Author: Maulana, A.
Title: Many objective optimization and complex network analysis
Issue Date: 2018-12-05

Summary

This thesis is combining two fields of computer science: Multicriteria optimization and complex network analysis. On the one hand, it investigates methods from complex network analysis to understand conflicts and dependencies between objective functions in many objective optimization, that is multicriteria optimization with a large number of criteria. On the other hand, it proposes the use of multicriteria optimization to analyze and solve problems in complex networks. In particular, problems to modularity and centrality in multiplex networks are addressed. Applications of the proposed methods are found in facility location problems, economic and biological network analysis, as well as in epidemiology.

Many objective optimization problems occur in settings where various interest groups and a diverse set of criteria have to be considered in order to find an optimal decision. Moreover, the number of possible decisions is very high and search algorithms are needed to explore the decision space. An example could be an urban planning problem, such as finding optimal locations of hospitals and schools. Ideally each citizen of the town should be satisfied and costs (both economic and environmental) should be minimized. If the number of criteria is very big, it is very difficult for human decision makers to understand and to discuss the problem. Moreover, optimization algorithms have difficulties to process large number of objectives. The Community Detection for Many-Objective Optimization (CoDeMo) technique is developed in this thesis for structuring: First we have a problem with many objectives, then evaluate this problem to determine correlations between objective functions. Based on the correlation matrix we can construct a network and detect the communities (clusters) of the network. Then we decompose the problem into independent subproblems and aggregate clusters with complementary objectives. Then the lower dimensional subproblems can be solved with state-of-the-art techniques in multi-objective optimization, and the solutions of the subproblems are merged again. The CoDeMo process has been successfully exemplified on a problem with 30 and a problem with 50 objectives in facility location, and for a model problem on gene regulatory network synthesis (NK- landscapes) with 10 objectives.

A challenging problem in complex network analysis is the detection of communities

and the assessment of centrality of nodes in multiplex networks. A multiplex network is a network with a fixed set of nodes, but different sets of edges (links). The different sets of edges form different layers of the multiplex network. For instance, in an economical networks one could consider different countries as nodes, and trade networks in different commodities, where for each type of commodity (e.g. oil and coffee) corresponds to one layer. It is in general more difficult to detects communities or clusters in these networks, because countries that form a trade cluster concerning one layer (e.g. oil trade) may not form a cluster in another layer (e.g. coffee trade). More important, the importance (centrality) of countries or nodes might differ from layer to layer, that is a country that is important in oil trade might not be important in coffee trade. The novel Pareto Front Modularity for Multiplex Networks (PaMoPlex) has been developed in the thesis as a tool to analyze the community structures Multiplex networks and to which extent they overlap between layers or differ. For this multi-objective combinatorial optimization is an essential ingredient. It yields a matrix of Pareto fronts that can then be grouped based on layers that are similar in their community structure. Also independence and the degree of conflict between communities can be assessed by means of PaMoPlex. It is applied to the world trading economy network layered into 11 commodity groups. Moreover, a similar approach was used to analyse the centrality of nodes. Here, enumeration algorithms can be used to find all Pareto optimal solutions, that is nodes that when compared to any other node are more important at least in one layer with respect to that node. It was found for the world economic trade network that there were seven Pareto optimal countries, of which five of them were in the G8 at this time.

Finally, the problem of immunization of complex networks was investigated. This problem has recently achieved a lot of attention, due to the increased risk of pandemia in a globalized world. The outbreak of Ebola and Zika are two examples. Moreover, similar spread can be observed when a computer virus spreads in the internet.

The problem of immunization is to modify networks in such a way, that the spreading of a virus is inhibited. Moreover, the cost of implementing the modification of the network should be minimized. The idea is to immunize or quarantine a subset of nodes in the network in order to decelerate or prevent the spread of a virus. However the number of possible subsets of nodes in a network is very large and heuristical combinatorial search needs to be applied, also due to the proven NP hardness of the resulting subset selection problem.

We developed a problem specific genetic algorithms that can solve the problem in some cases much better than the Netshield Plus algorithm, that so far was the state-of-the-art approach and based on a greedy node removal strategy. In this thesis it was also suggested for the first time to use bi-objective combinatorial optimization to solve this problem, that is, to find all Pareto optimal solutions for it, trading-off the cost of a immunization strategy and its effect. The approaches were tested on example networks, including the US flight network, pandemics board game network, and networks from the common benchmark networks used in literature.