Competitive Opinion Maximization in Social Networks

Jianjun Luo
Worcester Polytechnic Institute
jluo@wpi.edu

Xinyue Liu
Worcester Polytechnic Institute
xliu4@wpi.edu

Xiangnan Kong
Worcester Polytechnic Institute
xkong@wpi.edu

***Abstract*—Influence maximization in social networks has been intensively studied in recent years, where the goal is to find a small set of seed nodes in a social network that maximizes the spread of influence according to a diffusion model. Recent research on influence maximization mainly focuses on incorporating either user opinions or competitive settings in the influence diffusion model. In many real-world applications, however, the influence diffusion process can often involve both real-valued opinions from users and multiple parties that are competing with each other. In this paper, we study the problem of competitive opinion maximization (COM), where the game of influence diffusion includes multiple competing products and the goal is to maximize the total opinions of activated users by each product. This problem is very challenging because it is #P-*hard* and no longer keeps the property of submodularity. We propose a novel model, called ICOM (Iterative Competitive Opinion Maximization), that can effectively and efficiently maximize the total opinions in competitive games by taking user opinions as well as the competitor's strategy into account. Different from existing influence maximization methods, we inhibit the spread of negative opinions and search for the optimal response to opponents' choices of seed nodes. We apply iterative inference based on a greedy algorithm to reduce the computational complexity. Empirical studies on real-world datasets demonstrate that comparing with several baseline methods, our ICOM approach can effectively and efficiently improve the total opinions achieved by the promoted product in the competitive network.**

***Index Terms*—Influence maximization, social networks, competitive opinion maximization**

I. Introduction

Motivated by many real-world applications (such as viral marketing), the problem of influence maximization has been studied intensively in recent years. The goal of influence maximization is to identify a small set of influential users in a social network, so that the coverage of an item (*e.g.*, a product or a political view) is maximized in the influence diffusion. More formally, given a social network and a diffusion model, the influence maximization problem aims at selecting a small set of seed nodes that maximizes the spread of influence in the social network [17]. This problem is especially important in viral marketing. For example, with a limited budget for product promotion, a company may want to selectively choose a small set of users to distribute free samples, hoping that they will recommend the product to their friends, consequently increase the product sales or brand awareness by word-of-mouth marketing.

Conventional methods on influence maximization mainly focus on maximizing the user coverage in a social network, i.e., assuming that all users have identical utility scores. The

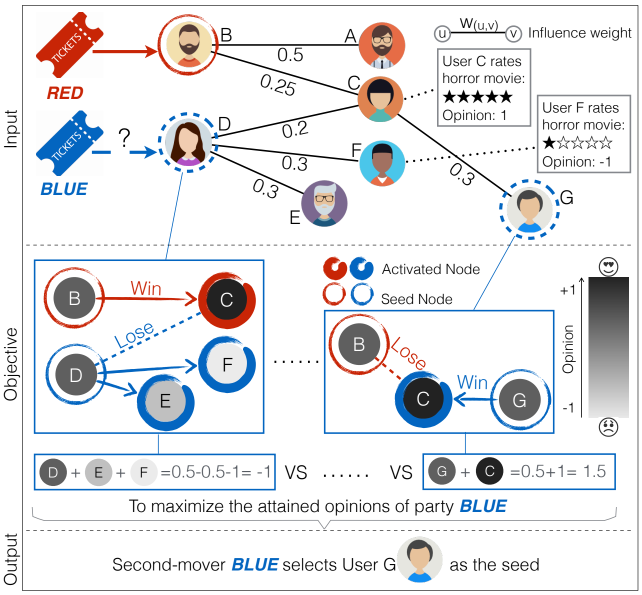
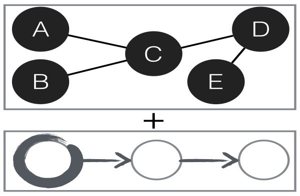
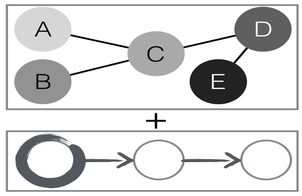


Fig. 1. An example of competitive opinion maximization. Given a social network and two parties (*blue* and *red*) promoting competing products (e.g., two horror movies), which users should the promoter (*blue*) pick as the seeds for viral marketing in two separate cases: (1) second-mover case: the seed users picked by the competitor (*red*) is *known*; (2) first-mover case: the promoter (*blue*) picks the seeds first, then the competitor (*red*) picks. Each user in the social network has a real-valued opinion (rating) towards the products. The two parties are competing to infect users in a competitive diffusion process. The goal of the promoter (*blue*) is to maximize the total ratings of its infected users. For simplicity, we ignore the first-mover case in this example.

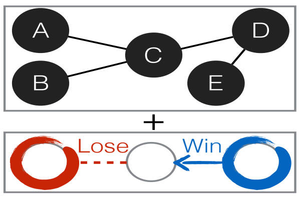
more users get infected in the diffusion, the better the performance (*e.g.*, higher sales or better reputation) is. Furthermore, most of existing works focus on non-competitive settings, i.e., assuming that there is only one party in the diffusion process. However, in many real-world applications, neither of the assumptions above holds. The diffusion process often involves multiple parties promoting competing items (e.g., products or political views) simultaneously and different users may have different opinions (or ratings) once infected in the diffusion. Each opinion (rating) is often a real-valued number, that can be negative. For example, in Figure 1, two movie studios (*blue* and *red*) released two horror films at the same weekend. The two studios may want to compete for high ratings in movie reviews (e.g., IMDB ratings) in order to attract more viewers in later weeks. Each rating may be positive or negative, depending on whether the viewer loves or hates horror movies. An ideal viral-marketing strategy should focus on finding horror movie fans and those who can



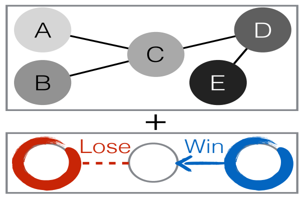
(a) Influence Maximization [17]: All users (nodes) have identical importance and the diffusion is non-competitive.



(b) Opinion Maximization [24]: Different users (nodes) can have different opinions and the diffusion is non-competitive.



(c) Competitive Influence Maximization [2]: All users (nodes) have identical importance, but the diffusion is competitive.



(d) Competitive Opinion Maximization (this paper): Consider both users' opinions and competitive diffusion.

Fig. 2. Comparison of four related problems.

influence the most fans. Therefore, we define a new problem to solve these real-world problems.

