
MATT HURSHMAN, Dalhousie University

Model Selection for Social Networks using Graphlets

Several network models have been proposed to explain the link structure observed in online social networks. Such models are usually justified by showing that they can replicate certain observed features of real networks, such as a power law degree distribution, a high degree of clustering, and the small world property. This talk will address the problem of choosing the model that best fits a given real world network. We implement a model selection method based on un-supervised learning. An alternating decision tree is trained using synthetic graphs generated according to each of the models under consideration. Graphs are represented as feature vectors incorporating the frequency counts of all connected subgraphs on 3 and 4 vertices as well as features describing the degree distribution and small world property. We show that the subgraph counts alone are sufficient in separating the training data. For the test data, we take four Facebook graphs from various American Universities. Our results show that models incorporating some form of the preferential attachment mechanism tend to perform the best.