A Cost Optimized Reverse Influence Maximization in Social Networks

Ashis Talukder, Md. Golam Rabiul Alam, Nguyen H. Tran, and Choong Seon Hong
Department of Computer Engineering, Kyung Hee University, South Korea
Email: {ashis, robi, nguyenth, cshong}@khu.ac.kr

***Abstract*—In recent years, Influence Maximization (IM) has gained great research interest in the field of social network research. The IM is a viral marketing based approach to find the influential users on the social networks. It determines a small seed set that can activate a maximum number of nodes in the network under some diffusion models such as Linear Threshold model or Independent Cascade model. However, previous works have not focused on the opportunity cost defined by the minimum number of nodes that must be motivated in order to activate the initial seed nodes. In this work, we have introduced a Reverse Influence Maximization (RIM) problem to estimate the opportunity cost. The RIM, working in opposite manner to IM, calculates the opportunity cost for viral marketing in the social networks. We have proposed the Extended Randomized Linear Threshold RIM (ERLT-RIM) model to solve the RIM problem. The ERLT-RIM is a Linear Threshold (LT)-based model which is an extension to the existing RLT-RIM model. We also have evaluated the performance of the algorithm using three real-world datasets. The result shows that the proposed model determines the optimal opportunity cost with time efficiency as compared to existing models.**

***Index Terms*—reverse influence maximization, opportunity cost, RIM, viral marketing, influence maximization, linear threshold model, social network.**

I. Introduction

In this day and age, social networks have become the ideal platform for disseminating and exchanging information, ideas, new reports, trends etc. due to the rapid growth of the number of social sites and their usage. Information origination and diffusion in the social networks are growing sharply every day and hence, social networks are becoming the most attractive medium for marketing and research field as well [1]. Social network research has been conducted in many directions and viral marketing based Influence Maximization (IM) is a prominent direction among them.

In the last decade, the IM has become attractive social network research field. The IM problem estimates influential users in the social network such that the spread of influence is maximized. In other words, it selects a seed set such that the total number of activated nodes is maximized assuming that the seed nodes are initially activated [2]. In the activation process, all the activated nodes try to activate their out-neighbors, that is, motivated users try to influence their friends or followers to make some specific decision. The process is also called viral marketing where influence spreads in the *word-of-mouth* effect [3]. It mirrors the human behavior in real-life scenarios that people always consult with the family members, friends, colleagues, or other experts before taking any decision ($viz.$ any purchase decision) [4]. The enormous applications of the IM problem include finding community leaders [5], profit maximization or maximizing product adaptation [6], [7], searching experts in some fields [8], rumor spread and detection [9], [10], e-commerce and media industry [11], contamination and outbreak detection [12], online recommendation [13] etc.

Most of the researchers in this field have conducted research to identify the influential seed set subject to maximize the influence [3], [14], [15] or maximize the profit [6], [16] or maximize product adaptation [7]. But estimating the minimum cost of influence maximization has not been addressed deeply. Many authors have described the cost of motivating the seed users of influence maximization problem in a trivial way. They have just offered free sample products or free tickets of a concert to the seed users [3], [6], [7]. The approach is not rational since those influential users are human and they also might be motivated by some other icon people they follow. Some authors have tried to find influential users with a cost budget and they considered the cost of all activated users including seeds. Zhu *et al.* [15] have balanced the influence and the profit but did not consider the cost in their profit maximization research. But none of the studies has addressed the cost of activating those influential users. Unlike the previous studies, we have tried to compute the minimum opportunity cost to activate the influential target users.

In this research, we have formulated a Reverse Influence Maximization (RIM) problem to calculate the opportunity cost of viral marketing in the social network. The opportunity cost [17] is identified by the minimum number of nodes that we need to activate in order to motivate the given set of target nodes in a social network. In IM problem, a small seed set of users that maximizes the spread of influence is determined [2]. On the other hand, RIM finds the minimum number of in-neighbors that are necessary to activate a given set of target nodes. Thus, the RIM works in the reverse fashion as compared to IM problem as illustrated in the Fig. 1. This is the logic of the naming of RIM. In this work, we propose the Extended Randomized Linear Threshold RIM (ERLT-RIM) model which is a variation of the classical Linear Threshold (LT) [2] model and an extension to the existing RLT-RIM model [18]. The ERLT-RIM considers already activated nodes