UNIVERSIDADE FEDERAL DO CEARÁ

CENTRO DE CIÊNCIAS

DEPARTAMENTO DE COMPUTAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

DOUTORADO EM CIÊNCIA DA COMPUTAÇÃO

TATIANE FERNANDES FIGUEIREDO

TEAM FORMATION PROBLEMS: AN INTEGER LINEAR OPTIMIZATION APPROACH

FORTALEZA

2021

TATIANE FERNANDES FIGUEIREDO

TEAM FORMATION PROBLEMS: AN INTEGER LINEAR OPTIMIZATION APPROACH

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Ciência da Computação. Área de Concentração: Algoritmos e Otimização

Orientador: Prof. Dr. Manoel Campêlo Bezerra Neto

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

F493t Figueiredo, Tatiane Fernandes.
Team formation problems: an integer linear optimization approach / Tatiane Fernandes Figueiredo. – 2021.
130 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Ciência da Computação , Fortaleza, 2021.
Orientação: Prof. Dr. Manoel Bezerra Campêlo Neto.

1. Team formation problem. 2. Social balance. 3. Branch & Cut. 4. Valid inequality. 5. Reformulation linearization technique. I. Título.

CDD 005

TATIANE FERNANDES FIGUEIREDO

TEAM FORMATION PROBLEMS: AN INTEGER LINEAR OPTIMIZATION APPROACH

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Ciência da Computação. Área de Concentração: Algoritmos e Otimização

Aprovada em: 30/03/2021

BANCA EXAMINADORA

Prof. Dr. Manoel Campêlo Bezerra Neto (Orientador)
Universidade Federal do Ceará (UFC)

Profa. Dra. Cláudia Linhares Sales
Universidade Federal do Ceará (UFC)

Profa. Dra. Rosa Maria Videira de Figueiredo
Avignon Université

Profa. Dra. Rosiane de Freitas Rodrigues
Universidade Federal do Amazonas (UFAM)

To my sister, Taciana (in memoriam).

ACKNOWLEDGEMENTS

To my parents and to my sister Taciana (in memoriam), for all the encouragement and support, always being my inspiration and motivation to move forward.

To my dear ones: Pablo Luiz, Márcio, Rafael, Eurinardo and Bonfim for helping me to overcome challenges that life imposed on me during these years of study and that, sometimes, seemed impossible. Thank you for being present at all stages of this doctorate, always supporting me unconditionally.

To all the ParGO research group members, especially my friends Rennan, Luiz Alberto and Mardson, for have given me the opportunity to be part of their lives and to present me with their friendship.

To the professor Manoel Campêlo, tutor of this work, which many times went further, being an advisor and motivator, becoming a reference for me. Thank you very much for the enormous dedication and commitment, for the didactic excellence and knowledge transmitted, for the patience and for the confidence placed in the accomplishment of this work.

To the professor Cláudia Linhares, my female reference in Brazilian science, thank you for the great support and encouragement, for the knowledge transmitted and for all the moments of relaxation promoted during all these years of research.

To the professor Rosa Figueiredo, for all the welcome during my internship in France, for presenting the application of the Structural Balance Theory in optimization problems and for inspiring me to be a great researcher.

To the professors Rosiane Rodrigues and Laura Bahiense, for the fundamental suggestions and for accepting the invitation to participate in the examination board of this work.

RESUMO

Dado um conjunto de indivíduos, cada um com uma habilidade única, e uma rede social captando a afinidade mútua entre eles, o Problema de Formação de Equipe (TFP -Team Formation Problem) visa encontrar uma única equipe que reúna as habilidades necessárias para realizar uma tarefa, enquanto busca otimizar os custos de comunicação entre os indivíduos envolvidos. Na primeira parte deste trabalho, estudamos uma generalização do TFP denominada como Problema de Formação de Múltiplas Equipes (MTFP - Multiple Team Formation Problem), que permite demandas distintas de trabalhadores para cada tipo de habilidade, assim como requisições de múltiplas equipes de trabalho e possibilidade de fracionamento do tempo de dedicação de cada indivíduo entre os times. Nesse caso, o custo total de comunicação é dado pela soma dos pesos das relações dos pares de indivíduos de um mesmo time. Na segunda parte, introduzimos uma nova variante do TFP, a ser chamada Problema de Formação de Equipes Competitivas (CTFP - Competitive Teams Formation Problem). Utilizando a teoria do equilíbrio social, neste problema, representamos a rede social que conecta os indivíduos envolvidos por meio de um grafo de sinal e consideramos simultaneamente custos de comunicação intra-equipe e inter-equipes, requisitando que haja apenas relações positivas entre indivíduos de uma mesma equipe e apenas relações negativas entre indivíduos de equipes diferentes. Para o MTFP, propomos uma formulação de Programação Linear Inteira (ILP) e famílias de desigualdades válidas. Experimentos computacionais atestam que o modelo ILP fortalecido por desigualdades válidas supera consistentemente a formulação quadrática apresentada na literatura para resolução do MTFP. Também consideramos uma versão generalizada do MTFP em que os indivíduos podem ter múltiplas habilidades. Para lidar com esta versão, adaptamos o modelo ILP inicial, gerando dois novos modelos, e apresentamos um outro conjunto de desigualdades válidas que fortalecem os dois modelos. Para o CTFP, também propomos uma formulação ILP, além de desigualdades válidas derivadas da teoria do equilíbrio estrutural que melhoram o desempenho computacional do modelo. Por fim, encerramos este trabalho com conclusões gerais e direções para trabalhos futuros.

Palavras-chave: Problema de formação de equipes. Equilíbrio social. Branch & Cut. Desigualdade válida. Técnica de linearização-reformulação.

ABSTRACT

Given a group of individuals, each one with a single skill, and a social network capturing the mutual affinity among them, the Team Formation Problem (TFP) aims to find a single team that meets the skills needed to perform a task while seeking to optimize the communication costs between the involved individuals. In the first part of this work, we study a generalized version of the TFP denominated as Multiple Team Formation Problem (MTFP), which allows distinct demands of workers per ability as well as multiple work teams and fractions of dedication time per team for each individual. In this case, the total communication cost is given by the sum of weighted pairwise relations between members within a same team. In the second part, we introduce a new variant of the TFP to be called Competitive Teams Formation Problem (CTFP). Using the theory of social balance, in this problem, we represent the social network that connects the involved individuals as a signed graph and consider both intra-team and inter-teams communication costs by asking to have only positive relationships between individuals of a same team and only negative relationships between individuals of different teams. For the MTFP, we propose an Integer Linear Programming (ILP) formulation and sets of valid inequalities. Computational experiments attest that the ILP model strengthened by valid inequalities consistently outperforms the existing quadratic formulation for MTFP. We also consider a generalized version of the MTFP where individuals may have multiple skills. To handle this version, we adapt the initial ILP model into two new models and present other valid inequalities. For the CTFP, we also propose an ILP formulation and valid inequalities derived from the structural balance theory that enhance the computational performance of the model. Finally, we close this work with general conclusions and directions for future works.

Keywords: Team formation problem. Social balance. Branch & Cut. Valid inequality. Reformulation linearization technique.

LISTA OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

In this initial chapter, we present a general description of team formation problems along with a brief review of the associated literature. In this context, we introduce the two main problems studied in this work: the Multiple Team Formation Problem (MFTP) and the Competitive Teams Formation Problem (CFTP). Our main contributions and the text organization are presented in the sequel.

1.1 Team formation

In the last decades, work teams formation has been shown as pivotal to the organizational success of companies (COHEN; BAILEY, 1997; BARRICK *et al.*, 1998). In the face of hard competition and technological challenges, an increasing number of companies has used group work strategies as an effective approach to resolving employee motivation issues and accomplishing organizational productivity goals (NEUMAN *et al.*, 1999).

Although much scientific research has proven that the technical competences are fundamental attributes on teams design (FITZPATRICK *et al.*, 2001; TSENG *et al.*, 2004; HLAOITTINUN *et al.*, 2007), the success or failure of a work team is also strongly influenced by the health of the relationships amongst its members (BALLESTEROS-PÉREZ *et al.*, 2012). The social identity theory suggests that the more members identify with their respective teams, the more likely they are to actively contribute to the welfare of the group and work towards common goals (GUNDLACH *et al.*, 2006). Indeed, empirical research indicates that members of these teams display higher affective commitment and have higher performance when compared to teams that have low cohesion (ANDREWS *et al.*, 2008).

Thus, the challenge of assembling successful work teams can be defined as the problem of finding sets of workers who have certain desired skills, while simultaneously considering the social compatibility among them (LAPPAS *et al.*, 2009). This kind of problem has received great attention from the OR community over the past years. In the literature, several ways to measure the quality of relationships among team members have been proposed, most of them based on psychology tools such as Myers-Briggs Type Indicator (MYERS, 1962; MYERS, 1997), Kolbe Conative Index (KOLBE, 1993), Sociometry (MORENO, 1941).

The first attempt to incorporate skills requirements and expectations of group synergy into a quantitative model was presented in 2001. Askin e Huang (2001) introduced a mixed-

integer goal programming model to form teams for cellular manufacturing, where the analysis of social relationships was accomplished through the Kolbe Conative Index (KCI) (KOLBE, 1993). KCI was used to study the behavior or instinctive drive of each individual as a way to evaluate the interpersonal style of workers. In the sequel, Chen e Lin (2004) presented a quadratic programming model that receives as input the personality profile of each worker. For such, the Myers–Briggs type indicator (MYERS, 1997) was used to identify the personality and preferences of each team member.

With the advent of online social networks, Wi *et al.* (2009) and Lappas *et al.* (2009), in a pioneering way, introduced the use of Social Network Analysis (SNA) techniques to measure relationships between members of possible work teams. Using a graph that represents a social network, basic SNA concepts such as degree, centrality and closeness were employed in algorithms and mathematical models to identify a team of experts. Additionally, Lappas *et al.* (2009) formally defined the problem of finding a work team in a social network. Given a task that requires a set of skills and a pool of workers organized in a social network, each one possessing one of the skills, the *Team Formation Problem* (TFP) refers to finding a subset of the workers that cover the skill requirements and can communicate effectively with each other. The concept of effective communication can be measured by various *communication cost* metrics. In their work, Lappas *et al.* (2009) computed the communication cost based on the distances between team members in the social network. They proved that, similar to many other grouping optimization problems, the TFP is NP-Hard (for single team instances).

Due to the current relevance of work teams formation in the business context and the popularization of online social networks, the seminal paper by Lappas *et al.* (2009) has triggered an increasing number of researches on the TFP or related problems. Many papers have been published at a fast pace in literature (LI; SHAN, 2010; KARGAR; AN, 2011; MAJUMDER *et al.*, 2012; ANAGNOSTOPOULOS *et al.*, 2012; BALLESTEROS-PÉREZ *et al.*, 2012; RANGAPURAM *et al.*, 2013; GUTIÉRREZ *et al.*, 2016; CAMPÊLO *et al.*, 2020; KOUVATIS *et al.*, 2020). Some of them present extensions for the TFP, for instance dealing with multiple team formation, part-time allocation and/or multiple skilled workers, as well as with different metrics to evaluate the communication costs.

In this scenario, problems involving the formation of multiple teams have been the focus of more recent works. Ballesteros-Pérez *et al.* (2012) introduced an optimization process carried out by means of matrix calculations that seeks both to maximize positive interactions and

to minimize negative interactions between pairs team members. For such, the authors assume that all employees have the same skills and knowledge to carry out the required tasks. By their turn, Gutiérrez *et al.* (2016) presented a generalization of the TFP which allows distinct demands of workers per skill, as well as fractions of dedication time by worker in each team. In this generalized version, called *Multiple Team Formation Problem* (MTFP), it is still assumed that each worker possesses a unique skill. The authors presented a quadratic optimization model, as well as constraint programming and heuristic solution approaches. In Campêlo *et al.* (2020), another variant of TFP was considered, where disjoint teams have to be formed from a set of individuals with multiple skills. The problem was show to be NP-hard even when a single team and unique skill are considered. An integer linear program (ILP) along with a polyhedral study were presented. Computational experiments with the ILP and a simulated annealing heuristic are reported in that work.

In the three aforementioned papers, the communication costs between workers are measured using *sociometric matrices*. These matrices, derived from the sociometric theory developed by Moreno (1941), reveal the "social atoms" formed by each individual in his/her social network. Through this theory, the relationships between individuals can be modeled by an edge-weighted graph. By considering the workers as the vertices of a graph, the affinity of a team can be defined as a function of the weights of the edges linking the team members. The most common way to edge-weighting is the use of simple ordinal rating scales (ROISTACHER, 1974). In particular, the rates may simply lie in the set $\{-1, +1\}$, for example to represent like/dislike or trust/distrust (HEIDER, 1946). In this case, signs $+/-$ can be used instead so as to obtain a *signed graph*.

Also studied by social psychologist Heider (1946), the signed graphs were used to describe the relationships between individuals belonging to a same group, providing a systematic statement of the principle of *structural balance*. The structural balance theory is based on the notion of cognitive consistency between friendship and hostility. For example, an enemy of a friend is probably my enemy as well, while a friend of a friend is probably my friend or can become one. In simple terms, the interaction of individuals follows the tendency to create stable (albeit not certainly conflict-free) social groups (LEVORATO; FROTA, 2017).

Although the structural balance analysis has proven to be an attractive tool for social network researchers (YANG *et al.*, 2007; DOREIAN; MRVAR, 2009; LESKOVEC *et al.*, 2010; TANG *et al.*, 2016) and has motivated the definition of Integer Linear Programming (ILP)

models to solve clustering problems (BANSAL *et al.*, 2004; BRUSCO; STEINLEY, 2011; FIGUEIREDO; FROTA, 2013; FIGUEIREDO; FROTA, 2014), its use to ground communication costs functions in team formation problems is still recent. The only work by Kouvatis *et al.* (2020) that provided some definitions of compatibility between pairs of users in a signed network, and algorithms for computing it. The authors also defined new versions of the TFP in signed networks and presented heuristic algorithms for solving them.

1.2 Our contributions

This thesis has two main parts. In the first one, we focus on the MTFP, as defined by Gutiérrez *et al.* (2016). Besides, we consider a generalization of the problem, where individuals may have multiple skills, as in the extension of TFP proposed by Campêlo *et al.* (2020). Similarly to these two works, the communication costs are calculated by considering signed and weight sociometric matrices. More precisely, the communication cost of a team is the summation of the weights of edges induced by its members, taking proportionally to their fractions of dedication times.

We model the MTFP as an integer linear program and present valid inequalities that are derived via reformulation-linearization techniques - RLT (SHERALI; ADAMS, 1990). Computational experiments demonstrate that the ILP formulation strengthened with valid inequalities consistently outperforms the quadratic model by Gutiérrez *et al.* (2016). The tests were carried out on a large variety of MTFP instances. They are built from both synthetic and real-world social network data.

Still in the first part, we use extra variables to straightforwardly adapt our MTFP model to deal with the case where individuals may have multiple skills. Besides, we project out some of the variables to get a second ILP model with fewer variables but an exponential number of constraints. We show how to separate these constraints in polynomial time. The correctness of this model and the separation procedure is demonstrated via max-flow/min-cut arguments. We also present valid inequalities for both formulations, again using RLT techniques. We apply the separation routine to solve the second program with the branch-and-cut method. We computationally compare the two formulations and show their potential to solve instances generated from the original MTFP instances.

Preliminary results that we have obtained for the MTFP were presented at CLAIO 2018. The complete work for this problem was published in Computers & Operations Re-

search (CAMPÊLO; FIGUEIREDO, 2021). A second paper with the results for the MMTFP is about to be submitted. It is worth to mention that our approach to MTFP and MMTFP is similar to the one we have used in our Master's dissertation to another variant of the TFP, where disjoint teams have to be formed from a set of multi-skilled individuals that can dedicate their working time to at most one project. Some results for this latter problem obtained during the Master's have been published in Annals of Operations Research (CAMPÊLO *et al.*, 2020).

In the second part of the thesis, we introduce a new version of the TFP that considers the concepts presented by Kouvatis *et al.* (2020) about compatible team of individuals, which aim to forming competitive work teams. Partitioning problems in signed graphs that use similar concepts have already been presented in the literature, some of them with the aim of studying and creating groups (BANSAL *et al.*, 2004; INOHARA, 2003; FIGUEIREDO; FROTA, 2013; FIGUEIREDO; FROTA, 2014; LI *et al.*, 2015b; SHI *et al.*, 2016; ARINIK *et al.*, 2017). However, to our best knowledge, none of them considers the analysis of the individual skills.

In the version of the TFP we introduce here, to be called *Competitive Teams Formation Problem* (CTFP), we simultaneously consider skill requirements and structural balance constraints when forming multiple teams. It is required to exist only positive relationships between individuals of the same team and only negative relationships between individuals of different teams. This intends that workers within a same team have no disgust to each other whereas teams are considerably competitive with each other. The objective is to include as many individuals as possible in these groups. We model the CTFP as an integer linear program and use the principle of structural balance to derive valid inequalities that enhance its computational performance.

Regarding this second problem, there is a paper in preparation to be submitted to a journal. Part of the results was obtained during a scientific visit to the Avignon Univertité et des Pays de Vaucluse, in collaboration with Rosa Figueiredo.

