Reinforcement-Learning-Based Dynamic Opinion Maximization Framework in Signed Social Networks

Qiang He, Yingjie Lv, Xingwei Wang, Jianhua Li, Min Huang, Lianbo Ma, and Yuliang Cai

***Abstract*—Dynamic opinion maximization (DOM) is a significant optimization issue, whose target is to select some nodes in the network and prorogate the opinions of network nodes, and produce the optimum node opinions. Until now, the node opinions of related researches are unchanged and seldom focus on social relationships. In the real scenario, the dynamic process of network nodes over time and user preference have existed. Therefore, this article proposes the *Q*-learning-based DOM (QDOM) framework in signed social networks to solve the OM problem, which is made up of two phases: 1) the activated dynamic opinion model and 2) the *Q*-learning-based seeding process. We propose the activated dynamic opinion model based on stateless *Q*-learning theory to derive the opinion propagation process. Moreover, we design the *Q*-learning-based seeding algorithm to obtain the seed nodes. The experimental results on the four signed social network data sets demonstrate that the proposed framework outperforms the state-of-the-art approaches on positive opinions, the ratio of positive opinions, and activated nodes.**

***Index Terms*—Activated dynamic opinion model, *Q*-learning, seeding algorithm, social trust.**

Manuscript received 16 August 2021; revised 28 November 2021; accepted 6 January 2022. Date of publication 11 January 2022; date of current version 13 March 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 61872073; in part by the Liaoning Revitalization Talents Program under Grant XLYC1902010; in part by the Key Project of Natural Science Foundation of China Under Grant 62032013; in part by NFSC through the Major International (Regional) Joint Research Project under Grant 71620107003; in part by the Science and Technology Planning Project of Shenyang City under Grant 21-108-9-19; in part by the Doctor Startup Foundation of Liaoning Province under Grant 2021-BS-055; in part by the Fundamental Research Funds for the Central Universities under Grant N2119004 and Grant N2119007; and in part by the National Natural Science Foundation of China under Grant 62002055. *(Corresponding author: Xingwei Wang.)*

Qiang He, Yingjie Lv, and Jianhua Li are with the College of Medicine and Biological Information Engineering, Northeastern University, Shenyang 110169, China.

Xingwei Wang is with the State Key Laboratory of Synthetical Automation for Process Industries, College of Computer Science and Engineering, Northeastern University, Shenyang 110169, China (e-mail: heqiangcai@gmail.com).

Min Huang is with the State Key Laboratory of Synthetical Automation for Process Industries, College of Information Science and Engineering, Northeastern University, Shenyang 110819, China.

Lianbo Ma is with the College of Software Engineering, Northeastern University, Shenyang 110169, China.

Yuliang Cai is with the College of Information Science and Engineering, Northeastern University, Shenyang 110819, China.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCDS.2022.3141952.

Digital Object Identifier 10.1109/TCDS.2022.3141952

I. Introduction

With the rapid growth of the Internet, millions of network users can communicate and chat with each other using social media (such as Facebook, Weibo, and Wechat). As a result, social network analysis becomes a potential and prevailing field, which can apply useful social information and relationships to business, political stability, and culture [1], [2]. As an essential content, the influence maximization (IM) problem have been studied and developed due to its widespread applications, e.g., viral marketing [3], [4], and recommendation system [5], [6] and rumor blocking [7], [8]. In general, the main target of IM is to leverage some seed nodes, propagate the product to others through their social relationship, and finally achieve the largest influenced nodes [9]–[11].

Recently, some varieties of the IM have been explored, such as opinion maximization (OM) [12], [13], dynamic IM (DIM) [14], [15], and even dynamic OM (DOM) [16], [17]. Specifically, the OM problem studies the maximization of overall opinions and the DIM problem studies the IM in a dynamic situations. In this case, the DOM studies the dynamic opinion process and the maximization of overall opinions. As illustrated in Fig. 1, the DOM problem is the extension of the IM, OM, and DIM. We can formulate the DOM as an optimization problem in the social networks, i.e., we can use some seed nodes to spread positive opinions toward the object, and promote the positive opinion spread and enhance the social reputations of the object. From the above analysis, we can find that the DOM has potential application values in business and politics promotion [18], [19]. For example, some companies would like to improve the reputation of a certain new product. Normally, the company will leverage some effective methods to promote the new product, arise the attention of customers and receive a good reputation from the customers. Some research results on the DOM can be used to achieve the promotion of a new product through seeding some efficient seeding algorithms at this point.

The core components of DOM mainly include dynamic opinion propagation and initial seed nodes. The related researches on the DOM mainly include IM, OM, and DIM, which focus on the choice of seed nodes. To be specific, the IM mainly adopts greedy algorithms, heuristic algorithms, or their improved algorithms to select seed nodes with the independent-cascade (IC) model [20], [21]

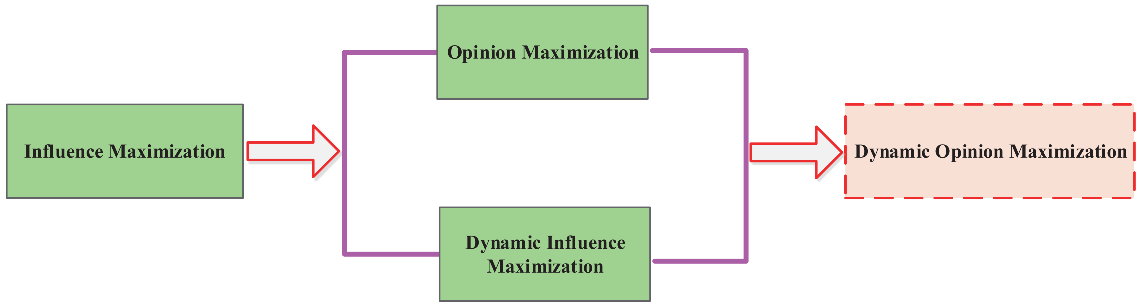


Fig. 1. Relationship between influence maximziation and dynamic opinion maximization.

or linear-threshold (LT) model [22], [23]. The OM mainly leverages similar algorithms by improving the IC and LT model with user opinions. The DIM mainly utilizes adaptive seeding algorithms to determine seed nodes. In this article, the three problems (IM, OM, and DIM) are merged to the DOM problem through using dynamic user opinions and user preferences (positive opinions, neutral opinions, and negative opinions), which make the optimization problem more complex and more challenging. Presently, the researchers concentrate on the maximization of the whole opinions or the maximization of positive opinions. They mainly consider the dynamic process of user opinions and propose some algorithms to select seed nodes.

In a word, although some propagation models and seeding algorithms have been proposed to settle the DOM problem, they are facing the following problems: 1) current propagation models mainly concentrate on the activation process or dynamic opinion process. However, the combination of them has not been explored in depth. In addition, current simple opinion formation models (such as the bounded confidence model [25], Hegselmann–Krause model [26], voter model [27], and Axelrod model [28]) hardly reflect nodes' decision intelligently and 2) the dynamic opinion problem is similarly used in social networks, but little attention to signed social networks. The trust/distrust relationships between edges are also an integral part of the DOM [29], [30]. Therefore, in this article, we present a Q-learning-based DOM (QDOM) framework in signed social networks, which is made up of two parts: 1) the activated dynamic opinion model and 2) the Q-learning-based seeding algorithm. This article has the following three contributions.

