

Article

Local-Forest Method for Superspreaders Identification in Online Social Networks

Yajing Hao^{1,2}, Shaoting Tang^{2,3,4,5,6,7,8}, Longzhao Liu^{2,3,5,6,*}, Hongwei Zheng^{5,9,*} and Xin Wang^{2,3,5,6}
and Zhiming Zheng^{2,3,4,5,6,7,8}

- ¹ School of Mathematical Sciences, Beihang University, Beijing 100191, China
 - ² Key Laboratory of Mathematics, Informatics and Behavioral Semantics (LMIB), Beihang University, Beijing 100191, China
 - ³ Institute of Artificial Intelligence, Beihang University, Beijing 100191, China
 - ⁴ State Key Laboratory of Software Development Environment (NLSDE), Beihang University, Beijing 100191, China
 - ⁵ Beijing Advanced Innovation Center for Future Blockchain and Privacy Computing, Beihang University, Beijing 100191, China
 - ⁶ PengCheng Laboratory, Shenzhen 518055, China
 - ⁷ Institute of Medical Artificial Intelligence, Binzhou Medical University, Yantai 264003, China
 - ⁸ School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China
 - ⁹ Beijing Academy of Blockchain and Edge Computing (BABEC), Beijing 100085, China
- * Correspondence: longzhao@buaa.edu.cn (L.L.); hwzheng@pku.edu.cn (H.Z.)

Abstract: Identifying the most influential spreaders in online social networks plays a prominent role in affecting information dissemination and public opinions. Researchers propose many effective identification methods, such as k-shell. However, these methods are usually validated by simulating propagation models, such as epidemic-like models, which rarely consider the Push-Republish mechanism with attenuation characteristic, the unique and widely-existing spreading mechanism in online social media. To address this issue, we first adopt the Push-Republish (PR) model as the underlying spreading process to check the performance of identification methods. Then, we find that the performance of classical identification methods significantly decreases in the PR model compared to epidemic-like models, especially when identifying the top 10% of superspreaders. Furthermore, inspired by the local tree-like structure caused by the PR model, we propose a new identification method, namely the Local-Forest (LF) method, and conduct extensive experiments in four real large networks to evaluate it. Results highlight that the Local-Forest method has the best performance in accurately identifying superspreaders compared with the classical methods.

Keywords: complex network; online social network; spreading; influential nodes



Citation: Hao, Y.; Tang, S.; Liu, L.; Zheng, H.; Wang, X.; Zheng, Z. Local-Forest Method for Superspreaders Identification in Online Social Networks. *Entropy* **2022**, *24*, 1279. <https://doi.org/10.3390/e24091279>

Academic Editor: Federico Vazquez

Received: 1 August 2022

Accepted: 8 September 2022

Published: 11 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The development of the Internet and Web 2.0 has spawned many online social platforms, usually modeled as online social networks [1–4], which have become the main underlying platforms of information dissemination [5] and public opinion formation [6]. It is widely believed that some nodes in social networks have a stronger spreading influence, which is related to their privileged locations [4,7]. These nodes with outstanding spreading ability can affect and even dominate the spreading processes. These nodes are called superspreaders [8], influencers [4], influential spreaders [9], or influential nodes [10]. Finding influential nodes has an important impact on many issues related to spreading processes, such as innovation diffusion [11], viral marketing [12], decision making [13], predicting popular scientific publications [14], and so on. Therefore, identifying influential nodes is of practical significance and has attracted extensive attention from scholars.

Nowadays, there have been many methods for superspreader identification, which have become an important field in network science. The mainly used methods are centrality-

based heuristic methods [4,15], namely centrality methods. Specifically, these methods use the ranking of nodes' centrality measure values to predict their spreading influence ranking. Here, the spreading influence of a node is generally quantified by the final propagation range it causes as the source of propagation. The classical measures include degree centrality (the number of the neighbors of a node) [16], betweenness centrality (the number of the shortest path through a node) [17], closeness centrality (reciprocal of the average shortest path distance to a node) [18], PageRank (the probability of accessing each node when the random walk process on the network reaches the steady state) [19], ks index generated by the k-shell decomposition (here we call it k-shell centrality) [9], and so on. Due to the great success of the k-shell method in epidemic-like models, such as the Susceptible–Infectious–Recovered (SIR) model, there are many follow-up studies, including improvement based on information entropy [20,21], improvement based on gravity equation [22,23], the hybrid method [24], improvement based on neighborhood information [25,26], improved decomposition method [27,28], and so on [29,30]. Moreover, there are other centrality methods proposed continuously. The node propagation entropy measure [31] considers the neighbors within two hops and the clustering coefficients of nodes. The semi-global triangular centrality method [32] uses the triangular patterns around nodes to identify influential spreaders. The generalized gravity centrality [33] combines the clustering coefficient and degree to evaluate nodes based on the gravity formula. The coupling-sensitive centrality [34] identifies the superspreaders on multiplex networks. Centrality measures can not only be used to identify individual spreaders but also can be applied to identify multiple spreaders, which is an important problem in social networks and is also known as the Influence Maximization (IM) problem [1,12,15,35–37]. For example, centrality methods combined with some selection strategies can be directly used to identify multiple influencers [9,15,38,39]. Other methods for IM problems include greedy algorithms [35,40], community-based algorithms [37,41], optimization-based meta-heuristic algorithms [1,42], and so on [43,44].

Despite the great progress of the identification methods, we find that the effectiveness of them is usually validated under the assumption that the spreading dynamic follows the epidemic-like models. However, this dynamical assumption has some deviations from spreading processes in online social media. The main difference is the Push-Republish mechanism with attenuation characteristic, which widely exists when information diffuses [3,45–48]. In short, the Push-Republish mechanism [3,45] is that the online users receive the messages passively and then may republish the messages depending on how attractive the message is. In addition, Ruan et al. [47] think that online users decide whether to spread the information mainly according to its content. This inspires them to propose an information contagion model where individuals make the final decisions to spread or not when being exposed to the information for the first time. In other words, unlike the passive disease infection process, people have the initiative to keep their personal tendency in the information diffusion. Furthermore, Kempe et al. [46] propose the decreasing cascade model, which sets that people's probability of spreading information decreases with repeated stimuli and reflects a 'marketing-saturated' phenomenon. Both the model of Ruan et al. and the model of Kempe et al. portray the special attenuation characteristic in online spreading processes. Note that the difference between dynamical models can significantly affect the effectiveness of identification methods [49]. For example, the previously well-concerned k-shell method, which performs well under the epidemic-like models, is not good enough in the rumor dynamics [50]. These naturally raise the following problems: What is the performance of classical identification methods under the above mechanisms? Are there new effective methods inspired by the characteristics of the propagations in online social media?

To solve these problems, we consider the Push-Republish mechanism and Ruan et al.'s model together to describe the online spreading processes. For the convenience of description, we name the process the Push-Republish (PR) model. We find that, compared to the case under the SIR model, the classical identification methods have a worse perfor-

mance when identifying superspreaders under the PR model. Furthermore, inspired by the features of the PR model, we propose a new centrality measure called the Local-Forest (LF) measure. To verify its effectiveness, we apply the proposed LF method to four real online social networks and compare it with the classical methods. Results show that the LF method performs better.

