Engineering Applications of Artificial Intelligence 41 (2015) 195–206

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Trust based latency aware influence maximization in social networks

Rezvan Mohamadi-Baghmolaei, Niloofar Mozafari, Ali Hamzeh *

Department of Computer Science and Engineering and Information Technology, School of Electrical Computer Engineering, Shiraz University, Shiraz Avenue, Iran

ARTICLE INFO

Article history:
Received 22 June 2014
Received in revised form
21 October 2014
Accepted 9 February 2015
Available online 17 March 2015

Keywords:
Influence maximization
Information diffusion
Independent cascade model
Trust path
Distrust path
TLIM

ABSTRACT

Influence maximization is the problem of finding a small set of nodes that maximizes the aggregated influence in social networks. The problem of influence maximization in social networks has been explored in many previous researches. They have mainly relied on similar temporal chances for every node to influence another; whereas in reality, time plays a major role in pairwise propagation rates in social networks. However, there is little research done on influence maximization considering temporal dynamics of the networks and existing approaches merely offers a mediocre performance due to ignoring trust aspects of the diffusion process. In this paper, we propose a Trust based Latency aware Influence Maximization model, abbreviated as TLIM, which selects the most influential nodes in social networks with considering time and trust simultaneously. To the best of our knowledge, we are the first to study trust in classic IC model and also the first to consider both important time and trust factors jointly in influence maximization problem. The main contributions of this paper are listed as follows: first, we extend the classic IC model to include time and trust simultaneously, which is more applicable in existing social networks. Second, we find the most influential nodes in social networks with considering time and trust together; and the last but not the least, it is applicable to well-known real social networks such as Epinions, Slashdot and Wikipedia. To explore the advantages of our approach, quite a lot of experiments with different specifications are conducted. The obtained results are very promising.

1. Introduction

Nowadays, the rapid growth and popularity of online social networking sites have brought a great deal of attention to social networks (Qiao et al., 2012; Chen et al., 2011; Barbieri et al., 2012; Lu and Lakshmanan, 2012; Belák et al., 2012; Rodriguez and Schölkopf, 2012; Ahmed and Ezeife, 2013). Beside a means of communication, online social networks provide immense sources of information, experience and innovations that enable everyone from everywhere to create, exploit, or spread content through the network via internet links. More importantly, social interaction plays a central role in shaping political or social movements and debates, such as critical role of Facebook or Twitter in the 2010 Arab Spring (Howard and Duffy, 2011). Social network analysis can help extracting worthwhile knowledge, controlling methods of exchanging data, maximizing acquisition of information in the network and also designing improved social networks with appropriate facilities of dissemination.

A lot of studies have been done in the context of information diffusion in social networks. Precisely, information diffusion is a research domain that concerns with the processes of dissemination of information and opinion sharing among members of a social network. According to a recent survey (Guille et al., 2013), studies conducted in this field include three general branches as following: "Detecting Interesting Topics", "Modeling Diffusion Processes" and "Identifying Influential Spreaders". The latter branch is what we study in this article, which is known as "influence maximization".

When news and innovations arise in a social network, they usually begin to spread through the network from person to person, in a virus manner, to achieve as large individuals as possible. The extent of diffusion in the network mainly relates to the mutual relationships of its members. Consider the phone network of a group of individuals. If we send a message including a rumor to someone in this network, he will inform those in his contacts about the new message. They will either accept or ignore the rumor. In fact, if they are affected by the sender, they will believe the new announcement and begin to spread it through the rest of the network using their influence on their friends. Depending on the initial receiver of the message, final amount of receivers in the network would be different. This process is called "word-of-mouth" effect in social networks and is so much operational in commercial intentions. Recently, online social communities have become the target of many companies as a way of advertising new products (Schivinski and Dabrowski, 2013). These companies aim to find the most influential individuals who are most suitable for promoting their brands and absorbing the most customers. They give free

* Corresponding author. Tel.: +98 7136133175.
E-mail address: ali@cse.shirazu.ac.ir (A. Hamzeh).

http://dx.doi.org/10.1016/j.engappai.2015.02.007
0952-1976/

or discounted samples of their product to these particular persons and trigger a large cascade of further adoptions in the whole network subsequently. Therefore, proficient adoption of an exact strategy for specifying potential trendsetters is of great interest; in order to gain the most profit utilizing their influence. These are some examples which the influence maximization problem can cope with. Given a social network graph, a seed size k, and a known influence cascade model, influence maximization is the problem of choosing a set of k influential seed nodes in the graph. Starting from these source nodes, we aim to maximize the spread of information – rumor, innovation, disease – under the influence propagation model in the network. As mentioned above, identifying the influential spreaders (seeds) has a different meaning in each kind of diffusion network. For example in the blogosphere, it means selecting a set of blogs and websites that broadcast the information in a broad range of others (Leskovec et al., 2007). In epidemiology, it consists of finding a set of persons that together are most likely to spread a virus to the greatest number of persons in the population. Identifying these individuals can help controlling the disease transmission in the network (Wallinga and Teunis, 2004). Finally, in viral marketing, the problem reduces to finding a set of trendsetters that absorb the most number of customers to adopt a product (Domingos and Richardson, 2001; Richardson and Domingos, 2002).

