Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 20(6): 11260–11280.
DOI: 10.3934/mbe.2023499
Received: 23 November 2022
Revised: 18 March 2023
Accepted: 03 April 2023
Published: 25 April 2023

Research article

A hot topic diffusion approach based on the independent cascade model and trending search lists in online social networks

Yuqi Chen, Xianyong Li*, **Weikai Zhou, Yajun Du, Yongquan Fan, Dong Huang and Xiaoliang Chen**

School of Computer and Software Engineering, Xihua University, Chengdu 610039, China

* **Correspondence:** Email: lixy@mail.xhu.edu.cn, xian-yong@163.com; Tel: +8602887726572; Fax: +8602887726572.

Abstract: In online social networks, users can quickly get hot topic information from trending search lists where publishers and participants may not have neighbor relationships. This paper aims to predict the diffusion trend of a hot topic in networks. For this purpose, this paper first proposes user diffusion willingness, doubt degree, topic contribution, topic popularity and the number of new users. Then, it proposes a hot topic diffusion approach based on the independent cascade (IC) model and trending search lists, named the ICTSL model. The experimental results on three hot topics show that the predictive results of the proposed ICTSL model are consistent with the actual topic data to a great extent. Compared with the IC, independent cascade with propagation background (ICPB), competitive complementary independent cascade diffusion (CCIC) and second-order IC models, the Mean Square Error of the proposed ICTSL model is decreased by approximately 0.78%–3.71% on three real topics.

Keywords: hot topic; trending search list; topic diffusion; independent cascade (IC) model; topic popularity; social networks

1. Introduction

Information exchange is one of people's most critical social activities in the real world. With the rapid development of Internet technology and the diversified choices of online social network platforms, a piece of information can be easily and quickly spread from one user to other users. Netizens can easily publish new events in the real world through social network platforms. In particular, public emergencies may be widely and quickly propagated on online social networks, finally forming hot topics. Therefore, diffusion processes of hot topics on social networks are worth discussing.

Recently, there have been many information diffusion models [1–3] for social networks. Considering user mobility in online social networks, Wang et al. [4] classified users and deduced the information diffusion system with the non-uniform mean field method. The final experimental results were better than the traditional SIR (susceptible-infected-recovered) model. For rumor propagation, Xiao et al. [5] quantified the influence of counter rumors on rumor propagation. They also studied the evolutionary trend of a complete rumor life cycle by analyzing the group behavior at different periods. Inspired by current image restoration techniques, Xiao et al. [6] pioneered the pixelation of diffusion networks by considering implicit or potential factors and information confrontation among users and proposed a rumor-anti-rumor game diffusion model. Considering the interaction among rumors, counter-rumors and inflammatory speeches, Li et al. [7] proposed a three-party cognitive game dissemination model based on disseminators, receivers and defenders. Meanwhile, they proposed some practical strategies to reduce the impact of rumor propagation. Jin et al. [8] judged key nodes by the number of direct forwarding user posts. They superimposed the influence of these key nodes, obtaining more accurate results. Sanjay et al. [9] found that previous research overestimated infected users. They defined the followers of infected users as exposed users. They further proposed the SEI (susceptible-exposed-infected) model. Using their model, the prediction of infected users was more accurate. Zhang et al. [10] divided infected nodes into structural hole infected nodes, high influential infected nodes and ordinary infected nodes. Then they provided an SEI^{3}R information propagation model. Experiments showed that the performance of the SEI^{3}R model is superior to other baseline models. It can be seen that these compartment models are also important research points for social networks.

Scholars also proposed many influence spread models for social networks. The optimization problem of selecting the most influential nodes is an NP-hard problem. Kempe et al. [11] adopted an analytical framework based on submodular functions to solve this problem effectively. They proved that the solution obtained by a natural greedy strategy is the optimal solution within 63%. By a binary multi-objective optimization paradigm, Olivares et al. [12] modeled the impact maximization problem (IMP) and the least cost influence (LCI) to find an effective solution. Inspired by the efficient evolutionary mechanism of swarm intelligence, Tang et al. [13] proposed an effective discrete shuffled frog-leaping algorithm (DSFLA). Experiments showed that this model is more efficient and effective than baseline models in solving the influence maximization problem. Considering that the information diffusion in existing networks is carried out asynchronously, Saito et al. [14] studied the problem of estimating the parameters for a continuous time delay independent cascade (CTIC) model. Their method better predicted the ranking of influential nodes. It also proved that the propagation speeds of different topics are different. Alasadi et al. [15] proposed a diffusion model based on the shared friend-aware independent cascade (IC) model. Their model improved the cascade of information propagation compared with the IC model. Considering that a node may have multiple active edges due to numerous active neighbor nodes, but it can only be activated by one of them to become an active node, Tang et al. [16] divided the edges between nodes into out-edges and in-edges and proposed a second-order IC model. Their model outperformed the IC model in simulating influence diffusion. Therefore, the improved IC models based on network characteristics perform better than the original information spread models.

To solve the problem of information overload in social networks, Li et al. [17] proposed an ICPB model, which depicted the users' real-time attention to simulate the propagation background. Its performance exceeded the IC model. Sexena et al. [18] assigned neutral, false and correct states to each node and updated them during the information propagation process. Then, they proposed the competitive independent cascade model with user bias (CICMB) which is superior in any case in the latest method. Considering the relationship between communicators and the competitive or complementary relationship with information dissemination, Huang et al. [19] proposed a competitive complementary independent cascade diffusion (CCIC) model. The prediction accuracy of this model is superior to existing methods. These models are all based on the independent cascade model, striving to expand online social networks' communication characteristics and the influence of the communication environment, background, nodes and edges on communication. However, these studies did not consider the impact of popular search lists on information dissemination in social networks.

Furthermore, some scholars have studied the influence of the life cycle of a topic on information diffusion in social networks. The propagation process of a hot topic can be regarded as a life cycle, including the networks' incubation, development, explosion, fading and decay. Zhang et al. [20] combined the tweets of the Hurricane Irma event with the life cycle of the evolution of network public opinion and found that the emotion in a tweet and the life cycle of the evolution of network public opinion had an interactive impact on the target users. Through investigation, Wei et al. [21] found a relationship between the sentiment of posts on social networks and users' information-sharing behavior. They also found a prominent role in the initial and explosive stages of public opinion evolution. To monitor the influence of hot topics, He et al. [22] analyzed the life cycle of hot topics on Weibo. They obtained four characteristics of hot topics: short cycle, rapid growth, rapid decline and short duration.

For various social networks, such as Weibo, Toutiao and Twitter, their trending search lists enumerate the most popular events and topics for a period in real-time, quickly sharing the information with other persons. For some hot topics, nodes could get information from their neighbors and trending search lists on social networks. Through label classifications and hot topics, users conveniently communicate, self-organize and effectively promote users' responses in social networks [23]. Thus, the trending search lists significantly impact users' social communication, interpersonal communication and study life.

To the best of our knowledge, there are few results on information diffusion based on trending search lists. This paper aims to establish a hot topic diffusion approach based on the IC model and trending search lists (ICTSL model). For this purpose, the user diffusion willingness, doubt degree, topic contribution, topic popularity and the number of new users are defined. Based on these proposed terms and the IC model, this paper proposes the ICTSL model to simulate information diffusion in social networks. Some experiments will verify the validity of predicting topics in Weibo networks.

