Knowledge-Based Systems 213 (2021) 106600

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Competitive and complementary influence maximization in social network: A follower's perspective

Huimin Huang [a], Zaiqiao Meng [b,*], Hong Shen [c,d]

[a] Oujiang College, Wenzhou University, Wenzhou, China
[b] University of Cambridge, Cambridge, United Kingdom
[c] School of Computer Science, University of Adelaide, Adelaide, Australia
[d] School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China

ARTICLE INFO

Article history:
Received 5 August 2019
Received in revised form 26 September 2020
Accepted 5 November 2020
Available online 13 November 2020

Keywords:
Influence maximization
Recurrent neural networks
Social networks

ABSTRACT

The problem of influence maximization is to select a small set of seed users in a social network to maximize the spread of influence. Recently, this problem has attracted considerable attention due to its applications in both commercial and social fields, such as product promotion, contagion prevention and public opinion forecasting. Most of prior work focuses on the diffusion model of single propagating entity, purely-complementary entities or purely-competitive entities. However, in reality, the influence diffusion in the social network is certainly more general, involving multiple propagating entities, which are competitive or complementary rather than single entity, purely-complementary entities or purely-competitive entities.

In this paper, we consider the problem that a company (follower) intends to promote a new product into the market by maximizing the influence of a social network, where multiple competitive and complementary products have been spreading. We propose a Competitive and Complementary Independent Cascade (CCIC) diffusion model, and propose a novel optimization problem, follower-based influence maximization that aims to select top-K influential nodes as seed nodes, which can maximize the influence of a social network where multiple competitive and complementary products have already been propagated. To solve follower-based influence maximization problem, we propose a Deep Recursive Hybrid model (DRH) and an approximation algorithm (DRHGA). The DRH model dynamically tracks entity correlations, cascade correlations, causalities between ratings and next-period adoption through a deep recursive network and computes influence probabilities between nodes on target product. Then, with the influence probabilities predicted through DRH model, the DRHGA algorithm can efficiently find the seed node set for the target product under the CCIC diffusion model. Experimental results conducted on several public datasets show that our method outperforms the state-of-the-art methods on prediction accuracy and efficiency.

1. Introduction

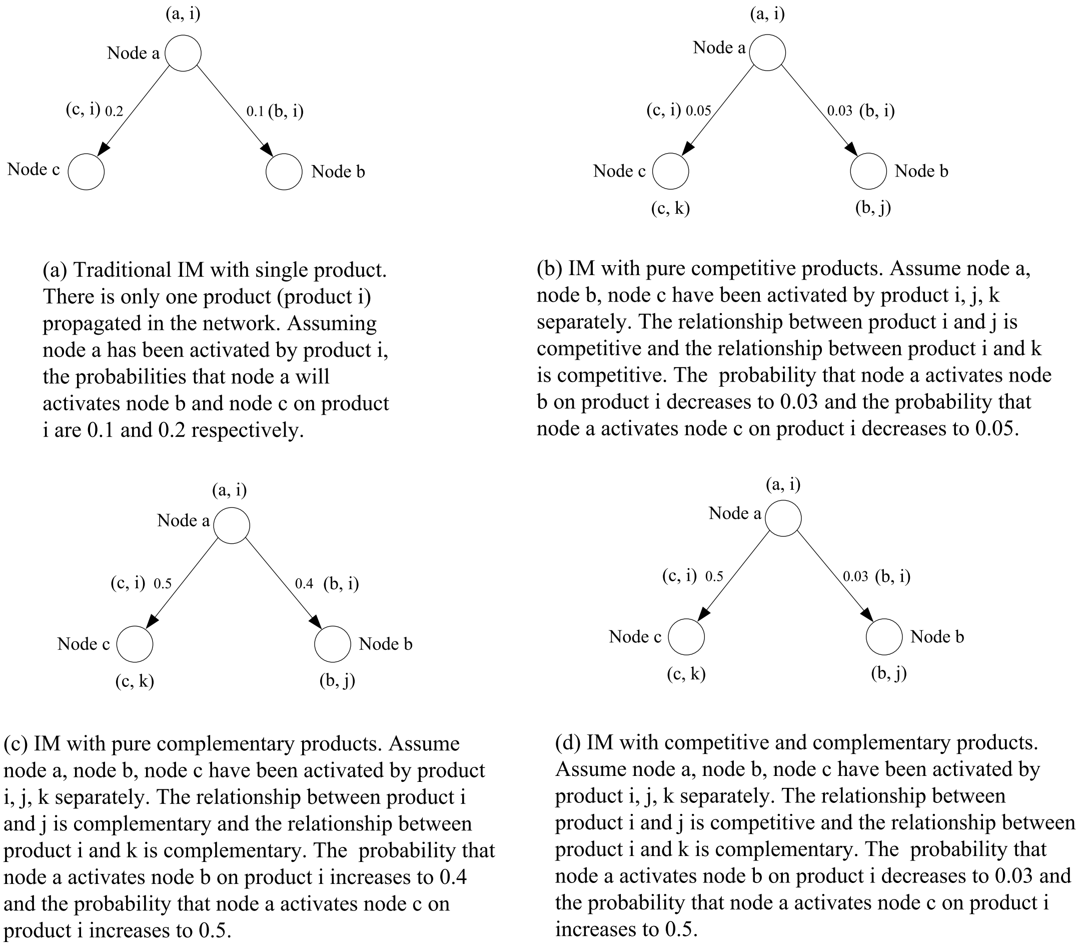
Human beings, as a kind of social animal, involve various interactive behaviors, such as listening music and purchasing products, which can affect the behaviors of others. The study of diffusion influence in social networks can benefit a number of applications, such as promoting a new product into the social network, predicting the spreading of rumors [1]. The influence maximization problem aims to select top-K influential nodes such that the spread of influence can be maximized. With the rapid development of social network platforms, the theory of maximizing influence has been widely used in practice. For example, e-commerce sites have achieve many successes by exploring the viral marketing based on influence maximization techniques in social networks [2].

Recently, a number of influence maximization algorithms have be proposed, most of which focus on the diffusion model of single propagating entity [3–9]. Also, considerable work has been done to extend single-entity influence maximization to pure competitive influence maximization [10–14] and pure complementary influence maximization [15]. However, these models and algorithms ignore more complex social interactions among multiple propagating entities that involve both competition and complementarity.

Fig. 1 presents how the Influence Maximization (IM) with competitive and complementary entities differs from traditional IM with single-entity, IM with pure complementary entities, and

[*] Corresponding author.
E-mail address: zm324@cam.ac.uk (Z. Meng).

https://doi.org/10.1016/j.knosys.2020.106600
0950-7051/© 2020 Elsevier B.V. All rights reserved.



(a) Traditional IM with single product. There is only one product (product i) propagated in the network. Assuming node a has been activated by product i, the probabilities that node a will activates node b and node c on product i are 0.1 and 0.2 respectively.

(b) IM with pure competitive products. Assume node a, node b, node c have been activated by product i, j, k separately. The relationship between product i and j is competitive and the relationship between product i and k is competitive. The probability that node a activates node b on product i decreases to 0.03 and the probability that node a activates node c on product i decreases to 0.05.

(c) IM with pure complementary products. Assume node a, node b, node c have been activated by product i, j, k separately. The relationship between product i and j is complementary and the relationship between product i and k is complementary. The probability that node a activates node b on product i increases to 0.4 and the probability that node a activates node c on product i increases to 0.5.

(d) IM with competitive and complementary products. Assume node a, node b, node c have been activated by product i, j, k separately. The relationship between product i and j is competitive and the relationship between product i and k is complementary. The probability that node a activates node b on product i decreases to 0.03 and the probability that node a activates node c on product i increases to 0.5.

Fig. 1. A simple example to describe how the Influence Maximization (IM) with competitive and complementary entities differs from traditional single-entity IM, pure competitive IM, pure complementary IM.

IM with pure competitive entities. Suppose we have a toy network with three nodes a, b & c and two edges $\langle a, b\rangle$ and $\langle a, c\rangle$. In the traditional single-entity IM (Fig. 1(a)), there is only one entity, product i, propagated in the network. Node b can be activated by product i through the influence from node a to node b, and node c can be activated by product i through the influence from node a to node c. In the IM with pure competitive entities (Fig. 1(b)), assume nodes a, b and c have been activated by products i, j and k respectively, where products i and j are competitive and products i and k are also competitive. From node a, the activation probabilities of product i through edges $\langle a, b\rangle$ and $\langle a, c\rangle$ will be weaken due to the competitive relationship among the three products. In the IM with pure complementary entities (Fig. 1(c)), assume nodes a, b and c have been activated by products i, j and k respectively. The relationship between product i and j is complementary and the relationship between product i and k is also complementary. The activation probability of node b on product i will be strengthened due to the complementary relationship between product i and j, and similarly, the activation probability of node b on entity i will be strengthened because of the complementary relationship between product i and k. In the IM with competitive and complementary entities (Fig. 1(d)), assume nodes a, b and c have been activated by products i, j and k respectively. The relationship between product i and j is competitive while the relationship between product i and k is complementary. The activation probability of node b on product i will be weakened due to the competitive relationship between product i and j, while the activation probability of node b on entity i will be strengthened because of the complementary relationship between product i and k.

