Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 83 (2016) 82 – 89

Procedia Computer Science

The 7th International Conference on Ambient Systems, Networks and Technologies (ANT 2016)

Link Prediction Based on Common-Neighbors for Dynamic Social Network

Lin Yao[a], Luning Wang[a], Lv Pan[a], Kai Yao[b]

[a]*School of Software, Dalian University of Technology. Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, China*
[b]*Shenyang University of Technology. Liaoning Province, China*

Abstract

Link prediction is an important issue in social networks. Most of the existing methods aim to predict interactions between individuals for static networks, ignoring the dynamic feature of social networks. This paper proposes a link prediction method which considers the dynamic topology of social networks. Given a snapshot of a social network at time t (or network evolution between $t1$ and $t2$), we seek to accurately predict the edges that will be added during the interval from time t (or $t2$) to a given future time t'. Our approach utilizes three metrics, the time-varied weight, the change degree of common neighbor and the intimacy between common neighbors. Moreover, we redefine the common neighbors by finding them within two hops. Experiments on DBLP show that our method can reach better results.

Peer-review under responsibility of the Conference Program Chairs

Keywords: link prediction; Common Neighbors; dynamic network

1. Introduction

A social network is a social structure made up of a set of social actors, whose nodes represent people or other entities embedded in a social context, and whose edges represent interaction, collaboration, or influence between entities. The associations are usually driven by mutual interests that are intrinsic to a group. Because the relationship between people are always changing, new edges and vertices are added to the graph over time and old ones may be deleted. Consequently, social network is generally complex and highly dynamic. As a key issue of social networks, link prediction has attracted more and more attention because link prediction is important for mining and analyzing the evolution of social networks[1].

The existing link prediction approaches can be classified into similarity-based ones and learning-based ones[1]. Similarity-based approaches are to compute the similarities between a pair of nodes by various graph-based similar-

* Corresponding author. Tel.: 008641162274396 ; fax: 008641162274455
E-mail address: yaolin@dlut.edu.cn

Peer-review under responsibility of the Conference Program Chairs
doi:10.1016/j.procs.2016.04.102

ity metrics and to use the ranking on the similarity scores to predict the link between two vertices[2,3,4]. Learning-based approaches are to treat the link prediction problem as a binary classification task. Therefore, some typical machine learning models such as classifier and probabilistic model can be used for solving this problem[5,6]. Compared with similarity-based approach, the latter approach usually suffer from model capacity and computational problems[7]. Moreover, most existing methods are designed for static network without considering the dynamics and evolutionary patterns of social networks. They would rather predict links from one static snapshot of the graph. However, graph data sets often show dynamic characteristics because of addition and deletion of nodes and edges in the networks.

In this paper, we propose a link prediction method for dynamic social networks. In a social network, people tend to create new relationships with people that are closer to them. The idea of using the size of Common Neighbors(CN) is just an attestation to the network transitivity property. Therefore, we design our prediction method based on the metric of common neighbors. While, the major problem is that most existing methods based on CN just focus on topological structure alone to predict the links in social networks without considering the dynamic network. In our approach, we consider three metrics, the time-varied weight, the change degree of common neighbor, and the intimacy between common neighbors. The time-varied weight reflects the change of topological structure with time. The closer the time is to us, the bigger the weight is. The change degree of common neighbor reflects the stability of every neighbor in the current period, and the weight of a common neighbor gets more with the smaller degree (i.e. more stable). In the CN algorithm, every common neighbor has the same weight for the final prediction. However, every neighbor may not have the same change degree. By adopting this metric, we can kick off some abnormal nodes. The intimacy between common neighbors is used to judge the similarity of two nodes. If their common neighbors are more closely, the similarity between these two nodes are higher. Furthermore, the common neighbors are defined as some nodes within two hops, while CN considers the common neighbors within only one hop.

The contributions of this work are:
(1) We propose a new link prediction method for dynamic networks. To improve the predictive accuracy, we adopt three metrics, the time-varied weight, the change degree of common neighbor, and the intimacy between common neighbors.
(2) We redefine the common neighbors by considering the nodes within two hops to achieve better performance.
(3) We demonstrate experimental results on the effectiveness of the proposed approach with DBLP database. Results show that our approach is competitive to (and sometimes are better than) those of the exact ones.

The rest of this paper is organized as follows. In Section 2, we discuss the related work. Our proposed method is given in Section 3. Experimental results are presented in Section 4. Finally, we conclude our work in Section 5.

2. Related Work

Most of the existing link prediction approaches can be classified into learning-based ones and similarity-based ones[1]. Learning-based approaches adopt the classifier, such as Markov chains[8,9], SVM[5], etc.. Some methods use probabilistic models such as Markov Random Fields, Bayers model, etc. to predict the link association[10,11]. Compared with similarity-based approaches, learning-based approaches have the difficulties in feature selection and unbalancing output classes and is suffered from computational cost and limitation of capacity, therefore it is not suitable for large-scale and dynamic networks[12]. Similarity-based approaches can adopt nodes' information, network topology, etc. to link prediction. In[13], the keyword distance is adopted to define similarity functions between a pair of users. In[14], users' interests are used to measure the similarity. In[15], authors propose a novel user similarity measure for online social networks, which combines both network and profile similarity. They also propose a method to infer a portion of the missing items from profile of the users contacts. Liben-Nowell and Kleinberg proposed one of the earliest topology-based prediction models that works explicitly on a social network[2]. They tested the predictive power of some proximity metrics, including Common neighbours(CN), Preferential Attachment(PA), Katz measure, etc.. The ranking on the similarity scores is used to predict the link between two vertices. In[16], nine common algorithms of link prediction are compared and the results show that the CN algorithm possesses the best performance.

However, most of the existing link prediction methods are aim to predict links from one static snapshot of the network graph, ignoring the underlying additional temporal information in pace with the evolution of the network. Some researchers have attempted to predict links by changing the dynamic network into several static networks. Then, they try to design algorithms based on static prediction methods. In[17], authors summarized the dynamic graph