

Title: Graph optimization problems related with social network balance

Advisors: Philippe Michelon (philippe.michelon@univ-avignon.fr)
Rosa Figueiredo (rosa.figueiredo@univ-avignon.fr)
Vincent Labatut (vincent.labatut@univ-avignon.fr)

Description:

This thesis aims at tackling clustering problems defined on signed networks as well as finding and solving relevant (from the applicative perspective) variants of this problem. We propose to do so by combining two complementary approaches: on the one hand, optimization methods designed for signed graphs, and on the other hand, tools developed for the analysis of complex networks.

In a *signed* graph, edges are labeled with either a negative or a positive sign. This type of graph was introduced by Heider in 1946 with the purpose of describing sentiment relations between people pertaining to the same social group. In the last decades, signed graphs have shown to be a very attractive discrete structure for social network researchers but also for researchers from other scientific areas, including portfolio analysis in risk management and biological systems.

A real-world network is rarely perfectly balanced, no matter how the concept of balance is defined. One big challenge is to evaluate how much balance there is in a social network. For this purpose, one first defines a measure of balance and then solve a related graph optimization problem: the optimal solution found allows us to quantify the balance of the network being treated.

According to *structural* balance, a signed graph is balanced if it can be partitioned into two or more mutually hostile subgroups, each having internal solidarity. This concept can be extended in various ways. Variations of structural balance conduct to the solution of different graph clustering problems. Optimization methods can be successfully used to solve such problems. Some recent works started tackling the problems related with structural balance, but there is still much to do. A deep investigation of mathematical formulations and exact solution approaches to problems related with structural balance is still missing.

A complex network is the graph-based representation of a complex system. As such, it has non-trivial topological properties like scale-freeness or small-worldness (high transitivity and low average distance). Consequently, specific tools were developed in order to analyze them and understand the systems they represent. The domain has been very active during the last two decades, especially regarding the community detection problem. This task consists in partitioning the node set of a network, with the constraint of obtaining groups of nodes with dense intra-group and sparse inter-group connections. Although it comes from a distinct field, this problem can be considered as a more general version of the balance-related ones introduced above. In the context of complex network analysis, the efforts have been focused almost exclusively on the processing of unsigned networks; only a very few, and recent, works tried to deal with structural balance.

This thesis aims at tackling the balancing problem with a complex network perspective. The PhD candidate will investigate a number of real-world signed networks, in order to identify the various forms the balancing problem can take in practice. This will allow us to develop mathematical formulations for each variant, and to propose efficient methods to solve them. The thesis goal is to answer practical questions such as: Which is the minimum number of individuals that should be removed from a group in order to obtain perfect balance? Is it possible to identify a mediation set in a social group? Does a dynamic network evolves towards a more balanced state? Are some subgroups more balanced than others? Which is the minimum number of relations that should change in a group in order to obtain a perfectly balanced network?

Another important aspect the candidate will have to study is the effect of the network structure on the complexity of the problems being solved, as well as on the efficiency of the resolution methods proposed. For community detection, the presence of certain topological properties is known to negatively affect some algorithms. The identification of such properties will allow us to define a model aiming at generating realistic signed networks.