Influence Maximization in Social Network with Negative Influence

Weijia Ju and Ling Chen

Abstract—**Influence maximization is the problem of finding a small set of seed nodes in the social community network and then maximizes the spread of influence under certain influence models. Under the current social network, influence has become an indispensable idea, and it has become the mainstream idea of people. How to expand their influence can be a problem worthy of discussion. Of course, the current influence not only includes positive influences, but also negative influences. There is an ancient Chinese saying that good things do not go out and the things in response to this statement go out, we use the negative effects that may occur in the network and minimize the negative effects of negative influences so that the positive impact can be maximized. Therefore, we introduce a quality factor q to change the negative influence and positively influence the influence of propagation in the network. We use our algorithm in the IC-N model and indicates that seed selection is related to the quality factor q. As the presenter of IC-N model shows that the model maintains a number of nice properties such as submodularity, which allows a greedy approximation algorithm for maximizing positive influence within a ratio of $1 - 1/e$. We also improve the way based on spreading paths. Our experiment results demonstrate the efficiency of our algorithm and show that our algorithm can spread more positive influence.**

Index Terms—**Social influence, influence maximization, negative and positive influence, independent cascade model.**

I. INTRODUCTION

Nowadays, influence plays a very important role in the social. How to maximize the influence propagation become a mainstream issue now. Social networks are becoming important dissemination platforms as they allow efficient interchange of ideas and information. The process of influence diffusion in social networks has been studied in many domains e.g. epidemiology, social median and economics. [1] It has shown that the study of the influence of diffusion in many great use ways such as designing marketing strategy [2], [3], analyzing human behavior [4] and rumor blocking. With the growing popularity of social networks such as Facebook, Myspace and Twitter, the power of viral marketing is more than ever. Therefore, it is important to understand the effective ways to use viral marketing. Influence maximization problem aims to find a seed set which size at most k and the expected number of vertices actived from seed set called influence spread should be the largest. What we do are all based on two models called Independent Cascade Model(ICM) and Linear Threshold

Weijia Ju is with the Department of Computer Science, Yangzhou University, Yangzhou, 225127, China (e-mail: 1069025168@qq.com).
Ling Chen is with the State Key Lab of Novel Software Tech, Nanjing University, Nanjing, 210093, China (e-mail: yzulchen@163.com).

Model(LTM). They are all originally provided by Kemp et.al in [5] and their extensions. The research community has recently studied the algorithmic aspects of maximizing influence in social networks for viral marketing ([5]-[18]). Actually, most of these works ignore one important role of influence maximization. That is the state of influence. We divide the state of influence into positive and negative. These researches just talk about the positive influence but ignoring the negative influence. We all know that negative influence can be fast accepted. For examples, first, children can learn everything very quickly whatever positive or negative. Second, once a restaurant opened, a person may tell you how delicious and also tell you the shortcomings. You may just remember the shortcomings and tell your friends do not go there. Third, actually, people prefer to adopt negative news. That is why the bad news can spread faster and further. In some literature, researchers have shown that when shaping people's decisions, the negative influence or emotions are usually stronger and more dominant. Those persons who disseminate negative opinions are called critics, and those who spread positive impacts are called promoter. Incorporating the generation and dissemination of negative influence into the influence cascade model is very important. So we should try our best to decrease the negative influence spread and expand the positive influence with our ability.

Recently, a lot of researches and related works are expanded on the classic influence maximization. Tong et.al [19] approved new strategies that select seed users in an adaptive manner. They also show that a simple greedy adaptive seeding strategy finds an effective solution with a provable performance guarantee. Wang *et al.* [20] put forward an independent cascade-based model for influence maximization, called IMIC-OC, was proposed to calculate positive influence. They first assumed that influential users spread positive opinions and held positive or negative opinions as their initial opinions at the beginning. Shan et.al [21] study a new type of activation named cumulative activation. They combine the independent cascade model and linear threshold model. A node is cumulatively active if its active probability is beyond a given threshold. They solved two optimization problems which are seed minimization with cumulative activation(SM-CA) problem and influence maximization with cumulative activation(IM-CA) problem.

