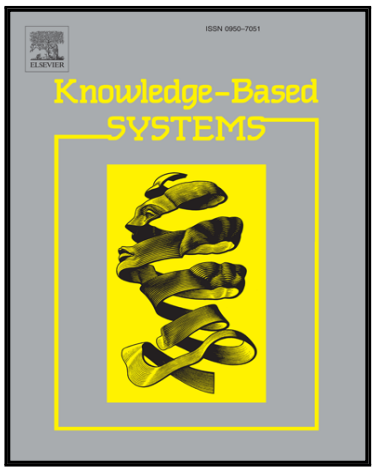

Efficient Incremental Dynamic Link Prediction Algorithms in Social Network

Zhongbao Zhang, Jian Wen, Li Sun, Qiaoyu Deng, Sen Su, Pengyan Yao

PII: S0950-7051(17)30314-3
DOI: 10.1016/j.knosys.2017.06.035
Reference: KNOSYS 3963

To appear in: *Knowledge-Based Systems*

Received date: 9 January 2017
Revised date: 23 June 2017
Accepted date: 26 June 2017

Please cite this article as: Zhongbao Zhang, Jian Wen, Li Sun, Qiaoyu Deng, Sen Su, Pengyan Yao, Efficient Incremental Dynamic Link Prediction Algorithms in Social Network, *Knowledge-Based Systems* (2017), doi: 10.1016/j.knosys.2017.06.035

Efficient Incremental Dynamic Link Prediction Algorithms in Social Network

Zhongbao Zhang[a], Jian Wen[a], Li Sun[a], Qiaoyu Deng[a], Sen Su[a], Pengyan Yao[b]

[a]*State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China*
[b]*Baidu Inc., Beijing, China*

Abstract

To enhance customers' loyalty and experience, link prediction in social networks can help service providers to predict the friendship between users in the future, according to the network structure and personal information. However, most of prior studies consider link prediction in the static scenario while ignoring that the social network generally is updated over time. In this paper, to address this problem, we design two efficient incremental dynamic algorithms that can predict the relationship between users according to the updated social network structure. The first one, instead of using classic prediction index, creates a latent space for each node in the network, and adopts the incremental calculation to predict the future links according to the position of each node in the latent space. The second one is a dynamic improved algorithm based on the resource allocation index, which only recalculates updated part of the social network structure instead of the whole social network. Extensive experiments show that our first algorithm has high prediction accuracy while the second algorithm incurs low running time cost at the expense of less prediction accuracy.

Keywords: Social network; link prediction; dynamic; latent space; resource allocation

1. Introduction

In the era of the web 2.0, more and more people have joined in a variety of social networks for establishing different kinds of relationships. For example, a person may join twitter for connecting to his close friends and LinkedIn for his career development. Due to the anonymity of social network, most people of the same social service actually have not set up the connection in such a virtual network with each other. Therefore, it is very important to predict social links and recommend users with some people who they might be interested in. Link prediction not only plays a decisive role in increasing user stickiness of service providers, but also improves enthusiasm of their users.

Previous studies [1], [2], [3], [4], with known network nodes, network structure and other related information, can predict the possibility of generating a link between two irrelevant users and find out the existence of the relationship between users that has not been reflected in the current network link owing to information loss. However, these prior studies have the following limitation: the graph used to predict the algorithm is static, with no new added nodes and no new generated edges, which is unpractical in real-world social networks.

To address this problem, we propose to study the problem of incremental dynamic link prediction. The naive method first in mind is that it can split the time into several discrete time slots and employ traditional link prediction technique for each time slot. However, it needs to recalculate the prediction on whole graph for each time slot, which may incur high time complexity. In this paper, our goal is to only recalculate the updated partial graph of social network instead of the whole social network when the social network is updated. To achieve this goal, we design two incremental dynamic algorithms, i.e., dynamic link prediction algorithms based on improved latent space (DLP-ILS) and dynamic link prediction algorithms based on improved resource allocation (DLP-IRA). In DLP-ILS, instead of using classic prediction index, it considers the relations between nodes and edges as the interaction among nodes. Specifically, it creates a latent space for every node in the network and predicts the future links incrementally according to the position of each node in the latent space. The large distance between nodes in latent space would decrease the possibility of the link generating.

In DLP-IRA, its core idea is that the more common neighbors the two nodes have, the more correlations there are, and they are more likely to generate the connection by the introduction of their neighbors. Then, it employs resource allocation metric (RA)in the following two ways: i) It is based on the degree of the common neighbors and the large degree has little contribution. ii) It considers the impact of the end node of the link, and the large degree of the node would decrease the possibility of the link generating. The advantage of the algorithm is that it only needs to recalculate the updated part of the social network structure instead of the whole network.

Email addresses: zhongbaozb@bupt.edu.cn (Zhongbao Zhang), wenjian@bupt.edu.cn (Jian Wen), l.sun@bupt.edu.cn (Li Sun), dengqiaoyu@bupt.edu.cn (Qiaoyu Deng), susen@bupt.edu.cn (Sen Su), yaopengyan@baidu.com (Pengyan Yao)