University of Salerno

Department of Mathematics

PhD in Mathematics, Physics and Applications
Curriculum: Mathematics
CYCLE: XXXV

PhD Thesis
Hidden Structures and Conflicting Items in Combinatorial Optimization Problems

Tutor
Prof. Raffaele Cerulli

Candidate
Domenico Serra

Co-Tutor
Prof. Francesco Carrabs

Coordinator
Prof.ssa Patrizia Longobardi

2021-2022

Abstract

Combinatorial optimization-based methods are widely used to solve complex real life problems. In this thesis, we use some of these methods for addressing several emerging combinatorial optimization problems on graphs that can be classified in two macro areas: *i)* graph substructures identification problems; *ii)* combinatorial optimization problems with conflict constraints.

In graph theory, graphs are defined as mathematical structures that describe entities of interest (as nodes) and their relationships (as edges). Many real-world problems can be described through the use of a graph. For example, graph theory finds application in: *i)* the context of social network analysis, where graphs are used to represent the interactions (edges) between users (nodes); and in *ii)* the study of biological networks, such as protein-protein interaction, where the nodes are proteins, and the edges represent their physical interactions. In graph analysis one common application is the identification of clusters (or communities) of nodes that are tightly connected. In social networks, a community could represent a set of users sharing the same interest, while in the protein-protein interaction networks it could represent a set of very similar proteins forming a protein complex.

In this thesis, three problems related to identifying graph substructures have been tackled.

The first problem addresses the 2-Edge-Connected Minimum Branch Vertices, that finds application in the design of optical networks. A graph is 2-Edge-Connected if by removing one edge, the graph is still connected. The problem looks for a spanning 2-edge connected subgraph having the minimum number of branch vertices that is vertices with degree strictly greater than two. In these networks, branch vertices are associated with switch devices that split the light signals and send them to the adjacent vertices. For this NP-complete problem we developed a genetic algorithm using ad-hoc designed operators.

The second addressed problem arise in the social network analysis and aims to study how users influence the choices of their neighbours. In particular, we addressed the Collapsed k-Core Problem that seeks to identify a subset of critical users in the network whose choices would alter the cohesiveness of a community. To the best of our knowledge, this is the first attempt to formulate this problem using mathematical programming. We implemented multiple solution approaches and compared them on a set of benchmark instances.

The last case studied is related to network clustering, where a cluster graph is a disjoint union of cliques. The Cluster Deletion problem is defined as the identification of the minimum number of edges to remove from a network to produce a cluster graph. This is a well-known NP-hard problem and we faced it using integer linear programming formulation and a heuristic approach based on edge contraction operation. Our results show the effectiveness of our methodology both on artificial and real-world biological networks.

Currently, the definition of many real-life problems, doesn't always fully capture their complexity. Indeed, classical optimization problems encountered over the years, although extensively studied, do not always take into account additional limitations, such as incompatibility situations, encountered in real-world problems. In this context, two or more elements of the problem cannot be chosen together to compose a feasible solution. Such incompatibilities are modelled by introducing conflict constraints in classical combinatorial optimization problems, leading to more realistic but often harder problems.

Finally, three different combinatorial optimization problems with conflicts constraints are addressed.

The first one is a variant of the set cover problem where pairwise conflicts are added among the subsets. In the formulation of this problem, two sets in conflict can belong to the same solution, provided that a non-negative penalty is paid. We introduced two mathematical formulations for the problem and offered a parallel Greedy Randomised Adaptive Search Procedure for its solution. The performance of our algorithm was evaluated through an extensive set of experiments. The results shown the effectiveness and efficiency of our methodology compared to the mathematical model solutions.

The second problem is related to the Maximum Flow Problem with Conflict constraints on the edges, for which we present a matheuristic method based on the combination of two different approaches: Carousel Greedy and Kernel Search. The results shown that the Carousel Greedy selection substantially improves the effectiveness of the Kernel Search.

The last problem is the Minimum Spanning Tree with Conflicts, that we solved by using a Kernel Search method. Also in this case, the results on benchmark instances shown that our methods identifies a better solution compared with existing methods.