Chaos, Solitons and Fractals 211 (2026) 118826

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

journal homepage: www.elsevier.com/locate/chaos

Hierarchical propagation for rumor source localization in signed networks

Wei-Wei Zhu [a,b], Qiang Guo [a,*], Matjaž Perc [c,d,e,f,*]

[a] Business School, University of Shanghai for Science and Technology, 200093, Shanghai, China
[b] School of Economics and Management, Chuzhou University, 239000, Chuzhou, China
[c] Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška Cesta 160, 2000 Maribor, Slovenia
[d] Community Healthcare Center Dr. Adolf Drolc Maribor, Ulica Talcev 9, 2000 Maribor, Slovenia
[e] Department of Physics, Kyung Hee University, 26 Kyungheedae-ro, Seoul 02447, Republic of Korea
[f] University College, Korea University, 145 Anam-ro, Seoul 02841, Republic of Korea

ARTICLE INFO

Keywords:
Complex networks
Source localization
Signed networks
Rumor propagation
Propagation dynamics model

ABSTRACT

We study rumor source localization in signed networks, where social ties may represent trust and support or distrust and opposition. These positive and negative relationships strongly influence how rumors spread, yet existing source localization methods for signed networks remain limited. In particular, current propagation models do not fully capture structural balance and may suffer from reverse infection, which overestimates the probability that some nodes become infected. We develop a source localization framework for signed networks when only a snapshot of observed node states is available and neither infection times nor propagation directions are known. First, we introduce a signed propagation model that incorporates structural balance and assigns more realistic transmission probabilities to positive and negative edges. Second, we propose a hierarchical propagation mechanism that follows the outward spread of rumors from each candidate source. By updating node states only through parent-layer and same-layer neighbors, our method prevents information from unrealistically flowing back toward the source. We show through experiments on scale-free and random networks that the proposed method improves localization accuracy and computational efficiency compared with current state-of-the-art propagation-based baselines. The advantage is especially clear in large networks, early stages of rumor spreading, sparse structures, low infection rates, and networks with a lower proportion of positive ties. Our results suggest that hierarchical propagation provides an effective and interpretable way to trace rumor sources in signed social networks.

1. Introduction

With the rapid development of online social networks such as social media and instant messaging, the efficiency and scope of information dissemination have achieved exponential growth. At the same time, the unregulated spread of online rumors and misinformation has become a core challenge in cyberspace governance. Rumor not only triggers public cognitive confusion and intensifies social group antagonism, but also poses a substantial threat to public health security, social stability, and national ideological security. Therefore, characterizing the propagation rules of online rumors and locating the rumor sources has been a research hotspot in the fields of network science, public opinion governance, and information security [1–4].

* Corresponding authors.
E-mail addresses: qiang.guo@usst.edu.cn (Q. Guo), matjaz.perc@um.si (M. Perc).

https://doi.org/10.1016/j.chaos.2026.118826
Received 30 June 2026; Received in revised form 3 July 2026; Accepted 13 July 2026

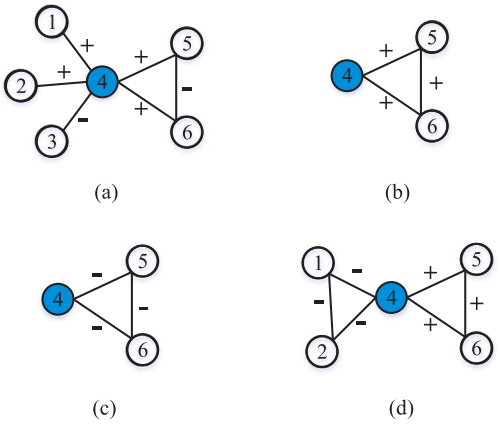

Fig. 1. Different types of signed networks structures. (a) A general signed network structure. (b) A triangle composed entirely of positive edges. (c) A triangle composed entirely of negative edges. (d) A triangle composed of positive edges and the other one composed of negative edges.

In early rumor propagation research, most relevant work was carried out based on the structure of unsigned complex networks, focusing on analyzing the influence of topological features on the propagation threshold, speed, and scope of rumors [5–7]. However, this type of research has a core limitation: in real social networks, the connections between nodes are not homogeneous and undifferentiated, but have significant signed characteristics. Specifically, the edges between nodes can be divided into positive relationships (such as trust, support, and identification) and negative relationships (such as distrust, opposition, and resistance). These two types of relationships have essentially different effects on rumor propagation: positive relationships promote the diffusion of rumors, while negative relationships inhibit the spread of rumors and even block the extension of propagation paths [8,9].

To more accurately characterize the laws of information propagation in real social networks, researchers have introduced the concept of signed networks into the field of rumor propagation research. A signed network is a complex network whose edges are assigned positive or negative signs, which can simultaneously characterize the topological structure of the network and the semantic features of the relationships between nodes [10]. It is perfectly suitable for the binary opposition relationship scenarios in social networks, such as trust/distrust and support/opposition, as shown in Fig. 1. The propagation heterogeneity brought by such signed characteristics cannot be accurately depicted by traditional unsigned network models, which also leads to a large deviation between the analysis results of traditional propagation models and the real propagation scenarios. As the core axioms state that the enemy of one's enemy is one's friend, the friend of one's friend is also one's friend, and the enemy of one's friend is one's enemy, structural balance is a critical feature that influences information propagation in signed networks. To better study information propagation in signed networks, Li et al. took structural balance into account, proposed an information propagation dynamics modeling method for signed network structures, and investigated propagation behaviors such as the information propagation threshold based on mean-field theory and the SIS propagation dynamics model [11]. However, this propagation dynamics modeling method has two limitations, which will be discussed in detail in Section 3 (Model and Methods).

As the research on the forward dynamics of rumor propagation continues to mature, researchers have gradually shifted their focus from the forward problem of "how rumors spread" to the inverse problem of "where rumors originate", namely the rumor source localization problem. The core goal of rumor source localization is to retroactively identify the initial propagation nodes of rumors based on the topological structure of the network, the infection states of nodes, and the dynamic rules of propagation. Traditional source localization methods are mostly based on the assumption of unsigned networks, and identify propagation sources through maximizing the propagation centrality of nodes [12,13], back-propagation [14–16], combining node state inference and maximum likelihood estimation (MLE) based on propagation dynamics models [17–19], and machine learning [20–23]. Among them, the source localization method based on combining node state inference and MLE via propagation dynamics models does not require to obtain the infection timestamp and direction of the observed nodes. For example, Lokhov et al. proposed the DMP source localization algorithm for the SIR propagation dynamics model based on the states of observed nodes at a specific moment [17,18]. However, this method suffers from the reverse infection phenomenon, which leads to an overestimation of the calculated node infection rate. For example, in Fig. 2, node 7 cannot spread information back to node 4 since node 1 is the source. Recently, Liu et al. established an SNIR false information propagation dynamics model by introducing the asymptomatic state, and inferred the initial propagation sources using MLE based on the states of observed nodes at a specific moment. However, this method adopts linear approximation to replace the nonlinear propagation process, and also has the reverse infection issue, thus leading to the overestimation of node infection states [24]. The kind source localization method only requires the state of observed nodes at a specific moment, and has strong interpretability. Therefore, this paper also studies source localization from this perspective.

However, the above source localization studies fail to account for the signed characteristics of social networks. At present, the relevant research on source localization in signed networks is still in its infancy. Through extensive literature search, only 3 academic papers have been published in this field, all from the same research team. Ma et al. took the lead in studying information source localization in signed networks [25]. They assumed that observed nodes can obtain the timestamp of information reception, proposed

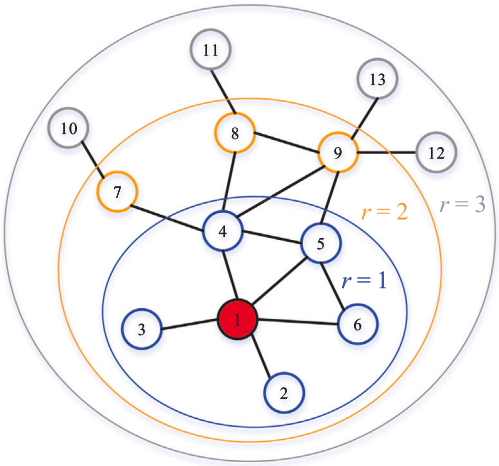

Fig. 2. Schematic diagram of rumor diffusion based on the hierarchical propagation. Assume that at the initial moment $t = 0$, node 1 is the source. When $t = 1$, only nodes within the blue region of $r = 1$ may be infected by node 1, and nodes with $r > 1$ will not be infected. Therefore, when $t = r$, nodes whose shortest path to node 1 is greater than r will not be infected. Under the hierarchical propagation rule, the state of a node is only affected by its parent-layer and same-layer nodes. For example, node 4 is only affected by node 1 and node 5, and not by nodes 7, 8, and 9. In fact, since node 1 is the initial spreader, node 7 cannot propagate rumor back to node 4. Nodes 8 and 9 need to propagate rumor to node 4 through long paths, so their influence on node 4 is negligible compared to that of node 1 and node 5.

the FONC method to select observer nodes, and finally performed source localization via MLE by continuously collecting information from observer nodes. However, this method does not take into account the structural balance property of signed networks. In addition, it simply sets the propagation rate of positive edges to a fixed value and that of negative edges to a smaller fixed value [25]. Subsequently, the team continued in-depth research, took the structural balance of signed networks into account, and proposed the DISLPSI source localization algorithm [26] and the S-DMP source localization algorithm [27]. Both algorithms infer propagation sources based on the states of observed nodes at a specific moment, but the former requires knowledge of infection direction of infected observers, while the latter leads to an overestimation of the calculated node infection probability due to the failure to eliminate reverse infection. However, in real-world networks with high noise, it is difficult to accurately obtain the infection timestamp and propagation direction of infected nodes. Therefore, inferring propagation sources only based on the node states at a specific observation moment is a more universal method.

To this end, aiming at the two limitations of the propagation dynamics modeling for signed networks proposed by Li et al. [11], this paper first re-establishes a propagation dynamics method that is more consistent with real-world scenarios. Unlike the higher-order structures in simplicial complexes [28–32], the triangles defined in our work are derived from the structural balance theory. The propagation rate of each edge takes into account not only the nature of positive and negative edges, but also the proportion of balanced or unbalanced triangles that the edge participates in. Second, aiming at the node state overestimation problem existing in the DMP-based source localization method [27] and the linear approximation-based diffusion model source localization framework [24], this paper proposes a hierarchical propagation-based node state estimation method, and infers the rumor source using MLE. Hierarchical propagation enables rumor to diffuse outward layer by layer starting from the propagation source, and the propagation is limited to parent nodes and same-layer nodes, thus completely eliminating the reverse propagation issue. The contributions of this work are as follows:

(1) A new rumor propagation dynamics modeling method between nodes in signed networks that is more consistent with real-world scenarios is proposed.