The outline of the paper is as follows. Section 2 introduces the PR model and shows the essential difference between the PR model and the SIR model. Section 3 gives a detailed description of the proposed LF measure. In Section 4, extensive experiments are conducted to evaluate the performance of the LF method and some classic methods. Finally, Section 5 summarizes the main results of this paper and prospects for future research.

2. Push-Republish Model

2.1. Description of the Push-Republish Model

The information diffusion procedure in online social media can be described as the following steps [3]: A user releases a message. Then friends of the user will learn of this message under the notification of the online social site, and this mechanism is called “push”. Affected by the interesting degree of the message, some friends will comment, cite or reprint the message. This behavior is called “republish”. The above steps will repeat recursively for each user who chooses to republish the message. In addition, some dynamical models, such as the decreasing cascade model [46] and Ruan et al.’s model [47], consider the attenuation characteristic to describe the influence propagation in social networks. In the decreasing cascade model, the probability that a node becomes activated decreases as the times of being stimulated increase. It indicates that the first contact with information is the most promising to lead to dissemination. Ruan et al. [47] divide nodes into three states: susceptible, adopted, or immune. In their model, a susceptible node converts its state to the adopted state or immune state immediately after receiving the information and then keeps its state. In other words, each node can be activated only when it is exposed to the information the first time.

The Push-Republish mechanism can directly describe the propagation mechanism in online social media and is compatible with Ruan et al.’s model, which considers the attenuation characteristic. Thus, we take the Push-Republish mechanism and Ruan et al.’s model into consideration and redescribe the above dynamical processes. For convenience, we name it the Push-Republish model. Online media can naturally be represented by networks where nodes represent users’ accounts, and edges represent the relationship between them. We divide all nodes into two states: S (do not know the message) and R (know the message). Nodes in set I_t represent those who publish the message at the time t , i.e., those who can spread the message. Set the initial time t as 0. Choose a single node s as the origin of the propagation, and obtain $I_0 = \{s\}$. Set the propagation probability as β . The PR model is as follows:

1. Step 1: At the time t , for node i in the set I_t , its neighbors in S state will become R state (i.e., the “push” mechanism). Meanwhile, such neighbors who receive the message for the first time will choose to republish the message with probability β (i.e., the “republish” behavior). Add the neighbors who choose to republish the message to the set I_{t+1} .
2. Step 2: Remove i from I_t .
3. Step 3: Perform steps 1–2 until $I_t = \emptyset$, then perform step 4.
4. Step 4: Update the propagation process to the next time step, i.e., set $t = t + 1$. If there is no node in set I_t , the propagation process ends. Denote the ending time as the propagation duration T . Otherwise, repeat step 3.

If edges in the network are directed (suppose that the user i follows the user j , then there is an edge from node i to node j in the network), the neighbors mentioned in the above PR model refer to the in-neighbors of nodes. We record the number of nodes with the state R when propagation ends to represent the propagation range. It can be seen that nodes with the state R in the PR model are the combination of the adopted nodes and

the immune nodes in Ruan et al.'s model. The PR model has two main characteristics. First, the PR model introduces the push mechanism, which is consistent with the message transmission mechanism in online social media. Second, if the node that did not republish the message meets the message again, it will keep its decision, i.e., different from the passivity of individuals towards virus infection, individuals have the initiative to choose whether to spread the information or not, reflecting the attenuation characteristic.

2.2. Comparison between PR Model and SIR Model

We further compare the PR model and the SIR model (see details in Appendix A) from two aspects: (1) How does information propagate under these two models? (2) How do the classical degree and k-shell methods (see details in Appendix B) perform under these two models? The network used in this section is the largest connected component of the Facebook network [51], which is an undirected network with 63,392 nodes.

We use the distance between the source node and the nodes that republish the message to compare the characteristic of the two specific propagation processes. Specifically, we select a fixed node chosen from the inner-most shell (i.e., nodes with the highest ks index) of the network as the propagation origin and use the PR model and the SIR model as the propagation model to conduct 10,000 simulations, respectively. To facilitate the comparison between the PR model and the SIR model, in the PR model, we say nodes in set I_t are in state I at time t (thus state I coexists with state R). For each time t , we calculate the average distance and minimum distance from the nodes in the state I to the source node. Figure 1 presents the time evolution of the average distance and minimum distance of the PR model (see Figure 1a) and the SIR model (see Figure 1b). We find that in the PR model, the minimum distance is not less than 2 when $t > 1$, while this index is equal to 1 for a long time in the SIR model. This phenomenon can be directly explained by the dynamic mechanisms. In the PR model, neighbors obtain the message from the source node, and then if they decide not to republish the message, they will keep their tendency. Thus, the edges between the neighbors do not participate in the propagation process, and the potential propagation paths in the neighborhood of the original node form a local tree-like structure. As a result, the minimum distance is not less than 2 when $t > 1$. However, in the SIR model, messages can be delivered among neighbors of the original node, and the edges between neighbors lead to repetitive stimuli, which promote the propagation. In reality, the inner k-shell is a cluster of nodes that are closely connected with each other. Kitsak et al. [9] have pointed out that one of the reasons why the k-shell method performed well in the SIR model was that the neighborhood of nodes in the high k-shell enabled them to effectively sustain propagation in the early stage of the propagation process, and then the epidemic could fully develop. Both the analysis of Kitsak et al. [9] and our experimental results here indicate the positive effect of closely connected neighbors on the propagation processes of the SIR model and explain the superiority of the k-shell method. However, due to the connections between neighbors becoming redundant in the PR model, the recognition ability of the k-shell method in it might not be as great as the case in the SIR model. The next experiment will confirm this claim.

Next, we will compare the performance of the classic degree method and k-shell method in the PR model and the SIR model. By convention, we carry out 1000 times Monte Carlo propagation simulation experiments in the PR model and the SIR model for each node in the Facebook network to estimate the real spreading influence of nodes in the two models, respectively [9,52]. Specifically, we use the average number of the R -state nodes when the propagation ends over all corresponding simulations to represent the source node's real spreading influence. Then, sort all nodes in descending order according to their spreading influence. In this paper, we use RK_{PR} and RK_{SIR} to represent the sorting of nodes under the PR model and under the SIR model, respectively.

