Identifying Influencers in Social Networks

Xinyu Huang 📧, **Dongming Chen *** 📧, **Dongqi Wang and Tao Ren**

Software College, Northeastern University, Shenyang 110169, China; neuhxy@163.com (X.H.);
wangdq@swc.neu.edu.cn (D.W.); rent@swc.neu.edu.cn (T.R.)
*   Correspondence: chendm@mail.neu.edu.cn

Abstract: Social network analysis is a multidisciplinary research covering informatics, mathematics, sociology, management, psychology, etc. In the last decade, the development of online social media has provided individuals with a fascinating platform of sharing knowledge and interests. The emergence of various social networks has greatly enriched our daily life, and simultaneously, it brings a challenging task to identify influencers among multiple social networks. The key problem lies in the various interactions among individuals and huge data scale. Aiming at solving the problem, this paper employs a general multilayer network model to represent the multiple social networks, and then proposes the node influence indicator merely based on the local neighboring information. Extensive experiments on 21 real-world datasets are conducted to verify the performance of the proposed method, which shows superiority to the competitors. It is of remarkable significance in revealing the evolutions in social networks and we hope this work will shed light for more and more forthcoming researchers to further explore the uncharted part of this promising field.

Keywords: complex network; social network analysis; multilayer network; node influence

1. Introduction

The research of network science is experiencing a blossom in the last decade, which provides profound implications in very different fields, from finance to social and biological networks [1]. Considering the enormous data scale, most studies merely focus on a small group of influential nodes rather than the whole network. Take social networks for instance, influential nodes are those that have the most spreading ability, or playing a predominant role in the network evolution. Notably, a popular star in online social media may remarkably accelerate the spreading of rumors, and a few super spreaders [2] could largely expand the epidemic prevalence of a disease (e.g., COVID-19) [3]. The research of influencer identification is beneficial to understanding and controlling the spreading dynamics in social networks with diverse applications such as epidemiology, collective dynamics and viral marketing [4,5].

Nowadays, individuals interact with each other in more complicated patterns than ever. It is a challenging task to identify influencers in social networks for the various kinds of interactions. As we have known, the graph model is widely utilized to represent social networks, however, it is incapable of dealing with the multiple social links. For example, people use Facebook or WeChat to keep communication with family members or friends, use Twitter to post news, use LinkedIn to search for jobs, and use TikTok to create and share short videos [6]. It is easy to represent each social scenario via a graph model separately, in spite of they are belonging to the same group of individuals. The neglect of the multiple relationships between social actors may lead to an incorrect result of the most versatile users [7]. With the proposal of multilayer networks [8,9], we are able to encode the various interactions, which is of great importance and necessity of identifying influencers in multiple social networks.

In this paper, we design a novel node centrality measure for monolayer network, and then apply it to multilayer networks to identify influencers in multiple social networks. This method is solely based on the local knowledge of a network's topology in order to be fast and scalable due to the huge size of networks, and thus suitable for both real-time applications and offline mining.

The rest of this paper is organized as follows. Section 2 introduces the related works on influencers identification in monolayer network and multilayer networks. Section 3 presents the mathematical model and the method for detecting influencers. Section 4 exhibits the experiments and analysis, including comparison experiments on twenty-one real-world datasets, which verifies the feasibility and veracity of the proposed method. Section 5 summarizes the whole paper and provides concluding remarks.

2. Related Works

The initial research on influencers identification may date back to the study of node centrality, which means to measure how "central" a focal node is [10]. A plethora of methods for influencers identification are proposed in the past 40 years, which can be mainly classified into centrality measures, link topological ranking measures, entropy measures, and node embedding measures [11,12]. Some of these measures take only the local information into account, while others even employ machine learning methods. Nowadays, it has been one of the most popular research topics and yielded a variety of applications [7] such as identifying essential proteins and potential drug targets for the survival of the cell [13], controlling the outbreak of epidemics [14], preventing catastrophic outages in power grids [15], driving the network toward a desired state [16], improving transport capacity [17], promoting cooperation in evolutionary games [18], etc. This paper investigates the problem of identifying influencers in social networks, by introducing a family of centrality-like measures and gives a brief comparison in Table 1.

Degree Centrality (DC) [19] is the simplest centrality measure, which merely counts how many social connections (i.e., the number of neighbors) a focal node has, defined as

$$DC(i) = \sum_{j}^{N} a_{ij},$$

(1)

where N is the total number of nodes, a_{ij} is the weight of edge (i, j) if i is connected to j, and 0 otherwise. The degree centrality is simple and merely considers the local structure around a focal node [20]. However, this method is probably mistaken for the negligence of global information, i.e., a node might be in a central position to reach others quickly although it is not holding a large number of neighbors [21]. Thus, Betweenness Centrality (BC) [22] is proposed to assess the degree to which a node lies on the shortest path between two other nodes, defined as

$$BC(i) = \sum_{s \neq i, s \neq t, i \neq t} \frac{g_{st}(i)}{g_{st}},$$

(2)

where g_{st} is the total number of shortest paths, $g_{st}(i)$ is the shortest path between s and t that pass through node i. The betweenness centrality considers global information and can be applied to networks with disconnected components. However, there is a great proportion of nodes that do not lie on the shortest path between any two other nodes, thereby the computational result receives the same score of 0. Besides, high computational complexity is also a limitation of applying for large-scale networks. Analogously, Closeness Centrality (CC) [23] is proposed to represent the inverse sum of shortest distances to all other nodes from a focal node, defined as

$$CC(i) = \frac{N - 1}{\sum_{j \neq i} d_{ij}},$$

(3)

where N is the total number of nodes, d_{ij} is the shortest path length from node i to node j. The closeness centrality is capable of measuring the core position of a focal node via the utilization of global shortest path length, while it suffers from the lack of applicability to networks with disconnected components, e.g., if two nodes that belong to different components do not have a finite distance between them, it will be unavailable. Besides, it is also criticized by high computational complexity.

Eigenvector Centrality [24] (EC) is a positive multiple of the sum of adjacent centralities. Relative scores are assigned to all nodes in a network based on an assumption that connections to high-scoring nodes contribute more to the score of the node than connections to low-scoring nodes, defined as

$$EC(i) = k_1^{-1} \sum_j A_{ij} x_j, \tag{4}$$

where k_i depicts the eigenvalue of adjacency matrix A, $x = k_1 A x$ depicts the eigenvector stable state of interactions with eigvenvalue k_1^{-1}. This measure considers the number of neighbors and the centrality of neighbors simultaneously, however, it is incapable of dealing with non-cyclical graphs. In 1998, Brin and Page developed the PageRank algorithm [25], which is the fundamental search engine mechanism of Google. PageRank (PR) is a positive multiple of the sum of adjacent centralities, defined as

$$PR_k(i) = \sum_{j=1}^{N} a_{ji} \frac{PR_{k-1}(j)}{k_j^{out}}, i = 1, 2, \ldots, N. \tag{5}$$

where N depicts the total number of nodes, $\sum_{i=1}^{N} PR_0(i) = 0$, k_j^{out} is the number of edges from node j point to i. Likewise, this method is efficient but also criticized by non-convergence in cyclical structures. As we have known, the clustering coefficient [26,27] is a measure of the degree to which nodes in a graph tend to cluster together, defined as

$$C_i = \frac{\sum_{j \neq i, k \neq j, k \neq i} a_{ij} a_{ik} a_{jk}}{\sum_{j \neq i, k \neq j, k \neq i} a_{ij} a_{ik}}. \tag{6}$$

It is widely considered that a node with a higher clustering coefficient may benefit forming communities and enhancing local information spreading. However, Chen et al. expressed contrary views that the local clustering has negative impacts on information spreading. They proposed a ClusterRank algorithm for ranking nodes in large-scale directed networks and verified its superiority to PageRank and LeaderRank [28]. Therefore, the effect of clustering coefficient on information spreading is uncertain, which may benefit local information spreading but prohibit global (especially directional network) information spreading. In 2016, Ma et al. proposed a gravity centrality [29] (GR) by considering the interactions comes from the neighbors within three steps, defined as

$$G(i) = \sum_{j \in \psi_i} \frac{k_s(i) k_s(j)}{d_{ij}^2}, \tag{7}$$

$$G_+(i) = \sum_{j \in \Gamma_i} G(j), \tag{8}$$

where $k_s(i)$ and $k_s(j)$ are the k-shell index of i and j, respectively. ψ_i is the neighborhood set whose distance to node i is less than or equal to 3, d_{ij} is the shortest path length between i and j. These methods consider semi-local knowledge of a focal node, i.e., the neighboring nodes within three steps, which are successful in many real-world datasets, such as Jazz [30], NS [31] and USAir network [32], etc. However, they are also with high computational complexity by globally calculating k-shell. In 2019, Li

et al. improved the gravity centrality and proposed a Local-Gravity centrality (LGR) [33] by replacing k-shell computing and merely considering the neighbors within R steps, defined as

$$LG_R(i) = \sum_{d_{ij} \leq R, j \neq i} \frac{k_i k_j}{d_{ij}^2}, \tag{9}$$

where k_i and k_j are the degrees of i and j, respectively, d_{ij} is the shortest path length between i and j. This method had been extremely successful in a variety of real-world datasets, however, the parameter R requires the calculating of network diameter, which is also a time-consuming process.