In this paper, we study the competitive opinion maximization (COM) problem with a competitive linear threshold (CLT) model. In CLT model, each party propagates the same way as it does in the linear threshold model [17], but an activated node cannot be influenced again by another party, since people seldom adopt another homogeneous product in a short time. An inactive node which receives influence from different parties at the same time will be activated by the one who sends the highest influence weight. Formally, the competitive opinion maximization problem corresponds to selecting a small set of seed users as the optimal response to the observed or assumed opponent's choices of seeds. The objective of selection is to maximize the total opinions gained after a competitive diffusion. Any two parties of the competition can be divided as a first mover and a second mover. A second mover can simply make its selection based on known opponent's selection, but the first mover needs to search for optimal choices to maximize the total opinions under the disadvantage of being observed by its opponent. Since the influence spread computation is #P-*hard* [10], the computation of opinions (spread with different utility scores) is at least #P-*hard*. The problem of opinion maximization in CLT diffusion model has not been studied in this context so far.

The major research challenges on competitive opinion maximization can be summarized as follows:

• **User Opinion:** One major challenge of the competitive opinion maximization problem is due to the real-valued opinions of the users. Conventional methods for influence maximization [3]–[5], [16] (as shown in Figures 2(a) and 2(c)) assume that all users in the social network have identical and positive utility scores. However, in competitive opinion maximization, user ratings (opinions) can be negative and different users can have different real-valued ratings. In such case, without considering opinions, a promotion strategy can lead to bad overall ratings. For example, horror films or cult films can be highly-rated by some viewers, but unpopular among mainstream audiences. It is important to target the appropriate customers because maximizing the spread of influence during the diffusion is no longer the optimal solution.

• **Rational Competitors:** Another challenge comes from the fact that multiple parties are competing in the market. If a user has already adopted an item, he or she will not accept another of the same type at the same time. We assume all the parties in the market are rational, which means they are likely to choose similar ideal customers if ignoring the competitors. The failure in a competition can block the diffusion of influence and significantly lower the total opinions. A wise decision should take opponents' choices into consideration and estimate the outcome based on the possible failure. As shown in Figures 2(c) and 2(d), a competitive diffusion model is necessary for the promoters to simulate the propagation of influence with competition.

• **Scalability:** Different from the problems that maximize the influence, our COM problem is neither submodular or monotonic, which means a standard greedy algorithm has no guarantee to the approximation ratio in [3]. Still, if the opponents' choices of seeds are observed, we can use a greedy algorithm as a heuristics to search through the network, since the standard greedy algorithm can achieve good empirical performance in non-submodular problems [11]. However, in some cases, the promoter can only make decisions based on the known opponent's strategy and will be observed by competitors as the first mover. Due to the passive position, a naive greedy algorithm for a first mover requires a lot of simulations, which is very slow for large networks. To address the problem, we utilize the inference from iterative simulation and design an efficient and effective heuristic algorithm.

In order to address the above challenges, we propose a novel solution, called ICOM (Iterative Competitive Opinion Maximization) method, to solve COM problem. By explicitly exploiting the users' opinions and opponents' information, our ICOM method can effectively find a set of seeds to compete with other parties for total opinions with an iterative inference procedure. Different from conventional influence maximization methods, the proposed ICOM model can exploit following information: (1) users' opinions; (2) observed competitor's choice of seeds; (3) competitor's strategy. Empirical studies on real-world tasks demonstrate that the proposed iterative opinion maximization approach can significantly boost the performances in terms of total opinions achieved by the promoter with competitors in real-world datasets.

II. Problem Formulation

A. Concept Definitions

The defined concepts will be used throughout this paper.

Definition 1(Social Network): An online social network can be represented as $G(\mathcal{V}, \mathcal{E}, \mathcal{W})$, where nodes $\mathcal{V} = \{u_1, \cdots, u_n\}$ is the set of users, $\mathcal{E}$ is the set of social links among users in $\mathcal{V}$ and $\mathcal{W} = \{w_{ij} | i = 1, \cdots, n; j = 1, \cdots, n\}$ is the set of

TABLE I
IMPORTANT NOTATIONS.

Symbol	Definition
$\mathcal{V} = \{u_1, \cdots, u_n\}$	the set of nodes
$\mathcal{E} = \{e_i \in \mathcal{V} \times \mathcal{V}\}$	the set of edges or links
$\mathcal{W} = \{w_{ij} \mid i, j = 1, \cdots, n\}$	the set of link weights
$\mathcal{P} = \{p_1, \cdots, p_c\}$	the set of parties competing in the network
$\mathcal{S} = (\mathcal{S}^1, \mathcal{S}^2, \cdots, \mathcal{S}^c)$	the set of seed user selections of c parties
u_i	user(node) i
w_{ij}	influence weight from u_j to u_i
i_t	type of competing item(product) t
p_v	party(promoter) v
$o_{it} \in [-1, 1]$	u_i's opinion towards i_t
$s_i^v \in [0, 1]$	the active status of u_i promoted by p_v
$\mathcal{S}^v$	the seed user selection of p_v
$\mathbf{o}_t = (o_{1t}, o_{2t}, \cdots, o_{nt})$	the vector of opinions on i_t assigned by all users in $\mathcal{V}$
$\mathbf{s}_v = (s_1^v, s_2^v, \cdots, s_n^v)$	the vector of status of users activated by p_v

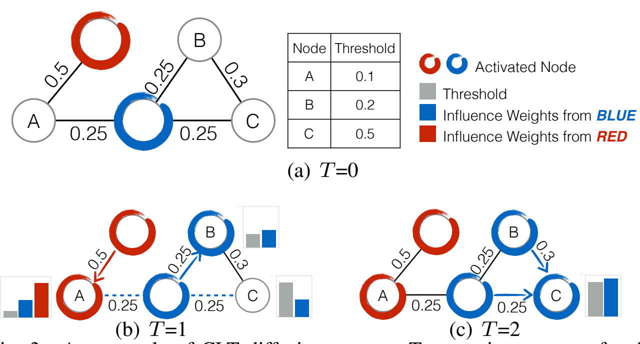


Fig. 3. An example of CLT diffusion process. Two parties compete for the node A, the one with higher influence weight wins. Propagation stops when no new nodes can be activated.

link weights. In the network, user(node) u_i can be influenced by its neighbor u_j according to the weight w_{ij}.

Definition 2 (Opinions): $\mathbf{o}_t = (o_{1t}, o_{2t}, \cdots, o_{nt})$ is the opinion vector of an item type i_t, where $o_{it} \in [-1, 1]$ represents the opinion of user u_i towards the type of items that compete with item t.

Definition 3 (Party): There are multiple parties $\mathcal{P} = \{p_1, \cdots, p_c\}$ promoting competing products in the network. These competing products are homogeneous and share common features, thus users will express the same opinions (preferences) on the competing items from different parties. In other words, these parties share the same opinion vector when they promote competing items.

