

## ABSTRACT

Link prediction in complex networks (e.g., social networks, biological networks, citation networks, etc.) has attracted increasing attention from both physical and computer science communities. The algorithms can be used to extract missing information, identify spurious interactions, evaluate network evolving mechanisms, and so on. Its study is crucial to the analysis of the evolution of networks. Lots of works employing different types of methodologies of link prediction are available. Most of them are based on structural or topological properties as extracting these features are easy in computation. Though, not all of them are efficient to extract. Most social networks exhibit some basic features like Small-world phenomenon, clustering and scale-free. Their corresponding measures are average path length, clustering coefficient and degree distribution respectively. In this thesis, these features are explored for calculating similarity measures of node-pairs in link prediction.

Many real-world networks show tendency of being organized in clusters that are quantified by clustering coefficient. This measure extracts local structural or topological information which are efficient to compute. The notion of mutual relationships, captured by common neighbors, are building blocks of many existing seminal works like Adamic-Adar index, resource allocation index, etc. The notion of common neighbors is further expanded to higher level. Based on clustering coefficients of level-2 common neighbors, a new algorithm CCLP2 is proposed to predict missing links in networks. CCLP2 extracts higher level clustering information of nodes which proved to be more informative and discriminating feature for link prediction as shown by the empirical results.

Exploring level-2 clustering information are useful discriminating feature but confined to neighbors of neighbors information. This might limit the prediction capability and hence, more local information are extracted using path feature. By employing higher order paths as discriminating features missing link are predicted in networks. The proposed method, called SHOPI, is based on resource allocation process in networks where the source node sends some resources as information to a destination node. The amount of information received by the destination derives the similarity score between them. Higher the information received by destination from the source represents higher similarity. SHOPI ensures to reach maximum information by restricting the information leaks through their common neighbor nodes. Empirical results on several networks validate the performance of SHOPI.