(2) Based on the states of observed nodes at a specific moment, a hierarchical propagation-based node state estimation method is proposed, and a rumor source localization method for signed network structures is developed accordingly.

(3) Extensive experiments show that the proposed source localization algorithm outperforms current state-of-the-art similar source localization methods in both computational efficiency and ranking accuracy. In addition, the proposed method is suitable for source localization at early stage and the ubiquitous sparse network structures. Meanwhile, it achieves higher localization accuracy when the proportion of positive edges is low and the infection rate is small.

2. Related works

Source localization methods are mainly divided into the following three categories. The first category is based on the states of observed nodes at a specific moment (snapshot), adopting either the maximum estimation method based on infection topological centrality [12,13], or the source localization method combining node state inference based on statistical physics with MLE [17–19,24,27]. The second category is based on the known infection timestamps of observed nodes, and locate the source via

back-propagation or MLE [14,33,34]. The third category is source localization using machine learning [20,35–37]. Machine learning-based source localization methods require a large number of simulations to obtain propagation data, and have poor interpretability. Recently, Wang et al. proposed a source localization method combining propagation dynamics and deep learning, which improved the interpretability [23]. In view of the particularity of information acquisition via machine learning, this paper will not conduct in-depth discussion on this category of methods.

In the field of source localization in complex networks, Shah et al. laid the foundational framework by proposing rumor centrality for rumor identification in complex networks, marking a cornerstone development in information source localization research [12]. Building on this work, Luo et al. developed the Jordan center estimation algorithm for source localization, which generally outperforms the classic rumor centrality algorithm in source localization accuracy [38], while Chen et al. extended the Jordan center concept for multi-source detection [39]. For practical multi-source localization scenarios, Jiang et al. divided the infected region into k subsets and applied Jordan centrality to each infected sub-graph for multi-source localization [40].

Considering that nodes surrounding the sources are usually also infected nodes, Wang et al. proposed the Label Propagation-based Source Identification (LPSI) method based on the source prominence [41]. Since the method requires full network observation, Hu et al. selected an approximately optimal observer node set via maximum entropy sampling, inferred the states of unobserved nodes based on the states of the observed nodes, and finally located the source based on the LPSI framework, which significantly improved the source localization accuracy [42]. Zhu et al. assumed that observed nodes can acquire the infection direction, and combined with the LPSI framework, investigated source localization scenarios applicable to a low infection rate [43]. The idea with label propagation can infer multiple sources based on the state of all nodes in the network at a given moment, which has been widely applied and extended in subsequent research. For example, Peng et al. incorporated information such as government reports, asymptomatic status, and quarantine measures, and applied this method to source localization for SEIR disease transmission [44]. Hu et al. constructed a risk propagation model for bank-enterprise financial networks based on lending-borrowing relationships, and adopted the LPSI method for risk source localization [45]. Ali et al. considered a weighted network and studied the rumor source localization problem in heterogeneous propagation environments [46]. Cheng et al. combined the path information of epidemic propagation with the LPSI method, and proposed the PSLM to deal with the problem of multiple source identification in multiplex networks [47].

Different from the aforementioned source localization methods based on the centrality of the topological structure of the infected sub-graph or label propagation, many scholars have recently combined propagation dynamics models to conduct source localization from the perspective of node state estimation. For example, Lokhov et al. proposed a DMP source localization algorithm for the SIR propagation dynamics model based on the states of observed nodes at a certain moment. However, this method suffers from the reverse infection phenomenon, which leads to an overestimation of the calculated node infection probability [17]. Yu et al. further applied the DMP method to source localization in complex hyper-graphs, achieving higher accuracy than traditional centrality-based methods [48]. Jiang et al. took into account the signed characteristics of social networks, extended the DMP algorithm based on the theory of structural balance, and developed the S-DMP source localization algorithm for signed networks [27]. As the DMP algorithm is suitable for single-source localization, Altarelli et al. proposed a BP source localization method, which can be applied to multi-source localization and has high innovativeness, however, its performance is not ensured when there are a large number of loops in the network structure [19]. Liu et al. established an SNIR misinformation propagation dynamics model by introducing the asymptomatic characteristics of nodes, and performed source localization based on the states of observed nodes and MLE. This work also conducted community division on the infected region, extending single-source localization to multi-source localization. But it adopts linear approximation to replace the nonlinear propagation process, and therefore also has the problem of overestimating the node infection state [24].

Distinct from the aforementioned approaches, a growing body of recent research has leveraged the infection moment and direction information of observed nodes to infer the source via back-propagation [14] or MLE [49]. For instance, Pinto et al. randomly selected observers and located the source by back-propagation [14]. To reduce the number of required observers, subsequent studies have demonstrated that early-infected observers can be effectively utilized for source localization tasks [50]. Spinelli et al. first deployed static observer nodes, then dynamically added additional observer nodes based on the observed propagation information, and finally performed source localization via back-propagation using the infection timestamp of the observers [51]. By calculating the delay matrix of information propagation, Hu et al. proposed a greedy algorithm to determine the optimal observer set, paired with a back-propagation algorithm for source localization [15]. Xu and Wang et al. selected observer nodes via topological centrality, and performed source localization by leveraging the correlation between topological distance and propagation delay based on the infection timestamps of the observer nodes [52,53]. Building on back-propagation, Yang et al. further incorporated diffusion path information and observer infection timestamps to develop an observer-similarity-based source localization method [16]. Ma et al. incorporated the infection direction information of observer nodes, selected observer nodes to achieve full network coverage via a greedy algorithm, and proposed a source localization method with network structure pruning based on infection rate paths and MLE, which significantly improved the source localization accuracy [4]. Based on degree and distance, Hou et al. proposed a greedy algorithm to select observer nodes for network coverage, and performed source localization combined with a penalty strategy based on the state of observer nodes at a specific moment, their information reception timestamps, and the propagation direction [54]. Devarapalli et al. proposed the Radial Observer Source Estimator (ROSE) algorithm, which first uses the two most informative observers to identify critical network regions, then applies the ROSE algorithm to locate source nodes within these critical regions [55]. However, all the aforementioned methods are not applicable to the source identification problem in multi-source triggered propagation scenarios. To address this gap, Fu et al. proposed the Maximum–Minimum Method (MMM) based on back-propagation, which can solve the multi-source localization problem with limited observation nodes [56].

Nevertheless, the performance of the MMM is susceptible to extreme values and degrades significantly in scenarios with a small number of observation nodes. Hu et al. improved the MMM by combining forward propagation modeling and integer programming, though the modified algorithm cannot handle non-integer time delay cases or strong noise environments [57]. Yuan et al. divided the infected region into communities, and located the source using the back-propagation algorithm in each community [58].

The only three existing studies on source localization in signed networks have been discussed in detail in the Introduction, and therefore, no further elaboration will be provided here.

3. Model and methods

A signed network structure is defined as G = (V, E, **W**), which consists of $N = |V|$ nodes and |E| edges, where V is the set of nodes in the network, E is the set of edges, and **W** is the weight matrix with elements $w_{ij} \in \{1, 0, -1\}$. Specifically, $w_{ij} = 1$ indicates that the edge e_{ij} between node i and node j is a positive edge, $w_{ij} = -1$ indicates a negative edge, and $w_{ij} = 0$ means there is no direct connection between node i and node j. The set of observer nodes is denoted as $O = \{o_1, o_2, \ldots, o_{|O|}\}$, where |O| is the number of observer nodes, satisfying $|O| = p_{obs} \times N$. Here, p_{obs} is the proportion of observer nodes, and the states of these nodes at a specific moment are recorded.

The research objective of this paper is to address two core challenges of source localization in signed networks based on the states of selected observer nodes at a specific moment. (1) How to characterize the differential impact of positive and negative signs on propagation paths, and construct a signed propagation dynamics model that conforms to real-world propagation laws. (2) In a high-noise environment where infection timestamps and directions are unknown, how to infer the initial spreaders of rumors in signed networks based on the states of observers at a specific moment, while achieving high source localization accuracy and computational efficiency.

3.1. Propagation dynamics model for signed networks

This paper adopts the classic SIR propagation dynamics model to simulate rumor propagation. In this model, there are three states: the susceptible state S, the infected state I, and the recovered state R. A node can only be in one of these states at any given moment. In the unsigned complex networks, the probability that a susceptible node gets infected when in contact with an infected node is called the propagation rate λ, and the probability that an infected node recovers is called the recovery rate μ.

Since this study is highly relevant to the signed network propagation dynamics modeling method proposed by Li et al. [11], we first restate this method here. The probability that an uninfected node i gets infected at time t is given by [11]

$$\xi_i(t) = \xi_i^+(t-1) \sum_{j=1}^{N} a_{ij} + \xi_i^-(t-1) \sum_{j=1}^{N} b_{ij}, \tag{1}$$

where a_{ij} is the element of the positive edge adjacency matrix, and b_{ij} is the element of the negative edge adjacency matrix. $\xi_i^+(t)$ and $\xi_i^-(t)$ are the probabilities that the uninfected node i gets infected via positive and negative edges at time t, respectively, which are calculated by

$$\xi_i^+(t) = \lambda \left(\frac{\sum_{j=1}^{N} \beta_{ij}^+ x_j(t)}{k_i^+} \right), \tag{2}$$

and

$$\xi_i^-(t) = \lambda \left(\frac{\sum_{j=1}^{N} \beta_{ij}^- x_j(t)}{k_i^-} \right), \tag{3}$$

where $x_j(t)$ is the state of node j at time t, the denominators are the number of positive and negative edges of node i, respectively. λ is the constant infection rate of models such as the SIS or SIR model. k_i^+ and k_i^- denote the number of positive and negative edges of node i, respectively. β_{ij}^+ and β_{ij}^- are the propagation rates of the positive and negative edge e_{ij}, respectively, calculated by

$$\beta_{ij}^+ = \frac{\sum_{k=1}^{N_+} \max\left(0, w_{ij} w_{ik} w_{jk}\right)}{\sum_{k=1}^{N} \left| w_{ij} w_{ik} w_{jk} \right|}, \tag{4}$$

and

$$\beta_{ij}^- = \frac{\sum_{k=1}^{N_-} \min\left(0, w_{ij} w_{ik} w_{jk}\right)}{\sum_{k=1}^{N} \left| w_{ij} w_{ik} w_{jk} \right|}, \tag{5}$$

where N_+ and N_- are the number of balanced and unbalanced triangles that edge e_{ij} participates in, respectively.

