MinT-Net: Novel and Scalable Network-enabled Comparative Tools for Stress Studies of Microbiomes in Transition

Alexis Gomez[1], Arun Sathanur[2], Mahantesh Halappanavar[2]

[1] Texas Tech University, Lubbock, TX
[2] Pacific Northwest National Laboratory, Richland, WA

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by **Battelle** *Since 1965*

Background and Motivations

Community detection is the process of analyzing graphs to distinguish different groups of nodes from one another. These communities exist within the species, gene, and protein networks of a microbiome. Many different algorithms have been developed to detect these communities. The project as a whole is intended to track communities in dynamic networks using known community detection algorithms. An initial effort created implementations of different algorithms for community detection to test for community quality with respect to computational time, focusing on the Girvan-Newman algorithm and the Louvain algorithm.

Initial Algorithm Analysis

Girvan-Newman Algorithm

$$c_B(e) = \sum_{s,t \in V} \frac{\sigma(s,t|e)}{\sigma(s,t)}$$

e represents each edge
s, t are nodes in the set of nodes V
σ(s, t) is the number of shortest (s, t) paths

Calculate the betweenness of each edge → Remove the edges with the highest betweenness → Repeat until there are no edges remaining

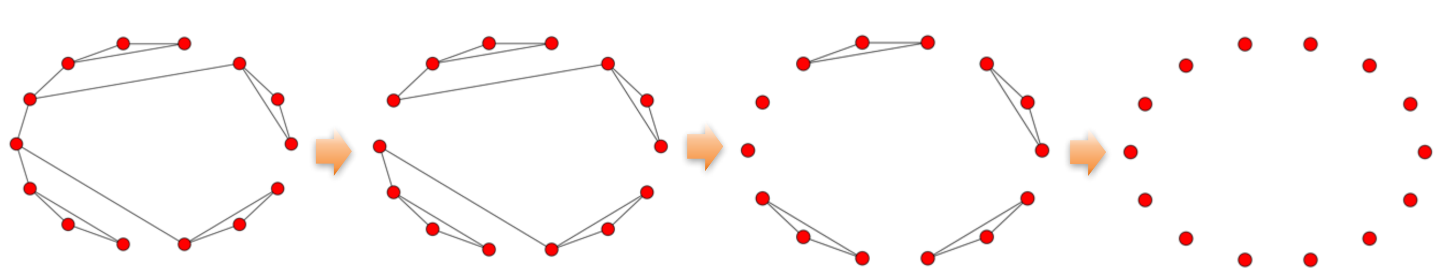


Figure 1. Transition from a full network to an edgeless network using the Girvan-Newman algorithm.

Louvain Algorithm

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j)$$

A_{ij} is the edge weight between i and j
k_i is the sum of the weights of the edges attached to i
c_i is the community of node i
m is the sum of edge weights in the graph

Calculate the modularity of the network → Calculate the change in modularity if a node changes to a different community → Move the node to the community that results in the greatest increase in modularity → Repeat until the modularity of the network cannot be increased any more

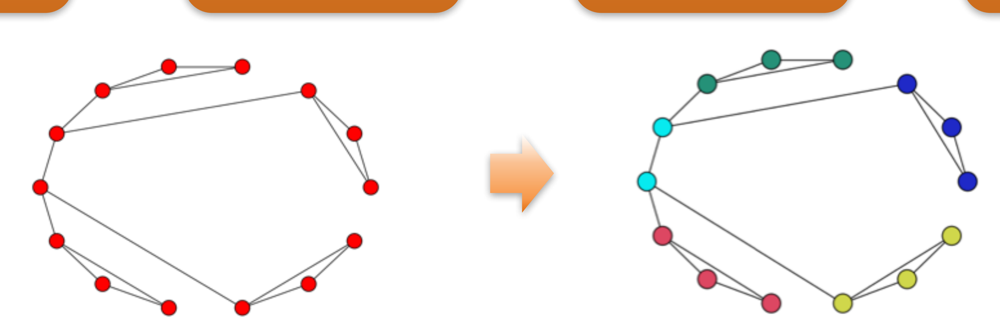


Figure 2. Transition from an un-partitioned network to a partitioned network using the Louvain algorithm.

Initial Algorithm Analysis

Methods

Trials were run on assortative planted partition models to test the accuracy of the algorithms with respect to their computational time.