Recently, competitive and complementary influence maximization problem has been studied in literature [16–18]. Their models achieved many success in the competitive and complementary IM, however, suffering from a number of major drawbacks: (a) they only focus on maximizing the influence spread on all competitive and complementary products in the network, without considering the problem from the perspective of the follower who intends to maximize influence spread on the target product in the social network where competitive and complementary products have being propagated; (b) most of the existing methods only take the interactions between two products into account, resulting in that they ignore more complex relationship between target product and entire activation sequence for individual user; (c) they ignore the side information such as user rating history, while rating information plays an important role in affecting user's next-period purchase.

Unlike the existing methods focusing on maximizing the influence spread on all products in the social network where competitive and complementary products are propagated, in this paper, we study the problem from the perspective of the follower who intends to maximize influence spread on the target product in the social network where competitive and complementary products have being propagated. The lack of models systematically describing influence diffusion of competitive and complementary products motivated us to propose a more powerful and expressive diffusion model to reasonably represent not only correlations of various products but also influence probabilities between two nodes on different products. Therefore, we propose a Competitive and Complementary Independent Cascade (CCIC) diffusion model and the follower-based influence maximization problem under CCIC diffusion model. We design an approach to capture dynamic influence probabilities between nodes, and supply a greedy approximation algorithm to solve the follower-based influence maximization problem.

How to effectively prompt a new product into a network when other competitive or complementary products have already been introduced into it is still a challenge [18]. In particular, modeling entity interactions and how these interactions change with time of new cascade generating, is extremely challenging. Not only is there potentially significant interactions between any two entities, but also these interactions could change over time when a user is activated by either entity through links in network. So the influence that another entity may have on a user may change considerably with each new entity activating the user [19].

To address those difficulties, we consider an entire sequence of entities by which the user was activated, and argue it leads to the current-time-step product activation. We observe that there could be substitute goods and complementary goods propagated through the social network. Buying a substitute good could set back buying another, which is typically of the same kind, whereas buying a complementary good could boost buying another, which tends to be adopted together. Buying a product could affect buying the others to the degree, which can be characterized by a generation probability $p(a_u^i|a_u^j)$, where $p(a_u^i)$ denotes the probability that user u is activated by product i and $p(a_u^j)$ denotes the probability that user u is activated by product j. Moreover, the effect degree is user specific and cannot be estimated as uniform and static. Similarly, buying a sequence of multiple products across timesteps could affect buying the target product to some degree, which can be characterized by a generation probability $p(a_u^{i_r}|a_u^{i_1}, \ldots, a_u^{i_{l-1}}, a_u^{i_l})$, assuming that before activated by product i_r, user u has been activated by l products $i_1, i_2, \ldots, i_l \in \mathcal{I} \setminus \{i_r\}$. That is, a user's activation probability on target product is determined by the entire sequence of entity activation of this user.

Considering the dynamics and sequentiality of conditional cascades as well as the non-linear, subtle relationships in the data, we propose a Deep Recursive Hybrid model (DRH) to capture entity correlations, cascade correlations, causalities between ratings and next-period purchase as well as dynamics of these correlations and causalities. The objective of our model is to predict the generation probability of target-product activation by learning hidden knowledge in activation sequence and rating sequence of user. User's ratings play an important role in affecting user's next-period purchase, as ratings directly reflect other user's feedback to the purchased products [20]. We construct our deep recursive network by using a novel Long Short-Term Memory (LSTM) learning framework. With two hidden temporal sequences: purchase sequence and rating sequence, the proposed deep recursive network can efficiently capture the relations between products, such as independent, complementary and competitive relations, which are hidden in the activation sequences and rating history. The relations are not limited to pairwise nodes. Even more important, the relations of various products are dynamic, namely, a relation will change considerably when a new product activates the user.

Our contributions can be summarized as follows:

(1) We propose a Competitive and Complementary Independent Cascade (CCIC) diffusion model and propose the follower-based influence maximization problem under this diffusion model. We propose a greedy algorithm (DRHGA) efficiently computes the seed node set for the target product.

(2) We propose a deep recursive hybrid model, DRH, to capture the entity correlations, cascade correlations, causalities between ratings and next-period adoptions as well as dynamics of these correlations and causalities. It brings up new insights into the influence diffusion. To the best of our knowledge, this is the first work to integrate deep recursive network and the side information, users' ratings, to capture generation distribution of target product activation in the IM problem.

(3) We perform experiments in multi-entity networks with competitive and complementary relations, to evaluate the effectiveness and efficiency of the proposed DRHGA algorithm in comparison with state-of-the-art algorithms. The experimental results verify the superiority of our algorithm to state-of-the-art algorithms.

2. Background and previous work

Influence diffusion takes advantage of the interactions between nodes in the network to spread ideas, infections and natural hazards etc. Traditional influence maximization problem, i.e., single-entity influence maximization, has been studied extensively. Kempe et al. [21] proposed two propagation models, the Linear Threshold (LT) and Independent Cascade (IC) models, to address the influence maximization problem, and proved this problem is NP-hard under these two models. Specially for the IC model, given the social network $\mathcal{G}(\mathcal{U}, \mathcal{E}, p)$ and $K \in Z_+$, where $\mathcal{U}$ denotes the set of nodes, $\mathcal{E}$ denotes the set of directed edges and $p \in [0, 1]$ specifies influence strength of each pairwise nodes (product-independent), influence maximization problem is to select a subset $\mathcal{S} \in \mathcal{U}$ ($|\mathcal{S}| = K$) as seed, through activating which the expected number of active nodes of the whole network (denoted as $\sigma(\mathcal{S})$) will be maximized at the end of influence diffusion process. In [21], Kempe et al. also proved that the influence function $\sigma(\mathcal{S})$ under both LT model and IC model is submodular and monotone, which allows influence function $\sigma(\mathcal{S})$ can be approximated by a greedy algorithm with approximation ratio of $1 - 1/e$. Formally, a non-negative real valued function f on subset of set $\mathcal{U}$ is submodular if for all $x \in \mathcal{U} \setminus \mathcal{T}$ and all $\mathcal{S} \subseteq \mathcal{T} \subseteq \mathcal{U}, f(\mathcal{S} \cup x) - f(\mathcal{S}) \geq f(\mathcal{T} \cup x) - f(\mathcal{T})$, and f is monotone if for all $\mathcal{S} \subseteq \mathcal{T} \subseteq \mathcal{U}, f(\mathcal{S}) \leq f(\mathcal{T})$. Leskovec et al. [8] proposed the CELF algorithm, which employed lazy-forward optimization in the process of selecting new seed and speeded up 700 times as their experiments demonstrated. Chen et al. [7] proposed the NewGreedy and MixedGreedy algorithms using the idea of reducing the scale of the propagation map. Tang et al. proposed approximation algorithms TIM+ [5] and IMM [6], which sampled nodes from the network propagation graphs to establish a reverse reachable set and speeded up orders of magnitude compared with the origin greedy algorithm. Huang et al. proposed a community-based and topic-aware influence maximization method [22]. Qiu et al. took a user's local network as the input to a graph neural network for learning her latent social representation [23].

Recent studies on influence diffusion have extended single-entity propagation to multiple-entity propagation. Multiple-entity influence maximization algorithms have been proposed, but most of them focus on either pure competitive relationship or pure complementary relationship between products. Carnes et al. [24], from the follower's perspective, proposed two models for the propagation of two competing entities, given the fixed budget of the follower and the initial seed node set for each entity. From the perspective of social network host, Lu et al. [25] proposed an algorithm which takes account of both collective expected spread maximization and fairness of seed allocation of each entity by setting optimization objective as a MinMax function. [26,27] used the game theoretic strategy to solve the competitive influence maximization problem. [12,13], studied the problem of competitive profit maximization from the host perspective. Lu et al. [15] studied the relationship between two entities, proposed the complementary influence maximization problem, and designed an approximation algorithm, which assigns deterministic values to the entities such that the complex correlations among entities can be modeled efficiently.

For competitive and complementary influence maximization, Ou et al. [14] studied the problem of competitive and cooperative influence maximization with game theory. Litou et al. [18] proposed a greedy algorithm for competitive and complementary influence maximization, assigning correlations of entities randomly. [17] used a Hawkes process to model spread of cascades in competitive and complementary networks, and proved the objective function is convex. However this method has a very high computational complexity. [16] used the Bayesian method to model the process of competitive and complementary influence spread, and assumed the product activation sequences are conditionally independent which is infeasible in the real world.