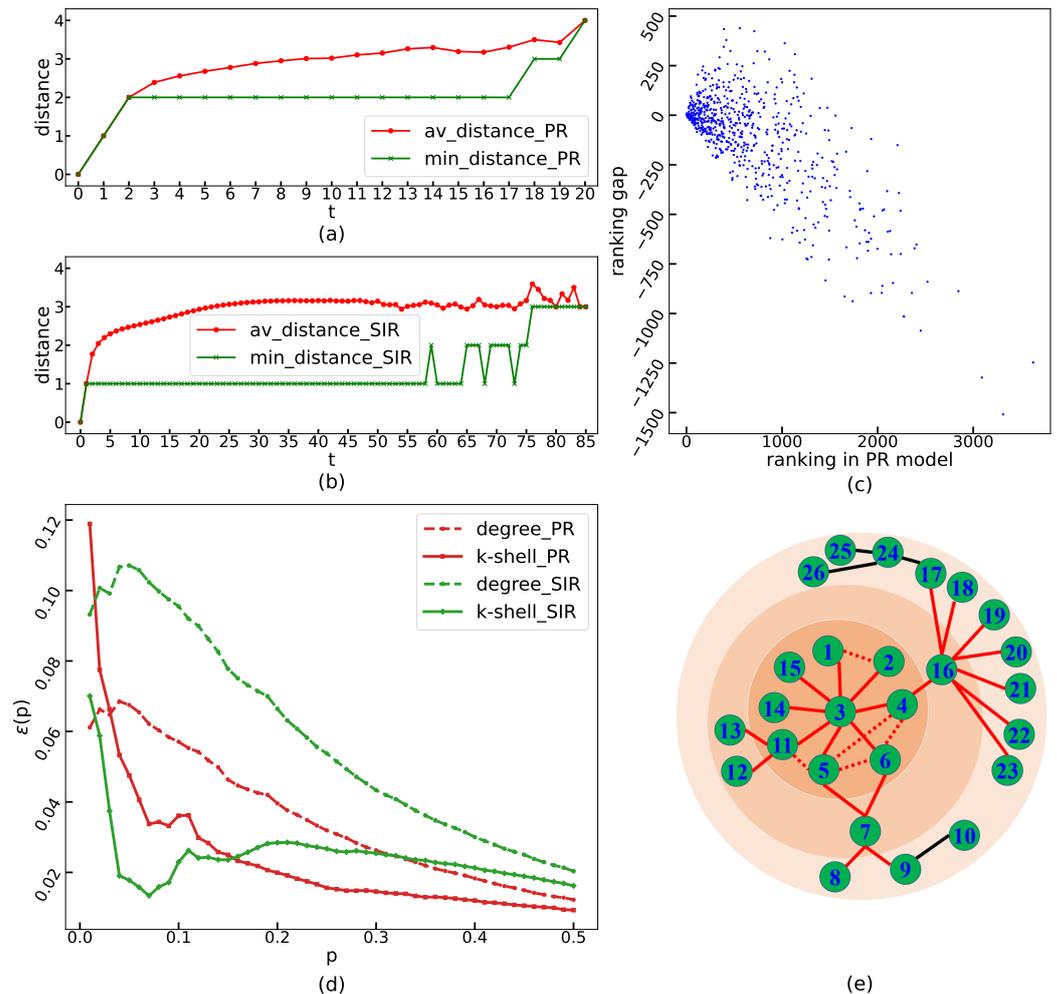


Figure 1. (a,b) The average distance and minimum distance from *I*-state nodes to the source node. (c) The ranking gap of nodes’ spreading influence between the case in the PR model and that in the SIR model. (d) The imprecision function of the degree method and the k-shell method under the PR model and the SIR model, respectively. We set $\beta = 0.012$ in (a–d). (e) Schematic diagram of the Local-Tree measure and the Local-Forest measure.

First, for each node in the inner-most shell (i.e., nodes with the largest k_s index), we compare its ranking of the spreading influence in the PR model and that in the SIR model, as shown in Figure 1c. The horizon axis means the ranking of the node in the ranking list RK_{PR} . The vertical axis represents the ranking gap, which is calculated by the ranking of the node in RK_{SIR} minus the ranking of the node in RK_{PR} . It can be seen that rankings of these nodes in the SIR model are generally higher (i.e., possess smaller ranking numbers) than that in the PR model, which means that the k-shell method performs not so well in the PR model.

Furthermore, we use the imprecision function [9,20,52,53] to measure the difference in the spreading influence between the top $[pN]$ nodes ($p \in [0, 1]$) predicted by the recognition method and the top $[pN]$ most influential nodes given by RK_{PR} or RK_{SIR} . Here p denotes the ratio of nodes, and N denotes the number of nodes in the network. The imprecision function [9,52] $\epsilon(p)$ for a given recognition method is defined as the following Equation (1):

$$\epsilon(p) = 1 - \frac{M(p)}{Meff(p)}, \tag{1}$$

where $M(p)$ is the average spreading influence of the top $[pN]$ nodes in the ranking list RK_M (represents the nodes’ ranking list identified by centrality methods), and $Meff(p)$ is

the average spreading influence of the top $[pN]$ nodes in sorting RK_{PR} or RK_{SIR} . For each value of p , the smaller the $\varepsilon(p)$, the better the recognition effect of the method. Figure 1d gives the imprecision function of the degree method and the k-shell method under the PR model and the SIR model, respectively. It can be seen that the recognition ability of the k-shell method (solid curves) in the SIR model (green curves) is significantly better than that of the degree method (dashed curves). However, in the PR model (red curves), the k-shell method is obviously inferior to the degree method in the small p region. Furthermore, compared to the case in the SIR model, the performance of the k-shell method decreases in the PR model, especially when identifying the top 10%.

In summary, the unique Push-Republish mechanism with attenuation feature in online social networks delivers different evolutionary results from those caused by the classical epidemic-like models. In addition, the performance of the traditional effective k-shell method decreases in the PR model. Therefore, a more effective method is necessary.

3. Local-Forest Method

As the simulations in the above section show, in the PR model, messages will not propagate between neighbors of the source node. To some extent, this leads to a locally tree-like structure, which is consistent with the reality that the majority of cascades on online social media are small and are described by a handful of simple tree structures [2]. Specifically, in the PR model, each node only has one choice to decide whether to republish the message or not and keeps the decision. Such a characteristic indicates that edges between neighbors of the source node are redundant. Inspired by this thought, we propose the Local-Tree (LT) measure and the Local-Forest (LF) measure. The LT measure is an intermediate measure to calculate the LF measure, whose basic idea is to remove the edges in the neighborhood of nodes that are redundant for propagation. For undirected networks, the LT centrality of a node is the sum of the degree of its neighbors minus twice the number of edges between neighbors. The LF centrality of a node is the sum of the LT values of its neighbors. The LF centrality $LF(i)$ of node i is calculated as:

$$LT(j) = \sum_{w \in N(j)} k(w) - 2m(j), \quad (2)$$

$$LF(i) = \sum_{j \in N(i)} LT(j), \quad (3)$$

where $N(j)$ is the set of neighbors of node j , $k(w)$ is the degree centrality of node w , and $m(j)$ is the number of edges between neighbors of node j . For directed networks, only the in-neighbors of nodes are considered, and the in-degree of nodes is used to replace the degree. Moreover, subtract the number of edges itself but not double. This is because one edge in an undirected network is equivalent to two opposite edges in a directed network. The summation of neighbors' LT measure in the LF measure highlights the importance of the node's degree and makes use of more information. Calculating $m(j)$ requires two levels of traversal of j 's neighbors, so the time complexity of calculating the LF centrality for all nodes in the network is $O(N\langle k \rangle^2)$, where N is the number of the nodes in the network and $\langle k \rangle$ is the average degree of all of the nodes in the network. Average degree $\langle k \rangle$ of sparse networks is much smaller than N . Thus in large-scale sparse networks, the time complexity of the LF measure is similar to the degree.