The above propagation dynamics modeling method has two critical limitations:

Limitation 1: The influence of nodes that do not participate in any triangle is ignored. For example, in Fig. 1(a), nodes 1, 2, and 3 do not form a triangular structure with any other nodes. Even if all three nodes are infected, the susceptible node 4 will not be infected by them.

Limitation 2: The designs of the infection rates β_{ij}^{+}, β_{ij}^{-} for positive and negative edges participating in triangles, and the node infection probability $\xi_i(t)$ are unreasonable. As shown in Fig. 1(a), for the unbalanced triangle formed by nodes (4,5,6), according to Eqs. (4) and (5), we can obtain $\beta_{45}^{+} = \beta_{46}^{+} = 0$ and $\beta_{45}^{-} = \beta_{46}^{-} = -1$. Combined with Limitation 1 and Eqs. (1)–(3), even if all neighbors of node 4 are infected, the infection probability of the uninfected node 4 is 0. Next, consider Fig. 1(b), a balanced triangle composed entirely of positive edges. According to Eqs. (4) and (5), we have $\beta_{45}^{+} = \beta_{46}^{+} = 1$ and $\beta_{45}^{-} = \beta_{46}^{-} = 0$. Therefore, if both neighbors of node 4 are infected, the infection probability of the uninfected node 4 is 2λ. Then look at Fig. 1(c), an unbalanced triangle composed entirely of negative edges. According to Eqs. (4) and (5), we get $\beta_{45}^{+} = \beta_{46}^{+} = 0$ and $\beta_{45}^{-} = \beta_{46}^{-} = -1$. Thus, if both neighbors of node 4 are infected, the infection probability of the uninfected node 4 is -2λ, which is an absurd negative infection rate. Finally, for Fig. 1(d), based on the analysis of Fig. 1(b) and (c), based on Eq. (1), if all four neighbors of node 4 are infected, the infection probability of node 4 is 0. All these conclusions are far from real-world scenarios.

In summary, it is clear that this information propagation dynamics modeling method for signed networks has inherent flaws and deviates from reality. To address this, we modify the method and design an information propagation dynamics model that is more consistent with real-world scenarios.

First, for the Limitation 1, we consider edges that do not participate in any triangle. If e_{ij} is a positive edge (i.e., $w_{ij} = 1$), the propagation rate of this positive edge is set as

$$\lambda_{ij}^{+} = 2\lambda. \tag{6}$$

If e_{ij} is a negative edge (i.e., $w_{ij} = -1$), the propagation rate of this edge is set as

$$\lambda_{ij}^{-} = \lambda/2. \tag{7}$$

The rationale for this design is as follows: for edges without reference triangles, the connected nodes often act as bridges or have few friends. When an edge serves as a bridge or a node is isolated, the node is more willing to share information. We set a lower propagation rate for negative edges due to the hostile relationship, which reduces the willingness to propagate information. In practice, other values can also be used, such as setting the positive edge propagation rate to λ and the negative edge propagation rate to $\lambda/4$.

Second, Limitation 2, we analyze the edges participating in triangles. If e_{ij} is a positive edge, then the propagation rate of this positive edge is

$$\lambda_{ij}^{+} = \min\left(1, \lambda\left(1 + \beta_{ij}^{+}\right)\right). \tag{8}$$

If e_{ij} is a negative edge, then the propagation rate of this negative edge is

$$\lambda_{ij}^{-} = \lambda\left(1 + \beta_{ij}^{-}\right). \tag{9}$$

The two terms β_{ij}^{+} and β_{ij}^{-} in the above two formulas are calculated by Eqs. (4) and (5). It should be noted that the term $+\beta_{ij}^{-}$ corresponds to subtracting a probability, given that β_{ij}^{-} is inherently negative in Eq. (5).

Consistent with the analysis of Limitation 2, we demonstrate the validity of our method using concrete examples. As shown in Fig. 1(a), nodes 4, 5 and 6 form an unbalanced triangle, where edges e_{45} and e_{46} are positive. According to Eqs. (8) and (9), the propagation rates of e_{45} and e_{46} are both equal to λ. Meanwhile, edge e_{56} is a negative edge with $w_{56} = -1$, yielding a propagation rate $\lambda_{56}^{-} = 0$. Even so, nodes 5 and 6 can still disseminate information to node 4 with a probability of λ. In Fig. 1(b), all edges are positive. Hence, the propagation rate of each edge is $\min(1, 2\lambda)$. For the structure in Fig. 1(c), all edges are negative. In accordance with Eqs. (5) and (9), the propagation rate of every edge is 0. This setting is reasonable given that all nodes in this structure are mutually hostile. Finally, we examine Fig. 1(d): the propagation rate for edges within the unbalanced triangle consisting of negative edges is 0, while all edges in the positive-edge triangle have a propagation rate of 2λ. Therefore, the illustrative cases under various conditions all verify that our model is more reasonable.

In this work, we assume that information propagation across different edges is mutually independent. Different from the additive probability formulation adopted in Eq. (1), we adopt a nonlinear function to calculate the infection probability of node i. This design guarantees that the calculated probability never exceeds 1 and achieves better rationality, as defined in the following equation

$$\xi_i(t) = 1 - \prod_{j \in \Gamma_i^{r=1}} \left(1 - \Lambda_{ij} x_j(t-1)\right), \tag{10}$$

where $\Gamma_i^{r=1}$ denotes the first-order neighbors of node i, and r stands for the neighbor order. Λ_{ij} represents the propagation rate of the edge connecting node i and node j, which is defined as

$$\Lambda_{ij} = \begin{cases} \lambda_{ij}^{+} & \text{if } w_{ij} = 1, \\ \lambda_{ij}^{-} & \text{if } w_{ij} = -1. \end{cases} \tag{11}$$

Therefore, based on the proposed propagation model and corresponding illustrative examples, we can conclude that the model reduces to simple pairwise propagation in the absence of triangular structures, where the transmission rate between node pairs is given by Eqs. (6) or (7). When triangular structures are involved, the propagation rate between nodes defined in this paper differs fundamentally from pairwise propagation, as well as the reinforcement effect featured in simplicial complexes. This is because it is also associated with structural balance in signed networks, and additionally depends on the number of triangles the edge participates in.

3.2. Source localization method for signed networks

Although Jiang et al. [27] considered structural balance and proposed one source localization algorithm based on DMP for signed network structures, the signed network propagation dynamics model they used is still based on the work of Li et al. [11]. Meanwhile, the DMP algorithm suffers from the reverse infection phenomenon. As shown in Fig. 2, assume node 1 is the initial propagation source: the infected node 4 will propagate to node 7 with a certain probability, but node 7 cannot propagate back to node 4. The DMP-based source localization algorithm ignores this fact, leading to an overestimation of the marginal probability.

In this paper, we only replace the propagation dynamics model proposed by Li et al. (used by Jiang et al.) with our proposed signed network propagation dynamics model, with no other modifications. Therefore, in this paper, we still refer to this method as "Jiang 2024 CSF".

Liu et al. proposed an SNIR epidemic propagation dynamics model based on the diffusion idea, which introduced the asymptomatic state to study misinformation source localization [24]. For comparison, this paper does not consider the asymptomatic state during rumor propagation, and modifies the model to the classic SIR propagation dynamics combined with the properties of signed networks. The details of the modified source localization algorithm are as follows.

Define $P_i(t) = \left[p_i^{\mathrm{S}}(t), p_i^{\mathrm{I}}(t), p_i^{\mathrm{R}}(t)\right]$ as the probability that node i is in each state at the time t. Let the source be node u = source. Initially, at $t = 0$, the state of the source is $P_u(t = 0) = [0, 1, 0]$, corresponding to the probabilities of the three states in the SIR model, respectively. When $t > 0$, the source recovers to the R state with probability μ at each moment, so we have

$$P_u(t) = \left[0, (1-\mu)^t, 1-(1-\mu)^t\right]. \tag{12}$$

For other non-initial source nodes, i.e., $i \in \mathrm{V} \setminus \{u\}$, the initial state at $t = 0$ is

$$P_i(t = 0) = [1, 0, 0]. \tag{13}$$

When $t > 0$, based on the mean-field theory, the probability that node i is in the susceptible S state is

$$p_i^{\mathrm{S}}(t) = p_i^{\mathrm{S}}(t-1) - \sum_{j \in \Gamma_i^1} p_i^{\mathrm{S}}(t-1) \times p_j^{\mathrm{I}}(t-1) \times \Lambda_{ij}, \tag{14}$$

where Γ_i^1 denotes the first-order neighborhood of node i, i.e., its directly connected neighbors. Here, we replace the constant propagation rate in the original unsigned network model of Liu et al. [24] with the inter-node propagation rate Λ_{ij} calculated by Eq. (11), which reflects the characteristics of signed networks.

The probability that node i is in the infected I state is

$$p_i^{\mathrm{I}}(t) = \sum_{j \in \Gamma_i^1} p_i^{\mathrm{S}}(t-1) \times p_j^{\mathrm{I}}(t-1) \times \Lambda_{ij} + (1-\mu) \times p_i^{\mathrm{I}}(t-1). \tag{15}$$

The probability that node i is in the recovered R state is

$$p_i^{\mathrm{R}}(t) = 1 - p_i^{\mathrm{S}}(t) - p_i^{\mathrm{I}}(t), \tag{16}$$

which can also be written as

$$p_i^{\mathrm{R}}(t) = p_i^{\mathrm{I}}(t-1) \times \mu + p_i^{\mathrm{R}}(t-1). \tag{17}$$

Since the above iterative method for calculating the probability of each node being in each state at each moment is the core idea of Liu et al. [24], in this paper, we refer to this method as "Liu 2025 IMP".

It can be seen from the above formulas that the diffusion-based "Liu 2025 IMP" calculates the infection probability of a node as the sum of the propagation probabilities of its surrounding infected neighbors. In fact, this linear approximation is only accurate when the infection rate approaches 0, which leads to an overestimation of the node infection probability.

Aiming at the problems existing in "Jiang 2024 CSF" and "Liu 2025 IMP", this paper proposes a hierarchical propagation-based node state estimation method, referred to as "Our method". The details of the method are as follows.

Initially, at $t = 0$, the state of the initial propagation source is $P_u(t = 0) = [0, 1, 0]$, corresponding to the probabilities of the three SIR states, respectively. When $t > 0$, it satisfies $P_u(t) = \left[0, (1-\mu)^t, 1-(1-\mu)^t\right]$, which is consistent with the previous setting.

For other non-initial source nodes, i.e., $i \in \mathrm{V} \setminus \{u\}$, the initial state is $P_i(t = 0) = [1, 0, 0]$.