This paper is organized as follows. Section 2 detailedly establishes a hot topic diffusion approach based on trending search lists to predict the evolution trend of the topic in social networks. This section introduces some terms, including user diffusion willingness, doubt degree, topic contribution, topic popularity and the number of new users. Then, a hot topic diffusion approach based on trending search lists is proposed. Section 3 shows that simulation and empirical analysis confirm the proposed approach's effectiveness in predicting the topic's evolution trend in social networks. Finally, the summary of this paper and future work are given in Section 4.

2. The hot topic diffusion approach based on IC model and trending search lists

This section will propose user diffusion willingness, doubt degree, topic contribution, topic popularity and the number of new users. Then, it establishes a hot topic diffusion approach based on the IC model and trending search lists.

Table 1. Distribution of users of different ages in the Internet sector.

Age	Social media	News	Shopping	Game	Music	Videos	Others
15–24	17.3%	13.5%	10.5%	10.8%	8.9%	16.1%	22.7%
25–34	17.0%	14.5%	11.6%	7.9%	7.4%	14.9%	26.7%
35–44	19.7%	16.6%	9.7%	6.9%	6.6%	15.0%	25.3%
45–54	23.4%	19.9%	6.5%	6.6%	4.9%	15.7%	22.9%
≥55	25.3%	23.5%	3.8%	7.6%	4.0%	15.4%	20.3%

2.1. User's diffusion willingness

According to the primary research data of 12 cities researched by CSM Media Research *, it is found that users of different ages have different abilities to accept information due to factors such as social status, psychology, personality and educational background (see Table 1).

In a real network, considering that users of different ages accept and propagate a piece of information with different abilities, we define the average receptivity of the user to the topic, denoted by a_0, namely, user diffusion willingness, as the following:

$$a_0 = \frac{\sum_{i=1}^{n} \frac{|age_i - a|}{age_i + a}}{n}, \tag{2.1}$$

where a is the age factor, and age_i represents the age of the user i, which means that users with age a are more likely to spread information on the topic.

In social networks, the user's influence is the degree that essential users affect other users' opinions [6, 24]. Users can quickly get information from their followers and are also influenced by them. The more followers a user has, the better his ability to get information is. The more influential the users in social networks that propagate information are, the more the topic can attract more users to propagate. In this paper, we simulate this ability as the speed at which the user's willingness to spread over time decreases, named b_0.

$$b_0 = \frac{N_{\text{leader}}}{N_{\text{all}}}, \tag{2.2}$$

where N_{leader} denotes the number of nodes with a high amount of disseminated information, and N_{all} denotes the total number of nodes in the social network. Considering the temporal features [25], as a user decreases their interest in some information, the user diffusion willingness declines over time. So, the user's diffusion willingness can be defined as follows.

$$Y(t) = -\gamma_1 a_0 t + \gamma_2 b_0, \tag{2.3}$$

where $\gamma_1, \gamma_2 > 0$ and are the weights of the user's willingness to diffuse a piece of information, which is influenced by the user's age factor and influence power. t denotes time, and the user diffusion willingness decreases with time. Different age users accept information differently; each topic corresponds to the most receptive age. The more significant the difference between a user's age and the most receptive age, the less likely the user is to spread information. So, the parameter γ_1 is negative in Eq (2.3).

*https://www.csm.com.cn/content/2016/11-11/1054081033.html

Table 2. Topic contribution rules for a topic in a Weibo network.

Rule	Users' contribution values
Posting original microblogs with the topic	five points each time
Forwarding related microblogs with the topic	one point each time
Participating in the topic and being forwarded by others	one point for every ten times

Furthermore, the higher the degree of a high influence node is, the easier it is to spread and maintain high returns at every moment. Thus, the parameter γ_2 is positive in Eq (2.3).

2.2. Doubt degree

Based on confirmation bias in sociology [26], users who face social life and social events in social networks are more willing to think independently. They always doubt new events or topics. For a new event or topic, a user j will have a doubt degree when examining the information from another user i, denoted by D_{ij}. Due to the complexity of human psychology and the arbitrariness of the relationship between users attracted based on trending search topics, the doubt degree can be regarded as a certain probability, i.e., $D_{ij} \in [0, 1]$. The more significant D_{ij} is, the more doubtful the user j is of the user i, and the greater the diffusion willingness of the user j is for the topic, and vice versa. Then, the state of the node j at time t, denoted by $S_j(t)$, can be defined as follows.

$$S_j(t) = \begin{cases} 0, Y(t) < D_{ij}, \\ 1, Y(t) \geq D_{ij}, \end{cases} \tag{2.4}$$

where $S_j(t) = 1$ indicates that the node j is activated, and $S_j(t) = 0$ indicates that the node j is not activated at time t. It is possible to predict whether potential users will doubt the active users the next time and to predict the next propagation trend.

2.3. Topic contribution

Based on the hot search management rules[†] for Weibo networks, the topic contribution value of Weibo users for a topic is determined by all users' contribution values on posting original microblogs with the topic, forwarding related microblogs with the topic, and participating in the topic and being forwarded by others. Table 2 shows the topic contribution rules for a topic in a Weibo network.

For a topic in a Weibo network, let $O(t)$, $F(t)$ and $PF(t)$ denote the increments of posting original microblogs with the topic, forwarding related microblogs with the topic, and participating in the topic and being forwarded by others from the time $t-1$ to t, respectively. Then, the topic contribution value of Weibo users for the topic at the time t, denoted by $C(t)$, is as follows.

$$C(t) = (\alpha_1 C(t-1) + \alpha_2 F(t) + \alpha_3 PF(t)) * \alpha_4 O(t), \tag{2.5}$$

where α_1, α_2, α_3 and α_4 are the weights of $C(t-1)$, $F(t)$, $PF(t)$ and $O(t)$, respectively.

[†]https://weibo.com/1934183965/KuKyPkp8Y

2.4. Topic popularity

Newton's cooling law [27] shows that when the temperature of an object is higher than its surroundings, the object transfers its heat to the surroundings, gradually cooling down. This law stipulates that the temperature difference between the cooling object and the surroundings decreases exponentially with time. Then, the object's surface temperature at time t, denoted by $T(t)$, is as follows.

$$T(t) = (T(t-1) - T_a(t))\exp(-\lambda\Delta t) + T_a(t), \tag{2.6}$$

where $T_a(t)$ represents the surroundings' temperature of the object at time t.

