Mechanical & Industrial Engineering
UNIVERSITY OF TORONTO

MIE498H1: Research Thesis 2023-2024

Supervisor	Samin Aref
Supervisor email	aref@mie.utoronto.ca
Number of Positions	1
Open to	Undergraduate Industrial Engineering Students
Term Offered	Winter (S)
Research Area	Operations Research
Research Topic	Analyzing and improving network clustering and community detection algorithms using discrete optimization

Project Description

This project focuses on the assessment and enhancement of the performance of an existing algorithm designed for a computationally demanding unsupervised learning task on networks (community detection). Community detection, which involves identifying groups within networked systems, is a crucial process. The Bayan algorithm, developed recently at the U of T, currently surpasses all other methods in accurately extracting communities. This algorithm is publicly accessible as the "bayanpy" library through Python's package installer (pip). In this project, the student will receive supervision and complete a series of regular research tasks which all involve Python programming. These tasks will encompass activities such as data analysis, computational experiments, and the implementation and testing of new algorithm enhancements in a GitHub environment. This project leverages cutting-edge techniques in computing and optimization to advance the solution of a fundamental NP-complete problem with precision and efficiency. The results of this research project contribute to the development of a reliable, open-source, and reproducible algorithm for the robust and theoretically grounded identification of communities, thereby enhancing a widely used optimization-based tool in the field of data science. Required skills of the student Background and experience in: -Python -Machine learning -Python libraries for data analysis (pandas, numpy, seaborn, scikit-learn) Other desired skills (to have or acquire during the project): -Python libraries for large-scale/network data analysis and optimization (networkX, CDlib, igraph, Gurobi, joblib) -collaborative software development -discrete optimization (e.g., branch and bound) -data science methods -graph theory and network science -open science -reproducible research.

Additional Information	N/A
Application Instructions	Interested MIE students should send one email to Prof. Samin Aref (aref@mie.utoronto.ca) with the following items combined into one pdf file: CV, unofficial transcript(s), one paragraph describing their interest in the project, one paragraph describing how they match the skills listed in the description of the project.