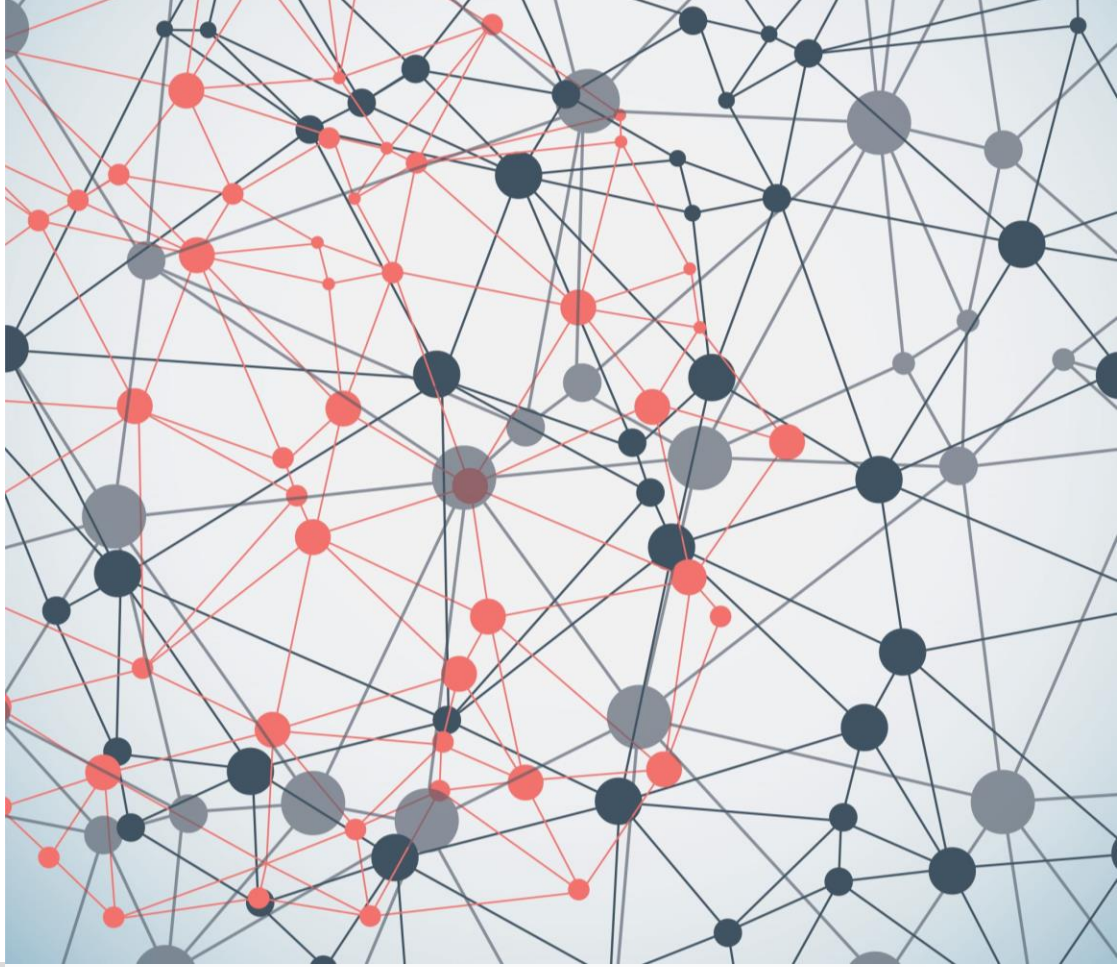
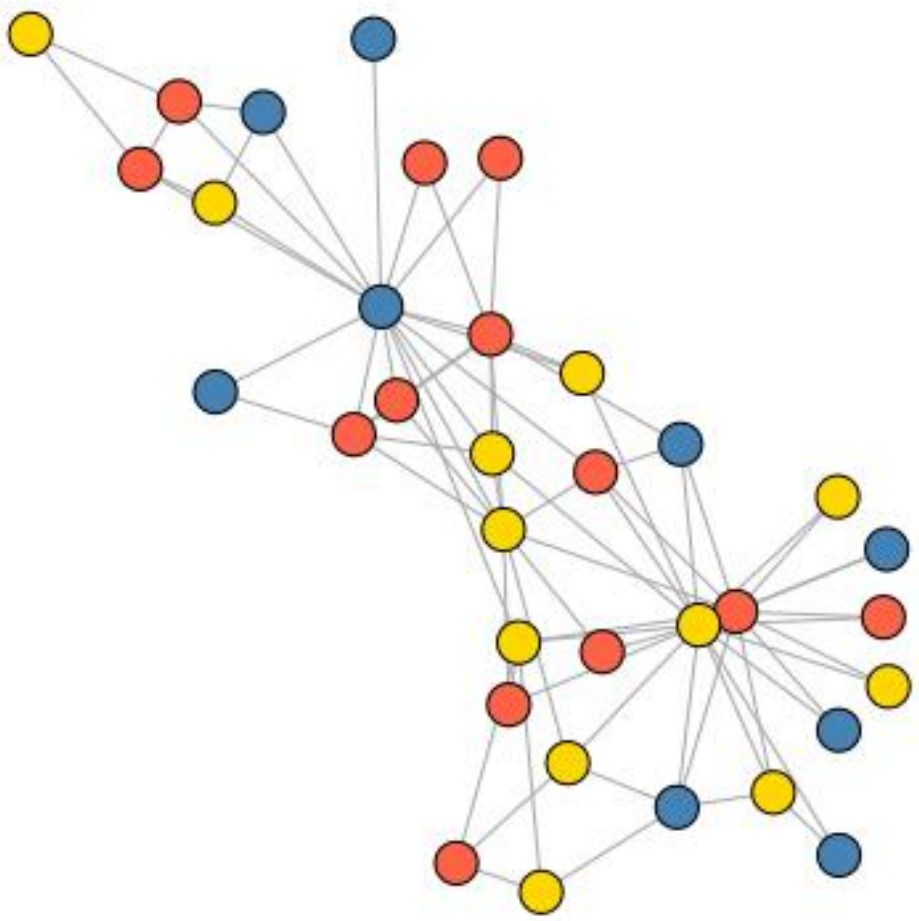
Analysis of Methods for Automatic Graph Clustering