3. Problem formulation

In this section, we first briefly review the traditional independent cascade diffusion model which is designed to solve the problem of single-entity influence maximization, then introduce our Competitive and Complementary Independent Cascade (CCIC) diffusion model, and then give the formulation of follower-based influence maximization, which aims to solve the problem of target-product influence maximization in the network where competitive and complementary products are already being propagated.

3.1. Reviewing traditional independent cascade model

In traditional independent cascade model, given the social network $\mathcal{G}(\mathcal{U}, \mathcal{E})$, where $\mathcal{U}$ denotes the set of nodes and $\mathcal{E}$ denotes the set of directed edges, each directed edge in the graph, $(v, u) \in \mathcal{E}$, has a corresponding probability $p(u, v) \in [0, 1]$. Intuitively, $p(u, v)$ represents the probability that node u independently activates node v through edge (u, v) when node u is activated. The dynamic propagation process in the independent cascade model unfolds in discrete timesteps as follows: At time $t = 0$, a preselected initial set S_0 is activated first, while other nodes are in an inactive state. This initial set of nodes is called a seed set. Let S_t represent the set of all active nodes up to timesteps $t \geq 1$. At any timesteps $t \geq 1$, for any node $u \in S_{t-1} \setminus S_{t-2}$ $(set\ S_{-1} = \phi)$ that was just activated at timesteps $t - 1$, node u attempts to activate the inactive out-neighbor $v \in N(u) \setminus S_{t-1}$ once, where $N(u)$ denotes the out-neighbors of node u, the probability of success of this attempt is $p(u, v)$, and this activation attempt is independent of all other activation attempts. If the attempt is successful, node v is activated at time t, that is, $v \in S_t \setminus S_{t-1}$; if the attempt is unsuccessful, and other out-neighbors of node v have not successfully activated node v at time t, then node v remains inactive in time t, that is $v \in \mathcal{U} \setminus S_t$. The propagation process runs until there are not any more nodes being activated.

3.2. Competitive and complementary independent cascade model

The traditional Independent Cascade Model assumes that there is only one product propagated in the social network and ignores the more complex social influence involving multiple products propagated in the network, where influence diffusion may occur for more than one product or with more than one mode of pure competition (including independence, pure competition or complementarity as well as competition and complementarity coexisting).

Definition 1 (***Independent, Competitive and Complementary Propagating Entities***). Let $p(a_u^i)$ denote the probability that user u is activated by product i and $p(a_u^i|a_u^j)$ denote the conditional probability that user u is activated by product i on the condition that user u has been already activated by product j. For product $i, j \in \mathcal{I}$, we say that: (1) product i is independent to product j iff $\forall u \in \mathcal{U}$, $p(a_u^i|a_u^j) = p(a_u^i)$, (2) product i competes with product j iff $\forall u \in \mathcal{U}$, $p(a_u^i|a_u^j) < p(a_u^i)$, (3) product i complements product j iff $\forall u \in \mathcal{U}$, $p(a_u^i|a_u^j) > p(a_u^i)$.

The diverse propogating eneties motivate us to propose a more expressive and powerful diffusion model, which is able to reasonably capture not only competition but also complementarity. To this end, we design the CCIC diffusion model, which follows the same process as traditional IC model, except two major differences. First, there are multiple entities propagated in the social network, and the nodes can sequentially be activated by multiple products, which may be independent, competitive or complementary. Second, influence probabilities between two nodes on different target products are different, depending on not only the influence strength between two nodes $p(u, v)$ but also the correlations of the various products [18,28]. Besides, this kind of correlation is dynamic, namely, the influence that an entity may have on a user may change considerably with each new entity activating the user.

The correlations of the various products can be characterized as the ratio of generation probability of conditional activation to probability of independent activation. E.g., for the case of only two products propagated in the network, for the user v, the correlation between product i and j can be represented as ratio $\frac{p(a_v^i|a_v^j)}{p(a_v^i)}$. For the case of more than two products propagated in the network, the dynamics of the correlations of various products need to be taken into account, since these correlations could change over time, namely, the influence that another product may have on a user may change considerably with each new product activating the user. Assuming product sequence $i_1, i_2, \ldots, i_l \in \mathcal{I} \setminus \{i_r\}$ has activated node v, the dynamics of the correlations between product i_r and others products can be characterized as $\frac{p(a_v^{i_r}|a_v^{i_1},\ldots,a_v^{i_{l-1}},a_v^{i_l})}{p(a_v^{i_r})}$.

Following [18], we consider the influence probability of pairwise nodes on the product i_r depends on not only the influence strength (product-independent and often given) between two nodes but also the correlations between product i_r and others products. Assuming product sequence $i_1, i_2, \ldots, i_l \in \mathcal{I} \setminus \{i_r\}$ has activated node v, the influence probability that node u activates node v on product i_r, $p_{u,v}^{i_r}$, is expressed as

$$p_{u,v}^{i_r} = p(u, v)\frac{p(a_v^{i_r}|a_v^{i_1}, \ldots, a_v^{i_{l-1}}, a_v^{i_l})}{p(a_v^{i_r})}, \tag{1}$$

where $p(u, v)$ denotes influence strength from node u to node v. Specially, for the case of $p(u, v)\frac{p(a_v^{i_r}|a_v^{i_1},\ldots,a_v^{i_{l-1}},a_v^{i_l})}{p(a_v^{i_r})} > 1$, we set $p_{u,v}^{i_r} = 1$.

3.3. Follower-based influence maximization

Based on the CCIC diffusion model, we propose the problem of follower-based influence maximization. The motivation of introducing the new problem is that we observe in reality, there is follower (company), who wants to use viral marketing to promote a new product in the market where competitive and complementary products are already being propagated. The problem of follower-based influence maximization is to find the top-K influential nodes as seed nodes for the target product to maximize the expected number of all active nodes on the target

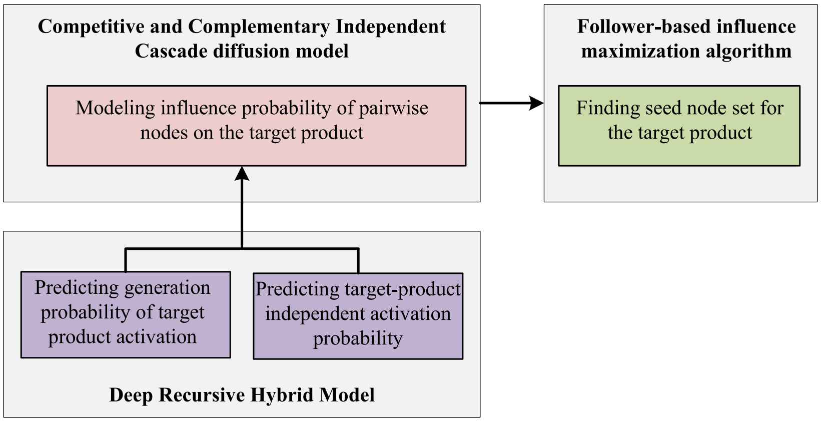

Competitive and Complementary Independent Cascade diffusion model
Modeling influence probability of pairwise nodes on the target product
Follower-based influence maximization algorithm
Finding seed node set for the target product
Predicting generation probability of target product activation
Predicting target-product independent activation probability
Deep Recursive Hybrid Model

Fig. 2. Overview of the system architecture.

product at the end of the diffusion process. The key to solving the problem is to compute the influence probability that node u activates node v on target product i^*, i.e., $p^{i^*}_{u,v}$ expressed as

$$p^{i^*}_{u,v} = p(u, v)\frac{p(a^{i^*}_v | a^{i_1}_v, \dots, a^{i_{l-1}}_v, a^{i_l}_v)}{p(a^{i^*}_v)}, \tag{2}$$

assuming product sequence $i_1, i_2, \dots, i_l \in \mathcal{I} \setminus \{i^*\}$ has activated node v. Specially, for the case of $p(u, v)\frac{p(a^{i^*}_v | a^{i_1}_v, \dots, a^{i_{l-1}}_v, a^{i_l}_v)}{p(a^{i^*}_v)} > 1$, we set $p^{i^*}_{u,v} = 1$.

Definition 2. (**Follower-based Influence Maximization**). Given a social network $\mathcal{G} = (\mathcal{V}, \mathcal{E}, p)$, a log of past purchasing $\mathbb{D}$ (activation sequence of each user can be obtained from $\mathbb{D}$), a record of rating $\mathbb{R}$ (rating sequence of each user can be obtained from $\mathbb{R}$) and the target product i^*, where p denotes influence strength of each pairwise nodes (product-independent), follower-based influence maximization under CCIC diffusion model is to find a subset of $\mathcal{V}$ for i^*, i.e., $\mathcal{S}eed_{i^*}$, such that

$$\mathcal{S}eed_{i^*} = \arg\max_{\mathcal{S}_{i^*}} E(\mathcal{S}_{i^*} | \mathcal{V}, \mathcal{E}, p^{i^*}_{u,v}) = \arg\max_{\mathcal{S}_{i^*}} E(\mathcal{S}_{i^*} | \mathcal{V}, \mathcal{E}, p, \mathbb{D}, \mathbb{R}), \tag{3}$$

where $p^{i^*}_{u,v}$ denotes influence probability of arbitrary pairwise nodes on target product i^* and $E(\mathcal{S}_{i^*} | \mathcal{V}, \mathcal{E}, p, \mathbb{D}, \mathbb{R})$ denotes the expected spread (the expected number of active nodes of the whole network at the end of influence diffusion process) under seed set $\mathcal{S}_{i^*}$, given the social network, the activation sequences and the rating sequences of each user.

