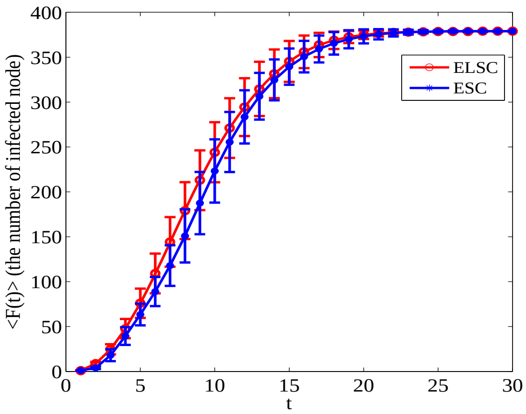


(a) Comparison by ELSC or ESC

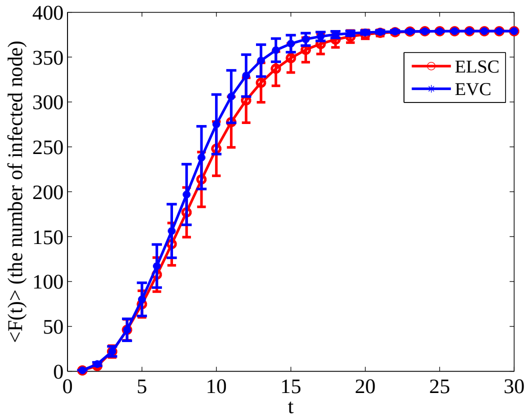


(b) Comparison by ELSC or EVC

FIGURE 7. The number of infected nodes by initially infected nodes in the top-5 list under $\beta = 0.6$



(a) Comparison by ELSC or ESC



(b) Comparison by ELSC or EVC

FIGURE 8. The number of infected nodes by initially infected nodes in the top-10 list under $\beta = 0.4$

by taking actual degree distribution. Secondly, local structure centrality combined with modified evidential centrality is extended to be applied in weighted networks. In order to verify the performance of ELSC, we make the experiments on three real networks. From the experimental results, we observe that ranked list of spreading ability of nodes by ELSC is more accurate than that by other centrality measures such as ESC, EVC. Moreover, we adopt the susceptible-infected (SI) model to simulate the epidemic spreading process of the top-L nodes, and it shows that ELSC is effective under the SI model. Experimental results on three real networks show that our approach can well identify influential nodes in weighted networks.

Acknowledgment. This work is supported by the National Natural Science Foundation of China under Grant No. 61170190, No. 61472341 and the Natural Science Foundation of Hebei Province China under Grant No. F2013203324, No. F2014203152 and No. F2015203326.

REFERENCES

[1] Y. Qin, X. X. Zhong, H. Jiang and Y. Y. Ye, An environment aware epidemic spreading model and immune strategy in complex networks, *Applied Mathematics and Computation*, vol.261, pp.206-215, 2015.

[2] L. Hebert-Dufresne, A. Allard, J. G. Young and L. J. Dube, Global efficiency of local immunization complex networks, *Nature Scientific Reports*, vol.3, no.2171, 2013.

[3] W. H. Li, S. T. Tang, S. Pei, S. Yan, S. J. Jiang, X. Teng and Z. M. Zheng, The rumor diffusion process with emerging independent spreaders in complex networks, *Physica A*, vol.397, pp.121-128, 2014.

[4] X. Huang, L. Vodenska, F. Wang, S. Havlin and H. E. Stanley, Identifying influential directors in the unitedstates corporate governance network, *Phys. Rev. E*, vol.84, no.4, 2011.

[5] B. S. Sohn and J. E. Jung, A novel ranking model for a large-scale scientific publication, *Mobile Networks and Applications*, 2014.

[6] S. Wang, S. H. Xie, X. M. Zhang, Z. J. Li, P. S. Yu and X. Y. X. Shu, Future influence ranking of scientific literature, *Eprint Arxiv*, 2014.

[7] T. Opsahl, F. Agneessens and J. Skvoretz, Node centrality in weighted networks: Generalizing degree and shortest paths, *Social Networks*, vol.32, pp.245-251, 2010.

[8] S. Gao, J. Ma, Z. M. Chen, G. H. Wang and C. M. Xing, Ranking the spreading ability of nodes in complex networks based on local structure, *Physica A*, vol.403, pp.130-147, 2014.

[9] Q. Li, T. Zhou, L. Y. Lv and D. B. Chen, Identifying influential spreaders by weighted LeaderRank, *Physica A*, vol.404, pp.47-55, 2014.

[10] J. Bae and S. Kim, Identifying and ranking influential spreaders in complex networks by neighborhood coreness, *Physica A*, vol.395, pp.549-559, 2014.

[11] B. Wei, J. Liu, D. J. Wei, C. Gao and Y. Deng, Weighted k-shell decomposition for complex networks based on potential edge weights, *Physica A*, vol.420, pp.277-283, 2015.

[12] A. P. Dempster, Upper and lower probabilities induced by a multi-valued mapping, *Annals of Mathematics and Statistics*, vol.38, no.2, pp.325-339, 1967.

[13] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, 1976.

[14] D. Wei, X. Deng, X. Zhang, Y. Deng and S. Mahadevan, Identifying influential nodes in weighted networks based on evidence theory, *Physica A*, vol.392, pp.2564-2575, 2013.

[15] C. Gao, D. Wei, Y. Hu, S. Mahadevan and Y. Deng, A modified evidential methodology of identifying influential nodes in weighted networks, *Physica A*, vol.392, pp.5490-5500, 2013.

[16] D. J. Watts and S. H. Strogatz, Collective dynamics of small-world networks, *Nature*, vol.393, no.6684, pp.440-442, 1998.

[17] M. Bessedik, R. Laib, A. Boulmerka and H. Drias, Ant colony system for graph coloring problem, *CIMCA*, vol.1, pp.786-791, 2005.

[18] T. Harko, F. S. N. Lobo and M. K. Mak, Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates, *Applied Mathematics and Computation*, vol.236, pp.184-194, 2014.

[19] G. Yan, T. Zhou, J. Wang, Z. Fu and B. Wang, Epidemic spread in weighted scale-free networks, *Chinese Physics Letters*, vol.22, pp.510-513, 2005.

[20] T. Zhou, J. G. Liu, W. J. Bai, G. Chen and B. H. Wang, Behaviors of susceptible-infected epidemics on scale-free networks with identical infectivity, *Phys. Rev. E*, vol.74, 2006.

[21] R. Yang, B. H. Wang, J. Ren, W. J. Bai, Z. W. Shi, W. X. Wang and T. Zhou, Epidemic spreading on heterogeneous networks with identical infectivity, *Phys. Lett. A*, vol.364, no.3, pp.189-193, 2007.

[22] M. G. Kendall, A new measure of rank correlation, *Biometrika*, vol.30, no.1, pp.81-93, 1938.

[23] W. Zachary, An information flow model for conflict and fission in small groups, *Journal of Anthropological Research*, vol.33, 1977.

[24] *http://moreno.ss.uci.edu/data.html.lesmis.*

[25] M. E. J. Newman, Finding community structure in networks using the eigenvectors of matrices, *Phys. Rev. E*, vol.74, 2006.