In some other studies ([14]-[17]), researchers show that negative impact is usually stronger and much more dominant than positive impact in shaping people's decisions. In our paper, we use the model IC-N which extends the IC. IC-N is proposed by Wei *et al.* [18]. They show that their heuristic algorithm has matching influence with a standard greedy approximation algorithm while being orders of magnitude faster. We take advantage of this model and decrease the

machine simulation. Through calculating all the paths with range length and the probabilities in all the edges to justify the quality factor (a parameter to maintain the negative influence).

For influence maximizing, we focus on maximizing the positive nodes of social networks after cascading. We call it positive influence diffusion. What we do is to suppress the negative influence and maximize the positive influence.

In this paper, we utilize the independent cascade with negative model (IC-N model) [18] which extends the independent cascade model. We add the calculation of paths on this basis. We first calculate the paths of all nodes pair. The q we introduced is to calculate the probability of node v can positively active its neighbors. Second, we use the probability values of these node pairs to calculate the influence index $F_q(S, x)$ of each node. The influence index $F_q(S, x)$(explained in therom1) shows the positive influence that x added to seed set *S*. Third, we determine the seed set *S* based on the number of seeds. To sum up, our approach avoids large computational simulation time and has good experiment results in the propagation of positive influence.

II. IC AND IC-N

There are two basic models in influence maximization. One is Linear Threshold Model (LTM), and another is Independent Cascade Model (ICM). This paper uses the ICM as basis model. For a classic independent cascade mode, ICM denoted given a network G (V, E, B), V is all the vertex in the G, E denotes all the edges in the G, B is the probabilities of each edge. The propagation is that each vertex v tries to activate its neighbors u at time t, if the probabilities of is smaller than the activate probabilities, the u will be activated and will activate its neighbors. If u cannot be activated, v will not activate *u*. The other neighbors of u also attempt to activate *u*. That means each vertex can only activate its neighbors once.

The propagation above just talks about positive nodes. If we put the negative nodes in it, the propagation will be different. Wei *et al.* [18] put forwards an idea about influence cascade with negative called IC-N. Based on this, we raise an idea like this. Given a network $G = (V, E, P)$. *V* is the set of vertex. *E* is the set of all edges. Each edges have a probabilities p as the propagate probabilities. The constant called quality factor. Each node in time *t* has one state of three states: nature (inactivated), positive (activated as positive state) and negative (activated as negative state). We select k vertices as activated seeds. And put them as positive state, other nodes as nature state. We set the probability of edge as p, if v is activated in time t. if v is negative activated, then it will use the probability of p to activate u negatively in time *t*+1; if *v* is positive activated, then it will activate its neighbors in time *t*+1. If u is activated, it will has probability p to become positive state or the probability 1-*q* to become negative state; each node v has only one chance to activate its neighbors *u*, *v* cannot activate *u* if fail; an activated vertices will maintain its state. The simulated propagation above ends until no new nodes activated. Under the q, the positive spread range of seed set S is. Our goal is to find a seed set which size is *k* and maximize the positive spread range.

Obviously, the propagation benefits negative influence. In the real social networks, the phenomena exist everywhere. For example, first when people buy something, they always concern its shortcomings. They will focus on the defects of the product when telling others. Second, in some news spread, people always concern some negative news so the negative news spread faster than the positive news. Third, to evaluate a new restaurant, the negative evaluation can be spread easier and can people are also willing to believe that. Although they don't experience the restaurant in person. The examples above all include a question: what if the information they provide are supposititious. Not everyone has the correct judgement about the information. This has led to many cases of intentional falsification of false information (that is, negative influence). That is why the negative influence spread faster and easier than positive influence. We can also find the fact that in the real social network, what other says, they will follow suit. According to these situations, people's opinions will not be static. If a person is actived by negative influence, maybe a time later, the person will change his mind by his own consideration. Such a situation is worth considering in the future work. In fact, some researchers approved the voter model [22], [23] which explained that users can randomly pick up the opinions of their neighbors at each time step. Wang *et al.* [24] add a freezing period based on the voter model. This operation has an advantage that users can change their opinions during this freezing time. After this, users can easily reach a consensus on the social networks.