When $t > 0$, the state $P_i(t) = \left[p_i^{\mathrm{S}}(t), p_i^{\mathrm{I}}(t), p_i^{\mathrm{R}}(t)\right]$ of node i is calculated as follows.

For the rth order neighbor of the source node u, i.e., $i \in \Gamma_u^r$:

(1) If the propagation moment $t < r$, the outward propagation from node u has not reached the rth order neighbor, and the propagation step has not reached the shortest path length between node u and node i, then node i must be in the susceptible S state. Therefore, the state of node i is

$$P_i(t \mid t < r) = [1, 0, 0]. \tag{18}$$

(2) If the propagation moment $t \geq r$, according to the mean-field theory, the probability that node i is in the susceptible S state is expressed by

$$p_i^{\mathrm{S}}(t \mid t \geq r) = p_i^{\mathrm{S}}(t-1) - \left(1 - \prod_{j \in \Gamma_i^1 \cap \partial_u^r} \left[1 - p_i^{\mathrm{S}}(t-1) \times p_j^{\mathrm{I}}(t-1) \times \Lambda_{ij}\right]\right), \tag{19}$$

where ∂_u^r denotes the set of nodes with a shortest path length of r centered at node u, i.e., the set of $[0, 1, \ldots, r]$-th order neighbors of node u, where the 0-th order neighbor is node u itself.

If $t \geq r$, the probability that node i is in the infected I state is

$$p_i^{\mathrm{I}}(t \geq r) = 1 - \prod_{j \in \Gamma_i^1 \cap \partial_u^r} \left[1 - p_i^{\mathrm{S}}(t-1) \times p_j^{\mathrm{I}}(t-1) \times \Lambda_{ij}\right] + (1-\mu) \times p_i^{\mathrm{I}}(t-1). \tag{20}$$

The probability that node i is in the recovered R state is the same as Eqs. (16) and (17).

As shown in Fig. 2, assume node 1 is the initial propagation source, and node 4 is in the $r = 1$ layer. Therefore, when calculating the state of node 4, we only analyze the influence of node 1 and node 5 on node 4, and do not consider the other three nodes 7, 8, and 9 with $r > 1$. This hierarchical propagation idea eliminates the reverse infection phenomenon. Although node 7 cannot infect node 4, nodes 8 and 9 may infect node 4. But the probability of infecting node 4 through multi-hop paths is much lower than that through direct neighbors (such as node 1 and node 5). Assume node 4 is infected by node 9, then the infection path is node 1 → node 5 → node 9 → node 4 or node 1 → node 6 → node 5 → node 9 → node 4. However, the latter occurs with a relatively low probability. Therefore, we focus our analysis on the former. Node 1 starts propagating at $t = 0$, and the earliest moment for node 9 to infect node 4 through this specified path is $t = 3$.

At $t = 1$, the probability that node 1 infects node 5 and does not infect node 4 is $(1 - \Lambda_{14}) \times \Lambda_{15}$.

At $t = 2$, the probability that node 1 has not recovered and has not infected node 4, node 5 has not infected node 4, and node 5 has infected node 9 is $\left[1 - (1-\mu)\Lambda_{14}\right] \times (1 - \Lambda_{54}) \times \Lambda_{59}$.

At $t = 3$, the probability that node 1 has not recovered and has not infected node 4, node 5 has not recovered and has not infected node 4, and node 9 has infected node 4 is $\left[1 - (1-\mu)^2\Lambda_{14}\right] \times \left[1 - (1-\mu)\Lambda_{54}\right] \times \Lambda_{94}$.

Node 4 can be infected by node 9 at $t = 3$ if and only if the above three events occur simultaneously, so the probability of this event is the product of the three probabilities.

Obviously, this probability is much lower than the probability of node 1 directly propagating to node 4 and node 5 propagating to node 4. Although we can calculate the probability of the event that node 1 → node 5 → node 9 → node 4 occurring at $t > 3$, the probability is even lower. If we consider this kind of reverse infection of parent-layer nodes through multi-hop cyclic paths, the improvement in the accuracy of node state estimation is limited, while the computational complexity will increase significantly. Therefore, this paper ignores this scenario.

By treating each node as the initial propagation source, we can use the above method to infer the probability of each node being in each state. Finally, combined with the states of the observed nodes, we use Bayesian theorem and MLE to locate the source. According to Bayesian theorem, the probability that node i is the propagation source satisfies $P(i \mid \mathrm{O}) \propto P(\mathrm{O} \mid i)$. Since it is almost impossible to calculate the joint probability of the observed node states, we adopt the mutual independence assumption as in [17,24,27], which gives

$$P(\mathrm{O} \mid i) \approx \prod_{\substack{k \in \mathrm{O} \\ q_k(t)=\mathrm{S}}} P_k^{\mathrm{S}}(t,i) \times \prod_{\substack{l \in \mathrm{O} \\ q_l(t)=\mathrm{I}}} P_l^{\mathrm{I}}(t,i) \times \prod_{\substack{m \in \mathrm{O} \\ q_m(t)=\mathrm{R}}} P_m^{\mathrm{R}}(t,i), \tag{21}$$

where $P_k^{\mathrm{S}}(t,i)$ denotes the probability that node k is in the S state at moment t when node i is the propagation source. The other two terms are defined similarly. However, the product of probabilities can easily lead to a value of 0, making it difficult to distinguish the probability magnitudes calculated by the above formula. Therefore, we take the logarithm of Eq. (21)

$$\log P(\mathrm{O} \mid i) \approx \sum_{\substack{k \in \mathrm{O} \\ q_k(t)=\mathrm{S}}} P_k^{\mathrm{S}}(t,i) + \sum_{\substack{l \in \mathrm{O} \\ q_l(t)=\mathrm{I}}} P_l^{\mathrm{I}}(t,i) + \sum_{\substack{m \in \mathrm{O} \\ q_m(t)=\mathrm{R}}} P_m^{\mathrm{R}}(t,i). \tag{22}$$

Finally, the identified propagation source satisfies

$$\mathrm{S}^* = \arg\max_{i \in \mathrm{V}} \log P(\mathrm{O} \mid i). \tag{23}$$

4. Experimental results and analysis

In the paper, we adopt the ranking value as the evaluation metric for the accuracy of source localization methods. Since the proposed method calculates the probability of each node being the propagation source, we sort these probabilities and take the corresponding rank as the localization accuracy. A rank of 1 indicates perfect source localization, and a larger rank value means lower accuracy. For example, if the true source is Node 1 and it is ranked 10th by the algorithm, the corresponding rank is 10.

The baseline algorithms for comparison include "Jiang 2024 CSF" and "Liu 2025 IPM". The time complexity of "Jiang 2024 CSF" is $O(t_{\mathrm{obs}} \times N \times N \times \langle k \rangle \times 3\langle k \rangle)$, while both "Liu 2025 IPM" and "Our" methods have a time complexity of $O(t_{\mathrm{obs}} \times N \times N \times \langle k \rangle)$. Here, the first N refers to the computation for each node as a potential propagation source, and the second N represents the state update of all nodes at each time step. $\langle k \rangle$ denotes the average node degree. The term $3\langle k \rangle$ in "Jiang 2024 CSF" arises because the algorithm requires computations for the neighbors of nodes i, j and k [27].

The experimental hardware is configured with an Intel(R) Core(TM)i7-7700 CPU and 16.0 GB RAM. We conduct running time experiments on BA networks [59] with the following default parameters: average node degree $\langle k \rangle = 8$, observation time $t_{\mathrm{obs}} = 5$, infection rate $\lambda = 0.15$, positive edge ratio pos_ratio = 0.5, and observer proportion $p_{\mathrm{obs}} = 0.7$.

The basic characteristics of the datasets used in this paper are shown in Table 1.

Table 1
Average running time (seconds) of different algorithms for calculating the probability of a single node being the propagation source with 10 independent runs.

N	Time (s) for "Liu 2025 IPM"	Time (s) for "Jiang 2024 CSF"	Time (s) for "Our"
100	0.0082	0.1353	0.0061
1000	0.1173	2.2816	0.0683

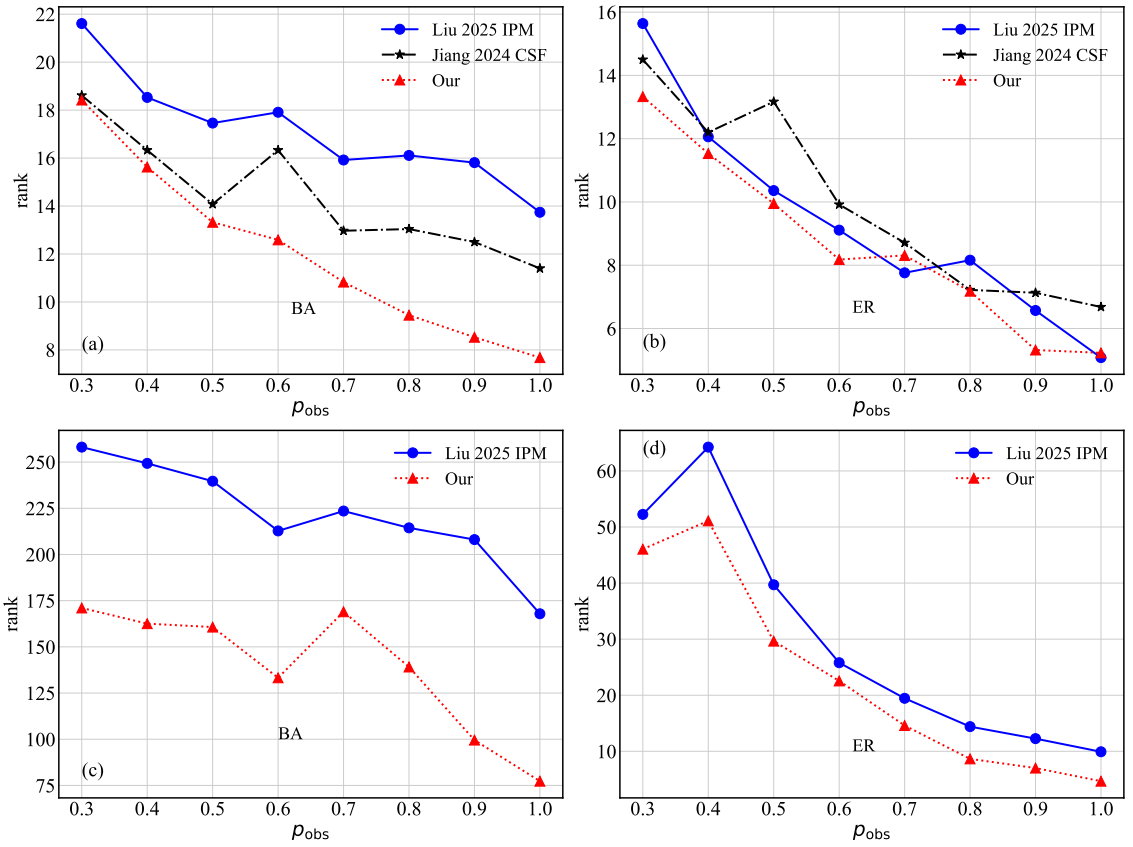

Fig. 3. Relationships between observer proportion p_{obs} and the rank of different algorithms. (a) and (c) are results on BA networks; (b) and (d) are results on ER networks. The network size is $N = 100$ for (a, b) and $N = 1000$ for (c, d). The average node degree $\langle k \rangle = 8$, infection rate $\lambda = 0.15$, observation time $t_{obs} = 5$, and positive edge ratio pos_ratio $= 0.5$. All results are averaged with 100 independent runs.