Table 1. Classical node centrality metrics comparison.

Metric	Topology	Complexity	Advantages	Disadvantages
DC [19]	Local	$O(n)$	simple	incapable of dealing with "bridge" nodes
BC [22]	Global	$O(nm + n^2 \log n)$	finding "bridge-like" nodes	cannot differentiate most marginal nodes
CC [23]	Global	$O(nm + n^2 \log n)$	finding "nearest" nodes	incapable in disconnected graphs
EC [24]	Global	$O(n + m)$	consider both of the quality and quantity of neighbors	may be non-convergent
PR [25]	Global	$O(n + m)$	efficient, widely applied in search engine	may be non-convergent
H-index [34]	Semi-local	$O(n + m)$	famous for academic evaluation	lack of global information
k-shell [14]	Global	$O(n + m)$	suitable for large-scale networks	indistinguishable
LR [35]	Global	$O(n + m)$	no parameters; robustness	may be non-convergent
GR, GR+ [29]	Global	$O(n^3)$	accuracy	high complexity in k-shell
LGR [33]	Semi-local	$O(n^2)$	simple and capable in most cases	additional parameters R determination

The above-mentioned centrality measures have been utilized to rank nodes' spreading abilities in monolayer networks. The ranking of nodes in multilayer networks is a more challenging task and is still an open issue. The information propagation process over multiple social networks is more complicated, and conventional models are incapable without any modifications. Zhuang and Yağan [36] proposed a clustered multilayer network model, where all constituent layers are random networks with high clustering to simulate the information propagation process in multiple social networks. Likewise, Basaras et al. [37] proposed an improved susceptible–infected–recovered (SIR) model with information propagation probability parameters (i.e., λ_{ii} for intralayer connections and λ_{ij} for interlayer connections). Most of the recent endeavors concentrated on the multiplex networks, (e.g., clustering coefficient in multiplex networks [38]), where all layers share the identical set of nodes but may have multiple types of interactions. Rahmede et al. proposed a MultiRank algorithm [39] for the weighted ranking of nodes and layers in large multiplex networks. The basic idea is to assign more centrality to nodes that are linked to central nodes in highly influential layers. The layers are more influential if highly central nodes are active in them. Wang et al. proposed a tensor decomposition method (i.e., EDCPTD centrality) [7] , which utilize the fourth-order tensor to represent multilayer networks and identify essential nodes based on CANDECOMP/PARAFAC (CP) tensor decomposition. They also exhibited the superiority to traditional solutions by comparing the performance of the proposed method with the aggregated monolayer networks. In a word, it is of great significance in identifying influencers in multiplex networks. Our purpose in this work is to devise a measure that can accurately detect influential nodes in a general multilayer network.

3. Modeling and Methods

3.1. Network Modeling

The problem of finding influential nodes is described as extracting a small set of nodes that can bring the greatest influence on the network dynamics. With a given network model $G = (V, E)$, where $V = \{v_1, v_2, ..., v_n\}$ is the node set and $E = \{(v_i, v_j)\}, (v_i, v_j \in V)$ is the edge set. The identification of influential nodes is to pick a minimum of nodes as the initial seeds, which can achieve the maximum influenced scope, described as

$$A = \arg\min_{A \in V} \max\{\sigma(A)\}, \tag{10}$$

where A is the initially infected nodes, $\sigma(A)$ denotes the final influenced node set. This problem is simplified as top-k influencers identification by additional setting $|A| = k$, which has recently attracted great research interests [40–42]. A variety of real-world social networks are, in fact, interconnected by different types of interactions between nodes, forming what is known as multilayer networks. In this paper, we employ a multilayer network model [9], which can represent nodes sharing links in different layers. The multilayer network model is defined as

$$\mathcal{M} = (\mathcal{G}, \mathcal{C}), \tag{11}$$

where $\mathcal{G} = \{G_\alpha; \alpha \in \{1, \dots, L\}\}$ is a family of (directed or undirected, weighted or unweighted) graphs $G_\alpha = (V_\alpha, E_\alpha)$, which represents layers of $\mathcal{M}$ and $\mathcal{C}$ depicts the interactions between nodes of any two different layer, given by

$$\mathcal{C} = \{E_{\alpha\beta} \subseteq V_\alpha \times V_\beta; \alpha, \beta \in 1, \dots, L, \alpha \neq \beta\}, \tag{12}$$

The corresponding supra-adjacency matrix can be represented as

$$\mathcal{M} = \begin{bmatrix} A_1 & I_{12} & \cdots & I_{1L} \\ I_{21} & A_2 & \cdots & I_{2L} \\ \vdots & \vdots & \ddots & \cdots \\ I_{L1} & I_{L2} & \cdots & A_L \end{bmatrix} \in \mathcal{R}^{N \times N}, \tag{13}$$

where $A_1, A_2, ..., A_L$ are the adjacency matrix of layer $1, 2, ..., L$, respectively. N is the total number of the nodes, which can be calculated by $N = \sum_{1 \leq l \leq L} |V^l|$. The non-diagonal block $I_{\alpha\beta}$ represents the inter-layer edges of layer α and layer β. Thus, the interlayer edges can be represented as

$$I = \bigcup_{\alpha, \beta = 1, \alpha \neq \beta}^{L} I_{\alpha\beta}. \tag{14}$$

Take the 9/11 terrorists network [43] for instance, the edges are classified into three categories (i.e., layers) according to the observed interactions which are plotted in Figure 1.

Figure 1. Network of 9/11 terrorists. (**a**) The monolayer network representation, the size of a node represents its degree; (**b**) The 9/11 terrorists' interactions represented by a multilayer network model, where L_1 presents confirmed close contact, L_2 layer shows various recorded interactions, L_3 contains potential or planed or unconfirmed interactions; (**c**) The super-adjacency matrix representation.

3.2. Methods

We employ the susceptible–infected–recovered (SIR) spreading model [44] as the influence analysis model. It has three possible states:

- Susceptible (S) state, where a node is vulnerable to infection.
- Infectious (I) state, where a node tries to infect its susceptible neighbors.
- Recovered (R) state, where a node has recovered (or isolated) and can no longer infect others.

In a network, if two nodes are connected then they are considered to have "contact". If one node is "infected", and the other is susceptible, then with a certain probability the latter may become infected through contact [45]. A node is considered to be recovered if it is isolated or immune to the disease. In detail, to check the spreading influence of one given node, we set this node as an infected node and the other nodes are susceptible nodes. At each time step, each infected node can infect its susceptible neighbors with infection probability β, and then it recovered from the diseases with probability γ, the differential equations are shown in Figure 2. For simplicity, we set $\gamma = 1$. The process of the SIR model is plotted in Figures 3 and 4 with the famous Krackhardt's Kite network [46].