III. BASIC PROPERTIES

Definition 1: Probability of positive influence spread in paths

We define path from v to u is connected by $e_1, e_2, \cdots e_l$, the probability of e_i is p_i, then the path $L = (e_1, e_2, \cdots e_l)$ from v positively activate *u*'s probability is

$$P(L) = q^l * \prod_{i=1}^{l} p_i$$

Definition 2: Independent paths

We define that if two paths L_1 and L_2 don't have the same edges, we call L_1 and L_2 are independent paths.

Definition 3: Define $P_q(v, u)$ as the probability which vertex *v* can positively activate *u*:

$$P_q(v, u) = \sum_{w \in Tin(u)} P_q(v, w) * Pr(w, u) * q \qquad (1)$$

Here: *Tin(u)* is the set of neighbors of in-edge of *u*. $Tin(u) = \{w| (w, u) \in E\}$,

$Pr(w, u)$ is the spread probability in edge (w, u).

From formula (1), we can see that $P_q(v, u)$ is the sum of all the independent paths from *v* to *u*.

Definition 4: We define $w \subseteq v$ as the subset of the vertex set *v*, the subgraph G_w of G is constructed by *w* and its related

connected edges. Then we can define $P_q(v,u|w)$ as the probability of that which v can activate successfully in subgraph G_w. Correspondingly, $P_q(v,u)$ in definition 3 can be signed as $P_q(v,u|v)$.

Definition 5: We define seed set $S \subseteq V$, for each vertex $u \in V$, when put S as the seed set, the probability of u which can be positive active successfully is:

$$G_q(S,u) = \sum_{V \in S} P(v,u|V - S + v) \qquad (2)$$

Here: $V - S + v$ is the set $(V \backslash S) \cup \{v\}$.

Definition 6: Under the quality factor q, the positive influence of the seed set S estimates:

$$\sigma_G(S,q) = \sum_{u \in V} G_q(S,u) \qquad (3)$$

Lemma 1: First we use $w + x$ to denote the set $w \cup \{x\}$:

$$P_q(v,u|w) = P_q(v,u|w + x) - P_q(v,x|w + x) * P_q(x,u|w + x) \qquad (4)$$

Proof: We can get all the paths from deleting all the paths about x from $w + x$. All the paths about x from v to u have $v \to x \to u$ form. Its probability is the product of $v \to x$ and $x \to u$.

Due to the different activated states of paths $v \to x$ and $x \to u$, we can know that the positive activated probability of $v \to x \to u$ from v is:

$$P_q(v,x|w + x) * P_q(x,u|w + x)$$

Then from deriving, we can get the result:

$$P_q(v,u|w) = P_q(v,u|w + x) - P_q(v,x|w + x) \\ * P_q(x,u|w + x)$$

Proof finished

Lemma 2:

$$G_q(S + x,u) = G_q(S,u) + P_q(x,u|V - S) * [1 - G_q S,x] \qquad (5)$$

Here is the proof:

$$G_q(S + x,u) - G_q(S,u) = \sum_{v \in S+x} P_q(v,u|V - S - x + v) - $$

$$\sum_{v \in S} P_q(v,u|V - S + v) = P_q(x,u|V - S - x + x)$$

$$+ \sum_{v \in S} [P(v,u|V - S - x + v) - P_q(v,u|V - S + v)]$$

$$= P_q(x,u|V - s) + \sum_{v \in S} P_q(v,x|V - S + v) * P_q(x,u|V - S + v)$$

(according to lemma 1)

In the above formula, $P_q(v,x|V - S + v) * P_q(x,u|V - S + v)$ is all the paths of V through x to u are considered in the set $V-S + v$. Since there is no loop link in the directed network (edges represent circular propagation, and such paths are generally not considered), that is, there is no repeated vertex in the path. So v can't be in the path of x to u,

$P(x,u|V - S + v) = P(x,u|V - S)$ is given, so the proof can be rewrite as below:

$$G_q(S + x,u) - G_q(S,u) = P_q(x,u|V - S) + \\ \sum_{v \in S} P_q(V,x|V - S + v) * P_q(x,u|V - S)$$

$$= P_q(x,u|V - S)\left[1 - \sum_{v \in S} P_q(v,x|V - S + v)\right]$$

$$= P_q(x,u|V - S)[1 - G_q(s,x)]$$

From the algorithm, we use greedy select vertex x to join the set S, maximize $I(S + x) - I(S)$. Then we can get theorems.