1) We mathematically define the DOM problem through integrating positive, neutral, and negative opinions. We also construct the objective function (i.e., overall positive opinions).
2) We solve the DOM problem using the QDOM framework, including an activated dynamic opinion model and Q-learning-based seeding algorithm. In particular, the activated dynamic opinion model adopts stateless Q-learning theory to derive the opinion propagation process. We propose the Q-learning-based seeding algorithm to adaptively obtain seed nodes and achieve large positive opinion propagation.
3) The experiments are proved on four signed data sets. The results show that our framework obtains more superior positive opinions, the ratio of positive opinions, and activated nodes.

The remainder of this article is organized as follows. Section II reviews the related work. In Section III, we describe the DOM problem. We propose the DOM framework in Section IV. The experimental results are given in Section V. Eventually, Section VI summaries our work.

II. Related Work

Our work is associated with three kinds of studies: 1) influence maximization; 2) dynamic influence maximization; and 3) opinion maximization. Moreover, we give Table I to clearly show the properties of different methods.

A. Influence Maximization

Kempe *et al.* [31] first investigated the IM problem and designed a greedy algorithm to obtain the seed nodes. Then, Shen *et al.* [32] studied the influence of heterogeneous networks, devised a cross-network learning model, and focused on the selection scheme of seed nodes and graph sparsification. The seed selection was described as a cross-network node prediction and was derived by the greedy algorithm. Chen *et al.* [33] designed a similarity-aware IM (SIM) model to estimate the influence propagation by joining the nodes' spatio temporal behavior. Based on historical check-ins, the authors calculated the similarity between nodes and devised the propagation to the consumption (PTC) model to catch user behavior. Yin *et al.* [34] studied the IM and designed an information propagation framework for advertisement recommendation. A novel signed-Pagerank (SPR) algorithm was designed to select the initial seed nodes through joining positive and negative relationships.

To conclude, the IM problem assumes that each node has positive opinions toward the new product and aims to maximize the influenced nodes.

B. Dynamic Influence Maximization

Vaswani and Lakshmanan [35] studied adaptive offline strategies for two problems: 1) MaxSpread (to maximize the spread of influence given the number of seeds and a time horizon) and 2) MinTss (to minimize the number of seeds and an expected number of target users to be influenced given a time horizon). In particular, the authors proposed theoretical bounds and empirical results for an adaptive strategy and quantified its practical

TABLE I
COMPARISONS OF DIFFERENT APPROACHES

Methods	User opinions	Dynamics	User preferences	Signed networks
[31, 32, 33]	×	×	×	×
[34]	×	×	×	√
[35, 36, 37, 38, 39]	×	√	×	×
[40, 41, 42]	√	×	√	×
[19, 43, 45]	√	√	√	×
[44, 46]	√	√	×	×
This paper	√	√	√	√

benefit over the nonadaptive strategy. Tong *et al.* [36] studied the seed user selection in an adaptive way. A greedy strategy was proposed to find the approximate solution. Moreover, the authors designed a heuristic algorithm for a better scalability guarantee. Song *et al.* [37] explored the influential node tracking (INT) problem in the dynamic scene. The main goal of the INT problem was to select some network nodes to obtain the largest influence propagation. Through dynamic network structure, the authors proposed an upper bound interchange greedy algorithm. Huang *et al.* [38] studied the adaptive target profit maximization (TPM) problem, and proposed two adaptive greedy algorithms to obtain the approximate solution. Guo and Wu [39] studied the revenue maximization problem by selecting some initiators and designed an adaptive greedy algorithm to determine the seed nodes.

In summary, the DIM problem usually adopts adaptive algorithms to generate the seed nodes and achieve the maximum influenced nodes.

C. Opinion Maximization

As presented, some methods have been proposed for the OM problem. For static user opinions, Chen *et al.* [12] first designed a negative opinion-based IC model. Chen *et al.* [40] explored the negative-aware IM problem to maximize positive nodes while minimizing negative nodes. Liu *et al.* [41] explored the active OM problem to achieve the largest opinion propagation.

For dynamic user opinions, Zhang *et al.* [42] designed the opinion-based cascading model with the LT model and considered the opinion dynamics of activated nodes. Ahmadinejad *et al.* [43] investigated the generation of opinion leaders to guide the others' opinions and proposed a polynomial-time algorithm to generate the maximum opinion spread.

Hudson and Khamfroush [44] devised a behavioral IC (BIC) model by integrating the personalities and opinions of user nodes in the process of information propagation. The authors leveraged this opinion propagation model to study the OM problem. Nayak *et al.* [45] proposed a novel approach to solve the OM problem through effective information propagation. The social interactions and evolution of opinions were modeled as a dynamic Bayesian network (DBN). The authors developed centralized and decentralized learning methods to get approximate results.

Moreover, reinforcement learning has been used in some dynamic scenarios. For example, Kabra *et al.* [59] proposed the multimodel contextual reinforcement learning for real-time and customized recommendations. First, the feature was user-item interactive state embedding which leveraged item information and its weightage. In addition, the authors devised the contextual cluster exploration strategy to improve the item-choice recommendations. The third novelty was an item-based multiagent framework to solve the case of sparsely chosen items. Huang *et al.* [60] proposed a hierarchically structured reinforcement learning approach to address the challenges of planning for generating coherent multisentence stories for the visual storytelling task. Within the framework, the task of generating a story given a sequence of images was divided across a two-level hierarchical decoder. Abhyuday *et al.* [61] studied the properties of off-policy reinforcement learning algorithms when applied to a real-world clinical scenario. The authors evaluated standard off-policy training methods on ventilation and sedation control and off-policy evaluation methods in the context of this problem.

From the above analysis, we can observe that those researches considered the static opinion and the dynamic opinion process. In the real scenario, the activation process and the opinion dynamic process of network nodes should be intimately connected. The current OM problem is focused on social networks, but little attention to signed social networks. Therefore, we study the DOM problem in the signed scene using an activated dynamic opinion model and Q-learning-based seeding algorithm by simultaneously addressing the above challenges.

III. PROBLEM FORMATION

In general, the social network can be mapped to a directed graph $G = (V, E, W)$, in which $V = \{v_1, v_2, \ldots, v_N\}$ (N is the number of nodes), $E = \{e_1, e_2, \ldots, e_M\}$ (M is the number of edges), and $W = \{w_{ij}, i, j \in [1, N]\}$ represents the weight from v_i to v_j. In this article, we consider the signed social network, thus $-1 \leq w_{ij} \leq 1$. Moreover, the node opinion is varied from -1 to 1. If an individual is positive toward the product, its opinion is between $x \in (0, 1]$, if an individual is negative toward the product, its opinion is between $x \in (-1, 0)$ and if an individual is neutral toward the product, its opinions is 0. To further describe the DOM problem, we give the following two definitions.

Definition 1 (Seed Nodes): They are some selected nodes from the network by the specific algorithms to propagate the positive opinion to other nodes. Here, the seed nodes adopt a positive opinion value (i.e., $+1$) on the product and their opinions cannot be changed.

