Procedia Computer Science

Volume 29, 2014, Pages 422–431

ICCS 2014. 14th International Conference on Computational Science

Maximizing the Cumulative Influence through a Social Network when Repeat Activation Exists

Chuan Zhou[1], Peng Zhang[1], Wenyu Zang[2], and Li Guo[1]

[1] Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
{zhouchuan,zhangpeng,guoli}@iie.ac.cn
[2] Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
zangwenyu@software.ict.ac.cn

Abstract

We study the problem of employing social networks for propagate influence when repeated activations are involved. While *influence maximization* has been extensively studied as the fundamental solution, it neglects the reality that a user may purchase a product/service repeatedly, incurring cumulative sales of the product/service. In this paper, we explore a new problem of *cumulative influence maximization* that brings the *influence maximization* a step closer to real-world viral marketing applications. In our problem setting, repeated activation exists and we aim to find a set of initial users, through which the maximal cumulative influence can be stimulated in a given time period. To describe the repeated activation behavior, we adopt the voter model to reflect the variation of activations over time. Under the voter model, we formulate the maximization problem and present an effective algorithm. We test and compare our method with heuristic algorithms on real-world data sets. Experimental results demonstrate the utility of the proposed method.

Keywords: Social Networks, Repeated Activations, Voter Model, Cumulative Influence Maximization

1 Introduction

The rapidly increasing popularity of online social networks has created immense opportunities for social network marketing, where the key issue is how to maximize the total sales profit (i.e., maximize the cumulate sales). Viral marketing, among existing solutions, represents one of the most effective ways to promote products and generate more profit on social networks. Viral marketing was first introduced to the data mining community by Domingos and Richardson [10, 22], where cost-effective methods are proposed to promote a new product or technology by giving free or discounted samples to a select group of influential individuals, in the hope that through the word-of-mouth effects over the social network, a large number of product adoptions will occur. Motivated by viral marketing, influence maximization has emerged as a fundamental research problem concerning the propagation of innovations through social networks [6, 7, 8, 13, 16, 17, 18].

 Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2014

doi:10.1016/j.procs.2014.05.038

However, existing influence maximization formulation, without considering the reality that a product can be purchased many times in a given time period, cannot precisely describe the cumulative sales (i.e., cumulative influence) maximization problem. Actually, in the literature of influence maximization, it is often assumed that once a user is active, the user will keep the active state forever and cannot be activated again. But the reality is that a user may purchase a product or service repeatedly under social influence. For example, daily necessities, such as toothpastes, makeups and cigarettes, are generally purchased continually by users. Therefore, how to maximize the cumulative sales cannot be simply narrowed down to the influence maximization problem in these scenarios, but demands a new solution framework.

We discuss the problem of *maximizing the cumulative influence* under repeated activations in this paper. Consider a social network as a graph $G = (V, E)$, where V represent users and E represent relationships. Let $S_t \subseteq V$ be the set of nodes who get activated at time t and $|S_t|$ be the size of set S_t. The problem can be described as **how to maximize the cumulative influence $\sum_{t=0}^{T} |S_t|$, when repeated activation exists?** Here, T denotes the time interval that we observe, and similar time constrain can also be found in [5, 20].

The above problem, inspired by influence maximization, can be converted to target an initial subset nodes in the network, such that cumulative influence through the effect of social influence can be maximized. Formally, given a positive integer k, we aim to solve the discrete optimization problem in Eq. (1), where $\mathbb{E}^S$ is the expectation operator with the initial set S.

$$S^* = \arg \max_{|S|=k, S \subseteq V} \mathbb{E}^S \Big[\sum_{t=0}^{T} |S_t| \Big] \tag{1}$$

To solve Eq. (1), the first *challenge* is how to model the variation of S_t when repeated activations happen. Obviously, existing popular propagation models such as Independent Cascade (IC) and Linear Threshold (LT) models [16] are incapable of describing the repeated activation phenomenon, as they mainly describe the range of influence without considering repeated activations.

The second *challenge* is how to develop an efficient algorithm for the optimization problem in Eq. (1). We all know that the influence maximization problem is NP-hard [16] under both IC and LT models, where a line of greedy or heuristic algorithms are proposed. In the same token, we need to answer the question whether the cumulative influence maximization is computationally tractable or an NP-hard problem?

To overcome these challenges, we *first* adopt the framework of the voter model to simulate the diffusions with repeated activation involved in social networks. The voter model, which was first proposed in [9], is probably one of the most basic and natural probability models to represent opinion diffusions in which people may switch opinions back and forth from time to time due to the interactions with other people in the network. The voter model assumes that a node can be activated repeatedly by its newly activated parents and the propagation probability depends on its newly active parents. *Second*, we explore cumulative influence maximization problem from the perspective of stochastic processes and present a simplified calculation procedure for the cumulative influence under the voter model. Furthermore, based on the analysis we propose an *ExactSolution* algorithm for the cumulative influence maximization problem (1), which seeks initial potential nodes with acceptable time cost by experimental results.

The *contribution* of the paper are summarized as follows:

1. We present the cumulative influence maximization problem in social networks, which aims to find the most potential nodes in case of repeated purchase behavior exists. (Section 2)
2. We theoretically analyze the cumulative influence maximization problem and present an efficient solution under the voter model. (Section 3,4)