1.3 Text organization

The two aforementioned parts of the thesis are structured in the format of scientific articles. This means that each one is self-contained. It includes its own introduction and bibliographic review of the approached problems, in addition to the obtained results and specific conclusions.

Chapter 2 is devoted to the Multiple Team Formation Problem (MTFP) defined by

Gutiérrez *et al.* (2016) and its generalization to comprise the case of multi-skilled individuals, denominated MMTFP. After the introductory Section 2.1, we present a review of the literature on team formation problems in Section 2.2. Then, Sections 2.3 and 2.5 formally define the MTFP and the MMTFP, respectively. The basis to our solution approaches are presented in Sections 2.4 to 2.6 which comprise the proposition of ILP formulations, valid inequalities and separation procedures. In the sequel, we present information related to our computational experiments: generation of instances (Section 2.7), computational results for MTFP and MTFP (Sections 2.8 and 2.9), and analysis of instances hardness (Section 2.10). Finally, in Section 2.11, we close the chapter with some concluding remarks.

In Chapter 3, we introduce and mathematically define the Compatible Teams Formation Problem. It uses the concept of structural balance to form teams of skilled individuals. This chapter has a structure similar to Chapter 2. We start with a introduction (Section 3.1) and then present a bibliography review on combinatorial problems based on the principle of structural balance (Section 3.2). The CTFP is then formally defined in Section 3.3. Structural properties and polyhedral results on balance are revisited and extended in Sections 3.4 and 3.5, respectively. In Section 3.6, we propose an ILP model for the CTFP together with valid inequalities. In Section 3.7, we describe and analyze our computational results. Section 3.8 ends the chapter with concluding remarks.

Finally, in Chapter 4, we close this thesis with general conclusions and directions for future works.

2 MULTIPLE TEAM FORMATION PROBLEM

In this chapter, we study the Multiple Team Formation Problem (MTFP), introduced by Gutiérrez *et al.* (2016). We also consider a generalization of the problem suggested by the same authors but still not studied in the literature. In the original version of the problem, each individual has a single skill whereas as multi-skilled individuals are considered in the more general version.

2.1 Introduction

The advent of many social networks on the internet has motivated several studies on people grouping and their interactions by the scientific community. In particular, grouping in work environments stands out. The social relationships among team members take on an important role that directly influences the success of a project (BALLESTEROS-PÉREZ *et al.*, 2012). Interactions within work teams, as well as task assignment and personal motivation, can significantly affect the final achievements (CAMPION *et al.*, 1993).

For several decades, the OR literature has dedicated different studies to the general problem of forming worker teams. However, the first attempt to incorporate functional requirements (skills) and expectation of group synergy into a comprehensive quantitative model was only presented in 2001 by Askin e Huang (2001). The authors proposed an Integer Linear Programming (ILP) formulation and heuristic methods for an aggregate worker assignment and training problem in cellular manufacturing. In the sequel, Fitzpatrick e Askin (2005) presented another ILP formulation and a metaheuristic algorithm for a similar problem. It is assumed that each individual in the labor pool has a single predetermined skill, and team skill requirements have been clearly defined. The Kolbe Conative Index (KOLBE, 1993) was used as the psychometric system to measure conative tendencies of individuals (which affect the teams dynamics).

Social networks provide an alternative for evaluating interpersonal relationships. Gaston *et al.* (2004) studied the influence of the graph structure among the individuals of a team in its performance. However, the first computational approach that explores an online social network in order to form a work team was presented by Lappas *et al.* (2009). The authors introduced the *Team Formation Problem* (TFP), whose objective is to find a single team meeting the skills needed to perform a task while maximizing the social compatibility of the involved

individuals. Let us remark that the TFP is restricted to create a single team of individuals with distinct skills. Requiring more than one expert per skill, allowing several skills per individual or considering the formation of multiple teams are possibilities not taken into account in Lappas *et al.* (2009).

Recently, two other papers deal with such alternatives. Gutiérrez *et al.* (2016) generalized the TFP by introducing the *Multiple Team Formation Problem* (MTFP), which allows distinct demands of workers per ability as well as multiple work teams and fractions of dedication time per team for each individual. In the MTFP, it is assumed that each individual possesses a unique skill. The authors presented a quadratic optimization model as well as constraint programming and heuristic approaches for the problem. In Campêlo *et al.* (2020), we considered a variant of this problem where disjoint teams have to be formed from a set of individuals with multiple skills, and each team may demand several individuals per skill. In these two studies, people's affinity is measured by sociometric matrices (MORENO, 1941). As in Lappas *et al.* (2009), the relationships between individuals can be modeled by an edge-weighted graph, and the affinity of a team is a function of the weights of edges between pairs of individuals in it.

In this work, we consider the MTFP, as defined by Gutiérrez *et al.* (2016), as well as a generalized version where each individual may have multiple skills, like in Campêlo *et al.* (2020). The characteristics of the mentioned problems based on networks are summarized in the Table 1.

Table 1 – Summary of the characteristics of the central problems related to MTFP.

Reference	# teams		Individuals per skill		Skills per individual		Fractions of dedication time	
	$= 1$	≥ 1	$= 1$	≥ 1	$= 1$	≥ 1	No	Yes
Lappas *et al.* (2009)	✓		✓		✓		✓	
Gutiérrez *et al.* (2016)		✓		✓	✓			✓
Campêlo *et al.* (2020)		✓		✓		✓	✓	
This work		✓		✓	✓	✓		✓

We model the MTFP as an integer linear program and present valid inequalities that greatly enhance its computational performance. These inequalities are derived via reformulation-linearization techniques - RLT (SHERALI; ADAMS, 1990). We carried out computational experiments with 162 synthetic instances generated according to Gutiérrez *et al.* (2016), as well

as with 324 instances adapted from real-world datasets available on the internet. The results show that our Integer Linear Programming (ILP) formulation consistently outperforms the quadratic model by Gutiérrez *et al.* (2016). They also demonstrate the effectiveness of the proposed valid inequalities.

By using extra variables, we adapt our model to deal with the case where individuals may have multiple skills. Such an extension of MTFP was mentioned by Gutiérrez *et al.* (2016) as an interesting line of research. For this more general case, we also present a second ILP model with fewer variables but an exponential number of constraints. We show how to separate these constraints in polynomial time and use the separation routine to solve the program. We computationally compare the two formulations and show their potential to solve 486 instances generated from the original MTFP instances.

The remainder of this text is structured as follows. In Section 2.2, we present a review of the literature related to team formation problems based on sociotechnical criteria. The MTFP is then formally defined in Section 2.3. In Section 2.4, we propose an ILP model for this problem. Section 2.5 is devoted to the case with multi-skilled individuals. Two ILP models and separation procedures are presented. In Section 2.6, we derive some classes of valid inequalities for both versions of the problem. Section 2.7 describes the MTFP instances used in the experiments. They are based on sociometric matrices generated according to Gutiérrez *et al.* (2016) or extracted from real-world social networks. Section 2.7.1 shows how these matrices are obtained. In Section 2.8, we describe and analyze numerical experiments with the formulations and valid inequalities proposed for MTFP in this work as well as with the existing formulation in the literature. In turn, the computational results for MMTFP are presented in Section 2.9. Additionally, in Section 2.10, we use the results of the computational experiments to analyse the hardness of MTFP and MMTFP instances. Finally, in Section 2.11, we close the chapter with general conclusions and future works.

2.2 Related work

It is well-known that productivity depends not only on the technical skills but also on effective interaction among team members. Researches that seek to study and quantify the effectiveness of worker teams through the analysis of the interpersonal relationships of the individuals and their respective skills have received much attention in the psychological and behavioral literature for a long time (see e.g. Gladstein (1984), Goodman *et al.* (1988), Guzzo

e Shea (1992), Colarelli e Boos (1992), Campion *et al.* (1993), Little e Madigan (1997) and references therein).

In order to measure the synergy of worker teams and predict their performance, different methodologies have been applied, such as Myers-Briggs (LIND, 1995; KUIPERS *et al.*, 2009), Sociometry (LODAHL; PORTER, 1961; COLARELLI; BOOS, 1992; WHEELAN, 2009; BALLESTEROS-PÉREZ *et al.*, 2012; GUTIÉRREZ *et al.*, 2016) and Kolbe Conative Index (ASKIN; HUANG, 2001; FITZPATRICK; ASKIN, 2005; FITSILIS *et al.*, 2015). They have been used to prove that the productivity of a team is strongly influenced by the health of the social relations between its members.

Although there are several pieces of research about this subject, the use of computational techniques to solve team formation problems, considering both social and technical requirements, is still recent. Possibly, the first work in this vein was presented only in 2001. In that work, Askin e Huang (2001) introduced an Integer Linear Programming (ILP) formulation that takes into account functional requirements (skills) and conative measures to form worker teams for cellular manufacturing. A similar problem in the same context was studied by Fitzpatrick e Askin (2005), who presented an ILP formulation and a metaheuristic algorithm. In both works, an individual is assigned to a team with a single skill he/she possesses in order to satisfy its technical requirements. Besides, the Kolbe Conative Index is used for the behavioral analysis of individuals.

Later, in 2009, Lappas *et al.* (2009) proposed to use a social network to assess how effectively professionals can collaborate. They introduced the problem of gathering a group of individuals with specific skills while maximizing their social/professional compatibility (which was measured by a metric/function known as communication cost). Such a problem was named Team Formation Problem (TFP). The TFP is NP-Hard (LAPPAS *et al.*, 2009), like many other grouping optimization problems (e.g. Falkenauer (1998), Majumder *et al.* (2012), Kargar *et al.* (2012), Campêlo *et al.* (2020)).

Following this paper, a long list of other works have addressed problems closely related to the TFP in the sense that they aim to select harmonious groups of people with desired technical skills and bases the analysis of interpersonal affinities on a social network. This list includes the works by Farhadi *et al.* (2011), Anagnostopoulos *et al.* (2012), Kargar *et al.* (2012), Majumder *et al.* (2012), Chhabra *et al.* (2013), Awal e Bharadwaj (2014), Li *et al.* (2015a). All these works focus on heuristic or approximation algorithms. We refer to Li *et al.* (2015a) and

Gutiérrez *et al.* (2016) for detailed comparisons of specific characteristics of the approached team formation problems.

Regarding the metrics used to assess a group's affinity, we underline those based on sociometry. Created by Moreno (1941), the sociometry studies the structure of groups through positive and negative links between individuals in a social network. From the application of sociometric tests, it is possible to measure the health of groups and their social status across quantitative sociometric tools (HARDING, 1952). The use of sociometry together with computational techniques has been studied for the creation and analysis of different types of worker teams with significant results about their performance and effectiveness (FERSHTMAN, 1997; ARINGHIERI, 2009; CHEN *et al.*, 2013). In the context of the TFP, the sociometry has been used by Ballesteros-Pérez *et al.* (2012), Gutiérrez *et al.* (2016), Baghel e Bhavani (2018), Silva e Krohling (2018). These authors consider variants of the TFP and propose solution methods grounded on metaheuristics.

Here, we highlight the work by Gutiérrez *et al.* (2016), who define the *Multiple Team Formation Problem* (MTFP), the focus of our work. This generalized version of TFP considers the requisition of multiple teams, the possibility of time allocation fractions for each individual, and different demands of workers per skill in each team. It is assumed that each individual has a single skill. The analysis of interpersonal relationships is based on a "sociometric matrix" given as input. Then, the goal is to maximize the "global efficiency", which can be understood as the percentage of positive links between all pairs of individuals assigned to a same team. To solve the MTFP, Gutiérrez *et al.* (2016) proposed an Integer Quadratic Programming formulation as well as three heuristic algorithms. The solution approach based on the Variable Neighborhood Research (VNS) metaheuristic shown to be the most promising one according to the computational experiments.

To the best of our knowledge, we are the first to model the MTFP via ILP although the literature has presented ILP models for related problems. For instance, Kargar *et al.* (2012) dealt with the following variant of TFP. Given a set of experts with a pairwise communication cost among them, and each one having a set of skills and an associated personnel cost, find a unique team that covers a set of required skills and minimizes a linear combination of the total personnel and communication costs of the team. When allocated to the team, an expert can be responsible for more than one required skill. More recently, Campêlo *et al.* (2020) presented the problem of finding disjoint teams, as harmonious as possible, where each team requires a

minimum number of individuals per skill. In this problem, each individual may have multiple skills but can be assigned to at most a work team with one of his/her skills. An integer linear program was presented along with a polyhedral study. Computational experiments with the formulation strengthened by proposed facet-defining inequalities and with a simulated annealing heuristic are reported.

2.3 Notation and problem definition

In this section, we present the Multiple Team Formation Problem (MTFP), as defined in Gutiérrez *et al.* (2016). We adopt the same basic notation of the original reference. Let the 8-tuple (P,H,K,Q,D',R,S,W) stand for:

- a *set of projects* $P = \{p_1,\ldots,p_m\}$ to which individuals have to be allocated;
- a *set of available individuals* $H = \{h_1,\ldots,h_n\}$ that can be allocated to P (each individual is considered to have a single skill);
- a *set of skills* $K = \{k_1,\ldots,k_f\}$ that individuals in H possess (each skill is required by at least one project in P);
- a *set of lists of individuals* $Q = \{Q_1,Q_2,...,Q_f\}$, where Q_a contains the individuals from H who share skill $k_a \in K$;
- a *set of individual's allowed time allocation fractions* $D' = \{d_1,\ldots,d_t\}$, $d_1 < d_2 < \cdots < d_t$ – for example, $D' = \{0,1\}$ (full-time allocation) or $D' = \{0,0.5,1\}$ (half-time allocations); we also define $D = D' \setminus \{0\}$;
- a *project requirements matrix* $R = [r_{al}]_{f\times m}$, where each entry r_{al} is a non-negative real number that specifies the fractions of time of individuals with skill $k_a \in K$ are needed for the project $p_l \in P$ – for example, $r_{al} = 1.5$ could be fulfilled with a full-time allocation of an individual plus a half-time allocation of another one;
- a *sociometric matrix* $S = [s_{ij}]_{n\times n}$ that contains the predisposition of each individual $h_i \in H$ to work with individual $h_j \in H$; the matrix elements can take on three possible values, $s_{ij} \in \{-1,0,+1\}$, where value $+1$ is chosen if h_i is willing to work with h_j, value -1 if h_i would prefer not to work with h_j, and 0 if h_i is neutral to work with h_j; elements s_{ii} are always equal to $+1$;
- a *list of weights* $W = \{w_1,w_2,\ldots,w_m\}$ that describes the *priorities of the projects*. For the sake of maintaining the objective function in the range $[0,1]$, the sum of all elements in W equals 1.

As defined by Gutiérrez *et al.* (2016), the MTFP consists in determining m teams, each one associated with a project, such that: (i) each individual $h_i \in H$ can dedicate a fraction $x_{il} \in D'$ of his/her working time to each project $p_l \in P$, performing his/her unique skill (of course not exceeding 100% of dedication time in total); (ii) each project $p_l \in P$ requires r_{al} units of working time for each skill $k_a \in K$, that can be partitioned among several workers; and (iii) the global harmony of the teams, expressed by the *global efficiency* $E = \sum_{l=1}^{m} w_l e_l$, has to be maximized, where e_l is the *project efficiency* of p_l, defined by:

$$e_l = \frac{1}{2}\left[1 + \frac{\sum_{h_i,h_j \in H} s_{ij} x_{il} x_{jl}}{\left(\sum_{k_a \in K} r_{al}\right)^2}\right]. \tag{2.1}$$

Note that $s_{ij}x_{il}x_{jl}$ represents the relationship between the pair h_i and h_j weighted by their dedications to project p_l. Thus, the numerator of the fraction in (2.1) sums these scores for p_l. Besides, since $\sum_{h_i \in H} x_{il} = \sum_{k_a \in K} r_{al}$ and $s_{ij} \in [-1,+1]$, such a fraction lies in the interval $[-1,1]$. The additive term 1 together with the multiplicative term $1/2$ leads then e_l to be a value in $[0,1]$. Therefore, the global efficiency is a weighted average of the project efficiencies, and so it is also a value between 0 and 1. It is assumed that the projects run simultaneously during the time period considered.

To present a mathematical program for MTFP, Gutiérrez *et al.* (2016) consider the following assumption:

A1 The elements in $D' = \{d_1, d_2, \ldots, d_t\}$ are equally spaced, i.e. $d_u = (u-1)\alpha$ for some constant $\alpha > 0$, for all $d_u \in D'$.