We consider a small network [9] as an example to illustrate the calculation of the Local-Tree measure and Local-Forest method (see Figure 1e). We take node 3 in Figure 1e as an example to calculate the LT centrality and the LF centrality. First calculate $LT(3) = \sum_{w \in N(3)} k(w) - 2m(3) = 23 - 2 \times 5 = 13$, where $N(3) = \{1, 2, 4, 5, 6, 11, 14, 15\}$. It can be found that when calculating the LT centrality of node 3, the information in the inner-most shadow area in Figure 1e is used. Subtracting the item $2m(3)$ is equivalent to deleting the dotted line edge. Then a local tree structure is formed in the first-order neighborhood of the node, which is why we call this measure Local-Tree centrality. Next, calculate

$LF(3) = \sum_{j \in N(3)} LT(j) = LT(1) + LT(2) + LT(4) + LT(5) + LT(6) + LT(11) + LT(14) + LT(15) = 8 + 8 + 19 + 14 + 13 + 13 + 8 + 8 = 91$. The red edges in Figure 1e contribute to the calculation of the LF centrality of node 3. It can be found that the LF measure contains the information of the node’s second-order neighbors.

Taking each node in the network shown in Figure 1e as the origin of propagation, we conduct one hundred thousand propagation simulations for each case based on the PR model. The first row in Table 1 arranges these nodes in descending order according to their real spreading influence and merges the nodes with equivalent topological positions in the network into the same column. Table 1 gives the degree centrality, k-shell centrality, LT centrality, and LF centrality of each node. The LF centralities of nodes in adjacent columns are significantly different, indicating that the LF measure can better discriminate the nodes. On the whole, the LF centralities of nodes decrease with the real spreading influence decreasing, which indicates the excellent performance of the LF method.

Table 1. The effectiveness of different identification methods in the sample of Figure 1e.

| Centrality Measure | 3 | 16 | 4 | 5 | 6 | 11 | 7 | 17 | 1,2 | 24 | 9 | 14,15 | 18–23 | 12,13 | 8 | 25,26 | 10 |
|--------------------|----|----|----|----|----|----|----|----|-----|----|----|-------|-------|-------|----|-------|----|
| k-shell | 3 | 1 | 3 | 3 | 3 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| degree | 8 | 8 | 4 | 5 | 4 | 4 | 4 | 2 | 2 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| LT | 13 | 12 | 19 | 14 | 13 | 13 | 10 | 11 | 8 | 4 | 5 | 8 | 8 | 4 | 4 | 3 | 2 |
| LF | 91 | 78 | 52 | 68 | 56 | 35 | 36 | 16 | 21 | 17 | 12 | 13 | 12 | 13 | 10 | 4 | 5 |

4. Simulation Results

4.1. Experimental Setup

We conduct experiments on four large-scale real online social networks to compare the LF method with five widely-concerned centrality methods, including the degree [16], k-shell [9], closeness centrality [18], PageRank [19], and Mixed Degree Decomposition (MDD) method [27]. We give a brief introduction to these methods in Appendix B. We investigate the friendship networks gathered from (1) Brightkite [54,55], (2) Facebook [51,54], (3) DouBan [54,56], and (4) Twitter [54,57]. Twitter is a directed network, and other networks are undirected networks. Each network used in this paper is obtained by extracting the largest connected component from the original data set and the largest weakly connected component for the directed network. To precisely estimate the real spreading influence of each node, we simulate the propagation process based on the PR model via independent Monte Carlo experiments. Take the average of the propagation ranges (i.e., the number of nodes in the state R when propagation ends) of the Monte Carlo experiments as the real spreading influence of the source node. Still use RK_{PR} to represent the ranking (in descending order) of nodes based on experiments, and use RK_M to represent the ranking list based on identification methods.

Table 2 lists some basic characteristics of the networks and the parameters in the experiments. N is the number of nodes in the network, M is the number of edges in the network, and $\langle k \rangle$ is the average degree of nodes in the network. MC is the number of Monte Carlo experiments for each node. Related works in recent years usually set MC as 100 [32–34,53] or 1000 [23,31,39,52]. A relatively small MC (i.e., 100) is acceptable for large networks (i.e., with $N > 10,000$) [32]. Here, we set larger MC parameters for each network as far as possible within the range of the computational power. β is the propagation probability in the PR model. Note that the position of the source node makes no difference in the final propagation size under a large value of β . Thus, complying with previous studies, we also select the value of β around its threshold. Considering that the PR model has no corresponding threshold, here we select a β slightly larger than the threshold of the SIR model (β_{th}). Here, the threshold of the SIR model is calculated as $\beta_{th} = \langle k \rangle / \langle k^2 \rangle$ [26], where $\langle k \rangle$ is the average degree, and $\langle k^2 \rangle$ is the second-order average degree. Under this parameter setting, the proportion of propagations with duration 1 (i.e., with $T = 1$) in all propagation simulations in each network ($P_{T=1}$ in Table 2) is 78–92%.

This is in line with the characteristic of propagation flow on social networks described by Goel et al. [2], i.e., across seven online domains they investigated, in which 73% to 95% of instances show no diffusion. Here the propagation duration is the time when the propagation ends. For example, when there no neighbor of the source node republishes the message, the propagation duration is recorded as 1.

Table 2. Networks attributes, experimental settings, and some statistics for simulation results.

| Network | N | M | $\langle k \rangle$ | MC | β_{th} | β | $P_{T=1}$ |
|------------|---------|---------|---------------------|------|--------------|---------|-----------|
| Brightkite | 56,739 | 425,890 | 7.506 | 1000 | 0.015618 | 0.016 | 90.9% |
| Facebook | 63,392 | 816,831 | 25.771 | 1000 | 0.011358 | 0.012 | 78.9% |
| DouBan | 154,908 | 654,324 | 4.224 | 500 | 0.027100 | 0.028 | 91.8% |
| Twitter | 465,017 | 834,797 | 1.795 | 200 | 0.116572 | 0.120 | 82.6% |

4.2. Methods Evaluation

Based on the experiments introduced above, we use three indicators [8,9,27,52] to analyze and compare the recognition ability of each method, including the imprecision function, the recognition rate, and the Kendall coefficient.

4.2.1. Imprecision Function

The definition of the imprecision function $\varepsilon(p)$ is shown as Equation (1) in Section 2. For any value of p , the smaller the value of the imprecision function, the better the recognition effect of the method. Figure 2 presents the imprecision function of each method, with p ranging from 0.01 to 0.5, and each subplot corresponds to a real online network. We can observe that our LF method, represented by the red curve in all subplots, has the smallest value of the imprecision function, which indicates the great performance of our method. In particular, when identifying the top 10% of superspreaders, our LF method significantly improves the efficiency of identification compared to other classical methods. The closeness method has a good performance, which is near the LF method in the DouBan network and the Twitter network, but the time complexity of the closeness measure is higher [15,18]. In the other two networks, our LF method has a significant advantage in the efficiency of identification. We conducted the experiments on Ubuntu Linux 18.04 LTS equipped with Intel(R) Xeon(R) CPU E5-2690 v3 @ 2.60GHz, and the algorithms were implemented in Python 3.8.5. The average consumed time of the 10 runs for each method on the four large networks is shown in Table 3. The results show that the time complexity of the LF method is the second-smallest among these methods (only larger than that of the degree method). Considering both the performance and time complexity, our LF method is the best choice among these methods.