Theorem1: $F_q(S,x) = \sigma_G(S + x,q) - \sigma_G(S,q)$

$$F_q(S,x) = \sigma_G(S + x,q) - \sigma_G(S,q) \qquad (6)$$

$$= [1 - G_q(S,x)] * \sum_{u \in V \backslash S} P_q(x,u|V - S)$$

Here, we can understand the $F_q(S,x)$ as the influence benefit value of the added seed x. And $F_q(S,x)$ is the key variable to select seed set. We will sort $F_q(S,x)$ of all nodes, and select the large values according to the size of seed set which we defined as k.

Theorem2: Sub modeling

For any function $f: 2^v \to R$ that maps any subset of the finite set V to a real value, f is said to satisfy the submodule, for any subset $S \subseteq V$ and any one of its supersets $T(S \subseteq T \subseteq V)$ and any element other than $T: u \in V \backslash T$, f satisfies $f(S \cup \{u\}) - f(S) \geq f(T \cup \{u\}) - f(T)$. Obviously, our method is consistent with this theorem.

IV. ALGORITHM

A. Variables

$P(v,u)$: $P_q(v,u|V - S)$. Initial value is $\Pr(v,u)$. It updates by using formula (4) when S changes.

G_u: $G_q(S,u)$. Initial value is 0 because the size of S is null. It updates by using formula (5) when S changes.

$F(x)$: $F_q(S,x)$. Initial value is $\sum_{u \in V} P(x,u)$. It updates by using formula (6) when S changes.

B. Algorithm Frame

Algorithm1: main algorithm for influence maximization
Input: $G = (V,E,B)$, directed graph G, probabilities of edges are $\Pr(v,u)$ k: size of seed set q: quality factor L: considered length of paths   Output: S: set of seed set
1、For all $x \in V$, set $G(x) = 0$
2、$S = \phi$
3、For all (v,u), $P(v,u) = \Pr(v,u)$
4、For all vertex x, $F(x) = \sum_{u \in V} P(x,u)$
5、Sort $F(x)$

6、while $/S/<k$ do
   6.1 select the first x in $F(x)$
   6.2 For all $v \in V \backslash S$ and $P(x,v) \neq 0$ update $G(x)$
   6.3 For all (v,u) update $P(v,u)$ by formula (5)
   6.4 For all $x \in V \backslash S$ update $F(x)$ and use dichotomy to insert $F(x)$
 end while

Algorithm 1 shows the main frame of our algorithm. From step 1 to step5 is the initial work. Step 1 initials influence value of all **vertex**. Step 2 defines the seed set S. Step 3 sets probability of each node pair (v, u). Step 4 calculate the influence bounds $F(x)$ of each node. Step 5 sort the $F(x)$ and select the biggest number as first seed. Step 6 has 4 steps and donates the way we select seed set. We will introduce one by one in the next three algorithms.

Algorithm2: update $G(x)$

Input: S: seed set;
   x: new seed in S;
   G: initial value of $G(v)$
Output: G: updated G
1、For all $v \in V \setminus S$ and $P(x, v) \neq 0$ do
  $G(v) = G(v) + P(x, v)[1 - G(x)]$
  End for

Algorithm 2 is the Step 6.2 in algorithm 1. We use the new seed in seed set S update each node v ∈ V\S influence value G (v)

Algorithm3: update P

Input: S: seed set
   X: new seed in S
   P: probability matrix
Output: P: updated probability matrix
1、For all $v \in V$ and $P(v, x) \neq 0$ do
2、 For all $u \in V$ and $P(x, u) \neq 0$ do
  $R(v, u) = P(v, u) - P(v, x) * P(x, u);$
  $P = R;$
  End for
  End for

Algorithm 3 is the step 6.3 of algorithm 1. We use the new seed in seed x set S **update** the probability of each node pair. It is worth mentioning that we need to discard the path that contains the seed node in the node pair.

Algorithm4: update $F(x)$

Input: S: seed set;
   x: new seed;
   F: $F(x)$;
Output: F: updated F
1、For all $v \in V \setminus S$ do
  Get rid of $F(v)$ of F; $F(v) = 0$
2、For all $v \in V$ and $P(v, u) \neq 0$ do
  $F(v) = F(v) + P(v, u)$
  End for
3、$F(v) = F(v) * (1 - G(v))$
4、Insert F into $F(v)$ using dichotomy
  End for

Algorithm 4 shows the step 6.4 of algorithm 1. We use the new seed x and updated $G(v)$ **and** updated $P(v)$ in seed set S update the influence bounds $F(v)$. Every time we calculate the $F(v)$ and rearrange the $F(v)$ to select the biggest node.