Since the natural decline process of topic popularity is similar to an object's temperature cooling process, the topic popularity at time t, denoted by $M(t)$, can be defined as follows.

$$M(t) = (M(t-1) - M_a(t))\exp(-\partial\Delta t) + M_a(t), \tag{2.7}$$

where $M_a(t)$ is constant, representing the surrounding popularity at time t in social networks. ∂ is a topic hot cooling coefficient. Considering that the topic popularity will drop to 0 at the final time, we assume $M_a = 0$ for simplicity. Then, the topic popularity at time t is simplified as follows.

$$M(t) = M(t-1)\exp(-\partial\Delta t). \tag{2.8}$$

Meanwhile, the topic popularity is influenced by the topic contribution value of network users for a topic at time t. Therefore, we further revise the topic popularity as follows.

$$M(t) = M(t-1)\exp(-\partial\Delta t) + C(t). \tag{2.9}$$

2.5. Number of new users

In a Weibo network, the users who have followed a topic will influence topic popularity. When topic popularity increases, new users will diffuse the topic. Since a social network has scale-free features, the number of new users obeys a power-law distribution. In the early diffusion stage, when an initial spreader releases a topic, topic popularity is always shallow. As the topic continuously spreads on the social network by different spreaders, the topic popularity increases with the topic hot. Considering the influence of topic popularity on the new users, the number of new users at time t, denoted by $P(t)$, is the following.

$$P(t) = \begin{cases} c(M(t)-\varepsilon)^{\alpha}, M(t) > \varepsilon, \\ 0, M(t) < \varepsilon, \end{cases} \tag{2.10}$$

where $c > 0$, $1 < \alpha < 2$ and ε is a topic popularity threshold. If topic popularity reaches ε, new users will join the social network. Based on the new user's distribution, most topics have low topic popularity and can only attract a small number of users. In contrast, a few topics with great topic popularity can attract many users. The reason is that when a topic appears in the trending search list of a social network, many users who have no contact with its spreaders will follow it in the social network. When the topic popularity reaches a certain value, the number of users will increase sharply.

In the following hot topic diffusion approach, all nodes have two states in a social network: activated and inactivated. The active nodes try to influence their inactive neighbors in a social network.

However, isolated nodes are excluded under the IC model. The reason is that these nodes are never involved in the information diffusion process of a topic, which would underestimate its propagation breadth. Therefore, the users behavior of the social network cannot be represented by the IC model. So we expand the IC model, add five parts, including user diffusion willingness, doubt degree, topic contribution, topic popularity and the number of new users, and formalize a trending topic in the hot search list as a trending topic node in the following hot topic diffusion algorithm.

2.6. Description of a hot topic diffusion approach based on trending search lists

A social network can be denoted by a graph $G = (V, E)$, where V stands for the users, and E represents the relationships between users. We assume that each edge between the related nodes v_i and v_j has a propagation probability $Y(t)$. In trending search lists, a trending topic is regarded as a TT node. Its topic popularity $M(t)$ is the contribution value of the network users for the topic at time t. At time t, let the set of nodes outside the social network denote by $W(t)$; let $U(t)$ stand for the set of the influential activated nodes. Considering the impact of a single topic on users, the hot topic diffusion approach based on the IC model and the trending search lists (ICTSL model) is shown in Figure 1. At the initial time t_0, initial propagation nodes are activated and stored in $U(t_0)$. All nodes of $W(t)$ are connected with the TT node at time t. Meanwhile, each node of $U(t)$ activates neighbor nodes with a probability $Y(t)$. Then, all activated nodes are stored in $U(t + 1)$. These activated nodes are influential activated nodes at time $t + 1$. Repeat this step until the value of $M(t)$ reaches the topic popularity threshold ε. At this moment, $P(t)$ new users from $W(t)$ will join the social network, which are activated by the TT node with the probability $Y(t)$. After the activated nodes tend to peak, the value of $M(t)$ decreases as topic popularity increases. When topic popularity is stable, the topic diffusion ends in the social network. The hot topic diffusion approach can be represented as the ICTSL algorithm (see Algorithm 1).

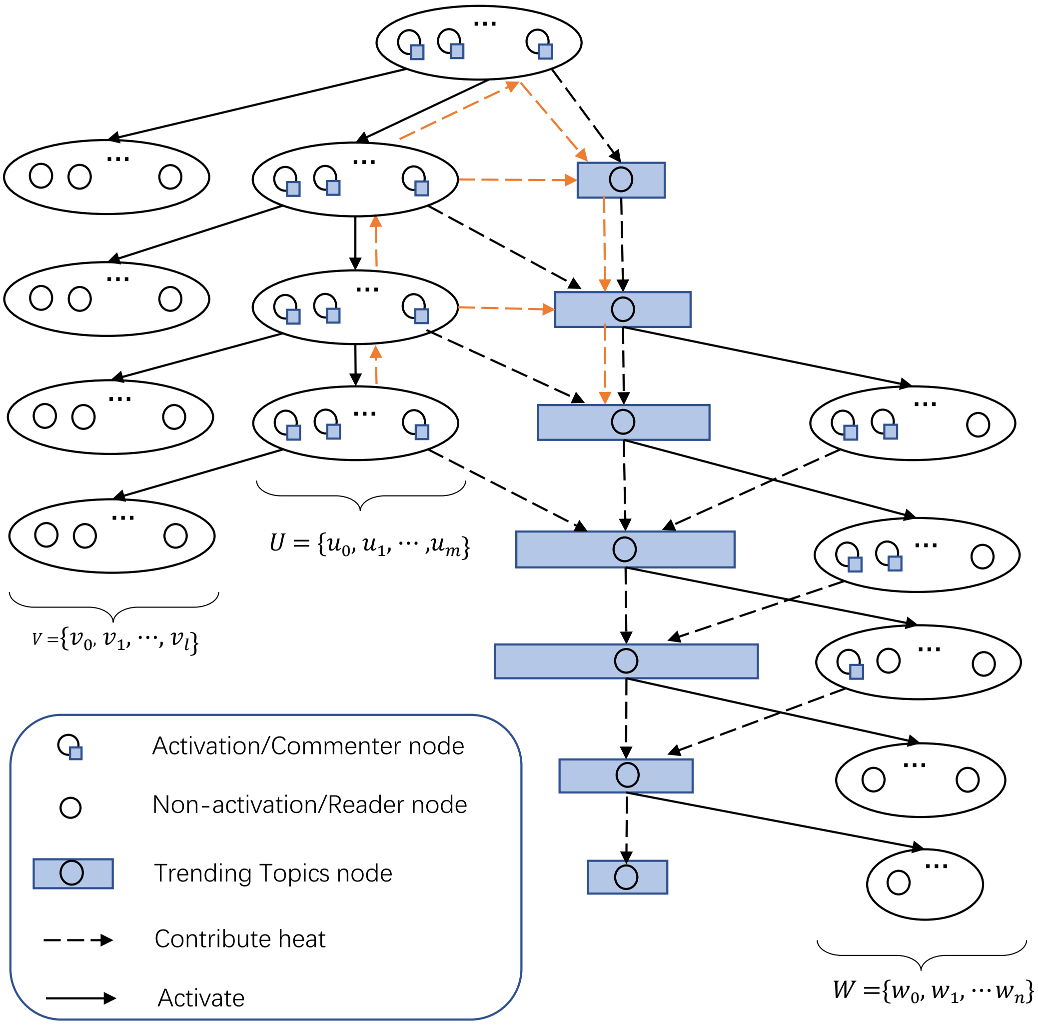


Figure 1. The frame of the ICTSL model.

Algorithm 1: ICTSL algorithm.

Input : The social network $G = (V, E)$ with a topic; The activated nodes set $U(t_0)$; The number of activated nodes array L; The popularity threshold ε; The TT node with a topic popularity $M(t_0) = 1$; The set of nodes outside the network G, i.e., $W(t_0)$; the doubt degree D_{ij} for the nodes i and j.

Output: The number of activated nodes; the complete life cycle of the hot topic.