Definition 2 (Dynamic Opinion Maximization): Given node set V and S ($S \subset V$), we select a subset of initial seed nodes from V by the seeding algorithm, leverage the seed nodes to propagate the positive opinions on the product, and achieve the largest overall positive opinions at time t, i.e., $\sum_{v_i \in V^+} x_i^t$. The mathematical form is as follows:

-maximize:

$$F(S) = \sum_{v_i \in V^+} x_i^t \tag{1}$$

-subject to:

$$S = \bigcup_{i \in I^*} \{\gamma_i v_i\} \tag{2}$$

$$\gamma_i = \begin{cases} 1, & \text{if } v_i \text{ is seed node} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

$$\sum_{i \in I^*} \gamma_i = m. \tag{4}$$

Here, $F(S)$ is the overall positive opinions at time t. $I^* = \{1, 2, \ldots, N\}$, $-1 \leq x_i^t \leq 1$, m is the size of seed nodes and x_i represents the opinion of v_i (i.e., node i), which is computed and updated in Section IV-A. In this article, we propose the activated dynamic opinion model to obtain the dynamic opinion of network nodes. Therefore, the node opinions of the DOM problem are dynamically adjusted over time. From the introduced OM problem above, compared with the traditional IM problem, we can find that the node opinions are continuous and the target does not have monotonic and submodule properties. Therefore, to address the OM problem, we subsequently propose the QDOM framework.

IV. Dynamic Opinion Maximization Framework

Similar to the conventional IM problem, the DOM problem is also NP-hard. Therefore, to get the approximate solution, we propose the QDOM framework, including an activated dynamic opinion model and Q-learning-based seeding algorithm. Specifically, the activated dynamic opinion model adopts the stateless Q-learning theory while we further propose a Q-learning seeding algorithm to adaptively select the seed nodes. Next, we will introduce the activated dynamic opinion model and Q-learning-based seeding algorithm in detail.

Algorithm 1 Activated Dynamic Opinion Model

Input: Network node set G and time T;
Output: Activated nodes and the node opinions;
Obtain the activated nodes by using LT model;
for $t = 1$ to T, **do**
  for $i = 1$ to N, **do**
    Compute the Q value of node i at time t;
    Choose x_i^t by $1 - \varepsilon$ greedy algorithm;
  end for
  Acquire the opinions of N nodes at step t;
end for

A. Activated Dynamic Opinion Model

In general, IC and LT models only consider the activation process, but focus less on dynamic opinion propagation. Thus, we propose the activated dynamic opinion model, which combines the activation process with dynamic opinion. In Algorithm 1, we give the pseudo code of activated dynamic opinion model and the detailed descriptions are shown as follows.

In LT model, each user u_i is associated with a threshold $\theta_i \in [0, 1]$. At time t of the LT model, an inactive user v_i can be activated when $\sum_{v_j \in C^{(t-1)}} w_{ji} > \theta_i$, where $C^{(t-1)}$ is the set of active users before stage t starts. Then, we describe the whole process of dynamic opinion propagation.

Moreover, we strive to address this problem through the reinforcement learning-based method [46], [47], which is a type of machine learning technique that enables each node to learn in an interactive environment by trial and error based on the feedback of its actions and experiences and thus eventually obtains its optimal rewards (utility).

We leverage the stateless Q-learning approach [48], [49] to show continuous interaction between nodes for the following reasons. First, each node has its action (whether the opinion difference between it and its in-degree neighbor node is less than a specific threshold) over time. During the whole process, there is no state change, that is, the position of each node is unchanged. Therefore, the dynamic opinion model can be derived

$$Q\left(x_i^{t+1}\right) = \underbrace{Q\left(x_i^t\right)}_{\text{old value}} + \underbrace{\beta_t}_{\text{learning rate}} \overbrace{\left[\underbrace{R\left(x_i^t\right)}_{\text{reward}} - Q\left(x_i^t\right)\right]}^{\text{learned value}} \tag{5}$$

where x_i^t is the selected opinion of node i at time t; $Q(x_i^t)$ is the Q-value for x_i^t and the opinion of node i at time t; for any node x_i^t, $-1 \leq x_i^t \leq 1$. Therefore, if $Q(x_i^t) \geq 1$, we set it as 1; if $Q(x_i^t) \leq -1$, we set it as -1. $R(x_i^t)$ is the expected reward for x_i^t. $R(x_i^t) = \sum_{j \in N_i^{\text{in}}} w_{ji} * r(x_j^t)$. If the difference of x_i and its in-degree neighbor x_j at time t is less than a specific threshold δ, (i.e., $|x_i^t - x_j^t| \leq \delta$), the payoff $r(x_i^t)$ is $+1$, otherwise -1.

At each time, node v_i selects the best-response action with the highest Q-value with a probability of 1-ε (i.e., exploitation), or select any other node randomly with a probability of ε (i.e., exploration).

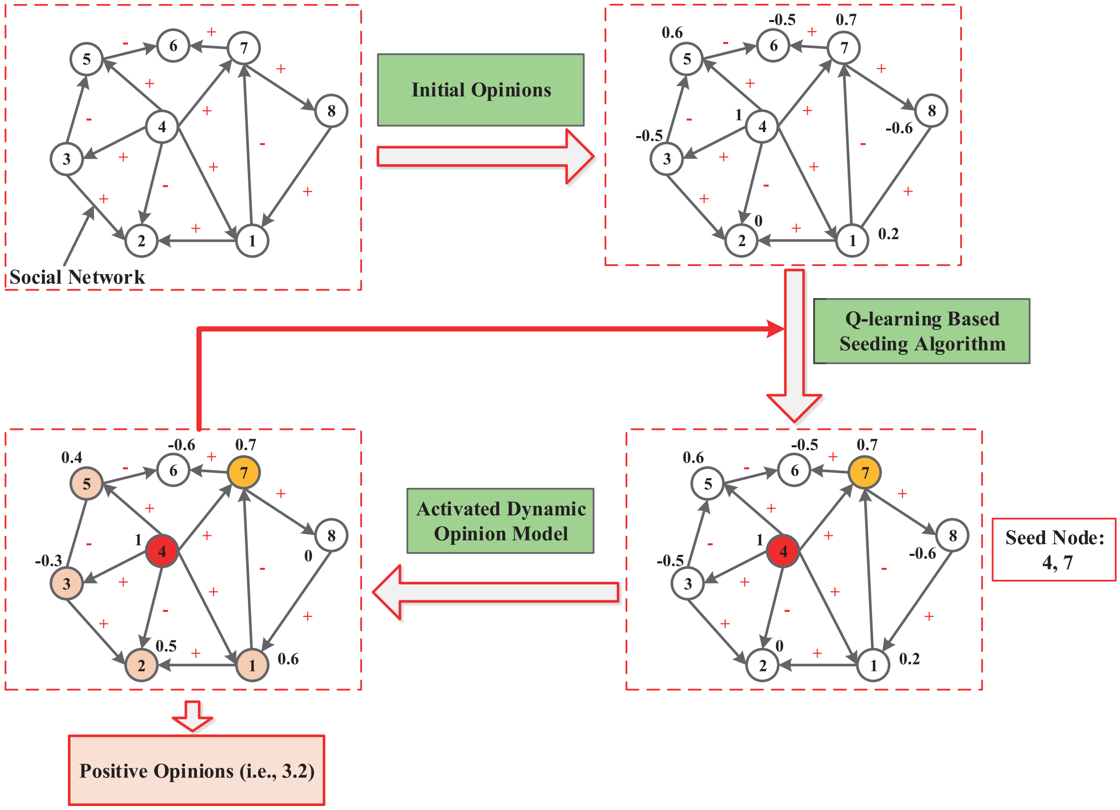


Fig. 2. Example of the OM process.