$$
\begin{cases}
\dfrac{dS}{dt} = -\beta S I \\[2mm]
\dfrac{dI}{dt} = \beta S I - \gamma I \\[2mm]
\dfrac{dR}{dt} = \gamma I
\end{cases}
$$
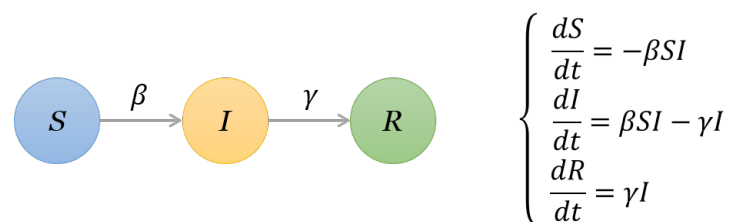
Figure 2. The illustration of parameters in the susceptible–infected–recovered (SIR) model and corresponding differential equations.
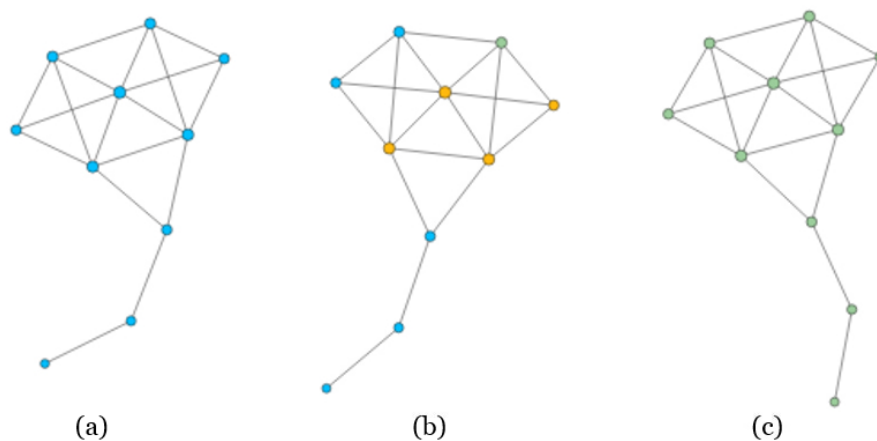
Figure 3. The process of SIR model on Krackhardt's Kite network. In panel (**a**), all the nodes are in Susceptible state; while we select one node to be infected, and the neighbors will be infected soon, as shown in panel (**b**); Finally, the network will reach a stable state, i.e., the number of recovered nodes will reach a maximum, as shown in panel (**c**).
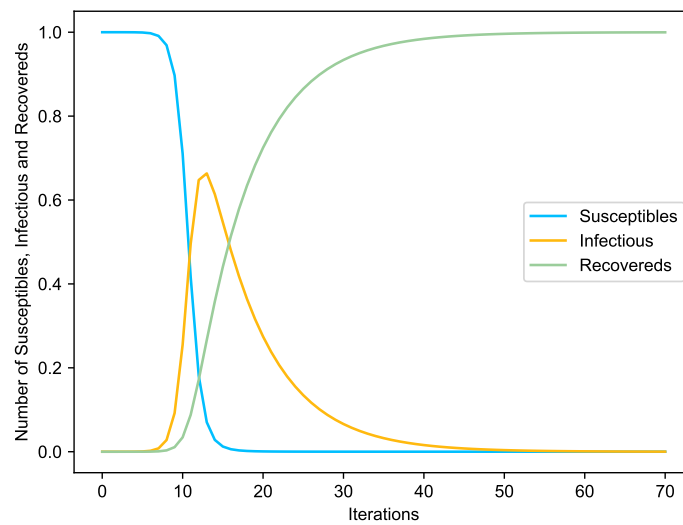
Figure 4. The varying susceptible, infectious and recovered nodes with the increasing of iterations.

In this paper, we define the node influence (INF, for short) as the energy derived from the neighbors, given as

$$INF_R(i) = \sum_{d_{ij} \leq r} \sum_{j \in \Gamma(i)} \frac{d_{ij}^2 w_{ij}}{k_j}, \tag{15}$$

where R is the truncation radius, $\Gamma(i)$ is the set of neighbors of node i, d_{ij} is the shortest path length between node i and node j, k_j is the degree of node j, w_{ij} is the weight of edge e_{ij}. For unweighted networks, $w_{ij} = 1$. Analogously, we apply the proposed INF measure to multilayer networks (represented as $INF_R^{\mathcal{M}}$) by the following modifications

$$INF_R^{\mathcal{M}}(i) = \sum_{d_{ij} \leq R} \sum_{\alpha < L} \sum_{j \in \Gamma^\alpha(i)} \frac{d_{ij}^2 w_{ij}}{k_j^\alpha}, \tag{16}$$

where R is the truncation radius, $\Gamma^\alpha(i)$ is the set of neighbors of node i at layer α, k_j^α is the degree of node j at layer α, d_{ij} is the shortest path length between node i and node j. For simplicity, we choose $R = 1$, thus $d_{ij} = 1$ if node i and node j is connected through an intralayer edge or interlayer edge, and 0 otherwise.

To explain the effect, we take the above-mentioned Krackhardt's Kite network (as plotted in Figure 5) and the 9/11 terrorists network (as plotted in Figure 1) for examples. The nodes centralities in Krackhardt's Kite network are shown in Table 2.
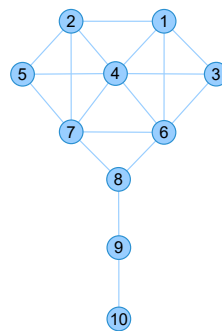
Figure 5. The network structure of Krackhardt's Kite network.

Table 2. Classic node centralities comparison of kite network. The maximum centralities are marked in bold.

| ID	Name	Degree	Betweenness	Closeness	Katz	Eigenvector	INF
1	Andre	0.4444	0.0231	0.5294	0.3307	0.3522	0.7211
2	Beverley	0.4444	0.0231	0.5294	0.3307	0.3522	0.7211
3	Carol	0.3333	0.0000	0.5000	0.3006	0.2858	0.6495
4	Diane	**0.6667**	0.1019	0.6000	**0.3907**	0.4810	**0.8273**
5	Ed	0.3333	0.0000	0.5000	0.3006	0.2858	0.6495
6	Fernando	0.5556	0.2315	**0.6429**	0.3595	**0.3977**	0.7830
7	Garth	0.5556	0.2315	**0.6429**	0.3595	**0.3977**	0.7830
8	Heather	0.3333	**0.3889**	0.6000	0.2887	0.1959	0.7109
9	Ike	0.2222	0.2222	0.4286	0.2431	0.0481	0.7914
10	Jane	0.1111	0.0000	0.3103	0.2168	0.0112	0.6225

As shown in Table 2, Node 4 is considered to be the most important node under the Degree, Katz and the proposed INF measure, while Node 8 has greater Betweenness, Node 6 and node 7 has greater Closeness or (Eigenvector). Thus, the node list (i.e., [4, 6, 7, 8]) is considered to be the influencers. Furthermore, to evaluate the nodes' influence, we set each node as the initially infected and recorded the final recovered nodes, respectively. This process is repeated for 10,000 times and the results are shown in Table 3.

Table 3. Averaging recovered nodes and iterations times of each node as initially infected spreaders under 10,000 times SIR stimulations with parameters setting $\beta = 0.35$, $\gamma = 1$.

ID	Name	Recovered Nodes	Iterations
1	Andre	4.8015	3.3399
2	Beverley	4.7902	3.3392
3	Carol	4.3485	3.1529
4	Diane	**5.3182**	3.3684
5	Ed	4.3336	3.1455
6	Fernando	5.0856	3.3251
7	Garth	5.0338	3.2994
8	Heather	4.0765	2.9625
9	Ike	2.6060	2.1806
10	Jane	1.8086	1.7019

As shown in Table 3, Node 4 (i.e., Diane), which is considered to be more influential under Degree, Katz, and INF centrality, shows more recovered nodes (i.e., 5.3182) after 10,000 times SIR stimulations. This experiment is available at https://neusncp.com/api/sir. Analogously, we conduct experiments on the three-layer 9/11 terrorists network. Particularly, we set the infected probability between intralayer edges as β and the probability between interlayer edges as $\beta^{\mathcal{M}} = w_{ij}\beta$. The experimental results are plotted in Figure 6.

