Breaking the Barrier to Transferring Link Information across Networks

Abstract:

Link prediction is one of the most fundamental problems in graph modeling and mining. It has been studied in a wide range of scenarios, from uncovering missing links between different entities in databases, to recommending relations between people in social networks. In this problem, we wish to predict unseen links in a growing target network by exploiting existing structures in source networks. Most of the existing methods often assume that abundant links are available in the target network to build a model for link prediction. However, in many scenarios, the target network may be too sparse to enable robust inference process, which makes link prediction challenging with the paucity of link data. On the other hand, in many cases, other (more densely linked) auxiliary networks can be available that contains similar link structure relevant to that in the target network. The linkage information in the existing networks can be used in conjunction with the node attribute information in both networks in order to make more accurate link recommendations. Thus, this paper proposes the use of learning methods to perform link inference by transferring the link information from the source network to the target network. We also note that the source network may contain the link information irrelevant to the target network. This leads to cross-network bias between the networks, which makes the link model built upon the source network misaligned with the link structure of the target network. Therefore, we resample the source network to rectify such cross-network bias by maximizing the cross-network relevance measured by the node attributes, as well as preserving as rich link information as possible to avoid the loss of source link structure caused by the re-sampling algorithm. The link model based on the re-sampled source network can make more accurate link predictions on the target network with aligned link structures across the networks. We present experimental results illustra- ing the effectiveness of the approach.