1 Compute the diffusion willingness $Y(t)$ by Eq (2.3), the topic contribution $C(t)$ by Eq (2.5), and the topic popularity $M(t)$ by Eq (2.9), at time t;
2 **if** $M(t) < \varepsilon$ **then**
3   Visit the nodes u_i in the set $U(t)$ and all their neighbors;
4   **if** u_i*.neighbor is not activated* **then**
5   **if** $Y(t) \geq D_{ij}$ **then**
6   u_i.neighbor is activated;
7   **else**
8   u_i.neighbor is not be activated;
9   **end**
10   **else**
11   Visit the next neighbor node of u_i;
12   **end**
13   Use the new nodes to replace the ones in the set $U(t)$;
14   Compute the number of new nodes at time t and store them in L;
15     $t = t + 1$;
16 **else**
17   Compute the number of new nodes $P(t)$ at time t and store them in $W(t)$;
18   Get the doubt degree of each node w_i with the TT node, denoted by D_{ti};
19   **if** $Y(t) \geq D_{ti}$ **then**
20   w_i is activated;
21   **else**
22   w_i is not be activated;
23   **end**
24   Continue to spread information according to steps 4–16 in G;
25   Use the new nodes to replace the ones in $U(t)$;
26   Compute the number of new nodes at this time and store them in L;
27     $t = t + 1$;
28   Get L and complete the life cycle of the hot topic when $M(t)$ is stable.
29 **end**

Table 3. Three topics in Weibo networks.

Topics	Type	Opinion leaders	Number of new users	Propagation time
Topic 1	Music	1938	931779	16h
Topic 2	News	127	4640	13h
Topic 3	Sports	2829	21726	11h

3. Experiments

To analyze the hot topic diffusion process in social networks, we distinguish the life cycle of a hot topic into six stages, including the initial stage, change and development stage, mature development stage, degrading stage, buffer stage and fading stage (see Figure 2). The following simulation experiments will analyze the influence of the user diffusion willingness and different network scales on the ICTSL model. Compared with the IC, ICPB, CCIC and second-order IC models on the data of three actual topics, experiments will further confirm that the ICTSL model better predicts the topics' diffusion trends. Meanwhile, the mean square error (MSE) of the ICTSL model is significantly lower than these baseline models, implying that the accuracy of the ICTSL model is the best.

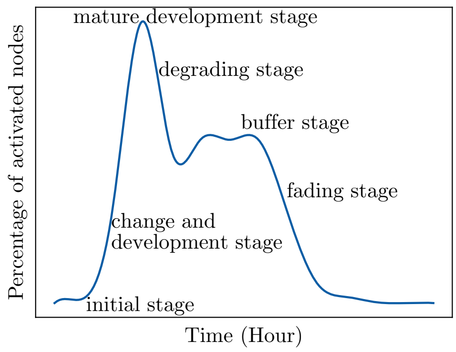


Figure 2. The life cycle of a hot topic.

3.1. Datasets

To verify the accuracy of predictive results under the proposed ICTSL model in social networks, we conduct experiments on three different topics of actual datasets (see Table 3) from Weibo networks.

Topic 1: "Together for a Shared Future" is a promotional song for the theme slogan of the Beijing 2022 Winter Olympics, which is a music topic. There were 1938 opinion leaders on this topic, which attracted 931779 new users after 16 hours of diffusion.

Topic 2: "U.S. Embassy in China apologizes" is a news topic. On this topic, 127 opinion leaders attracted 4640 new users to participate in the topic through 13 hours of diffusion.

Topic 3: "National Games Table Tennis" belongs to the sports category. On this topic, 2829 opinion leaders attracted 21726 new users to participate in the topic through 11 hours of diffusion.

3.2. Baselines

To verify the ICTSL model's role in promoting topic dissemination, we adopt four baselines to compare the percentages of the activated nodes for different topics in social networks.

IC model [11]: A node u activating its adjacent node v is a probability event. The probability that an inactive node is activated by an activated neighbor node is independent of the state of its neighbor who has tried to activate the node before.

Second-order (2nd) IC model [16]: It is an improved version of the IC model, which considers the impact of secondary forwarding in the process of information dissemination. In the second-order IC model, the propagation channel can transmit information to its neighbor nodes, and these neighbor nodes can forward the information to their neighbor nodes.

ICPB model [17]: It adds the influences of users' behavior and information diffusion to the IC model. Users' behavior is the reaction of the communication channel or users' response to a piece of information. Actions can be forwarded, commented on, liked or ignored. The user influence refers to the influence of communication channels or audiences in social networks. It can measure social relationships, user activity and the number of fans.

CCIC model [19]: It considers connection coefficients between nodes. Compared with the IC model, it more accurately describes the process of information dissemination in social networks. In the CCIC model, information sources deliver information to communication channels with higher connection coefficients. Then, these communication channels deliver information to their neighbor nodes and forward it in order of connection coefficients.

3.3. Experimental setup and evaluation index

In all experiments, we adopt a time slice method, taking the number of users participating in a topic every hour as a time slice. The total number of social network nodes is the number of users participating in a topic and the users' fans. After a topic is published, the users who participate within two hours are regarded as the initial activated nodes, while others are not activated. We conduct data analysis on these three topics and find that each topic has an active period of information propagation. The active period is taken as a research object. According to the detailed data, we uniformly set the number of users $N = 1000$ and the popularity threshold $\varepsilon = 100$. Table 4 shows the specific parameters at the period of initial information propagation.

For the sake of qualitative analysis and comparisons between the proposed method and baselines, we use the mean square error (MSE) as the evaluation index of all methods.

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2, \tag{3.1}$$

Table 4. Partial experimental parameter setting.

Topics	a_0	b_0	c	α	∂
Topic 1	0.11	0.01	0.21	1.19	0.39
Topic 2	0.04	0.02	0.05	1.30	0.22
Topic 3	0.09	0.13	0.65	1.02	0.36

where y_i represents the ratio of activated nodes to all nodes at time t in Weibo networks, $\hat{y}_i$ represents the ratio of activated nodes predicted at the time i to all nodes in Weibo networks, and m represents the total time of Weibo networks. MSE reflects the degree of difference between a predictive method's predictive value and a real dataset's accurate value. The smaller the value of MSE is, the better the accuracy of a predictive method is in a Weibo network.

3.4. Performance comparison

From Figures 3–5, in the initial stage, the percentages of activated nodes under the ICTSL model and baselines are almost consistent with the actual values for Topics 1, 2 and 3. In the change and development stage and the mature development stage, the ICTSL model well predicts the trends of the topic diffusion for Topics 1, 2 and 3. However, the percentages of activated nodes predicted by the baseline models are insufficient. The main reason is that they do not consider a trending search list that will attract other nodes. In the degrading and buffer stages, the percentages of activated nodes under the ICTSL model can predict the trends of activated nodes for Topics 1, 2 and 3, even reaching the peak at the second time for Topic 1. In this stages, the IC, ICPB, CCIC and second-order IC models cannot predict the percentages of activated nodes for Topics 1, 2 and 3. Although the ICTSL model perform poorly in the fading stage on Topics 1, 2 and 3, it is generally superior to other comparison models. The reason is that the nodes attracted by the trending search list are activated, steadily creating a degree of information hot and making information diffusion more widely. Therefore, the ICTSL model can well predict the trends of the topic diffusion on actual datasets.