Definition 4 (Active User Vector): Users who are influenced by Party v to adopt the promoted item are defined to be activated by p_v, while others are inactive. Active user vector $\mathbf{s}_v = (s_1^v, s_2^v, \cdots, s_n^v)$ represents the activated status of all users in the network by party p_v. Given the target item type i_t, when user u_i is activated by party p_v and adopt the product with opinion o_{it}, $s_i^v = 1$, otherwise $s_i^v = 0$ (inactivated).

Definition 5 (Seed Set): The decisions made by c competing parties can be represented as seed user set list $\mathcal{S} = (\mathcal{S}^1, \mathcal{S}^2, \cdots, \mathcal{S}^c)$, while $\mathcal{S}^{-v} = (\mathcal{S}^1, \cdots, \mathcal{S}^{v-1}, \mathcal{S}^{v+1}, \cdots, \mathcal{S}^c)$ defines the known or predicted seed sets selected by all parties except p_v.

B. Problem Definition

This paper focuses on the competitive LT model, but the proposed model shall be able to be adapted to competitive IC model.

Definition 6 (Competitive Linear Threshold): Competitive Linear Threshold (CLT) model is similar to the naive LT model, but with multiple participants. Each node u_i in the network has an activation threshold θ_i chosen uniformly at random in [0,1] and is possibly influenced by each neighbor j with weight w_{ij} such that $\sum_j w_{ij} \leq 1$. The influences of different parties start to propagate at the same time. Given a random choice of thresholds, at any timestamp T, if the total received weight of an inactive node i from party p_v's activated neighbors is higher than that from other parties and θ_i, then u_i is activated by p_v at timestamp $T+1$ (Figure 3(b)). Once a node is activated by one party, it cannot be activated by any other party. When more than one party selects the same node as seed node or is eligible to activate the same node at the same timestamp T, *e.g.*, p_1 and p_2, having $\sum_j w_{ij} s_j^1 = \sum_j w_{ij} s_j^2 > \theta_i$, node i will be equally activated by both parties p_1 and p_2 at $T+1$, denoted as $s_i^1 = s_i^2 = \frac{1}{2}$. The influence process repeats until no new node becomes active by any party (Figure 3(c)).

Definition 7 (Competitive Opinion Maximization): Given a network G, an item type i_t, corresponding opinion vector $\mathbf{o}_t$, the budget of seeds k and the CLT diffusion model, the goal of the party p_v in COM is to select a set of seed nodes $\mathcal{S}^v(|\mathcal{S}^v| = k)$ among G to propagate the influence to maximize the expected total opinions towards its item achieved at the end of the diffusion:

$$\mathcal{S}^v = \arg\max_{\mathcal{S}^v} \sigma(\mathcal{S}^v)$$

We define

$$\sigma(\mathcal{S}^v) = \sum_{\text{outcomes } X} \text{Prob}(X) \cdot \sigma_X(\mathcal{S}^v)$$
$$\sigma_X(\mathcal{S}^v) = s_v^X \cdot o_t$$

where $\mathbf{s}_v^X$ is the active user vector $\mathbf{s}_v$ subject to a choice X of node thresholds, denoting the set of all nodes reachable from any node in $\mathcal{S}^v$ for p_v, with the selections of other parties $\mathcal{S}^{-v}$. In other words, for p_v, the objective is to select seed set $\mathcal{S}^v$ to maximize the expected total opinions of eventually-influenced users $\sigma(\mathcal{S}^v)$.

For simplicity, we set $|\mathcal{P}| = 2$ in the following method description and experiment. We assume the opinions towards the items promoted by different parties are given and share the same opinion vector $\mathbf{o}_t$ for the specific item type t. This is consistent with the general case that users have relatively fixed preferences towards homogeneous products.

III. THE PROPOSED METHOD

We first propose an adapted greedy algorithm (GCOM) by incorporating the competitive setting and opinion objective into the greedy IM algorithm. Then we propose an improved method (ICOM) to efficiently and effectively maximize the total opinions under CLT model.

Algorithm 1 Greedy Alogorithm for COM (GCOM)

Input: social network G, opinion vector $\mathbf{o}_t$, seed user set size k, competitors' strategy (algorithm) M, competitors' choices of seed users $\mathcal{S}^{-v}$ (optional)

Output: the seed user selection $\mathcal{S}^v$

initialize $\mathcal{S}^v \leftarrow \emptyset$
if $\mathcal{S}^{-v}$ is known **then**
  while $\mathcal{V} \setminus \mathcal{S}^v \neq \emptyset \wedge |\mathcal{S}^v| \neq k$ **do**
    $u_{\text{best}} \leftarrow \arg\max_{u \in \mathcal{V} \setminus \mathcal{S}^v} \sigma(\mathcal{S}^v \cup \{u\}) - \sigma(\mathcal{S}^v)$
    $\mathcal{S}^v \leftarrow \mathcal{S}^v \cup \{u_{\text{best}}\}$
else
  if M ignores competition **then**
    $\mathcal{S}^{-v} \leftarrow M(G, \mathbf{o}_t, k)$
    go to 2
  else
    while $\mathcal{V} \setminus \mathcal{S}^v \neq \emptyset \wedge |\mathcal{S}^v| \neq k$ **do**
      $O_{\max} \leftarrow -\infty$
      for each u in $\mathcal{V} \setminus \mathcal{S}^v$ **do**
        $\mathcal{S}_{temp} \leftarrow \mathcal{S}^v \cup \{u\}$
        $\mathcal{S}^{-v} \leftarrow M(G, \mathbf{o}_t, k, \mathcal{S}_{temp})$
        if $\sigma(\mathcal{S}_{temp}) > O_{\max}$ **then**
          $O_{\max} \leftarrow \sigma(\mathcal{S}_{temp})$, $u_{\text{best}} \leftarrow u$
      $\mathcal{S}^v \leftarrow \mathcal{S}^v \cup \{u_{\text{best}}\}$
Return $\mathcal{S}^v$

Algorithm 2 Iterative Alogorithm for COM (ICOM)

Input: social network G, opinion vector $\mathbf{o}_t$, seed user set size k, competitors' strategy(algorithm) M, competitors' choices of seed users $\mathcal{S}^{-v}$ (optional), maximum round of iteration r