Table 3. The consumed time of the LF method and other classical methods in the four real networks.

| Network | Degree | k-Shell | Closeness | PageRank | MDD | LF |
|------------|---------|---------|--------------|----------|-----------|---------|
| Brightkite | 0.04799 | 1.73308 | 11,845.25111 | 14.45678 | 28.02683 | 0.95804 |
| Facebook | 0.03003 | 2.15947 | 24,302.80110 | 29.29709 | 44.02255 | 1.19346 |
| DouBan | 0.12007 | 1.50177 | 59,277.99047 | 27.76005 | 50.87384 | 1.72524 |
| Twitter | 0.24067 | 4.83157 | 1813.34151 | 14.19303 | 192.96852 | 2.10739 |

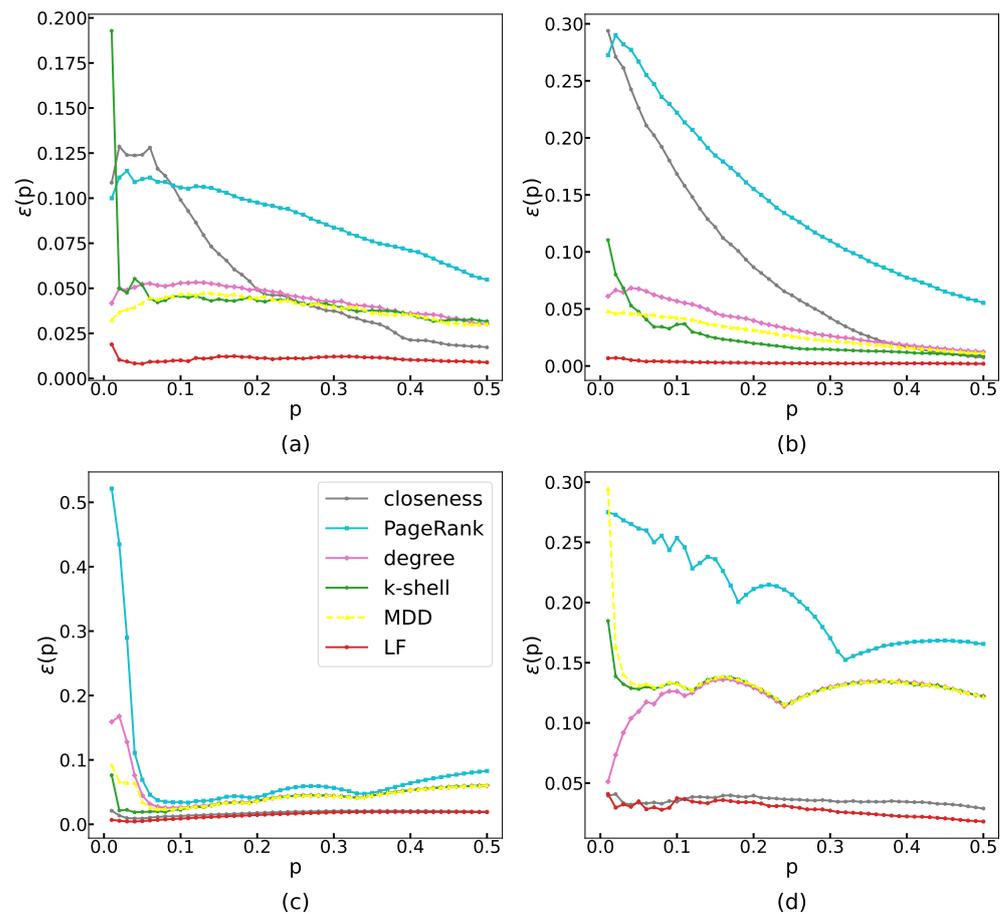


Figure 2. The imprecision function of the LF method and other classical methods under different p in four real networks, including (a) Brightkite, (b) Facebook, (c) DouBan, and (d) Twitter. Curves with different colors correspond to different methods: closeness (grey), PageRank (cyan), degree (pink), k-shell (green), MDD (yellow), and LF (red).

4.2.2. Recognition Rate

The recognition rate reflects the ability of methods to find important nodes. Here we use Pei et al.'s [8] definition of the recognition rate $r(p)$, which reads as:

$$r(p) = \frac{|I_p \cap P_p|}{|I_p|}, \quad (4)$$

where I_p is the set of nodes ranking in the top p fraction of the RK_{PR} , P_p is the set of nodes ranking in the top p fraction of the RK_M , and $|I_p|$ is the number of nodes in I_p . For any value of p , the greater the recognition rate $r(p)$, the higher the accuracy of the method for influential node recognition. Figure 3 shows the curves of the recognition rate of different methods in the four real networks, with p ranging from 0.01 to 0.5. It can be seen that the red curve corresponding to the LF method is higher than other curves, which indicates that the LF method has better recognition accuracy in terms of identifying specific nodes than other methods.

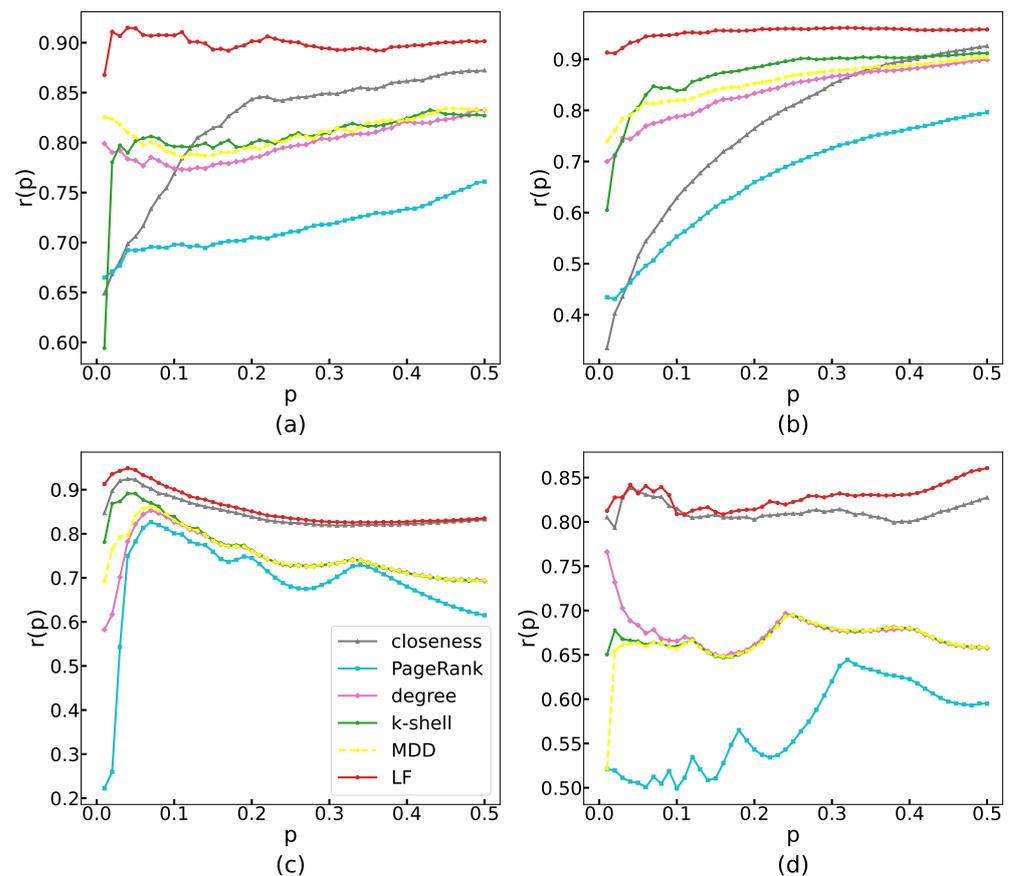


Figure 3. The recognition rate of the LF method and other classical methods under different p in four real networks, including (a) Brightkite, (b) Facebook, (c) DouBan, and (d) Twitter. Curves with different colors correspond to different methods: closeness (grey), PageRank (cyan), degree (pink), k-shell (green), MDD (yellow), and LF (red).