Randa Elshafei

Avery Peiffer



Introduction

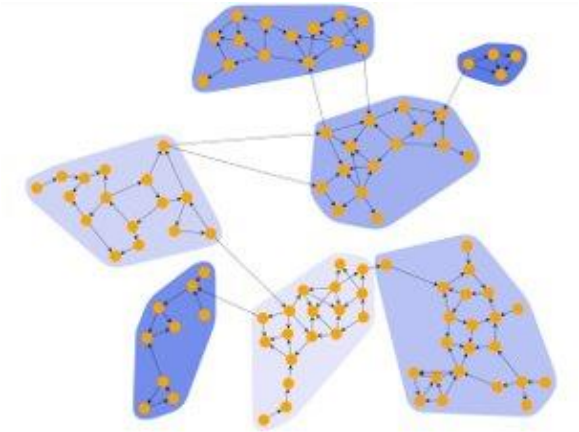
- Many systems can be modeled as networks consisting of nodes and edges
- Creating such a graph by hand requires a large time investment
- Emphasis on automating the graph creation process

CLARINET

- Automatically extracts information from literature and creates graph to model intracellular signaling
- Uses output of automatic machine reading tool to formulate possible candidates for graph

CLARINET (cont.)

- Two scores assigned to nodes: individual assessment and pair assessment
  - Calculated using a frequency class metric (how often the event appears throughout the literature)
- Clusters nodes based on these metrics, then adds to model

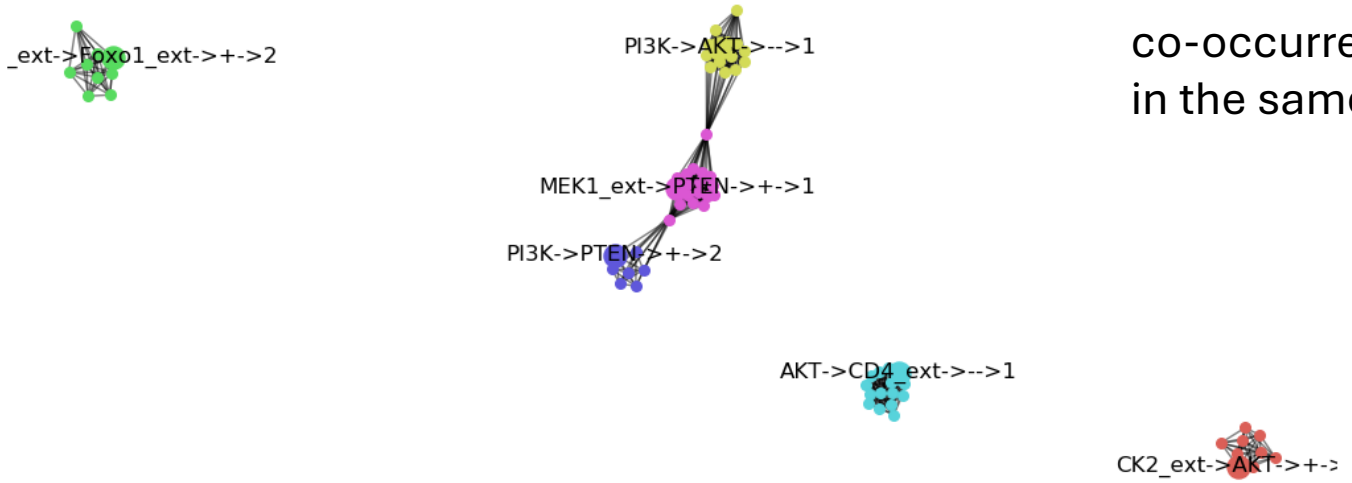


Project Progress

- Downloaded and ran existing CLARINET materials
  - Worked through errors with executing Jupyter Notebook/Python script
  - Spoke to Yasmine Ahmed and were able to get code to a workable state

Project Progress (cont.)