Output: the seed user selection $\mathcal{S}^v$

initialize $\mathcal{S}^v \leftarrow \emptyset$
if $\mathcal{S}^{-v}$ is known **then**
  $\mathcal{S}^v \leftarrow GCOM(G, \mathbf{o}_t, k, M, \mathcal{S}^{-v})$
else
  if M ignores competition **then**
    $\mathcal{S}^{-v} \leftarrow M(G, \mathbf{o}_t, k)$
    go to 2
  else
    Randomly generate an initial $\hat{\mathcal{S}}^v$
    $\mathcal{S}^{-v} \leftarrow M(G, \mathbf{o}_t, k, \hat{\mathcal{S}}^v)$
    $O_{\max} \leftarrow -\infty$
    for *round* in $1, \cdots, r$ **do**
      $\mathcal{S}^v \leftarrow GCOM(G, \mathbf{o}_t, k, M, \mathcal{S}^{-v})$
      $\mathcal{S}^{-v} \leftarrow M(G, \mathbf{o}_t, k, \mathcal{S}^v)$
      if $\sigma(\mathcal{S}^v) > O_{\max}$ **then**
        $O_{\max} \leftarrow \sigma(\mathcal{S}^v)$, $\mathcal{S}^v_{\text{best}} \leftarrow \mathcal{S}^v$
    $\mathcal{S}^v \leftarrow \mathcal{S}^v_{\text{best}}$
Return $\mathcal{S}^v$

It is worth noting that there are many improved algorithm to solve IM problem. Algorithms like CELF++ [13] and LDAG [8] are able to dramatically improve the efficiency under LT model. However, in this paper, we do not focus on improving the efficiency of seed selection algorithm in influence maximization as these studies. Instead, we propose a strategy that can solve COM problem effectively with opinions and opponent's information better than conventional IM strategies without additional consideration.

A. Greedy Algorithm on COM

Based on the accessible information, the parties can be classified into the first mover and the second mover. A second mover in the competition knows the competitor's exact choice of seed users $\mathcal{S}^{-v}$, while a first mover only knows the competitor's seed user selection strategy M.

We use an analogous greedy algorithm GCOM to search the optimal response to the competitors' observed or simulated choices directly (Algorithm 1). If p_v is a second mover who has observed the competitor's choices of seed users $\mathcal{S}^{-v}$, to form its own seed set, the promoter will search through all the nodes in $\mathcal{V}$ to greedily select the seeds. When searching for the qth ($q \leq k$) seed user, it tries every node that is not selected in the previous $q-1$ seeds to form different temporary set $\mathcal{S}^v \cup \{u\}$. Let the qth seed of p_v be the available node with the maximum simulated total opinions. The selecting process of a first mover who only knows the opponent's strategy M is similar, but it needs to simulate the $\mathcal{S}^{-v}$ based on the temporary $\mathcal{S}^v \cup \{u\}$ repeatedly. To simplify the calculation, a first mover whose competitor uses fix strategy (ignores competition) can be seen as a second mover after one simulation. However, it will be impractical and not scalable for the first mover due to the extra cost of estimating the second mover's choice when the opponent's model considers competition. So we adapt GCOM by using iterative inference and propose an efficient and effective method to address the COM problem.

B. The ICOM Method

Now that we propose the ICOM method (Algorithm 2) to improve the GCOM algorithm by simplifying the work flow for the first mover. Reducing the simulation times required by a first mover whose opponent seed set is not fixed is a key point to tackle the limitation of the simple greedy method. Inspired by the tit for tat strategy in game theory, ICOM is a heuristic method based on a small batch of multi-round simulation to approximate the strategic dominance.

Instead of considering every available node for the qth seed to form the possible temporary seed set and then simulating the possible responses for the second mover, ICOM randomly chooses k nodes as the very first seed set of the first mover p_v. Given the opponent's strategy M, we can easily infer the possible response $\mathcal{S}^{-v}$ from the competitor to the temporary $\mathcal{S}^v$. Then the iterative selection begins, and it includes the following steps: First, a new temporary $\mathcal{S}^v$ can be selected based on GCOM and the inferred $\mathcal{S}^{-v}$, *i.e.*, first mover p_v can search for an optimal response towards $\mathcal{S}^{-v}$ as a second mover. Then a new inferred $\mathcal{S}^{-v}$ can be predicted based on the new temporary $\mathcal{S}^v$ and the strategy M. The final active users

TABLE II

SUMMARY OF EXPERIMENTAL DATASETS.

	Data Sets		
Characteristics	CiaoDVD	Flixster	Filmtrust
# Nodes	2,740	5,320	1,642
# Links	20.8k	44.3k	44.6k
# Items	13.1k	3,470	2,071
# Ratings	34.3k	110k	35.5k
Link type	directed	undirected	directed
Rating scale	[1,5]	[0.5,5]	[0.5,4]

of p_v diffused by $\mathcal{S}^v$ after competing with $\mathcal{S}^{-v}$ will be used to estimate the total opinion $\sigma(\mathcal{S}^v)$ achieved in this round. Repeat such iterative inference until the maximum round r is reached. We choose the $\mathcal{S}^v$ with the highest total opinion $O_{\max}$ in all these rounds as the seed set of p_v.

To simplify the problem, we assume competitor's selection model is known. If the model is not given, we can assume the competitor will take optimal strategy. According to the empirical results (section IV), this assumption generally can be seen as the case that both parties use the strategy combining opinions and competition.

IV. EXPERIMENTS

A. Data Collection

We test our approach on three real-world networks with ratings (Summarized in Table II).

CiaoDVD: Ciao is a website for product reviews and price comparison. The dataset is crawled from the entire category of DVDs from the website [15]. The trust relationships between users are represented by the directed edges of the networks.

Flixster: Flixster is a website and a mobile app for movie information and ratings [1]. The users of Flixster link each other in the form of "friendship".

Filmtrust: FilmTrust is a small dataset crawled from the entire FilmTrust website in June, 2011. The website integrates "trust" social networks with movie ratings and reviews.

B. Experiment Setting

In CLT model, a user u_i can influence neighbors with certain weights. The weight of a directed link e_{ij} measures the influence from u_j to u_i. We calculate the weight of e_{ij} using Jaccard coefficient, which is widely used in social influence analysis. The strength of relationship, *i.e.*, the weight of link is defined as $\frac{\Gamma(u_i)\cap\Gamma(u_j)}{\Gamma(u_i)\cup\Gamma(u_j)}$. The threshold of users $[\theta_1,...,\theta_m]$, is randomly generated from a uniform distribution within $[0,1]$. We can get complete ratings of an item from all the users by the incomplete rating vector in datasets. To predict unknown ratings based on observed ratings, we use the matrix factorization method for collaborative filtering following [18], [21], [23]. The rating of user u_i towards item t can be approximated by the inner product of user profile $\mathbf{u}_i$ and item profile $\mathbf{v}_j$, which we can learn given the observed ratings. The approximated ratings which exceed the scope are replaced by the highest or lowest rating allowed in corresponding data set. We then convert the ratings to opinions using minmax normalization. The opinion $o_{it}=2\cdot\frac{r_{it}-r^t_{\min}}{r^t_{\max}-r^t_{\min}}-1$, where $r^t_{\max}$ and $r^t_{\min}$ are the minimum and maximum rating of product t. This operation maps the range of ratings to the range -1 to 1, which distinguishes the influence of positive or negative opinion towards the item, while original ratings in real-world dataset are always positive.