Using the variable transformation $x' = x/\alpha$, they get

$$e_l = \frac{1}{2}\left[1 + \frac{\sum_{h_i,h_j \in H} \alpha^2 s_{ij} x'_{il} x'_{jl}}{\left(\sum_{k_a \in K} r_{al}\right)^2}\right], \tag{2.2}$$

where x' is an integer valued vector with entries in the range $[0,...,t-1]$. Thus, they propose the following integer quadratic model for MTFP:

$$\max \quad \sum_{p_l \in P} w_l e_l \tag{2.3}$$

$$\text{s.t:} \sum_{p_l \in P} x'_{il} \leq 1/\alpha, \qquad \forall h_i \in H, \tag{2.4}$$

$$\sum_{h_i \in Q_a} x'_{il} = r_{al}/\alpha, \qquad \forall p_l \in P, \forall k_a \in K, \tag{2.5}$$

$$x'_{il} \in \{0,1,\ldots,t-1\}, \qquad \forall h_i \in H, \forall p_l \in P. \tag{2.6}$$

Although not mentioned in Gutiérrez *et al.* (2016), let us observe that the feasibility of this program requires another assumption:

A2 For all $p_l \in P$ and $k_a \in P$, r_{al}/α is an integer;

It is also worth remarking that, since $\sum_{h_i,h_j \in H} x_{il}x_{jl}$ is constant in any feasible solution (it is equal to $(\sum_{k_a \in K} r_{al})^2$), we can take any scalar δ, shift s_{ij} to $s_{ij}+\delta$, for every pair h_i and h_j, and still get an equivalent objective function. In particular, using $\delta = -2$, we can conclude that maximizing $E = \sum_{l=1}^{m} w_l e_l$ is equivalent to *minimizing* $E' = \sum_{l=1}^{m} w_l e'_l$, where

$$e'_l = \frac{1}{2}\left[-1 + \frac{\sum_{h_i,h_j \in H} s'_{ij}x_{il}x_{jl}}{\left(\sum_{k_a \in K} r_{al}\right)^2}\right], \tag{2.7}$$

and $s'_{ij} = 2 - s_{ij}$, for every pair $h_i, h_j \in H$. More precisely, by (2.7) and (2.1) we have

$$\begin{aligned} e'_l &= \frac{1}{2}\left[-1 + \frac{\sum_{h_i,h_j \in H}(2-s_{ij})x_{il}x_{jl}}{\left(\sum_{k_a \in K} r_{al}\right)^2}\right] \\ &= \frac{1}{2}\left[-1 - \frac{\sum_{h_i,h_j \in H} s_{ij}x_{il}x_{jl}}{\left(\sum_{k_a \in K} r_{al}\right)^2}\right] + \frac{1}{2}\frac{\sum_{h_i,h_j \in H} 2x_{il}x_{jl}}{\left(\sum_{k_a \in K} r_{al}\right)^2} = -e_l + 1. \end{aligned}$$

Since $\sum_{l=1}^{m} w_l = 1$, we obtain $E' = 1 - E$. Again, note that the multiplicative and additive terms in (2.7) were taken to have $E' \in [0,1]$. Now, the coefficients of the quadratic terms belong to $\{1,2,3\}$. Having only positive coefficients will be useful in our modeling.

The reason for weighting the relation between two individuals h_i and h_j with $s_{ij} \in \{-1,0,1\}$ comes from the sociometric theory (MORENO, 1941). As mentioned earlier, this theory has been successfully applied to predict team performance in different areas (see e.g. Aringhieri (2009), Chen *et al.* (2013)) and has been used in the definition of team formation problems (BALLESTEROS-PÉREZ *et al.*, 2012; BAGHEL; BHAVANI, 2018; SILVA; KROHLING, 2018), in addition to Gutiérrez *et al.* (2016). However, the formulation directly adapts to a general weighting matrix $S = [s_{ij}]$. In particular, by scaling s_{ij}/σ, where $\sigma = \max\{|s_{ij}|\}$, we get $s_{ij} \in [-1,1]$ and so can similarly calculate e_l and e'_l by (2.1) and (2.7). Thus, we keep the objective function value between 0 and 1.

2.4 Integer linear programming model for MTFP

To represent the allocation matrix $X = [x_{il}]$ in our ILP model for MTFP, we define the following decision variables. For each $h_i \in H$, $p_l \in P$ and $d_u \in D$, variable $y_{ilu} \in \{0,1\}$ is such that $y_{ilu} = 1$ if and only if individual h_i is allocated to team p_l with time allocation fraction d_u. In addition, variable $z_{ijluv} \in \{0,1\}$ indicates whether (value 1) or not (value 0) the pair of

individuals $h_i, h_j \in H$ is in a same team $p_l \in P$ with time allocation fractions equal to $d_u \in D$ and $d_v \in D$, respectively. For simplicity, we define the matrices $Y = [y_{ilu}]$ and $Z = [z_{ijluv}]$ with dimensions $n \times m \times t$ and $n \times n \times m \times t \times t$, respectively. However, we can fix a priori $y_{ilu} = 0$, if $h_i \in H \setminus H_l$, and $z_{ijluv} = 0$, if $h_i \in H \setminus H_l$ or $h_j \in H \setminus H_l$, where $K_l = \{k_a \in K : r_{al} > 0\}$ and $H_l = \{h_i \in H : k_a \in K_l, h_i \in Q_a\}$ respectively are the subset of skills required by project p_l and the subsets of individuals having one of these skills. Besides, we use variables z_{ijluv} and z_{jilvu} indistinctly (only one of them is effectively needed).

Using these variable, we can rewrite (2.7) as

$$e'_l = \frac{1}{2}\left(-1 + \frac{\sum\limits_{h_i,h_j \in H} \sum\limits_{d_u,d_v \in D} s'_{ij} d_u d_v z_{ijluv}}{(\sum\limits_{k_a \in K} r_{al})^2}\right) \tag{2.8}$$

In addition, we obtain the following ILP model for MTFP:

$$\min \quad \sum_{p_l \in P} w_l e'_l \tag{2.9}$$

$$\text{s.t:} \quad y_{ilu} + y_{jlv} - z_{ijluv} \leq 1, \qquad \forall h_i, h_j \in H, i \neq j, \forall p_l \in P, \forall d_u, d_v \in D, \tag{2.10}$$

$$\sum_{p_l \in P} \sum_{d_u \in D} d_u y_{ilu} \leq 1, \qquad \forall h_i \in H, \tag{2.11}$$

$$\sum_{h_i \in Q_a} \sum_{d_u \in D} d_u y_{ilu} \geq r_{al}, \qquad \forall p_l \in P, \forall k_a \in K, \tag{2.12}$$

$$y_{ilu} \in \{0,1\}, \qquad \forall h_i \in H, \forall p_l \in P, \forall d_u \in D, \tag{2.13}$$

$$z_{ijluv} \in \{0,1\}, \qquad \forall h_i, h_j \in H, i \neq j, \forall p_l \in P, \forall d_u, d_v \in D. \tag{2.14}$$

Constraints (2.10) guarantee that, if a given pair of individuals $h_i, h_j \in H$ is assigned to a same team $p_l \in P$ with certain time allocation fractions $d_u, d_v \in D$, which means setting $y_{ilu} = y_{jlv} = 1$, then the corresponding z-variable linking these individuals (z_{ijluv}) will also be set to 1, thus properly counting the sociometric value s'_{ij} weighted by the time allocation fractions in the objective function. Since the shift in the sociometric matrix yields positive coefficients for all z-variables, if one of the y-variables in (2.10) is zero, the objective sense leads the corresponding z-variable to zero too. Thus, it is needless to add constraints $z_{ijluv} \leq y_{ilu}$ and $z_{ijluv} \leq y_{jlv}$.

Constraints (2.11) and (2.12) are the counterparts of (2.4) and (2.5), respectively. The former state that each individual is not allowed to work more than 100% of his/her dedication time. The latter ensure that the demands of individuals per skill (project requirements) are met by each team. Constraints (2.13) and (2.14) define the binary domain of all decision variables. We leave it implicit that some variables are set to zero a priori.

Remark 1. *We use $\geq$-inequalities in* (2.12) *instead of equality. This way we do not need assumptions A1-A2, used in formulation* (2.3)-(2.6)*. On the other hand, under these assumptions, we can replace $\geq$ by $=$ in* (2.12)*. Indeed, equality tends to be attained in an optimal solution, since the objective function coefficients are positive. If a strictly inequality holds in* (2.12) *for some team and skill, we can always reduce (by a multiple of α) the fraction of dedication time of any individual assigned to such a team and skill to get another feasible solution of strictly better objective value, unless this team has a single person, which would lead to a solution with the same value.*

Formulation (2.9)-(2.14) can be seen as a generalization of the model presented by Campêlo *et al.* (2020) for the multiple disjoint teams formation problem. Modifications were made only to enable an individual to participate in multiple teams. It is worth reinforcing that our formulation does not assume that the elements in D are equally spaced nor that they are the same for every individual. However, without loss of generality, we can assume that each individual may be allocated to a project with a unique fraction of dedication time. Indeed, it would be a matter of enlarging set D, if necessary, to contain the allowed combinations of fractions. In other words, we could have $d_u + d_v \in D$ if $d_u, d_v \in D$ and $d_u + d_v \leq 1$. Thus, we could add the constraints $\sum_{d_u \in D} y_{ilu} \leq 1$, for all $h_i \in H$ and $p_l \in P$, to the model.

2.5 Extension for multi-skilled individuals (MMTFP)

Gutiérrez *et al.* (2016) point out that an interesting extension of MTFP consists in allowing people to be categorized under more than one skill. They claim that dealing with such an extension would involve a significantly higher level of complexity. Using our approach, however, this is not exactly the case.

Let us now admit that any individual in H may have multiple skills in K. Precisely, let $L_i = \{k_a \in K : h_i \in Q_a\}$ be the set of skills of $h_i \in H$, for all $i \in \{1, \ldots, m\}$. We consider the extension of MTFP where any individual h_i may be allocated to several teams (up to 100% of dedication time) with possibly different skills from L_i. Let us denote MMTFP this extended version of the problem.

An adaptation of formulation (2.9)-(2.14) for MMTFP can be made by defining new binary variables y'_{ilua}, for all $p_l \in P$, $h_i \in H_l$, $d_u \in D$ and $k_a \in L_i \cap K_l$, to indicate whether (value 1) or not (value 0) individual h_i is allocated to project p_l with time fraction d_u and skill k_a. Then,

we get the following model:

$$\min \quad \sum_{p_l \in P} w_l e'_l \tag{2.15}$$

$$\text{s.t: } (2.10), (2.11), (2.13), (2.14), \tag{2.16}$$

$$\sum_{d_u \in D} d_u y_{ilu} \geq \sum_{d_u \in D} \sum_{k_a \in L_i \cap K_l} d_u y'_{ilua} \qquad \forall p_l \in P, \forall h_i \in H_l, \tag{2.17}$$

$$\sum_{h_i \in Q_a} \sum_{d_u \in D} d_u y'_{ilua} \geq r_{al}, \qquad \forall p_l \in P, \forall k_a \in K_l, \tag{2.18}$$

$$y'_{ilua} \in \{0,1\}, \qquad \forall p_l \in P, \forall h_i \in H_l, \forall d_u \in D, \forall k_a \in K_l. \tag{2.19}$$

Constraints (2.17) establish that the dedication of every individual to a team includes his/her dedication to this team in his/her diverse skills. We could use equality in (2.17) under some assumptions on D, as we discuss in the end of this section. Constraints (2.18) replace (2.12) in order to only count the dedication time of individual h_i for demand r_{al} if he/she indeed assumes skill k_a in project p_l.

This straightforward adaptation is obtained at the expense of introducing additional variables. However, we claim that we could use the y-variables only. Precisely, under assumption A1, we show next that we can replace (2.17)-(2.19) by

$$\sum_{h_i \in Q_{K'}} \sum_{d_u \in D} d_u y_{ilu} \geq \sum_{k_a \in K'} r_{al}, \quad \forall p_l \in P, \forall K' \subseteq K, \tag{2.20}$$

where $Q_{K'} = \bigcup_{k_a \in K'} Q_a$ is the subset of people having at least one skill in K'.

Let (YY') and (Y) denote formulation (2.15)-(2.19) and formulation (2.15), (2.16) and (2.20), respectively. Before proving the equivalence between the two formulations, we establish some auxiliary results. They will be used to show how to transform a binary point $y = [y_{ilu}]$ satisfying (2.20) into $y' = [y'_{ilua}]$ satisfying (2.17)-(2.19), and vice-versa.

Given a vector $y \in \mathbb{R}_+^{n \times m \times t}$ and $p_l \in P$, let us construct an arc-capacitated network, denoted by $G^l(y) = (V, A, c)$, as illustrated in Figure 1. This network has an arc from a source node s to every individual, an arc from every individual to all skills he/she possesses, and an arc from every skill to a target node t. Precisely, the node set and arc set of $G^l(y)$ respectively are:

$$V = \{s,t\} \cup H \cup K \text{ and } A = (\{s\} \times H) \cup (K \times \{t\}) \cup \{(h_i, k_a) : h_i \in H, k_a \in L_i\}.$$

The capacity of an arc $e \in A$ is

$$c_e = \begin{cases} \sum_{d_u \in D} d_u y_{ilu}, & \text{if } e = (s, h_i), h_i \in H, \\ r_{al}, & \text{if } e = (k_a, t), k_a \in K, \\ +\infty, & \text{if } e = (h_i, k_a), h_i \in H, k_a \in L_i. \end{cases}$$

Given the allocation in project p_l defined by y, the idea is to partition the dedication time of the individuals to their skills by a flow in $G^l(y)$ from s to t. Precisely, given $y = [y_{ilu}]$, which defines the capacities of the arcs entering H, the flow from h_i to k_a will define $y' = [y'_{ilua}]$.

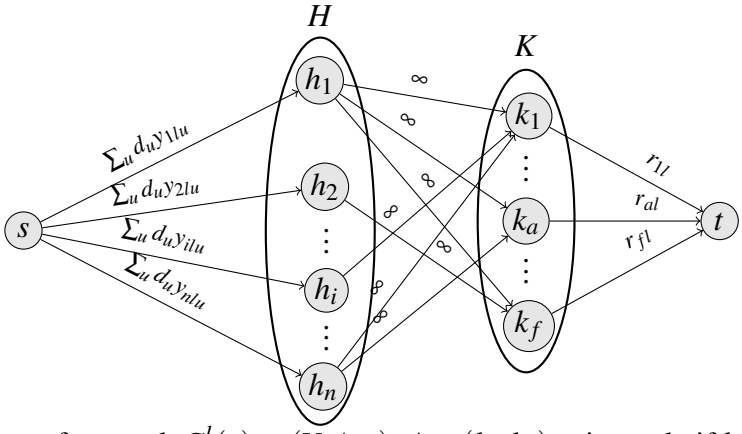


Figure 1 – Illustration of network $G^l(y) = (V, A, c)$. Arc (h_i, k_a) exists only if $k_a \in L_i$ (equivalently $h_i \in Q_a$).

Formally, an st-flow in $G^l(y)$ is a function f on the arcs that satisfies the capacity and flow conservation constraints, i.e. $0 \leq f_e \leq c_e$ for all $e \in A$, and $\sum_{(u,v)\in A} f_{uv} = \sum_{(v,u)\in A} f_{vu}$ for all $v \in V \setminus \{s,t\}$. The value of f, referred to as $|f|$, is the flow leaving node s (or entering node t), i.e. $|f| = \sum_{h_i \in H} f_{sh_i} = \sum_{k_a \in K} f_{k_t}$. Given $S \subset V$ with $s \in S$ and $\bar{S} = V \setminus S$ with $t \in \bar{S}$, the subset of arcs $[S,\bar{S}] := \{(u,v) \in A : u \in S, v \in \bar{S}\}$ is called an st-cut and its capacity is $c[S,\bar{S}] := \sum_{(u,v)\in[S,\bar{S}]} c_{uv}$. In general, for disjoint subsets $S \subset V$ and $S' \subset V$, let $c[S,S'] := \sum_{(u,v)\in[S,S']} c_{uv}$.

A flow of maximum value is simply called maximum flow whereas a cut of minimum capacity is a minimum cut. By the Max-Flow Min-Cut Theorem (FORD; FULKERSON, 1956), it is known that $|f| = c[S,\bar{S}]$, if f and $[S,\bar{S}]$ are a maximum st-flow and a minimum st-cut, respectively.

First, we prove that satisfying constraints (2.20) of formulation (Y) is equivalent to having a flow saturating the arcs that enter the target node t.

Proposição 2.5.1. *Let $y \in \mathbb{R}_+^{n\times m\times t}$, $p_l \in P$ and f be a maximum st-flow in $G^l(y)$. Then, $|f| = \sum_{k_a \in K} r_{al}$ if, and only if, all constraints (2.20) related to p_l are satisfied by y.*

Proof. Due to the capacity and flow conservation constraints, we have that:

$$\sum_{d_u \in D} d_u y_{ilu} \geq f_{sh_i} = \sum_{k_a \in L_i} f_{h_i k_a}, \forall h_i \in H, \tag{2.21}$$

$$r_{al} \geq f_{k_a t} = \sum_{h_i \in Q_a} f_{h_i k_a}, \forall k_a \in K. \tag{2.22}$$

First suppose that $|f| = \sum_{k_a \in K} r_{al}$. Then, $|f| = \sum_{k_a \in K} f_{k_a t}$, and equality must occur in (2.22), that is

$$\sum_{h_i \in Q_a} f_{h_i k_a} = r_{al}, \forall k_a \in K. \tag{2.23}$$

Let $K' \subseteq K$. Adding the inequalities in (2.21) related to $h_i \in Q_{K'}$, we obtain

$$\sum_{h_i \in Q_{K'}} \sum_{d_u \in D} d_u y_{ilu} \geq \sum_{h_i \in Q_{K'}} \sum_{k_a \in L_i} f_{h_i k_a}. \tag{2.24}$$

Note that the right hand side of (2.24) can be bounded as follows:

$$\sum_{h_i \in Q_{K'}} \sum_{k_a \in L_i} f_{h_i k_a} \geq \sum_{h_i \in Q_{K'}} \sum_{k_a \in L_i \cap K'} f_{h_i k_a} = \sum_{k_a \in K'} \sum_{h_i \in Q_a} f_{h_i k_a}. \tag{2.25}$$

By (2.24), (2.25) and (2.23), we conclude that Constraint (2.20) related to p_l and K' is satisfied by y.

