

Using Cyclic Structure to Improve Inference on Network Community Structure

Authors : Behnaz Moradijamei, Michael Higgins

Abstract : Identifying community structure is a critical task in analyzing social media data sets often modeled by networks. Statistical models such as the stochastic block model have proven to explain the structure of communities in real-world network data. In this work, we develop a goodness-of-fit test to examine community structure's existence by using a distinguishing property in networks: cyclic structures are more prevalent within communities than across them. To better understand how communities are shaped by the cyclic structure of the network rather than just the number of edges, we introduce a novel method for deciding on the existence of communities. We utilize these structures by using renewal non-backtracking random walk (RNBRW) to the existing goodness-of-fit test. RNBRW is an important variant of random walk in which the walk is prohibited from returning back to a node in exactly two steps and terminates and restarts once it completes a cycle. We investigate the use of RNBRW to improve the performance of existing goodness-of-fit tests for community detection algorithms based on the spectral properties of the adjacency matrix. Our proposed test on community structure is based on the probability distribution of eigenvalues of the normalized retracing probability matrix derived by RNBRW. We attempt to make the best use of asymptotic results on such a distribution when there is no community structure, i.e., asymptotic distribution under the null hypothesis. Moreover, we provide a theoretical foundation for our statistic by obtaining the true mean and a tight lower bound for RNBRW edge weights variance.

Keywords : hypothesis testing, RNBRW, network inference, community structure

Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development

Conference Location : Chicago, United States

Conference Dates : December 12-13, 2020