CALIFORNIA SCIENCE & ENGINEERING FAIR
2019 PROJECT SUMMARY

Name(s)

Alex Levine

Project Number

S1406

Project Title

Efficiently Identifying High-Centrality Nodes in Modular Networks

Abstract

Objectives

Identifying high-centrality nodes in large networks can be computationally inefficient, given that the most efficient exact algorithm runs in O(MN) time. A number of randomized approximations have been proposed based on sampling. We aim to identify more efficient ways to compute centrality based on the properties of real-world networks. We are also interested in understanding the dynamics of networked systems, such as social network behavior.

Methods

We initially conjecture that the highest centrality nodes in a network lie on the border between two communities.

Using the NetworkX library in Python, we test our conjecture on various network models, namely the Erdos-Renyi Model (ER) and the Barabasi-Albert (BA) Model and some real-world models. For each trial, we generate two communities using the Stochastic Block Model. To verify our results, we use the K-Means clustering algorithm and the Louvain Community Detection Method.

Results

Our conjecture holds 100% of the time on the ER Model. We initially do not yield the same results on the BA Model, however, as we increase the number of nodes and the average degree, the probability of our conjecture tends towards 100%. On the graphs where are conjecture seemingly doesn t hold, we apply the Louvain Method and reveal a very different community structure than we originally defined and find that our conjecture does in fact hold true. We obtain similar results when applying the Louvain method to real-world networks.

We provide theoretical evidence for two structured cases: complete graphs and star graphs. We find that our conjecture does not hold on the star graph when one community is larger than the other, however, we conjecture that there are more natural community structures than initially imposed, which may explain some of our results for the BA model.

Conclusions

Our findings have the potential to significantly speed up the computation of betweenness centrality. Rather than computing the betweenness centrality of every node in the network, one can uncover a community structure using the efficient Louvain Method and only compute the centrality of the inter-community nodes. In addition, our results can be used to better understand the dynamics of many social and biological networks.

Summary Statement

We find a very strong correlation between high-centrality nodes and inter-community nodes and demonstrate the implications of this result.

Help Received

Dr. Behrouz Touri of UCSD provided the research idea and supported me throughout the research process. He most notably helped with the theoretical evidence.