4. The proposed method

According to our problem formulation, we can see that under CCIC diffusion model, to evaluate expected spread of one seed set, the key is to predict influence probabilities of pairwise nodes on target product i^*. To predict these influence probabilities, we propose a method that is composed of an approach to infer dynamic influence probabilities of pairwise nodes on target product, a deep recursive hybrid model to predict generation probability of target product activation and a greedy algorithm to compute seed nodes. Fig. 2 presents the overview of system architecture of our method. We will detail our method in the following subsections.

4.1. DRH : Deep recursive hybrid model to predict influence probability on target product

CCIC diffusion model involves multiple-entity propagating, competitive and complementary, as well as complex correlations among entities and cascades. Furthermore, this kind of correlation is dynamic because the interactions among products will change with the dynamics of cascades (entity activating the user through the influence between nodes). Thus, we need to build a model to capture entity correlations, cascade correlations and dynamics of these correlations.

The goal of DRH model is to predict influence probability of pairwise nodes on target product. Considering the dynamics and sequentiality of conditional cascades as well as the non-linear, subtle relationships in the data, we propose a deep recursive hybrid model to learn how different cascades interact with each other and then use these interactions combining rating sequences to more accurately predict generation probability of target product activation. Assume $u \in \mathcal{U} = \{1, \dots, m\}$ be a user and $i \in \mathcal{I} = \{1, \dots, n\}$ be a item by which the user can be activated. Let $\boldsymbol{X} \in \{0, 1\}^{n \times m}$ be activation matrix and x_u be the column (binary) with activation status of each item for u. Let $\boldsymbol{Y}^{n \times m}$ be rating matrix and y_u be the column (integer, e.g., $\in \{1, 2, 3, 4, 5\}$) with ratings on each item for user u. Let $x_{u,t}$ indicate the activation status of user u on tth item, where $x_{u,t} = 1$ if user u is activated by tth item; $x_{u,t} = 0$, otherwise. Let $y_{u,t}$ indicate the rating of user u on tth item, where $y_{u,t} \in \{1, 2, 3, 4, 5\}$ if user u rates on tth item; $y_{u,t} = 0$, otherwise.

The aim of our deep recursive hybrid model-DRH is to train a machine learning model $f(\cdot, \cdot)$, which can predict the generation probability of target-item activation for a user according to the user's entire activation sequence and rating sequence information. That is,

$$\mathbf{z}_t = f(\mathbf{x}_{(1:t-1)}, \mathbf{y}_{(1:t-1)}), \tag{4}$$

where $\mathbf{x}_{(1:t-1)}$ is an aggregation of historical activations from $\mathbf{x}_1$ to $\mathbf{x}_{t-1}$, and $\mathbf{y}_{(1:t-1)}$ is an aggregation of historical ratings from $\mathbf{y}_1$ to $\mathbf{y}_{t-1}$.

To realize the deep recursive hybrid model, we design a novel Long Short-Term Memory (LSTM) learning framework. LSTM [29], as a particular Recurrent Neural Network (RNN), has been proposed to alleviate the problem of the vanishing and exploding gradient of RNN [30]. In basic LSTM, a temporal sequence $x_1, x_2, \dots, x_t$ is taken as input, the hidden state of step t is updated by combining the current input x_t and the previous hidden state h_{t-1}, with three functions, forget gate, input gate and output gate, controlling the input and output of the network. While in our DRH model, we construct the proposed LSTM with two hidden temporal sequences, one hidden temporal sequence for modeling entity interactions and another hidden temporal sequence for modeling user evaluations (ratings). With two different types of hidden temporal sequences embedded in the model, the dynamics of each type can be captured well. Studies on psychology and marketing has revealed that user ratings reflects the user' feedback to the adopted items, which plays an important role in affecting user' next-period adoption [20]. Thus, we taken both item activation sequence and user rating sequence into account when constructing our deep recursive network.

Furthermore, considering the hidden temporal sequences of activation and rating do not contribute equally to the output, we design a deep recursive network structure with activation sequence as main sequence and rating sequence as additional sequence. Only the main sequence has the hidden state, and the additional sequence, having no hidden state, only affects cell state through a called decomposition gate. The structure of our deep recursive network architecture is presented in Fig. 3 and mathematical expressions are detailed as:

$$\begin{aligned}
\mathbf{f}_t &= \sigma(\mathbf{W}_f \cdot [\mathbf{x}_t, \mathbf{h}_{t-1}] + \mathbf{b}_f)\\
\mathbf{i}_t &= \sigma(\mathbf{W}_i \cdot [\mathbf{x}_t, \mathbf{h}_{t-1}] + \mathbf{b}_i)\\
\mathbf{o}_t &= \sigma(\mathbf{W}_o \cdot [\mathbf{x}_t, \mathbf{h}_{t-1}] + \mathbf{b}_o)\\
\mathbf{d}_t &= \sigma(\mathbf{W}_d \cdot \mathbf{C}_{t-1} + \mathbf{b}_d)\\
\widetilde{\mathbf{C}}_t^r &= \mathbf{d}_t * \sigma(\mathbf{W}_r \cdot \mathbf{y}_t + \mathbf{b}_r)\\
\widetilde{\mathbf{C}}_t &= \tanh(\mathbf{W}_c \cdot [\mathbf{x}_t, \mathbf{h}_{t-1}] + \mathbf{b}_c)\\
\mathbf{C}_t &= \mathbf{f}_t * (\mathbf{C}_{t-1} + \widetilde{\mathbf{C}}_t^r) + \mathbf{i}_t * \widetilde{\mathbf{C}}_t\\
\mathbf{h}_t &= \mathbf{o}_t * \tanh(\mathbf{C}_t),
\end{aligned} \tag{5}$$

where σ is the sigmoid function, $\cdot$ and $*$ respectively denote matrix product and element-wise product, $\mathbf{f}_t$, $\mathbf{i}_t$, $\mathbf{o}_t$ and $\mathbf{d}_t$ are forget gate, input gate, output gate and decomposition gate respectively, $\mathbf{C}_t$ is cell state, $\mathbf{h}_t$ is hidden state, $\mathbf{W}_*$ is the weight of different gates and $\mathbf{b}_*$ represents the bias. In our deep recursive network, only memories related with activation are preserved. The additional sequences (rating sequences) are controlled by the previous memory (as the fourth formula of Eq. (5) shows), and are imposed on the cell state (as the fifth and seventh formulas of Eq. (5) show).

Through the proposed deep recursive network, the probability that user is activated by target item under condition of entire activation sequence and rating sequence, i.e., $p(\mathbf{o}_t|\mathbf{x}_{(1:t-1)}, \mathbf{y}_{(1:t-1)})$ can be predicted well. To compute the target-item independent activation probability of arbitrary user v, i.e., $p(a_v^{i^*})$, we assume $p(a_v^{i^*})$ subjects to Gauss distribution $\mathcal{N}(0, 1)$. Then, corresponding to Eq. (2), the influence probabilities of pairwise nodes on target item can be obtained as:

$$p_{u,v}^{i^*} = p(u, v)\frac{p(\mathbf{o}_t|\mathbf{x}_{(1:t-1)}, \mathbf{y}_{(1:t-1)})}{p(a_v^{i^*})}. \tag{6}$$

4.2. DRHGA : Follower-based influence maximization algorithm

We prove monotonicity and submodularity of follower-based influence maximization problem (Please see the Appendix) and propose a greedy-based DRHGA Algorithm for to compute the top-K influential nodes for the target item. As outlined in Algorithm 1, DRHGA Algorithm has an approximation ratio of $(1 - \frac{1}{e})$ from the result of [21], where e is the base of the natural logarithm. With the influence probabilities of pairwise nodes inferred from DRH model, the algorithm selects K influential nodes greedily that maximize the marginal expected spread, and finally returns the seed node set $\mathcal{S}_{i^*}$.

5. Experimental setup

In this section, our research questions, datasets, baselines, evaluation metrics and experimental settings are presented.