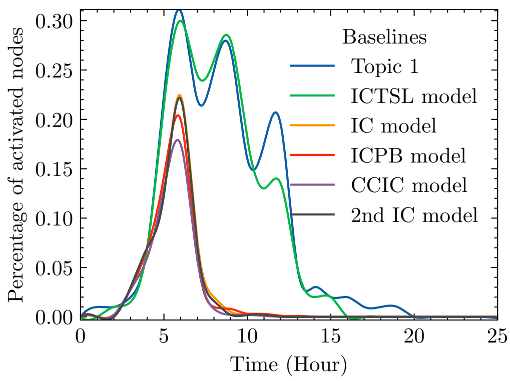


Figure 3. Comparisons of the predictive results of the IC, second-order IC, ICPB model, CCIC and ICTSL models on Topic 1.

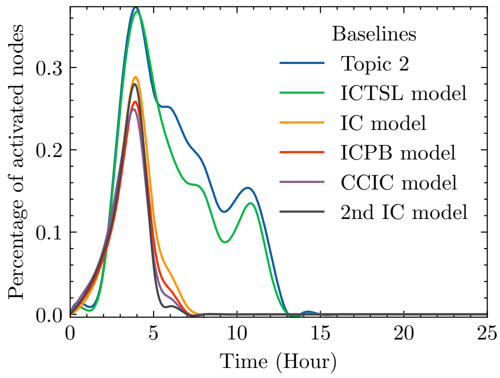


Figure 4. Comparisons of the predictive results of the IC, second-order IC, ICPB model, CCIC and ICTSL models on Topic 2.

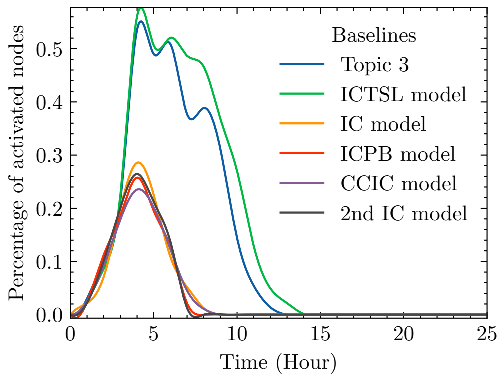


Figure 5. Comparisons of the predictive results of the IC, second-order IC, ICPB model, CCIC and ICTSL models on Topic 3.

From Table 5 and Figure 6, the ICTSL model has the smallest MSE (0.46%) compared with other models for Topics 1, 2 and 3. For Topic 1, the MSE of the IC model is the largest (1.26%). The MSEs of the ICPB and CCIC models are the second smallest (1.24%). The MSE of the ICTSL model is 0.80%, 0.78%, 0.78% and 0.79% lower than the IC, ICPB, CCIC and second-order IC models, respectively. For Topics 2 and 3, the MSE of the IC model is the largest, while the MSE of the second-order IC model is the second largest. The MSE of the ICTSL model is 2.09% and 2.19% lower than the CCIC and ICPB models for Topic 2, respectively. Meanwhile, the MSE of the ICTSL model is 3.40% and 3.42% lower than the CCIC and ICPB models for Topic 3, respectively.

Compared with the IC, ICPB, CCIC and second-order IC models, the MSE of the proposed ICTSL model is decreased by approximately 0.78%–3.71% for three real topics. Overall, the accuracy of the proposed ICTSL model is the best for Topics 1, 2 and 3.

The above experiments conclude that the ICTSL model can effectively predict the trend of topic diffusion in social networks. Suppose we need to lead the trend of topic diffusion. Then, we can change the user diffusion willingness and network scale to control the trend of the topic diffusion according to the predictive result of the ICTSL model on the topic in social networks.

Table 5. MSE values of different models.

Topics	IC model	ICPB model	CCIC model	2nd IC model	ICTSL model
Topic 1	1.26%	1.24%	1.24%	1.25%	0.46%
Topic 2	3.68%	3.62%	3.52%	3.72%	1.43%
Topic 3	5.62%	5.33%	5.31%	5.37%	1.91%

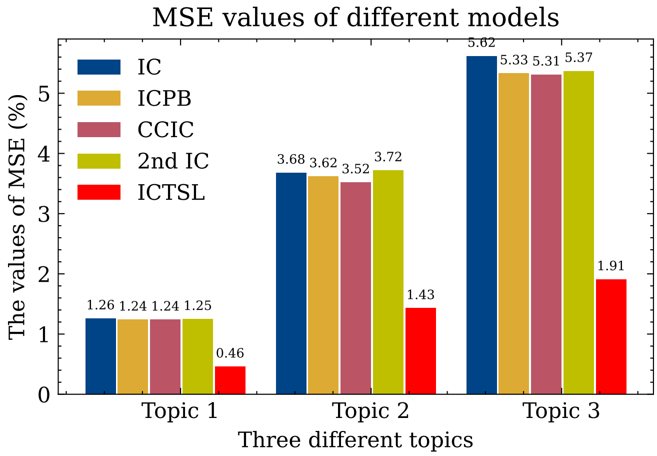


Figure 6. MSE values of different models.

3.5. Parameter analysis

First, we compare the percentages of activated nodes under the ICTSL model on different initial scale networks. The age factor b_0 is set to 30; the influence factor a_0 is set to 0.03; the topic popularity threshold ε is set to 100; a is set to 0.044; b is set to 1.35 and the hot cooling coefficient ∂ is set to 0.28. From Figure 7, in the initial stage and the change and development stage, the larger a network's initial scale is, the faster the percentage of activated nodes increases. The peak of the large initial scale network is much higher than that of a small initial scale network. The main reasons are as follows. (1) The users of the large initial scale network have more opportunities to get the topic information from trending search lists than the small initial scale network. (2) The large initial scale network is easier to reach the topic popularity threshold ε than the small initial scale network. (3) Topics on the large initial scale network easily enter trending search lists.

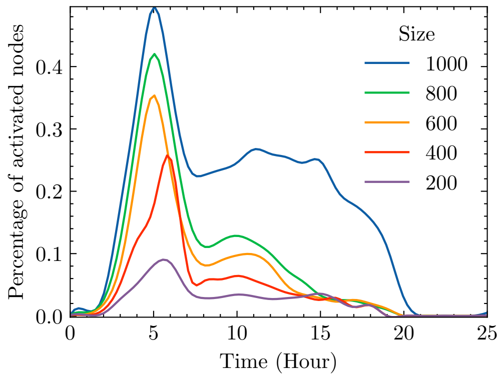


Figure 7. Comparisons of predictive results under the ICTSL model on different initial scale networks.

Next, we discuss the influence of different user diffusion willingness on the ICTSL model (see Figure 8). On an initial diffusion network, the number of users N is set to 1000; the age factor b_0 is set to 30; the influence factor a_0 is set to 0.03; the topic popularity threshold ε is set to 100; a is set to 0.041; b is set to 1.33 and the hot cooling coefficient ∂ is set to 0.21.

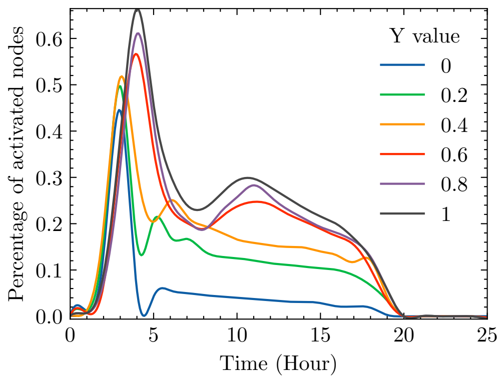


Figure 8. Influence of different user diffusion willingness on the ICTSL model.