By conducting SIR simulations on the three-layer 9/11 terrorists network, we can obtain the influential nodes of each layer by calculating the number of finally recovered nodes. Afterward, we sort the nodes by the averaging recovered nodes, and compare the order with the results computed from the proposed INF indicator. It is shown in Figure 6 that the compared values (i.e., recovered nodes and INF) are in the same tendency, which verifies the feasibility of the proposed INF measure. Notably, several influential nodes, such as "Essid Sami Ben Khemais", "Mohamed Atta", and "Marwan Al-Shehhi" are also in the central position of the network, as shown in Figure 1a.
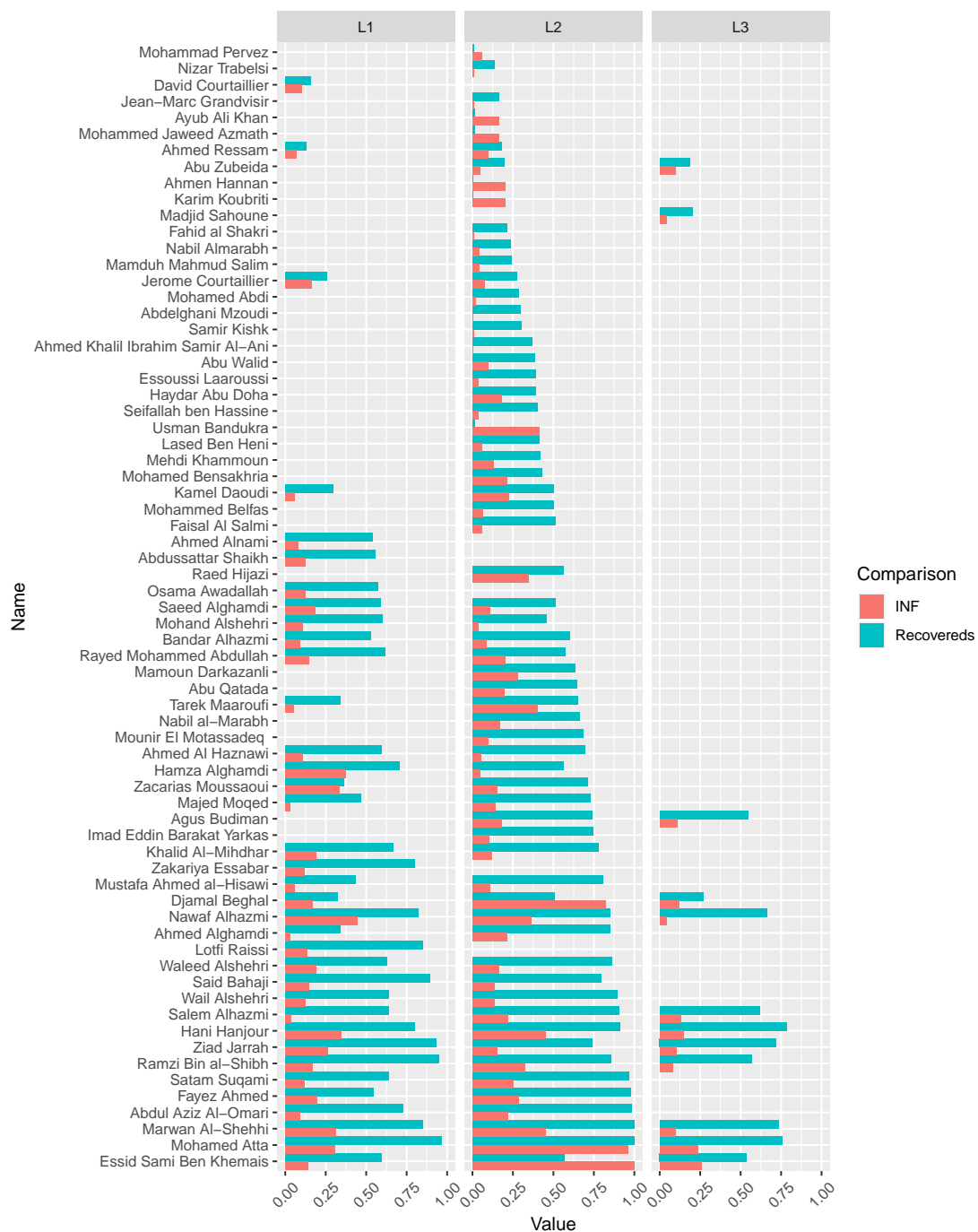
Figure 6. The comparison of recovered nodes in 9/11 terrorists network by initially setting every node infected separately. The recovered nodes and the corresponding INF values are normalized, which are marked in cyan and red, respectively. The comparison of the nodes in three layers are plotted in facet L_1, L_2 and L_3.

The experimental results on the two sample networks show the feasibility of the proposed measure on monolayer and multilayer networks, respectively. Experiments on more real-world networks will be given in Section 4.

3.3. Complexity Analysis

Suppose m and n are the numbers of edges and nodes, respectively, L is the number of layers, the average degree of nodes is d, R is the truncation radius (commonly setting as $R = 1$). The complexity

of INF for monolayer network is $O(n + d^R)$. As for multilayer networks, the computational complexity is $O(n + Ld^R)$, where L is also a small positive integer. Thus, the time complexity is acceptable as $O(n + Ld)$. Overall, the proposed measure considers more neighboring information than the degree centrality and has a lower computational complexity than betweenness centrality and closeness centrality (i.e., $O(nm + n^2 \log n)$).

4. Experiments and Discussion

The experimental environment was with Intel(R) Core (TM) i5-7200U CPU @ 2.50 GHz (4 CPUs), 2.7 GHz, the memory was 8 GB DDR3. The operating system was Windows 10 64 bit, the programming language was Python 3.7.1, and the relevant libs were NetworkX 2.2 and Multinetx. The goal of the experiments was to compare the performance of the proposed INF measure with competitive indicators.

4.1. Experimental Datasets

In this paper, 21 real-world datasets were employed to verify the performance of the proposed method, which were classified into two groups. The first group covered 12 monolayer networks, which comprised four social networks (i.e., Club, Dolphins, 911 and Lesmis), three biological networks (i.e., Escherichia, C.elegans and DMLC), collaboration networks (i.e., Jazz and NS), a communication network (i.e., Eron), a power network (i.e., Power) and a transport network (i.e., USAir), as shown in Table 4.

Table 4. Statistics of 12 real-world monolayer networks.

Dataset Name	$|V|$	$|E|$	$<k>$	$<d>$	$|C|$	r	$|H|$	β_c
Club [47]	34	78	4.5882	2.4082	0.5706	−0.4756	1.6933	0.1477
Dolphins [48]	62	159	5.1290	3.3570	0.2590	−0.0436	1.3268	0.1723
911 [49]	69	159	4.6087	2.4672	0.4698	−0.0380	1.7304	0.1434
Lesmis [50]	77	254	6.5974	2.6411	0.5731	−0.1652	1.8273	0.0905
Escherichia [51]	97	212	4.3711	5.5369	0.3675	0.4116	1.2367	0.2270
Eron [52]	143	623	8.7133	2.9670	0.4339	−0.0195	1.4829	0.0839
Jazz [30]	198	2742	27.6970	2.2350	0.6175	0.0202	1.3951	0.0266
USAir [32]	332	2126	12.8072	2.7381	0.6252	−0.2079	3.4639	0.0231
NS [31]	379	914	4.8232	6.0419	0.7412	−0.0817	1.6630	0.1424
C.elegans [53]	453	2032	9.0066	2.6638	0.6465	−0.2197	4.4782	0.0254
DMLC [54]	659	1570	4.7648	2.6370	0.3279	−0.1914	14.8897	0.0143
Power [26]	4941	6594	2.6691	18.9892	0.0801	0.0035	1.4504	0.3483

Note: $|V|$ and $|E|$ denotes the number of nodes and edges, respectively. $<k>$ is the average degree; $<d>$ is the average shortest path length; $|C|$ is the average clustering index; $<r>$ is the assortativity coefficient; $|H|$ is the degree heterogeneity and β_c represents the epidemic threshold of the SIR model. **Club** contains the friendships between the 34 members of a karate club at a US university. **Dolphins** dataset is a animals social network. **911** represents a monolayer terrorist network of September 11 attacks. **Lesmis** is the coappearance network of characters in the novel Les Miserables. **Escherichia** represetns transcriptional regulation networks in cells orchestrate gene expression, where nodes are operons, and each edge is directed from an operon that encodes a transcription factor to an operon that it directly regulates (an operon is one or more genes transcribed on the same mRNA). **Eron** is a email network collected from Eron company. **Jazz** lists the collaboration patterns of jazz musicians. **USAir** is an undirected weighted network as obtained by considering the 500 US airports with the largest amount of traffic from publicly available data. Nodes represent US airports and edges represent air travel connections among them. **NS** represents coauthorships between 379 scientists whose research centers on the properties of networks of one kind or another. **C.elegans** represents the edges of the metabolic network of C.elegans. **DMLC** represents the inferred Links by small/medium-scale rotein-protein interactions (collected from protein-protein interaction data bases). **Power** is a power grid of the western United States.

The second group covered nine multilayer networks, which comprised six social networks (i.e., Padgett, Krackhardt, Vickers, Kapferer, Lazega and CS-Aarhus), two transport networks (i.e., LondonTransport and EUAirTransportation) and a biological network (i.e., humanHIV), as shown in Table 5. Data availability: http://www.neusncp.com/user/file?id=12&code=data.

Table 5. Statistics of nine real-world multilayer networks.

Dataset Name	$|L|$	$|V|$	$|E|$	$|E_A|$	$|E_C|$	$<k>$	$<d>$	$|C|$
Padgett [55]	2	26	46	35	11	3.5385	2.6923	0.1441
Krackhardt [56]	3	63	307	244	63	9.746	2.1731	0.3943
Vickers [57]	3	87	605	518	87	13.908	2.1802	0.4823
Kapferer [58]	4	150	769	552	217	10.2533	2.5889	0.3002
Lazega [59,60]	3	211	2051	1842	209	19.4408	2.3958	0.3938
humanHIV1 [61]	5	1195	1504	1269	235	2.5172	4.1385	0.0221
CS-Aarhus [62]	5	224	948	620	328	8.4643	3.1847	0.3603
LondonTransport [63]	3	399	472	441	31	2.3659	14.2989	0.0243
EUAirTransportation [64]	37	2034	15199	3588	11611	14.9449	3.5087	0.5969

Note: $|L|$ denotes the number of layers; $|V|$ and $|E|$ are the total number of nodes and edges, respectively; $|E_A|$ and $|E_C|$ denote the number of intralayer edges and interlayer edges, respectively. $<k>$ is the average degree; $<d>$ is the average shortest path length; $|C|$ is the average clustering index; **Padgett** consists of 2 layers (marriage alliances and business relationships) describing florentine families in the Renaissance; **Krackhardt** consists of 3 kinds of relationships (Advice, Friendship and "Reports to") between managers of a high-tech company; **Vickers** is collected by Vickers from 29 seventh grade students in a school in Victoria, Australia. Students are asked to nominate their classmates on a number of three kinds of relations; **Kapferer** exhibits interactions in a tailor shop in Zambia (then Northern Rhodesia) over a period of ten months, where layers represent two different types of interaction, recorded at two different times (seven months apart) over a period of one month; **Lazega** consists of three kinds of interactions (Co-work, Friendship and Advice) between partners and associates of a corporate law partnership; **humanHIV1** represents the multiplex genetic and protein interactions network of the human HIV type 1; **CS-Aarhus** consists of five kinds of online and offline relationships (Facebook, Leisure, Work, Co-authorship, Lunch) between the employees of Computer Science department at Aarhus; **LondonTransport** is collected from the official website of Transport for London (https://www.tfl.gov.uk/). Nodes are train stations in London and edges encode existing routes between stations; **EUAirTransportation** is composed by thirty-seven different layers each one corresponding to a different airline operating in Europe.