B. Q-Learning-Based Seeding Process

Different from the existing IM problem, the objective function of DOM has no monotonicity and submodularity, which makes the DOM problem more challenging. Moreover, considering the dynamics of network nodes over time in Section IV-A, we should design an effective seeding scheme. In this view, we propose a Q-learning-based seeding algorithm to adaptively generate seed nodes. To better explain the seeding process, we first define the reward, action, and the state of Q-learning.

Reward: The opinion propagation process is the environment. Reward represents the positive opinions of network nodes using the strategy at the current state, that is $F(S)$.

Action: It is significant for the DOM framework to design the action. Normally, the action is designed to select certain activated nodes. However, if we select one action from all possible strategies for each step, it will cost too much computation space. Therefore, we design a simple and effective action strategy, that is, select some potential strategies. Here, we consider four potential strategies (i.e., MaxDegree, Blocking, Subgreedy, and Mixstrategy). In our simulation, we use the four potential strategies, which effectively decrease the complexity of action space. Given the current state, we can choose the optimal potential strategy to effectively solve the OM problem.

State: Each node can be activated or inactivated by other activated nodes. For example, one node and two activated nodes, that is, node i, active nodes a and b. There exist three states: 1) a activates the node i; 2) b activate node i; and 3) node i is not activated by active nodes a and b.

Here, we assume that our proposed activated dynamic opinion model can select any node to make it adjust an opinion. It should be noted that our constructed Q-learning method can decrease the computation complexity dramatically and achieve good scalability through using the potential action strategies rather than all possible action strategies. We leverage the Q-learning function to obtain the whole computation process and $Q_{t+1}(s_t, a_t)$ is updated in the following:

$$Q_{t+1}(s_t, a_t) = (1-\alpha)Q_t(s_t, a_t) + \alpha\left[R_{t+1}(s_t, a_t) + \gamma \max_{a_{t+1}}(Q(s_{t+1}, a_{t+1}))\right] \quad (6)$$

where $R_{t+1}(s_t, a_t)$ is the expected reward of choosing action a, which represents the positive opinions of network nodes using the action a at time t. $\max_{a_{t+1}}(Q(s_{t+1}, a_{t+1})$ is the expected accumulated influence gain at time $t+1$.

The pseudo-code of the Q-learning-based seeding algorithm is summarized in Algorithm 2, which contains four steps: 1) we select our own strategy by using (8) (lines 3 and 4); 2) we obtain the ith seed node and the reward $R_{t+1}(s_t, a_t)$ by the opinion propagation; (lines 5 and 6); 3) we update the seed node set S and the state to obtain the next Q value (lines 7 and 8); 4) we stop steps 2–4 when seed nodes exceed the upper limit.

In Fig. 2, we show an overall process of the DOM. Here, the signed social network is composed of eight nodes and 13 directed edges. We initialize node opinions randomly and start to select the first seed node (node 4) by using a Q-learning-based algorithm, and use the activated dynamic opinion model to update the node opinions. We will complete the above process when the second seed node (node 7) is obtained. Finally, the generated opinion set $\{0.6, 0.5, -0.3, 1, 0.4, -0.6, 0.7, 0\}$ and finally, the number of positive opinions equals to 3.2.

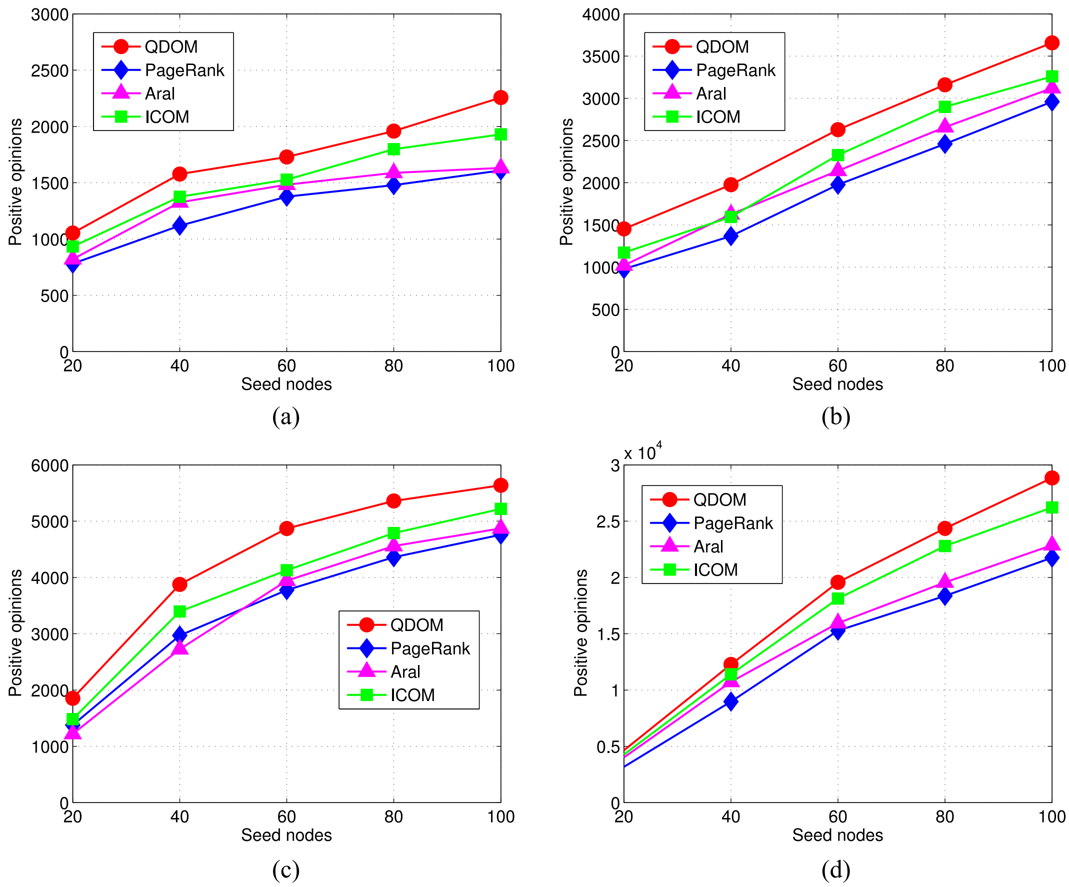


Fig. 3. Positive opinions over the seed nodes on four signed social networks. (a) BA. (b) WE. (C) Wiki-RfA. (d) Slashdot.

Algorithm 2 Q-Learning-Based Seeding Algorithm

Input: graph G, seed node size m;
Output: Seed node set S;
Initialize $S = \emptyset$;
for $i = 1$ to m, **do**
  Calculate own strategy by using Equation 6 exploration scheme;
  Obtain the $i - th$ seed node;
  Obtain the reward $R_{t+1}(s_t, a_t)$;
  $S = S \cup S_i$;
  Update the state and obtain the next Q value;
end for

V. Experimental Evaluations

In this section, we show the data and experiment setup, benchmarks, and experimental results.

A. Data and Experiment Setup

To estimate the performance of our framework, in Table II, we utilize four typical signed social network data sets (including Bitcoin Alpha (BA) [51], Wikipedia Elections (WEs) [52], Wiki-RfA [53], and Slashdot [54]), which have been widely used in influence propagation.