As shown in Table 1, "Liu 2025 IPM" and "Our" method have comparable time consumption. When $N = 100$, the running time of "Jiang 2024 CSF" based on DMP is more than 20 times that of our method. When $N = 1000$, "Liu 2025 IPM" takes about twice as long as our method, while "Jiang 2024 CSF" is over 30 times slower.

To compute the probability of every node being the propagation source, the total running time equals the single-node runtime multiplied by the network size N. For instance, for a BA network with $N = 1000$, the runtime of "Jiang 2024 CSF" for all nodes is approximately 2.2816×1000 seconds. If the experiment is repeated 100 times to obtain the average result, the total time reaches $2.2816 \times 1000 \times 100 = 228160$ seconds, equivalent to 63.38 h. When further analyzing the impacts of different parameters (e.g., observation time t_{obs} ranging from 1 to 10), the total runtime will rise to 633.8 h. However, the computational efficiency can be greatly improved if GPU acceleration is adopted. Due to the limitations of our computational hardware, therefore, we exclude "Jiang 2024 CSF" from subsequent experiments when $N = 1000$.

To comprehensively explore the effects of observer proportion, observation time, infection rate, positive edge ratio and network sparsity on localization performance, we adopt two classic synthetic networks: BA networks with power-law degree distribution [59] and ER random networks with normal degree distribution [60].

4.1. Influence of observer proportion

First, we analyze how the proportion of observer nodes affects the performance of each algorithm. A low observer proportion means limited information can be collected from a single snapshot. Accordingly, we set the range of the observer proportion from

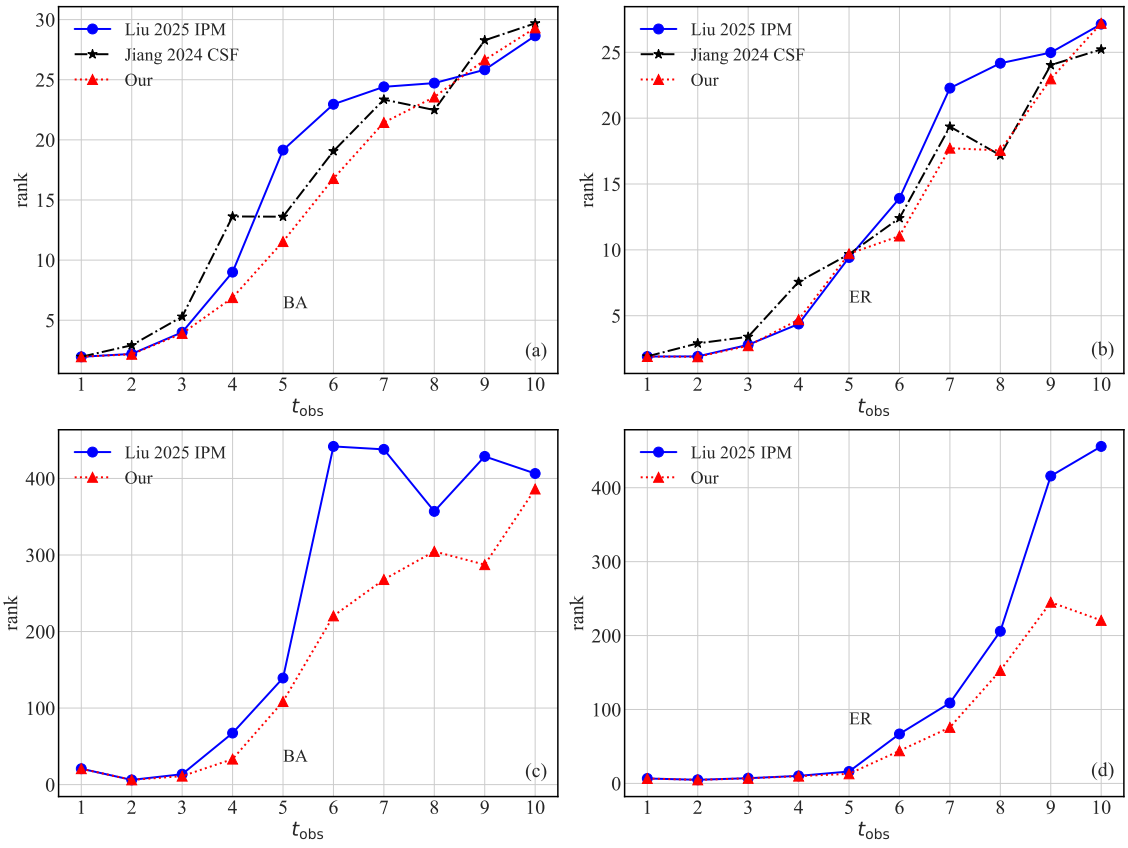

Fig. 4. Relationships between observation time t_{obs} and the rank of different algorithms. (a) and (c) are results on BA networks; (b) and (d) are results on ER networks. The network size is $N = 100$ for (a,b) and $N = 1000$ for (c,d). The average node degree $\langle k \rangle = 8$, infection rate $\lambda = 0.15$, observer proportion $p_{obs} = 0.7$, and positive edge ratio pos_ratio $= 0.5$. All results are averaged with 100 independent runs.

0.3 to 1.0 with a step size of 0.1. As shown in Fig. 3, it can be clearly observed that the rank decreases as the observer proportion increases, which indicates an improvement in localization accuracy.

Overall, our method achieves the lowest rank in most cases, demonstrating its superior performance. This advantage is particularly prominent for large-scale networks with $N = 1000$, as presented in Fig. 3(c, d). On BA networks, the rank of our algorithm is far lower than that of the other two methods. On ER networks, our approach still performs best, while the performance gap among the three algorithms is relatively small. Furthermore, all algorithms yield lower ranks on ER networks compared with BA networks, which verifies that higher localization accuracy can be achieved on more homogeneous network structures.

4.2. Influence of observation moment

Next, we analyze the impact of observation time on source localization performance. As clearly shown in Fig. 4, localization accuracy decreases slightly at first as the observation moment elapses, and then drops sharply once it exceeds a certain threshold. Specifically, for networks with $N = 100$, a dramatic decline occurs when $t_{obs} > 3$; for networks with $N = 1000$, the critical threshold is $t_{obs} > 4$. The results also indicate that "Jiang 2024 CSF" delivers the lowest accuracy at early observation times, while "Liu 2025 IPM" performs worst at late observation times. For small networks with $N = 100$, the ranks of the three algorithms are comparable when t_{obs} is greater than 8. Overall, our algorithm achieves the best performance.

Consistent with previous findings, our method advantages become more prominent on large-scale networks. In addition, all algorithms achieve higher localization accuracy on ER networks than on BA networks, as illustrated in Fig. 4(c, d). When the observation time is extremely late, rumors have spread across the entire network with a large number of infected nodes, making accurate source localization challenging for all methods. For instance, on BA networks with $N = 1000$, the rank of "Liu 2025 IPM" accounts for approximately 40% of the total network size when $t_{obs} > 6$. In this case, its performance is only marginally better than random guessing, where the expected rank equals $N/2$.

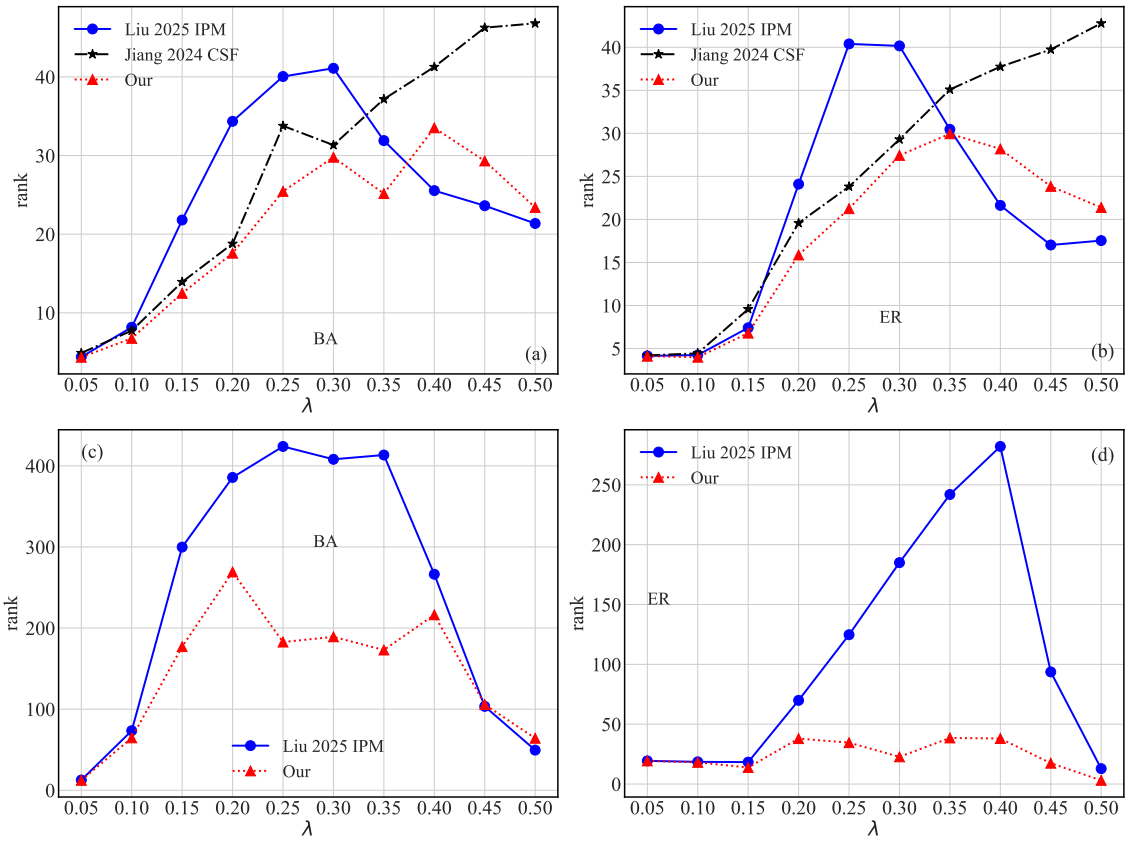

Fig. 5. Relationships between infection rate λ and the rank of different algorithms. (a) and (c) are results on BA networks; (b) and (d) are results on ER networks. The network size is $N = 100$ for (a, b) and $N = 1000$ for (c, d). The average node degree $\langle k \rangle = 8$, observer proportion $p_{\text{obs}} = 0.7$, observation time $t_{\text{obs}} = 5$, and positive edge ratio pos_ratio = 0.5. All results are averaged with 100 independent runs.

