# FINAL PROJECT FOR MODERN OPTIMIZATION METHODS
## FALL SEMESTER 2020

The main goal of this final project is to apply algorithms learnt from this course, and to gain the ability to judge and compare these algorithms to the existing ones. Here is a list of components that will be required in the final project.

1. **Social Network Analysis.** *(10 points)* The first part of this project is to gain some background knowledge on social network analysis:

   (a) What is a Social Network? State the definitions of the basic components (node and edge).

   (b) What is a center of a network? State the definition of the so-called "Degree Centrality".

   (c) What is a community of a network? State the definition of a community using a criterion called "Modularity".

   (d) State the community detection problem of a social network as an optimization problem.

2. **Commmunity Detection via Metaheuristics.** *(25 points)* This final project aims at tackling community detection problems using metaheuristics optimization algorithms learnt in this course. In this section, please implement the following algorithms in Community Detection problem:

   (a) Genetic Algorithm.

   (b) Simulated Annealing.

   (c) Tabu Search.

   (d) Particle Swarm Optimization (You cannot use SIB method).

   (e) Ant Colony Optimization.

   For each method, state clearly the optimization problem, the definition of a particle (data structure, use the one I mentioned in the lecture), and all necessary parameters with adequate explanations. The objective function is the modularity. One specific constraint is that the number of members in a community must be less than 10% of the total number of nodes in the network. You can choose to use any programming language, but it is necessary for each algorithm to be written in a function form.

3. **A Small Test on the Karate Network.** *(15 points)* We will use this famous Karate network to test your program. The Zachary Karate Club Network, or the karate network in short, is a dataset that contains social ties among the members of a university karate club collected by Wayne Zachary in 1977. There are 34 nodes and 78 edges in this network. In this problem, the specific constaint changes from 10% to 50%.

   (a) Write the formal statement of the optimization problem.

   (b) Calculate how many potential combinations of results in this problem. Note that a community must have at least 2 nodes, but the maximum number of nodes is restricted to be less than 50% of the total number of nodes.

   (c) Implement all five algorithms for the community detections of the karate network. Do only 10 iterations and use only 5 particles for group methods.

(d) Check the performance using a graphical representation.

4. **The Community Detection on the Coauthorship Network of a Specific Subject.** *(45 points)* After running your algorithms with a small testing network, we will study a large-scale citation .

   (a) Send me an email with title "MATH 5080 - (your student ID) - FINAL EXAM DATA". You will receive a dataset via email.

   (b) Provide some basic statistics of the network. List the number of nodes and the number of edges in the network.

   (c) Among five algorithms, choose four of your choices and implement them for the community detection of the coauthorship networks. The maximum number of iteration is set at 100. For grouped methods, choose an appropriate number of particles. Choose appropriate parameter values for these methods too.

   (d) Check the performance using a graphical representation, and explain the strength and weakness of each method.

5. **Conclusion.** *(5 points)* In the last part of this final project, please provide the following information:

   (a) Summarize your comparison results on four methods of your choice, both graphical and written in sentences.

   (b) State the advantages and disadvantages of all four methods shown in this application to Community Detection Problem.

   (c) State at least one potential improvement on the best method to make the algorithm even better on this problem.

6. **How about Doing all Known Method.** *(Bonus 10 points)* A bonus score will be given if you implement all five methods (instead of four) in Question 4, and answer all questions in Question 5 based on five methods.