TABLE II
Statistics of Three Signed Social Networks

Networks	Nodes	Links	Average degree
Bitcoin Alpha	3,783	24,186	12.787
Wikipedia elections	7,118	103,675	29.130
Wiki-RfA	10,835	159,388	14.710
Slashdot	77,350	516,575	13.028

Besides, the Jaccard Coefficient [55] is leveraged to estimate the weight between v_i and v_j

$$w_{ij} = \frac{\left|N_i^{\text{in}} \cap N_j^{\text{out}}\right|}{\left|N_i^{\text{in}} \cup N_j^{\text{out}}\right|} \tag{7}$$

where N_i^{in} is the in-degree node set of node i and N_j^{out} is the out-degree node set of node j. It should be noted that our method can be used in both directed and undirected signed social networks. In this article, we mainly introduce the propagation model and seeding algorithm in directed signed social networks. For undirected social networks, we only need to modify the in-degree neighbors of each node in (5) and the Jaccard Coefficient weight in (8) into its degree.

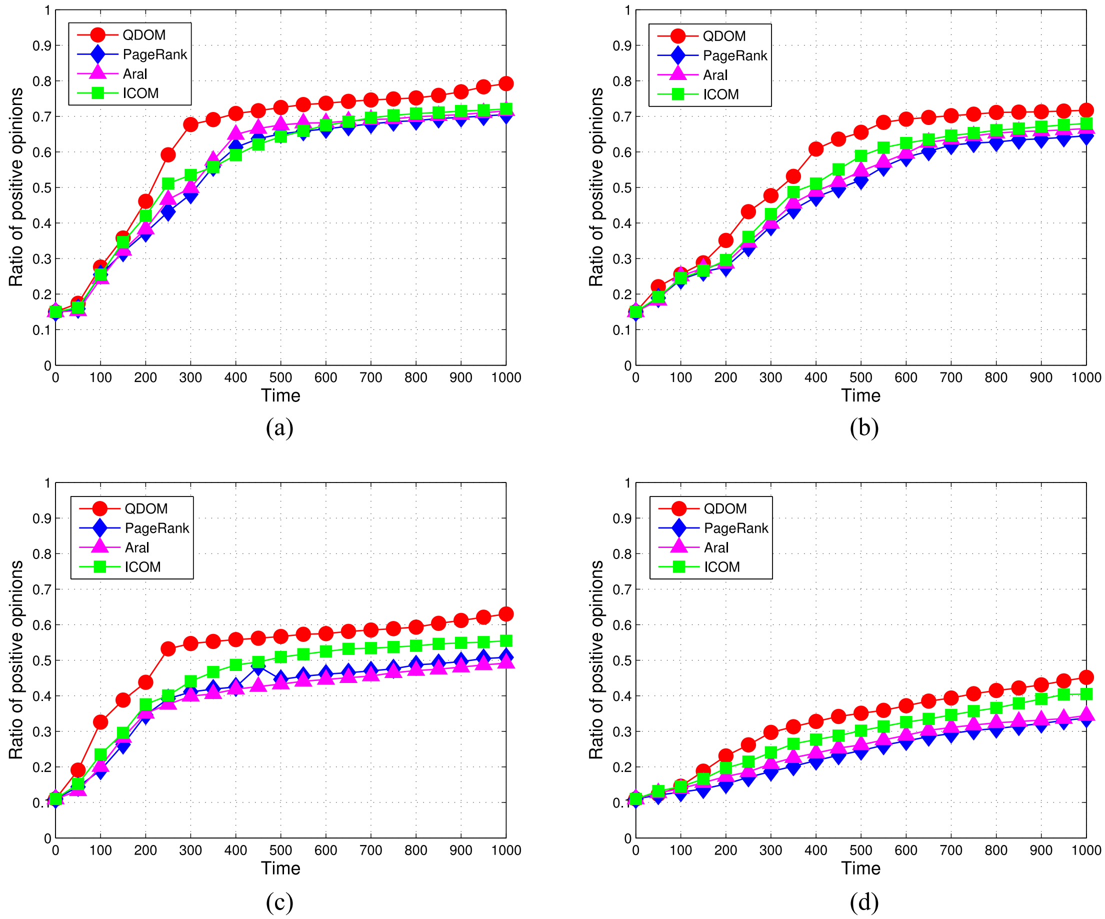


Fig. 4. Ratio of positive opinions over the seed nodes on four signed social networks. (a) BA. (b) WE. (C) Wiki-RfA. (d) Slashdot.

In the experiments, we use another measurement index, named ratio of positive opinions and it is calculated as follows:

$$\mathrm{PR} = \frac{\sum_{v_i \in V^+} x_i}{N}. \tag{8}$$

The ratio of positive opinions represents the proportion of individuals taking positive opinions. To show the superiority of our framework, the initial ratio of positive opinions is set less than 20%. We can explain that when a new product comes into the market, most people are not familiar with the new product. As a result, they may have negative or neutral opinions toward the product.

For each experiment, the initial opinion of each node is randomly generated ranging at $[-1, +1]$. The experiments are repeated 20 times and the results are the average results of all positive opinions. We vary seed nodes between [0, 100] to obtain the different positive opinions. Furthermore, threshold θ_i is set to 0.7, threshold δ is set to 0.8, α and β_t are set to 0.01, and threshold γ is set to 0.05.

Moreover, we compare the performance of our method with three prevailing OM algorithms (i.e., PageRank [56], Aral [57], and ICOM [58]). The related information of the three baseline algorithms is as follows.

1) *PageRank [56]:* PageRank is an effective scheme used to estimate the importance of a particular Web page on the Internet. It has been also a significant index to appraise the influence of network nodes, and the influence score of v_i is acquired by $I_i = [(1-\eta)/N] + \eta * \sum_{v_j \in V} [(a_{ij} * B_{ij})/|N_i^{\text{out}}|]$, where η is a damping factor, a_{ij} denotes the connection relationship between v_i and v_j and if node i points to node j, a_{ij} is 1, otherwise 0, and B_{ij} is the number of the shortest paths connecting v_i and v_j.
2) *Aral and Dhillon [57]:* It selects optimal seed nodes with substantially low degrees and high Burt's constraints.
3) *Iterative-Based Competitive Opinion Maximization (ICOM) [58]:* ICOM leverages iterative inference based on a greedy algorithm to determine the seed nodes.

B. Experimental Results

We first show the performance of different approaches over the seed nodes in terms of positive opinions. Then, we compare the ratio of positive opinions of our approach with other baseline algorithms over time. Finally, we compare the performance of our approach with benchmark algorithms in terms of activated nodes.