Now, assume that $|f| < \sum_{k_a \in K} r_{al} = c[K, \{t\}]$. Let $[S, \bar{S}]$ be a minimum st-cut, $H_s = H \cap S$, $H_t = H \cap \bar{S}$, $K_s = K \cap S$, $K_t = K \cap \bar{S}$. Since $c[S, \bar{S}] = |f| < +\infty$ and every arc in $H \times K$ has infinite capacity, we can conclude that there is no arc from H_s to K_t. Therefore,

$$c[K_s, \{t\}] + c[K_t, \{t\}] = c[K, \{t\}] > |f| = c[S, \bar{S}] = c[\{s\}, H_t] + c[K_s, \{t\}],$$

and so $c[K_t, \{t\}] > c[\{s\}, H_t]$. It follows that

$$\sum_{k_a \in K_t} r_{al} = c[K_t, \{t\}] > c[\{s\}, H_t] \geq \sum_{h_i \in H_t} \sum_{d_u \in D} d_u y_{ilu}.$$

Since there is no arc from H_s to K_t, it must be $Q_{K_t} \subseteq H_t$. Then,

$$\sum_{k_a \in K_t} r_{al} > \sum_{h_i \in Q_{K_t}} \sum_{d_u \in D} d_u y_{ilu},$$

which means that y violates Constraint (2.20) related to p_l and K_t. □

Using the flow provided by Proposition 2.5.1, we can show how to transform a feasible point to the linear relaxation of (Y) into a feasible to the linear relaxation of (YY'), and vice-versa.

Lemma 1. *If* (y, y') *satisfies* (2.17)-(2.18)*, then* y *satisfies* (2.20)*. Conversely, if* $y \geq 0$ *satisfies* (2.20)*, then there is* $y' \geq 0$ *such that* (y, y') *satisfies* (2.17)*, as equality, and* (2.18)*.*

Proof. Suppose that (y, y') satisfies (2.17)-(2.18). Let $p_l \in P$ and $K' \subseteq K$. Summing up Constraints (2.18) related to p_l and all $k_a \in K'$, we get

$$\sum_{k_a \in K'} r_{al} \leq \sum_{k_a \in K'} \sum_{h_i \in Q_a} \sum_{d_u \in D} d_u y'_{ilua} = \sum_{h_i \in Q_{K'}} \sum_{k_a \in L_i \cap K'} \sum_{d_u \in D} d_u y'_{ilua}.$$

Then, upper bounding the rightmost hand side of the above expression and using (2.17), we obtain

$$\sum_{k_a \in K'} r_{al} \leq \sum_{h_i \in Q_{K'}} \sum_{k_a \in L_i} \sum_{d_u \in D} d_u y'_{ilua} \leq \sum_{h_i \in Q_{K'}} \sum_{d_u \in D} d_u y_{ilu}.$$

Therefore, y satisfies (2.20).

Now, suppose that y satisfies (2.20). Let $p_l \in P$. By Proposition 2.5.1, $G^l(y)$ has a maximum flow f such that $|f| = \sum_{k_a \in K} r_{al}$. So, by (2.22), we have

$$r_{al} = f_{k_a t} = \sum_{h_i \in Q_a} f_{h_i k_a}, \quad \forall k_a \in K. \tag{2.26}$$

Moreover, as stated in (2.21),

$$\sum_{k_a \in L_i} f_{h_i k_a} = f_{sh_i} \leq \sum_{d_u \in D} d_u y_{ilu}, \quad \forall h_i \in H. \tag{2.27}$$

Observe that we can split the flow entering and leaving h_i into

$$f_{sh_i} = \sum_{d_u \in D} f_{sh_i d_u}, \tag{2.28}$$

$$f_{h_i k_a} = \sum_{d_u \in D} f_{h_i k_a d_u}, \forall k_a \in L_i, \tag{2.29}$$

such that $f_{sh_i d_u} \geq 0$, $f_{h_i k_a d_u} \geq 0$ and

$$\sum_{k_a \in L_i} f_{h_i k_a d_u} = f_{sh_i d_u} \leq d_u y_{ilu}, \quad \forall d_u \in D. \tag{2.30}$$

Indeed, it is a matter of splitting vertex h_i into $|D|$ vertices (say $h_i d_u$, for all $d_u \in D$) and partitioning the entering and leaving arcs and their capacities accordingly. See Figure 2, where it is illustrated the original flow and the splitted flow.

For every $h_i \in H$, $k_a \in L_i$ and $d_u \in D$, let $\delta_{iu} = d_u y_{ilu} - f_{sh_i d_u} \geq 0$ and $y'_{ilua} = \frac{1}{d_u}(f_{h_i k_a d_u} + \delta_{iu}/|L_i|)$. Then, (2.30) and the definition of δ_{iu} imply

$$\sum_{k_a \in L_i} y'_{ilua} = (\sum_{k_a \in L_i} \frac{f_{h_i k_a d_u}}{d_u}) + \frac{\delta_{iu}}{d_u} = \frac{f_{sh_i d_u}}{d_u} + y_{ilu} - \frac{f_{sh_i d_u}}{d_u} = y_{ilu}, \tag{2.31}$$

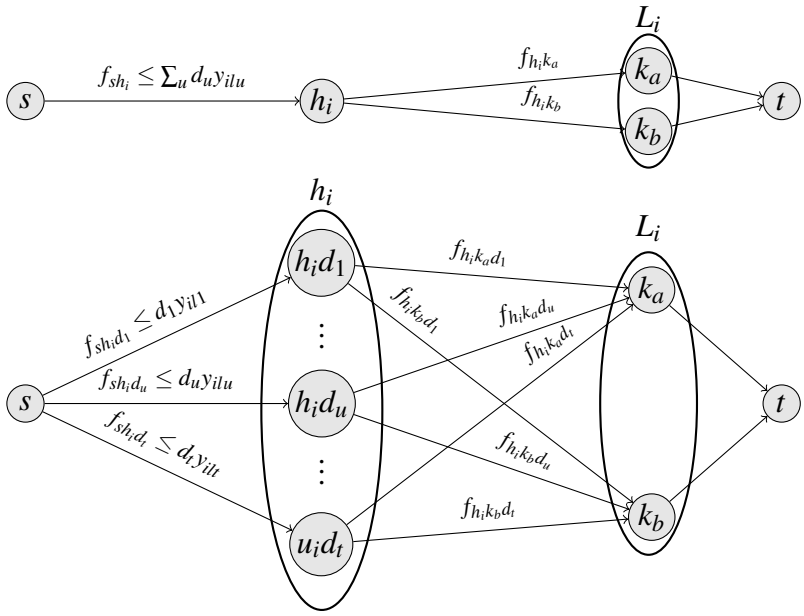


Figure 2 – Splitting the flow at h_i.

for all $h_i \in H$ and $d_u \in D$. Besides, for all $k_a \in K$,

$$\sum_{h_i \in Q_a} \sum_{d_u \in D} d_u y'_{ilua} = \sum_{h_i \in Q_a} \sum_{d_u \in D} (f_{h_i k_a d_u} + \delta_{iu}/|L_i|) \geq \sum_{h_i \in Q_a} f_{h_i k_a} = r_{al},$$

where the inequality is due to (2.29) and $\delta_{iu} \geq 0$, and the last equality comes from (2.26). It follows that (y, y') satisfies (2.17), as equality, and (2.18). □

We use Proposition 2.5.1 and Lemma 1 to compare formulations (Y) and (YY') and their linear relaxations. Given a set C comprising points of the form $(x,y) \in \mathbb{R}^p \times \mathbb{R}^q$, let $\text{proj}_x(C)$ be the set of points $x \in \mathbb{R}^p$ such that $(x,y) \in C$ for some $y \in \mathbb{R}^q$. For any ILP formulation X, let $\mathscr{F}(X)$ and $\overline{\mathscr{F}}(X)$ be the feasible sets of X and its linear relaxation, respectively. In addition, let $\mathscr{F}^*(X)$ be the optimal set of X.

Proposição 2.5.2. $\overline{\mathscr{F}}(Y) = \text{proj}_y(\overline{\mathscr{F}}(YY'))$. *Moreover, under assumptions A1-A2,* $\mathscr{F}(Y) = \text{proj}_y(\mathscr{F}(YY'))$ *and* $\mathscr{F}^*(Y) = \text{proj}_y(\mathscr{F}^*(YY'))$.

Proof. The inclusions $\text{proj}_y(\overline{\mathscr{F}}(YY')) \subseteq \overline{\mathscr{F}}(Y)$ and $\text{proj}_y(\mathscr{F}(YY')) \subseteq \mathscr{F}(Y)$ are direct consequences of Lemma 1. Now, let $(y,z) \in \overline{\mathscr{F}}(Y)$. By Lemma 1, there is $y' \geq 0$ such that (y,y',z) satisfies (2.15)-(2.18). Besides, by (2.31), we can deduce that $0 \leq y'_{ilua} \leq y_{ilu} \leq 1$, for all $h_i \in H$, $p_l \in P$, $d_u \in D$ and $k_a \in K$. Therefore, $(y,y',z) \in \overline{\mathscr{F}}(YY')$.

Additionally, assume A1-A2, that is, suppose there is $\alpha > 0$ such that $d_u = (u-1)\alpha$, for all $d_u \in D'$, and r_{al} is a multiple of α, for all $p_l \in P$ and $k_a \in K$. Let $(y,z) \in \mathscr{F}(Y)$. We want to find y' such that $(y,y',z) \in \mathscr{F}(YY')$. Let $p_l \in P$ and consider the network $G^l(y)$. By

Proposition 2.5.1, $G^l(y)$ has a maximum flow f such that $|f| = \sum_{k_a \in K} r_{al}$. Let us also consider the auxiliary network $\hat{G}^l(y) = (V, A, \hat{c})$ with the same nodes and arcs as $G^l(y)$ but with arc capacities $\hat{c}_e = c_e/\alpha$, for all $e \in A$. By the assumption on the d_u's and r_{al}'s and the fact that $y \in \mathbb{B}^{n\times m\times t}$, we have that these scaled capacities are now integers. Then, $\hat{G}^l(y)$ has an optimal integer flow $\hat{f}$, and we can assume that $f = \alpha\hat{f}$. Let us recall that f satisfies (2.26)-(2.30). As a counterpart of (2.30), we get

$$\sum_{k_a \in L_i} \hat{f}_{h_i k_a d_u} = \hat{f}_{sh_i d_u} \leq (u-1) y_{ilu}, \quad \forall d_u \in D \; \forall h_i \in H, \tag{2.32}$$

where $\hat{f}_{sh_i d_u} = f_{sh_i d_u}/\alpha \geq 0$ and $\hat{f}_{h_i k_a d_u} = f_{h_i k_a d_u}/\alpha \geq 0$. Let $h_i \in H$, $k_a \in L_i$ and $d_u \in D$. Then, (2.32) together with the integrality of y and $\hat{f}$ imply that $\hat{f}_{sh_i d_u}, \hat{f}_{h_i k_a d_u} \in \{0, \ldots, u-1\}$. It means that $f_{sh_i d_u}, f_{h_i k_a d_u} \in D'$ with $f_{sh_i d_u} \leq d_u$. Let $\delta_{ia} = \sum_{d_u \in D} f_{h_i k_a d_u}$. Using (2.29), (2.27) and (2.11), we obtain

$$\delta_{ia} = \sum_{d_u \in D} f_{h_i k_a d_u} = f_{h_i k_a} \leq \sum_{d_u \in D} d_u y_{ilu} \leq 1.$$

Therefore, δ_{ia} is a summation of elements of D' not exceeding 1, and so $\delta_{ia} \in D'$.

Let us define y' as follows. For every $h_i \in H$, $k_a \in L_i$ and $d_w \in D$, we set $y'_{ilwa} = 1$, if $d_w = \delta_{ia}$, and $y'_{ilua} = 0$, otherwise. Then, for all $h_i \in H$, we have

$$\sum_{k_a \in L_i} \sum_{d_w \in D} d_w y'_{ilwa} = \sum_{k_a \in L_i} \delta_{ia} = \sum_{d_u \in D} \sum_{k_a \in L_i} f_{h_i k_a d_u} \leq \sum_{d_u \in D} d_u y_{ilu}, \tag{2.33}$$

where the inequality is due to (2.30). In addition, for all $k_a \in K$, we obtain

$$\sum_{h_i \in Q_a} \sum_{d_w \in D} d_w y'_{ilwa} = \sum_{h_i \in Q_a} \delta_{ia} = \sum_{h_i \in Q_a} \sum_{d_u \in D} f_{h_i k_a d_u} = \sum_{h_i \in Q_a} f_{h_i k_a} = r_{al}, \tag{2.34}$$

where the two last equalities are due (2.29) and (2.26), respectively. Therefore, (y, y') satisfies all constraints (2.17)-(2.18). Besides, since $(y, z) \in \mathscr{F}(Y)$ and y' is binary, we conclude that $(y, y', z) \in \mathscr{F}(YY')$.

Once we have proved $\mathscr{F}(Y) = \text{proj}_y(\mathscr{F}(YY'))$ and the formulations have the same objective function, we also get $\mathscr{F}^*(Y) = \text{proj}_y(\mathscr{F}^*(YY'))$ □

Proposition 2.5.2 states that formulations (Y) and (YY') are equivalent under assumptions A1-A2. Particularly when $D' = \{0, 1\}$ (full-time allocation), (YY') becomes the formulation proposed by Campêlo *et al.* (2020), and so (Y) is an alternative for such a formulation with fewer variables. In general, (Y) is a relaxation of (YY'). Actually, (YY') models a variant of MMTFP

where an individual can be assigned to multiple skills while satisfying the following condition: the fraction of time dedicated to each project must lie in D but it can be divided arbitrarily among the skills assumed by the individual in the project.

As shown next, the above results allows us to deduce that equality can be used in Constraints (2.17) or (2.18), if A1-A2 are assumed. In any case, the modified formulation (YY') and formulation (Y) still have the same optimal value.

Proposição 2.5.3. *Consider the assumptions A1-A2. With equality in* (2.18)*, we have*

$$\overline{\mathscr{F}}(Y) \supseteq \text{proj}_y(\overline{\mathscr{F}}(YY')),\ \mathscr{F}(Y) = \text{proj}_y(\mathscr{F}(YY'))\ \textit{and}\ \mathscr{F}^*(Y) = \text{proj}_y(\mathscr{F}^*(YY')).$$

With equality (2.17)*, we have*

$$\overline{\mathscr{F}}(Y) = \text{proj}_y(\overline{\mathscr{F}}(YY')),\ \mathscr{F}(Y) \supseteq \text{proj}_y(\mathscr{F}(YY'))\ \textit{and}\ \mathscr{F}^*(Y) = \text{proj}_y(\mathscr{F}^*(YY')).$$

Proof. Similarly to Remark 1, we can deduce that using equality in (2.18) keeps formulation (YY') correct, although with a smaller feasible set. Then, Proposition 2.5.2 implies $\text{proj}_y(\overline{\mathscr{F}}(YY')) \subseteq \overline{\mathscr{F}}(Y)$ and $\text{proj}_y(\mathscr{F}(YY')) \subseteq \mathscr{F}(Y)$. Besides, since equality holds in (2.34), we still get $\mathscr{F}(Y) \subseteq \text{proj}_y(\mathscr{F}(YY'))$. Therefore, $\mathscr{F}(Y) = \text{proj}_y(\mathscr{F}(YY'))$ and so $\mathscr{F}^*(Y) = \text{proj}_y(\mathscr{F}^*(YY'))$.

Now consider (YY') with equality in Constraints (2.17). Because Lemma 1 guarantees equality in (2.17), we get $\overline{\mathscr{F}}(Y) = \text{proj}_y(\overline{\mathscr{F}}(YY'))$. The first part of this lemma also implies $\text{proj}_y(\mathscr{F}(YY')) \subseteq \mathscr{F}(Y)$ and $\text{proj}_y(\mathscr{F}^*(YY')) \subseteq \mathscr{F}^*(Y)$. Finally, to show $\mathscr{F}^*(Y) \subseteq \text{proj}_y(\mathscr{F}^*(YY'))$, we can slightly modify the second part of the proof of Proposition 2.5.2 to obtain equality in (2.33). For, it is enough to take $(y,z) \in \mathscr{F}^*(Y)$ such that the maximum flow in $G^l(y)$, for each $p_l \in P$, satisfies $f_{sh_i} = \sum_{d_u \in D} d_u y_{ilu}$ for all $h_i \in H$. This is always possible since f_{sh_i} is a multiple of α. Thus, if $(y,z) \in \mathscr{F}^*(Y)$ does not have this property, we can change y by $\tilde{y}$ such that $\tilde{y}_{ilu} = 1$ only for the dedication time fraction $d_u = f_{sh_i}$ and change z accordingly. Such an $\tilde{y}$ yields the same maximum flow f and defines another optimum solution to (Y) (recall that the objective function coefficients are positive). □

Remark 2. *Under assumptions A1-A2, Proposition 2.5.3 implies that the optimum value of* (YY') *does not change if equality is used in* (2.17)-(2.18)*. Then, summing up* (2.18) *for* $k_a \in K_l$ *and using* (2.17) *lead to*

$$\sum_{k_a \in K_l} r_{al} = \sum_{k_a \in K_l} \sum_{h_i \in Q_a} \sum_{d_u \in D} d_u y'_{ilua} = \sum_{h_i \in H_l} \sum_{k_a \in L_i \cap K_l} \sum_{d_u \in D} d_u y'_{ilua} = \sum_{h_i \in H_l} \sum_{d_u \in D} d_u y_{ilu}.$$

It means that equality holds in (2.20) *for* $K' = K_l$.