4.2.3. Kendall Correlation Coefficient

Kendall’s *tau* correlation coefficient [58] is a statistic that measures the rank correlation of two random variables. We use it to quantify the rank correlation between the measure and the real spreading influence obtained from Monte Carlo experiments. There are three calculation forms of Kendall’s *tau* coefficient. Considering that the number of values of the spreading influence (i.e., an average of propagation ranges) is more than that of the centrality measures, we use τ_c [59,60] in our paper. For each node i , record its centrality x_i and its real spreading influence y_i as (x_i, y_i) . For any pair of (x_i, y_i) and (x_j, y_j) , they are said to be concordant if $sgn((x_i - x_j) \times (y_i - y_j)) = 1$, discordant if $sgn((x_i - x_j) \times (y_i - y_j)) = -1$. If $sgn((x_i - x_j) \times (y_i - y_j)) = 0$, they are neither concordant nor discordant. τ_c is defined as:

$$\tau_c = \frac{2(n_c - n_d)}{n^2 \frac{m-1}{m}}, \tag{5}$$

where n_c is the number of concordant pairs, n_d is the number of discordant pairs, r and c are numbers of the values of the x and y respectively, and $m = \min(r, c)$. Clearly, the higher the τ_c , the higher the ranking correlation between the two variables. Briefly speaking, the above imprecision function evaluates the method in terms of the real spreading range, the recognition rate evaluates methods from the aspect of the specific nodes, and here, the τ_c gives a concise quantification of the ranking correlation between the measures and the real spreading influence of the nodes from a global aspect. Table 4 gives Kendall’s τ_c correlation coefficient between each measure and the real spreading influence. Apparently,

the LF measure we proposed has the highest τ_c in each network considered, which means that it has the highest correlation with the real spreading influence of nodes.

Table 4. τ_c between centrality measures and real spreading influence in four real networks. The best value is in bold (higher is better).

| Network | Degree | k-Shell | Closeness | PageRank | MDD | LF |
|------------|---------|---------|-----------|----------|---------|----------------|
| Brightkite | 0.42768 | 0.45701 | 0.68762 | 0.22887 | 0.42658 | 0.69770 |
| Facebook | 0.79352 | 0.83248 | 0.80367 | 0.55818 | 0.80476 | 0.90089 |
| DouBan | 0.65134 | 0.65002 | 0.72248 | 0.45388 | 0.65896 | 0.78665 |
| Twitter | 0.34802 | 0.35625 | 0.63684 | 0.17299 | 0.34434 | 0.70861 |

5. Discussion

Identifying superspreaders is of great significance in dealing with many real problems, especially related to the spreading processes [61,62], such as promoting innovation diffusion and controlling rumor spreading. Nowadays, a large number of identification methods have been proven to be effective on epidemic-like propagation models. However, these propagation models rarely consider the Push-Republish dynamics with attenuation characteristics, which exist widely in online spreading processes. This naturally leads to the question of whether classical identification methods perform well when considering the new dynamics. In addition, proposing new effective identification methods is always a meaningful and challenging problem.

To address these issues, we use the Push-Republish model to describe the message propagation process in online social media based on the work of Zhao et al. [3] and Ruan et al. [47]. In this model, the message is pushed to users passively, and users keep their spreading decision. Through model simulations, we find that a local tree-like spreading structure in the neighborhood of the source node occurs and that the vast majority of information diffusion will end in one step, which is consistent with the previous discovery [2]. These phenomena indicate that the edges between neighbors of the source node may be redundant, i.e., make no difference in diffusion. Then, we test the performance of some classical identification methods by simulating our PR model. Results show the poorer performance of the widely concerned k-shell method when identifying the top 10% of superspreaders. Considering the redundant edges and tree-like structure, we define the Local-Tree measure and propose the Local-Forest method to identify superspreaders in online social media. We apply our LF method and other classical identification methods to four real large networks. Through comparison, we highlight that the LF method performs the best in all networks from the aspect of all indicators, including imprecision function, recognition rate, and Kendall's correlation.

The LF method is not limited to the PR model because it is proposed based on the consideration of the redundant edges and tree-like structure, which are common features in various domains such as spreading processes. It indicates that the LF method has the potential to accurately identify superspreaders in other domains with the above features. In the future, we will explore the extension of our method in these fields. Moreover, the combination of the LF measure and the research about multi-source node identification is worth studying.

Author Contributions: Conceptualization, S.T. and Y.H.; methodology, Y.H.; software, Y.H. and X.W.; writing—original draft preparation, Y.H.; writing—review and editing, L.L., S.T., and Y.H.; supervision, S.T. and Z.Z.; project administration, H.Z. and S.T.; funding acquisition, Z.Z., H.Z., and S.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Program of National Natural Science Foundation of China Grant No. 11922102, 11871004, 62141605 and National Key Research and Development Program of China Grant No. 2018AAA0101100, 2021YFB2700304.

Institutional Review Board Statement: Not applicable

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. SIR Model

There are three states in the SIR model, including S (susceptible), I (infectious), and R (recovered). Initially, only the source node is in the state I , and all the other nodes are in the susceptible state (S). At each time step, nodes in the state I infect their neighbors in the state S with probability β , and they enter the R state at the next time step. Nodes with state R cannot be infected again.

Appendix B. Classic Centrality Measures

1. Degree.

The degree (denoted as k) of a node is the number of nodes connected to it. For a directed network, the in-degree (denoted as k_{in}) of a node is the number of nodes pointing to it (such nodes are called in-neighbor), and the out-degree (denoted as k_{out}) is the number of nodes pointed by it.

2. K-shell centrality.

In k-shell decomposition, firstly, remove all nodes with degree $k = 1$ (set $k = k_{in} + k_{out}$ for directed networks). Then, there may appear nodes with $k = 1$ again. Continue to remove all nodes with degree $k = 1$ iteratively until there is no node with $k = 1$. The ks index of the removed nodes is 1, and here we define the k-shell centrality of these nodes as $ks = 1$. We say the nodes with $ks = 1$ are in the 1-shell. Iteratively, remove the nodes with $k = 2$ in a similar way and obtain the 2-shell. Continue the removing process to obtain higher- k shells until there is no more node in the network. Finally, the k-shell centrality of each node is clarified, and the network can be viewed as the union of all shells. Nodes in the inner-most shell possess the highest k-shell centrality (i.e., highest ks index).

3. Closeness centrality.

The closeness centrality of node i is defined as

$$c(i) = \frac{1}{\sum_{j \in V} d_{ij}}, \quad (\text{A1})$$

where V is the set of all the nodes in the network and d_{ij} is the distance between node i and node j .

