TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214
DOI: 10.26599/TST.2025.9010081

Research Article

Dynamic Linear Threshold Model with Decay for Influence Spread

Xunlian Wu, Anqi Zhang, Yining Quan, Qiguang Miao, and Penggang Sun*

Abstract: Influence propagates more widely over social networks, and strongly affects people's thoughts, opinions, and even behaviors in human life. Influence spread is a more complicated problem that makes the modeling of this process even more challenging. The linear threshold (LT) model provides a simple way to describe the problem. However, this model neglects the decay of individuals' influence as time lapses. This paper develops a dynamic linear threshold model with decay (DLTD) in which influence decays based on three strategies: (1) Random decay, (2) linear decay, and (3) nonlinear decay. The results on real-world datasets show that compared with the LT model and the independent cascade (IC) model, the DLTD tends to be more accurate especially for the decay with a linear function, and more concentrated to converge to the ground-truth as the size of seed set increases. We also observe that our model is more stable to the networks' structures. Furthermore, an interesting finding indicates that the balance between community structures' internal links and external links is more effective for influence spread.

Key words: influence spread; dynamic linear threshold model; linear threshold model; social networks; decay of influence

1 Introduction

With the development of information technologies, people linked by social networks communicate more frequently, and individuals' influence propagates more quickly, which will strongly affect people's thoughts, opinions, even behaviors in human livings. Influence spread is widely observed on social networks and plays an important role in many applications, e.g., viral marketing[1–4], personalized recommendations[5], rumor control[6–8], selection of influential blog posts[9, 10], product promotion[11], and behavior adoption[12, 13]. Recently, influence spread has also been used in the field of community detection[14–17]. Influence maximization[18–20] is to find the most influential seed set with a given size that can maximize the area for influence spread. For example, the most influential users' opinions on a new product propagate on social networks, and further motivate more users to buy the product, and tend to maximize the spread of influence.

Influence spread is a more complicated problem that makes the modeling of this process even more challenging. As Kazemzadeh et al.[21] mentioned, the most commonly used diffusion models include the classical epidemic outbreak model[22], independent cascade (IC) model[23], weight cascade (WC) model[24], linear threshold (LT) model[23], and trigger model[25]. The IC model and LT model are two well-known models which provide different rules to activate nodes in the spreading process respectively. The IC model neglects the heterogeneity of nodes, i.e., each active node has the same probability to activate its neighbors

• Xunlian Wu, Yining Quan, Qiguang Miao, and Penggang Sun are with School of Computer Science and Technology, Xidian University, Xi'an 710071, China. E-mail: xunlianwu@stu.xidian.edu.cn; ynquan@xidian.edu.cn; qgmiao@xidian.edu.cn; psun@mail.xidian.edu.cn.
• Anqi Zhang is with School of Computer and Artificial Intelligence, Henan Finance University, Zhengzhou 450046, China. E-mail: zhanganqi@hafu.edu.cn.
* To whom correspondence should be addressed.
Manuscript received: 2024-07-02; revised: 2025-02-18; accepted: 2025-04-18

independently. To solve this problem, some works such as Refs. [23, 24], assume that each active node has a random or weighted probability to activate its neighbors independently. The LT model addresses that each inactive node has its own threshold, and is more likely to be activated if more neighbors are active. Over the past several years, many variants of the LT model were developed[11, 23, 26–30]. For instance, Pathak et al.[27] proposed a generalized LT model, which can simulate multiple cascades by allowing a node to switch its states. Yang and Leskovec[28] introduced a LT model that can measure the global influence of a node by adopting a non-parametric influence function. In particular, Kempe et al.[23] provided a well-accepted formulation for the LT model by considering random thresholds, which has been widely used for influence spread[31–38]. In the LT model, initially, only the seed nodes are active, and each inactive node has a random threshold. At each time, each active node tries to activate inactive nodes in its neighborhood, and an inactive node can be activated if the accumulated influence from its active neighbors exceeds the inactive node's threshold at least.

Intuitively, individuals' influence decays over time. Influence decay in social networks refers to the phenomenon where the impact or influence of an individual's actions, opinions, or behaviors on others diminishes over time. This decay is a natural consequence of several factors that are commonly observed in real-world social interactions. **(1) Novelty wearing off:** When a new piece of information or behavior is introduced (e.g., a new product purchase or a viral post), its initial impact is strong because it is novel. However, as time passes, the novelty diminishes, and the influence weakens[39]. For example, if a user buys a new phone and shares it with friends, the initial excitement may prompt friends to consider buying the same phone, but this effect fades as time goes on (as shown in Fig. 1). **(2) Diminishing marginal impact:** As more people are exposed to the information or behavior, the additional impact on new individuals decreases. This is similar to the economic concept of diminishing returns, where the incremental benefit of additional exposure becomes smaller over time[40]. **(3) Competing influences:** Over time, other competing influences (e.g., alternative products, conflicting opinions) may dilute the original impact[41]. For instance, if multiple users recommend different products, the influence of the initial recommendation may be overshadowed by newer ones. As shown in Fig. 1: User A is an active node who bought a cell phone, and user B is user A's friend, and the behavior of user A was not adopted by user B at $t = 0$, then the influence of user A on user B will decay as time lapses, where the strength of influence is described by link weights. However, this problem is often neglected by the LT model.

To this end, this paper develops a dynamic linear threshold model with decay (DLTD) in which influence decays based on three strategies: (1) Random decay, (2) linear decay, and (3) nonlinear decay. We conduct extensive experiments on real-world datasets as well as synthetic networks to investigate the performance of our model.

The main contributions of this work are as follows:

(1) Introduction of the DLTD: We propose a novel DLTD that introduces three distinct decay strategies during the influence propagation process. This model enhances the traditional linear threshold model by accounting for the time-dependent decay of influence, offering a more realistic and adaptable framework for influence spread.

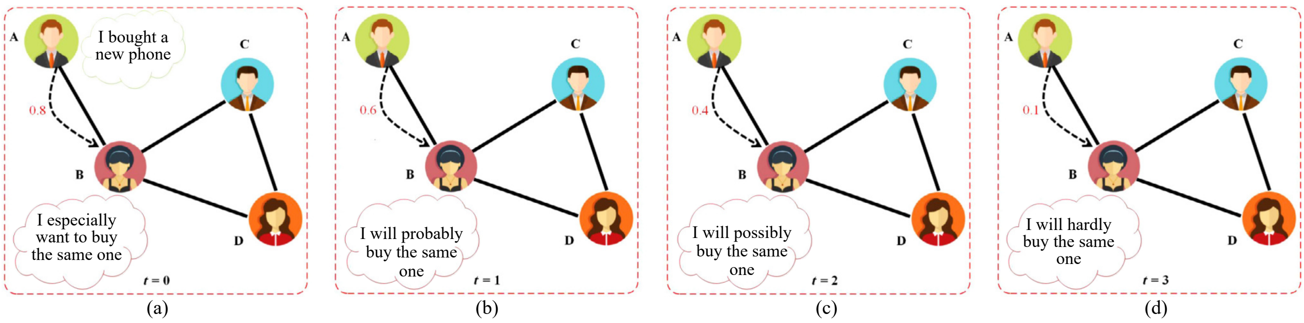


Fig. 1 Illustration of the decay of influence as time lapses. (a) shows that user A is an active node who bought a cell phone, and user B is user A's friend, and the behavior of user A was not adopted by user B at *t* = 0. (b), (c) and (d) show the that influence of user A on user B will decay as time lapses, where the strength of influence is described by link weights.

(2) Experimental validation on real-world datasets: We conduct extensive experiments on real-world datasets to evaluate the effectiveness of the DLTD model. Our experiments show that, compared to the traditional LT model, DLTD particularly with linear decay provides higher accuracy and better convergence towards the ground truth, especially as the seed set size increases. These results demonstrate the practical advantages of incorporating decay into the influence spread model.

(3) Stability and robustness analysis on synthetic networks: We investigate the performance of DLTD on synthetic networks with varying characteristics. The results indicate that our model is more robust and stable across different network structures. Furthermore, we reveal that the balance between internal and external links within community structures significantly enhances the effectiveness of influence spread. This insight provides a theoretical basis for designing more efficient influence propagation strategies in complex networks.

The remainder of this paper is organized as follows. In Section 2, we describe preliminary for influence spread. In Section 3, we introduce our DLTD. In Section 4, we test our model on real-world datasets as well as synthetic networks. Section 5 is the conclusion.

2 Preliminary

2.1 Notions

$G(V, E)$ is a directed graph, where V is the node set, and E is the edge set. $N^{\text{in}}(v)$ and $N^{\text{out}}(v)$ are the set of incoming neighbors and outcoming neighbors of node v, respectively. The adjacency matrix $\boldsymbol{A}$ can be used to describe nodes' interactions in the graph, i.e., $a_{uv} = 1$ if node u and node v are linked by an edge, and $a_{uv} = 0$ otherwise. Formally,

$$\boldsymbol{A} = (a_{uv})_{n\times n} = \begin{cases} a_{uv} = 1, & (u, v) \in E; \\ a_{uv} = 0, & (u, v) \notin E \end{cases} \tag{1}$$

where $u, v \in V$, and $|V| = n$.

For influence spread on G, it is generally transformed into its directed version G^d, i.e., each edge in G becomes bidirectional, weighted in G^d, and can be denoted by $\boldsymbol{W}$.

$$\boldsymbol{W} = (w_{u\mapsto v})_{n\times n} = \begin{cases} w_{u\mapsto v} \in [0, 1], & (u, v) \in E; \\ w_{v\mapsto u} \in [0, 1], & (u, v) \in E \end{cases} \tag{2}$$

where $\boldsymbol{W}$ is an asymmetric matrix, $w_{u\mapsto v}$ is used to describe the strength of influence from u to v, and u, $v \in V$.

The sum of link weights of incoming neighbors of a node satisfies,

$$\sum_{u\in N^{\text{in}}(v)} w_{u\mapsto v} \leqslant 1 \tag{3}$$

2.2 LT model

For the LT model, the core idea is that a user's state can change from inactive to active when there are enough activated incoming neighbors, and sum of the incoming weights is at least the user's threshold.