It is important to mention that model (Y) presents an exponential number of constraints, since there is an inequality in (2.20) for every subset $K' \subseteq K$. However, Proposition 2.5.1 ensures that these constraints can be separated in polynomial time.

Corolário 2.5.1. *Let $y \in \mathbb{Q}_+^{n\times m\times t}$. Checking whether y satisfies all Constraints in* (2.20) *or determining one of these constraints violated by y can be done in polynomial time.*

Proof. Since y is rational, network $G^l(y)$ has rational capacities for every $p_l \in P$. Then, a maximum flow in each of these m networks can be found in polynomial time. By the proof of Proposition 2.5.1, with this flow in hand we can identify a violated constraint or conclude that all of them are satisfied. □

2.6 Valid inequalities

In this section, we present some valid inequalities for MTFP (and MMTFP). First of all, let us recall that the following inequalities could be added to the formulations while keeping the optimal value:

$$z_{ijluv} \leq y_{ilu} \qquad \forall p_l \in P, \forall h_i, h_j \in H_l, i \neq j, \forall d_u, d_v \in D, \tag{2.35}$$

$$z_{ijluv} \leq y_{jlv} \qquad \forall p_l \in P, \forall h_i, h_j \in H_l, i \neq j, \forall d_u, d_v \in D, \tag{2.36}$$

$$\sum_{d_v \in D} y_{jlv} \leq 1 \qquad \forall p_l \in P, \forall h_j \in H_l, \tag{2.37}$$

where $K_l = \{k_a \in K : r_{al} > 0\}$ and $H_l = \{h_i \in H : k_a \in K_l, h_i \in Q_a\}$. Thus, using (2.10), the integer solutions satisfy:

$$z_{ijluv} = y_{ilu} y_{jlv} \qquad \forall p_l \in P, \forall h_i, h_j \in H_l, i \neq j, \forall d_u, d_v \in D, \tag{2.38}$$

$$0 = y_{ilu} y_{ilv} \qquad \forall p_l \in P, \forall h_i \in H_l, \forall d_u, d_v \in D, u \neq v. \tag{2.39}$$

Based on these equations, we use the reformulation-linearization technique - RLT (ADAMS; SHERALI, 1986; SHERALI; ADAMS, 1990) to derive valid inequalities. Basically, we multiply a valid linear inequality by x or $(1-x)$, where x is a binary expression, and linearize the obtained quadratic inequality. It is a well-known way for tightening the relaxation. First, we use (2.37) as base inequalities.

Proposição 2.6.1. *Let $p_l \in P$, $h_i, h_j \in H_l$, with $i \neq j$, and $d_u \in D$. The following inequalities are*

valid for MTFP:

$$\sum_{d_v \in D} z_{ijluv} \leq y_{ilu}, \tag{2.40}$$

$$\sum_{d_v \in D} z_{ijluv} \geq y_{ilu} + \sum_{d_v \in D} y_{jlv} - 1. \tag{2.41}$$

Proof. Inequalities (2.40) and (2.41) are obtained by multiplying (2.37) by y_{ilu} and $(1 - y_{ilu})$, respectively, and then using (2.38). □

Instead of multiplying the base inequality by y_{ilu} or $(1 - y_{ilu})$, we can consider its product by $\sum_{d_u} y_{ilu}$ or $(1 - \sum_{d_u} y_{ilu})$. In the first case, we get inequalities dominated by (2.40). In the second one, we obtain inequalities stronger than (2.41), as follows.

Proposição 2.6.2. *Let $p_l \in P$, and $h_i, h_j \in H_l$ with $i \neq j$. The following inequalities valid for MTFP:*

$$\sum_{d_u \in D} \sum_{d_v \in D} z_{ijluv} \geq \sum_{d_u \in D} y_{ilu} + \sum_{d_v \in D} y_{jlv} - 1. \tag{2.42}$$

Proof. The product of (2.37) by $(1 - \sum_{d_u} y_{ilu}) \geq 0$ together with (2.38) lead to (2.42). □

Let us remark that the summation of (2.41) and $-z_{ijluw} \geq -y_{jlw}$, for all $w \neq v$, is exactly (2.10). Therefore, these constraints of the formulation are dominated by (2.41) and (2.35)-(2.36). By its turn, any inequality (2.41) is dominated by (2.40) and (2.42). Indeed, it is the summation of (2.42) and $-\sum_{d_v \in D} z_{ijlwv} \geq -y_{ilw}$, for all $w \neq u$. In addition, (2.40) clearly dominates (2.35)-(2.36).

Next we derive the counterparts of the two above propositions where (2.12) are now taken as base inequalities.

Proposição 2.6.3. *Let $p_l \in P$, $k_a \in K_l$, $h_i \in H_l$ and $d_u \in D$. If $h_i \notin Q_a$, the following inequalities are valid for MTFP:*

$$\sum_{h_j \in Q_a} \sum_{d_v \in D} d_v z_{ijluv} \geq r_{al} y_{ilu}, \tag{2.43}$$

$$\sum_{h_j \in Q_a} \sum_{d_v \in D} d_v z_{ijluv} \leq r_{al} y_{ilu} + \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v y_{jlv} - r_{al}. \tag{2.44}$$

If $h_i \in Q_a$, the following inequalities are valid for MTFP:

$$\sum_{\substack{h_j \in Q_a \\ j \neq i}} \sum_{d_v \in D} d_v z_{ijluv} \geq (r_{al} - d_u) y_{ilu}, \tag{2.45}$$

$$\sum_{\substack{h_j \in Q_a \\ j \neq i}} \sum_{d_v \in D} d_v z_{ijluv} \leq (r_{al} - d_u) y_{ilu} + \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v y_{jlv} - r_{al}. \tag{2.46}$$

Proof. By (2.12), $\sum_{h_j \in Q_a} \sum_{d_v \in D} d_v y_{jlv} \geq r_{al}$ is valid. If $h_i \notin Q_a$, the product of this inequality by y_{ilu} yields (2.43), after using (2.38). Similarly we get (2.44) by multiplying such an inequality by $(1 - y_{ilu})$. If $h_i \in Q_a$, the product of the same inequality by y_{ilu} can now be written as $\sum_{d_v \in D} d_v y_{ilv} y_{ilu} + \sum_{\substack{h_j \in Q_a \\ j \neq i}} \sum_{d_v \in D} d_v y_{jlv} y_{ilu} \geq r_{al} y_{ilu}$. By applying (2.39) and (2.38) to the first and second terms of the left hand side and noting that $y_{ilu} y_{ilu} = y_{ilu}$, we obtain (2.45). Analogously, the product by $(1 - y_{ilu})$ leads to (2.46). □

Proposição 2.6.4. *Let $p_l \in P$, $k_a \in K_l$, $h_i \in H_l$. If $h_i \notin Q_a$, the following inequality is valid for MTFP:*

$$\sum_{d_u \in D} \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v z_{ijluv} \leq \sum_{d_u \in D} r_{al} y_{ilu} + \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v y_{jlv} - r_{al}. \tag{2.47}$$

If $h_i \in Q_a$, the following inequality is valid for MTFP:

$$\sum_{d_u \in D} \sum_{\substack{h_j \in Q_a \\ j \neq i}} \sum_{d_v \in D} d_v z_{ijluv} \leq \sum_{d_u \in D} (r_{al} - d_u) y_{ilu} + \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v y_{jlv} - r_{al}. \tag{2.48}$$

Proof. Let us multiply $\sum_{d_u} y_{ilu} \leq 1$ by $\Delta := \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v y_{jlv} - r_{al} \geq 0$ to get $\sum_{d_u \in D} \Delta y_{ilu} \leq \Delta$, or still $\sum_{d_u \in D} \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v y_{jlv} y_{ilu} \leq \Delta + \sum_{d_u \in D} r_{al} y_{ilu}$. If $h_i \notin Q_a$, we can use equations (2.38) to obtain exactly (2.47). If $h_i \in Q_a$, again (2.38) yields $\sum_{d_u \in D} \sum_{d_v \in D} d_v y_{ilv} y_{ilu} + \sum_{d_u \in D} \sum_{\substack{h_j \in Q_a \\ j \neq i}} \sum_{d_v \in D} d_v z_{ijluv} \leq \Delta + \sum_{d_u \in D} r_{al} y_{ilu}$. Using (2.37), we can conclude that the leftmost hand side is equal to $\sum_{d_u} d_u y_{ilu}$, thus leading to (2.48). □

It is worth observing that, if equality is used in (2.12), then equality also holds in (2.43)-(2.48). Moreover, in this case, inequalities (2.43) are equivalent to (2.44) (since $\sum_{h_j \in Q_a} \sum_{d_v \in D} d_v y_{jlv} = r_{al}$), and they dominate (2.47), which are the summation of them in $d_u \in D$. A similar comment can be made for inequalities (2.45), (2.46) and (2.48), respectively.

Therefore, when using equality in (2.12), it would be preferable to apply (2.44) and (2.46) as cuts (instead of (2.47) and (2.48)). However, the number of those inequalities can make their computational use prohibitive when solving large instances. It can be the case of

aggregating them somehow while trying to keep their strength. The following result appears in this setting. It has shown to be useful in our computational experiments, to be presented in Section 2.8.

Proposição 2.6.5. *Assume A1-A2. Let $p_l \in P$, $k_a \in K_l$, $h_i \in H_l$ and $d_w \in D$. If $h_i \notin Q_a$, the following inequality is valid for MTFP:*

$$\sum_{\substack{d_u \in D \\ u \neq w}} \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v z_{ijluv} \leq r_{al}(1 - y_{ilw}). \tag{2.49}$$

If $h_i \in Q_a$, the following inequality is valid for MTFP:

$$\sum_{\substack{d_u \in D \\ u \neq w}} \sum_{\substack{h_j \in Q_a \\ j \neq i}} \sum_{d_v \in D} d_v z_{ijluv} + \sum_{\substack{d_u \in D \\ u \neq w}} d_u y_{ilu} \leq r_{al}(1 - y_{ilw}). \tag{2.50}$$

Proof. Under A1-A2, equality can be used in (2.12), according to Remark 1. Then, we get equality in (2.43). Summing up these equations in $d_u \in D$ leads to

$$\sum_{\substack{d_u \in D \\ u \neq w}} \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v z_{ijluv} + \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v y_{ilw} y_{jlv} = \sum_{\substack{d_u \in D \\ u \neq w}} r_{al} y_{ilu} + r_{al} y_{ilw}.$$

Again equality in (2.12) simplifies this expression to

$$\sum_{\substack{d_u \in D \\ u \neq w}} \sum_{h_j \in Q_a} \sum_{d_v \in D} d_v z_{ijluv} = \sum_{\substack{d_u \in D \\ u \neq w}} r_{al} y_{ilu}.$$

Since $\sum_{\substack{d_u \in D \\ u \neq w}} y_{ilu} \leq 1 - y_{ilw}$ due to (2.37), we get (2.49). Similarly, we obtain (2.50). □

In the multi-skill case, we can also use (2.20) as base inequalities. Similarly to the above three propositions, we can derive the following valid inequalities.

Proposição 2.6.6. *Let $p_l \in P$, $K' \subseteq K_l$, $h_i \in H_l$ and $d_u \in D$. If $h_i \notin Q_{K'}$, the following inequalities are valid for MMTFP:*

$$\sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v z_{ijluv} \geq \sum_{k_a \in K'} r_{al} y_{ilu}, \tag{2.51}$$

$$\sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v z_{ijluv} \leq \sum_{k_a \in K'} r_{al} y_{ilu} + \sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v y_{jlv} - \sum_{k_a \in K'} r_{al}. \tag{2.52}$$

If $h_i \in Q_{K'}$, the following inequalities are valid for MMTFP:

$$\sum_{\substack{h_j \in Q_{K'} \\ j \neq i}} \sum_{d_v \in D} d_v z_{ijluv} \geq \left(\sum_{k_a \in K'} r_{al} - d_u \right) y_{ilu}, \tag{2.53}$$

$$\sum_{\substack{h_j \in Q_{K'} \\ j \neq i}} \sum_{d_v \in D} d_v z_{ijluv} \leq \left(\sum_{k_a \in K'} r_{al} - d_u \right) y_{ilu} + \sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v y_{jlv} - \sum_{k_a \in K'} r_{al}. \tag{2.54}$$

Proof. By (2.20), we have $\sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v y_{jlv} \geq \sum_{k_a \in K'} r_{al}$. The product of this inequality by y_{ilu} yields $\sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v y_{jlv} y_{ilu} \geq \sum_{k_a \in K'} r_{al} y_{ilu}$. Then, using (2.38) we directly get (2.51), if $h_i \notin Q_{K'}$, or $\sum_{\substack{h_j \in Q_{K'} \\ j \neq i}} \sum_{d_v \in D} d_v z_{ijluv} + \sum_{d_v \in D} d_v y_{ilv} y_{ilu} \geq \sum_{k_a \in K'} r_{al} y_{ilu}$, if $h_i \in Q_{K'}$. In the second case, since $y_{ilu} y_{ilu} = y_{ilu}$ and $y_{ilu} y_{ilv} = 0$ for $u \neq v$, we get (2.53). Similarly we get (2.52) and (2.54) by multiplying (2.20) by $(1 - y_{ilu})$. □

Proposição 2.6.7. *Let $p_l \in P$, $K' \subseteq K_l$, $h_i \in H_l$. If $h_i \notin Q_{K'}$, the following inequality is valid for MMTFP:*

$$\sum_{d_u \in D} \sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v z_{ijluv} \leq \sum_{d_u \in D} \sum_{k_a \in K'} r_{al} y_{ilu} + \sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v y_{jlv} - \sum_{k_a \in K'} r_{al}. \tag{2.55}$$

If $h_i \in Q_{K'}$, the following inequality is valid for MMTFP:

$$\sum_{d_u \in D} \sum_{\substack{h_j \in Q_{K'} \\ j \neq i}} \sum_{d_v \in D} d_v z_{ijluv} \leq \sum_{d_u \in D} \left(\sum_{k_a \in K'} r_{al} - d_u \right) y_{ilu} + \sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v y_{jlv} - \sum_{k_a \in K'} r_{al}. \tag{2.56}$$

Proof. Let $\Delta := \sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v y_{jlv} - \sum_{k_a \in K'} r_{al}$ so that $\Delta \geq 0$ is valid by (2.20). Its product by $1 - \sum_{d_u \in D} y_{ilu}$ is $\sum_{d_u \in D} \sum_{h_j \in Q_{K'}} \sum_{d_v \in D} d_v y_{jlv} y_{ilu} \leq \Delta + \sum_{d_u \in D} \sum_{k_a \in K'} r_{al} y_{ilu}$. If $h_i \notin Q_{K'}$, we can use (2.38) to get exactly (2.55). If $h_i \in Q_{K'}$, again (2.38) together with (2.37) and $y_{ilu} y_{ilu} = y_{ilu}$ yield (2.56). □

Observe that Propositions 2.6.3 and 2.6.4 are particular cases of Propositions 2.6.6 and 2.6.7 when $K' = \{k_a\}$. Regarding Proposition 2.6.5, its counterpart for MMTFP needs (2.20) to be satisfied at equality, which can be considered when $K' = K_l$ (see Remark 2). Thus, we can show the following result.