4. PageRank.

The PageRank value of node i is calculated iteratively. The PageRank value of node i in step t is calculated as:

$$PR_t(i) = \alpha \sum_{j \in N(i)} \frac{PR_{t-1}(j)}{k_{out}(j)} + \frac{1 - \alpha}{N}, \quad (\text{A2})$$

where $N(i)$ is the set of the neighbors of node i , $k_{out}(j)$ is the out-degree of node j , α is a constant, and N is the number of the nodes in the network. Conventionally, set $PR_0(i) = 1$ for each node i , and $\alpha = 0.85$. Finally, the $PR_t(i)$ converges to $PR(i)$, which is the PageRank centrality of node i .

5. Mixed degree decomposition (MDD).

Use the mixed degree $k^{(m)} = k^{(r)} + \lambda * k_i^{(e)}$ to replace the degree k to decompose the network in the same way as the k-shell decomposition, where $k^{(r)}$ is the residual degree (number of links connecting to the remaining nodes) and $k^{(e)}$ is the exhausted degree (number of links connecting to the removed nodes). λ is a constant, which is set as 0.7 commonly.

References

1. Wang, L.; Ma, L.; Wang, C.; Xie, N.g.; Koh, J.M.; Cheong, K.H. Identifying influential spreaders in social networks through discrete moth-flame optimization. *IEEE Trans. Evol. Comput.* **2021**, *25*, 1091–1102. [\[CrossRef\]](#)
2. Goel, S.; Watts, D.J.; Goldstein, D.G. The structure of online diffusion networks. In Proceedings of the 13th ACM Conference on Electronic Commerce, Valencia, Spain, 4–8 June 2012; pp. 623–638.
3. Zhao, J.; Wu, J.; Xu, K. Weak ties: Subtle role of information diffusion in online social networks. *Phys. Rev. E* **2010**, *82*, 016105. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Pei, S.; Morone, F.; Makse, H.A. Theories for influencer identification in complex networks. In *Complex Spreading Phenomena in Social Systems*; Springer: Cham, Switzerland, 2018; pp. 125–148.
5. Liu, L.; Wang, X.; Zheng, Y.; Fang, W.; Tang, S.; Zheng, Z. Homogeneity trend on social networks changes evolutionary advantage in competitive information diffusion. *New J. Phys.* **2020**, *22*, 013019. [\[CrossRef\]](#)
6. Wang, X.; Sirianni, A.D.; Tang, S.; Zheng, Z.; Fu, F. Public discourse and social network echo chambers driven by socio-cognitive biases. *Phys. Rev. X* **2020**, *10*, 041042. [\[CrossRef\]](#)
7. Pei, S.; Makse, H.A. Spreading dynamics in complex networks. *J. Stat. Mech. Theory Exp.* **2013**, *2013*, P12002. [\[CrossRef\]](#)
8. 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. [\[CrossRef\]](#)
9. 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*, 888–893. [\[CrossRef\]](#)
10. Chen, D.; Lü, L.; Shang, M.S.; Zhang, Y.C.; Zhou, T. Identifying influential nodes in complex networks. *Phys. A Stat. Mech. Its Appl.* **2012**, *391*, 1777–1787. [\[CrossRef\]](#)
11. Abrahamson, E.; Rosenkopf, L. Social network effects on the extent of innovation diffusion: A computer simulation. *Organ. Sci.* **1997**, *8*, 289–309. [\[CrossRef\]](#)
12. Richardson, M.; Domingos, P. Mining knowledge-sharing sites for viral marketing. In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, AB, Canada, 23–26 July 2002; pp. 61–70.
13. Muchnik, L.; Aral, S.; Taylor, S.J. Social influence bias: A randomized experiment. *Science* **2013**, *341*, 647–651. [\[CrossRef\]](#)
14. Lü, L.; Chen, D.; Ren, X.L.; Zhang, Q.M.; Zhang, Y.C.; Zhou, T. Vital nodes identification in complex networks. *Phys. Rep.* **2016**, *650*, 1–63. [\[CrossRef\]](#)
15. Erkol, Ş.; Castellano, C.; Radicchi, F. Systematic comparison between methods for the detection of influential spreaders in complex networks. *Sci. Rep.* **2019**, *9*, 15095. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Albert, R.; Jeong, H.; Barabási, A.L. Error and attack tolerance of complex networks. *Nature* **2000**, *406*, 378–382. [\[CrossRef\]](#)
17. Freeman, L.C. A set of measures of centrality based on betweenness. *Sociometry* **1977**, *40*, 35–41. [\[CrossRef\]](#)
18. Sabidussi, G. The centrality index of a graph. *Psychometrika* **1966**, *31*, 581–603. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Brin, S.; Page, L. The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. ISDN Syst.* **1998**, *30*, 107–117. [\[CrossRef\]](#)
20. Liu, Y.; Tang, M.; Zhou, T.; Do, Y. Core-like groups result in invalidation of identifying super-spreader by k-shell decomposition. *Sci. Rep.* **2015**, *5*, 9602. [\[CrossRef\]](#)
21. Wang, M.; Li, W.; Guo, Y.; Peng, X.; Li, Y. Identifying influential spreaders in complex networks based on improved k-shell method. *Phys. A Stat. Mech. Its Appl.* **2020**, *554*, 124229. [\[CrossRef\]](#)
22. 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. Its Appl.* **2016**, *451*, 205–212. [\[CrossRef\]](#)
23. Li, Z.; Huang, X. Identifying influential spreaders in complex networks by an improved gravity model. *Sci. Rep.* **2021**, *11*, 22194. [\[CrossRef\]](#)
24. Hou, B.; Yao, Y.; Liao, D. Identifying all-around nodes for spreading dynamics in complex networks. *Phys. A Stat. Mech. Its Appl.* **2012**, *391*, 4012–4017. [\[CrossRef\]](#)
25. Lin, J.H.; Guo, Q.; Dong, W.Z.; Tang, L.Y.; Liu, J.G. Identifying the node spreading influence with largest k-core values. *Phys. Lett. A* **2014**, *378*, 3279–3284. [\[CrossRef\]](#)
26. Bae, J.; Kim, S. Identifying and ranking influential spreaders in complex networks by neighborhood coreness. *Phys. A Stat. Mech. Its Appl.* **2014**, *395*, 549–559. [\[CrossRef\]](#)
27. Zeng, A.; Zhang, C.J. Ranking spreaders by decomposing complex networks. *Phys. Lett. A* **2013**, *377*, 1031–1035. [\[CrossRef\]](#)
28. Wei, B.; Liu, J.; Wei, D.; Gao, C.; Deng, Y. Weighted k-shell decomposition for complex networks based on potential edge weights. *Phys. A Stat. Mech. Its Appl.* **2015**, *420*, 277–283. [\[CrossRef\]](#)
29. Wang, Z.; Zhao, Y.; Xi, J.; Du, C. Fast ranking influential nodes in complex networks using a k-shell iteration factor. *Phys. A Stat. Mech. Its Appl.* **2016**, *461*, 171–181. [\[CrossRef\]](#)
30. Fu, Y.H.; Huang, C.Y.; Sun, C.T. Identifying super-spreader nodes in complex networks. *Math. Probl. Eng.* **2015**, *2015*, 675713. [\[CrossRef\]](#)
31. Yu, Y.; Zhou, B.; Chen, L.; Gao, T.; Liu, J. Identifying Important Nodes in Complex Networks Based on Node Propagation Entropy. *Entropy* **2022**, *24*, 275. [\[CrossRef\]](#)
32. Namtirtha, A.; Dutta, B.; Dutta, A. Semi-global triangular centrality measure for identifying the influential spreaders from undirected complex networks. *Expert Syst. Appl.* **2022**, *206*, 117791. [\[CrossRef\]](#)