The problem of influence maximization in social networks has been explored in many previous researches such as Kempe et al. (2003), Leskovec et al. (2007), Kimura and Saito (2006), Kimura et al. (2007), Chen et al. (2009), Chen et al. (2010, 2011), Barbieri et al. (2012), Lu and Lakshmanan (2012) and Ahmed and Ezeife (2013). They have mainly relied on similar temporal chances for every node to influence another; whereas in reality, time plays a major role in pairwise propagation rates in social networks. However, there is little research conducted on the influence maximization considering temporal dynamics of the networks (Rodriguez and Schölkopf, 2012; Liu et al., 2012; Chen et al., 2012) and the existing approaches merely offers a mediocre performance due to ignoring trust aspects of the diffusion process (Such et al., 2011; Barzilai, 2013). In this paper, we study influence maximization problem considering both trust and time simultaneously. We proceed to illustrate our idea using an example about voting. Suppose a person p hesitates to elect a particular candidate or not. As shown in Fig. 1, p interacts with some neighbors; he trusts some of them (n_1, n_2) which specified by positive links and does not trust the others (n_3, n_4) which specified by negative links. Undoubtedly, each neighbor who takes an action (here, decision to vote for a nomination) has some degree of positive or negative influence on p's opinion. This influence does not propagate immediately to p and is accompanied with some delay. In fact, it may affect p in each subsequent times with different intensities. In Fig. 1, if neighbor n_1 votes for a candidate c, he can persuade p to vote for c as well. However, this positive influence on p varies in each time unit. For example, if we consider a day as a time interval, n_1 may positively influence p in the first day with probability $P_{n_{11}}$. The next day, n_1's influence would be different (probably less than before) and he would influence p with probability $P_{n_{12}}$. In this example, after these two days, n_1 cannot affect p any more. In the case of neighbor n_4, he dissuade p by voting for c and his negative influence

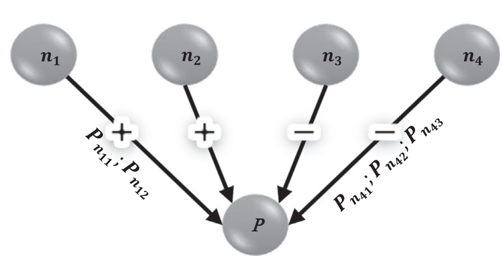

Fig. 1. An example of positive and negative influence spread within different time intervals.

spread to p with probabilities $P_{n_{41}}$, $P_{n_{42}}$ and $P_{n_{43}}$ in the first, second and third days respectively. This example expresses the role of time and trust in the process of influence propagation in a social network, while previous studies fail to model this problem.

To the best of our knowledge, we are the first to study trust in the classic "independent cascade model" and also the first to consider both important time and trust factors jointly in the influence maximization problem. The main contributions of this paper are listed as follows: (1) we extend the classic IC model to include time and trust simultaneously, which is more applicable in existing social networks. (2) We find the most influential nodes with considering two major time and trust factors together in social networks. (3) We evaluate the proposed model on three large scale trust networks of Wikipedia, Slashdot and Epinions. The results of our experiment prove the advantage of considering both time and trust aspects of the influence maximization problem.

The remainder of this paper is organized as follows. Section II, provides an overview of the related work. In Section III, we define the trust-based time constrained influence maximization problem. Section IV describes our new diffusion model. We then present a new framework devised to detect influential nodes under our model. Section V reports experimental results and analysis. Finally, in Section VI we provide the conclusion of our studies and discuss the future research.

2. Related work

In this section, we first introduce the existing basic models of influence in the area of influence maximization. After that, the most related work for this article is reviewed.

2.1. Diffusion models

Independent Cascade Model (ICM) (Goldenberg et al., 2001) and Linear Threshold Model (LTM) (Granovetter, 1978) are two classic graph based approaches which exhibit the process of influence propagation in a social network (Guille et al., 2013). Both approaches model a social network as a directed graph in which nodes and links represent individuals and relations between them respectively. A node is active if it has adopted the content spreading through the graph or it is inactive otherwise. Each link of the graph is associated with an influence probability showing the degree of influence of the tail node on the head of the link. When a node becomes active, it can send the content to its neighbors by its outgoing links and affect them with a specific probability. In the following, we define these models in more details.

ICM is a popular diffusion model which has been widely studied in the context of influence maximization (e.g., Kempe et al., 2003; Kimura and Saito, 2006; Kimura et al., 2007; Chen et al., 2009,2010,2011,2012; Barbieri et al., 2012; Liu et al., 2012). It describes an iterative propagation process, focusing on the sender of the information. Flow of diffusion starts from the initial seed nodes which are all active at first. In each step, the newly activated nodes try to influence their inactive neighbors with pre-defined probabilities on the corresponding links. In the case of success, the receiver becomes active in the next time step; otherwise nothing will happen to that node. The process continues in discrete time units, till no more nodes transfer to active state.

In the case of LTM, in addition to the influence probability on each link, we need an influence threshold for each node in the network. As well as ICM, the propagation proceeds iteratively in discrete time steps while focusing on the receiver node. In fact, this model investigates the simultaneous impressions of a node's neighbors on its status. Thus, a node becomes active by its activated neighbors, if the summation of influence degrees on the incoming links exceeds its threshold amount. Again, the diffusion process ends when no transition happens to active state. In both cases of LTM and ICM, an activated node remains active