Proposição 2.6.8. *Assume A1-A2. Let $p_l \in P$, $h_i \in H_l$, and $d_w \in D$. The following inequality is valid for MMTFP:*

$$\sum_{\substack{d_u \in D \\ u \neq w}} \sum_{\substack{h_j \in H_l \\ j \neq i}} \sum_{d_v \in D} d_v z_{ijluv} + \sum_{\substack{d_u \in D \\ u \neq w}} d_u y_{ilu} \leq \sum_{k_a \in K_l} r_{al} (1 - y_{ilw}). \tag{2.57}$$

Proof. Under A1-A2, similarly to Remark 1, we can conclude that equality can be used in (2.20) whenever $K' = K_l$. Thus, let $K' = K_l$ and so $Q_{K'} = H_l$. As the two rightmost terms in (2.54) vanish, equality holds in (2.53). The summation of the obtained equations in $d_u \in D$ together

with (2.38) and then (2.39) lead to

$$\begin{aligned}
&\sum_{\substack{d_u\in D\\ u\neq w}}\sum_{\substack{h_j\in H_l\\ j\neq i}}\sum_{d_v\in D} d_v z_{ijluv} + \sum_{\substack{d_u\in D\\ u\neq w}} d_u y_{ilu} \\
&= \sum_{k_a\in K_l} r_{al} y_{ilw} - \sum_{\substack{h_j\in H_l\\ j\neq i}}\sum_{d_v\in D} d_v y_{ilw} y_{jlv} - d_w y_{ilw} + \sum_{\substack{d_u\in D\\ u\neq w}}\sum_{k_a\in K_l} r_{al} y_{ilu} \\
&= \sum_{k_a\in K_l} r_{al} y_{ilw} - \sum_{\substack{h_j\in H_l\\ j\neq i}}\sum_{d_v\in D} d_v y_{jlv} y_{ilw} - \sum_{d_v\in D} d_v y_{ilv} y_{ilw} + \sum_{\substack{d_u\in D\\ u\neq w}}\sum_{k_a\in K_l} r_{al} y_{ilu}.
\end{aligned}$$

Again equality in (2.20) implies $\sum_{h_j\in H_l}\sum_{d_v\in D} d_v y_{jlv} = \sum_{k_a\in K_l} r_{al}$, which simplifies the righthand side of above expression to $\sum_{k_a\in K_l} r_{al}\sum_{\substack{d_u\in D\\ u\neq w}} y_{ilu}$. Therefore, using $\sum_{\substack{d_u\in D\\ u\neq w}} y_{ilu} \leq 1 - y_{ilw}$ due to (2.37), we get (2.57). □

2.7 MTFP Instances

No real-world dataset is available for MTFP in the literature. On the other hand, a synthetic dataset was generated by Gutiérrez *et al.* (2016) to use in their experiments. The authors organize the instances in three groups, each one related to a fixed percentage p of positive relationships in the sociometry matrix S. In each group, the instances are divided into 9 classes, according to the values chosen for the following parameters: number of projects (m), number of available individuals (n), number of skills (f), minimum time allocation fraction allowed (α). See Table 2. Note that these instances satisfy A1-A2.

Table 2 – Parameter values for each instance class.

Class	Projects (m)	Individuals (n)	Skills (f)	Time allocation fractions (D')
1	2	25	10	0.0-1.0
2	5	50	5	0.0-1.0
3	10	100	10	0.0-1.0
4	2	25	10	0.0-0.5-1.0
5	5	50	5	0.0-0.5-1.0
6	10	100	10	0.0-0.5-1.0
7	2	25	10	0.0-0.25-0.5-0.75-1.0
8	5	50	5	0.0-0.25-0.5-0.75-1.0
9	10	100	10	0.0-0.25-0.5-0.75-1.0

Note. Reprinted from Gutiérrez *et al.* (2016).

As the dataset is no longer accessible, we have used the same parameters to generate similar instances. Recall that an instance is given by a tuple (P,H,K,Q,D,R,S,W) or still $(m=|P|,n=|H|,f=|K|,Q,1/\alpha=|D|,R,S,W)$. As no detail on the distribution of skills among individuals (Q), the project requirements matrix (R), and the list of project priorities (W) for each instance were given by Gutiérrez *et al.* (2016), we adopted equal priorities ($1/m$), and

each pair (matrix R, list Q) was built as described next, while taking care that the number of people required per project and skill is within the limit of availability.

For each class, three requirements matrices R were generated with increasing values $\sum_{k_a} \sum_{p_l} r_{al}$, corresponding to a low, medium and high requirement of skilled individuals for the projects. Thus, in total, we have 27 distinct requirements matrices which cover different application scenarios by also varying the density and range of non-null entries. It refers to different levels of homogeneity/heterogeneity among projects regarding number and types of required skills as well as amount of time fractions per skill. In every requirement matrix R, the total dedication time per project (sum of a column) is less than n/m in order to guarantee feasibility.

Once matrix R is generated, the assignment of each individual's skill is as follows. The minimum number of needed individuals with each skill (sum of a row of R) is calculated, and they are randomly chosen within the set of available individuals. In the end, if there is still some individual left, his skill is randomly chosen from the set of skills.

For each matrix R, two lists Q were generated in this way. Thus, we have six pairs (R, Q) for each class. A correlated pair (R, Q) along with a matrix S and the other already specified parameters define an MTFP instance. Details on the generation of matrices S are presented below.

2.7.1 Generation of sociometric matrices

The sociometric matrices reveal the "social atoms" formed by each individual and his/her social network, which can be apprehended by the choices/ rejections made or received in a sociometric test (MORENO, 1941). Since the paper by Forsyth e Katz (1946), the majority of works on the analysis of group structure involves the matrix representation of sociometric data. Although a number of techniques are used in the sociometric research, by far the most common way to rank personal relationships is to use simple ordinal rating scales. Besides making it easier for a group member to assess his/her teammates, limited scales usually produce satisfactory quality of data (ROISTACHER, 1974).

Considering the standards used in the literature to create sociometric matrices, we generated three datasets that are available at https://github.com/figueiredoft/MTFP.

Synthetic signal sociometric matrices

Given the number of individuals (n), three $n \times n$ symmetric $\{-1,0,+1\}$-matrices S were created as suggested by Gutiérrez *et al.* (2016), each one related to a percentage p of positive $(+1)$ relationships. Precisely, we used $p = 30\%$, 50% and 70% to define the instances related to groups I, II and III, respectively. After randomly choosing the $+1$'s entries of S, the remaining entries are set to 0 (neutral) or -1 (negative) with probability 85% and 15%. As $n \in \{25,50,100\}$ and $p \in \{0.3,0.5,0.7\}$, there are nine distinct sociometric matrices of this type. With them, we get six instances for each combination of group and class. The i-th instance is related to the i-th pair (R,Q) of the class and the matrix S corresponding to the group and number of individuals in the class. Since there are nine classes and three groups, it results in 162 MTFP instances of this type, generated similarly to those proposed by Gutiérrez *et al.* (2016).

Real-world signal sociometric matrices

To create $\{-1,0,+1\}$ matrices based on real-world networks, we used a dataset from Epinions social network where links are classified as positive or negative. A signed link indicates whether a member of the network "trusts" or "does not trust" the other. The dataset is available at https://snap.stanford.edu and has 131,828 vertices and 841,372 (directed) edges. From this dataset, three $n \times n$ sociometric matrices were generated, for each $n \in \{25,50,100\}$, as described below. Replacing the nine synthetic signal matrices by these nine real-world matrices, we obtained other 162 MTFP instances in the manner.

Each matrix created is basically the adjacency matrix of the network induced by a subset of n vertices selected by a heuristic algorithm. Since the density of the Epinions network is low, the algorithm seeks to maximize the number of edges between the chosen individuals. First, it selects beforehand the set H_0 of individuals who gave at least a certain number of indications to the other members of the network. Then, those n individuals with the greatest number of indications given or received within H_0 are chosen. From this initial subset, a local search is carried out. It randomly selects one individual in the subset and another one out of the subset (possibly outside H_0), and exchanges them if the number of edges increases. This step is repeated 5,000 times. The final subset of individuals defines a symmetric sociometric matrix S: for two selected individuals h_i and h_j, $s_{ij} = 0$ if there is no link between them, $s_{ij} = +1$ if all links between them are positive, and $s_{ij} = -1$ otherwise. Once an $n \times n$ matrix has been created, the corresponding individuals are discarded from H_0 in order to define the next matrix of the same order. This allows us to get unrelated matrices with different densities.

Table 3 shows some information on the generated matrices, particularly density (percentage of non-null off-diagonal entries) and percentages of the non-null entries which are positive and negative. Unlike the synthetic matrices, neither the number of positive entries nor negative entries is controlled. So, the instances are no longer organized in groups I, II and III. Now, we arrange them in 3 sets (H,M,L), associated with the highest, medium and lowest density. The i-th instance of a pair class-set is given by the i-th pair (R,Q) of the class and the matrix S of the set with size defined by the class.

Table 3 – Signal matrices generated with Epinions social network dataset.

Set	Size (n)	Density (%) (a)	Negative entries (%) (b)	Positive entries (%) (c)	$\rho = a\frac{c-b}{100}$
L	25	36.00	12.04	87.96	27.33
M	25	47.33	2.11	97.89	45.33
H	25	69.67	27.75	72.25	31.00
L	50	26.69	17.74	82.26	17.22
M	50	33.22	12.28	87.72	25.06
H	50	74.45	8.22	91.18	61.76
L	100	22.48	18.51	81.49	14.16
M	100	42.12	11.85	88.15	32.14
H	100	61.86	17.41	82.59	40.32

Note. The Epinions network has density equals 0.08%.

Weighted sociometric matrices

In order to analyze the performance of the proposed formulations with a larger scale of values to rate relationships, we used another real-world dataset, also available at https://snap.stanford.edu. It is associated with the BitCoin OTC social network, which has 5,881 vertices and 35,592 (directed) edges. In this network, each member rates other members in a scale of -10 (total distrust) to +10 (total trust) in steps of 1. It is the first explicit weighted signed directed network available for research. From this dataset, for each $n \in \{25, 50, 100\}$, other three $n \times n$ sociometric matrices were generated by applying the same heuristic used with the Epinions network and using the weights $w_{ij} \in [-10, 10]$ of the links to define the sociometric matrix entries $s_{ij} = s_{ji} = \frac{w_{ij}+w_{ji}}{2}$. Similarly, we used these nine matrices to replace the signal matrices and obtain 162 new MTFP instances. Details on each matrix can be seen in Table 4. They are also organized in 3 sets (L,M,H) according to the density.

Table 4 – Weighted matrices with BitCoin OCT social network dataset.

Set	Size (n)	Density (%) (a)	Negative entries (%) (b)	Positive entries (%) (c)	$\rho = a\frac{c-b}{100}$
L	25	39.34	28.81	71.19	16.67
M	25	51.00	6.54	93.46	44.33
H	25	69.00	4.83	95.17	62.33
L	50	36.73	41.11	58.89	6.53
M	50	42.73	6.88	93.12	36.85
H	50	56.98	7.88	92.12	48.00
L	100	20.72	5.12	94.88	18.60
M	100	25.17	39.89	60.11	5.09
H	100	39.94	7.69	92.31	33.80

Note. The BitCoin OCT network has density equals 0.10%.

2.8 Computational Experiments and Results

We report on computational results for different solution approaches based on the formulations presented in Sections 2.3 and 2.4. They were coded in C language using the CPLEX 12.8 callable library with parallelism enabled and four threads. The experiments were carried out on a machine equipped with Intel Core i5 Dual Core, 4 $\times$ 1.8Ghz, 8GB of RAM under macOS Catalina, version 10.15.5. To assess the efficiency of the solutions approaches, we calculated the average computational time (in seconds) required to exactly solve the 6 instances within the same class (1 to 9) and group (I, II, III). Each instance is composed by a tuple $(m,n,f,Q,1/\alpha,R,S,1/m)$, where (m,n,f,α) are given by Table 2 and (Q,R,S) are generated according to Section 2.7.

We tested the integer quadratic formulation (2.3)-(2.6) as well as the binary quadratic formulation defined by (2.9),(2.11)-(2.13) and using directly the product $y_{ilu}y_{jlv}$ instead of z_{ijluv} in the expression (2.8) that defines the objective function. From now on, these two formulations will be simply referred to as "integer quadratic" and "binary quadratic". We also tested the ILP formulation (2.9)-(2.14) with or without the inclusion of the presented valid inequalities. Table 5 shows the number of variables and constraints in each formulation. Furthermore, we considered the direct addition of each set of valid inequalities (2.37), (2.40), (2.42), (2.43), (2.45), (2.49) and (2.50) separately as well as of all combinations of them. Table 6 shows the number of constraints in each set of inequalities. The other sets of inequalities were not evaluated since they are dominated by or equivalent to the tested ones. This is a consequence of that fact that we used equality in (2.12), which is possible because the time allocation fractions are equally spaced in the tested instances (Assumption A1). This also allowed us to use inequalities (2.43), (2.45) and (2.49) at equality.

Table 5 – Number of variables and constraints in each of the MTFP formulations.

Formulations		Number of Variables	Number of Constraints
Integer Quadratic	(2.3)-(2.6)	$\mathcal{O}(n \times m)$	$\mathcal{O}(n+m \times f)$
Binary Quadratic	(2.9),(2.11)-(2.13) with z_{ijluv} replaced by $y_{ilu}y_{jlv}$	$\mathcal{O}(n \times m \times t)$	$\mathcal{O}(n+m \times f)$
ILP	(2.9)-(2.14)	$\mathcal{O}(m(n \times t + n^2 \times t^2))$	$\mathcal{O}(n+m(n^2 \times t^2 + f))$

Table 6 – Number of constraints generated from each set of the valid inequalities.

Valid Inequalities	Number of Constraints Generated
(2.37)	$\mathcal{O}(n \times m)$
(2.40)	$\mathcal{O}(n^2 \times m \times t)$
(2.42)	$\mathcal{O}(n^2 \times m)$
(2.43)	$\mathcal{O}(n \times m \times t \times f)$
(2.45)	$\mathcal{O}(n \times m \times t \times f)$
(2.49)	$\mathcal{O}(n \times m \times t \times f)$
(2.50)	$\mathcal{O}(n \times m \times t \times f)$

2.8.1 MTFP results for synthetic signal sociometric matrices dataset

In this subsection, we report on computational results for MTFP solution approaches with the instances based on synthetic signal sociometric matrices. For the smallest instances (those with $m = 2$ teams and $n = 25$ individuals), we tested the ILP formulation with all possible combinations of the inequalities (2.37), (2.40), (2.42), (2.43), (2.45), (2.49) and (2.50). See Table 25 in Appendix A for detailed results. Table 7 shows an extract of these results. It compares the computational times of the integer and binary quadratic formulations, the ILP formulation (2.9)-(2.14) and this ILP formulation with the combinations of inequalities whose average computational time for every class and group was less than 3 seconds. Each row relates to the formulation indicated in the first column and presents the average for the 6 instances within the same class and group. The last two rows show the averages for these selected combinations of inequalities and for all combinations, respectively.

Clearly, the strengthened ILP model reached an optimal solution within a considerably shorter time in comparison with pure ILP model. Besides, it consistently outperformed the quadratic formulations. It can also be seen that every combination of inequalities presented in Table 7 (those yielding the best performances) include inequalities (2.37), (2.43) and (2.45).

Table 7 – Average computational times for the quadratic models, the linear model and the linear model with inclusion of valid inequalities - Class 1, Class 4 and Class 7 instances with the synthetic sociometric matrices.

25 Individuals - Synthetic Sociometric Matrices									
Model	Class 1			Class 4			Class 7		
	I	II	III	I	II	III	I	II	III
Integer Quadratic	0.177	0.297	0.549	24.359	78.058	564.577	470.557	591.871	1333.226

Binary Quadratic	0.169	0.355	0.559	11.614	34.155	22.324	377.069	439.082	1083.731
ILP	0.348	0.407	0.376	21.039	35.034	14.414	569.736	260.326	202.214
37-43-45	0.171	0.151	0.143	0.634	0.926	0.812	0.858	1.253	0.635
37-40-43-45	0.171	0.153	0.136	0.667	0.866	0.669	0.806	0.767	0.595
37-43-45-49	0.116	0.166	0.137	0.554	0.938	0.591	0.993	1.857	0.670
37-43-45-50	0.135	0.143	0.112	0.648	0.772	0.870	0.729	1.448	0.679
37-40-42-43-45	0.183	0.204	0.196	0.667	0.926	0.739	1.206	1.382	0.977
37-40-43-45-49	0.199	0.170	0.162	0.927	1.325	0.700	1.214	1.137	0.806
37-40-43-45-50	0.123	0.130	0.118	0.664	1.192	0.618	0.947	0.813	0.643
37-42-43-45-49	0.214	0.176	0.147	0.619	0.749	0.698	1.382	1.114	0.717
37-42-43-45-50	0.103	0.139	0.123	0.530	0.841	0.575	0.822	0.843	0.517
37-40-42-43-45-49	0.275	0.183	0.230	0.618	0.811	0.738	1.231	1.027	0.827
37-40-42-43-45-50	0.119	0.156	0.121	0.612	1.197	0.782	0.818	0.989	0.677
37-40-43-45-49-50	0.201	0.147	0.139	0.695	1.212	0.538	1.166	0.956	0.652
All inequalities	0.279	0.268	0.208	0.553	0.861	0.667	1.277	1.216	0.954
Avg. (less than 3s comb.)	0.176	0.168	0.152	0.645	0.970	0.692	1.035	1.139	0.719
Avg. (all combinations)	0.425	0.422	0.391	14.279	16.299	9.915	85.972	48.155	40.092

Clearly, the strengthened ILP model reached an optimal solution within a considerably shorter time in comparison with pure ILP model. Besides, it consistently outperformed the quadratic formulations. It can also be seen that every combination of inequalities presented in Table 7 (those yielding the best performances) include inequalities (2.37), (2.43) and (2.45).

Table 8 shows similar information for the medium-sized instances (those with $m = 5$ teams and $n = 50$ individuals). It presents the average computational times for the instances (out of 6) solved within the time limit (10,000 seconds). The superscript in parentheses informs the number of resolved instances (if not 6). Cell value "n/a" (*not available*) indicates that none of the 6 instances were solved due to time/memory limitation (10,000s/7Gb). Out-of-memory cases occurred only with the integer quadratic formulation. As we can see, neither of the three pure formulations (integer and binary quadratic and ILP) was able to solve any instance of Class 8. Note also that the integer quadratic formulation did not solve 10 out of 18 instances of Class 5 and demanded much higher computing times for the others. The other two pure formulations did not perform well on these instances either.

