NETWORK MOTIFS IN SOCIAL NETWORKS

Claire Luciano | Summer REU Fellowship | University of Connecticut
OVERVIEW

- Graph Theory
- Network Motifs
- FANMOD
- Edge Sampling Algorithm
- Randomized Enumeration Algorithm
- FANMOD Performance and Network Size
- Network Motifs in Social Networks
WHAT IS GRAPH THEORY?

• Graphs are composed of vertices, or nodes, and edges that connect the pairs of vertices.

• In directed graphs, edges are directed from one vertex to another.

• Graphs are a useful mathematical model for representing various types of systems, including biological systems.

• Biological systems are modeled as graphs based on the biological relationships within each individual system.
NETWORKS

- Transcriptional Regulation in E. coli
- Protein-protein interaction networks
- Social networks
E. coli transcription regulation represented as a directed graph

- An operon is a group of contiguous genes that are transcribed into a single mRNA molecule.

- Node A $\rightarrow$ Node B
  - Node A: Operon that encodes a transcription factor
  - Node B: Operon directly regulated by TF
  - Edges represent direct transcriptional interactions.
E. COLI TRANSCRIPTION NETWORK MOTIFS

• Generating temporal expression programs, and governing the responses to fluctuating external signals.
• Feedforward loop
• Single-input module
• Dense overlapping regulons (DORs)
NETWORK MOTIF DETECTION SOFTWARE TOOLS

• MFINDER
  - Supports sampling using the edge sampling algorithm.

• MAVISTO
  - Visualizes occurrences of a motif in a network by a force-directed graph layout algorithm.

• PAJEK
  - Supports the search for all occurrences of a certain pattern in a network.
  - Does not sufficiently support the enumeration of subgraphs and statistical comparison with random graphs.

• FANMOD
  - Utilizes Random Enumeration sampling algorithm
EDGE SAMPLING

- MFINDER uses a network motif detection algorithm called edge sampling.
- In the edge sampling algorithm, we start by selecting a random edge in the input graph. Next, the edge is randomly extended until we obtain a connected subgraph with the desired number of vertices.
- Edge sampling results in a sampling bias. It samples certain subgraphs more often than others.
- The amount of bias cannot be estimated with only the number of edges neighboring the oversampled subgraph.
RANDOMIZED ENUMERATION

• FANMOD uses a new motif detection algorithm called randomized enumeration.

• Randomized enumeration first starts with an algorithm that enumerates all of the subgraphs with the desired properties.

• The algorithm is then modified to randomly skip over some of these subgraphs when the program is executed.

• The result is an unbiased subgraph sampling algorithm.
NETWORK MOTIFS IN WIKIPEDIA NETWORK

Full Runtime: ~20 hours
E-OPINIONS MOTIFS

507.59  
99.777  
90.379  
68.732  
34.143  
28.586  
2.4178
ACKNOWLEDGEMENTS

Thank You!

- My Advisor, Dr. Vincent Huang
- The University of Connecticut
- The National Science Foundation REU Programs