V. EXPERIMENT

A. Dataset

The first dataset, Wiki-Vote, is a voting history network from Wikipedia [25], where nodes represent Wikipedia users,

and a directed edge from u to v means v voted on u (for promoting u to adminship). The second dataset, Epinions, is a Who-trust-whom network of Epinions.com [26], where nodes are members of the site and a directed edge from u to v means v trusting u (and thus u has influence to v). The third dataset, Email [27], is a large set of email messages, the Enron corpus, was made public during the legal investigation concerning the Enron corporation. This network is merged by a dynamic network, so the edge weights represent the number of mails between two nodes in a certain period.

TABLE I: STATICS OF THE TWO REAL-WORLD NETWORKS

Dataset	Wiki-Vote	Epinions	Email
Number of nodes	7K	76K	1134
Number of edges	101K	509K	5451
Average degree	26.64	13.4	9.61
Maximal degree	1065	3079	71

B. Experiment Result

Fig. 1, Fig. 2 and Fig.3 donate that our algorithm runs in different datasets. The first one is in the wiki-vote, the second one is in the Epinions and the third one is in the Email. We set the length of paths as four. We use up to 50 seeds and with the quality taking values from 0 to 0.5. When the quality is 0, it means each one are accepted information positively. With the increase of q, negative influence also increases. Take the production of goods, we should always maintain the production of high-quality goods.
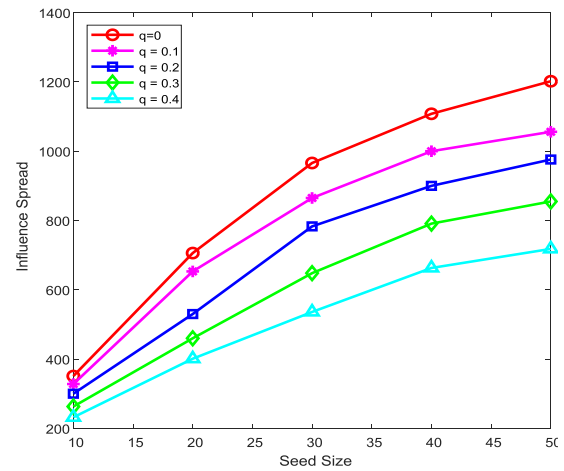

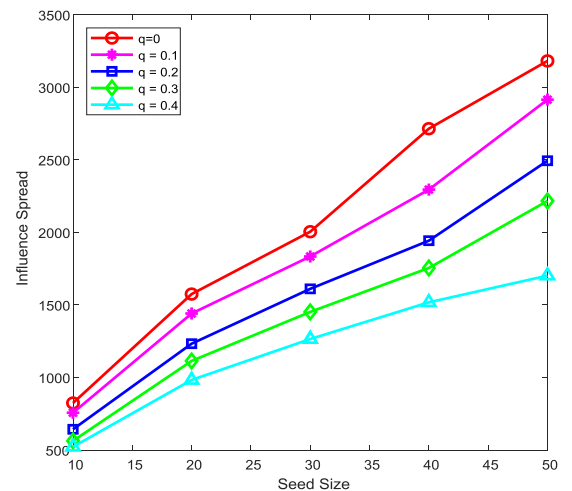Fig. 1. Influence Spread in Wiki-vote under different quality q.

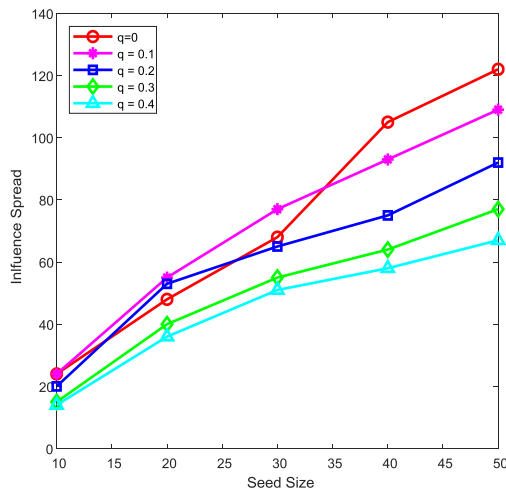Fig. 2. Influence Spread in Epinions under different quality q.