First of all, we initialize a seed set S by selecting some nodes from V, and only nodes within S are active. Each inactive node v has a threshold θ_v, and θ_v is uniformly chosen from $[0,1]$ at random. At each time step, an active node attempts to activate its inactive neighbor v, and v becomes active if the total weight of its active incoming neighbors is at least θ_v. The process stops until no more nodes can be activated. In Fig. 2, we illustrate the LT model for influence spread.

$$\sum_{u\in N^{\text{in}}(v)} w_{u\mapsto v} \geqslant \theta_v \tag{4}$$

Matrix $\boldsymbol{W}$ can be determined by the uniform method or random method[42]. For the uniform method,

$$\boldsymbol{W} = (w_{u\mapsto v})_{n\times n} = \begin{cases} w_{u\mapsto v} = \dfrac{1}{\left|N^{\text{in}}(v)\right|}, \\ w_{v\mapsto u} = \dfrac{1}{\left|N^{\text{in}}(u)\right|} \end{cases} \tag{5}$$

For the random method, the weight of every edge is generated at random, $r_{uv} \in [0, 1]$, after weights are assigned, we normalize the weights of all incoming edges of a node v so that they sum to 1.

$$\boldsymbol{W} = (w_{u\mapsto v})_{n\times n} = \begin{cases} w_{u\mapsto v} = \dfrac{r_{uv}}{\sum\limits_{m\in N^{\text{in}}(v)} r_{mv}}, \\ w_{v\mapsto u} = \dfrac{r_{vu}}{\sum\limits_{m\in N^{\text{in}}(u)} r_{mu}} \end{cases} \tag{6}$$

3 DLTD

The LT model provides a simple way to describe the problem. However, this model neglects the decay of individuals' influence as time lapses. This paper develops a DLTD in which influence decays based on

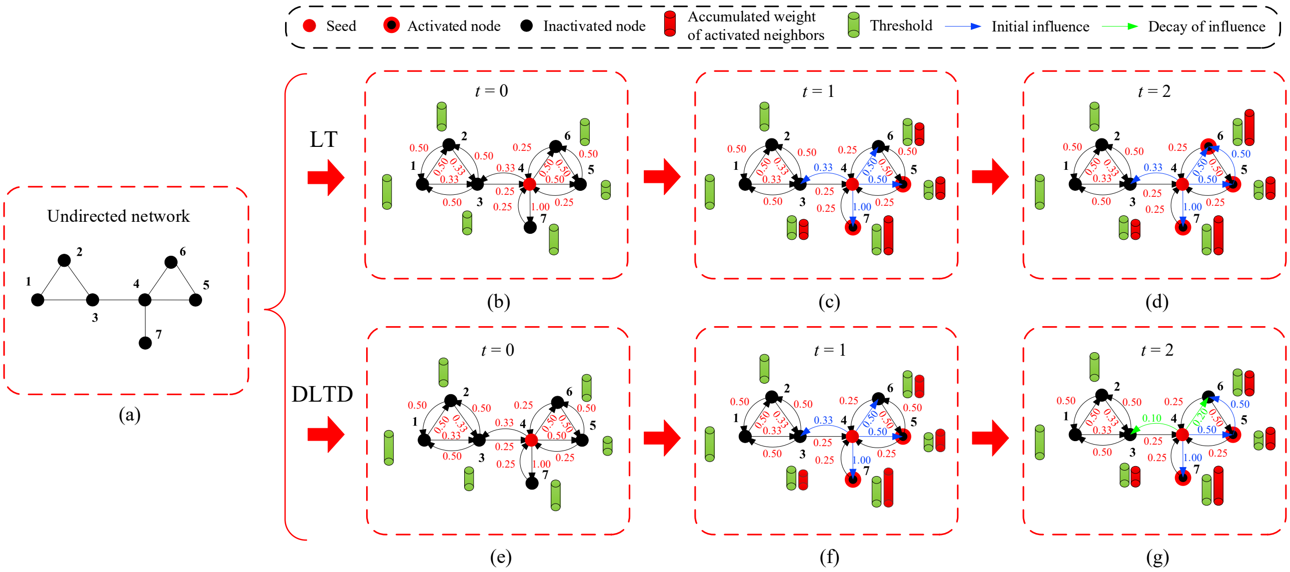


Fig. 2 Illustration of the LT model and the DLTD model for influence spread. (a) A toy network is used in this paper. (b)–(d) correspond to the LT model. (e)–(g) correspond to the DLTD model. In (b) and (e), only node 4 as a seed is active, and each inactive node is initialized with a threshold at random. At time $t = 0$, the strength of influence from one node to another node is described by edge weights, and the weights are determined by the uniform method. In (c) and (f), node 4 attempts to activate its inactive neighbors at time $t = 1$, and node 5 and node 7 are activated directly. In (d), at time $t = 2$, the accumulated influences from node 4 and node 5 make node 6 active based on the LT model, while in (g), it is totally different for the DLTD model due to the decay of influence. The process stops until no more nodes can be activated.

three strategies: (1) Random decay, (2) linear decay, and (3) nonlinear decay.

For the DLTD, at time $t = 0$, only seed nodes are in active state, and every inactive node is initialized with a threshold randomly. We also determine $\boldsymbol{W}^t$ at $t = 0$,

$$\boldsymbol{W}^{(t=0)} = \left(w^{(0)}_{u\mapsto v}\right)_{n\times n} = \begin{cases} w^{(0)}_{u\mapsto v} = \dfrac{1}{\left|N^{\text{in}}(v)\right|}; \\ w^{(0)}_{v\mapsto u} = \dfrac{1}{\left|N^{\text{in}}(u)\right|} \end{cases} \tag{7}$$

At each time step, an active node attempts to activate its inactive neighbors. Given that node u becomes active and node v is still inactive at time $t-2$, the influence of node u on node v is invariable at time $t-1$, $t-2$, ..., 1, 0. Accordingly, the corresponding influence weight can be expressed as

$$w^{(t-1)}_{u\mapsto v} = w^{(t-2)}_{u\mapsto v} = \cdots = w^{(1)}_{u\mapsto v} = w^{(0)}_{u\mapsto v} \tag{8}$$

For example, at time $t = 0$, node 4 is active, $w^{(0)}_{4\mapsto 6} = w^{(1)}_{4\mapsto 6} = 0.5$ in Figs. 2e and 2f.

At time t, node v is still inactive, the influence from node u to node v starts to decay based on a function $\varPhi$ as t increases, e.g., at time $t = 2$, $w^{(2)}_{4\mapsto 6} < w^{(1)}_{4\mapsto 6}$ in Fig. 2g. Accordingly, the influence weight from node u to node v at time t can be updated as

$$w^{t}_{u\mapsto v} = \varPhi\left(w^{(t-1)}_{u\mapsto v}\right) \tag{9}$$

For the function $\varPhi$, we propose three strategies to control the decay of influence in the process,

$$\varPhi(x) = \begin{cases} (1-\alpha)x, \\ (1-\beta)x, \\ x^{\gamma} \end{cases} \tag{10}$$

where $\alpha \in (0, 1]$ is chosen uniformly at random, and can be used to describe the random decay. In general, we set α between different nodes at each time, e.g., at time $t = 2$, α for $w^{(2)}_{4\mapsto 3}$ and $w^{(2)}_{4\mapsto 6}$ in Fig. 2g is different. In Eq. (10), β is chosen from (0, 1], and can be used to describe the linear decay. Here, for different nodes, β is fixed. In Eq. (10), it is an exponential function, and can be used to describe the nonlinear decay. Here, $\gamma = 2$. The DLTD model with random decay is called DLTD-R, with linear decay called DLTD-L, with nonlinear decay called DLTD-NL. These strategies are designed to capture different real-world scenarios and behaviors observed on social networks. In the following, we provide a detailed explanation of each strategy. (1) Random decay: This strategy captures the inherent randomness and unpredictability in human behavior and influence dynamics. (2) Linear decay: This reflects a steady and predictable reduction in influence over time, which is often observed in real-world scenarios where the novelty of information diminishes at a constant rate. (3) Nonlinear decay: This accounts for scenarios where influence decreases rapidly at first and then stabilizes, or vice versa, capturing more complex decay patterns.

By incorporating these three decay strategies, our DLTD model is versatile and capable of simulating a wide range of real-world influence spread behaviors. The choice of decay strategy can be tailored to the specific characteristics of the social network and the nature of the influence being modeled. This flexibility allows for more accurate and realistic predictions of influence spread, enhancing the applicability of our model in various domains such as viral marketing, information dissemination, and public health campaigns. In general, at time t, an inactive node v becomes active if the total weights of its active incoming neighbors is at least θ_v,

$$\sum_{u \in N_{\text{active}}^{\text{in},t}(v)} w_{u \mapsto v}^{t} \geqslant \theta_v \tag{11}$$

where $N_{\text{active}}^{\text{in},t}(v)$ consists of active incoming neighbors of inactive node v at time t.

This process stops until no more nodes can be activated. In Fig. 2d, at time $t = 2$, the accumulated influences from node 4 and node 5 make node 6 active for the LT model, while in Fig. 2g, it is totally different for the DLTD model due to the decay of influence. The pseudocode for this process is shown in Algorithm 1.

