

Root and community inference on preferential attachment networks

Min Xu

Rutgers University

Abstract:

Preferential attachment (PA) is a popular way of modeling random networks in which the network starts as a single node which we call the root node, and at every new time step, a new node and new edges are added to the network; this dynamic captures the growth/recruitment process that underlies many real-world networks.

Given only a single snapshot of the final network G , we study the problem of constructing confidence sets for the early history, in particular the root node, of the unobserved growth process; the root node can be patient zero in a disease infection network or the source of fake news in a social media network.

In the case where the graph is a PA tree contaminated with random noise edges, we propose an inference algorithm based on Gibbs sampling that scales to networks with millions of nodes and provide theoretical analysis showing that the expected size of the confidence set is small so long as the noise level is not too large. We also propose variations of the model in which multiple growth processes occur simultaneously, reflecting the growth of multiple communities, and we use these models to provide a new approach to community detection.