Fig. 3 shows positive opinions over the seed nodes on four signed social network data sets. Here, seed nodes vary from 0 to 100, and time is set 200. We can see that positive opinions of four algorithms continually increase over the seed nodes on BA, WE, Wiki-RfA, and Slashdot. The

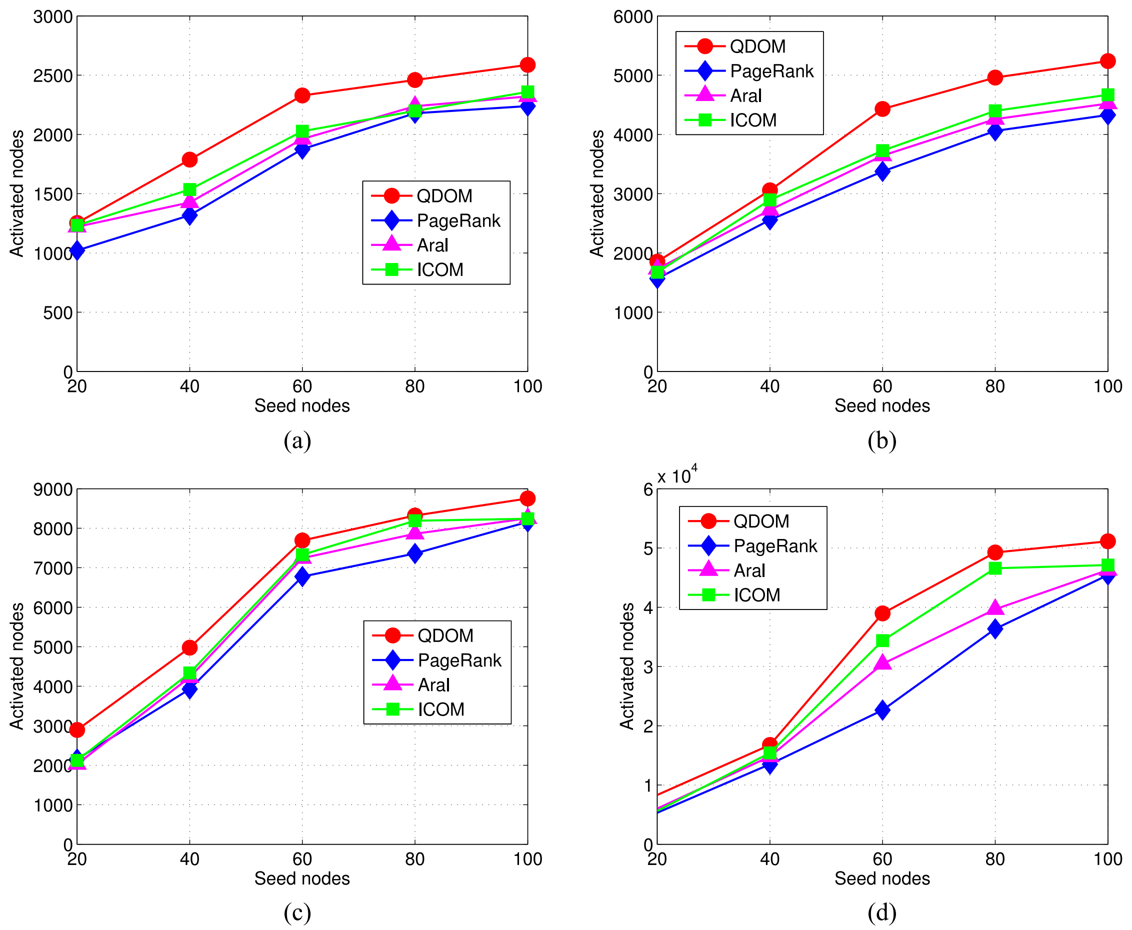


Fig. 5. Activated nodes over the seed nodes on four signed social networks. (a) BA. (b) WE. (C) Wiki-RfA. (d) Slashdot.

proposed QDOM consistently outperforms other approaches on each signed social data set, and CIOM obtains more superior positive opinions than PageRank and Aral on the whole on four data sets. In addition, positive opinions of PageRank are the worst over the seed nodes on BA, WE, and Slashdot. To sum up, the results reflect the validity of QDOM, which includes two components, i.e., the activated dynamic opinion model and Q-learning-based seeding process, respectively. We adopt the activated dynamic opinion process for each individual while proposing the selection of seed nodes with the Q-learning method, which dynamically adjusts seeding strategies and guarantees the quality of seeding process. In contrast, PageRank leverage one-hop measure (i.e., its neighbor nodes) to determine the seed nodes. Aral selects the optimal seed nodes using the heuristic method. ICOM uses an iterative method to determine seed nodes. However, the three baseline methods do not consider the opinion dynamics of network nodes, which cannot effectively adjust the selection of seed nodes.

Fig. 4 shows the ratio of positive opinions over the seed nodes where the time varies from 0 to 1000 and seed nodes are 50. As presented, QDOM outperforms the other compared approaches. In particular, on BA, the ratio of positive opinions of QDOM increases dramatically between [0, 500] and then keeps relatively stable between [500, 1000], and the ratio of positive opinions generated by our method and other approaches changes with a similar trend. On Slashdot, we can also see that although the ratio of positive opinions of QDOM grows slowly, QDOM still outperforms benchmark algorithms. Moreover, our framework promotes the ratio of positive opinions about 64.1% than initial opinions, improves about 8.7% than PageRank, improves about 7.6% than Aral, and improves about 7.1% than ICOM. On Slashdot, our framework improves the ratio of positive opinions about 34.2% than initial opinions, improves about 11.5% than PageRank, improves about 10.7% than Aral, and improves about 4.7% than ICOM.

Fig. 5 shows the positive opinions over the seed nodes on activated nodes. Seed nodes are 50 and time is set 200. From the results, we can see that the activated nodes of four algorithms are rising over the seed nodes and QDOM has larger activated nodes than the benchmark algorithms on four signed social network data sets. For instance, on Slashdot, the activated nodes of QDOM improve greatly, whereas other approaches gradually increase. In summary, from Fig. 5, we can observe that the increase of seed nodes can improve the activated nodes on the whole and the proposed QDOM has evident advantages in promoting positive opinion propagation and producing more superior activated nodes.

C. Discussion

We explore the DOM problem in signed social networks by considering the dynamic opinion process and user's negative opinions, which makes the problem more complex and more challenging than the existing IM problem. As introduced in Section IV, our QDOM devises the activated dynamic opinion models with a stateless Q-learning method to model node opinions and proposes the Q-learning-based seeding algorithm to select the seed nodes. From the above results, we can observe that QDOM obtains more superior positive opinions over the seed nodes, the ratio of positive opinions over time, and activated nodes over the seeds nodes than the compared algorithms, which demonstrates the effectiveness of our framework. In addition, it is important to point out that our proposed Q-learning-based seeding algorithm can easily adjust its strategies by changing the action and improve the scalability and performance of QDOM compared with baseline algorithms. Conversely, the baseline algorithms (PageRank, Aral, and ICOM) cannot effectively determine the seed nodes in the scene of negative opinion and dynamic opinion.

VI. Conclusion

In this article, we have designed the QDOM framework in signed social networks to solve the DOM problem. The proposed framework contains two phases: 1) the activated dynamic opinion model and 2) the Q-learning-based seeding process. We define the DOM problem and give the objective function (i.e., positive opinions of activated nodes). In particular, the objective function can be adjusted according to user's requirements. We proposed the activated dynamic opinion model with stateless Q-learning theory to derive the opinion propagation process. Moreover, based on the designed reward, action, and state, we propose the Q-learning-based seeding algorithm to obtain the seed nodes. The experimental results on four signed network data sets show that QDOM outperforms the baseline algorithms on positive opinions, the ratio of positive opinions, and activated nodes.

For future works, we will explore how to leverage the deep reinforcement learning technique to address the DOM problem in large-scale signed social networks. Moreover, we will also analyze the effectiveness of our proposed opinion models and seeding algorithms from the mathematical theory.

References

[1] E. Bakshy, I. Rosenn, C. Marlow, and L. A. Adamic, "The role of social networks in information diffusion," in *Proc. Int. Conf. World Wide Web*, 2012 pp. 519–528.