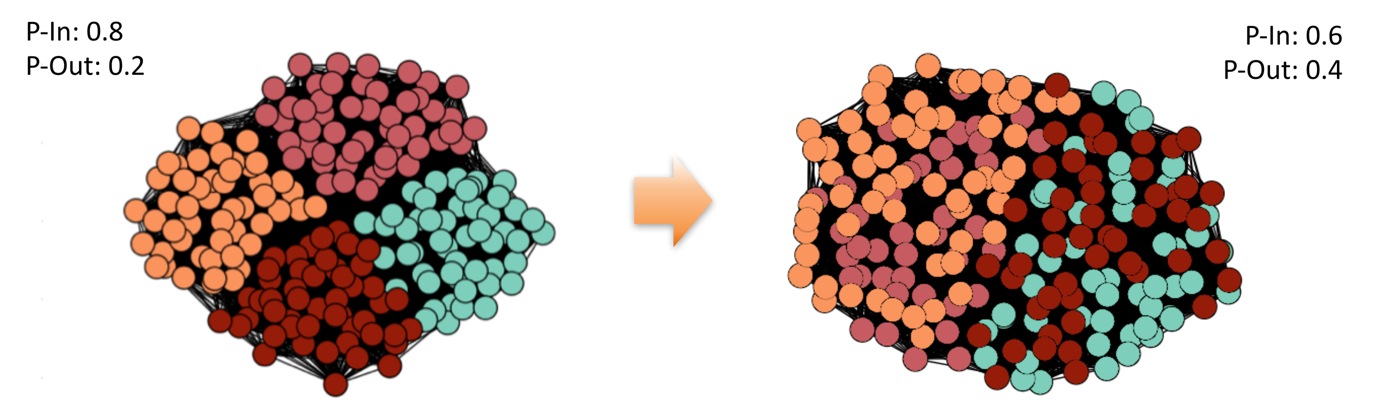
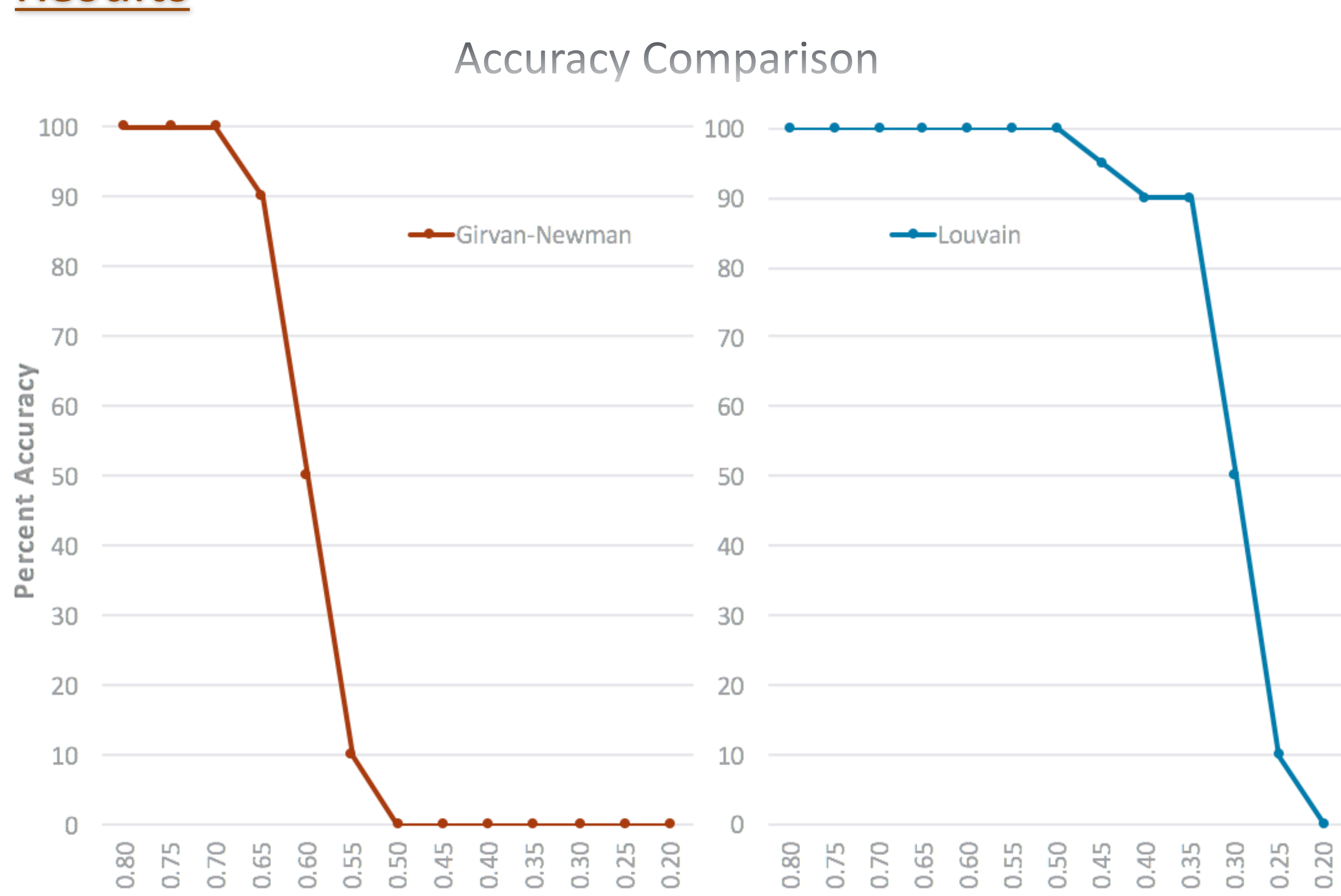