4.2. Performance Comparison

To verify the performance of the proposed node influence in networks, this paper carries out a comparison experiment on the above-mentioned datasets: The nodes were removed by a certain indicator in descending order, and the number of subgraphs was recorded. This process repeated until there were not any nodes left. The varying tendency of the subgraphs' number exhibited the influence of a focal centrality. The experimental results are shown in Figure 7.

As shown in Figure 7, with the nodes removing, the number of subgraphs was increasing and reached a maximum when the network was totally broken up, i.e., there were no edges at this moment. Afterward, the number of subgraphs (i.e., the number of nodes) was decreasing and finally reached zero when all the nodes were removed. The maximum numbers of subgraphs were obtained by the proposed INF measure on all the datasets except C.Elegans. However, the result of C.Elegans obtained by INF was very close to the best situation of BC, which suggests the feasibility of the proposed INF measure.

We applied the SIR model to compare the rankings of influences calculated by each indicator among the above-mentioned networks. Initially, one node was set as "infected" state to infect its neighbors with probability β. Afterward, the infected nodes were recovered and never be infected again with probability γ. This spreading process repeated until there were no more infected nodes in the network. The influence of any node i can be estimated by

$$P(i) = \frac{N_R}{N},\tag{17}$$

where N_R is the number of recovered nodes after the spreading process, and N is the total number of nodes in the network. For simplicity, we set $\gamma = 1$ and the epidemic threshold was

$$\beta_c \approx \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}\tag{18}$$

After having obtained the standard nodes' influence sequence via SIR model simulations, we employed the Kendall's Tau coefficient [65] to compare the performance of each indicator. The Kendall's Tau coefficient is an index measuring the correlation strength between two sequences. Suppose given the standard sequence $X = (x_1, x_2, \ldots, x_N)$, and we obtained the computational sequence $Y = (y_1, y_2, \ldots, y_N)$ by a certain indicator. Any pair of two-tuples (x_i, y_i) and (x_j, y_j) $(x \neq j)$ are concordant if both $x_i > x_j$ and $y_i > y_j$ or $x_i < x_j$ and $y_i < y_j$. Meanwhile, they are considered as discordant, if $x_i > x_j$ and $y_i < y_j$ or $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$, pairs are neither concordant nor discordant. Therefore, Kendall's Tau coefficient is defined as

$$\tau = \frac{N_c - N_d}{0.5n(n-1)},\tag{19}$$

where N_c and N_d indicate the number of concordant and discordant pairs, respectively. The range of τ is $[-1, 1]$. Table 6 shows the computational Tau results with the comparison of standard sequence from SIR model simulations.
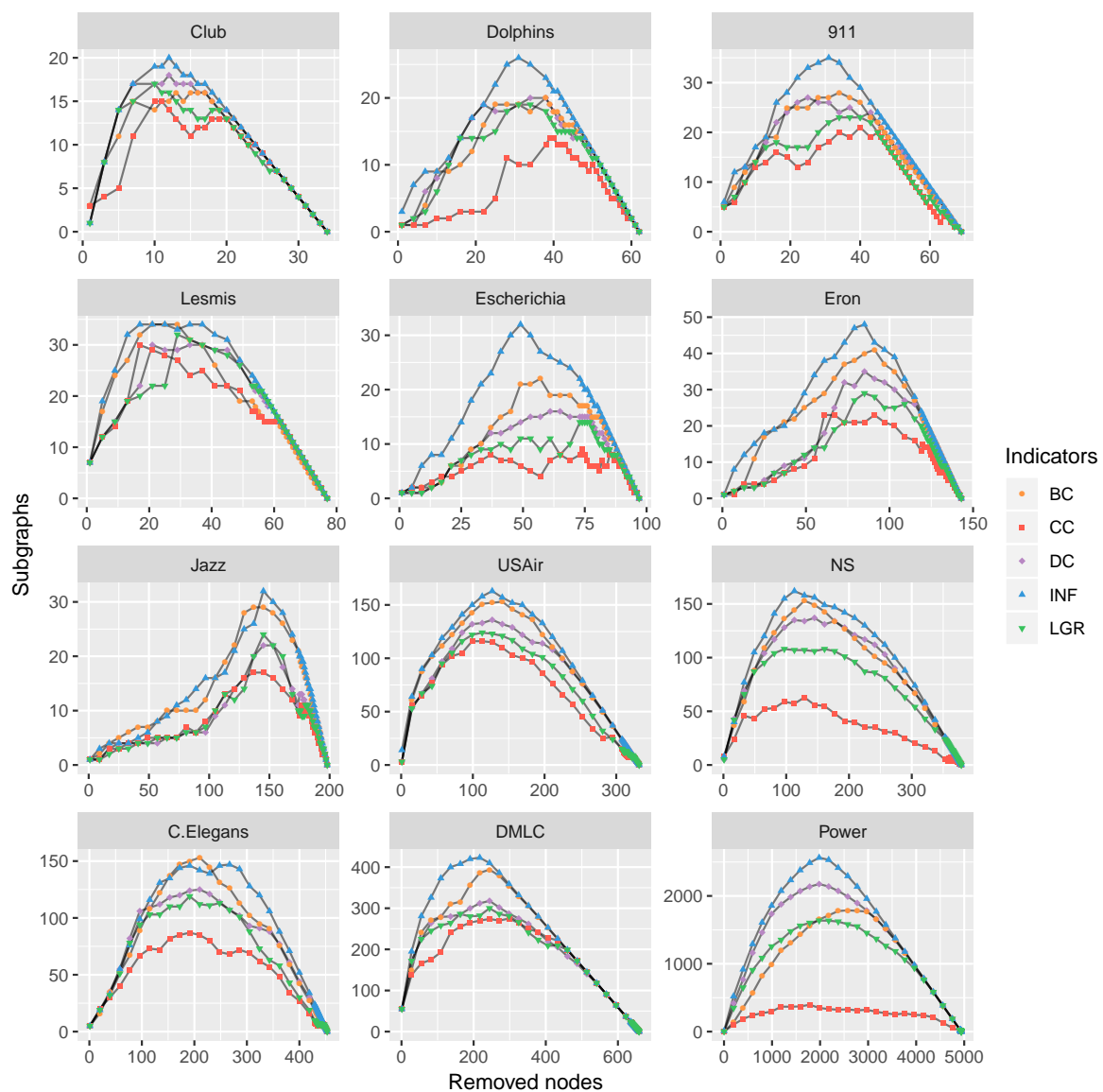
Figure 7. The varying subgraphs of 12 monolayer networks with the removal of the most influential nodes repeatedly with each centrality indicator.

As shown in Table 6, the proposed measure outperformed the competitors in most cases, even in the Escherichia network, the computed Tau result of INF (0.0692) was close to that of CC (0.0971). Thus, it was also competitive in this network.