Given the network represented by the weighted adjacency matrix, and the opinions converted from ratings, we randomly choose three items from each dataset to perform the experiments. For each experiment, there are two parties in the market: a first mover and a second mover. There are five strategies for each party to take, *i.e.*, for an item type in a network, there are 25 kinds of setting to simulate the possible result in a competitive market.

C. Compared Methods

In order to demonstrate the effectiveness of our inferred COM approach, we test with following methods.

• **Random**: Method Random is a baseline method that selects inactive nodes as seeds randomly.

• **Influence Maximization (IM)** [17]: Method IM is a greedy method for influence maximization problem under linear threshold model.

• **Opinion Maximization (OM)**: Method OM is an adaptive version of IM, which also greedily select seeds. Instead of selecting the node with largest marginal spread, it selects the one that can gain highest opinion.

• **Competitive Influence Maximization (ICIM)**: ICIM is a degenerative version of ICOM, which neglects the preferences of users on the target item, aiming to get wider spread. It takes the number of infected nodes as optimal objective instead of total opinions.

• **Competitive Opinion Maximization (ICOM)**: ICOM is the proposed method based on CLT diffusion model.

All methods except Random include a greedy method to select the next seed with the largest marginal objective, so IM, OM, ICIM are sensible to be the baseline of ICOM. We can improve the efficiency of these algorithms by adapting that common step in future study. The comparison shows the advantage of combining opinions and competitors information into the strategy for COM problem. All of the experiments are evaluated under competitive environment, even though Method Random, IM or OM selects seeds as if it is the only promoter in the diffusion. In different settings, we use corresponding strategies to get the seed nodes of the first mover and the second mover. We then use CLT diffusion model to propagate the influence at the same time and output the total opinions of each party at the end as results. For models with randomness, their performances are measured by the 5-time average results.

D. Performances on ICOM

We first study the effectiveness of the proposed ICOM method on competitive opinion maximization. We report the total opinions achieved by the first mover in Table III. The performance is grouped by the opponent's strategy. Experiments are conducted over 9 randomly chosen items from 3 different networks. The budget of seed sets k is 10, while the inferring methods will only do 5 rounds inference (r=5). Performance ranks of each model within the group are also listed. We use

TABLE III
RESULTS OF DIFFERENT MODELS AS **FIRST MOVERS**. THE RESULTS ARE REPORTED AS "AVERAGE OPINIONS + (RANK)".

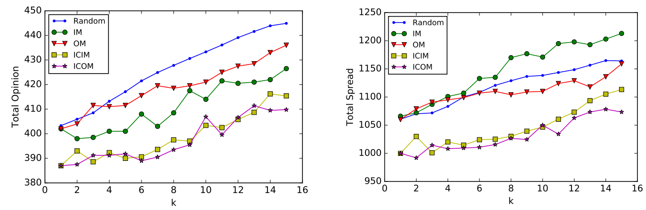
Competitor	Methods	Ciao			Flixster			Filmtrust			Avg.
		Item C_1	Item C_2	Item C_3	Item X_1	Item X_2	Item X_3	Item T_1	Item T_2	Item T_3	Rank
Random+	Random	310.1 (5)	210.8 (5)	197.6 (5)	67.3 (3)	49.7(4)	-90.0(4)	83.9 (5)	76.6 (5)	18.3 (5)	(4.6)
	IM	583.4 (3)	372.0 (3)	355.6 (2)	28.0 (5)	18.7(5)	-51.4(3)	130.5 (4)	120.7 (4)	28.3 (4)	(3.7)
	OM	595.6 (1)	389.4 (1)	378.5 (1)	66.3 (4)	56.9(3)	15.5(1)	137.7 (3)	126.5 (3)	36.9 (3)	(2.2)
	ICIM	564.0 (4)	353.7 (4)	341.0 (4)	100.1 (2)	91.2(2)	-200.7(5)	166.9 (2)	152.9 (2)	37.9 (2)	(3)
	ICOM	586.4 (2)	382.8 (2)	352.5 (3)	117.4 (1)	121.9(1)	-51.3(2)	173.6 (1)	222.1 (1)	39.4 (1)	(1.6)
IM+	Random	76.8 (5)	51.5 (5)	48.3 (5)	163.9 (4)	132.0 (4)	-276.9 (4)	261.0 (3)	238.5 (3)	58.7 (3)	(4)
	IM	255.5 (4)	170.2 (4)	164.5 (4)	5.9 (5)	4.3 (5)	-6.7 (3)	171.1 (5)	157.1 (5)	37.2 (5)	(4.4)
	OM	287.2 (3)	299.0 (3)	248.8 (3)	185.6 (3)	157.7 (3)	15.7 (2)	178.1 (4)	164.1 (4)	46.0 (4)	(3.2)
	ICIM	542.8 (2)	369.2 (2)	353.8 (2)	211.8 (2)	162.9 (2)	-348.8 (5)	398.1 (2)	364.4 (2)	89.6 (2)	(2.3)
	ICOM	553.2 (1)	378.5 (1)	364.2 (1)	226.6 (1)	179.7 (1)	20.4 (1)	402.4 (1)	368.9 (1)	96.3 (1)	(1)
OM+	Random	75.1 (5)	48.0 (5)	46.0 (5)	140.1 (3)	112.9 (4)	-343.5 (5)	261.0 (3)	238.3 (3)	58.7 (3)	(4)
	IM	347.2 (3)	125.5 (4)	154.8 (4)	20.9 (4)	10.0 (5)	-354.4 (3)	171.0 (5)	158.1 (5)	37.4 (5)	(4.2)
	OM	254.8 (4)	174.8 (3)	169.0 (3)	3.1 (5)	50.6 (3)	3.3 (2)	174.6 (4)	160.1 (4)	41.5 (4)	(3.6)
	ICIM	565.2 (2)	370.8 (2)	353.2 (2)	207.5 (2)	160.7 (2)	-354.4 (3)	395.0 (2)	362.7 (2)	88.4 (2)	(2.1)
	ICOM	558.5 (1)	380.2 (1)	358.2 (1)	226.1 (1)	184.9 (1)	12.9 (1)	402.4 (1)	368.9 (1)	96.3 (1)	(1)
ICIM+	Random	11.6 (5)	6.2 (5)	6.4 (5)	0.8 (4)	1.3 (4)	-1.4 (4)	75.3 (3)	68.8 (3)	15.8 (3)	(4)
	IM	92.2 (3)	37.2 (4)	38.2 (4)	-1.6 (5)	-1.6 (5)	-0.8 (3)	1.4 (5)	2.3 (5)	-0.3 (5)	(4.3)
	OM	78.2 (4)	55.8 (3)	60.2 (3)	5.6 (3)	10.1 (3)	15.7 (1)	8.6 (4)	8.3 (4)	8.3 (4)	(3.2)
	ICIM	306.0 (1)	210.8 (2)	196.0 (1)	80.4 (1)	59.1 (2)	-132.0 (5)	369.4 (1)	338.6 (2)	78.6 (2)	(1.9)
	ICOM	300.8 (2)	222.0 (1)	191.4 (2)	73.7 (2)	70.4 (1)	14.2 (2)	231.3 (2)	338.9 (1)	84.5 (1)	(1.6)
ICOM+	Random	9.8 (5)	3.6 (5)	4.7 (5)	-2.9 (4)	-0.3 (4)	-339.6 (3)	73.7 (3)	4.2 (4)	0.0 (4)	(4.1)
	IM	83.2 (4)	35.5 (4)	26.8 (4)	-4.6 (5)	-5.1 (5)	-356.7 (4)	1.4 (5)	2.1 (5)	-0.7 (5)	(4.6)
	OM	94.5 (3)	40.8 (3)	53.2 (3)	4.2 (3)	3.9 (3)	14.7 (1)	8.6 (4)	8.1 (3)	7.9 (2)	(2.8)
	ICIM	291.6 (1)	184.3 (2)	173.1 (2)	118.6 (1)	89.1 (1)	-358.1 (5)	362.8 (1)	74.2 (2)	7.6 (3)	(2)
	ICOM	287.6 (2)	195.5 (1)	184.4 (1)	105.2 (2)	81.4 (2)	13.8 (2)	294.3 (2)	334.5 (1)	67.5 (1)	(1.6)