4.3. Influence of infection rate

Fig. 5 presents the experimental results regarding infection rate. The recovery rate is set to 0.5, and an excessively high infection rate will lead to rapid network-wide rumor propagation. Thus we set the infection rate ranging from 0.05 to 0.5 with a step size of 0.05. From Fig. 5, we can see that when $\lambda \leq 0.1$, all three algorithms achieve high accuracy with similar performance, because the rumor spreads within a limited range at low infection rates. As the infection rate increases, the three algorithms exhibit different patterns in small-scale networks, as shown in Fig. 5(a, b). For example, for small networks ($N = 100$), "Liu 2025 IPM" and "Our" method show a trend of rank rising first and then falling as λ increases, whereas the rank of "Jiang 2024 CSF" keeps rising. When $\lambda \leq 0.35$, our method has the lowest rank and "Liu 2025 IPM" performs the worst. When λ exceeds 0.35, "Liu 2025 IPM" surpasses the others.

Interestingly, when the network scale is large, as shown in Fig. 5(c, d) for $N = 1000$, our algorithm consistently performs best, which is different from the case when $N = 100$. For large networks ($N = 1000$), both algorithms obtain high ranks (low accuracy) when λ is between 0.15 and 0.4. However, our method is more robust for relatively homogeneous ER network structures, while "Liu 2025 IMP" performs poorly when the infection rate is between 0.2 and 0.45. Overall, "Our" method performs best in all cases except when the infection rate is high and the network size is small, where its accuracy is slightly lower than that of "Liu 2025 IPM".

The above variation trends of ranking values for "Liu 2025 IPM" and "Our" method with changing infection rates can be explained as follows. At low infection rates, rumors spread slowly over a limited range, leading to low ranking values. As the infection rate rises, rumors propagate more widely alongside increased propagation uncertainty, which degrades localization accuracy and pushes up the rank. When the infection rate exceeds a certain threshold, the propagation process becomes more deterministic. Rumors can easily spread to neighboring nodes, and the localization accuracy gradually improves, accompanied by a decline in ranking values.

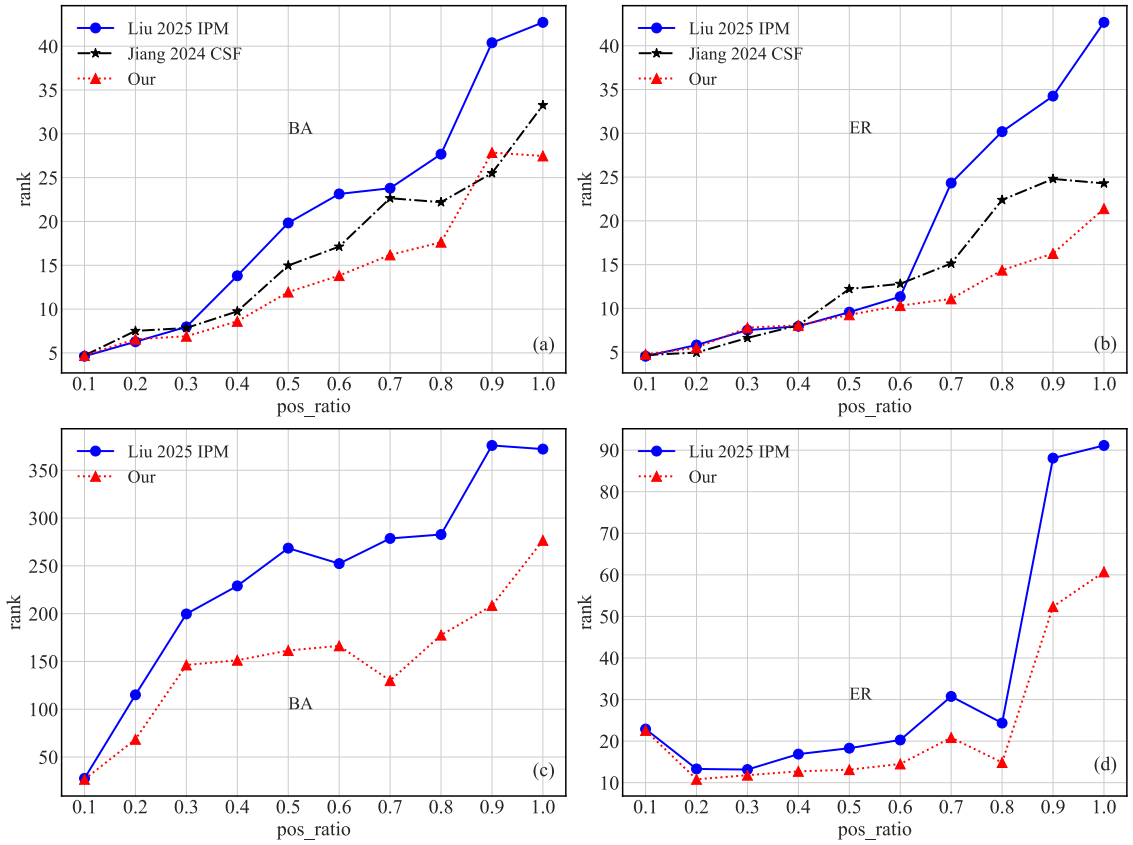

Fig. 6. Relationships between positive edge ratio pos_ratio and the rank of different algorithms. (a) and (c) are results on BA networks; (b) and (d) are results on ER networks. The network size is $N = 100$ for (a, b) and $N = 1000$ for (c, d). The average node degree $\langle k \rangle = 8$, infection rate $\lambda = 0.15$, observer proportion $p_{\text{obs}} = 0.7$, and observation time $t_{\text{obs}} = 5$. All results are averaged with 100 independent runs.

4.4. Influence of positive edge ratio

We further investigate the influence of the ratio of positive and negative edges on source localization performance, as illustrated in Fig. 6. The results reveal that the rank increases as the proportion of positive edges rises. Overall, our algorithm achieves the best performance among all compared methods. For small-scale networks (Fig. 6(a, b)), when the positive edge ratio is low, specifically below 30% for BA networks and 40% for ER networks, the three algorithms yield similar and low rank values. This indicates that higher proportions of negative edges contribute to better localization accuracy. Once the positive edge ratio exceeds the above thresholds, the rank climbs at an accelerated pace. A plausible explanation is that a high fraction of negative edges leads to a greater number of triangles formed by two or three negative edges. According to Eqs. (4), (5), (8) and (9), the propagation rates of such edges are relatively low, which confines rumors to a limited range and thus results in low rank values for all algorithms. As the proportion of negative edges decreases, more triangles composed of positive edges emerge. The corresponding propagation rates increase accordingly, allowing rumors to spread more widely and ultimately raising the difficulty of source localization.

For large-scale networks, as shown in Fig. 6(c, d), distinct trends are observed between BA and ER networks. In highly heterogeneous BA networks, the rank rises steadily from the very beginning as the positive edge ratio increases. By contrast, for homogeneous ER networks, the rank remains stable with no obvious degradation when the positive edge ratio is less than 0.6, but it surges sharply once the ratio exceeds 0.8. This demonstrates that the impact of positive and negative edge ratios on source localization varies significantly across network structures.

4.5. Influence of average node degree

Finally, we investigate the effect of average degree on source localization performance. As shown in Fig. 7, consistent with the previous experimental results, our method achieves the best performance overall. All three algorithms deliver higher localization accuracy on homogeneous ER networks than on heterogeneous BA networks. We also observe that the rank rises as the average node degree increases, which indicates that our method is better suited for sparse networks - a ubiquitous characteristic of real-world

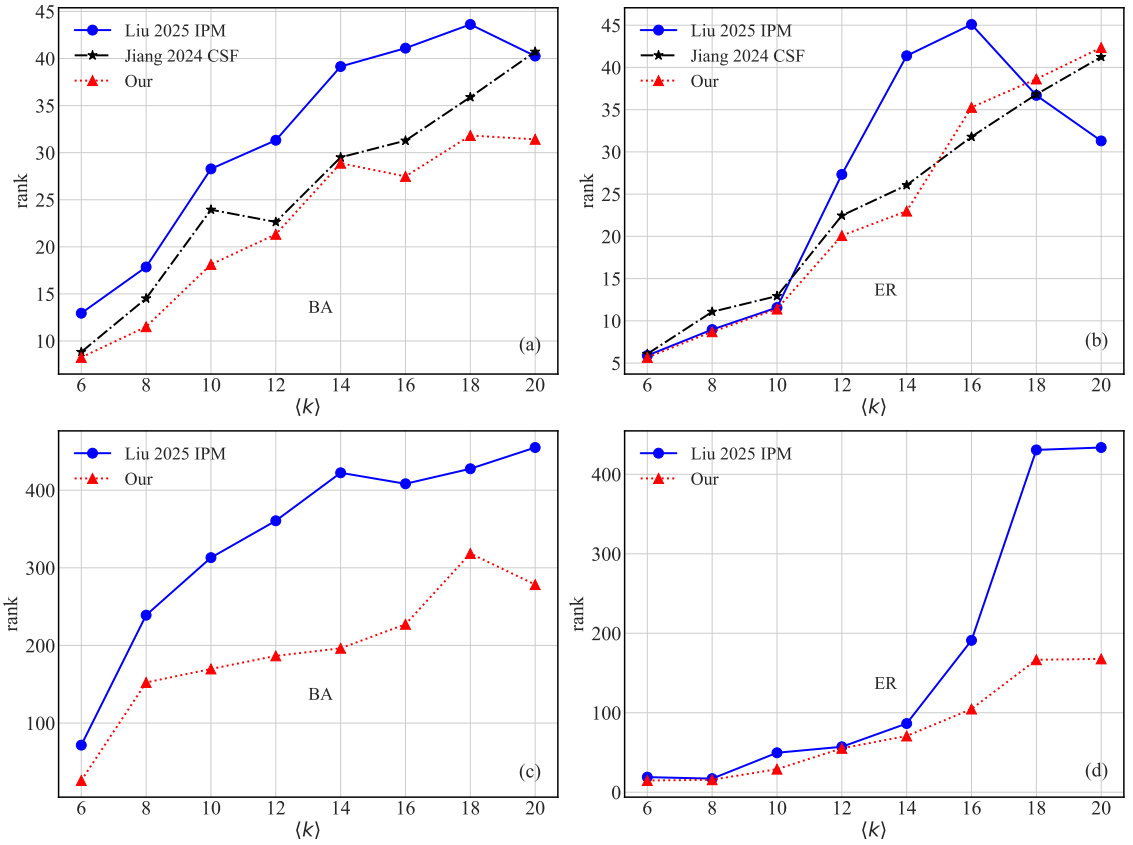

Fig. 7. Relationships between average node degree $\langle k \rangle$ and the rank of different algorithms. (a) and (c) are results on BA networks; (b) and (d) are results on ER networks. The network size is $N = 100$ for (a, b) and $N = 1000$ for (c, d). The infection rate $\lambda = 0.15$, positive edge ratio pos_ratio $= 0.5$, observer proportion $p_{\text{obs}} = 0.7$, and observation time $t_{\text{obs}} = 5$. All results are averaged with 100 independent runs.