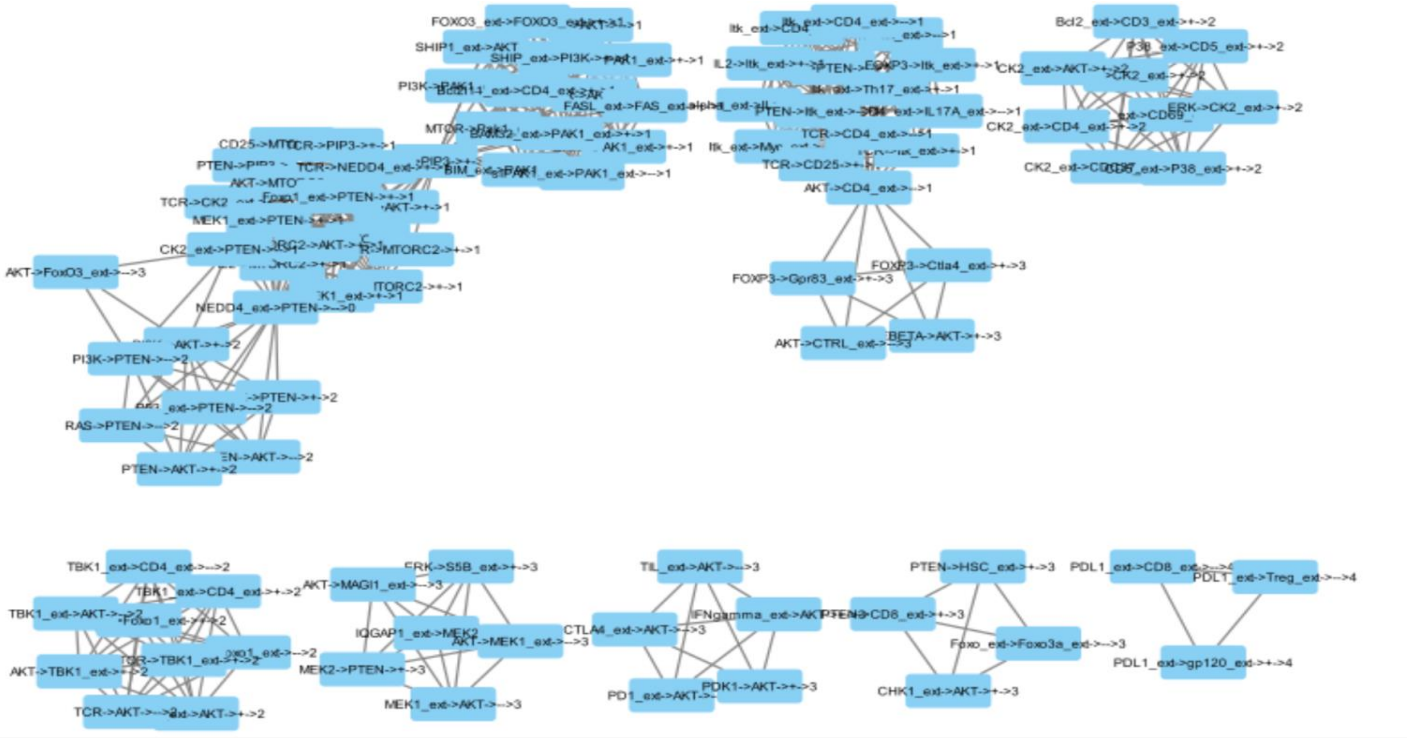
- Image showing clustering output from running CLARINET code for the case used in paper



Clusters represent the co-occurrence of events in the same paper

Project Progress (cont.)

- Cystoscope tool was used to help visualize the results of the network before and after applying CLARINET
- The clustering output from running the CLARINET code before removal of less frequent events



Project Progress (cont.)

- The clustering output from running the CLARINET code after removal of less frequent events

Project Progress (cont.)

- We have successfully used CLARINET code to generate results for two new cases that were provided by Yasmine

Future Work

- Analyzing the resulting output for the original case in CLARNET paper as well as the two new provided cases
- Exploring the Git tool, which uses intensity topology to cluster a graph
- Work with the yWorks tool, which guides us on clustering graphs using different methods
- These tools will help us to gain a greater understanding of graph clustering on a greater scale, instead of within the scope of the CLARINET tool

Image Citations

- https://stackoverflow.com/questions/58598186/counting-edges-in-a-graph-by-the-attribute-of-nodes-they-connect
- https://www.yworks.com/pages/clustering-graphs-and-networks