On the other hand, the strengthened ILP formulation yielded an optimal solution in considerably shorter time. When comparing the potential of each set of inequalities, again the combinations including (2.37), (2.43) and (2.45) performed better so that we do not show results for the other combinations. Besides, note that the best combinations (in bold) also include inequalities (2.42). The results on Class 8, groups II-III differentiate these combinations from

the others. The last row shows the averages for the combinations presented in Table 8. Observe that the highlighted versions are exactly those that have spent time below average on every class/group.

Table 8 – Average computational times for the quadratic models, the linear model and the linear model with inclusion of valid inequalities - Class 2, Class 5 and Class 8 instances with the synthetic sociometric matrices.

50 Individuals - Synthetic Sociometric Matrices									
Models	Class 2			Class 5			Class 8		
	I	II	III	I	II	III	I	II	III
Integer Quadratic	29.007	538.844	2239.436	2568.255$^{(4)}$	3182.475$^{(3)}$	8393.801$^{(1)}$	n/a	n/a	n/a
Binary Quadratic	28.386	556.249	2728.781	727.976	1726.498	5398.983$^{(4)}$	n/a	n/a	n/a
ILP	134.468	50.292	304.134	706.723	553.631	2809.245	n/a	n/a	n/a
37-43-45	2.102	0.855	0.407	8.276	4.593	5.753	50.153	174.507	179.563
37-40-43-45	2.548	0.406	0.383	5.166	2.925	3.453	39.677	183.688	128.640
37-43-45-49	2.311	0.367	0.199	5.429	3.173	4.035	44.738	151.192	145.300
37-43-45-50	2.324	0.441	0.179	5.882	3.646	4.550	44.621	95.168	128.752
37-40-42-43-45	**1.882**	**0.301**	**0.264**	**5.505**	**2.967**	**2.758**	**37.752**	**37.949**	**27.495**
37-40-43-45-49	2.270	0.331	0.278	5.892	3.102	3.604	53.762	228.489	111.606
37-40-43-45-50	2.251	0.462	0.297	5.112	3.179	3.293	38.438	136.426	140.677
37-42-43-45-49	**2.110**	**0.310**	**0.199**	**5.564**	**3.009**	**3.207**	**42.711**	**37.369**	**31.251**
37-42-43-45-50	**2.150**	**0.369**	**0.230**	**5.279**	**2.748**	**2.872**	**42.580**	**32.618**	**45.665**
37-40-42-43-45-49	**1.825**	**0.320**	**0.235**	**5.708**	**3.108**	**3.074**	**39.416**	**36.829**	**35.239**
37-40-42-43-45-50	**1.996**	**0.333**	**0.260**	**5.722**	**2.992**	**2.832**	**28.434**	**34.472**	**43.523**
37-40-43-45-49-50	2.346	0.356	0.278	5.840	3.302	4.263	50.649	85.228	144.259
All inequalities	**1.952**	**0.312**	**0.232**	**5.007**	**3.159**	**3.560**	**42.432**	**30.483**	**41.783**
Avg. (only comb.)	2.159	0.397	0.265	5.722	3.223	3.635	42.720	97.263	92.596

Finally, we present the results for classes 3, 6 and 9 (instances with $m = 10$ teams and $n = 100$ individuals). In Table 9, we show the results only for the best combinations of inequalities, exactly those highlighted in Table 8, which uses inequalities (2.37), (2.42), (2.43) and (2.45). We can see that the highlighted version is the only one that spent time below average on every class/group. It is worth mentioning that the quadratic formulations and the pure ILP formulation were not able to solve any of these instances in the time limit.

Table 9 – Average computational times for the linear model with inclusion of valid inequalities - Class 3, Class 6 and Class 9 instances with the synthetic sociometric matrices.

100 Individuals - Synthetic Sociometric Matrices									
Models	Class 3			Class 6			Class 9		
	I	II	III	I	II	III	I	II	III

37-40-42-43-45	287.260	12.009	17.331	618.082	27.994	74.238	589.247	17.334	87.147
37-42-43-45-49	**278.704**	**10.261**	**7.175**	**477.952**	**18.212**	**89.928**	**294.976**	**11.974**	**119.469**
37-42-43-45-50	265.457	9.971	6.972	688.120	17.014	102.450	626.048	27.314	129.248
37-40-42-43-45-49	286.955	12.072	17.184	661.578	29.626	163.988	590.948	17.341	86.204
37-40-42-43-45-50	312.493	12.730	18.129	1158.721	30.212	65.941	456.084	14.217	191.214
All inequalities	303.832	12.624	18.681	646.585	25.898	179.565	471.785	14.110	208.279
Avg.	289.117	11.611	14.245	708.506	24.826	112.685	504.848	17.048	136.927

From these experiments, we conclude that the ILP formulation strengthened by proposed valid inequalities consistently outperforms the basic ILP and quadratic formulations. Particularly, inequalities (2.37), (2.43) and (2.45) were crucial to drastically reduce the computational time (see Table 25). This combination was even improved when used together with (2.42) and (2.49). Thus, the inclusion of these valid inequalities to eliminate fractional and/or symmetric solutions provides an up-and-coming strategy. We can still observe that the use of more inequalities does not necessarily improves the overall performance. It makes the formulation computationally heavier, which does not compensate the tightening of the linear relaxation.

2.8.2 MTFP results for real-world sociometric matrices dataset

The same experiments reported in Subsection 2.8.1 were carried out with the instances based on real-world social networks, generated according to Section 2.7. An extract of the computational results is presented in this subsection. Again, we have selected the combinations of inequalities leading to best performances, which coincide with those for the synthetic instances.

Tables 10 and 11 present the results for the smallest instances (those with $m = 2$ teams and $n = 25$ individuals) using the social networks Epinions and BitCoin OCT, respectively. Each value is the average computational time for the 6 instances within the same class and set. Again, the strengthened ILP formulations consistently outperformed the other formulations on both types of instances. It solved all instances with average computational time less than 1 second for every class and group.

Table 10 – Average computational times for the quadratic models, the linear model and the linear model with inclusion of valid inequalities - Class 1, Class 4 and Class 7 instances with the Epinions sociometric matrices.

25 Individuals - Epinions Social Network									
Models	Class 1			Class 4			Class 7		
	L	M	H	L	M	H	L	M	H
Integer Quadratic	0.123	0.141	0.169	4.453	3.437	40.308	60.123	240.141	100.169
Binary Quadratic	0.132	0.182	0.239	0.481	0.979	1.517	70.031	229.614	171.820
ILP	0.352	0.273	0.272	2.768	3.950	5.766	84.058	84.764	56.052
37-43-45	0.099	0.096	0.108	0.219	0.200	0.255	0.657	0.431	0.420
37-40-43-45	0.105	0.094	0.085	0.323	0.198	0.228	0.609	0.478	0.438
37-43-45-49	0.103	0.111	0.099	0.242	0.200	0.316	0.722	0.435	0.456
37-43-45-50	0.104	0.092	0.100	0.234	0.204	0.291	0.536	0.371	0.452
37-40-43-45-49	0.096	0.099	0.097	0.213	0.153	0.270	0.703	0.428	0.562
37-40-43-45-50	0.113	0.094	0.095	0.280	0.169	0.264	0.619	0.518	0.523
37-42-43-45-49	0.083	0.096	0.092	0.272	0.180	0.289	0.695	0.557	0.548
37-42-43-45-50	0.104	0.092	0.095	0.248	0.187	0.235	0.653	0.454	0.368
37-40-42-43-45	0.123	0.114	0.121	0.204	0.205	0.191	0.603	0.473	0.487
37-40-42-43-45-49	0.110	0.106	0.110	0.267	0.226	0.215	0.632	0.560	0.558
37-40-42-43-45-50	0.104	0.102	0.120	0.227	0.160	0.195	0.579	0.499	0.434
37-40-43-45-49-50	0.120	0.086	0.100	0.219	0.200	0.251	0.646	0.494	0.706
37-40-42-43-45-49-50	0.143	0.097	0.117	0.245	0.202	0.217	0.600	0.590	0.641
Avg. (only comb.)	0.108	0.098	0.103	0.246	0.191	0.248	0.635	0.484	0.507

Table 11 – Average computational times for the quadratic models, the linear model and the linear model with inclusion of valid inequalities - Class 1, Class 4 and Class 7 instances with the BitCoin OCT sociometric matrices.

25 Individuals - BitCoin OCT Social Network									
Models	Class 1			Class 4			Class 7		
	L	M	H	L	M	H	L	M	H
Integer Quadratic	0.146	0.177	0.278	6.418	13.368	102.762	13.307	48.140	136.944
Binary Quadratic	0.145	0.144	0.263	2.423	11.132	24.763	30.630	27.942	78.303
ILP	0.504	0.628	0.536	26.573	22.763	31.955	403.501	303.902	332.532
37-43-45	0.142	0.134	0.105	0.216	0.266	0.255	0.302	0.328	0.303
37-40-43-45	0.090	0.079	0.093	0.206	0.260	0.166	0.274	0.416	0.373
37-43-45-49	0.065	0.074	0.088	0.174	0.269	0.262	0.332	0.351	0.363
37-43-45-50	0.077	0.080	0.080	0.173	0.303	0.225	0.262	0.382	0.257
37-40-43-45-49	0.074	0.065	0.076	0.170	0.233	0.205	0.370	0.393	0.404
37-40-43-45-50	0.083	0.099	0.090	0.180	0.243	0.189	0.325	0.375	0.330
37-42-43-45-49	0.108	0.097	0.102	0.193	0.258	0.156	0.457	0.636	0.494
37-42-43-45-50	0.100	0.105	0.089	0.153	0.242	0.214	0.299	0.421	0.376
37-40-42-43-45	0.089	0.106	0.091	0.169	0.225	0.158	0.321	0.426	0.390
37-40-42-43-45-49	0.089	0.088	0.087	0.148	0.221	0.201	0.387	0.456	0.457

37-40-42-43-45-50	0.090	0.099	0.092	0.127	0.263	0.152	0.345	0.467	0.376
37-40-43-45-49-50	0.095	0.093	0.077	0.170	0.255	0.221	0.373	0.465	0.384
37-40-42-43-45-49-50	0.084	0.099	0.092	0.143	0.198	0.214	0.366	0.541	0.389
Avg. (only comb.)	0.091	0.094	0.089	0.171	0.249	0.201	0.340	0.435	0.377

Tables 12 and 13 show similar information for the medium-sized instances (those with $m = 5$ teams and $n = 50$ individuals). Once more, the strengthened ILP formulation yielded an optimal solution in considerably shorter time. Similarly to the synthetic instances, when comparing the performance of the valid inequalities, the combinations whose computational time is below average on every class/group include (2.37), (2.42), (2.43) and (2.45). They are highlighted in bold. However, unlike what happens with synthetic instances, the formulation with all inequalities is no longer in this selected group.

Table 12 – Average computational times for the quadratic models, the linear model and the linear model with inclusion of valid inequalities - Class 2, Class 5 and Class 8 instances with the Epinions sociometric matrices.

50 Individuals - Epinions Social Network									
Models	Class 2			Class 5			Class 8		
	L	M	H	L	M	H	L	M	H
Integer Quadratic	2.773	11.315	2953.001	43.225	72.654	2672.137$^{(3)}$	n/a	n/a	n/a
Binary Quadratic	2.741	19.885	2965.151	1553.922$^{(3)}$	4928.948$^{(4)}$	5608.864$^{(3)}$	921.593$^{(5)}$	2110.564$^{(5)}$	n/a
ILP	13.771	25.645	1150.047	137.764	90.174	2155.344$^{(4)}$	654.289$^{(5)}$	964.866$^{(5)}$	n/a
37-43-45	0.936	0.446	0.169	1.972	1.783	0.832	56.413	7.760	4.615
37-40-43-45	0.776	0.454	0.294	2.177	1.661	0.954	34.934	4.191	5.673
37-43-45-49	0.910	0.426	0.169	1.939	1.734	1.968	47.197	5.663	4.202
37-43-45-50	0.961	0.436	0.141	2.037	1.594	1.066	58.129	5.974	2.949
37-40-43-45-49	0.741	0.434	0.231	2.483	1.567	1.940	41.706	5.818	5.947
37-40-43-45-50	0.769	0.431	0.227	2.473	1.672	1.857	44.994	6.617	5.432
37-42-43-45-49	**0.689**	**0.451**	**0.137**	**2.128**	**1.670**	**1.201**	**26.527**	**4.915**	**3.931**
37-42-43-45-50	**0.650**	**0.409**	**0.163**	**2.132**	**1.509**	**0.883**	**30.617**	**5.064**	**3.103**
37-40-42-43-45	**0.770**	**0.400**	**0.170**	**1.928**	**1.516**	**0.906**	**29.664**	**3.199**	**4.547**
37-40-42-43-45-49	**0.735**	**0.424**	**0.191**	**1.951**	**1.443**	**1.182**	**30.573**	**4.068**	**4.060**
37-40-42-43-45-50	**0.783**	**0.425**	**0.193**	**2.066**	**1.627**	**0.926**	**32.770**	**4.027**	**4.600**
37-40-43-45-49-50	0.731	0.445	0.261	2.311	1.936	1.347	53.061	5.166	5.838
37-40-42-43-45-49-50	0.806	0.447	0.195	2.592	1.989	1.046	31.853	3.742	5.100
Avg. (only comb.)	0.789	0.433	0.196	2.168	1.669	1.239	39.880	5.092	4.615

Table 13 – Average computational times for the quadratic models, the linear model and the linear model with inclusion of valid inequalities - Class 2, Class 5 and Class 8 instances with the BitCoin OCT sociometric matrices.

50 Individuals - BitCoin OCT Social Network									
Models	Class 2			Class 5			Class 8		
	L	M	H	L	M	H	L	M	H
Integer Quadratic	12.674	33.963	754.831	695.961	224.938$^{(3)}$	1032.985$^{(4)}$	971.013	n/a	n/a
Binary Quadratic	10.638	42.993	684.873	145.848	30.792742	438.854	200.961	1231.775	3907.413
ILP	723.794	978.161	1071.752	518.723	401.804	2525.441$^{(4)}$	1099.522	665.088	1670.375
37-43-45	0.759	0.992	0.969	2.883	3.266	2.641	9.556	9.914	12.090
37-40-43-45	0.718	0.989	1.010	2.991	2.668	2.781	8.781	9.861	18.266
37-43-45-49	0.738	0.933	0.915	2.938	3.229	2.515	9.315	8.827	15.819
37-43-45-50	0.760	0.903	0.947	2.856	2.315	2.870	9.780	8.589	12.822
37-40-43-45-49	0.742	0.897	1.054	2.395	2.727	3.190	10.195	13.027	22.998
37-40-43-45-50	0.768	0.937	1.091	2.160	2.554	3.803	8.493	9.892	17.291
37-42-43-45-49	**0.657**	**0.722**	**0.853**	**2.192**	**2.363**	**2.378**	**8.169**	**10.212**	**16.520**
37-42-43-45-50	**0.681**	**0.807**	**0.921**	**2.154**	**2.176**	**2.368**	**7.499**	**10.035**	**14.926**
37-40-42-43-45	**0.608**	**0.803**	**0.999**	**1.934**	**2.064**	**2.222**	**8.914**	**10.606**	**16.037**
37-40-42-43-45-49	**0.640**	**0.807**	**1.009**	**1.701**	**2.661**	**2.701**	**8.493**	**10.234**	**15.936**
37-40-42-43-45-50	**0.643**	**0.881**	**1.016**	**2.062**	**2.914**	**2.782**	**8.025**	**10.589**	**16.587**
37-40-43-45-49-50	0.874	1.081	1.316	2.292	2.908	3.152	11.002	14.748	23.144
37-40-42-43-45-49-50	0.687	1.083	1.209	2.133	3.026	3.142	7.998	11.618	18.491
Avg. (only comb.)	0.714	0.910	1.024	2.361	2.682	2.811	8.940	10.627	16.994

Lastly, Tables 14 and 15 present the results for classes 3, 6 and 9 (instances with $m = 10$ teams and $n = 100$ individuals). We show the results only for the best combinations of inequalities that were selected in Tables 12 and 13. Over again, similarly to synthetic instances, we can see that the highlighted version is the only one that spent time below average on every class/group.

Table 14 – Average computational times for the linear model with inclusion of valid inequalities - Class 3, Class 6 and Class 9 instances with the Epinions sociometric matrices.

100 Individuals - Epinions Social Network									
Models	Class 3			Class 6			Class 9		
	L	M	H	L	M	H	L	M	H
37-42-43-45-49	**13.995**	**4.203**	**4.259**	**51.454**	**19.693**	**14.549**	**18.326**	**9.486**	**11.142**
37-42-43-45-50	14.046	4.230	3.829	57.378	17.166	24.216	16.778	17.757	7.991
37-40-42-43-45	15.313	7.411	5.834	64.895	31.624	11.672	17.477	13.529	11.204
37-40-42-43-45-49	15.403	7.411	5.872	80.970	36.742	13.598	18.207	14.280	11.821
37-40-42-43-45-50	15.166	7.182	5.904	89.040	25.914	13.657	21.798	17.777	25.625
Avg.	14.785	6.087	5.140	68.748	26.228	15.539	18.517	14.566	13.556

Table 15 – Average computational times for the linear model with inclusion of valid inequalities - Class 3, Class 6 and Class 9 instances with the BitCoin OCT sociometric matrices.