Algorithm 1 Pseudocode for simulating the DLTD

Require: Graph $G(V, E)$, seed set S
Ensure: Set of active nodes S_T
$S_T \leftarrow S$, $s' \leftarrow S$, Snew $\leftarrow \varnothing$, **AW** $\leftarrow \varnothing$;
$t = 0$;
for $v \in V \backslash S$ **do**
  $\theta_v \leftarrow$ random[0, 1];
  $N_{\text{active}}^{\text{in},t}(v) \leftarrow \varnothing$;
while s' != $\varnothing$ **do**
  if t != 0 **do**
  for $u \in s'$ **do**
    for $v \in N^{\text{out}}(u)$ **do**
    if $v \notin S_T$ and $N_{\text{active}}^{\text{in},t}(v)$!= $\varnothing$ **then**
      $w_{u \mapsto v}^{t} = \Phi\left(w_{u \mapsto v}^{t-1}\right)$;
  for $u \in s'$ **do**
  for $v \in N^{\text{out}}(u)$ **do**
    if $v \notin S_T$ **then**
    $N_{\text{active}}^{\text{in},t}(v) \leftarrow N_{\text{active}}^{\text{in},t}(v) \cup \{u\}$;
    AW$[v] = \sum_{u \in N_{\text{active}}^{\text{in},t}(v)} w_{u \mapsto v}^{t}$;
    AW$[v]$ is the accumulated weights of v;
    if AW$[v] \geqslant \theta_v$ **then**
    v becomes active;
    Snew $\leftarrow$ Snew $\cup \{v\}$, $S_T \leftarrow S_T \cup \{v\}$;
  $s' \leftarrow$ Snew;
  t += 1;
return S_T

4 Experiment and Discussion

4.1 Datasets

- **Epinions**[43]: This dataset consists of product reviews with timestamps about 18 000 (18k) reviewers (nodes), and a directed trust network among reviewers. In the network, there exists a direct edge from reviewer B to reviewer A, if A trusts B. The basic statistics of Epinions are listed in Table 1.
- **Last2K-FM**[44]: It contains social networking and music artist listening information with 2k users from Last.fm online music system (http://www.lastfm.com). In this dataset, there are about 92 800 artist listening records from 1892 users. The basic statistics of Last2K-FM in Table 1.
- **Cora**[45]: It is a citation network that consists of 2708 nodes (machine learning papers) and 5429 edges (citation network). The nodes and edges represent papers and the citation relationship between them respectively. Note that Cora is regarded as an undirected graph, and can be downloaded from https://linqs.soe.ucsc.edu/data. The basic statistics of Cora are summarized in Table 2.
- **Citeseer**[45]: It is also a citation network that consists of 3312 nodes and 4732 edges. Citeseer is

Table 1 Summary of the datasets' topological properties and statistics. M denotes million.

Dataset	Number of nodes	Number of directed edges	Number of products/artists	Average degree	Maximum degree
Epinions	18k	1.2M	262k	66.7	4k
Last2K-FM	1892	25 434	17 632	13.4	238

Table 2 Detailed statistics of the citation network datasets.

Dataset	Number of nodes	Number of directed edges	Average degree	Maximum degree
Cora	2708	10 556	3.90	336
Citeseer	3312	9228	2.79	198

regarded as an undirected graph as Cora, and can be downloaded from https://linqs.soe.ucsc.edu/data. The basic statistics of Citeseer is summarized in Table 2.

• **LFR benchmark**[46]: This benchmark can generate artificial networks with community structure. The networks contain $n = 5000$ nodes with average degree $\langle k \rangle$ = 5, 10, 15; maximum degree maxk = 50; minimum community size minc = 10, maximum community size maxc = 50; the mixing parameter μ = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, where $\mu \in [0,1]$, and as μ increases[46], nodes within communities in the networks interact more weakly.

4.2 Baselines

• **IC[23]:** The IC model is a probabilistic model of influence spread. In this model, each activated node has one chance to activate its inactive neighbors. The activation attempt is successful with a predefined probability p. The process unfolds in discrete time steps: At each step, newly activated nodes from the previous step try to activate their neighbors, and the process continues until no more activations occur.

• **Independent Cascade with Invitation (ICI)[47]:** The ICI model is designed to address the limitations of traditional diffusion models by incorporating the effects of invitations on the behavior of individuals in a social network. While traditional models like the IC focus solely on the process of influence spread, the ICI model extends this by including the multi-stage behavior changes triggered by invitations.

• **LT[23]:** The LT model is another classic influence spread model. Each node has a threshold chosen uniformly at random from the interval $[0,1]$. A node becomes activated when the sum of the weights of its activated neighbors exceeds this threshold. The process is deterministic given the thresholds and edge weights. The spread continues until no more activations can occur.

• **LT with Colors (LTC)[11]:** The LTC model introduces three active states for users based on product ratings: adopted, promoted, and inhibited. Seeds are initially assigned to these states randomly. While LTC retains the influence process of the traditional LT model, it modifies the threshold function for inactive users v by combining the ratings from active in-neighbors u and the edge weight $w_{u,v}$. Upon activation, a user can either adopt with a certain probability or act as a message bridge, with the possibility of being promoted or becoming inhibited. The influence spread in LTC is measured by the expected number of adopted users.

• **Feature-aware Threshold Model (FTM)[48]:** The FTM enhances the LT model by incorporating additional factors into the threshold function. These factors include edge weight, the user's positive feelings towards each feature of an item, and their internal resistance to influence. The threshold function is then encapsulated in a logistic function to determine the activation probability. This enables FTM to capture more detailed influence dynamics by considering multiple influencing factors.

• **Continuous-Time Independent Cascade (CT-IC)[49]:** The CT-IC model extends the traditional IC model by incorporating time delays in the information diffusion process. In this model, once a user is activated, they attempt to influence their inactive neighbors based on an independent meeting probability. If a meeting occurs within the predefined deadline τ, the influence process follows the standard IC mechanism, granting the active user a single opportunity to activate their neighbor. The final influence spread in CT-IC is measured as the expected number of activated users before the deadline.

• **IC with Negative (IC-N) opinions[50]:** The IC-N model[9] introduces opinion diversity by distinguishing between positive and negative states. Each seed node can be activated as either positive or negative, with the state evolving through interactions with neighbors. Specifically, if a positive user influences a neighbor, the neighbor may become positive or negative, depending on a predefined probability. The model measures the spread of influence by counting the number of positive users in the network.

4.3 Determination of seed set

To test our model for influence spread, we determine the seed set through three indicators[51].

• **Degree:** Degree is a very effective indicator to measure the importance of a node, and is defined as the number of nodes linking to it[52, 53].

$$k_u = \sum_{v=1}^{n} a_{uv} \tag{12}$$

where k_u represents the degree of node u.

● **Betweenness centrality:** Betweenness centrality of node u can be defined as the sum of the fraction of all-pairs shortest paths that pass through[54–56].

$$\mathrm{bc}_u = \sum_{s,t \in V, s \neq t \neq u} \frac{\sigma_{st,u}}{\sigma_{st}} \tag{13}$$

where bc_u represents the betweenness centrality of node u, σ_{st} represents the number of shortest paths between node s and node t, and $\sigma_{st,u}$ is the number of those paths passing through node u.

● **Closeness centrality:** Closeness centrality measures how close a node in the network is to all other nodes, and can be defined as the average shortest path length from node u to all other reachable nodes in the network[52].

$$\mathrm{cc}_u = \frac{1}{\sum_{v \in V, v \neq u} l_{uv}} (n-1) \tag{14}$$

where cc_u represents the closeness centrality of node u, l_{uv} represents the shortest-path length between node u and node v.

Here, we select seed nodes with higher degree, betweenness centrality, and closeness centrality, and the seed set size is in [10, 20, 30, 40, 50].

4.4 Performance comparison

4.4.1 Experimental results on Epinions and LastFM-2K

In this subsection, we use two real-world datasets (Epinions and LastFM-2K) to evaluate the accuracy of our model for influence spread. The influence weights on all edges are generated by two methods: The uniform method and the random method. For a fair comparison, in each propagation process, we set users who are first to review among their friends as the seed set. The actual influence spread is the number of users who review the product or listen the artist on Epinions or LastFM-2K, respectively. We use our model to calculate the predicted spread, then we compare the predicted spread with the actual spread. In order to better display the experimental results, for the same actual spread, we only report the average predicted spread. In Figs. 3a–3c, the influence weights are generated using the uniform method. We compare the DLTD model with the LT model on Epinions. The green line represents the ideal spread, and we observe that influence spread based on the LT model tends to be more messy as the size of seed set increases, i.e., it overestimates the influence spread. This is consistent with the observations on other datasets[57]. The DLTD model tends to be more concentrated to converge to the ground-truth as the size of seed set increases, especially the DLTD-L that is more accurate for influence spread than the DLTD-R and the DLTD-NL. In Fig. 3b, we set the parameter $\beta = 0.05$ for the DLTD-L. In Figs. 3d–3f, the results on the random method are similar. Similar results are observed on LastFM-2K dataset as shown in Figs. 4a–4f. In Fig. 4d, we set the parameter $\beta = 0.02$ for the DLTD-L. Based on the findings above, we conclude that the DLTD with a linear function (DLTD-L) tends to be the best for influence spread compared with the LT model. In addition, we compared six models: IC, ICI, LTC, FTM, CT-IC, and IC-N. Figures 5a–5f show the propagation effects of DLTD-L in comparison to IC, ICI, LTC, FTM, CT-IC, and IC-N on the Epinions dataset, respectively. Figures 6a–6f, respectively display the propagation effects of DLTD-L compared with IC, ICI, LTC, FTM, CT-IC, and IC-N on the LastFM-2K dataset. It can be observed that our DLTD-L model is better able to fit the actual propagation range.

4.4.2 Experimental results on Cora and Citeseer

In this subsection, we choose seed nodes with higher degree, betweenness centrality, and closeness centrality, respectively, and further study influence spread on two real-world datasets (Cora and Citeseer). The size of seed nodes is in [10, 20, 30, 40, 50]. For the influence weights, the results on the random method are similar to the uniform method, so the results on the random method are omitted.

Figures 7a–7c correspond to the seeds based on degree, betweenness centrality, and closeness centrality for influence spread on Cora, respectively, and Figs. 7d–7f correspond to the seeds based on degree, betweenness centrality, and closeness centrality for influence spread on Citeseer, respectively. We can see that the LT model neglects the decay of influence, and hence obtains a larger area for influence spread. The area based on the model of DLTD-NL strongly decreases. Here we set the parameter $\beta = 0.1$ for the model of DLTD-L.

