The Thirty-Second AAAI Conference
on Artificial Intelligence (AAAI-18)

Negative-Aware Influence Maximization on Social Networks*

Yipeng Chen,[†] Hongyan Li,[†] Qiang Qu[‡,†]
[†]MOE Key Laboratory of Machine Perception
School of Electronics Engineering and Computer Science
Peking University
[‡]Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences
chenyipeng@pku.edu.cn, lihy@cis.pku.edu.cn, qiang@siat.ac.cn

Abstract

How to minimize the impact of negative users within the maximal set of influenced users? The Influenced Maximization (IM) is important for various applications. However, few studies consider the negative impact of some of the influenced users. We propose a negative-aware influence maximization problem by considering users' negative impact. A novel algorithm is proposed to solve the problem. Experiments on real-world datasets show the proposed algorithm can achieve 70% improvement on average in expected influence compared with rivals.

Introduction

The Influence Maximization (*IM*) problem is a fundamental social network problem that has wide applications including viral marketing. It aims to find a small set of seed users of size k that can eventually influence maximum users (Kempe, Kleinberg, and Tardos 2003).

Existing studies on *IM* problems mostly consider the cardinality of the influenced users with less attention paid to the negative impact of some users. This often poses problems in maximizing the influence. For example, in the marketing, not all the users receiving promotion can be adapted, some of which may even have negative impact on the promoted products, e.g., freeloaders or users with bad credits. It is hereby necessary to consider negative users in the IM problems. Although some researchers (Chen et al. 2011) realize the existence of negative users, they focus on the spread of negative opinion. In their models, users are negative when influenced by negative users. The studies find seeds that can maximize the influence of positive users. However, their models are ineffective to minimize the impact of negative users. This may cause problems since negative users would reduce the impact by positive users. For example, if too many freeloaders are involved in the promotion, the interests of companies may be reduced even if many users get influenced.

*The work was partially supported by CAS Pioneer Hundred Talents Program and the MOE Key Laboratory of Machine Perception at Peking University under grant number K-2017-02. Hongyan Li is the corresponding author.

To address the above problem brought by negative users in *IM*, this paper introduces *Negative-aware Influence Maximization* (*NIM*). Given a social network with labeled positive and negative users, the goal is to maximize the number of positive users while minimize the number of negative ones in the influenced user set. The problem can be reduced from IM thus it is NP-hard. Moreover, the objective function does not have properties as many other IM problems such as submodularity and monotonicity for exploration, which makes it even harder. To this end, we propose a Reverse Influence Set based algorithm for NIM (RIS-NIM). Experiments on large real datasets show promising results that *RIS-NIM* is able to achieve 70% improvement in expected influence compared with rivals.

Problem Formulation

A social network is modeled as a directed graph $G(V, E)$. The node set V represents users in the network G, which consists of two disjoint sets V^+ and V^- indicating positive and negative users, respectively. The edge set E represents follower/followee relationships between users. This paper considers Triggering Model as the underlying diffusion model, which is the generalization of many well-known models including *Independent Cascade* (IC) and *Linear Threshold* (LT) models, to characterize the diffusion process on network G (Kempe, Kleinberg, and Tardos 2003).

Given a social network $G(V, E)$, the *Negative-aware Influence Maximization* (NIM) problem is to find a size-k node set S (k is a user-specified parameter as constraint) that maximizes the expected influence $E[I(S)]$ defined as follows:

$$\mathbb{E}[I(S)] = \sum_{v \in V^+} \Pr(S \rightsquigarrow v) - \sum_{v \in V^-} \Pr(S \rightsquigarrow v). \quad (1)$$

Existing studies (Kempe, Kleinberg, and Tardos 2003) prove the NP-hardness of IM. Meanwhile, IM is a special case of NIM because for each instance of conventional IM, we can construct a corresponding equivalent NIM instance by defining $V^+ = V$ and $V^- = \emptyset$. Thus, NIM is NP-hard. Furthermore, the objective function $\mathbb{E}[I(S)]$ of NIM does not have the properties of monotone non-decreasing nor submodular as many other IM problems, which makes it more difficult to be solved.