Figure 3. Transition from a network with more assortativity to less assortativity

Results



After the trials, the Louvain algorithm was identified to not only be more computationally-time efficient, but more accurate when detecting communities in models with less assortativity. The accuracy and efficiency of the Louvain algorithm is promising for its future use in dynamic community detection in networks that model microbiomes in transition.

Application on Dynamic Graphs

Jaccard Index

$$J(C_m, C_n) = \frac{C_m \cap C_n}{C_m \cup C_n} = \frac{C_m \cap C_n}{C_m + C_n - (C_m \cap C_n)}$$

∩= intersection
∪= union

Pointwise Mutual Information

$$PMI(C_m, C_n) = \log \frac{P(C_m|C_n)}{P(C_m)} = \log \frac{P(C_n|C_m)}{P(C_n)}$$

P = probability

C_m is a community from the first graph
C_n is a community from the second graph

Application on Dynamic Graphs

Methods

Detection efforts on dynamic networks with community structure were performed on models using the framework of the Chinese Restaurant stochastic process.

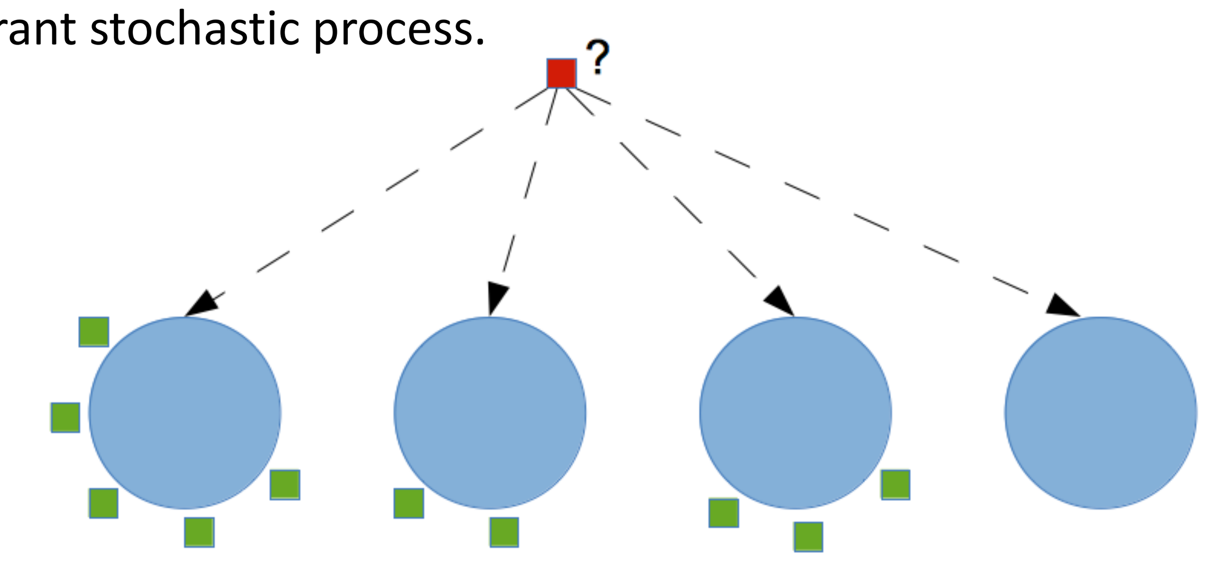


Figure 4. How communities are formed using the Chinese Restaurant stochastic process (https://medium.com/kifi-engineering/from-word2vec-to-doc2vec-an-approach-driven-by-chinese-restaurant-process-93d3602eaa31)

Results

Preliminary results using the Jaccard Index and Pointwise Mutual Information to track communities seem promising for identifying changes in the communities of dynamic networks, including the splitting and merging of communities, shown below. Leveraging these preliminary results, we plan on developing a set of formal rules to track communities in dynamic graphs.