In Table 3, we investigate influence spread as a function of β based on the model of DLTD-L compared with the LT, IC, and ICI models on Cora and Citeseer. From Table 3, we observe that the area for influence spread decreases as β increases, which is consistent with our expectations.

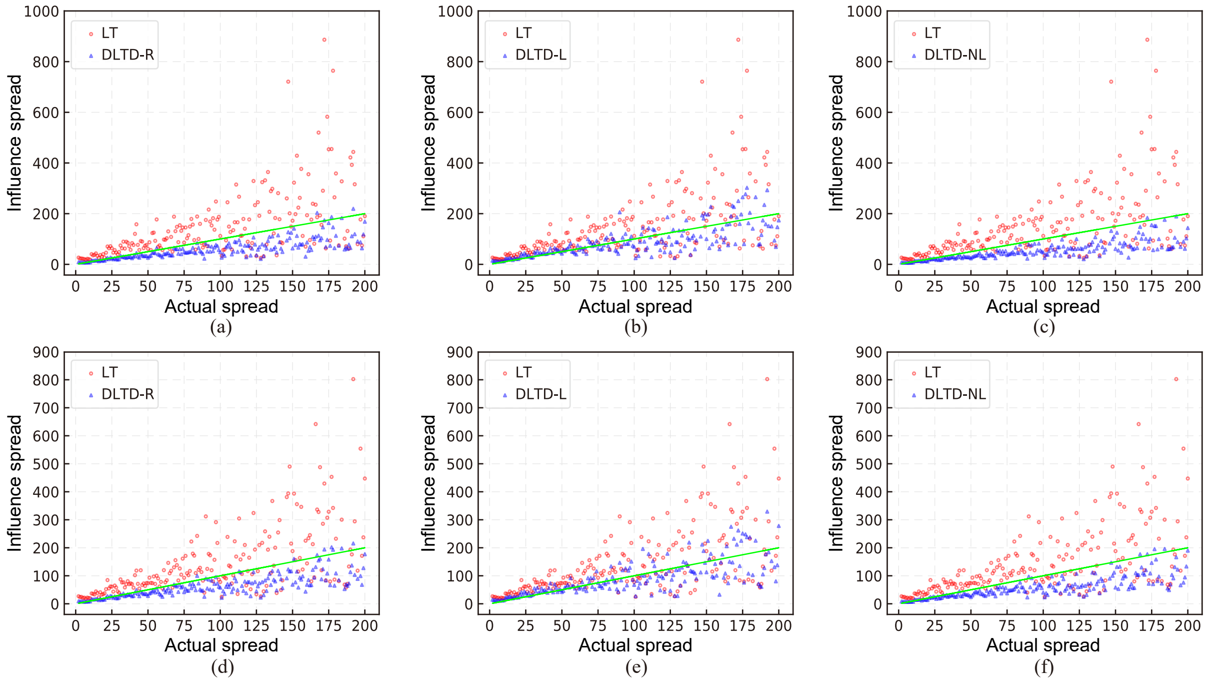


Fig. 3 Comparison of the predicted influence spread with the actual spread based on the models of LT, DLTD-R, DLTD-L, and DLTD-NL on Epinions. (a)–(c) correspond to influence spread with uniform weights on Epinions, and (d)–(f) correspond to influence spread with random weights on Epinions. The ideal spread is shown by green lines.

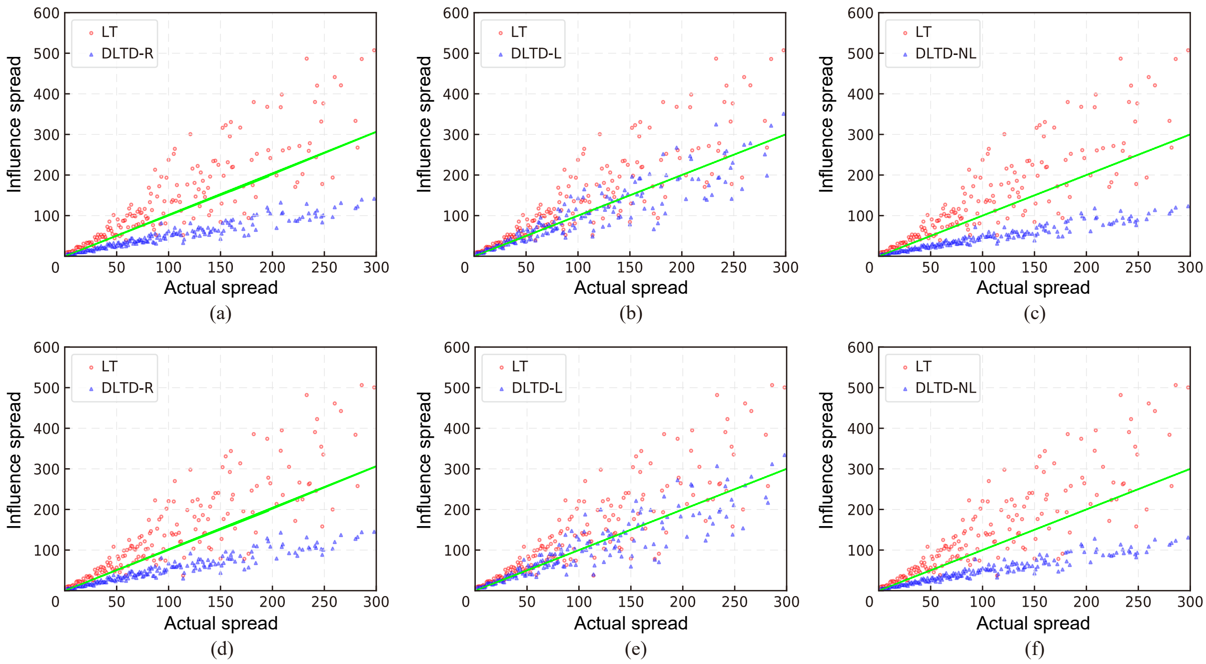


Fig. 4 Comparison of the predicted influence spread with the actual spread based on the models of LT, DLTD-R, DLTD-L, and DLTD-NL on LastFM-2K. (a)–(c) correspond to influence spread with uniform weights on LastFM-2K, and (d)–(f) correspond to influence spread with random weights on LastFM-2K. The ideal spread is shown by green lines.

4.4.3 Experimental results on LFR benchmark

In this subsection, we investigate the effect of community strength (μ) on influence spread based on the models of LT, DLTD-R, DLTD-L, and DLTD-NL, respectively. The nodes with higher closeness are selected as seeds, and we test them on the LFR benchmark[46].

Figures 8a–8c, 8d–8f, 8g–8i, and 8j–8l correspond to

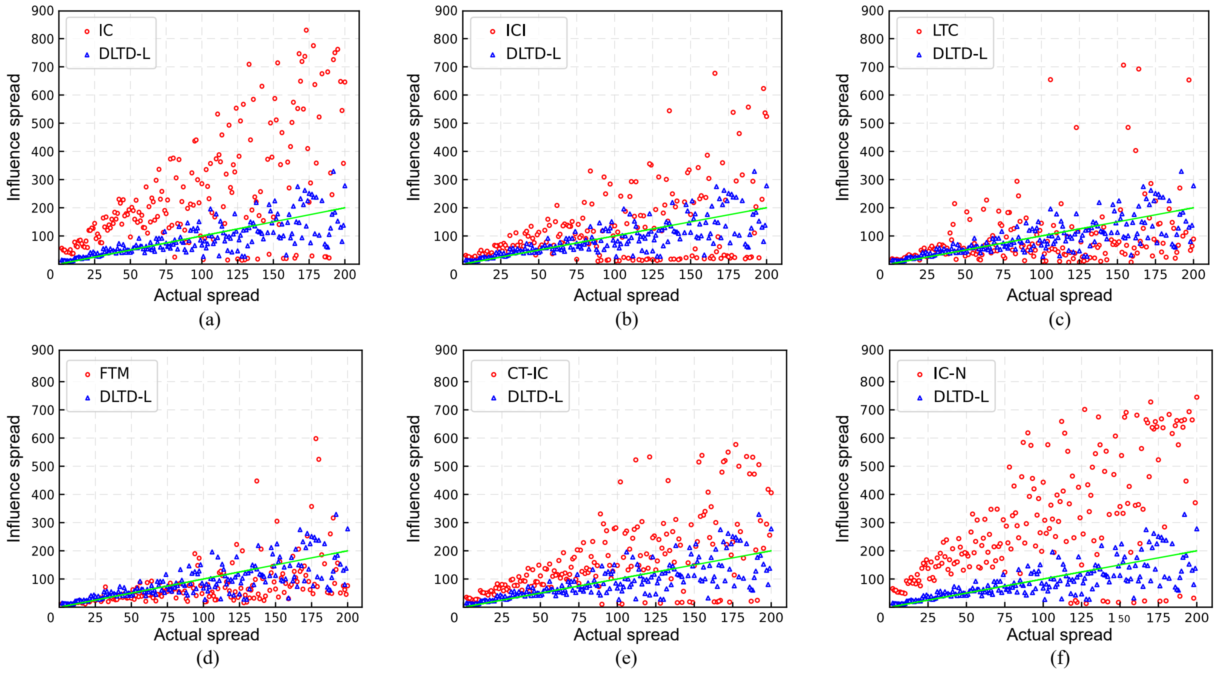


Fig. 5 Comparison of the predicted influence spread with the actual spread based on the models of IC, ICI, LTC, FTM, CT-IC, and IC-N on Epinions. (a)–(f) correspond to influence spread with random weights on Epinions. The ideal spread is shown by green lines.