If setting the limitation of identifying *k* influencers, we conducted experiments on the real-world datasets with top-*k* nodes by computational centrality nodes and compared the recovered nodes (i.e., the final number of nodes with recovered states). To compare the varying parameter *k* with the obtained *τ*, we conducted experiments on the above-mentioned datasets and set the ratio of β/β_c, as shown in Figure 8.
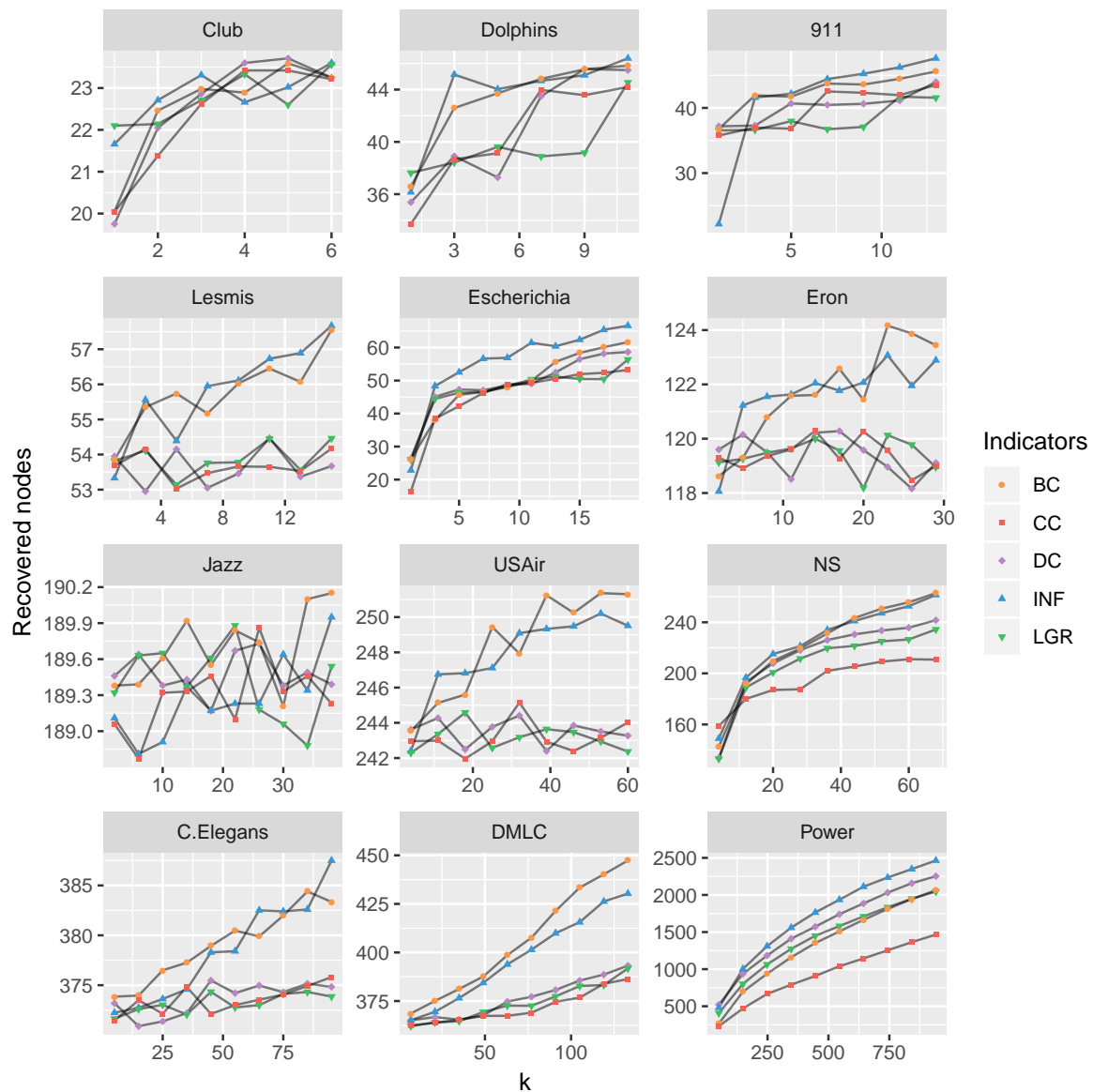
Figure 8. The algorithms' performance comparison for varying *k* (ranging from 0 to 20 percent of total nodes), measured by the recovered nodes. The betweenness centrality and the proposed INF measure are very competitive than the others.

Table 6. The indicators' accuracies measured by the Kendall's Tau (τ).

Dataset Name	DC	INF	LGR	BC	CC
Club	0.2442	**0.2513**	0.1515	0.1016	−0.0766
Dolphins	0.0238	**0.0344**	−0.0196	−0.0323	−0.0354
911	0.0878	**0.1918**	0.0426	0.0409	−0.0895
Lesmis	0.0909	**0.1114**	0.1032	−0.1839	−0.0519
Escherichia	0.0726	0.0692	0.0009	−0.0808	**0.0971**
Eron	0.0454	**0.0927**	−0.0056	0.0031	0.0107
Jazz	0.069	**0.0838**	0.0325	0.0244	0.0312
USAir	−0.0358	**0.0743**	−0.0402	−0.0491	−0.0311
NS	−0.0378	**0.0302**	−0.0089	−0.0143	−0.0462
C.elegans	0.0154	**0.0697**	0.0242	0.0185	0.0214
Power	−0.005	**0.0272**	0.0024	−0.0015	0.0209

Note: Given a network, the parameters of SIR model are given with the transmission probability $\beta = 0.35$ and recovering probability $\mu = 1$ for simplicity. To obtain the standard ranking of nodes' influences, we conducted 1000 independent simulations, in each process every node is selected once as the infect seed once. The best perfromed indicator for each network is emphasized by bold.

As shown in Figure 8, the proposed node influence method is quite competitive in most of the datasets, although second to the performance of betweenness indicator in DMLC and Jazz datasets.

Analogously, we conducted experiments on the nine multilayer networks by removing nodes with maximum centralities; the results are plotted in Figure 9.
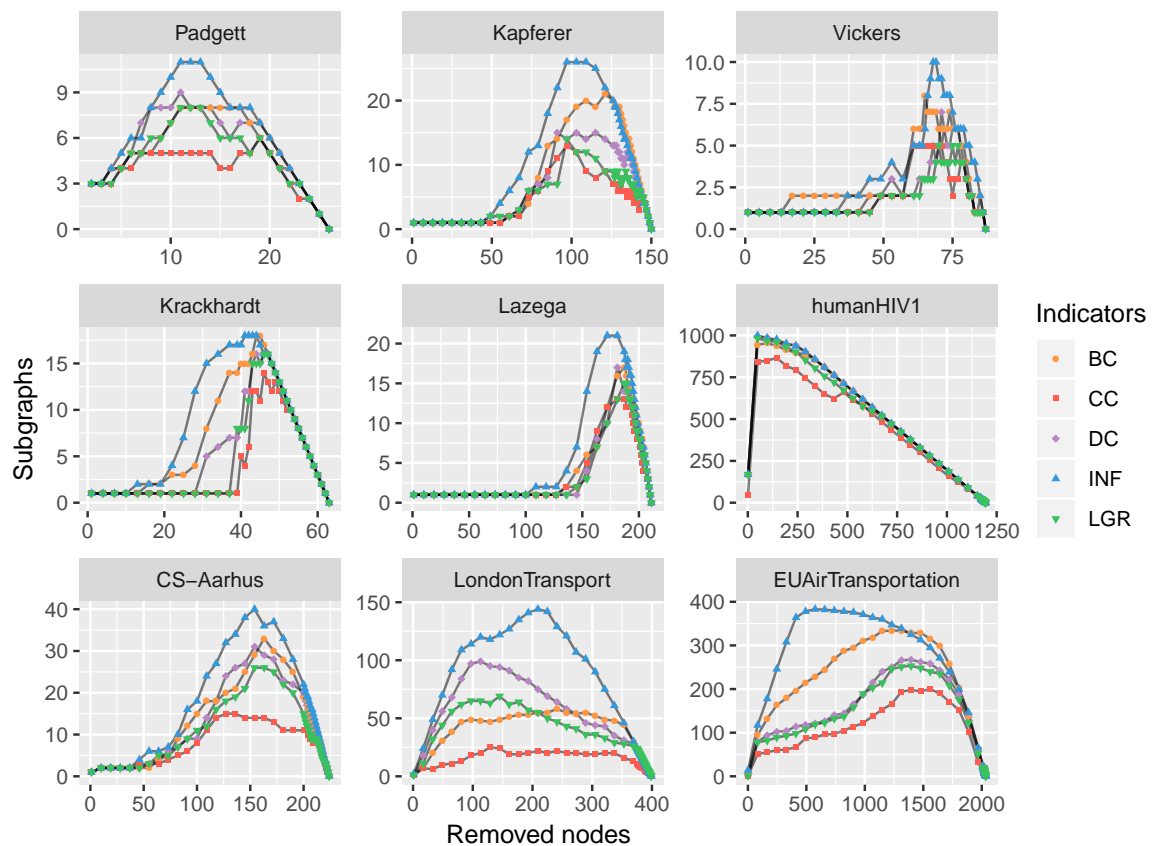
Figure 9. The computational time comparison of different indicators. The accumulation of running time on the 12 real-world datasets has exhibited that the proposed INF measure is much more efficient than the competitors.