Algorithm 1 DRHGA Algorithm

```
Input: social network $\mathcal{G} = (\mathcal{U}, \mathcal{E}, p)$, seed set size $K$, the target
    product $i^*$, purchasing log $\mathbb{D}$ and rating history $\mathbb{R}$
Output: Seed node set $\mathcal{S}_{i^*}$.
predicts influence probability of arbitrary pairwise nodes on
    target product, $p_{u,v}^{i^*}$ through DRH model;
for $j = 1 \dots K$ do
    $t^* \leftarrow \arg\max_{t \in \mathcal{U}} E(\mathcal{S}_{i^*} \cup t | \mathcal{V}, \mathcal{E}, p_{u,v}^{i^*})$
            $- E(\mathcal{S}_{i^*} | \mathcal{V}, \mathcal{E}, p_{u,v}^{i^*})$;
    $\mathcal{S}_{i^*} = \mathcal{S}_{i^*} \cup t^*$;
end for
return $\mathcal{S}_{i^*}$;
```

Table 1
Overview of the datasets used in this paper.

	Yelp 2013	Epinions	Flixster
Users	70,816	22,166	11,258
Items	15,584	296,277	14,296
Ratings	355,021	922,267	92,846
Links	622,873	355,813	84,606

5.1. Research questions

We organize the reminder of our paper by answering the research questions below:

(RQ1) Does our DRHGA algorithm has superiority in terms of prediction accuracy, compared with the state-of-art algorithms?
(RQ2) Can the proposed DRHGA algorithm outperform the state-of-art algorithms w.r.t. expected spread and running time?
(RQ3) How about the effect of hyper-parameters (neg-ratio and embedding size D) on the performance of the proposed DRHGA algorithm?
(RQ4) Is the proposed DRHGA algorithm scalable when the network size increases?
(RQ5) Does our DRH model contribute to improve the performance of the propose algorithm?

5.2. Datasets

We use three benchmark datasets: Yelp challenge 2013, Epinions and Flixster, which have been widely used for evaluating the performances of influence maximizing models [2,31,32] as well as the recommendation models [33]. In total, there are over 100,000 users, 300,000 items, 1,000,000 trusted social relations and 1,300,000 user reviews in our datasets. We show the statistics about our datasets in Table 1.

Yelp is a website about business reviewing, where people can comment on local restaurant services and add other persons as friends. Epinions is a social website about consumer opinion, where people share their reviews on products including food, books, and electronics etc. Flixster is a mobile and social rating website, specially rating on movies. Each of these three datasets contains a social graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a record of user rating $\mathbb{R}$ (each rating including user, item, timestamps etc.). Specifically, Yelp 2013 contains 70,816 users, 622,873 links, 15,584 items, 355,021 ratings [32]; Epinions contains 22,166 users, 355,813 links, 296,277 items, 922,267 ratings [2]; and Flixster contains 11,258 users, 84,606 links, 14,296 items, 92,846 ratings [31].

Kalish, the economist, argued that the purchase of a new product consists of two steps — awareness and purchase [31]. Product purchase is exactly the main goal in viral marketing. Thus, in influence diffusion, only when purchase happens, the

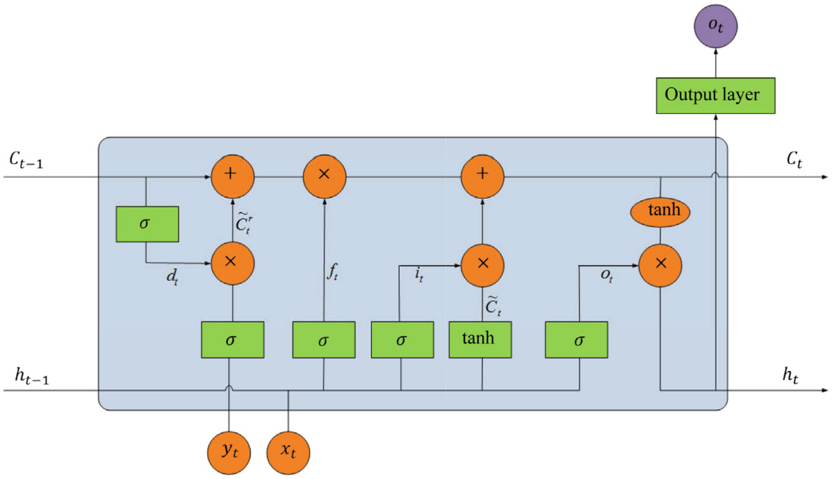


Fig. 3. The structure of our deep recursive network architecture.

node is considered to have been activated. In our experiments, we adopt the widely used idea that purchase means activation [31,34]. We need to use purchase (activation) log to predict generation probability of target-product activation through our DRH model. However, for the three datasets, we only have rating timestamp, thus we assume the purchase time is same as rating time. The same assumption is set in literature [2] to solve the problem of influence diffusion. This assumption will not affect the validity of our experiment, since we take the assumption all through the process of the experiments, including training, testing, validating and ground-truth obtaining (the detailed method to obtain ground truth will be presented in Section 5.4). With the assumption, we now have the purchase log, $\mathbb{D}$. Since these three datasets have no initial influence strength from node u to node v, $p(u, v)$, we generate $p(u, v)$ through the following way, which is widely used in literature [4,21]. we set the nonuniform influence strength $p(u, v)$ as $1/d(v)$, where $d(v)$ denote the in-degree of node v.

5.3. Baselines

We evaluate the performance of the proposed DRHGA algorithm comparing with the following state-of-the-art algorithms.

IMM: It is a single-entity influence maximization algorithm, with greedy strategy and Monte Carlo simulations. A classical statistical tool, martingale is used in IMM and a much higher efficiency can be obtained by IMM in practice.

Clash: It considers the competitive and complementary relation between products. It models the probability that a user is activated by a target product with the Bayes Formula, and make the assumption that each probability of the posteriori distribution of product activation is independent [16].

CorrelatedC: Competitive and complementary relation between products has been taken into account in this algorithm, and a Hawkes process is used to model the propagation of cascades [17].

CCDLT: It builds a spread model which considers different degrees of competition or complementarity among cascades, users' opinions to the products and the dynamics of the opinions, but randomly assigns the correlations of products. Then, it uses a greedy strategy to compute influence maximization [18].

5.4. Evaluation metrics and experimental settings

To evaluate prediction accuracy of the proposed DRHGA algorithm, we adopt three common evaluation metrics: **precision**, **recall** and **F1-score**. And to evaluate the overall performance of the proposed DRHGA algorithm, we adopt **expected spread** and **running time** as evaluation metrics, which are extensively used on influence maximization problem.

To evaluate the effectiveness of the proposed DRHGA algorithm, we obtain the ground truth with the purchase actions in the following cases: social influence - a user buys a product after at least one of his friends bought the product. Following [35], we set the threshold of time lag that the user buys a product after his friends.

We use 90% log of purchasing and rating as train set and 5% log of purchasing and rating as validation set, and set 5% log of purchasing and rating as well as all links as test set. For all the baseline algorithms and our proposed method, we repeat the experiments 10 times and report the average evaluation results.

Negative sampling is a technique used to train machine learning models that generally have several order of magnitudes more negative observations compared to positive ones. Negative sampling ratio is the ratio of negative observations to positive ones. For our DRH, we set negative sampling ratio, i.e., $Neg\text{-}ratio = 5$, and set the size of embeddings' dimension, i.e., $D = 50$. The minibatch is set to 128.

For the baselines, the optimal parameters are set according to the corresponding literatures.

6. Results and analysis

6.1. Evaluation of precision, recall and F1-Score

We first introduce how to calculate the evaluation metrics **precision**, **recall** and **F1-Score**.

Precision: it measures the ratio of the truly activated nodes in all the activated nodes that are searched by the algorithm, and is expressed as:

$$Precision = \frac{TP}{TP + FP}, \tag{7}$$

where TP denotes the number of the actually activated nodes in the activated nodes that are detected by the algorithm, and TP denotes the number of the nodes that are not actually activated in the activated nodes that are detected by the algorithm.

Recall: It indicates how many of the nodes that are actually activated are detected by the algorithm, and is expressed as:

$$Recall = \frac{TP}{TP + FN}, \tag{8}$$

Table 2
Comparison of precision, recall and F1-Score for all models on three datasets.

		Precision	Recall	F1-Score
Flixster	IMM	.2160	.4604	.2940
	Clash	.2255	.5185	.3138
	CCDLT	.2803	.5592	.3734
	CorrelatedC	.2902	.5761	.3860
	DRH	**.5725**	**.7761**	**.6589**
Epinions	IMM	.2467	.4925	.3289
	Clash	.2558	.5402	.3472
	CCDLT	.3126	.5799	.4062
	CorrelatedC	.3264	.5974	.4305
	DRH	**.6294**	**.7857**	**.6989**
Yelp 2013	IMM	.4205	.5015	.4574
	Clash	.4403	.5204	.4770
	CCDLT	.4621	.5316	.4944
	CorrelatedC	.4732	.5617	.5137
	DRH	**.8705**	**.6847**	**.7665**

where *FN* denotes the number of the actually activated nodes that are not detected by the algorithm.