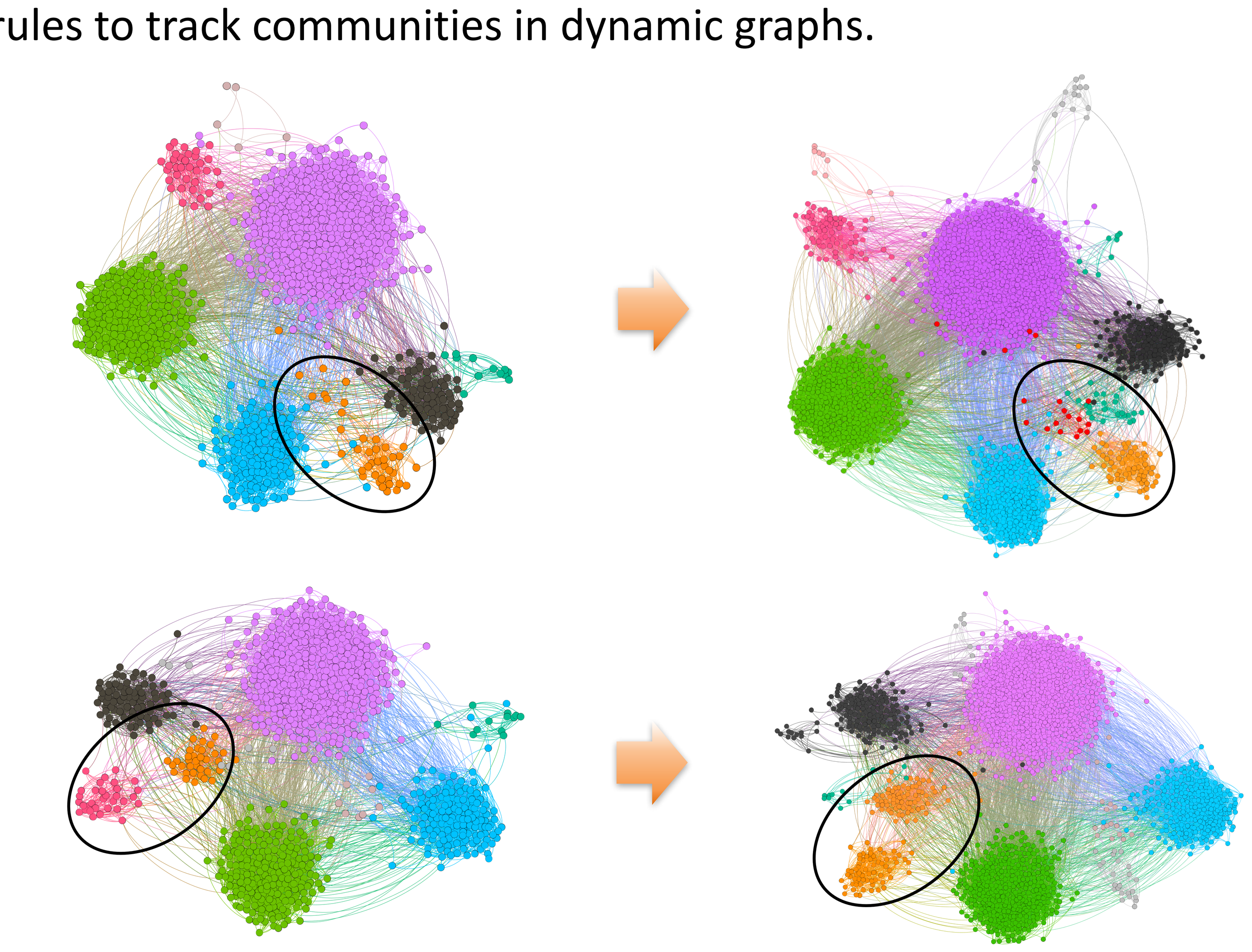


Figure 5. A visualization of a community splitting in two (above) and two communities merging (below).

Acknowledgements

This material is based upon work supported by the National Science Foundation through the Robert Noyce Teacher Scholarship Program under Grant #1418852. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The research was also made possible by the California State University STEM Teacher and Researcher Program, in partnership with Chevron (www.chevron.com), the National Marine Sanctuary Foundation (www.marinesanctuary.org), and Pacific Northwest National Laboratory.

U.S. DEPARTMENT OF ENERGY · Chevron · NATIONAL MARINE SANCTUARY FOUNDATION · CAL POLY SAN LUIS OBISPO · CSU · STAR

www.pnnl.gov