As shown in Figure 8, the percentages of activated nodes of the ICTSL model have roughly the same trends under different user diffusion willingness. When the user diffusion willingness values Y = 0, 0.2 and 0.4, the percentages of activated nodes simultaneously peak at time 3 in the mature developmental stage. When the user diffusion willingness values Y = 0.6, 0.8 and 1, the percentages of activated nodes increase and peak at time 4 in the change and developmental stage. Therefore, the larger the user diffusion willingness value is, the higher the percentage of activated nodes is at peak time under the ICTSL model.

In the degrading stage, the percentages of activated nodes of the ICTSL model all drop sharply under different user diffusion willingness. The lower the user diffusion willingness value is, the shorter the time of the degradation stage is. Thus, even if topic popularity decreases, one can increase the user diffusion willingness to activate nodes on the network.

When user diffusion willingness values are different, the durations of the buffer stages are also different under the ICTSL model. The higher the user diffusion willingness value is, the longer the duration of the buffer stage is. In the fading stage, the percentage of active nodes quickly drops to 0. At about time 20, the ICTSL model tends to stabilize. Currently, topic popularity tends to be 0; users will ignore topic information. While the user diffusion willingness value increases, more users will be attracted to the topic information, keeping it longer, and vice versa. Therefore, for a hot topic, if we need to speed up its topic diffusion, we can mobilize the relevant users of the topic with an enormous diffusion willingness to participate in the discussion of the relevant topic. In contrast, if we need to decrease the topic diffusion quickly, we can reduce the diffusion willingness by choosing different users of different ages on social networks.

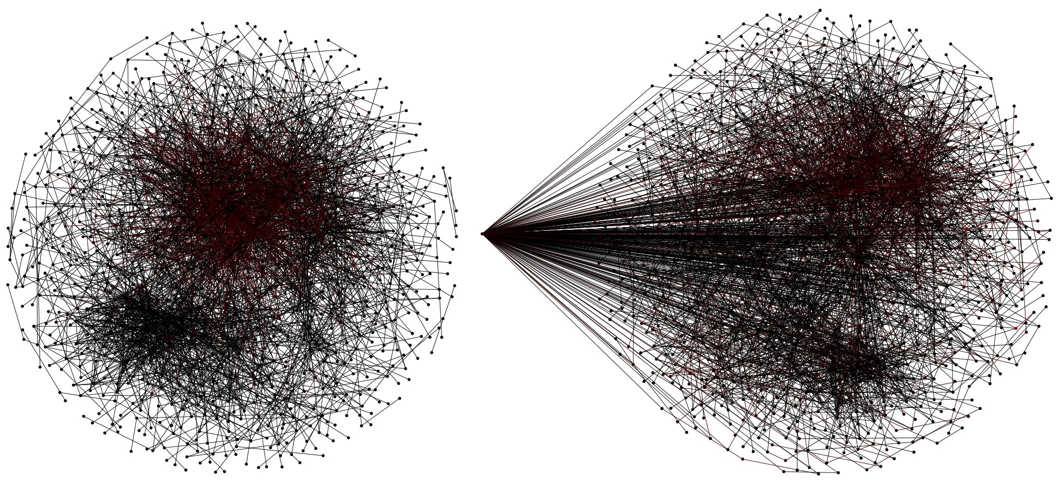
3.6. Case study

To analyze the diffusion process of hot topics on Weibo networks, we make a case study under the ICTSL and IC models for Topic 3. Compared with the IC model, we will verify the effectiveness of the ICTSL model in the life cycle of Topic 3.

Under the IC and ICTSL models, activated nodes are similar in the initial stage (see Figure 9). Since the topic popularity threshold of the ICTSL model is not reached 100 in the initial phase, the models propagate information through the propagation rule of the IC model at the initial propagation phase. The difference is that each information propagation is influenced by a trending topic node under the ICTSL model, increasing the hot degree of the trending topic node.

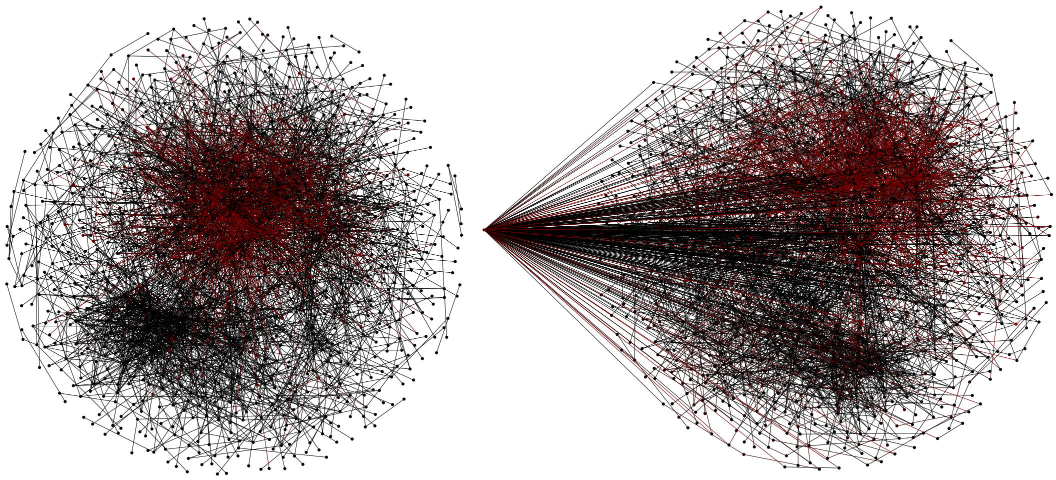
In the change and development stage, Topic 3 receives many nodes' attention over time. Under the IC model, the nodes with large degrees propagate the topic, maintaining a good propagation trend (see Figure 10(a)). Under the ICTSL model, topic propagation maintains the same propagation trend as the IC model. The topic propagation of the ICTSL model is wider than the IC model. The reason is that many nodes connect with the trending topic node, significantly raising the topic's hot degree under the ICTSL model (see Figure 10(b)).