[2] P. Liu, D. F. Glas, T. Kanda, and H. Ishiguro, "Two demonstrators are better than one—A social robot that learns to imitate people with different interaction styles," *IEEE Trans. Cogn. Devel. Syst.*, vol. 11, no. 3, pp. 319–333, Sep. 2019.

[3] H. T. Nguyen, M. T. Thai, and T. N. Dinh, "A billion-scale approximation algorithm for maximizing benefit in viral marketing," *IEEE/ACM Trans. Netw.*, vol. 25, no. 4, pp. 2419–2429, Aug. 2017.

[4] Q. He, H. Fang, J. Zhang, and X. Wang, "Dynamic opinion maximization in social networks," *IEEE Trans. Knowl. Data Eng.*, early access, May 4, 2021, doi: 10.1109/TKDE.2021.3077491.

[5] X. Meng *et al.*, "Personalized privacy-preserving social recommendation," in *Proc. AAAI Conf. Artif. Intell.*, 2018, pp. 3796–3803.

[6] L. Ma, X. Wang, X. Wang, L. Wang, L. Shi, and M. Huang, "TCDA: Truthful combinatorial double auctions for mobile edge computing in Industrial Internet of Things," *IEEE Trans. Mobile Comput.*, early access, Mar. 8, 2021, doi: 10.1109/TMC.2021.3064314.

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

[8] A. Manoj, G. K. Palshikar, and S. Baskaran, "Frauds in online social networks: A review," in *Social Networks and Surveillance for Society*, 2019, pp. 1–18.

[9] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in *Proc. ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2009, pp. 199–208.

[10] Q. He *et al.*, "Reinforcement learning based competitive opinion maximization approach in signed social networks," *IEEE Trans. Comput. Social Syst.*, early access, Nov. 13, 2021, doi: 10.1109/TCSS.2021.3120421.

[11] L. Ma, S. Cheng, and Y. I. Shi, "Enhancing learning efficiency of brain storm optimization via orthogonal learning design," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 11, pp. 6723–6742, Nov. 2021.

[12] W. Chen *et al.*, "Influence maximization in social networks when negative opinions may emerge and propagate," in *Proc. SIAM Int. Conf. Data Min. (SDM)*, 2011, pp. 379–390.

[13] Q. He, X. Wang, M. Huang, Z. Lei, Y. Cai, and L. Ma, "TIFIM: A two-stage iterative framework for influence maximization in social networks," *Appl. Math. Comput.*, vol. 354, pp. 338–352, Aug. 2019.

[14] Y. Yang and J. Pei, "Influence analysis in evolving networks: A survey," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 3, pp. 1045–1063, Mar. 2021.

[15] B. Cautis, S. Maniu, and N. Tziortziotis, "Adaptive influence maximization," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2019, pp. 3185–3186.

[16] S. Galhotra, A. Arora, and S. Roy, "Holistic influence maximization: Combining scalability and efficiency with opinion-aware models," in *Proc. Int. Conf. Manage. Data*, 2016, pp. 743–758.

[17] Q. He, X. Wang, B. Yi, F. Mao, and M. Huang, "Opinion maximization through unknown influence power in social networks under weighted voter model," *IEEE Syst. J.*, vol. 14, no. 2, pp. 1874–1885, Jun. 2020.

[18] L. Ma, M. Huang, S. Yang, R. Wang, and X. Wang, "An adaptive localized decision variable analysis approach to large-scale multiobjective and many-objective optimization," *IEEE Trans. Cybern.*, early access, Jan. 21, 2021. [Online]. Available: http://dx.doi.org/10.1109/TCYB.2020.3041212

[19] Y. Cai, H. Zhang, Y. Wang, Z. Gao, and Q. He, "Adaptive bipartite fixed-time time-varying output formation-containment tracking of heterogeneous linear multiagent systems," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 3, 2021, doi: 10.1109/TNNLS.2021.3059763.

[20] A. Goyal, W. Lu, and L. V. Lakshmanan, "CELF++: Optimizing the greedy algorithm for influence maximization in social networks," in *Proc. Int. Conf. Compan. World Wide Web*, 2011, pp. 47–48.

[21] C. Wang, W. Chen, and Y. Wang, "Scalable influence maximization for independent cascade model in large-scale social networks," *Data Min. Knowl. Disc.*, vol. 25, no. 3, pp. 545–576, 2012.

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

[23] C. Castellano, S. Fortunato, and V. Loreto, "Statistical physics of social dynamics," *Rev. Mod. Phys.*, vol. 81, pp. 591–646, May 2009.

[24] A. Gionis, E. Terzi, and P. Tsaparas, "Opinion maximization in social networks," in *Proc. SIAM Int. Conf. Data Min. (SDM)*, 2013, pp. 387–395.

[25] J. Lorenz, "Continuous opinion dynamics under bounded confidence: A survey," *Int. J. Modern Phys. C*, vol. 18, no. 12, pp. 1819–1838, 2007.

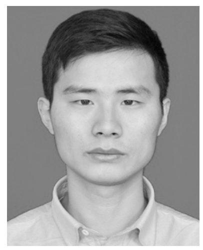
[26] S.R. Etesami and T. Başar, "Game-theoretic analysis of the Hegselmann-Krause model for opinion dynamics in finite dimensions," *IEEE Trans. Autom. Control*, vol. 60, no. 7, pp. 1886–1897, Jul. 2015.

[27] K. Suchecki, V. M. Eguíluz, and M. S. Miguel, "Voter model dynamics in complex networks: Role of dimensionality, disorder, and degree distribution," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*. vol. 72, no. 3, 2005, Art. no. 036132.

[28] R. Axelrod, "The dissemination of culture: A model with local convergence andglobal polarization," *J. Conflict Resolut.* vol. 41, no. 2, pp. 203–226, 1997.

[29] Y. Cai, H. Zhang, Y. Liu, and Q. He, "Distributed bipartite finite-time event-triggered output consensus for heterogeneous linear multi-agent systems under directed signed communication topology," *Appl. Math. Comput.*, vol. 378, Aug. 2020, Art. no. 125162.