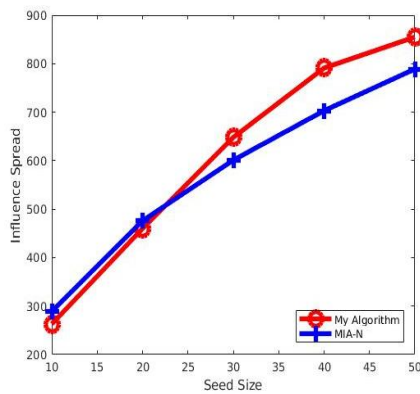Fig. 3. Influence Spread in Email under different quality q.

Fig. 4. Influence Spread vs. MIA-N in Wiki-vote under different the same q.
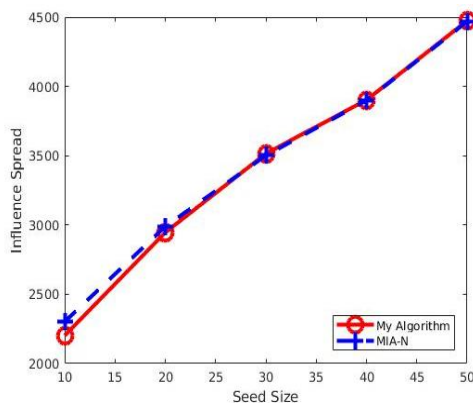

Fig. 5. Influence Spread vs. MIA-N in Epinions under different the same q.

Fig. 4 donates the experiment in wiki-vote dataset and Fig. 5 donates the experiment in Epinions. We can see that our algorithm can spread better than MIA-N [14] in wiki-vote but a little worse than MIA-N in Epinions. So we will continue updating our algorithm when we overcome big data.

VI. CONCLUSION

We can see that our experiment can suppress the negative influence spread and expand the positive influence spread. But we think this still have some inadequacy. First, each node actually should have its own quality factor, but we use the same quality factor. From social network, each person has its own judgement, so we will try to give each person a q to calculate influence in the future work. Second, once nodes successfully negative active its neighbors, the probabilities of propagation should be increased. We all know that the meaning of a person's words may be magnified a lot. Third, [14] allowing different propagation delays along different edges to model the nonuniform interaction frequency between individuals. In addition, we can tap for more parameters through the real social network. We can think of each single point as a matrix. Each node represents a person, a product or any other thing, so they can change even they are actived to positive or negative. Also, this model is based on ICM. If we use LTM to do this, maybe the affect will be better.

REFERENCES

[1] V. Mahajan, E. Muller, and F. M. Bass, "New product diffusion models in marketing: A review and directions for research," *J. Marketing*, vol. 54, no. 1, pp. 1–26, 1990.

[2] J. Goldenberg, B. Libai, and E. Muller, "Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata," *Acad. Marketing Sci. Rev*., vol. 2001, no. 9, pp. 1–19, 2001.

[3] R. M. Bond *et al*., "A 61-million-person experiment in social influence and political mobilization," *Nature*, vol. 489, no. 7415, pp. 295–298, 2012.

[4] L. Fan *et al*., "Least cost rumor blocking in social networks," in *Proc. IEEE 33rd Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jul. 2013, pp. 540–549.

[5] D. Kempe, J. M. Kleinberg, and E´. Tardos, "Maximizing the spread of influence through a social network," in *Proc. the 9th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2003, pp. 137–146.

[6] D. Kempe, J. M. Kleinberg, and E´. Tardos, "Influential nodes in a diffusion model for social networks," in *Proc. the 32nd International Conference on Automata, Languages, and Programming*, 2005, pp. 1127–1138.

[7] M. Kimura and K. Saito, "Tractable models for information diffusion in social networks," in *Proc. the 10th European Conference on Principles and Practice of Knowledge Discovery in Databases*, 2006, pp. 259–271.

[8] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. Van- Briesen, and N. S. Glance, "Cost-effective outbreak detection in networks," in *Proc. the 13th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2007, pp. 420–429.