In the mature development stage (see Figure 11), since the topic popularity threshold reaches 100



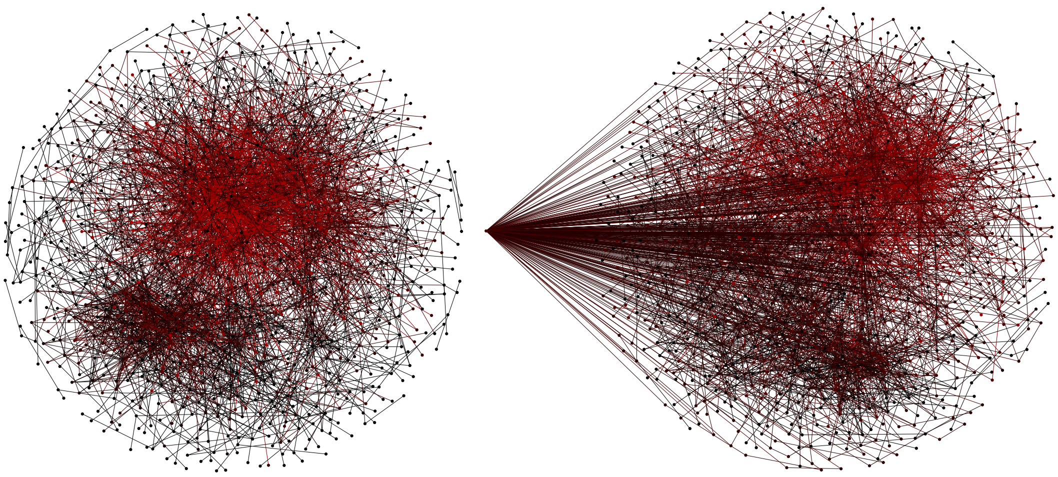
(a) IC Model (b) ICTSL Model

Figure 9. The topic propagation in the initial stage of Topic 3 under the IC and ICTSL models.



(a) IC Model (b) ICTSL Model

Figure 10. The topic propagation in the change and development stage of Topic 3 under the IC and ICTSL models.

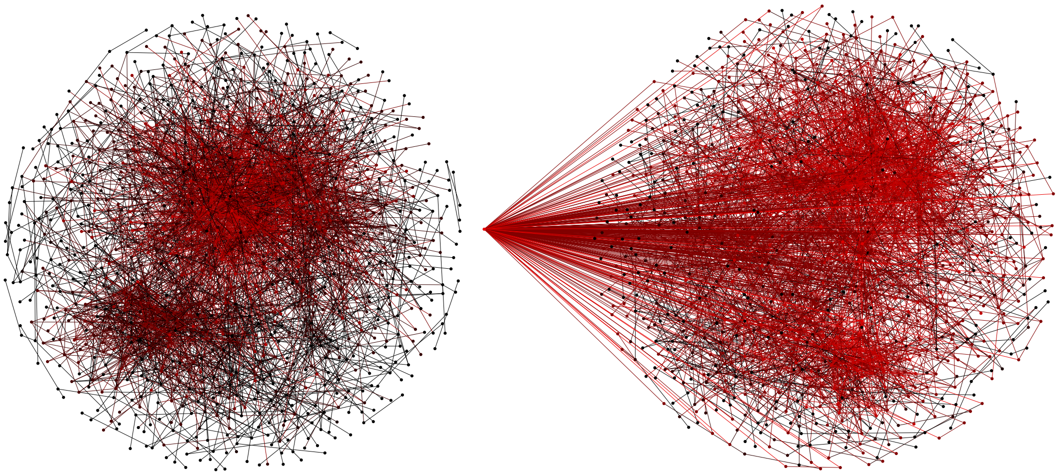


(a) IC Model (b) ICTSL Model

Figure 11. The topic propagation in the mature development stage of Topic 3 under the IC and ICTSL models.

under the ICTSL model, the trending topic node attracts a lot of non-neighbor nodes to propagate this topic information. Then, the nodes activated by the ICTSL model are more than the IC model.

In the degrading stage and the buffer stage, few nodes are activated under the IC model because the hot degree of Topic 3 declines. However, there are still some activated nodes under the ICTSL model. The reason is that there exists the trending topic node which also attracts new nodes to propagate the topic information. The number of activated nodes under the ICTSL model maintains a certain level in the buffer stage or even reaches a new peak (see Figure 12). In the fading stage, few new nodes are activated under the IC and ICTSL models, and the topic propagation will end over time (see Figure 13).



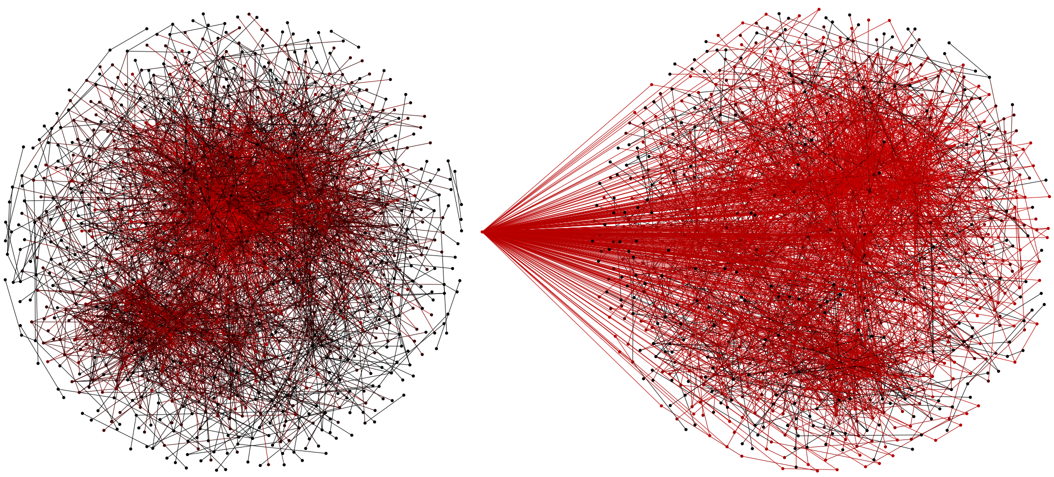
(a) IC Model (b) ICTSL Model

Figure 12. The topic propagation in the degrading stage and buffer stage of Topic 3 under the IC and ICTSL models.

Based on the above experiments, the propagation trends of Topics 1, 2 and 3 under the ICTSL model are more in line with the law of realistic diffusion than other comparison models. Therefore, the ICTSL model can well describe topic propagation in real social networks.

4. Conclusions

This paper proposed novel terms, including user diffusion willingness, doubt degree, topic contribution, topic popularity and the number of new nodes for social networks. Based on these terms, trending



(a) IC Model (b) ICTSL Model

Figure 13. The topic propagation in the fading stage of Topic 3 under the IC and ICTSL models.

search lists and the IC model, a hot topic diffusion approach based on the independent cascade model and trending search lists (ICTSL model) is proposed. Simulation experiments discuss the influence of user willingness and different initial scale networks for the proposed ICTSL model on information diffusion. Compared with the IC, ICPB, CCIC and second-order IC models, the ICTSL model could better predict the trend of topic diffusion on three actual topics of Weibo networks.

Since various hot search media platforms have launched hot topic trending search lists, the topic diffusion on online social media differs from the traditional diffusion methods. The proposed ICTSL model can better describe the communication trend of hot topics. The ICTSL model is helpful for information diffusion, public opinion analysis and rumor suppression. However, the ICTSL model does not consider the information diffusion of different unknown topics. In the future, we will consider the topic order and position in trending search lists and propose novel hot topic diffusion models to improve the accuracy of predicting the diffusion trend of multiple topics on social networks.