TABLE IV
RESULTS OF DIFFERENT MODELS AS **SECOND MOVERS**.THE RESULTS ARE REPORTED AS "AVERAGE OPINIONS + (RANK)".

Competitor	Methods	Ciao			Flixster			Filmtrust			Avg.
		Item C_1	Item C_2	Item C_3	Item X_1	Item X_2	Item X_3	Item T_1	Item T_2	Item T_3	Rank
Random+	Random	273.9 (5)	185.5 (5)	181.6 (5)	121.4 (4)	98.9 (4)	-230.1 (4)	315.3 (3)	287.1 (3)	69.0 (3)	(4)
	IM	583.4 (4)	372.0 (4)	355.6 (4)	28.0 (5)	18.7 (5)	-51.4 (3)	130.5 (5)	120.7 (5)	28.3 (5)	(4.4)
	OM	595.6 (3)	389.4 (3)	378.4 (3)	66.2 (3)	56.9 (3)	15.5 (2)	137.7 (4)	126.5 (4)	36.9 (4)	(3.2)
	ICIM	643.2 (2)	412.6 (2)	394.4 (2)	210.0 (2)	162.3 (2)	-345.1 (5)	321.9 (2)	294.7 (2)	72.3 (2)	(2.3)
	ICOM	655.8 (1)	429.4 (1)	416.3 (1)	229.2 (1)	182.5 (1)	20.6 (1)	328.8 (1)	364.2 (1)	93.0 (1)	(1)
IM+	Random	76.8 (5)	51.5 (5)	48.3 (5)	163.9 (4)	132.0 (4)	-276.9 (4)	261.0 (3)	238.5 (3)	58.7 (3)	(4)
	IM	255.5 (4)	170.2 (4)	164.5 (4)	5.9 (5)	4.3 (5)	-6.7 (3)	171.1 (5)	157.1 (5)	37.2 (5)	(4.4)
	OM	287.2 (3)	299.0 (3)	248.8 (3)	185.6 (3)	157.7 (3)	15.7 (2)	178.1 (4)	164.1 (4)	46.0 (4)	(3.2)
	ICIM	542.8 (2)	369.2 (2)	353.8 (2)	211.8 (2)	162.9 (2)	-348.8 (5)	398.1 (2)	364.4 (2)	89.6 (2)	(2.3)
	ICOM	553.2 (1)	378.5 (1)	364.2 (1)	226.6 (1)	179.7 (1)	20.4 (1)	402.4 (1)	368.9 (1)	96.3 (1)	(1)
OM+	Random	75.1 (5)	48.0 (5)	46.0 (5)	140.1 (3)	112.9 (4)	-343.5 (5)	261.0 (3)	238.3 (3)	58.7 (3)	(4)
	IM	347.2 (3)	125.5 (4)	154.8 (4)	20.9 (4)	10.0 (5)	-354.4 (3)	171.0 (5)	158.1 (5)	37.4 (5)	(4.2)
	OM	254.8 (4)	174.8 (3)	169.0 (3)	3.1 (5)	50.6 (3)	3.3 (2)	174.6 (4)	160.1 (4)	41.5 (4)	(3.6)
	ICIM	565.2 (2)	370.8 (2)	353.2 (2)	207.5 (2)	160.7 (2)	-354.4 (3)	395.0 (2)	362.7 (2)	88.4 (2)	(2.1)
	ICOM	558.5 (1)	380.2 (1)	358.2 (1)	226.1 (1)	184.9 (1)	12.9 (1)	402.4 (1)	368.9 (1)	96.3 (1)	(1)
ICIM+	Random	87.5 (4)	64.1 (3)	58.2 (4)	84.8 (2)	55.7 (3)	-124.7 (4)	230.1 (1)	210.0 (2)	49.7 (2)	(2.8)
	IM	92.2 (3)	37.2 (5)	38.2 (5)	-1.6 (5)	-1.6 (5)	-0.8 (3)	1.4 (5)	2.3 (5)	-0.3 (5)	(4.6)
	OM	78.2 (5)	55.8 (4)	60.2 (3)	5.6 (4)	10.1 (4)	15.7 (2)	8.6 (4)	8.3 (4)	8.3 (4)	(3.8)
	ICIM	299.1 (2)	193.4 (2)	188.0 (2)	101.0 (1)	83.8 (1)	-172.8 (5)	24.6 (3)	21.4 (3)	8.5 (3)	(2.4)
	ICOM	314.8 (1)	215.6 (1)	210.2 (1)	76.5 (3)	63.2 (2)	20.6 (1)	31.8 (2)	286.7 (1)	81.9 (1)	(1.4)
ICOM+	Random	73.6 (5)	49.2 (3)	62.4 (3)	90.7 (2)	42.3 (3)	-264.7 (3)	228.5 (1)	145.4 (1)	53.2 (1)	(2.4)
	IM	83.2 (4)	35.5 (5)	26.8 (5)	-4.6 (5)	-5.1 (5)	-356.7 (5)	1.4 (5)	2.1 (5)	-0.7 (5)	(4.9)
	OM	94.5 (3)	40.8 (4)	53.2 (4)	4.2 (4)	3.9 (4)	14.7 (1)	8.6 (4)	8.1 (4)	7.9 (3)	(3.4)
	ICIM	302.4 (2)	178.4 (2)	192.6 (2)	105.0 (1)	75.4 (1)	-353.6 (4)	162.3 (2)	22.7 (3)	4.4 (4)	(2.3)
	ICOM	310.3 (1)	208.4 (1)	198.3 (1)	83.6 (3)	70.9 (2)	14.3 (2)	100.7 (3)	26.7 (2)	19.4 (2)	(1.9)