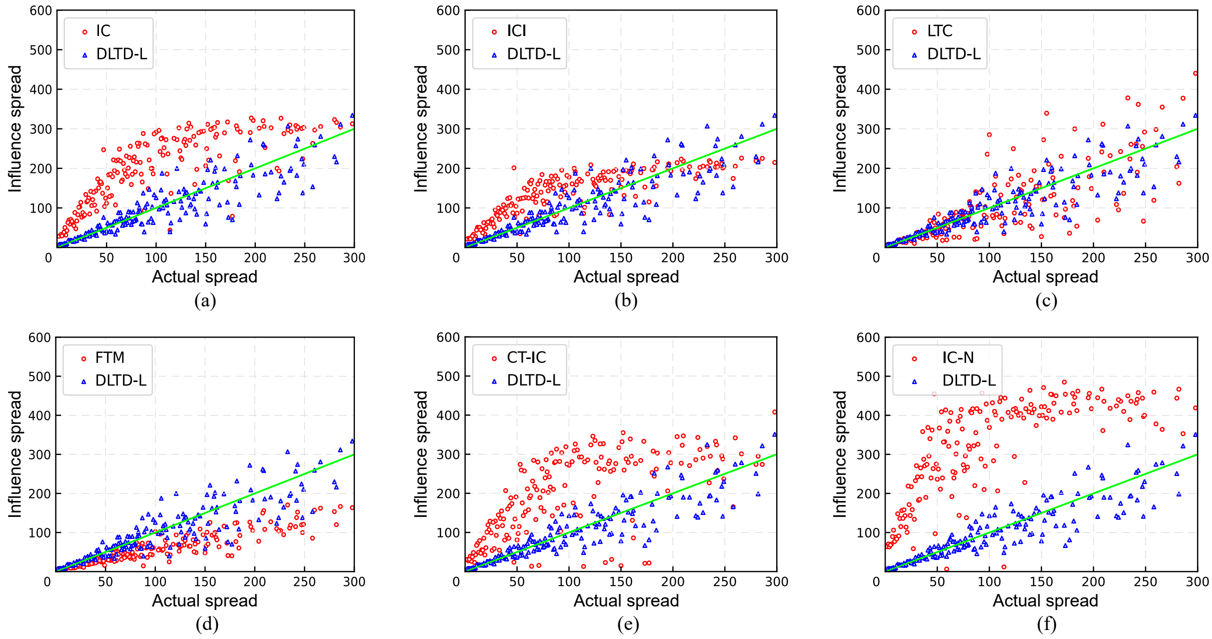


Fig. 6 Comparison of the predicated influence spread with the actual spread based on the models of IC, ICI, LTC, FTM, CT-IC, and IC-N on LastFM-2K. (a)–(f) correspond to influence spread with random weights on LastFM-2K. The ideal spread is shown by green lines.

influence spread based on the models of LT, DLTD-R, DLTD-L, and DLTD-NL, respectively. From Figs. 8a–8c, we can see that μ = 0.5 tends to be the best for influence spread (where μ indicates the networks structures), i.e., it is more effective for influence spread if we keep the balance between communities' internal links and external links in the networks, and strong communities can expedite the spread within communities, while hinder the spread over the whole networks. Similar findings can also be observed on the

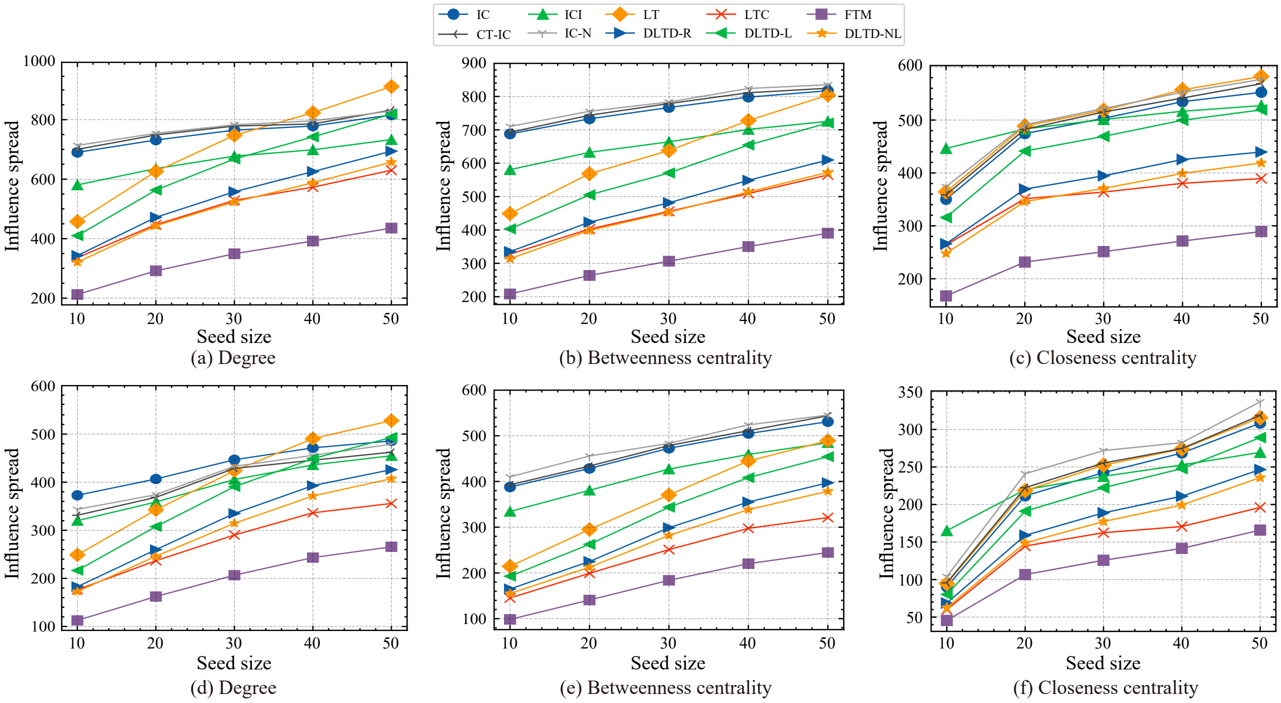


Fig. 7 Influence spread based on the models of IC, ICI, LT, LTC, FTM, CT-IC, IC-N, DLTD-R, DLTD-L, and DLTD-NL. (a)–(c) and (d)–(f) correspond to seed nodes with degree, betweenness centrality, and closeness centrality on Cora and Citeseer, respectively.

models of DLTD-R, DLTD-L, and DLTD-NL in Figs. 8d–8f, 8g–8i, and 8j–8l, respectively. In addition, we observe that compared with the LT model, our model is more stable to the networks' structures (variable μ) when the networks' average degree ($\langle k \rangle$) is fixed.

Figure 9 shows the effect of networks' average degree ($\langle k \rangle$) on influence spread based on the models of LT, DLTD-R, DLTD-L, and DLTD-NL respectively. In Fig. 9a, we can see that influence spread based on the LT model is insensitive to the changes of $\langle k \rangle$, while in the DLTD model, it is not suitable for influence spread as $\langle k \rangle$ increases, especially for DLTD-R and DLTD-NL (see Figs. 9b and 9d, i.e., it is harder to activate nodes on dense networks by considering the decay of influence because more active neighbors are required, and if a node has more neighbors, then each neighbor's influence on the node is weaker.

5 Conclusion

Social networks are becoming more and more popular, and making our life more convenient and efficient. Influence propagates more widely over social networks, and strongly affects people's thoughts, opinions, and even behaviors in human life. Influence spread is a more complicated problem that makes the modeling of this process even more challenging. The LT model provides a simple way to describe the problem. However, this model neglects the decay of individuals' influence as time lapses.

In this paper, we develop a dynamic linear threshold model with decay for influence spread. The results on real-world datasets show that compared with the LT model, the DLTD tends to be more accurate especially the decay with a linear function, and more concentrated to converge to the ground-truth as the size of seed set increases. We also observe that our model is more stable to the networks' structures when the networks' average degree is fixed, and is more sensitive to the networks' density. In addition, an interesting phenomenon indicates that it is more effective for influence spread if we keep the balance between community structures' internal links and external links in the networks.

There are many linear and nonlinear functions that can be used to describe the decay of influence as time lapses. In the future work, we will investigate this problem comprehensively to improve our model's accuracy, and use it to study influence spread on multiplex networks.

Table 3 Sensitivity analysis of β on the DLTD-L model for influence spread and comparison with the LT, IC, and ICI models on Cora and Citeseer.

Indicator	Model	Cora					Citeseer				
		10	20	30	40	50	10	20	30	40	50
Degree	LT	457.43	625.83	747.57	823.46	911.37	248.97	342.57	423.80	490.52	528.13
	IC	690.19	731.51	764.73	778.69	815.95	372.79	406.66	446.65	471.38	485.81
	ICI	580.66	635.52	677.35	698.66	733.06	320.02	358.73	405.56	436.00	455.00
	DLTD-L (β = 0.02)	450.89	606.76	727.39	802.79	888.22	241.20	334.15	418.93	481.03	523.22
	DLTD-L (β = 0.04)	438.27	591.43	708.54	780.29	862.01	229.30	326.79	409.03	475.17	509.82
	DLTD-L (β = 0.06)	424.79	584.41	690.75	767.16	849.41	229.26	319.57	401.14	465.81	505.58
	DLTD-L (β = 0.08)	417.17	569.18	673.32	755.18	836.51	222.69	312.60	395.78	455.52	496.50
	DLTD-L (β = 0.10)	410.15	563.27	667.79	742.13	818.82	216.56	307.74	390.58	448.87	493.29
Betweenness centrality	LT	448.95	568.46	637.84	727.8	804.13	214.48	295.03	370.71	444.76	489.18
	IC	687.76	732.98	766.89	797.94	817.03	387.85	428.42	472.52	504.80	530.67
	ICI	580.63	632.51	664.03	701.32	725.49	334.38	380.73	427.08	459.18	484.73
	DLTD-L (β = 0.02)	437.55	549.07	623.05	699.69	779.49	210.20	285.10	367.74	438.58	486.47
	DLTD-L (β = 0.04)	428.27	535.27	607.30	684.75	764.64	206.02	278.32	362.19	433.81	476.27
	DLTD-L (β = 0.06)	419.97	521.38	592.68	670.35	747.67	198.39	274.58	354.98	423.31	466.23
	DLTD-L (β = 0.08)	407.67	514.62	580.88	663.85	736.09	197.29	269.36	348.79	415.10	463.39
	DLTD-L (β = 0.10)	403.28	505.07	570.78	654.52	720.22	192.77	262.47	343.68	408.38	454.23
Closeness centrality	LT	364.23	488.55	518.45	557.14	582.29	94.65	218.54	252.43	274.02	315.59
	IC	350.13	474.29	503.38	534.73	552.43	90.29	211.51	242.44	268.62	308.39
	ICI	446.36	483.07	501.19	516.58	527.32	165.20	219.83	237.54	252.70	269.57
	DLTD-L (β = 0.02)	347.87	475.95	510.62	539.37	558.29	93.22	209.68	244.09	270.37	311.65
	DLTD-L (β = 0.04)	338.55	464.08	493.88	526.27	545.74	86.63	209.09	236.33	262.20	302.84
	DLTD-L (β = 0.06)	335.37	455.05	482.54	514.96	540.04	86.60	199.66	233.36	255.94	296.99
	DLTD-L (β = 0.08)	324.49	445.22	473.58	506.62	524.28	85.27	197.47	227.00	253.82	292.48
	DLTD-L (β = 0.10)	315.76	441.31	469.13	499.69	519.25	80.24	191.12	222.36	247.85	289.22

