Project 3: Data-driven reverse engineering of biological regulatory networks

Supervisors: Dr. Roland Somogyi, Dr. Parvin Mousavi

Contact: pmousavi@cs.queensu.ca

Gene regulatory networks represent the interaction of genes in cells and can be displayed as network graphs. When referring to gene regulatory networks in the context of the proposed project, we assume genes to be the nodes of the network graph and their interactions to be the connections of the network graph. Given the high dimensionality of gene expression data, any built gene interaction network from such data will be complex. A main goal of data-driven modeling of gene networks is to characterize the key network features and regulatory nodes.

In this project we will aim to analyze a given gene regulatory network, built using data-driven approaches, to answer the following questions:

- Networks of pair-wise gene connections can be described by a connectivity matrix and displayed as network graphs. How can we identify which elements of the network are the most important regulators?
- Key regulatory nodes likely stand out by having high connectivity. How should we defiine unusually high connectivity?
- Connectivity of nodes varies with the selected connectivity threshold. How should we compare nodes that have relatively high connectivity at high thresholds, but not at low thersholds, and vice versa?
- Are there e.g. nodes with strong connections to hubs, that themselves are not hubs, but could be controllers of hubs?
- Which nodes have the most influence on the overall behavior of the network using e.g. flux analysis?

The project will examine a combination of connectivity strength, flux, direct vs. indirect connections and null datasets to answer some of the above questions.

Requirements: Good mathematical and statistical background; familiarity with molecular biology.