33. Li, H.; Shang, Q.; Deng, Y. A generalized gravity model for influential spreaders identification in complex networks. *Chaos Solitons Fractals* **2021**, *143*, 110456. [[CrossRef](#)]
34. Zeng, Q.; Liu, Y.; Tang, M.; Gong, J. Identifying super-spreaders in information–epidemic coevolving dynamics on multiplex networks. *Knowl.-Based Syst.* **2021**, *229*, 107365. [[CrossRef](#)]
35. Kempe, D.; Kleinberg, J.; Tardos, É. Maximizing the spread of influence through a social network. In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 24–27 August 2003; pp. 137–146.
36. Morone, F.; Makse, H.A. Influence maximization in complex networks through optimal percolation. *Nature* **2015**, *524*, 65–68. [[CrossRef](#)] [[PubMed](#)]
37. Cao, T.; Wu, X.; Wang, S.; Hu, X. OASNET: An optimal allocation approach to influence maximization in modular social networks. In Proceedings of the 2010 ACM Symposium on Applied Computing, Sierre, Switzerland, 22–26 March 2010; pp. 1088–1094.
38. Estevez, P.A.; Vera, P.; Saito, K. Selecting the most influential nodes in social networks. In Proceedings of the 2007 International Joint Conference on Neural Networks, Orlando, FL, USA, 12–17 August 2007; pp. 2397–2402.
39. Ullah, A.; Wang, B.; Sheng, J.; Long, J.; Khan, N.; Sun, Z. Identifying vital nodes from local and global perspectives in complex networks. *Expert Syst. Appl.* **2021**, *186*, 115778. [[CrossRef](#)]
40. Goyal, A.; Lu, W.; Lakshmanan, L.V. Celf++ optimizing the greedy algorithm for influence maximization in social networks. In Proceedings of the 20th International Conference Companion on World Wide Web, Hyderabad, India, 28 March–1 April 2011; pp. 47–48.
41. Wang, Y.; Cong, G.; Song, G.; Xie, K. Community-based greedy algorithm for mining top-k influential nodes in mobile social networks. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 24–28 July 2010; pp. 1039–1048.
42. Zareie, A.; Sheikahmadi, A.; Jalili, M. Identification of influential users in social network using gray wolf optimization algorithm. *Expert Syst. Appl.* **2020**, *142*, 112971. [[CrossRef](#)]
43. Antelmi, A.; Cordasco, G.; Spagnuolo, C.; Szufel, P. Social influence maximization in hypergraphs. *Entropy* **2021**, *23*, 796. [[CrossRef](#)] [[PubMed](#)]
44. Fu, B.; Zhang, J.; Bai, H.; Yang, Y.; He, Y. An Influence Maximization Algorithm for Dynamic Social Networks Based on Effective Links. *Entropy* **2022**, *24*, 904. [[CrossRef](#)]
45. Zhao, J.; Wu, J.; Feng, X.; Xiong, H.; Xu, K. Information propagation in online social networks: A tie-strength perspective. *Knowl. Inf. Syst.* **2012**, *32*, 589–608. [[CrossRef](#)]
46. Kempe, D.; Kleinberg, J.; Tardos, É. Influential nodes in a diffusion model for social networks. In Proceedings of the International Colloquium on Automata, Languages, and Programming, Lisbon, Portugal, 11–15 July 2005; pp. 1127–1138.
47. Ruan, Z.; Wang, J.; Xuan, Q.; Fu, C.; Chen, G. Information filtering by smart nodes in random networks. *Phys. Rev. E* **2018**, *98*, 022308. [[CrossRef](#)] [[PubMed](#)]
48. Xuan, Q.; Shu, X.; Ruan, Z.; Wang, J.; Fu, C.; Chen, G. A self-learning information diffusion model for smart social networks. *IEEE Trans. Netw. Sci. Eng.* **2019**, *7*, 1466–1480. [[CrossRef](#)]
49. Watts, D.J.; Dodds, P.S. Influentials, networks, and public opinion formation. *J. Consum. Res.* **2007**, *34*, 441–458. [[CrossRef](#)]
50. De Arruda, G.F.; Barbieri, A.L.; Rodriguez, P.M.; Rodrigues, F.A.; Moreno, Y.; da Fontoura Costa, L. Role of centrality for the identification of influential spreaders in complex networks. *Phys. Rev. E* **2014**, *90*, 032812. [[CrossRef](#)] [[PubMed](#)]
51. Viswanath, B.; Mislove, A.; Cha, M.; Gummadi, K.P. On the evolution of user interaction in facebook. In Proceedings of the 2nd ACM Workshop on Online Social Networks, Barcelona, Spain, 17 August 2009; pp. 37–42.
52. Liu, D.; Nie, H.; Zhao, J.; Wang, Q. Identifying influential spreaders in large-scale networks based on evidence theory. *Neurocomputing* **2019**, *359*, 466–475. [[CrossRef](#)]
53. Zhang, J.; Zhang, Q.; Wu, L.; Zhang, J. Identifying Influential Nodes in Complex Networks Based on Multiple Local Attributes and Information Entropy. *Entropy* **2022**, *24*, 293. [[CrossRef](#)] [[PubMed](#)]
54. Kunegis, J. Konect: The koblenz network collection. In Proceedings of the 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil, 13–17 May 2013; pp. 1343–1350.
55. Cho, E.; Myers, S.A.; Leskovec, J. Friendship and mobility: User movement in location-based social networks. In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 21–24 August 2011; pp. 1082–1090.
56. Zafarani, R.; Liu, H. Social Computing Data Repository at ASU, 2009. Available online: <http://datasets.syr.edu/> (accessed on 7 September 2022).
57. De Choudhury, M.; Lin, Y.R.; Sundaram, H.; Candan, K.S.; Xie, L.; Kelliher, A. How does the data sampling strategy impact the discovery of information diffusion in social media? In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington, DC, USA, 23–26 May 2010.
58. Kendall, M.G. A new measure of rank correlation. *Biometrika* **1938**, *30*, 81–93. [[CrossRef](#)]
59. Berry, K.J.; Johnston, J.E.; Zahran, S.; Mielke, P.W. Stuart’s tau measure of effect size for ordinal variables: Some methodological considerations. *Behav. Res. Methods* **2009**, *41*, 1144–1148. [[CrossRef](#)]
60. Stuart, A. The estimation and comparison of strengths of association in contingency tables. *Biometrika* **1953**, *40*, 105–110. [[CrossRef](#)]

61. Wang, X.; Li, W.; Liu, L.; Pei, S.; Tang, S.; Zheng, Z. Promoting information diffusion through interlayer recovery processes in multiplex networks. *Phys. Rev. E* **2017**, *96*, 032304. [[CrossRef](#)]
62. Liu, L.; Wang, X.; Tang, S.; Zheng, H.; Zheng, Z. Multilayer social reinforcement induces bistability on multiplex networks. *J. Stat. Mech. Theory Exp.* **2021**, *2021*, 063402. [[CrossRef](#)]