Appendix

Appendix A Theoretical Analysis

A1 Impact of Balancing Internal and External Links in Community Structures on Influence Spread Effectiveness

In complex networks, community structures refer to groups of nodes with dense internal connections and sparse external connections. Understanding the impact of community structures on influence spread is crucial for designing effective propagation strategies. Our study reveals that balancing internal and external links can significantly enhance the effectiveness of influence spread, i.e., the coverage. This paper provides a simplified theoretical analysis of this phenomenon.

To theoretically explore the impact of balancing internal and external links on influence spread, we define the following simplified mathematical model:

Network model: Let $G = (V, E)$ represent the graph, where V is the set of nodes and E is the set of edges. The network consists of C communities, each containing n nodes. Let p_{in} denote the probability of forming an internal link within a community, and let p_{out} denote the probability of forming an external link between different communities. Accordingly, the total number of edges can be expressed as

$$m = C \times \frac{n(n-1)}{2} p_{\text{in}} + \frac{C(C-1)}{2} n^2 p_{\text{out}}.$$

Influence spread model: We use the DLTD-L, where the influence decay coefficient is β, and at each time step, the influence decays by $(1-\beta)$. The activation of a node depends on whether the accumulated influence exceeds its threshold θ.

Propagation effectiveness measure: We use two metrics to evaluate propagation effectiveness. The first is influence coverage, denoted by $\sigma(S)$, which represents the number of activated nodes. The second is propagation efficiency, denoted by ϕ, which is

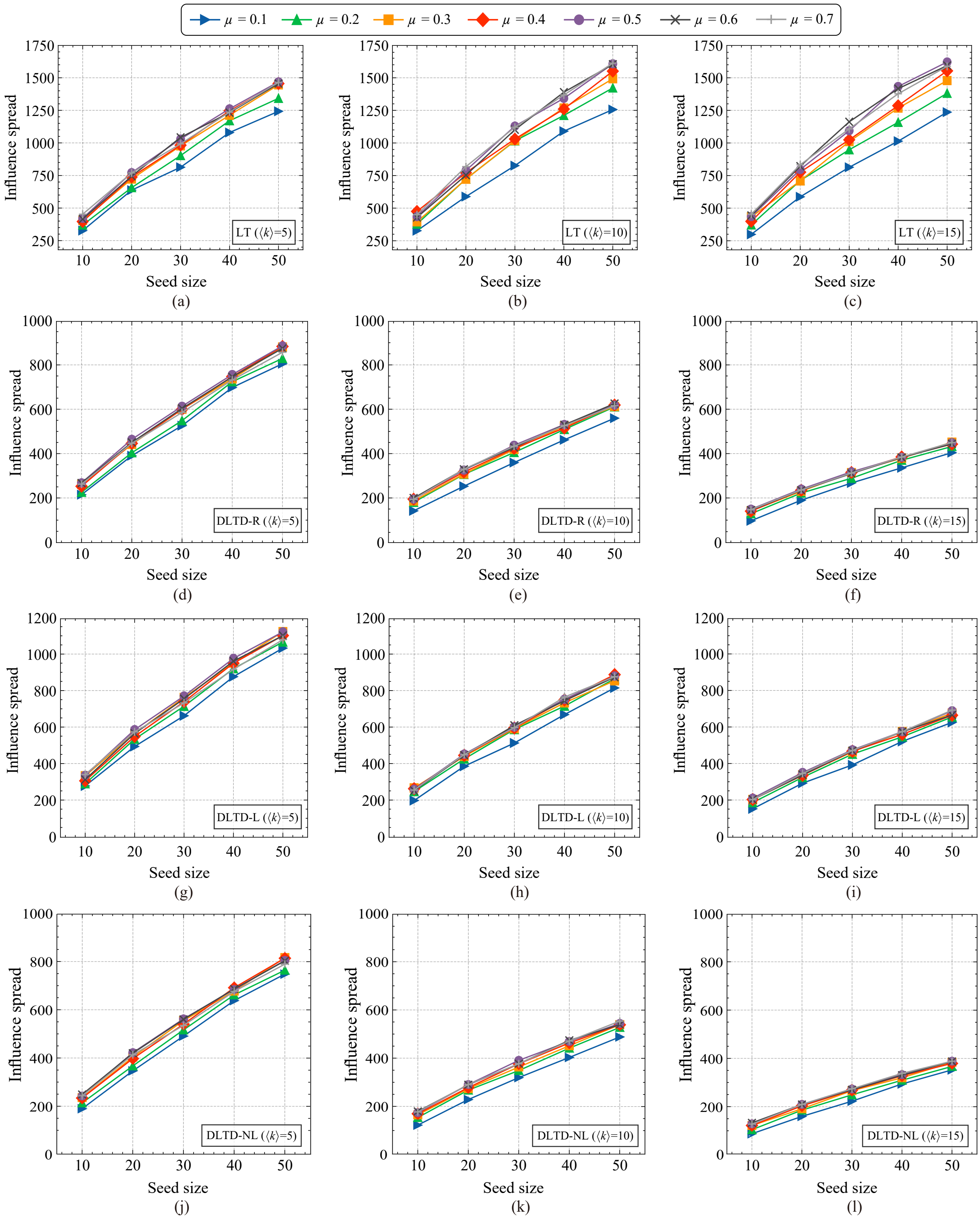


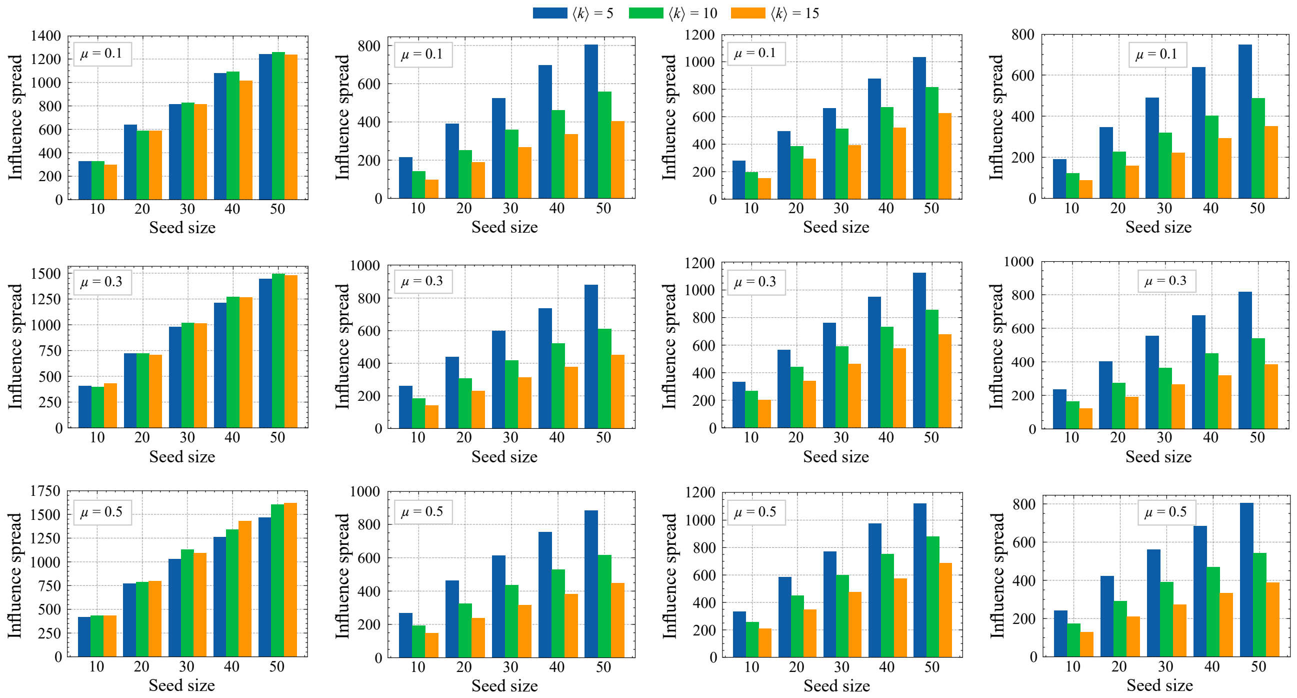
Fig. 8 Sensitivity analysis of networks' structures (μ) on influence spread. (a)–(c), (d)–(f), (g)–(i), and (j)–(l) correspond to the models of LT, DLTD-R, DLTD-L, and DLTD-NL for the networks generated by the LFR benchmark, respectively.

defined as the ratio of influence coverage to the number of time steps τ, i.e.,

$$\phi = \frac{\sigma(S)}{\tau},$$

where τ is the propagation time.

