Generalized framework for detecting communities of social networks by the scanning method

Tai-Chi Wang*, Frederick Kin Hing Phoa

Institute of Statistical Science, Academia Sinica

Abstract

With the growth in the big data regime and the popularity of social media, recognizing and analyzing social network patterns are important issues. In real world, society offers a wide variety of possible communities, such as schools, families, and firms. For this reason, community/cluster detection draws much attention as it is important to many applications in business and social sciences. The scan statistics have been verified as a useful tool to determine both structure and attribute clusters in networks. However, most of previous methods assumed that the baseline network model follows the Poisson distribution assumption. In this paper, we generalize the previous scan statistic to accommodate to random connection probability model and logit model. Simulation studies show that the generalized methods have better detection results, and empirical studies show the differences among the proposed methods and the previous methods.