Acknowledgments

The authors express our sincere thanks to the editors and reviewers. This work is partially supported by the Sichuan Science and Technology Program (Nos. 2022YFG0378, 2023YFS0424, 2023YFQ0044 and 2021YFQ0008), the National Natural Science Foundation of China (Nos. 61802316, 61872298 and 61902324), and the Innovation Fund of Postgraduate, Xihua University (Grant Nos. YCJJ2021124, YCJJ2021025 and YCJJ2021031).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. Y. Lu, J. Liu, The impact of information dissemination strategies to epidemic spreading on complex networks, *Phys. A* , **536** (2019), 120920. https://doi.org/10.1016/j.physa.2019.04.156
2. Y. Z. Ning, X. Liu, H. M. Cheng, Z. Y. Zhang, Effects of social network structures and behavioral responses on the spread of infectious diseases, *Phys. A* , **539** (2019), 378–4371. https://doi.org/10.1016/j.physa.2019.122907
3. T. Wang, J. He, X. Wang, An information spreading model based on online social networks, *Phys. A* , **490** (2018), 488–496. https://doi.org/10.1016/j.physa.2017.08.078
4. Y. Wang, J. Wang, H. Wang, L. Zhang, M. Li, Users' mobility enhances information diffusion in online social networks, *Inf. Sc.*, **546** (2020), 329–348. https://doi.org/10.1016/j.ins.2020.07.061
5. Y. Xiao, W. Li, S. Qiang, A rumor & anti-rumor propagation model based on data enhancement and evolutionary game, *IEEE Trans. Emerging Top. Comput.*, **10** (2020), 690–703. https://doi.org/10.1109/TETC.2020.3034188
6. Y. Xiao, Z. Huang, Q. Li, X. Lu, T. Li, Diffusion pixelation: A game diffusion model of rumor & anti-rumor inspired by image restoration, *IEEE Trans. Knowl. Data Eng.*, **1** (2022). https://doi.org/10.1109/TKDE.2022.3144310

7. Q. Li, T. Xiang, T. Dai, Y. Xiao, An information dissemination model based on the rumor & anti-rumor & stimulate-rumor and tripartite cognitive game, *IEEE Trans. Cognit. Develop. Syst.*, **1** (2022), 1–13. https://doi.org/10.1109/TCDS.2022.3193576

8. D. Jin, X. Ma, Y. Zhang, H. Abbas, H. Yu, Information diffusion model based on social big data, *Mobile Networks Appl.*, **23** (2018), 717–722. https://doi.org/10.1007/s11036-018-1004-4

9. S. Kumar, M. Sanini, M. Goel, N. Aggarwal, Modeling information diffusion in online social networks using SEI epidemic model, *Proc. Comput. Sci.*, **171** (2020), 672–678. https://doi.org/10.1016/j.procs.2020.04.073

10. Q. Zhang, X. Li, Y. Fan, Y. Du, An SEI3R information propagation control algorithm with structural hole and high influential infected nodes in social networks, *Eng. Appl. Artif. Intell.*, **108** (2022), 104573. https://doi.org/10.1016/j.engappai.2021.104573

11. D. Kempe, Maximizing the spread of influence through a social network, *Theory Comput.*, **11** (2015), 105–147. https://doi.org/10.1145/956750.956769

12. R. Olivares, F. Muoz, F. Riquelme, A multi-objective linear threshold influence spread model solved by swarm intelligence-based methods, *Knowl. Based Syst.*, **212** (2021), 106623. https://doi.org/10.1016/j.knosys.2020.106623

13. J. Tang, R. Zhang, P. Wang, Z. Zhao, L. Fan, X. Liu, A discrete shuffled frog-leaping algorithm to identify influential nodes for influence maximization in social networks, *Knowl. Based Syst.*, **187** (2020), 104833. https://doi.org/10.1016/j.knosys.2019.07.004

14. K. Saito, M. Kimura, K. Ohara, H. Motoda, Learning continuous-time information diffusion model for social behavioral data analysis, *Adv. Mach. Learn.*, (2009), 322–337. https://doi.org/10.1007/978-3-642-05224-8_25

15. M. K. Alasadi, H. N. Almamory, Diffusion model based on shared friends-aware independent cascade, *J. Phys.*, **1294** (2019), 1884–2023. https://doi.org/10.1088/1742-6596/1294/4/042006

16. W. Tang, G. Luo, Y. Wu, L. Tian, Z. Cai, A second-order diffusion model for influence maximization in social networks, *IEEE Trans. Comput. Social Syst.*, **6** (2019), 702–714. https://doi.org/10.1109/TCSS.2019.2921422

17. K. Li, T. Xu, S. Feng, L. Qiao, H. Shen, X. Cheng, et al., The propagation background in social networks: Simulating and modeling, *Int. J. Autom. Comput.*, **17** (2020), 353–363. https://doi.org/10.1007/s11633-020-1227-2

18. A. Saxena, H. Saxena, R. Gera, Competitive influence propagation and fake news mitigation in the presence of strong user bias, preprint, arXiv: 2011.04857v1.

19. H. Huang, Z. Meng, H. Shen, Competitive and complementary influence maximization in social network: a follower's perspective, *Knowl. Based Syst.*, **213** (2020), 106600. https://doi.org/10.1016/j.knosys.2020.106600

20. L. Zhang, J. Wei, R. J. Boncella, Emotional communication analysis of emergency microblog based on the evolution life cycle of public opinion, *Inf. Discovery Delivery*, **48** (2020), 151–163. https://doi.org/10.1108/IDD-10-2019-0074

21. J. Wei, L. Zhang, Analysis of information dissemination based on emotional and the evolution life cycle of public opinion, in *2019 International Conference on Robots & Intelligent System (ICRIS)*, 2019, 265–268. https://doi.org/10.1109/ICRIS.2019.00076
22. Y. He, J. Li, M. Zhu, D. Xu, Life cycle identification and analysis of microblog hot topics, in *2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, **2** (2018), 156–159. https://doi.org/10.1109/IHMSC.2018.10142
23. C. Wukich, A. Steinberg, Nonprofit and public sector participation in self-organizing information networks: Twitter hashtag and trending topic use during disasters, *Risk Hazards Crisis Pub. Policy*, **4** (2013), 83–109. https://doi.org/10.1002/rhc3.12036
24. P. Li, H. Nie, F. Yin, J. Liu, D. Zhou, Modeling and estimating user influence in social networks, *IEEE Access*, **8** (2020), 21943–21952. https://doi.org/10.1109/ACCESS.2020.2969361
25. Z. Chen, J. Wei, S. Liang, T. Cai, X. Liao, Information cascades prediction with graph attention, in *IEEE Transactions on Emerging Topics in Computing*, **9** (2021). https://doi.org/10.3389/fphy.2021.739202
26. H. Zhang, Y. Yao, W. Tang, J. Zhu, A Hegselmann-Krause model with social influence-aware bounded confidence threshold in social networks, in *IEEE Transactions on Network Science and Engineering*, **5** (2022), 22–24. https://doi.org/10.1109/DSIT55514.2022.9943930
27. S. Silva, Newton's cooling law in generalised statistical mechanics, *Phys. A*, **565** (2021), 125539. https://doi.org/10.1016/j.physa.2020.125539
28. J. Goldenberg, L. E. Muller, Talk of the network: A complex systems look at the underlying process of word-of-mouth, *Mark. Lett.*, **12** (2001), 211–223. https://doi.org/10.1023/A:1011122126881
29. J. Goldenberg, B. Libai, Using Complex systems analysis to advance marketing theory development: modeling heterogeneity effects on new product growth through stochastic cellular automata, *Acad. Mark. Sci. Rev.*, **9** (2001), 1–19. https://www.amsreview.org/articles/goldenberg09-2001.pdf

AIMS Press