Internal and external link ratios: Define the internal link ratio $\rho_{\text{int}} = \dfrac{p_{\text{in}}}{p_{\text{in}} + p_{\text{out}}}$ and the external link ratio $\rho_{\text{out}} = 1 - \rho_{\text{int}}$.



(a) LT (b) DLTD-R (c) DLTD-L (d) DLTD-NL

Fig. 9 Sensitivity analysis of networks' average degree ($\langle k \rangle$) on influence spread. (a–d) correspond to the models of LT, DLTD-R, DLTD-L, and DLTD-NL for the networks generated by the LFR benchmark, respectively.

A2 Mathematical Modeling of Influence Effectiveness and Link Ratios

To analyze the impact of ρ_{int} on propagation effectiveness ϕ, we construct the following model:

Influence coverage: The influence coverage of the seed set S can be expressed as

$$\sigma(S) = \sum_{c=1}^{C} \sigma_c(S),$$

where

$$\sigma_c(S) = n\left(1 - \mathrm{e}^{-\lambda_c |S_c|}\right),$$

with $\lambda_c = \alpha\rho_{\mathrm{int}} + \gamma\rho_{\mathrm{out}} = \alpha\rho_{\mathrm{int}} + \gamma(1-\rho_{\mathrm{int}})$, where α and γ represent the propagation contributions from internal and external links.

Propagation efficiency: Based on the above influence coverage, the propagation efficiency can be expressed as

$$\phi(\rho_{\mathrm{int}}) = \frac{\sigma(S)}{\tau} = \sum_{c=1}^{C} n\left(1 - \mathrm{e}^{-\lambda_c |S_c|}\right) \cdot \lambda_c.$$

Assuming that the propagation time τ is inversely proportional to λ_c, we get

$$\phi(\rho_{\mathrm{int}}) = \sum_{c=1}^{C} n\left(1 - \mathrm{e}^{-\lambda_c |S_c|}\right) \cdot \lambda_c.$$

To find the optimal internal link ratio ρ^*_{int} that maximizes $\phi(\rho_{\mathrm{int}})$, we optimize the function as follows:

$$\phi(\rho_{\mathrm{int}}) = \sum_{c=1}^{C} n\left(1 - \mathrm{e}^{-[\alpha\rho_{\mathrm{int}}+\gamma(1-\rho_{\mathrm{int}})]|S_c|}\right) \cdot [\alpha\rho_{\mathrm{int}} + \gamma(1-\rho_{\mathrm{int}})]$$

Taking the derivative with respect to ρ_{int}, we find

$$\frac{\mathrm{d}\phi}{\mathrm{d}\rho_{\mathrm{int}}} = (\alpha - \gamma) \sum_{c=1}^{C} n\left[|S_c| \mathrm{e}^{-\lambda_c |S_c|} \lambda_c + 1 - \mathrm{e}^{-\lambda_c |S_c|}\right].$$

Setting the derivative equal to zero, we solve for ρ^*_{int}. Since the function is typically unimodal, the solution ρ^*_{int} is unique, and we conclude that there exists an optimal internal link ratio that maximizes the propagation efficiency ϕ.

Through the simplified theoretical analysis, we demonstrate that balancing internal and external links is crucial for optimizing influence spread effectiveness. There exists an optimal internal link ratio ρ^*_{int} where the influence spread reaches its maximum, ensuring rapid spread within communities while maintaining broad coverage across communities.

Appendix B Submodularity

Submodularity is a crucial property for ensuring the approximation guarantees in influence maximization algorithms, which typically rely on this property for providing near-optimal solutions. However, the primary focus of the DLTD model is not on influence maximization but rather on the accuracy of influence spread prediction. While the absence of submodularity means that we cannot provide formal approximation guarantees, our experiments demonstrate that the DLTD model provides more accurate and stable predictions of influence spread compared to traditional models, especially when considering the time-dependent decay of influence. We believe that the model's ability to better reflect real-world influence propagation, despite lacking submodularity, justifies its use for tasks focused on accurately predicting influence spread rather than maximizing it. The improved realism in capturing decay dynamics is the key strength of the DLTD model, making it particularly useful for scenarios where precise influence spread prediction is critical. In addition to this, we find that under certain conditions, DLTD satisfies the submodularity.

Proof In the DLTD-L (linear decay) model, the influence from an active node u to an inactive node v at time t decays linearly over time. The influence from node u to node v is given by

$$w^{t}_{u\mapsto v} = (1-\beta)w^{(t-1)}_{u\mapsto v} \tag{15}$$

$w^{(0)}_{u\mapsto v}$ is the initial weight of influence, and β is the decay rate (with $0<\beta<1$) and β is fixed.

At each time step, an inactive node v becomes active if the sum of the weighted influence from its active neighbors exceeds its threshold θ_v, i.e.,

$$\sum_{u\in N^{\text{in},t}_{\text{active}}(v)} w^{t}_{u\mapsto v} \geqslant \theta_v \tag{16}$$

where $N^{\text{in},t}_{\text{active}}(v)$ is the set of active neighbors of node v at time t.

Now, to prove that the influence spread in the DLTD-L model is submodular, we need to show that the marginal gain in influence spread from adding a new node to a set decreases as the set grows.

Let S be the set of initially active nodes, and let $\sigma(S)$ denote the influence spread from this set. We want to show that the marginal gain in influence spread decreases as the set grows. That is, we need to prove that

$$\Delta P = \sigma(P\cup\{v\}) - \sigma(P) \geqslant \sigma(Q\cup\{v\}) - \sigma(Q) = \Delta Q \tag{17}$$

for any two seed sets $P, Q\subseteq V$ with $P\subseteq Q$, and any node $v\notin Q$.

The marginal gain in influence spread by adding node v to set P is

$$\Delta P = \sigma(P\cup\{v\}) - \sigma(P) \tag{18}$$

and the marginal gain in influence spread by adding node v to set Q is

$$\Delta Q = \sigma(Q\cup\{v\}) - \sigma(Q) \tag{19}$$

Since the influence from node v to other nodes decays over time, the marginal gain ΔP is larger when the set is smaller because fewer nodes are already activated to influence the activation of other nodes. As the set Q grows, the new nodes have already been influenced by a larger number of active nodes, leading to a smaller marginal increase in the influence spread. Thus, we have

$$\Delta P \geqslant \Delta Q \tag{20}$$

which satisfies the condition for submodularity. ■

Therefore, we conclude that the influence spread in the DLTD-L model is submodular, as the marginal gain in influence spread decreases as the set of active nodes grows. This implies that the DLTD-L model satisfies submodularity, and thus, greedy algorithms can be used to approximate the maximum influence spread effectively.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No. 61872432) and the Shaanxi Fundamental Science Research Project for Mathematics and Physics (No. 22JSY024).

References

[1] P. Domingos and M. Richardson, Mining the network value of customers, in *Proc. 7th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2001, pp. 57–66.

[2] M. Richardson and P. Domingos, Mining knowledge-sharing sites for viral marketing, in *Proc. 8th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, Edmonton, Canada, 2002, pp. 61–70.

[3] H. Huang, H. Shen, Z. Meng, H. Chang, and H. He, Community-based influence maximization for viral marketing, *Appl. Intell.*, vol. 49, no. 6, pp. 2137–2150, 2019.

[4] P. Huang, L. Guo, and Y. Zhong, Efficient algorithms for maximizing group influence in social networks, *Tsinghua Science and Technology*, vol. 27, no. 5, pp. 832–842, 2022.

[5] X. Song, Y. Chi, K. Hino, and B. L. Tseng, Information flow modeling based on diffusion rate for prediction and ranking, in *Proc. 16th Int. Conf. World Wide Web*, Banff,

Canada, 2007, pp. 191–200.

[6] C. Budak, D. Agrawal, and A. El Abbadi, Limiting the spread of misinformation in social networks, in *Proc. 20th Int. Conf. World Wide Web*, Hyderabad, India, 2011, pp. 665–674.

[7] Y. Li, Z. Fan, X. Yuan, and X. Zhang, Recognizing fake information through a developed feature scheme: A user study of health misinformation on social media in China, *Inf. Process. Manag.*, vol. 59, no. 1, p. 102769, 2022.

[8] A. Yang, X. Huang, X. Cai, X. Zhu, and L. Lu, ILSR rumor spreading model with degree in complex network, *Phys. A: Stat. Mech. Appl.*, vol. 531, p. 121807, 2019.

[9] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts, Everyone's an influencer: Quantifying influence on twitter, in *Proc. 4th ACM Int. Conf. Web Search and Data Mining*, Hong Kong, China, 2011, pp. 65–74.

[10] J. Weng, E. P. Lim, J. Jiang, and Q. He, TwitterRank: Finding topic-sensitive influential twitterers, in *Proc. 3rd ACM Int. Conf. Web Search and Data Mining*, New York, NY, USA, 2010, pp. 261–270.

[11] S. Bhagat, A. Goyal, and L. V. S. Lakshmanan, Maximizing product adoption in social networks, in *Proc. 5th ACM Int. Conf. Web Search and Data Mining*, Seattle, WA, USA, 2012, pp. 603– 612.

[12] J. E. Lönnqvist and F. große Deters, Facebook friends, subjective well-being, social support, and personality, *Comput. Human Behav.*, vol. 55, pp. 113–120, 2016.

[13] D. Centola, The spread of behavior in an online social network experiment, *Science*, vol. 329, no. 5996, pp. 1194–1197, 2010.

[14] P. G. Sun, X. Wu, Y. Quan, and Q. Miao, Influence percolation method for overlapping community detection, *Phys. A*: *Stat. Mech. Appl.*, vol. 596, p. 127103, 2022.

[15] P. G. Sun, X. Wu, Y. Quan, and Q. Miao, Rearranging 'indivisible' blocks for community detection, *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 6, pp. 6252–6263, 2023.

[16] L. Wang, Y. Zhu, X. Wan, Y. Zhang, S. Feng, C. Li, J. Hou, Y. Wei, F. Zhang, and B. Hu, Rapid identification of critical states in complex biological processes based on single-sample communityd detection, *Tsinghua Science and Technology*, DOI: 10.26599/TST.2025.9010076.

