

# Title: Graph optimization problems related with social network balance

**Advisers:** Rosa Figueiredo <[rosa.figueiredo@univ-avignon.fr](mailto:rosa.figueiredo@univ-avignon.fr)>  
Vincent Labatut <[vincent.labatut@univ-avignon.fr](mailto:vincent.labatut@univ-avignon.fr)>

**Context.** This internship position 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.

**Signed Graphs.** 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.

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. 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 select a measure of balance and then solve a related graph optimization problem: the optimal solution found allows to quantify the balance of the network being treated. Variations of the notion of structural balance conduct to the need to solve different variants of the graph clustering problem. 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 efficient solution approaches to problems related with structural balance is still missing.

**Complex Networks.** 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.

Community detection 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 focused almost exclusively on the processing of unsigned networks; only a very few, and recent, works tried to deal with structural balance.

**Expected Work.** This internship position aims at tackling the balancing problem with a complex network perspective. The student will **investigate a number of real-world signed networks** [2], in order to characterize their topological properties [3], and to identify the various forms the balancing problem can take in practice. This will allow us to develop mathematical formulations for variants of the problem and to propose efficient methods to solve it. The starting point of this internship will be the **investigation of a solution approach for the Relaxed Correlation Clustering problem** [1]. The goal is to answer practical questions such as: Is it possible to identify a mediation set in a social group? Does a dynamic network evolve towards a more balanced state? Are some subgroups more balanced than others? 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 proposed resolution methods.

## References

- [1] R. Figueiredo & G. Moura. Mixed Integer Programming Formulations for Clustering Problems Related to Structural Balance. *Social Networks*, 35:639-651, 2013.
- [2] I. Mendonça, R. Figueiredo, V. Labatut & P. Michelon: Relevance of Negative Links in Graph Partitioning: A Case Study Using Votes From the European Parliament, 2<sup>nd</sup> European Network Intelligence Conference (ENIC), 2015.
- [3] B. Kantarcı & V. Labatut. Classification of Complex Networks Based on Topological Properties, 3<sup>rd</sup> Conference on Social Computing and its Applications, 2013, p.297-304.

# Titre : Optimisation et équilibre structurel dans les graphes sociaux

**Encadrants :** Rosa Figueiredo <[rosa.figueiredo@univ-avignon.fr](mailto:rosa.figueiredo@univ-avignon.fr)>  
Vincent Labatut <[vincent.labatut@univ-avignon.fr](mailto:vincent.labatut@univ-avignon.fr)>

**Contexte.** Ce stage vise à traiter des problèmes relatifs au partitionnement de graphes signés, ainsi qu'à identifier et à résoudre des variantes de ce problème qui soient pertinentes du point de vue applicatif. On propose de procéder en combinant deux approches complémentaires : d'une part, des méthodes d'*optimisation* conçues pour les graphes signés, et d'autre part, des outils développés pour l'analyse de *réseaux complexes*.

**Graphes signés.** Dans un *graphe signé*, chaque lien est annoté avec un signe : soit positif, soit négatif. Ce type de graphe a été défini par Heider en 1946, dans le but de décrire des relations émotionnelles entre des personnes appartenant à un même groupe social. Depuis, les graphes signés ont été largement utilisés dans l'étude de réseaux sociaux, mais également dans d'autres domaines comme l'analyse de risque en gestion et l'étude de systèmes biologiques.

On dit qu'un graphe signé est *structurellement équilibré* s'il peut être partitionné en deux (ou plus) sous-groupes mutuellement hostiles (i.e. reliés seulement par des liens négatifs) tout en exhibant une solidarité interne (i.e. contenant uniquement des liens positifs). Ce concept d'équilibre peut être étendu de différentes façons. Mais quelle que soit la notion d'équilibre utilisée, un réseau réel (i.e. un graphe représentant un système du monde réel) est rarement *parfaitement* équilibré : on trouvera quelques liens positifs entre les groupes et quelques liens négatifs à l'intérieur.

L'un des défis du domaine est de quantifier le niveau d'équilibre d'un tel réseau. Pour ce faire, on doit résoudre un problème d'optimisation sur le graphe : estimer la partition permettant d'obtenir l'équilibre maximal pour une mesure d'équilibre donnée. Chaque variante de la notion d'équilibre est susceptible de mener à un problème de partitionnement différent. Quelques travaux récents se sont intéressés aux problèmes relatifs à l'équilibre structurel, mais il y a encore beaucoup à faire, notamment en termes de revue des formulations mathématiques existantes et des méthodes de résolution proposées.

**Réseaux complexes.** Un réseau complexe est la représentation sous forme de graphe d'un système complexe. Cela signifie que ce graphe possède des propriétés topologiques non-triviales, par exemple ils sont souvent sans-échelle et petit-monde. Pour cette raison, des outils spécifiques ont été développés pour les analyser et comprendre les systèmes qu'ils représentent. Ce domaine de recherche est extrêmement actif depuis les vingt dernières années, en particulier sur le problème de la détection de communautés.

La détection de communautés consiste à partitionner un graphe, avec la contrainte d'obtenir des sous-groupes de nœuds densément interconnectés, mais faiblement connectés au reste du réseau. Bien qu'il provienne d'un champ différent, on peut considérer ce problème comme une version plus générale de celui présenté ci-dessus pour les graphes signés. Dans le contexte de l'analyse des réseaux complexes, les efforts ont porté presque exclusivement sur le traitement de réseaux non-signés, et seuls quelques travaux très récents se sont intéressés à l'équilibre structurel.

**Travail demandé.** Ce stage vise à traiter le problème de l'équilibre d'un graphe signé en utilisant une approche réseaux complexes. L'étudiant devra **étudier un certain nombre de réseaux réels** [2], afin de caractériser leurs propriétés topologiques [3], et d'identifier les diverses formes que ce problème peut prendre en pratique. Ceci nous permettra de développer une formulation mathématique des variantes de ce problème et de proposer des méthodes de résolution efficace. Le point de départ sera notamment **l'étude d'une solution proposée pour le problème appelé *Relaxed Correlation Clustering*** [1]. Le but est de répondre à des questions pratiques telles que : Est-il possible d'identifier un sous-groupe de médiateurs dans un groupe social ? Un réseau dynamique évolue-t-il vers un état plus équilibré ? Certains sous-groupes sont-ils plus équilibrés que d'autres ? L'effet de la structure du graphe sur la complexité des problèmes et l'efficacité des méthodes de résolution constituent deux autres points importants que l'étudiant pourra aborder.

## Références

- [1] R. Figueiredo & G. Moura. Mixed Integer Programming Formulations for Clustering Problems Related to Structural Balance. *Social Networks*, 35:639–651, 2013.
- [2] I. Mendonça, R. Figueiredo, V. Labatut & P. Michelon: Relevance of Negative Links in Graph Partitioning: A Case Study Using Votes From the European Parliament, 2<sup>nd</sup> European Network Intelligence Conference (ENIC), 2015.
- [3] B. Kantarci & V. Labatut. Classification of Complex Networks Based on Topological Properties, 3<sup>rd</sup> Conference on Social Computing and its Applications, 2013, p.297-304.