[30] Q. He *et al.*, "Positive opinion maximization in signed social networks," *Inf. Sci.*, vol. 558, pp. 34–49, May 2021.
[31] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proc. ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2003, pp. 137–146.
[32] X. Shen, S. Mao, and F.-L. Chung, "Cross-network learning with fuzzy labels for seed selection and graph sparsification in influence maximization," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 9, pp. 2195–2208, Sep. 2019.
[33] X. Chen, Y. Zhao, G. Liu, R. Sun, X. Zhou, and K. Zheng, "Efficient similarity-aware influence maximization in geo-social network," *IEEE Trans. Knowl. Data Eng.*, early access, Dec. 18, 2020, doi: 10.1109/TKDE.2020.3045783.
[34] X. Yin, X. Hu, Y. Chen, X. Yuan, and B. Li, "Signed-PageRank: An efficient influence maximization framework for signed social networks," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 5, pp. 2208–2222, May 2021.
[35] S. Vaswani and L. V. S. Lakshmanan, "Adaptive influence maximization in social networks: Why commit when you can adapt?," 2016, *arXiv:1604.08171*.
[36] G. Tong, W. Wu, S. Tang, and D. Z. Du, "Adaptive influence maximization in dynamic social networks," *IEEE/ACM Trans. Netw.*, vol. 25, no. 1, pp. 112–125, Feb. 2017.
[37] G. Song, Y. Li, X. Chen, X. He, and J. Tang, "Influential node tracking on dynamic social network: An interchange greedy approach," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 2, pp. 359–372, Feb. 2017.
[38] K. Huang, J. Tang, X. Xiao, A. Sun, and A. Lim, "Efficient approximation algorithms for adaptive target profit maximization," in *Proc. IEEE 36th Int. Conf. Data Eng. (ICDE)*, 2020, pp. 649–660.
[39] J. Guo and W. Wu, "A k-hop collaborate game model: Adaptive strategy to maximize total revenue," *IEEE Trans. Comput. Social Syst.*, vol. 7, no. 4, pp. 1058–1068, Aug. 2020.
[40] Y. Chen, H. Li, and Q. Qu, "Negative-aware influence maximization on social networks," in *Proc. AAAI Conf. Artif. Intell.*, 2018, pp. 1–2.
[41] X. Liu, X. Kong, and P. S. Yu, "Active opinion maximization in social networks," in *Proc. ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2018, pp. 1840–1849.
[42] H. Zhang, N. T. Dinh, and M. T. Thai, "Maximizing the spread of positive influence in online social networks," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst. (ICDCS)*, 2013, pp. 317–326.
[43] A. M. Ahmadinejad, S. Dehghani, M. T. Hajiaghayi, H. Mahini, S. Seddighin, and S. Yazdanbod, "How effectively can we form opinions?," in *Proc. Int. Conf. World Wide Web*, 2014, pp. 213–214.
[44] N. Hudson and H. Khamfroush, "Behavioral information diffusion for opinion maximization in online social networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 1259–1268, Apr.–Jun. 2020.
[45] A. Nayak, S. Hosseinalipour, and H. Dai, "Smart information spreading for opinion maximization in social networks," in *Proc. IEEE Conf. Comput. Commun.*, 2019, pp. 2251–2259.
[46] S. Hansen, A. Pritzel, P. Sprechmann, A. Barreto, and C. Blundell, "Fast deep reinforcement learning using online adjustments from the past," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2018, pp. 10567–10577.
[47] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, pp. 237–285, May 1996.
[48] H. Yang and Q. Lu, "Dynamic contextual multi arm bandits in display advertisement," in *Proc. IEEE Int. Conf. Data Min. (ICDM)*, 2016, pp. 1305–1310.
[49] S. Wang and L. Huang, "Multi-armed bandits with compensation," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2018, pp. 5119–5128.
[50] N. Cesa-Bianchi, C. Gentile, G. Lugosi, and G. Neu, "Boltzmann exploration done right," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2017, pp. 6284–6293.
[51] S. Kumar, F. Spezzano, V. S. Subrahmanian, and C. Faloutsos, "Edge weight prediction in weighted signed networks," in *Proc. IEEE Int. Conf. Data Min. (ICDM)*, 2016, pp. 221–230.
[52] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Governance in social media: A case study of the Wikipedia promotion process," in *Proc. Int. AAAI Conf. Weblogs Social Media*, 2010, pp. 1–8.
[53] R. West, H. S. Paskov, J. Leskovec, and C. Potts, "Exploiting social network structure for person-to-person sentiment analysis," *Trans. Assoc. Comput. Linguist.*, vol. 2, pp. 297–310, Sep. 2014.
[54] P. Xu, W. Hu, J. Wu, W. Liu, B. Du, and J. Yang, "Social trust network embedding," in *Proc. IEEE Int. Conf. Data Min. (ICDM)*, 2019, pp. 678–687.
[55] R. Real and J. M. Vargas, "The probabilistic basis of Jaccard's index of similarity," *Systematic Biol.*, vol. 45, no. 3, pp. 380–385, 1996.
[56] S. Brin and L. Page, "Reprint of: The anatomy of a large-scale hypertextual Web search engine," *Comput. Netw.*, vol. 56, no. 18, pp. 3825–3833, 2012.
[57] S. Aral and P. S. Dhillon, "Social influence maximization under empirical influence models," *Nat. Human Behav.*, vol. 2, no. 6, pp. 375–382, 2018.
[58] J. Luo, X. Liu, and X. Kong, "Competitive opinion maximization in social networks," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Min.*, 2019, pp. 250–257.
[59] A. Kabra, A. Agarwal, and A. S. Parihar, "Potent real-time recommendations using multi-model contextual reinforcement learning," *IEEE Trans. Comput. Social Syst.*, early access, Aug. 5, 2021, doi: 10.1109/TCSS.2021.3100291.
[60] Q. Huang, Z. Gan, A. Celikyilmaz, D. Wu, J. Wang, and X. He, "Hierarchically structured reinforcement learning for topically coherent visual story generation," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 8465–8472.
[61] J. Abhyuday, P. Thomas, and H. Yu, "Towards high confidence off-policy reinforcement learning for clinical applications," in *Proc. CausalML Workshop (ICML)*, 2018, pp. 1–9.



Qiang He received the Ph.D. degree in computer application technology from Northeastern University, Shenyang, China, in 2020.

He is currently an Associate Professor with the College of Medicine and Biological Information Engineering, Northeastern University. He has published more than 40 journal articles and refereed conference papers. His research interests include social network analytic, machine learning, and data mining.

Yingjie Lv received the B.A. degree in Chinese language and literature from the Hebei University of Economics and Business, Sijiazhuang, China, in 2020. He is currently pursuing the M.E. degree with Northeastern University, Shenyang, China.

His research interests include social network, mobile edge comouting and reinforcement learning.

Xingwei Wang received the B.S., M.S., and Ph.D. degrees in computer science from Northeastern University, Shenyang, China, in 1989, 1992, and 1998, respectively.

He is currently a Professor with the College of Software, Northeastern University. He has published more than 100 journal articles, books, book chapters, and refereed conference papers. His research interests include cloud computing and future Internet.

Prof. Wang has received several best paper awards.

Jianhua Li received the Ph.D. degree from the School of Medicine and Biomedical Engineering, Northeastern University, Shenyang, China, in 2017.

He is currently a Lecturer with Northeastern University. His current research interests include biomedical data mining, complex networking, and artificial neural network.

Min Huang received the B.S. degree in automatic instrument, the M.S. degree in systems engineering, and the Ph.D. degree in control theory from Northeastern University, Shenyang, China, in 1990, 1993, and 1999, respectively.

She is currently a Professor with the College of Information Science and Engineering, Northeastern University. She has authored/coauthored more than 100 journal articles, books, and refereed conference papers. Her research interests include modeling and optimization for logistics, and supply chain system.

Lianbo Ma received the B.Sc. degree in communication engineering and the M.Sc. degree in communication and information system from Northeastern University, Shenyang, China, in 2004 and 2007, respectively, and the Ph.D. degree from the University of Chinese Academy of Sciences, Beijing, China, in 2015.

He is currently a Professor with Northeastern University. He has published over 90 journal articles, books, and refereed conference papers. His current research interests include computational intelligence and machine learning.

Yuliang Cai received the M.S. degree in control theory and control engineering from the Dalian University of Technology, Dalian, China, in 2017. She is currently pursuing the Ph.D. degree in control theory and control engineering with Northeastern University, Shenyang, China.

Her research interests include multiagent system, pattern recognition, and data mining.