F1-score: it is the harmonic mean of precision and recall, and is expressed as:

$$F1 - score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}. \tag{9}$$

RQ1: Table 2 presents the comparison of **precision**, **recall** and **F1-Score** of our DRHGA algorithm against the baselines for the three datasets.

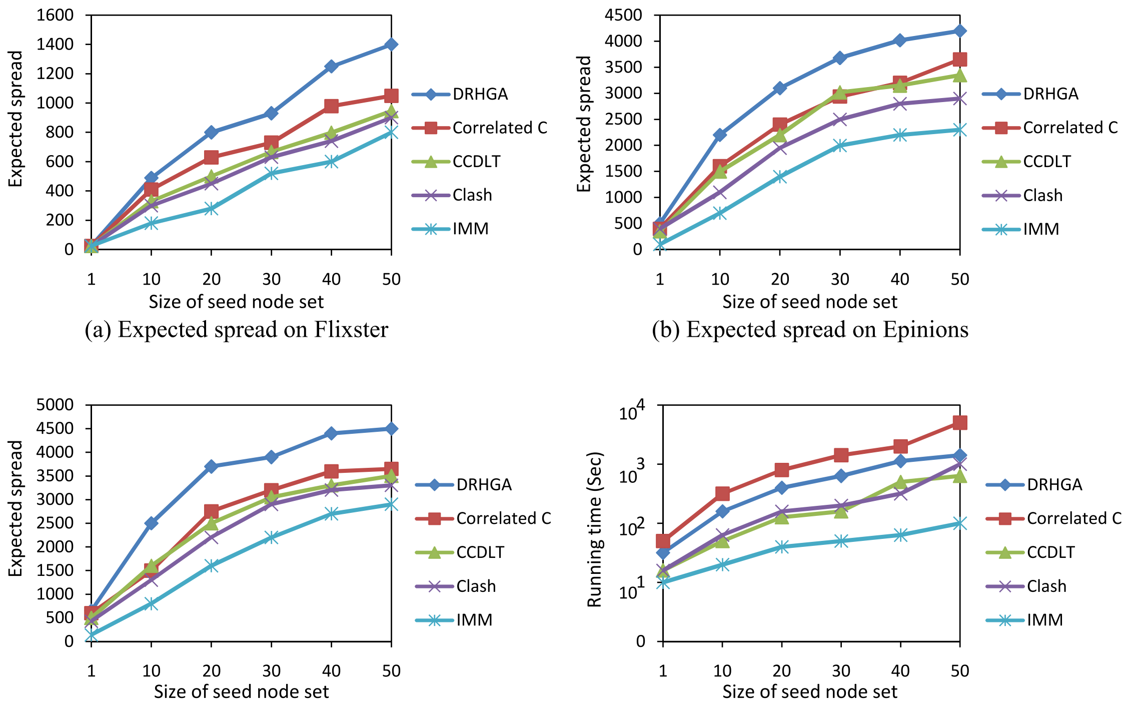
It can be seen from Table 2 that for the three datasets, all the algorithms have best performance of prediction accuracy on Yelp dataset because of the large size of this dataset. The most significant improvements of precision performance of our DRHGA respectively compared with IMM, Clash, CCDLT and CorrelatedC are 165%, 153%, 104% and 97%, which all happened on Dataset Flixster. The most significant improvements of Recall performance of our DRHGA respectively compared with IMM, Clash, CCDLT and CorrelatedC are 69%, 50%, 39% and 35%, which all happened on Dataset Flixster. The most significant improvements of F-Score performance of our DRHGA respectively compared with IMM, Clash, CCDLT and CorrelatedC are 121%, 110%, 76% and 71%, which all happened on Dataset Flixster. This illustrates that the proposed DRHGA algorithm presents significant superiority compared with all the other algorithms on all three datasets because it considers the dependency between the present-time-step activation and the activation history as well as the rating history. It also can be seen from Table 2 that IMM algorithm obtains lowest performance on all the three datasets, comparing with other algorithms. The reason is that IMM is an approximation algorithm with higher performance to solve the problem of single-entity influence maximization.

6.2. Algorithm overall performance

We evaluate the overall performance of all algorithms with two metrics, **expected spread** and **Running time**. **Expected spread** is the expected number of active nodes of the whole network at the end of influence diffusion process, given a set of seed nodes. In recent years, various algorithms to estimate **expected spread** have been proposed, which are more efficient than the general Monte Carlo simulation method. These algorithms fall into two categories — one category includes heuristic algorithms for Linear Threshold model or for Independent Cascade model, and another category includes improved greedy algorithms with Monte Carlo simulation. To estimate **expected spread** in our DRHGA Algorithm, we use the PMIA Algorithm [36], a typical heuristic algorithm for Independent Cascade model, since PMIA has better performance and scalability, especially for large-scale social networks.

RQ2: To evaluate the proposed DRHGA, we compare it with the baseline algorithms by setting the number of seeds K as 1, 20, 30, 40, 50. The expected spread of all algorithms are reported in Figs. 4(a)–(c). We can observe the following findings from Figs. 4(a)–(c): (1) When changing the seed size, the proposed DRHGA algorithm outperforms all the baseline algorithms with evaluation metric of expected spread, on the three datasets. E.g., on Yelp dataset challenge 2013, when setting the seed setsize K as 50, the expected spread of the proposed DRHGA increases by 1904, 1311, 852 and 549, respectively comparing with IMM, Clash, CCDLT and CorrelatedC. (2) On all the three datasets, all of the competitive and complementary influence maximization algorithms, i.e., Clash, CCDLT, CorrelatedC and DRHGA, outperform single-entity influence maximization algorithm, IMM, in terms of expected spread, which demonstrates that in the network with multiple propagation entities, it is helpful to capture a more complex form of correlation between different entities and discover how these correlations affect the propagation of a given entity, when considering the competitive and complementary relation of entities. (3) In general, Clash algorithm has the lowest performance among all competitive and complementary influence maximization algorithms, with evaluation metric of expected spread. E.g., on Yelp dataset challenge 2013, when setting the seed setsize K as 50, the expected spread of Clash, CCDLT, CorrelatedC and DRHGA are respectively 2892, 3339, 3645, 4203. This indicates that considering the dependencies of posteriori probabilities of activation does help to capture the correlations of different entities better. (4) The performances of DRHGA and CorrelatedC outperform that of CCDLT, which manifests that it is helpful to better capture the degrees of competitiveness or complementarity among different entities by modeling the relations of different generation probabilities of activation, without randomly assigning the correlations of entities. (5) The proposed DRHGA outperforms CorrelatedC in terms of expected spread, which indicates the effectiveness of the proposed DRH model that incorporates deep recursive network to catch the non-linear and subtle relations among complex data efficiently.

The running time of all algorithms are reported in Fig. 4(d), in which the experimental results upon Yelp Dataset are only presented due to the space constraints. Similar trends of performance of the algorithms are found for datasets Flixster and Epinions. From Fig. 4(d), it can be observed that upon Yelp Dataset, comparing with the baseline algorithms, the proposed DRHGA algorithm obtains a better performance of expected spread by spending more time. Specifically, when setting the seed setsize K as 50, the running time of our algorithm is about 0.808 h, and the running time of Clash, CCDLT and IMM is about 0.278 h, 0.175 h and 0.028 h, respectively. In fact, when the seed setsize K=1,10,20,30,40,50, the difference of running time between our algorithm and Clash/CCDLT/IMM is within one hour, and the efficiencies of these algorithms are similar. It also can be seen from Fig. 4(d) that Clash and CCDLT are the two algorithms having better performance of running time, within the four competitive and complementary influence maximization algorithms. The reason is that these two algorithms either assign the correlations of entities randomly or assume in the Bayes Formula, each probability of the posteriori distribution of product activation is independent, which simplifies the modeling but reduces the prediction accuracy.