the average rank to compare the general performance of the model on item types with different popularity and in various networks. The performance of second mover are shown in Table IV, grouped by the first mover's strategy.

The first observation we have in Table III and Table IV is as follows: almost all the methods that explicitly exploit the opinions of the item can achieve higher total opinions than the baseline Random, and the corresponding degenerate method IM or CIM which only chases for the spread. These results can support the importance of considering users' opinions and the opponent's seed selection jointly. Maximizing the number of activated users does not guarantee the total opinions gained for the items. In some cases, *e.g.*, when diffusing a type of items with distinct feature that makes opinions from different people in stark contrast, or a kind of unpopular items that many people host negative opinions, the attempt to maximize the spread can even lead to a result worse than to select seed nodes randomly. We can also observe that in many situations, the performance of models considering competition have a significant improvement compared with the ones that ignore. These results support our claim that in a multi-player propagation, taking the opponent's strategy into consideration usually enhance competitiveness. Making decision based on the known or simulated seeds of opponents can avoid the possible failure in the competition or weaken performances brought by sharing the same customers with opponents. Thus, a market with multiple players should take competition into account and apply a more flexible and effective strategy to choose its seed users. These two main observations indicate the necessity and superiority of ICOM to achieve maximum



(a) Total opinions in the market (b) Total spread in the market

Fig. 4. Performances of two parties in Ciao when the second mover uses ICOM and the first mover chooses different strategies ($r = 5$).

opinions in a competitive environment. The overall results showed that no matter as a first mover or a second mover and no matter what strategy the opponent takes, ICOM strategy, which exploits both opinion and competition information, is better than other baseline methods.

E. Discussion of Parameters

To have a better understanding of how the competitive strategy works in the market, we fix the second mover strategy and change the first mover strategy, to see how the overall market grows as the budget k grows. Figure 4 presents the performances from both participants in terms of opinions and influence spread. One can see that as the budget k grows, the total gain over opinions or spread got by the two parties increases, which supports the intuition that increasing the budget can improve the final performances. However, the choice of first mover model makes the outcome different, while the second mover's fixed strategy is ICOM. As discussed in the previous section, the first mover will achieve more total opinions if it takes a competitive strategy. But from the figure, a competitive strategy will significantly decreases the second mover's gain and the overall gain of both parties. Therefore when both parties try to take a competitive strategy, maximizing its own gain, the overall market in terms of opinions or spread would decline.

Then we present the influence of different maximum number of round r on the performance of ICOM model. Figure 5 shows the convergence of the proposed ICOM by testing the performance after each iteration step. We randomly select an item from each of the datasets to run an experiment that both players are using ICOM, and each setting repeats 5 times to measure the performance. The budget k is 10 and the maximum tested round r is 15. Figure 5(a) demonstrates the relationship between r and the cost of time in the experiment of Ciao dataset, while other two figures are similar and omitted. It indicates that the time cost of ICOM linearly grows along with the parameter r. Figure 5(b) to 5(d) shows the total opinions achieved along with different r. The performances of the ICOM converge very fast after a few iterations. From these results, we can see that the performance of ICOM is not sensitive to the maximum number of rounds as long as r is assigned with a modest number. Although a global optimal result is not guaranteed, it has a relatively good performance and is suitable for balancing the performance and the cost. Thus in our previous experiment, we use 5 as the default maximum number of rounds. This also supports our intuition that using inference makes the cost of time controllable and exploiting competitor's selection is important and effective for competitive opinion maximization.

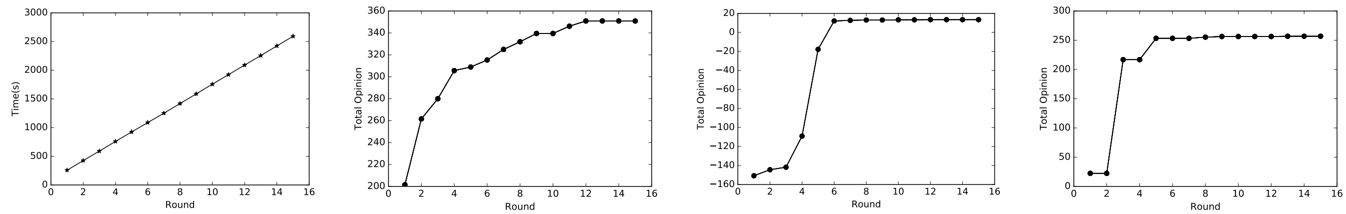
V. Related Work

To the best of our knowledge, this paper is the first work on competitive opinion maximization. Our work is related to influence maximization, opinion maximization and competitive influence maximization.

Influence Maximization: Influence maximization (IM) is to study how to choose a small set of seed nodes in a network, which has the best opportunity to influence the most number of nodes through a given diffusion model. In [17] Kempe *et al*. obtained the first provable approximation guarantee for the two basic stochastic influence cascade models they proposed, the independent cascade (IC) model and the linear threshold (LT) model. Chen *et al*. [8]–[10] designed a scalable algorithm for the IC model that can handle large-scale social networks and proposed the first scalable IM algorithm tailored for the LT model. Instead of using heuristics to estimate the spread like [14], an efficient algorithm called CELF proposed by Leskovec *et al*. [19] exploits submodularity and dramatically improves the efficiency of the greedy algorithm for IM problem.

Opinion Maximization: Instead of selecting the most influential nodes, the aim of opinion maximization is to make the item favorable and get more positive opinions. Chen *et al*. [7] proposed a model that incorporates the emergence and propagation of negative opinions into the IC model to maximize the expected number of positive activated nodes. Zhang *et al*. [24] considered the negative and neutral opinions, proposed an adapted IC model to maximize the total opinions of activated users. Gionis *et al*. [12] assumed the opinions of individuals get formed dynamically by the mutual influence of internal opinions and the neighbors' opinions. Liu *et al*. [22] studied a multi-round single-party opinion maximization problem, where they attempts to find the optimal set of seeds in each round of promotion to maximize the total opinion spread in the network based on the opinion observed.

