Physica A 444 (2016) 20–34

Contents lists available at ScienceDirect

Physica A

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

Influence maximization in social networks under an independent cascade-based model

Qiyao Wang [a], Yuehui Jin [a], Zhen Lin [b,c,*], Shiduan Cheng [a], Tan Yang [d]

[a] State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

[b] Department of Technology Development, Shanghai Clearing House, Shanghai 200002, China

[c] College of Computer Science & Technology, Zhejiang University, Hangzhou 310027, China

[d] School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

HIGHLIGHTS

- A novel independent cascade model based on opinion change, called IC-OC, was proposed for information spreading.
- Two probabilities were introduced to predict opinion change when users are exposed to bilateral opinions.
- The IMIC-OC model was proposed for influence maximization.
- Experiments were conducted on three real networks to verify that the IMIC-OC model has larger influence than two baseline methods.

ARTICLE INFO

Article history:
Received 19 April 2015
Received in revised form 23 July 2015
Available online 13 October 2015

Keywords:
Influence maximization
Positive influence
Independent cascade-based model
Opinion change
Social networks

ABSTRACT

The rapid growth of online social networks is important for viral marketing. Influence maximization refers to the process of finding influential users who make the most of information or product adoption. An independent cascade-based model for influence maximization, called IMIC-OC, was proposed to calculate positive influence. We assumed that influential users spread positive opinions. At the beginning, users held positive or negative opinions as their initial opinions. When more users became involved in the discussions, users balanced their own opinions and those of their neighbors. The number of users who did not change positive opinions was used to determine positive influence. Corresponding influential users who had maximum positive influence were then obtained. Experiments were conducted on three real networks, namely, Facebook, HEP-PH and Epinions, to calculate maximum positive influence based on the IMIC-OC model and two other baseline methods. The proposed model resulted in larger positive influence, thus indicating better performance compared with the baseline methods.

1. Introduction

Online social networks (OSNs) are beneficial for researchers and companies that study user behaviors and make profits, respectively. For example, advertisers found a subset of users to maximize the adoption of products [1], which is a way of using influence maximization. The goal of influence maximization is to maximize the number of users involved in discussions

* Correspondence to: No 2 East Beijing Road, Shanghai Clearing House, Shanghai 200002, China. Tel.: +86 21 23198875.
E-mail address: linzhen@shclearing.com (Z. Lin).

http://dx.doi.org/10.1016/j.physa.2015.10.020
0378-4371/

or adopt products in the network. The independent cascade-based model for influence maximization is proposed in this study to handle this problem.

In recent years, many researchers have focused on influence maximization. Kempe et al. [2] first used greedy algorithm to study influence maximization based on two models, namely, Independent Cascade (IC) and Linear Threshold (LT). Chen et al. [3] calculated positive influence amidst the existence of negative opinions within a network. However, the authors assumed that if users adopted one opinion, they would never change their opinions. As time passes, users are probably influenced by others to adopt opposite opinions. Thus, when users are exposed to different attitudes, probabilities are introduced to predict whether such users would change their minds until no neighbor changes opinions.

To better understand the proposed model for influence maximization, a novel IC model based on opinion change, called IC-OC, is proposed to show the process through which users build opinions. In the beginning of information spreading, the IC model is used to predict the initial opinions of the user. As several users become involved in the discussion, a transforming probability is introduced to predict whether users would change their initial opinions. The proposed model is verified by the dataset culled from Facebook.[1] Then, the IMIC-OC model is used to calculate maximum positive influence in three real networks.

We model social network as a directed influence graph. Nodes denote individuals and edges denote "who influences whom". In some social networks, such as Twitter[2] and Sina Weibo,[3] user relations are unidirectional, such that the structure of network is directly used as influence graphs. Other social networks provide bidirectional relations, such as Facebook and the co-author's network. In these networks, an undirected edge is divided into two directed edges to obtain the influence graph.

In the current study, the IC-OC model is proposed to explain how users build opinions during the process of information spreading. The IMIC-OC model is proposed to find influential users who have maximum positive influence. The following items summarize our contributions to the literature.

- Transforming probability is introduced to predict whether users change their initial opinions when bilateral opinions exist in the network.
- The IC-OC model for information spreading is proposed to gain a better explanation for opinion formation.
- The IMIC-OC model is proposed to tackle the influence maximization problem; the properties of the proposed model are also analyzed.

Experiments were conducted on a real network. First, the IC-OC model was verified for information spreading. Experimental results indicated that the proposed model had better performance than the baseline methods. Maximum positive influence was calculated by applying the IMIC-OC model on three real datasets. The findings showed that the proposed model had better performance compared with the baseline methods. Further explanations were given.

The reminder of this study is organized into sections. Prior works on influence maximization and opinion formation are discussed in Section 2. Section 3 presents the problem statement. Section 4 introduces the IC-OC model for information spreading and the IMIC-OC model for influence maximization. Section 5 shows the experimental results. Finally, Section 6 presents the conclusion.

2. Related works

2.1. Influence maximization

Many researchers have attempted to tackle the influence maximization problem. Kempe et al. [2] proposed two basic information diffusion models, namely, IC and LT. An algorithm was used to obtain an approximate solution for this NP-hard problem. Chen et al. [4] categorized the solutions to the problem of influence maximization into two classes: reducing running time of algorithms and using new methods to calculate influence.

Many studies have focused on improving the efficiency of algorithms on influence maximization [5–7]. Zhang et al. [8], for example, reduced running time and memory consumption by mapping several networks into one network. Goyal et al. [9] proposed CELF++, a novel CELF algorithm, to reduce running time. Heidari et al. [10] used Monte-Carlo simulation to improve time complexity, whereas Li et al. [11] considered wide influence spread to tackle location-aware influence maximization.

Other studies proposed new algorithms to calculate influence [12,13]. Zhou et al. [14] employed the voter model to handle the situation wherein different users activate the same user. Wu et al. [15] used core users in the forefront of a network to calculate influence, and Li et al. [16] proposed a novel conformity-aware cascade model, which used conformity theory to calculate influence. Meanwhile, Zhu et al. [17] combined continuous-time Markov chain and IC model for influence maximization. Guo et al. [18] proposed personal influence maximization, which aimed to find users who can activate targeted users. Lee et al. [19] maximized influence on specific users by query processing. In sum, the problem of influence maximization is based on information spreading. Researchers have proposed new models to exploit the structure of network

[1] http://www.facebook.com.
[2] http://www.twitter.com.
[3] http://www.weibo.com.