To compare the time complexity of the proposed INF measure with classic methods, the runtime of 21 networks are recorded and shown in Figure 10.
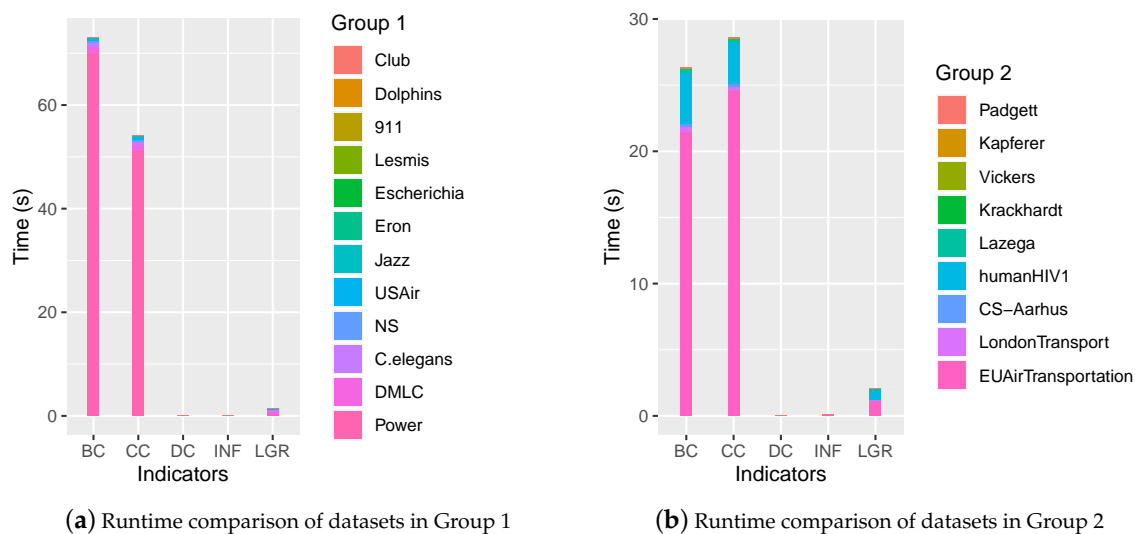

(**a**) Runtime comparison of datasets in Group 1

(**b**) Runtime comparison of datasets in Group 2

Figure 10. The runtime comparison of each indicator.

As shown in Figure 10, the runtime accumulated from either group indicated that the proposed INF measure was efficient, which was close to that of DC and superior to BC, CC and LGR.

4.3. Discussion

Influencers identification is a fundamental issue with wide applications in different fields of reality, such as epidemic control, information diffusion, viral marketing, etc. Currently, degree centrality [19] is the simplest method, which considers nodes with larger degrees are more influential. However, for the lack of global information, a node lying in a "bridge" position might be neglected for holding a small degree. The betweenness [22] and closeness [23] centrality consider global information, but they are holding a high complexity, which are not suitable for applications in large-scale networks. Local gravity is a balanced method, however, the determination of parameter *R* requires computing network diameter, which is also time-consuming. Thus, a novel node influence measure is proposed in this paper, which merely considers the local neighboring information of a focal node with the complexity of $O(n + Ld)$. Experimental results on 21 real-world datasets indicate the feasibility of the proposed measure.

Firstly, the experiments of counting subgraphs with removing influential nodes show that the capability of the proposed INF measure. By removing the nodes according to the INF indicator, the networks are more easily broken up, as shown in Figures 7 and 9. Secondly, we apply the SIR model to evaluate the node influence, which suggested the proposed INF measure is competitive to other indicators in most cases. Although inferior to BC on Jazz and DMLC networks, it is also competitive. By analyzing the structures of these two networks, we find that the nodes of Jazz network are densely connected (i.e., the average degree of 27.6970) and most of the nodes are holding the same number of neighbors (approximately 28 neighbors), which brings difficulties to identify which node is more influential. On the contrary, there is only one node (i.e., Node 2) holding a large number of neighbors (i.e., 439 neighbors) and the others only holding few neighbors (approximately four neighbors) in DMLC network, which is also difficult to identify influencers. Overall, the proposed method outperforms the other indicators in most cases. Finally, we compare the running time of each indicator on the 21 real-world datasets. Experimental results show the efficiency of the proposed measure.

5. Conclusions

Aiming at solving the problem of identifying influencers in social networks, this paper proposes a novel node influence indicator. This method merely considers the local neighboring information in order to be fast and suitable for applications in large-scale networks. Extensive experiments on 21 real-world datasets are conducted, and the experimental results show that the proposed method outperforms competitors. Afterwards, the time complexity is compared, and we verify the efficiency of the proposed indicator. Overall, the proposed node influence indicator is capable of identifying influencers in social networks. The contribution of this work is likely to benefit many real-world social applications, such as promoting network evolutions, preventing the spreading of rumors, etc.

As part of future works, the influencers in dynamic networks can be further studied by applying the proposed INF measure into a multilayer network model with numerous ordinal layers. The node's influence can be calculated by accumulating the local neighbors across all the layers. Besides, the effect of layers needs to be taken into consideration. In a word, we hope the findings in this work will help to improve the researches in this promising field.

References

1. Aldecoa, R.; Marín, I. Surprise maximization reveals the community structure of complex networks. *Sci. Rep.* **2013**, *3*, 1–9. [CrossRef]
2. Pei, S.; Muchnik, L.; Andrade, J.S., Jr.; Zheng, Z.; Makse, H.A. Searching for superspreaders of information in real-world social media. *Sci. Rep.* **2014**, *4*, 5547. [CrossRef] [PubMed]
3. Xu, Z.; Rui, X.; He, J.; Wang, Z.; Hadzibeganovic, T. Superspreaders and superblockers based community evolution tracking in dynamic social networks. *Knowl.-Based Syst.* **2019**, 105377. [CrossRef]
4. Malliaros, F.D.; Rossi, M.E.G.; Vazirgiannis, M. Locating influential nodes in complex networks. *Sci. Rep.* **2016**, *6*, 19307. [CrossRef] [PubMed]
5. Salavati, C.; Abdollahpouri, A.; Manbari, Z. BridgeRank: A novel fast centrality measure based on local structure of the network. *Phys. A Stat. Mech. Appl.* **2018**, *496*, 635–653. [CrossRef]
6. Huang, X.; Chen, D.; Ren, T. Social network coalescence based on multilayer network model. *J. Nonlinear Convex Anal.* **2019**, *20*, 1465–1474.
7. Wang, D.; Wang, H.; Zou, X. Identifying key nodes in multilayer networks based on tensor decomposition. *Chaos Interdiscip. J. Nonlinear Sci.* **2017**, *27*, 063108. [CrossRef]
8. Kivelä, M.; Arenas, A.; Barthelemy, M.; Gleeson, J.P.; Moreno, Y.; Porter, M.A. Multilayer networks. *J. Complex Netw.* **2014**, *2*, 203–271. [CrossRef]
9. Boccaletti, S.; Bianconi, G.; Criado, R.; Del Genio, C.I.; Gómez-Gardenes, J.; Romance, M.; Sendina-Nadal, I.; Wang, Z.; Zanin, M. The structure and dynamics of multilayer networks. *Phys. Rep.* **2014**, *544*, 1–122. [CrossRef]
10. Rodrigues, F.A. Network centrality: an introduction. In *A Mathematical Modeling Approach from Nonlinear Dynamics to Complex Systems*; Springer: Berlin, Germany, 2019; pp. 177–196.
11. Peng, S.; Zhou, Y.; Cao, L.; Yu, S.; Niu, J.; Jia, W. Influence analysis in social networks: A survey. *J. Netw. Comput. Appl.* **2018**, *106*, 17–32. [CrossRef]

12. Grover, A.; Leskovec, J. node2vec: Scalable feature learning for networks. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 855–864.