[17] J. Xie, X. Wang, Y. Liu, W. Gong, C. Yan, W. Rafique, M. Khan, and A. A. Khan, Social media-driven user community finding with privacy protection, *Tsinghua Science and Technology*, vol. 30, no. 4, pp. 1782–1792, 2025.

[18] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, Models of social networks based on social distance attachment, *Phys. Rev. E*, vol. 70, no. 5, p. 056122, 2004.

[19] Y. Li, J. Fan, Y. Wang, and K. L. Tan, Influence maximization on social graphs: A survey, *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 10, pp. 1852–1872, 2018.

[20] T. Lev, I. Ben-Gal, and E. Shmueli, Influence maximization through scheduled seeding in a real-world setting, *IEEE Trans. Computat. Soc. Syst.*, vol. 9, no. 2, pp. 494–507, 2022.

[21] F. Kazemzadeh, A. A. Safaei, and M. Mirzarezaee, Influence maximization in social networks using effective community detection, *Phys. A*: *Stat. Mech. Appl.*, vol. 598, p. 127314, 2022.

[22] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, *Rev. Mod. Phys.*, vol. 87, no. 3, pp. 925–979, 2015.

[23] D. Kempe, J. Kleinberg, and É. Tardos, Maximizing the spread of influence through a social network, in *Proc. 9th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, Washington, DC, USA, 2003, pp. 137–146.

[24] Z. Aghaee and S. Kianian, Efficient influence spread estimation for influence maximization, *Soc. Netw. Anal. Min.*, vol. 10, no. 1, p. 81, 2020.

[25] J. Guo and W. Wu, Influence maximization: Seeding based on community structure, *ACM Trans. Knowl. Discov. Data*, vol. 14, no. 6, p. 66, 2020.

[26] D. Liben-Nowell and J. Kleinberg, Tracing information flow on a global scale using internet chain-letter data, *Proc. Natl. Acad. Sci. USA*, vol. 105, no. 12, pp. 4633–4638, 2008.

[27] N. Pathak, A. Banerjee, and J. Srivastava, A generalized linear threshold model for multiple cascades, in *Proc. 2010 IEEE Int. Conf. Data Mining*, Sydney, Australia, 2010, pp. 965–970.

[28] J. Yang and J. Leskovec, Modeling information diffusion in implicit networks, in *Proc. 2010 IEEE Int. Conf. Data Mining*, Sydney, Australia, 2010, pp. 599– 608.

[29] C. Lagnier, L. Denoyer, E. Gaussier, and P. Gallinari, Predicting information diffusion in social networks using content and user's profiles, in *Proc. 35th European Conf. IR Research*, Sydney, Australia, 2013, pp. 74–85.

[30] R. Olivares, F. Muñoz, and F. Riquelme, A multi-objective linear threshold influence spread model solved by swarm intelligence-based methods, *Knowl.-Based Syst.*, vol. 212, p. 106623, 2021.

[31] W. Chen, Y. Yuan, and L. Zhang, Scalable influence maximization in social networks under the linear threshold model, in *Proc. 2010 IEEE Int. Conf. Data Mining*, Sydney, Australia, 2010, pp. 88–97.

[32] A. Goyal, W. Lu, and L. V. S. Lakshmanan, SIMPATH: An efficient algorithm for influence maximization under the linear threshold model, in *Proc. 11th Int. Conf. Data Mining*, Vancouver, Canada, 2011, pp. 211– 220.

[33] M. Samadi, R. Nagi, A. Semenov, and A. Nikolaev, Seed activation scheduling for influence maximization in social networks, *Omega*, vol. 77, pp. 96–114, 2018.

[34] M. Kimura, K. Saito, and R. Nakano, Extracting influential nodes for information diffusion on a social network, in *Proc. 22nd National Conf. Artificial Intelligence - Volume 2*, Vancouver, Canada, 2007, pp. 1371–1376.

[35] S. Banerjee, M. Jenamani, and D. K. Pratihar, A survey on influence maximization in a social network, *Knowl. Inf. Syst.*, vol. 62, no. 9, pp. 3417–3455, 2020.

[36] S. Li, F. Kong, K. Tang, Q. Li, and W. Chen, Online influence maximization under linear threshold model, in *Proc. 34th Int. Conf. Neural Information Processing Systems*, Sydney, Australia, 2020, p. 101.

[37] B. Xu and L. Liu, Information diffusion through online social networks, in *Proc. 2010 IEEE Int. Conf. Emergency*

Management and Management Sciences, Beijing, China, 2010, pp. 53–56.
[38] C. C. Zou, D. Towsley, and W. Gong, Modeling and simulation study of the propagation and defense of internet e-mail worms, *IEEE Trans. Depend. Secure Comput.*, vol. 4, no. 2, pp. 105–118, 2007.
[39] J. Berger and K. L. Milkman, What makes online content viral, *J. Mark. Res.*, vol. 49, no. 2, pp. 192–205, 2012.
[40] D. J. Watts and P. S. Dodds, Influentials, networks, and public opinion formation, *J. Consum. Res.*, vol. 34, no. 4, pp. 441–458, 2007.
[41] S. Aral and D. Walker, Identifying influential and susceptible members of social networks, *Science*, vol. 337, no. 6092, pp. 337–341, 2012.
[42] W. Chen, C. Wang, and Y. Wang, Scalable influence maximization for prevalent viral marketing in large-scale social networks, in *Proc. 16th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, Washington, DC, USA, 2010, pp. 1029–1038.
[43] J. Tang, H. Gao, H. Liu, and A. Das Sarma, eTrust: Understanding trust evolution in an online world, in *Proc. 18th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, Beijing, China, 2012, pp. 253–261.
[44] I. Cantador, P. Brusilovsky, and T. Kuflik, Second workshop on information heterogeneity and fusion in recommender systems (HetRec2011), in *Proc. 5th ACM Conf. Recommender Systems*, Chicago, IL, USA, 2011, pp. 387–388.
[45] Q. Lu and L. Getoor, Link-based classification, in *Proc. 20th Int. Conf. Int. Conf. Machine Learning*, Washington, DC, USA, 2003, pp. 496–503.
[46] A. Lancichinetti, S. Fortunato, and F. Radicchi, Benchmark graphs for testing community detection algorithms, *Phys. Rev. E*, vol. 78, no. 4, p. 046110, 2008.
[47] S. Zhang, J. Sun, W. Lin, X. Xiao, Y. Huang, and B. Tang, Information diffusion meets invitation mechanism, in *Proc. ACM on Web Conf. 2024*, Singapore, 2024, pp. 383–392.
[48] N. Barbieri and F. Bonchi, Influence maximization with viral product design, in *Proc. 2014 SIAM Int. Conf. Data Mining*, Philadelphia, PA, USA, 2014, pp. 55–63.
[49] B. Liu, G. Cong, D. Xu, and Y. Zeng, Time constrained influence maximization in social networks, in *Proc. 12th Int. Conf. Data Mining*, Brussels, Belgium, 2012, pp. 439–448.
[50] W. Chen, A. Collins, R. Cummings, T. Ke, Z. Liu, D. Rincon, X. Sun, Y. Wang, W. Wei, and Y. Yuan, Influence maximization in social networks when negative opinions may emerge and propagate, in *Proc. 11th SIAM Int. Conf. Data Mining*, Mesa, AZ, USA, 2011, pp. 379–390.
[51] P. G. Sun, W. Che, Y. Quan, S. Wang, and Q. Miao, Random networks are heterogeneous exhibiting a multi-scaling law, *Phys. A: Stat. Mech. Appl.*, vol. 587, p. 126479, 2022.
[52] L. C. Freeman, Centrality in social networks conceptual clarification, *Soc. Netw.*, vol. 1, no. 3, pp. 215–239, 1978.
[53] N. E. Friedkin, Theoretical foundations for centrality measures, *Am. J. Sociol.*, vol. 96, no. 6, pp. 1478–1504, 1991.
[54] U. Brandes, On variants of shortest-path betweenness centrality and their generic computation, *Soc. Netw*, vol. 30, no. 2, pp. 136–145, 2008.
[55] L. C. Freeman, A set of measures of centrality based on betweenness, *Sociometry*, vol, 40, no. 1, pp. 35–41, 1977.
[56] U. Brandes, A faster algorithm for betweenness centrality, *J. Math. Sociol.*, vol. 25, no. 2, pp. 163–177, 2001.
[57] A. Goyal, F. Bonchi, and L. V. S. Lakshmanan, A data-based approach to social influence maximization, *Proc. VLDB Endow.*, vol. 5, no. 1, pp. 73–84, 2011.

Xunlian Wu received the PhD degree in computer science and technology from Xidian University, Xi'an, China in 2025. He is currently a lecturer at School of Computer Science and Artificial Intelligence, Zhengzhou University, Zhengzhou, China. His research interests include graph machine learning, community detection, and clustering analysis.

Yining Quan received the PhD degree in cryptography from Xidian University, Xi'an, China in 2010. He is currently a professor at School of Computer Science and Technology, Xidian University, Xi'an, China. His research interests include data mining, social network analysis, and visualization and visual analysis of big data.

Anqi Zhang received the MS degree in biomedical engineering from Northwestern Polytechnical University, Xi'an, China in 2020. She is currently a lecturer at School of Computer and Artificial Intelligence, Henan Finance University, Zhengzhou, China. Her research interests include deep learning and graph clustering.

Qiguang Miao received the PhD degree in computer application technology from Xidian University, Xi'an, China in 2005. He is currently a professor at School of Computer Science and Technology, Xidian University, Xi'an, China. His research interests include machine learning, intelligent image processing, and malware behavior analysis and understanding.

Penggang Sun received the PhD degree in computer science and technology from Xidian University, Xi'an, China in 2011. He is currently a professor at School of Computer Science and Technology, Xidian University, Xi'an, China. His research interests include data mining, pattern recognition, graph theory, and complex network analysis.