networks. A reasonable explanation is that denser network topologies enable the diversity of propagation paths, thus increasing the difficulty of source localization. This also accounts for the similar trends observed in Figs. 6 and 7.

Based on the above analyses of influencing factors related to propagation models and network structures, our proposed source localization method achieves superior and more stable performance. One possible reason why all algorithms perform better on ER networks is that rumors propagate more extensively on BA networks under the same infection rate, making source localization more challenging. According to the SIR propagation threshold formula $\mu \frac{\langle k \rangle}{\langle k^2 \rangle}$, BA networks have a larger value of $\langle k^2 \rangle$ given the same average degree, which leads to a lower propagation threshold [61].

5. Conclusions and discussion

In this paper, we have studied rumor source localization in signed networks by combining structural balance with a hierarchical view of information propagation. We have first revised the propagation dynamics of signed networks so that positive and negative ties affect transmission in a manner that is more consistent with social relationships. We have then estimated node states through a layer-by-layer diffusion process from each candidate source. Because each node is influenced only by nodes in its parent layer and in the same layer, our method avoids the reverse infection problem that can overestimate infection probabilities in earlier propagation-based approaches. Finally, we have incorporated the inferred node-state probabilities into a maximum likelihood framework to identify the most likely rumor source from a snapshot of observed node states.

We have shown that the proposed approach achieves more accurate source ranking than the compared state-of-the-art methods while retaining favorable computational efficiency. Its advantages are particularly evident when rumor spreading is observed at an early stage, when the network is large, and when the underlying topology is sparse. These conditions are important because real social networks are often only partially observed and are typically sparse rather than densely connected. We have also shown that the method performs better on more homogeneous networks than on highly heterogeneous ones, and that its localization accuracy improves when the proportion of negative edges is higher or the infection rate is lower. These findings indicate that signed relationships are not merely additional edge attributes, but central determinants of both propagation dynamics and source identifiability.

The study also has limitations that define useful directions for future work. The current hierarchical diffusion mechanism deliberately excludes reverse infection through cyclic paths. This choice improves computational efficiency and removes a major source of overestimation, but it may also neglect some secondary paths in dense regions or around nodes embedded in many triangles. Future extensions could incorporate controlled corrections for such cyclic effects without losing the computational advantages of the present framework. We ignore the impact of group behaviors on propagation dynamics, incorporated group behaviors would render the model more consistent with real-world scenarios [62]. In addition, the method is designed for situations in which infection timestamps and propagation directions are unavailable. When such information can be reliably obtained, integrating it with the proposed hierarchical framework could further improve localization accuracy. Furthermore, our proposed source localization method that avoids overestimation of infection scale can be extended to epidemic source localization tasks, and even to source identification under the co-evolution of information and epidemic spread [63,64]. Finally, extending the method from static signed networks to time-varying signed networks is a natural next step, especially because online social relationships and information flows evolve continuously over time [65,66].

CRediT authorship contribution statement

Wei-Wei Zhu: Writing – original draft, Methodology, Investigation, Conceptualization. **Qiang Guo:** Writing – review & editing, Supervision, Methodology, Formal analysis. **Matjaž Perc:** Writing – review & editing, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant No. 72171150), and M.P. was supported by the Slovenian Research and Innovation Agency (Grant No. P1-0403).

Data availability

No data was used for the research described in the article.

References

[1] Oliveira KA, Traversa P, de Arruda G Ferraz, Moreno Y. Rumor propagation on hypergraphs. Nat Commun 2026;17(1):3253. http://dx.doi.org/10.1038/s41467-026-70096-w.
[2] Dong Y, Huo L, Perc M, Boccaletti S. Adaptive rumor propagation and activity contagion in higher-order networks. Commun Phys 2025;8(1):261. http://dx.doi.org/10.1038/s42005-025-02181-3.
[3] Devarapalli RK, Das S, Biswas A. Multiple rumor source identification in social networks leveraging community and monitor information. Inf Fusion 2024;111:102530. http://dx.doi.org/10.1016/j.inffus.2024.102530.
[4] Ma Z-W, Wang H-J, Hu Z-L, Zhu X-B, Peng H, L L-Y, Huang Y-Z, Li M. Rumor source localization in social networks based on the propagation direction of observers. Chaos: Interdiscip J Nonlinear Sci 2026;36(1):013145. http://dx.doi.org/10.1063/5.0306739.
[5] Moreno Y, Nekovee M, Pacheco AF. Dynamics of rumor spreading in complex networks. Phys Rev E 2004;69(6):066130. http://dx.doi.org/10.1103/PhysRevE.69.066130.
[6] Zhang W, Wu J, Feng M, Li Q, Perc M. Rumor propagation and supervision during confrontation: An importance-driven SIRQS network model. Chaos: Interdiscip J Nonlinear Sci 2026;36(5):053111. http://dx.doi.org/10.1063/5.0336448.
[7] Yu Z, Yan R, Wu S, Jiang M, Yang F. Dynamic analysis and immune control strategy of a rumor propagation model considering network topology. Chaos Solitons Fractals 2026;208:118287. http://dx.doi.org/10.1016/j.chaos.2026.118287.
[8] Leskovec J, Huttenlocher D, Kleinberg J. Signed networks in social media. In: Proceedings of the SIGCHI conference on human factors in computing systems. 2010, p. 1361–70. http://dx.doi.org/10.1145/1753326.1753532.
[9] Tian Y, Lambiotte R. Spreading and structural balance on signed networks. SIAM J Appl Dyn Syst 2024;23(1):50–80. http://dx.doi.org/10.1137/22M1542325.
[10] Shi G, Altafini C, Baras JS. Dynamics over signed networks. Siam Rev 2019;61(2):229–57. http://dx.doi.org/10.1137/17M1134172.
[11] Li H-J, Xu W, Song S, Wang W-X, Perc M. The dynamics of epidemic spreading on signed networks. Chaos Solitons Fractals 2021;151:111294. http://dx.doi.org/10.1016/j.chaos.2021.111294.
[12] Shah D, Zaman T. Detecting sources of computer viruses in networks: theory and experiment. In: Proceedings of the ACM sIGMETRICS international conference on measurement and modeling of computer systems. 2010, p. 203–14. http://dx.doi.org/10.1145/1811039.1811063.
[13] Zhu K, Ying L. Information source detection in the sir model: A sample-path-based approach. IEEE/ACM Trans Netw 2016;24(1):408–21. http://dx.doi.org/10.1109/TNET.2014.2364972.
[14] Pinto PC, Thiran P, Vetterli M. Locating the source of diffusion in large-scale networks. Phys Rev Lett 2012;109(6):068702. http://dx.doi.org/10.1103/PhysRevLett.109.068702.
[15] Hu Z-L, Wang L, Tang C-B. Locating the source node of diffusion process in cyber–physical networks via minimum observers. Chaos: Interdiscip J Nonlinear Sci 2019;29(6):063117. http://dx.doi.org/10.1063/1.5092772.
[16] Yang F, Li C, Peng Y, Liu J, Yao Y, Wen J, Yang S. Locating the propagation source in complex networks with observers-based similarity measures and direction-induced search. Soft Comput 2023;27(21):16059–85. http://dx.doi.org/10.1007/s00500-023-08000-7.
[17] Lokhov AY, Mézard M, Ohta H, Zdeborová L. Inferring the origin of an epidemic with a dynamic message-passing algorithm. Phys Rev E 2014;90(1):012801. http://dx.doi.org/10.1103/PhysRevE.90.012801.