13. Jeong, H.; Mason, S.P.; Barabási, A.L.; Oltvai, Z.N. Lethality and centrality in protein networks. *Nature* **2001**, *411*, 41. [CrossRef] [PubMed]

14. Kitsak, M.; Gallos, L.K.; Havlin, S.; Liljeros, F.; Muchnik, L.; Stanley, H.E.; Makse, H.A. Identification of influential spreaders in complex networks. *Nat. Phys.* **2010**, *6*, 888. [CrossRef]

15. Albert, R.; Albert, I.; Nakarado, G.L. Structural vulnerability of the North American power grid. *Phys. Rev. E* **2004**, *69*, 025103. [CrossRef] [PubMed]

16. Liu, Y.Y.; Slotine, J.J.; Barabási, A.L. Controllability of complex networks. *Nature* **2011**, *473*, 167. [CrossRef]

17. Yan, G.; Zhou, T.; Hu, B.; Fu, Z.Q.; Wang, B.H. Efficient routing on complex networks. *Phys. Rev. E* **2006**, *73*, 046108. [CrossRef]

18. Xie, W.; Yu, W.; Zou, X. Diversity-maintained differential evolution embedded with gradient-based local search. *Soft Comput.* **2013**, *17*, 1511–1535. [CrossRef]

19. Bonacich, P. Factoring and weighting approaches to status scores and clique identification. *J. Math. Sociol.* **1972**, *2*, 113–120. [CrossRef]

20. McPherson, M.; Smith-Lovin, L.; Cook, J.M. Birds of a feather: Homophily in social networks. *Annu. Rev. Sociol.* **2001**, *27*, 415–444. [CrossRef]

21. Borgatti, S.P. Centrality and network flow. *Soc. Netw.* **2005**, *27*, 55–71. [CrossRef]

22. Freeman, L.C. A set of measures of centrality based on betweenness. *Sociometry* **1977**, *40*, 35–41. [CrossRef]

23. Freeman, L.C. Centrality in social networks conceptual clarification. *Soc. Netw.* **1978**, *1*, 215–239. [CrossRef]

24. Bonacich, P. Power and centrality: A family of measures. *Am. J. Sociol.* **1987**, *92*, 1170–1182. [CrossRef]

25. Brin, S.; Page, L. The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. ISDN Syst.* **1998**, *30*, 107–117. [CrossRef]

26. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world' networks. *Nature* **1998**, *393*, 440. [CrossRef] [PubMed]

27. Barrat, A.; Barthelemy, M.; Pastor-Satorras, R.; Vespignani, A. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3747–3752. [CrossRef] [PubMed]

28. Chen, D.B.; Gao, H.; Lü, L.; Zhou, T. Identifying influential nodes in large-scale directed networks: The role of clustering. *PLoS ONE* **2013**, *8*, e77455. [CrossRef] [PubMed]

29. Ma, L.l.; Ma, C.; Zhang, H.F.; Wang, B.H. Identifying influential spreaders in complex networks based on gravity formula. *Phys. A Stat. Mech. Appl.* **2016**, *451*, 205–212. [CrossRef]

30. Gleiser, P.M.; Danon, L. Community structure in jazz. *Adv. Complex Syst.* **2003**, *6*, 565–573. [CrossRef]

31. Newman, M.E. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8577–8582. [CrossRef]

32. Colizza, V.; Pastor-Satorras, R.; Vespignani, A. Reaction-diffusion processes and metapopulation models in heterogeneous networks. *Nat. Phys.* **2007**, *3*, 276–282. [CrossRef]

33. Li, Z.; Ren, T.; Ma, X.; Liu, S.; Zhang, Y.; Zhou, T. Identifying influential spreaders by gravity model. *Sci. Rep.* **2019**, *9*, 8387. [CrossRef] [PubMed]

34. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16569–16572. [CrossRef] [PubMed]

35. Lü, L.; Zhang, Y.C.; Yeung, C.H.; Zhou, T. Leaders in social networks, the delicious case. *PLoS ONE* **2011**, *6*, e21202. [CrossRef] [PubMed]

36. Zhuang, Y.; Yağan, O. Information propagation in clustered multilayer networks. *IEEE Trans. Netw. Sci. Eng.* **2016**, *3*, 211–224. [CrossRef]

37. Basaras, P.; Iosifidis, G.; Katsaros, D.; Tassiulas, L. Identifying influential spreaders in complex multilayer networks: A centrality perspective. *IEEE Trans. Netw. Sci. Eng.* **2017**, *6*, 31–45. [CrossRef]

38. Cozzo, E.; Kivelä, M.; De Domenico, M.; Solé, A.; Arenas, A.; Gómez, S.; Porter, M.A.; Moreno, Y. Clustering coefficients in multiplex networks. *arXiv* **2013**, arXiv:1307.6780.

39. Rahmede, C.; Iacovacci, J.; Arenas, A.; Bianconi, G. Centralities of nodes and influences of layers in large multiplex networks. *J. Complex Netw.* **2018**, *6*, 733–752. [CrossRef]

40. Paidar, M.; Chaharborj, S.S.; Harounabadi, A. Identifying Top-k Most Influential Nodes by using the Topological Diffusion Models in the Complex Networks. *Network* **2017**, *4*, 5. [CrossRef]

41. Ohara, K.; Saito, K.; Kimura, M.; Motoda, H. Resampling-based predictive simulation framework of stochastic diffusion model for identifying top-K influential nodes. *Int. J. Data Sci. Anal.* **2020**, *9*, 175–195. [CrossRef]

42. Tang, J.; Zhang, R.; Yao, Y.; Yang, F.; Zhao, Z.; Hu, R.; Yuan, Y. Identification of top-k influential nodes based on enhanced discrete particle swarm optimization for influence maximization. *Phys. A Stat. Mech. Appl.* **2019**, *513*, 477–496. [CrossRef]

43. Silber, M.D. *The Al Qaeda Factor: Plots against the West*; University of Pennsylvania Press: Philadelphia, PA, USA, 2011.

44. Hethcote, H.W. The mathematics of infectious diseases. *SIAM Rev.* **2000**, *42*, 599–653. [CrossRef]

45. Guan-Rong, C.; Xiao-Fan, W.; Xiang, L. *Introduction to Complex Networks: Models, Structures and Dynamics*; Higher Education Press: Beijing, China, 2012.

46. Krackhardt, D. Assessing the political landscape: Structure, cognition, and power in organizations. *Adm. Sci. Q.* **1990**, *35*, 342–369. [CrossRef]

47. Zachary, W.W. An information flow model for conflict and fission in small groups. *J. Anthropol. Res.* **1977**, *33*, 452–473. [CrossRef]

48. Lusseau, D.; Schneider, K.; Boisseau, O.J.; Haase, P.; Slooten, E.; Dawson, S.M. The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. *Behav. Ecol. Sociobiol.* **2003**, *54*, 396–405. [CrossRef]

49. Tsvetovat, M.; Kouznetsov, A. *Social Network Analysis for Startups: Finding Connections on the Social Web*; OŔeilly Media, Inc.: Sebastopol, CA, USA, 2011.

50. Knuth, D.E. *The Stanford GraphBase: A Platform for Combinatorial Algorithms*; SODA: Austin, TX, USA, 1993; Volume 93, pp. 41–43.

51. Shen-Orr, S.S.; Milo, R.; Mangan, S.; Alon, U. Network motifs in the transcriptional regulation network of Escherichia coli. *Nat. Genet.* **2002**, *31*, 64–68. [CrossRef] [PubMed]

52. Rossi, R.; Ahmed, N. The network data repository with interactive graph analytics and visualization. In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015.

53. Duch, J.; Arenas, A. Community detection in complex networks using extremal optimization. *Phys. Rev. E* **2005**, *72*, 027104. [CrossRef] [PubMed]

54. Cho, A.; Shin, J.; Hwang, S.; Kim, C.; Shim, H.; Kim, H.; Kim, H.; Lee, I. WormNet v3: A network-assisted hypothesis-generating server for Caenorhabditis elegans. *Nucleic Acids Res.* **2014**, *42*, W76–W82. [CrossRef]

55. Action, R. The Rise of the Medici. *Am. J. Sociol.* **1993**, *98*, 1259–1319.

56. Krackhardt, D. Cognitive social structures. *Soc. Netw.* **1987**, *9*, 109–134. [CrossRef]

57. Vickers, M.; Chan, S. *Representing Classroom Social Structure*; Victoria Institute of Secondary Education: Melbourne, Australia, 1981.

58. Kapferer, B. *Strategy and Transaction in an African Factory: African Workers and Indian Management in a Zambian Town*; Manchester University Press: Manchester, UK, 1972.

59. Lazega, E. *The Collegial Phenomenon: The Social Mechanisms of Cooperation Among Peers in a Corporate Law Partnership*; Oxford University Press on Demand: Oxford, UK, 2001.

60. Snijders, T.A.; Pattison, P.E.; Robins, G.L.; Handcock, M.S. New specifications for exponential random graph models. *Sociol. Methodol.* **2006**, *36*, 99–153. [CrossRef]

61. Stark, C.; Breitkreutz, B.J.; Reguly, T.; Boucher, L.; Breitkreutz, A.; Tyers, M. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **2006**, *34*, D535–D539. [CrossRef] [PubMed]

62. Magnani, M.; Micenkova, B.; Rossi, L. Combinatorial analysis of multiple networks. *arXiv* **2013**, arXiv:1303.4986.

63. De Domenico, M.; Solé-Ribalta, A.; Gómez, S.; Arenas, A. Navigability of interconnected networks under random failures. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8351–8356. [CrossRef] [PubMed]

64. Cardillo, A.; Gómez-Gardenes, J.; Zanin, M.; Romance, M.; Papo, D.; Del Pozo, F.; Boccaletti, S. Emergence of network features from multiplexity. *Sci. Rep.* **2013**, *3*, 1344. [CrossRef] [PubMed]

65. Kendall, M.G. A new measure of rank correlation. *Biometrika* **1938**, *30*, 81–93. [CrossRef]