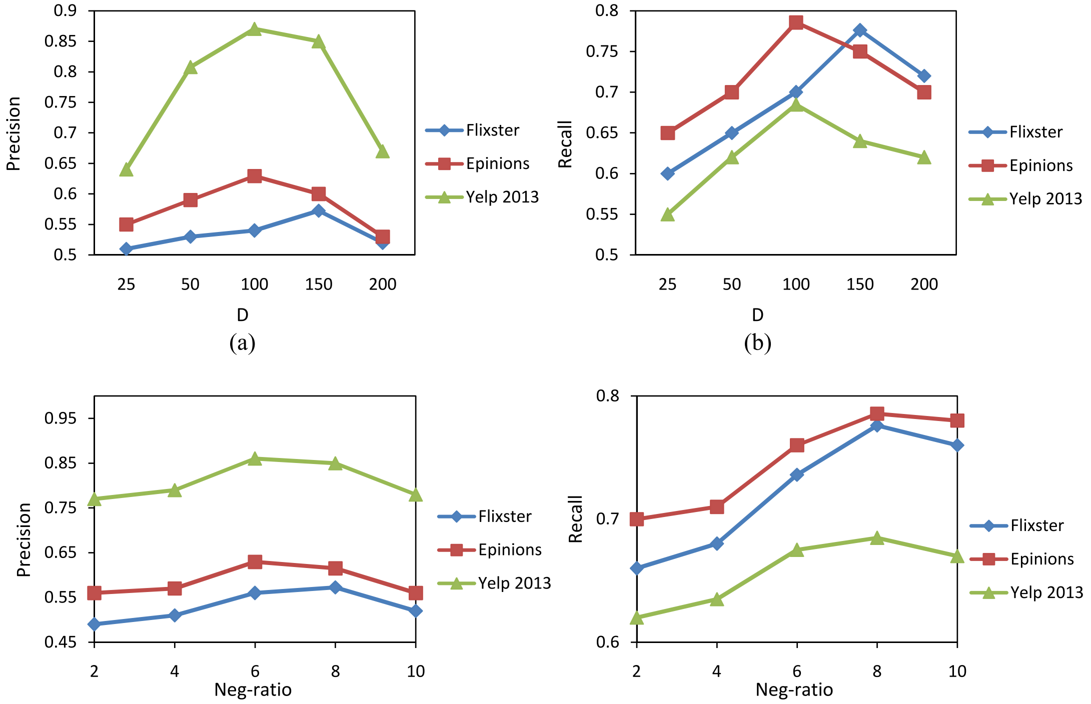
(a) Expected spread on Flixster

(b) Expected spread on Epinions

(c) Expected spread on Yelp 2013

(d) Running time on Yelp 2013

Fig. 4. Expected spread and running time with different size of seed node set K.



(a)

(b)

(c)

(d)

Fig. 5. Performance on Precision and Recall with different embedding sizes D (subfigures (a) and (b)) and negative sampling ratio Neg-ratio (subfigures (c) and (d)), respectively, on the three datasets, Fixster, Epinions and Yelp 2013.

6.3. Sensitivity analysis

RQ3: In this part, we analyze the sensitivity of the proposed DRHGA to hyper-parameters. To evaluate the impact of the dimension of latent space, we set the size of embedding dimension, D, as 25, 50, 100, 150 and 200 and compare the performance of Precision and Recall for different D upon the three datasets, with fixing Neg-ratio $= 5$. From Figs. 5(a)–(b), we can observe performance changes result from the different embedding sizes. It can be seen from Figs. 5(a)–(b) that in general, larger dimension leads to better performance. Specifically, the optimal embedding

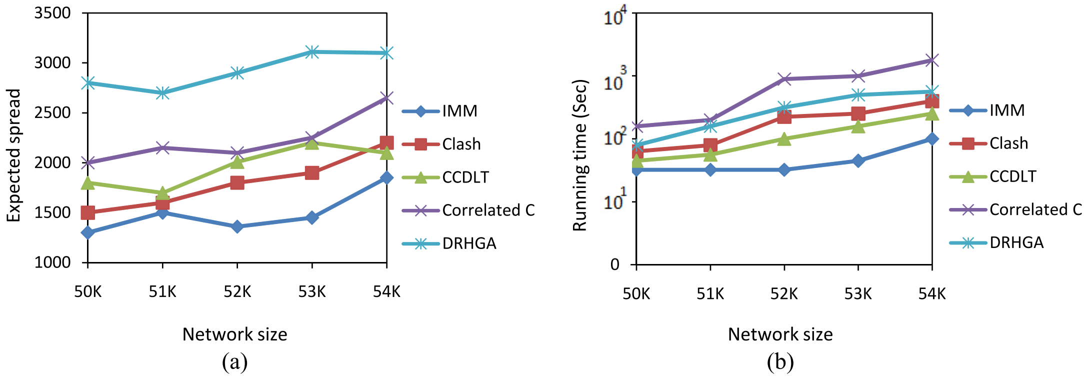


Fig. 6. Scalability on Yelp 2013.

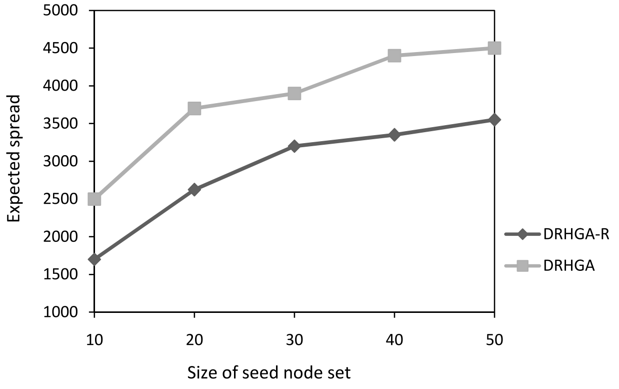


Fig. 7. Performance comparison of DRHGA and DRHGA-R on Yelp 2013.

size of the proposed DRHGA for dataset Flixster is 150, while it is 100 for datasets Epinions and Yelp 2013.

To understand the impact of the negative sampling ratio on the proposed DRHGA, we set Neg-ratio as 2, 4, 6, 8, 10 and compare the performance of Precision and Recall for different Neg-ratio upon the three datasets. Figs. 5(c)–(d) report the experimental results with different negative sampling ratios. It can be seen from Figs. 5(c)–(d) that: (1) In general, sampling more negative samples can lead to better prediction accuracy. (2) Upon the three datasets, the optimal negative sampling ratio for the proposed DRHGA is within [6,8], which means that best performance can be achieved through tuning the Neg-ratio.

6.4. Scalability evaluation

RQ4: To evaluate the scalability of DRHGA in comparison with the baseline algorithms, we need to generate subnetworks with different sizes. We use Yelp Dataset with a directed graph $\mathcal{G} = (\mathcal{U}, \mathcal{E})$ to generate subnetworks by the way of selecting a node randomly and executing the bread-first traversal. Fig. 6(a)–(b) demonstrate the performance of expected spread and running time of all algorithms, with the network size M varying from 50 K to 54 K and the seed set size K fixed as 50.

It can be seen from Fig. 6(a) that the expected spread of all algorithms appears to be relatively stable with the subnetwork size varying. The trends in Fig. 6(a) are consistent with the trends of experiments w.r.t. **RQ2**, i.e., the expected spread of the proposed DRHGA algorithm outperforms all baseline algorithms significantly. We can observe from Fig. 6(b) that the time cost of all algorithms tends to increase with the subnetwork size enlarging. It is notable in Fig. 6(b) that the proposed DRHGA has better scalability than CorrelatedC in terms of running time.

6.5. Contribution of the DRH model

RQ4: To evaluate the effect of the proposed DRH model, we compare the performance of expected spread of our DRHGA algorithm with the variant of DRHGA, DRHGA-R algorithm. The only difference between them is that DRHGA-R algorithm randomly assigns the generation probability of the target product activation and target-product independent activation probability. The strategy of assigning these probabilities randomly refers to the strategy used in literature [18]. From Fig. 7, it can be found that the expected spread of proposed DRHGA significantly outperforms that of DRHGA-R upon the dataset Yelp 2013, when the seed set size K is set as 10, 20, 30, 40 and 50. This manifests that through our DRH model to effectively modeling the latent relation between the target-product activation and activation history as well as rating history, the proposed DRHGA algorithm can achieve a better performance dramatically than DRHGA-R algorithm randomly assigning the correlations between entities.

7. Conclusions

We propose a novel optimization problem, follower-based influence maximization, which aims to find the top-K influential nodes for the target product in the network where other competitive and complementary products have already been propagating. To tackle the problem, we have proposed an approach to infer influence probabilities between nodes on target product, a Deep Recursive Hybrid model (DRH) and a greedy influence maximization algorithm (DRHGA). (1) The approach of inferring influence probability on target product can capture change of influence probability caused by the dynamics of cascades; (2) DRH dynamically can track entity correlations and cascade correlations, causalities between ratings and next-period purchase through a deep recursive network; (3) DRHGA is greedy-based and efficiently mines seed node set for the target product; (4) Experimental results validate the effectiveness of the proposed algorithm DRHGA, and the results show the superiority of our algorithm compared with the state-of-the-art algorithms.

As to future work, we intend to study online follower-based influence maximization under CCIC diffusion model, which is more challenging because of the dynamics and evolvement of the social network [37].

CRediT authorship contribution statement

Huimin Huang: Conceptualization, Methodology, Formal analysis, Software, Data curation, Writing - original draft. **Zaiqiao Meng:** Validation, Formal analysis, Writing - review & editing, Supervision. **Hong Shen:** Validation, Writing - review & editing.