RIS-NIM Algorithm

This paper proposes Reverse Influence Set based algorithm for Negative-aware Influence Maximization (RIS-NIM). Firstly, RIS-NIM samples a series of reverse reachable sets by RIS method (Tang, Shi, and Xiao 2015). RIS-NIM estimates the expected influence spread $E[I(S)]$ as

$$\mathbb{E}[I(S)] = |V| \cdot \mathbb{E}[Z(S)], \tag{2}$$

where

$$Z(S) = \begin{cases} 1, R(v) \cap S \neq \emptyset \;\&\; v \in V^+ & (3) \\ -1, R(v) \cap S \neq \emptyset \;\&\; v \in V^- & (4) \end{cases}$$

As $E[I(S)]$ can be seen as the difference between two submodular functions, next, RIS-NIM uses modular-modular procedure (Iyer and Bilmes 2012) to select seeds iteratively. We construct the upper and lower bound functions of two submodular functions as follows:

Lower bound function: (Narasimhan and Bilmes 2012) For a submodular $f : 2^V \to \mathbb{R}^+$ and π as an arbitrary permutation of V. Let $W_i = \{\pi(1), \pi(2), \ldots, \pi(i)\}$ where $\pi(i)$ denotes the i-th element of π, so $W_{|V|} = V$. We define lower bound function $h : 2^V \to \mathbb{R}^+$ as

$$h(\pi(i)) = \begin{cases} f(W_1), & \text{i=1} \quad (5) \\ f(W_i) - f(W_{i-1}), & \text{otherwise.} \quad (6) \end{cases}$$

For $A \subset V$, this paper further defines:

$$h(A) = \sum_{v \in A} h(v) \tag{7}$$

Upper bound function: (Iyer and Bilmes 2012) For a submodular function $f : 2^V \to \mathbb{R}^+$, and an arbitrary $B \subseteq V$, we define upper bound $H : 2^V \to \mathbb{R}^+$ as

$$H(A) = f(B) + \sum_{v \in A \backslash B} f(v|\emptyset) - \sum_{v \in B \backslash A} f(v|B \backslash \{v\}), \tag{8}$$

where $A \subset V$ and $f(A|B) = f(A \cup B) - f(B)$.

By maximizing the difference between h and H, we improve the expectation of the seeds iteratively.

Experiments

This study employs 4 public available datasets from Stanford Network Analysis Project. Table 1 presents the properties of the datasets. In order to simulate negative users in applications, we randomly selected 50% types of users as negative users. RIS-NIM supports triggering model in general. Without loss of generality, we report the results on IC model. However, we observe superiority of the proposed algorithm in other models such as LT. We estimate the expected influence of a seed set by taking its average influence of $10,000$ simulations. All the experiments are conducted on an Intel i5 3.2GHz CPU machine with 32GB RAM.

We compare RIS-NIM with IMM (Tang, Shi, and Xiao 2015) and degree discount greedy algorithm, which are well-known solutions for IM problems. Figure 1 shows that RIS-NIM can achieve much higher expected influence than the competitors on all datasets. That is because IMM and greedy algorithms try to maximize the number of influenced users no matter they are negative or positive. While RIS-NIM finds seeds that can influence maximal positive users while minimal negative ones.

Sources	Nodes	Edges
Hamsterster	1,858	12,534
BlogCatalog	10,312	333,983
Flickr	80,513	5,899,882
YouTube	1,138,499	2,990,443

Table 1: Dataset properties.

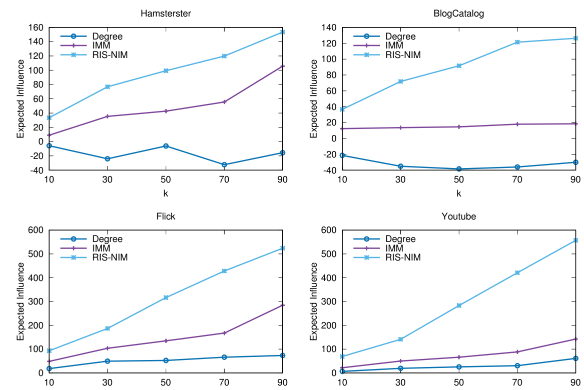

Figure 1: Expected influence varying k under the IC model.

Conclusion

In this study, we propose Negative-aware Influence Maximization (NIM) that takes users' negativity into consideration, and it aims to find users who can influence maximal positive users and minimal negative ones. RIS-NIM algorithm is presented to solve NIM. Preliminary results show that our algorithm can achieve higher expected influence than well-known IM algorithms.

In the future work, we are to improve the expected influence estimation of RIS-NIM and provide its theoretical bounds.

References

Chen, W.; Collins, A.; Cummings, R.; Ke, T.; Liu, Z.; Rincon, D.; Sun, X.; Wang, Y.; Wei, W.; and Yuan, Y. 2011. Influence maximization in social networks when negative opinions may emerge and propagate. In *SDM*, 379–390.

Iyer, R., and Bilmes, J. 2012. Algorithms for approximate minimization of the difference between submodular functions, with applications. In *UAI*.

Kempe, D.; Kleinberg, J.; and Tardos, E. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 137–146.

Narasimhan, M., and Bilmes, J. A. 2012. A submodular-supermodular procedure with applications to discriminative structure learning. *arXiv preprint arXiv:1207.1404*.

Tang, Y.; Shi, Y.; and Xiao, X. 2015. Influence maximization in near-linear time: A martingale approach. In *SIGMOD*, 1539–1554.