[18] Ke Q, Masuda N, Jin Z, Liu C, Zhan X-X. Source detection in epidemic dynamics on hypergraphs using a dynamic message passing algorithm. Chaos: Interdiscip J Nonlinear Sci 2026;36(1):013121. http://dx.doi.org/10.1063/5.0311821.
[19] Altarelli F, Braunstein A, DallAsta L, Lage-Castellanos A, Zecchina R. Bayesian inference of epidemics on networks via belief propagation. Phys Rev Lett 2014;112(11):118701. http://dx.doi.org/10.1103/PhysRevLett.112.118701.
[20] Cheng L, Zhu P, Tang K, Gao C, Wang Z. GIN-SD: source detection in graphs with incomplete nodes via positional encoding and attentive fusion. In: Proceedings of the AAAI conference on artificial intelligence, Vol. 38, 2024, p. 55–63. http://dx.doi.org/10.1609/aaai.v38i1.27755.
[21] Ali SS, Rastogi A, Anwar T, Rizvi SAM, Yang J, Wu J, Sheng QZ. Generalized local prominence for source detection in real-world rumor networks. IEEE Trans Knowl Data Eng 2025;37(8):4620–34. http://dx.doi.org/10.1109/TKDE.2025.3567282.
[22] Gopal GN, Kovoor BC. Motives and spread pattern in focus for rumor source detection using class-balanced embedded mlp. IEEE Trans Comput Soc Syst 2026;13(1):510–29. http://dx.doi.org/10.1109/TCSS.2025.3591846.
[23] Wang Y, Chai Q, Lin L, Jia T. PDSL: Propagation dynamics aware framework for source localization. IEEE Trans Netw Sci Eng 2026;13:9027–44. http://dx.doi.org/10.1109/TNSE.2026.3688551.
[24] Liu W, Shao Y, Chen Y, Chen L. False information sources detecting based on an epidemic diffusion model. Inf Process Manage 2025;62(5):104197. http://dx.doi.org/10.1016/j.ipm.2025.104197.
[25] Ma Z-W, Sun L, Ding Z-G, Huang Y-Z, Hu Z-L. Source localization in signed networks with effective distance. Chin Phys B 2024;33:028902. http://dx.doi.org/10.1088/1674-1056/ad1482.
[26] Ma Z-W, j. Wang H, Hu Z-L, Zhu X-B, Huang Y-Z, Huang F. DISLPSI: A framework for source localization in signed social networks with structural balance. Phys Lett A 2024;523:129772. http://dx.doi.org/10.1016/j.physleta.2024.129772.
[27] Jiang Z-X, Hu Z-L, Huang F. Source localization in signed networks based on dynamic message passing algorithm. Chaos Solitons Fractals 2024;188:115532. http://dx.doi.org/10.1016/j.chaos.2024.115532.
[28] Iacopini I, Petri G, Barrat A, Latora V. Simplicial models of social contagion. Nat Commun 2019;10(1):2485. http://dx.doi.org/10.1038/s41467-019-10431-6.
[29] Fan J, Zhao D, Xia C, Tanimoto J. Coupled spreading between information and epidemics on multiplex networks with simplicial complexes. Chaos: Interdiscip J Nonlinear Sci 2022;32(11):113115. http://dx.doi.org/10.1063/5.0125873.
[30] Zhang S, Zhao D, Xia C, Tanimoto J. Impact of simplicial complexes on epidemic spreading in partially mapping activity-driven multiplex networks. Chaos: Interdiscip J Nonlinear Sci 2023;33(6):063128. http://dx.doi.org/10.1063/5.0151881.
[31] Lucas M, Iacopini I, Robiglio T, Barrat A, Petri G. Simplicially driven simple contagion. Phys Rev Res 2023;5(1):013201. http://dx.doi.org/10.1103/PhysRevResearch.5.013201.
[32] Malizia F, Gallo L, Frasca M, Kiss IZ, Latora V, Russo G. A pair-based approximation for simplicial contagion. Chaos Solitons Fractals 2025;199:116776. http://dx.doi.org/10.1016/j.chaos.2025.116776.
[33] Sattarov O, Choi J. Detection of rumors and their sources in social networks: A comprehensive survey. IEEE Trans Big Data 2024;11(3):1528–47. http://dx.doi.org/10.1109/TBDATA.2024.3522801.
[34] Wang H-J, Zhang F-F, Sun K-J. An algorithm for locating propagation source in complex networks. Phys Lett A 2021;393:127184. http://dx.doi.org/10.1016/j.physleta.2021.127184.
[35] Wang J, Jiang J, Zhao L. An invertible graph diffusion neural network for source localization. In: Proceedings of the ACM web conference 2022. ACM; 2022, p. 1058–69. http://dx.doi.org/10.1145/3485447.3512155.
[36] Xu X, Qian T, Xiao Z, Zhang N, Wu J, Zhou F. PGSL: A probabilistic graph diffusion model for source localization. Expert Syst Appl 2024;238:122028. http://dx.doi.org/10.1016/j.eswa.2023.122028.
[37] Bao Q, Jiang Y, Zhang W, Jiao P, Su J. Graph contrastive learning for source localization in social networks. Inform Sci 2024;679:121090. http://dx.doi.org/10.1016/j.ins.2024.121090.
[38] Luo W, Tay WP, Leng M. How to identify an infection source with limited observations. IEEE J Sel Top Signal Process 2014;8(4):586–97. http://dx.doi.org/10.1109/JSTSP.2014.2315533.
[39] Chen Z, Zhu K, Ying L. Detecting multiple information sources in networks under the sir model. IEEE Trans Netw Sci Eng 2016;3(1):17–31. http://dx.doi.org/10.1109/TNSE.2016.2523804.
[40] Jiang J, Wen S, Yu S, Xiang Y, Zhou W. K-center: An approach on the multi-source identification of information diffusion. IEEE Trans Inf Forensics Secur 2015;10(12):2616–26. http://dx.doi.org/10.1109/TIFS.2015.2469256.
[41] Wang Z, Wang C, Pei J, Ye X. Multiple source detection without knowing the underlying propagation model. In: Proceedings of the AAAI conference on artificial intelligence, Vol. 31, 2017, http://dx.doi.org/10.1609/aaai.v31i1.10477.
[42] Hu Z-L, Wang H-J, Sun L, Tang C-B, Li M. Source localization in complex networks with optimal observers based on maximum entropy sampling. Expert Syst Appl 2024;256:124946. http://dx.doi.org/10.1016/j.eswa.2024.124946.
[43] Zhu P, Cheng L, Gao C, Wang Z, Li X. Locating multi-sources in social networks with a low infection rate. IEEE Trans Netw Sci Eng 2022;9(3):1853–65. http://dx.doi.org/10.1109/TNSE.2022.3153968.
[44] Peng S-L, Wang H-J, Peng H, Zhu X-B, Li X, Han J, Zhao D, Hu Z-L. NLSI: An innovative method to locate epidemic sources on the seir propagation model. Chaos: Interdiscip J Nonlinear Sci 2023;33(8):083125. http://dx.doi.org/10.1063/5.0152859.
[45] Hu Z-L, Jin Q, Sun L, Peng S. Source identification on financial networks with label propagation. Phys A 2025;659:130328. http://dx.doi.org/10.1016/j.physa.2024.130328.
[46] Ali SS, Rastogi A, Anwar T. Frost: Controlled label propagation for multisource detection. IEEE Trans Comput Soc Syst 2024;11(5):6217–28. http://dx.doi.org/10.1109/TCSS.2024.3390931.
[47] Cheng L, Li X, Han Z, Luo T, Ma L, Zhu P. Path-based multi-sources localization in multiplex networks. Chaos Solitons Fractals 2022;159:112139. http://dx.doi.org/10.1016/j.chaos.2022.112139.
[48] Yu X, Nie Y, Li W, Luo G, Lin T, Wang W. Source inference for misinformation spreading on hypergraphs. Chaos Solitons Fractals 2024;187:115457. http://dx.doi.org/10.1016/j.chaos.2024.115457.
[49] Cheng L, Zhu P, Tang K, Gao C, Wang Z. Efficient source detection in incomplete networks via sensor deployment and source approaching. IEEE Trans Inf Forensics Secur 2025;20:3705–16. http://dx.doi.org/10.1109/tifs.2025.3550069.
[50] Paluch R, Lu X, Suchecki K, B.K. Szymański JAHoł yst. Fast and accurate detection of spread source in large complex networks. Sci Rep 2018;8(1):2508. http://dx.doi.org/10.1038/s41598-018-20546-3.
[51] Spinelli B, Celis LE, Thiran P. A general framework for sensor placement in source localization. IEEE Trans Netw Sci Eng 2019;6(2):86–102. http://dx.doi.org/10.1109/TNSE.2017.2787551.
[52] Xu S, Teng C, Zhou Y, Peng J, Zhang Y, Zhang Z-K. Identifying the diffusion source in complex networks with limited observers. Phys A 2019;527:121267. http://dx.doi.org/10.1016/j.physa.2019.121267.
[53] Wang H-J, Hu Z-L, Tao L, Shao S, Wang S-Z. The locatability of pearson algorithm for multi-source location in complex networks. Sci Rep 2023;13(1):5692. http://dx.doi.org/10.1038/s41598-023-32832-w.
[54] Hou D, Gao C, Wang Z, Li X, Li X. Random full-order-coverage based rapid source localization with limited observations for large-scale networks. IEEE Trans Netw Sci Eng 2024;11(5):4213–26. http://dx.doi.org/10.1109/TNSE.2024.3406394.

[55] Devarapalli RK, Biswas A. Estimating rumor source in social networks using incomplete observer information. Expert Syst Appl 2024;249:123499. http://dx.doi.org/10.1016/j.eswa.2024.123499.
[56] Fu L, Shen Z, Wang W-X, Fan Y, Di Z. Multi-source localization on complex networks with limited observers. Europhys Lett 2016;113(1):18006. http://dx.doi.org/10.1209/0295-5075/113/18006.
[57] Hu Z-L, Shen Z, Tang C-B, Xie B-B, Lu J-F. Localization of diffusion sources in complex networks with sparse observations. Phys Lett A 2018;382(14):931–7. http://dx.doi.org/10.1016/j.physleta.2018.01.037.
[58] Yuan S, Liu W, Zeng H, Wang C. Multi-source localization on complex networks based on community detection. Europhys Lett 2023;141(6):60001. http://dx.doi.org/10.1209/0295-5075/acbfd7.
[59] Barabási A-L, Albert R. Emergence of scaling in random networks. Science 1999;286(5439):509–12. http://dx.doi.org/10.1126/science.286.5439.509.
[60] Erdős P, Rényi A, et al. On the evolution of random graphs. Publ Math Inst Hung Acad Sci 1960;5(1):17–60, https://pages.cs.wisc.edu/ cs809-1/ErdosRenyi.pdf.
[61] Pastor-Satorras R, Castellano C, Mieghem P Van, Vespignani A. Epidemic processes in complex networks. Rev Modern Phys 2015;87(3):925–79. http://dx.doi.org/10.1103/RevModPhys.87.925.
[62] Zhang J, Zhai S, Linghu H, Ma J. Neuro-driven cluster synchronization in the dynamics of collective human behavior. Communicat Nonlinear Science Numerical Simulation 2025;109226. http://dx.doi.org/10.1016/j.cnsns.2025.109226.
[63] Zhai S, Zhang Z, Ma J. Convergence of individual opinions during a bi-virus epidemic and its impact. Math Comput Simulation 2025;236:334–53. http://dx.doi.org/10.1016/j.matcom.2025.04.006.
[64] Hu Z-L, Peng S-L, Lü L, Zheng Z, Tang C-B, He X-L, Huang Y-Z, Li W. Motif-based epidemic source localization in information-epidemic multiplex networks. Phys A 2026;698:131748. http://dx.doi.org/10.1016/j.physa.2026.131748.
[65] Hu Z-L, Shen Z, Cao S, Podobnik B, Yang H, Wang W-X, Lai Y-C. Locating multiple diffusion sources in time varying networks from sparse observations. Sci Rep 2018;8:2685. http://dx.doi.org/10.1038/s41598-018-20033-9.
[66] Chai Y, Wang Y, Zhu L. Information sources estimation in time-varying networks. IEEE Trans Inf Forensics Secur 2021;16:2621–36. http://dx.doi.org/10.1109/TIFS.2021.3050604.