Mechanical & Industrial Engineering
UNIVERSITY OF TORONTO

MIE498H1: Research Thesis 2026-2027

Supervisor	Samin Aref
Supervisor email	s.aref@utoronto.ca
Number of Positions	1
Open to	Undergraduate Industrial Engineering Students
Term Offered	Fall (F)
Research Area	Operations Research
Research Topic	Developing an approximation algorithm for a network clustering problem using integer optimization

Project Description

Community detection, the process of identifying groups within a network, is a core problem in network science and data science [1], [2], [3]. Among the most common approaches for detecting communities are the modularity-maximization algorithms [4], [5], [6], which are designed to maximize a utility function, modularity [7], across all possible ways that the nodes of a network can be partitioned into communities. There are several integer programming formulations for the modularity maximization problem. One of the strongest formulations is proposed in [8]. In this thesis project, the student will develop a new approximation algorithm for a network clustering problem and makes it available as a part of the open-source python library cd-lib. Technical support and supervision will be provided during regular bi-weekly meetings. References [1] Z. Yang, R. Algesheimer, and C. J. Tessone, "A Comparative Analysis of Community Detection Algorithms on Artificial Networks," Sci. Rep., vol. 6, no. 1, p. 30750, Aug. 2016, doi: 10.1038/srep30750. [2] S. Fortunato and M. E. J. Newman, "20 years of network community detection," Nat. Phys., vol. 18, no. 8, Art. no. 8, Aug. 2022, doi: 10.1038/s41567-022-01716-7. [3] S. Aref, H. Chheda, and M. Mostajabdaveh, "The Bayan Algorithm: Detecting Communities in Networks Through Exact and Approximate Optimization of Modularity," ArXiv Prepr. ArXiv220904562, 2022. [4] M. Chen, K. Kuzmin, and B. K. Szymanski, "Community detection via maximization of modularity and its variants," IEEE Trans. Comput. Soc. Syst., vol. 1, no. 1, pp. 46–65, 2014. [5] A. Lancichinetti and S. Fortunato, "Limits of modularity maximization in community detection," Phys. Rev. E, vol. 84, no. 6, p. 066122, Dec. 2011, doi: 10.1103/PhysRevE.84.066122. [6] M. E. J. Newman, "Equivalence between modularity optimization and maximum likelihood methods for community detection," Phys. Rev. E, vol. 94, no. 5, p. 052315, Nov. 2016, doi: 10.1103/PhysRevE.94.052315. [7] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," Phys. Rev. E, vol. 69, no. 2, p. 026113, Feb. 2004, doi: 10.1103/PhysRevE.69.026113. [8] T. N. Dinh and M. T. Thai, "Toward optimal community detection: From trees to general weighted networks," Internet Math., vol. 11, no. 3, pp. 181–200, 2015.

Additional Information

Required skills:
1. Experience with formulating and solving linear programming models.
2. Experience with formulating and solving discrete optimization models (e.g., integer or mixed-integer programming). Python programming Gurobi Python

libraries for optimization 5. Python libraries for data analysis (pandas and numpy)

Other desired skills (to have or acquire during the project):
1. Familiarity with networks and graphs
2. Basic familiarity with network clustering, community detection, and modularity maximization
3. Academic writing and research experience

Application Instructions

Interested MIE INDY students should send one email to Prof. Samin Aref (s.aref@utoronto.ca) 2 months before the enrolment deadline, with the following combined into one pdf file: CV, unofficial transcript(s), one human-written paragraph describing their interest in the project, one human-written paragraph describing how they match the skills listed in the description of the project, grades from relevant courses.

A take-home assignment will be given to the applicants to assess the match between their skills and the project before approving their enrolment in the thesis project.