Declaration of competing interest

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

Appendix

Theorem 1. *follower-based influence maximization problem is NP-hard.*

Proof. Suppose we have influence probabilities of pairwise nodes on target item in hand, follower-based influence maximization problem is simplified to the classical influence maximization under Independent Cascade model. Therefore, the theorem holds. □

Theorem 2. *For CCIC diffusion Model, the objective function* $E(\mathcal{S}_{i^*}|\mathcal{V}, \mathcal{E}, p, \mathbb{D}, \mathbb{R})$ *satisfies monotonicity and submodularity.*

Proof. (1) monotonicity. After the diffusion process of the existing items in $\mathcal{I} - i^*$, for item i^*, $\forall\ \mathcal{S}_{i^*} \subseteq \mathcal{U}$, adding a new seed node $u \in \mathcal{U} \setminus \mathcal{S}_{i^*}$ into $\mathcal{S}_{i^*}$, will not cause the influence spread decrease, i.e., $E(\mathcal{S}_{i^*} \cup \{u\}|\mathcal{V}, \mathcal{E}, p, \mathbb{D}, \mathbb{R}) \geq E(\mathcal{S}_{i^*}|\mathcal{V}, \mathcal{E}, p, \mathbb{D}, \mathbb{R})$.

(2) submodularity. After the diffusion process of the existing items in $\mathcal{I} - i^*$, influence probabilities of pairwise nodes on target item are inferred through DRH. For $\mathcal{R} \subseteq \mathcal{T} \subseteq \mathcal{U}$ and $v \in \mathcal{U} \setminus \mathcal{T}$, it is easy to verify $E(\mathcal{R} \cup \{v\}|\mathcal{V}, \mathcal{E}, p, \mathbb{D}, \mathbb{R}) - E(\mathcal{R}|\mathcal{V}, \mathcal{E}, p, \mathbb{D}, \mathbb{R}) \geq E(\mathcal{T} \cup \{v\}|\mathcal{V}, \mathcal{E}, p, \mathbb{D}, \mathbb{R}) - E(\mathcal{T}|\mathcal{V}, \mathcal{E}, p, \mathbb{D}, \mathbb{R})$ with live-edge path following [21]. Therefore, the theorem holds. □

References

[1] H. Huang, H. Shen, Z. Meng, Item diversified recommendation based on influence diffusion, Inf. Process. Manage. 56 (3) (2019) 939–954.
[2] H. Hung, H. Shuai, D. Yang, L. Huang, W. Lee, J. Pei, M. Chen, When social influence meets item inference, in: KDD, ACM, 2016, pp. 915–924.
[3] S. Galhotra, A. Arora, S. Roy, Holistic influence maximization: Combining scalability and efficiency with opinion-aware models, in: SIGMOD, ACM, 2016, pp. 743–758.
[4] W. Chen, C. Wang, Y. Wang, Scalable influence maximization for prevalent viral marketing in large-scale social networks, in: KDD, ACM, 2010, pp. 1029–1038.
[5] Y. Tang, X. Xiao, Y. Shi, Influence maximization: near-optimal time complexity meets practical efficiency, in: SIGMOD, ACM, 2014, pp. 75–86.
[6] Y. Tang, Y. Shi, X. Xiao, Influence maximization in near-linear time: A martingale approach, in: SIGMOD, ACM, 2015, pp. 1539–1554.
[7] W. Chen, Y. Wang, S. Yang, Efficient influence maximization in social networks, in: KDD, ACM, 2009, pp. 199–208.
[8] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J.M. Vanbriesen, N.S. Glance, Cost-effective outbreak detection in networks, in: KDD, ACM, 2007, pp. 420–429.
[9] S. Cheng, H. Shen, J. Huang, G. Zhang, X. Cheng, Staticgreedy: solving the scalability-accuracy dilemma in influence maximization, in: CIKM, ACM, 2013, pp. 509–518.
[10] X. He, G. Song, W. Chen, Q. Jiang, Influence blocking maximization in social networks under the competitive linear threshold model, in: Proceedings of the 2012 Siam International Conference on Data Mining, SIAM, 2012, pp. 463–474.
[11] A. Bozorgi, S. Samet, J. Kwisthout, T. Wareham, Community-based influence maximization in social networks under a competitive linear threshold model, Knowl.-Based Syst. 134 (2017) 149–158.
[12] Y. Zhu, D. Li, Host profit maximization for competitive viral marketing in billion-scale networks, in: INFOCOM, IEEE, 2018, pp. 1160–1168.
[13] A. Khan, B. Zehnder, D. Kossmann, Revenue maximization by viral marketing: A social network host's perspective, in: ICDE, IEEE, 2016, pp. 37–48.
[14] H.-C. Ou, C.-K. Chou, M.-S. Chen, Influence maximization for complementary goods: Why parties fail to cooperate? in: CIKM, ACM, 2016, pp. 1713–1722.
[15] W. Lu, W. Chen, L.V.S. Lakshmanan, From competition to complementarity: comparative influence diffusion and maximization, very large data bases 9 (2) (2015) 60–71.
[16] S.A. Myers, J. Leskovec, Clash of the contagions: Cooperation and competition in information diffusion, in: ICDM, IEEE, 2012, pp. 539–548.
[17] A. Zarezade, A. Khodadadi, M. Farajtabar, H.R. Rabiee, H. Zha, Correlated cascades: Compete or cooperate. in: AAAI, 2017, pp. 238–244.
[18] I. Litou, V. Kalogeraki, D. Gunopulos, Influence maximization in a many cascades world, in: ICDCS, IEEE, 2017, pp. 911–921.
[19] F. Wu, B.A. Huberman, Novelty and collective attention, Proc. Natl. Acad. Sci. USA 104 (45) (2007) 17599–17601.
[20] M. Mcpherson, L. Smithlovin, J.M. Cook, Birds of a feather: Homophily in social networks, Rev. Sociol. 27 (1) (2001) 415–444.
[21] D. Kempe, J.M. Kleinberg, E. Tardos, Maximizing the spread of influence through a social network, in: KDD, ACM, 2003, pp. 137–146.
[22] H. Huang, H. Shen, Z. Meng, H. Chang, H. He, Community-based influence maximization for viral marketing, Appl. Intell. (2019) 1–14.
[23] J. Qiu, J. Tang, H. Ma, Y. Dong, K. Wang, J. Tang, Deepinf: Social influence prediction with deep learning, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, ACM, 2018, pp. 2110–2119.
[24] T. Carnes, C. Nagarajan, S.M. Wild, A. Van Zuylen, Maximizing influence in a competitive social network: a follower's perspective, in: Proceedings of the Ninth International Conference on Electronic Commerce, ACM, 2007, pp. 351–360.
[25] W. Lu, F. Bonchi, A. Goyal, L.V. Lakshmanan, The bang for the buck: fair competitive viral marketing from the host perspective, in: SIGKDD, ACM, 2013, pp. 928–936.
[26] H. Li, S.S. Bhowmick, J. Cui, Y. Gao, J. Ma, Getreal: Towards realistic selection of influence maximization strategies in competitive networks, in: SIGMOD, ACM, 2015, pp. 1525–1537.
[27] A. Borodin, M. Braverman, B. Lucier, Strategyproof mechanisms for competitive influence in networks, in: WWW, ACM, 2013, pp. 141–151.
[28] H. Huang, Z. Meng, S. Liang, Recurrent neural variational model for follower-based influence maximization, Inform. Sci. (2020).
[29] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.
[30] J. Schmidhuber, Gradient flow in recurrent nets: the difficulty of learning long-term dependencies, in: A Field Guide To Dynamical Recurrent Neural Networks, IEEE Press, 2001.
[31] N. Barbieri, F. Bonchi, G. Manco, Topic-aware social influence propagation models, in: 2012 IEEE International Conference on Data Mining, 2012, pp. 81–90.
[32] Z. Ren, S. Liang, P. Li, S. Wang, M. de Rijke, Social collaborative viewpoint regression with explainable recommendations, in: Proceedings of the tenth ACM international conference on web search and data mining, 2017, pp. 485–494.
[33] Z. Meng, R. McCreadie, C. Macdonald, I. Ounis, S. Liu, Y. Wu, X. Wang, S. Liang, Y. Liang, G. Zeng, et al., Beta-rec: build, evaluate and tune automated recommender systems, in: Fourteenth ACM Conference on Recommender Systems, 2020, pp. 588–590.
[34] S. Bhagat, A. Goyal, L.V. Lakshmanan, Maximizing product adoption in social networks, in: Proceedings of the fifth ACM international conference on Web search and data mining, 2012, pp. 603–612.
[35] J. Yang, J. Leskovec, Modeling information diffusion in implicit networks, in: 2010 IEEE International Conference on Data Mining, IEEE, 2010, pp. 599–608.
[36] C. Wang, W. Chen, Y. Wang, Scalable influence maximization for independent cascade model in large-scale social networks, Data Min. Knowl. Discov. 25 (3) (2012) 545–576.
[37] Z. Meng, S. Liang, X. Zhang, R. McCreadie, I. Ounis, Jointly learning representations of nodes and attributes for attributed networks, ACM Trans. Inf. Syst. 38 (2) (2020) 1–32.