[9] R. Narayanam and Y. Narahari, "A shapley value based approach to discover influential nodes in social net- works," *IEEE Transactions on Automation Science and Engineering*, 2010.

[10] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in *Proc. the 15th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2009.

[11] W. Chen, C. Wang, and Y. Wang, "Scalable influence maximization for prevalent viral marketing in large scale social networks," in *Proc. 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2010.

[12] Y. Wang, G. Cong, G. Song, and K. Xie, "Community-based greedy algorithm for mining top-k influential nodes in mobile social networks,"

in *Proc. the 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2010.

[13] W. Chen, Y. Yuan, and L. Zhang, "Scalable influence maximization in social networks under the linear threshold model," in *Proc. the 10th IEEE International Conference on Data Mining*, 2010.

[14] G. Peeters and J. Czapinski, "Positive-negative asymmetry in evaluations: The distinction between affective and informtional negativity effects," *European Review of Social Psychology*, 1990.

[15] S. E. Taylor, "Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis," *Psycho-Logical Bulletin*, vol. 110, no. 1, pp. 67–85, 1991.

[16] R. F. Baumeister, E. Bratslavsky, and C. Finkenauer, "Bad is stronger than good," *Review of General Psychology*, vol. 5, no. 4, pp. 323–370, 2001.

[17] P. Rozin and E. B. Royzman, "Negativity bias, negativity dominance, and contagion," *Personality and Social Psychology Review*, vol. 5, no. 4, pp. 296–320, 2001.

[18] W. Chen, A. Collens, R. Cummings, T. Ke, Z. Liu, D. Rincon, X. Sun, Y. Wang, W. Wei, and Y. Yuan, "Influence maximization in social networks when negative opinions may emerge and propagate," in *Proc. the 2011 SIAM International Conference on Data Mining (SDM'2011)*, Mesa, Arizona, U.S.A., April 2011

[19] G. M. Tong, W. L. Wu, S. J. Tang, and D. Z. Du, "Adaptive influence maximization in dynamic social networks," *IEEE/ACM Transactions on Networking*, vol. 25, no. 1, Feb. 2017

[20] Q. Y. Wang, Y. H. Jin, Z. Lin, S. D. Cheng, and T. Yang, "Influence maximization in social networks under an independent cascade-based model," *Physical A.*, vol. 444, pp. 20-34, Oct. 2015

[21] X. H. Shan, W. Chen, Q. Li, X. M. Sun, and J. L. Zhang, *Cumulative Activation in Social Networks*, 2016.

[22] P. Clifford and A. W. Sudbury, "A model for spatial conflict," *Biometrika*, vol. 60, pp. 581–588, 1973.

[23] R. A. Holley and T. M. Liggett, "Ergodic theorems for weakly interacting infinite systems and the voter model," *Ann. Probab.*, vol. 3, no. 4, pp. 643–663, 1975.

[24] Z. Wang, Y. Liu, Y. Zhang, and Z. Wang, "Freezing period strongly impacts the emergence of a global consensus in the voter model," *Sci. Rep.*, vol. 3597, pp. 1–10, 2014.

[25] J. Leskovec. Wikipedia vote network. [Online]. Available: http://snap.stanford.edu/data/wiki-Vote.html

[26] J. Leskovec. Epinions social network. [Online]. Available: http://snap.stanford.edu/data/soc-Epinions.html

[27] B. Klimt and Y. M. Yang. Language Technology Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA. [Online]. Available: http://www-2.cs.cmu.edu/~enron/

Weijia Ju is currently a postgraduate in the school of information engineering at Yangzhou University in Jiangsu, China. His main research interests are influence maximization in social networks, artificial intelligence. He has received the national scholarship and published a paper in an international journal before.

Ling Chen is a professor in the Computer Science Department, Yangzhou University, China. His research interests include data mining and computational intelligence. He has co-edited 6 books/proceedings, and published more than 300 research papers including over 120 journal papers. He has received many awards from government and agencies. His recent research has been supported by National Science Foundation of China, Foundation of Chinese National Key Lab of Novel Software Technology. He has served as the program committee chair or member for several major international conferences. Professor Chen is a member of IEEE CS society and ACM, a senior member of the Chinese Computer Society.