Competitive Influence Maximization: The problem is to study the simultaneous propagation of multiple items in a social network. The solutions for competitive influence maximization can basically be classified into two types: opponent strategy known or opponent strategy unknown. For the first type, a popular solution is to minimize the opponent's influence, *i.e*., Influence Blocking Maximization [3], [4], [16]. Carnes *et al*. [5] studied the competitive influence maximization problem from a follower's perspective, *i.e*., finding a best response to the first mover's selection. For the second type, Bharathi *et al*. [2] proposed a natural generalization of the IC model and used game theory to study the diffusion with multiple competing items. Chen *et al*. [6] proposed a data-driven approach to study the multi-player influence maximization and proposed a game to collect the picking strategies from human or AI to analysis. Lin *et al*. [20] proposed a learning-based framework using reinforcement learning and game theory to address the multi-round competitive influence maximization



(a) Time (b) CiaoDVD (c) Flixster (d) Filmtrust

Fig. 5. Influence of parameter round r

problem. Zhang *et al*. studied the maximization problem of multiple competing or complementary products in a social network at the same time in [25].

VI. CONCLUSION

In this paper, we study the competitive opinion maximization (COM) problem and propose a novel method ICOM to address it. Our method is based on the CLT diffusion model, where different parties compete to activate nodes. ICOM estimates the users' opinions towards the target type of items (competing items) and optimizes the seed selection collectively by exploiting the information from competitors. ICOM also utilizes the iterative inferences to improve the performance of opinions and reduce the time complexity when competitors' seed selections are unknown. Based on three real-world social networks, the experiment results validate the effectiveness and efficiency of our proposed model ICOM.

REFERENCES

[1] S. Bhagat, A. Goyal, and L. V. Lakshmanan. Maximizing product adoption in social networks. In *Proceedings of the tth ACM International Conference on Web Search and Data Mining (WSDM '12)*, pages 603–612, 2012.

[2] S. Bharathi, D. Kempe, and M. Salek. Competitive influence maximization in social networks. *Internet and Network Economics*, pages 306–311, 2007.

[3] A. Borodin, Y. Filmus, and J. Oren. Threshold models for competitive influence in social networks. In *Proceedings of International Workshop on Internet and Network Economics (WINE '10)*, volume 6484, pages 539–550, 2010.

[4] C. Budak, D. Agrawal, and A. El Abbadi. Limiting the spread of misinformation in social networks. In *Proceedings of the 20th International Conference on World Wide Web (WWW '11)*, pages 665–674, 2011.

[5] T. Carnes, C. Nagarajan, S. M. Wild, and A. Van Zuylen. Maximizing influence in a competitive social network: a follower's perspective. In *Proceedings of the 9th International Conference on Electronic Commerce (ICEC '07)*, pages 351–360, 2007.

[6] H.-H. Chen, Y.-B. Ciou, and S.-D. Lin. Information propagation game: a tool to acquire humanplaying data for multiplayer influence maximization on social networks. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '12)*, pages 1524–1527, 2012.

[7] 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 *Proceedings of the 11th SIAM International Conference on Data Mining (SDM '11)*, pages 379–390, 2011.

[8] W. Chen, C. Wang, and Y. Wang. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '10)*, pages 1029–1038, 2010.

[9] W. Chen, Y. Wang, and S. Yang. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '09)*, pages 199–208, 2009.

[10] W. Chen, Y. Yuan, and L. Zhang. Scalable influence maximization in social networks under the linear threshold model. In *Proceedings of the 10th IEEE International Conference on Data Mining (ICDM '10)*, pages 88–97, 2010.

[11] A. Das and D. Kempe. Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. *arXiv preprint arXiv:1102.3975*, 2011.

[12] A. Gionis, E. Terzi, and P. Tsaparas. Opinion maximization in social networks. In *Proceedings of the 13th SIAM International Conference on Data Mining (SDM '13)*, pages 387–395, 2013.

[13] A. Goyal, W. Lu, and L. V. Lakshmanan. Celf++: optimizing the greedy algorithm for influence maximization in social networks. In *Proceedings of the 20th International Conference Companion on World Wide Web (WWW '11 Companion)*, pages 47–48. ACM, 2011.

[14] A. Goyal, W. Lu, and L. V. Lakshmanan. Simpath: An efficient algorithm for influence maximization under the linear threshold model. In *Proceedings of the 11th IEEE International Conference on Data Mining (ICDM '11)*, pages 211–220. IEEE, 2011.

[15] G. Guo, J. Zhang, D. Thalmann, and N. Yorke-Smith. Etaf: An extended trust antecedents framework for trust prediction. In *Proceedings of the International Conference on Advances in Social Network Analysis and Mining (ASONAM '14)*, pages 540–547, 2014.

[16] X. He, G. Song, W. Chen, and Q. Jiang. Influence blocking maximization in social networks under the competitive linear threshold model. In *Proceedings of the 12th SIAM International Conference on Data Mining (SDM '12)*, pages 463–474, 2012.

[17] D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '03)*, pages 137–146, 2003.

[18] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8), 2009.

[19] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance. Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '07)*, pages 420–429. ACM, 2007.

[20] S.-C. Lin, S.-D. Lin, and M.-S. Chen. A learning-based framework to handle multi-round multi-party influence maximization on social networks. In *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15)*, pages 695–704, 2015.

[21] X. Liu, C. Aggarwal, Y.-F. Li, X. Kong, X. Sun, and S. Sathe. Kernelized matrix factorization for collaborative filtering. In *Proceedings of the 16th SIAM International Conference on Data Mining (SDM '16)*, pages 378–386. SIAM, 2016.

[22] X. Liu, X. Kong, and P. Yu. Active opinion maximization in social networks. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '18)*, 2018.

[23] X. Liu, Y. Song, C. Aggarwal, Y. Zhang, and X. Kong. Bicycle: Item recommedation with life cycles. In *Proceedings of the 17th IEEE International Conference on Data Mining (ICDM '17)*, pages 297–306, 2017.

[24] H. Zhang, T. N. Dinh, and M. T. Thai. Maximizing the spread of positive influence in online social networks. In *Proceedings of the 33rd IEEE International Conference on Distributed Computing Systems (ICDCS '13)*, pages 317–326, 2013.

[25] J. Zhang, S. Wang, Q. Zhan, and S. Y. Philip. Intertwined viral marketing in social networks. In *Proceedings of the International Conference on Advances in Social Network Analysis and Mining (ASONAM '16)*, pages 239–246, 2016.