100 Individuals - BitCoin Social Network									
Models	Class 3			Class 6			Class 9		
	L	M	H	L	M	H	L	M	H
37-42-43-45-49	**60.977**	**21.777**	**34.028**	**34.915**	**17.001**	**32.361**	**12.117**	**11.052**	**12.737**
37-42-43-45-50	61.399	22.044	34.020	36.082	18.078	32.419	13.877	12.638	12.638
37-40-42-43-45	72.740	26.275	38.360	35.601	20.011	36.218	17.099	12.637	17.835
37-40-42-43-45-49	72.465	25.860	38.083	32.841	20.041	37.095	17.130	12.875	17.719
37-40-42-43-45-50	72.052	26.295	38.309	33.283	18.052	30.709	15.649	12.201	16.223
Avg.	67.927	24.450	36.560	34.544	18.637	33.760	15.175	12.281	15.430

In general, when it comes to comparing the performance of the formulations, the experiments with instances based on real social networks has led to the same conclusions as the experiments with the synthetic instances.

It is worth remarking that the instances with sociometric matrices extracted from Epinions and BitCoin OCT social networks were solved in a shorter time by all formulations when compared to instances with synthetic sociometric matrices. We believe that this fact is due to the density and structure of the matrices, which directly determines the non-null quadratic terms in the objective function and how they are related. Large real social networks, as those used here, usually have low density. For example, Epinions and BitCoin OCT networks have densities equal to 0.08% and 0.010%, respectively. As we used a submatrix of the adjacency matrix of the social network as matrix S, the larger the order n of S, the lower its density tends to be. This compensates for the potential increase in the complexity of the instance due to the increase of n. In addition, the non-null elements are usually concentrated in certain portions of the matrix due the group patterns of the social networks. In a way, this favors the separability of the problem. By its turn, the synthetic sociometric matrices have their density controlled by parameters and present a more homogeneous distribution of non-null entries. This structure better reflects social networks where a large proportion of individuals know each other, such as companies and schools.

2.9 MMTFP Results

For the tests with MMTFP solution approaches, we generated new instances from some of those used in the previous subsections. Basically, we changed the lists $Q_1, Q_2, \ldots, Q_f$ of individuals related to each skill. In an MTFP instance, every individual belongs to exactly

one of these lists. To create a related MMTFP instance, it was chosen a predefined percentage $\sigma \in \{30\%, 50\%, 70\%\}$ of individuals to have 2 or 3 skills, i.e. to belong to 2 or 3 of these lists. Thus, we classify the new instances into three categories: A ($\sigma = 30\%$), B ($\sigma = 50\%$) and C ($\sigma = 70\%$). The multi-skilled individuals were randomly selected from set H. Then, for each of them, the number of multiple skills (2 or 3 skills) and the skills themselves were randomly chosen. The original MTFP instances with $n = 50$ individuals (classes 2, 5, and 8) were used as base instances. Precisely, each MTFP instance originates 3 MMTFP instances (one for each category) by changing the lists $Q_1, Q_2, \ldots, Q_f$ as explained. Thus, the 6 original MTFP instances for each class (2, 5 and 6) and group (I, II and III) result in 6 MMTFP instances for each category (A, B, and C). In total, we generated 162 MMTFP instances for each dataset – Synthetic, Epinions and Bitcoin OCT.

We compared the performances of formulation (YY'), given by (2.15)-(2.19), and formulation (Y), obtained from (YY') by replacing (2.17)-(2.19) by (2.20). The initial model (Y) contains only inequalities (2.20) related to $K' = K_l$ and $K' = \{k_a\}$ for all $k_a \in K_l$, for all $p_l \in P$. The remaining inequalities (2.20) are used as lazy constraints and separated according to Corollary 2.5.1. One constraint is determined for each project and, if violated, it is then included in the model. Starting with none of inequalities (2.20) in the model has usually led to worse performance. We have also tried to separate fractional solutions and used the violated constraints as cuts. Since this strategy did not improved the overall performance, it was not applied in the experiments reported here.

Both formulations are strengthened by inequalities (2.37), (2.40) and (2.42), as well as by (2.51) and (2.53) for each set $K' = \{k_a\}$ for all $k_a \in K_l$, and (2.51), (2.53), (2.55), (2.56) and (2.57) for $K' = K_l$. These inequalities are directly inserted in the original ILP formulations. Besides, (2.51), (2.53), (2.55) and (2.56) are used at equality for $K' = K_l$. This is possible because A1-A2 hold for the tested instances. We have tested some other combinations of inequalities, and sometimes they can perform better. For consistency, the computational times presented in the next tables were obtained by the version just specified.

Table 16 compares the performance of formulations with the synthetic instances. Tables 17 and 18 show the corresponding results for the Epinions and Bitcoin OCT datasets, respectively. Each value in the tables is the average computational time (in seconds) needed to exactly solve the instances within the same category, class and group/set. We remark that the use of inequalities (2.52) and (2.54) causes the solver to exceed the time limit (10,000s).

Table 16 – Average computational times for formulations (YY') and (Y) with inclusion of valid inequalities - MMTFP instances with synthetic sociometric matrices.

50 Individuals - Synthetic Sociometric Matrices									
Models	Category A								
	Class 2			Class 5			Class 8		
	I	II	III	I	II	III	I	II	III
YY'	20.079	5.249	5.820	68.849	51.102	80.911	56.402	277.869	25.256
Y	20.952	4.271	5.479	54.531	34.134	53.349	109.494	55.992	15.657
Models	Category B								
	Class 2			Class 5			Class 8		
	I	II	III	I	II	III	I	II	III
YY'	43.535	14.544	1.864	889.024	541.730	135.260	2563.741	647.223	67.502
Y	45.244	19.580	1.579	575.449	66.564	34.829	1072.840	125.190	22.987
Models	Category C								
	Class 2			Class 5			Class 8		
	I	II	III	I	II	III	I	II	III
YY'	212.561	34.017	2.063	546.805	55.403	32.203	2584.242	151.792	81.992
Y	162.808	29.147	2.436	283.705	58.371	24.022	1105.577	73.146	33.402

Table 17 – Average computational times for formulations (YY') and (Y) with inclusion of valid inequalities - MMTFP instances with Epinions sociometric matrices.

50 Individuals - Epinions Social Network									
Models	Category A								
	Class 2			Class 5			Class 8		
	L	M	H	L	M	H	L	M	H
YY'	5.773	6.629	1.545	18.544	88.489	2.828	17.879	5.728	11.143
Y	6.110	5.243	0.562	15.513	69.842	1.358	18.841	5.211	6.592
Models	Category B								
	Class 2			Class 5			Class 8		
	L	M	H	L	M	H	L	M	H
Y	6.227	3.343	1.322	18.227	19.794	8.367	22.460	19.647	21.103
Y	7.589	3.612	0.953	16.981	11.649	4.009	17.037	15.356	17.488
Models	Category C								
	Class 2			Class 5			Class 8		
	L	M	H	L	M	H	L	M	H
Y	23.238	1.128	1.408	25.804	73.583	4.983	39.906	45.661	32.385
Y	17.940	1.400	1.238	25.641	12.732	2.707	30.519	22.280	16.811

Table 18 – Average computational times for formulations (YY') and (Y) with inclusion of valid inequalities - MMTFP instances with BitCoin OCT sociometric matrices.

50 Individuals - BitCoin OCT Social Network									
Models	Category A								
	Class 2			Class 5			Class 8		
	L	M	H	L	M	H	L	M	H
YY'	6.781	6.845	12.826	11.314	14.653	21.115	16.836	18.116	24.250
Y	5.776	6.250	12.929	8.375	13.580	20.702	16.776	20.364	18.156
Models	Category B								
	Class 2			Class 5			Class 8		
	L	M	H	L	M	H	L	M	H
YY'	8.189	14.259	20.664	19.077	19.728	42.280	44.417	71.138	775.220
Y	10.416	11.635	22.594	19.958	14.928	42.266	19.958	24.928	52.266
Models	Category C								
	Class 2			Class 5			Class 8		
	L	M	H	L	M	H	L	M	H
YY'	13.182	14.537	37.942	24.319	38.643	61.787	182.075	41.072	1062.285
Y	15.146	17.041	30.067	21.923	39.327	111.886	177.709	35.333	763.375

It is worth highlighting that both formulations were not able to solve the MMTFP instances in the time limit without the use of additional valid inequalities. On the other hand, using less inequalities may shorten the computational times in some cases. This happens for example in Class 2 if inequalities (2.40) and (2.42) are disregarded. However, these inequalities are very useful in Class 8, where times would much longer without them. We can also remark that formulation (Y) outperforms or is comparable to formulation (YY') in the majority of the test cases. We believe that this happens because formulation (YY') has more variables and the strategy used to separate inequalities in formulation (Y) was effective.

2.10 Instances hardness

The input of the MTFP (or the MMTFP) is defined by a tuple $(m = |P|, n = |H|, f = |K|, Q, D, R, S, W)$. So, it is almost impossible to draw general conclusions on which parameters or combinations of parameters most affect the hardness of the problem. In order to offer some insights into this issue, we have plotted some graphics to compare the variation in the average computational times versus some specific parameters. We analyze the integer and binary quadratic formulations, the pure ILP formulation and its best strengthened version. Since the order of the times may vary too much from one formulation to another, we had to use different

scales in graphics, even when they belong to a same figure. So, we advertise the reader to pay attention to ranges in the y-axis.

Number of individuals (n), teams (m) and skills (f)

The triple (m,n,f) defines the size (number of variables and constraints) of the quadratic models and, together with $|D|$, the size of the linear models. So, an exponential increase in time is expected with an increase in the values of these parameters. Such an expectation is confirmed by the graphics of Figure 3. Each bar in the chart corresponds to the average time for the 45 synthetic instances associated with the same triple (m,n,f). The time limit (10,000s) was taken as the resolution time for the instances not solved within this limit. This means that the bars related to triple $(50,5,5)$ and mainly to triple $(100,10,10)$ can be much higher in Figure 3(a).

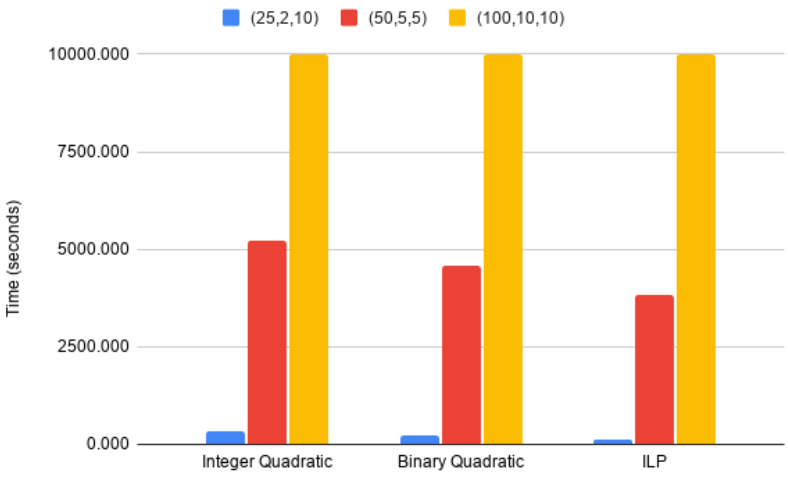
Set of time allocation fractions (D)

In the tested instances, this set is precisely defined by α, and its size is $1/\alpha$. This parameter impacts on the number of possible teams and allocation of dedication times among their members. Besides, it directly affects the values of the solutions and the number of symmetric solutions in the linear formulations. Thus, as Figure 4 shows, the resolution times increase exponentially as α decreases. Each bar describes the average computational time for the 36 synthetic instances related to the respective α and $n \in \{25,40\}$ (those with 100 individuals were not considered because they could not be solved by the pure formulations).

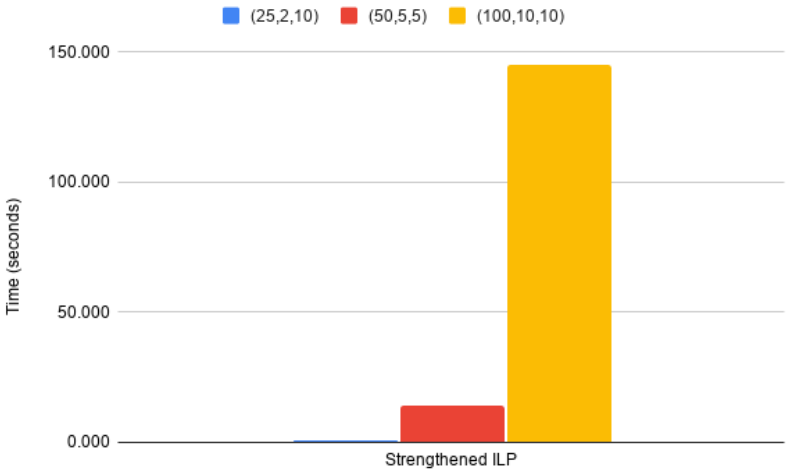
Project requirements matrix R

The structure of this matrix and the value distribution of its entries influence the "combinatorics" of the problem. We have evaluated this influence by two parameters: (i) $\sum_{k_a \in K, p_l \in P} r_{al}/n$, which measures the proportion of demanded (fractions of) individuals, (ii) density of R, i.e. proportion of non-null entries. Figures 5-6 summarize the results for the synthetic instances with $n \in \{25,50\}$.

Regarding (i), the synthetic instances were classified in 3 levels (low, medium and high) that correspond to matrices R_1, R_2 and R_3 used to generate the instances. Figure 5(a) shows that instances with medium requirement levels tend to be more difficult to solve by the pure formulations. Note that, for very low or very high requirement levels, the problem becomes very little restricted or quite restricted and tends to be easily solved. However, a different tendency was observed with the strengthened ILP formulation (see Figure 5(b)). We believe that the



(a) Pure formulations

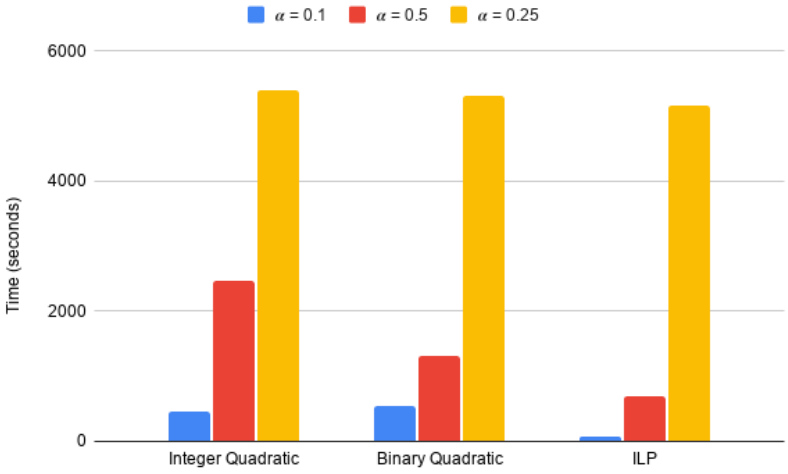


(b) ILP with inequalities 37,42,43,45,49

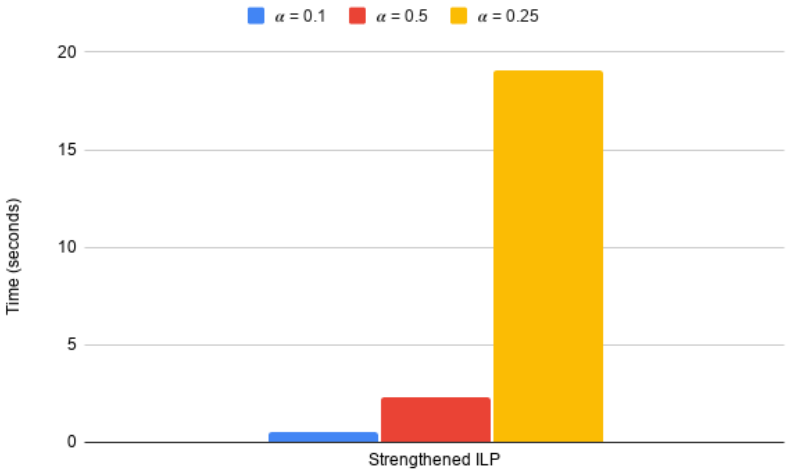
Figure 3 – Average computational times per number of individuals, teams and skills (n, m, f) for the synthetic instances. Scales are different in the two graphics.

valid inequalities, specially when used at equality, are even more effective if the feasible set has a greater number of symmetric solutions, which occurs for low and medium requirement levels. Although instances with high level requirements have demanded higher times, it is worth remarking that two outlier instances are leading to the discrepant height of the third bar in Figure 5(b). Disregarding them, the average drops to 1,244s, little higher than the average for the medium level (1,151s).

Regarding (ii), we have observed in our experiments that the size of the subset of skills required by each project and the intersection between these subsets influence the hardness of the instances. Trying to capture this influence through a single parameter, we consider the density of R and partition the instances into two groups, related to sparser and denser matrices



(a) Pure formulations



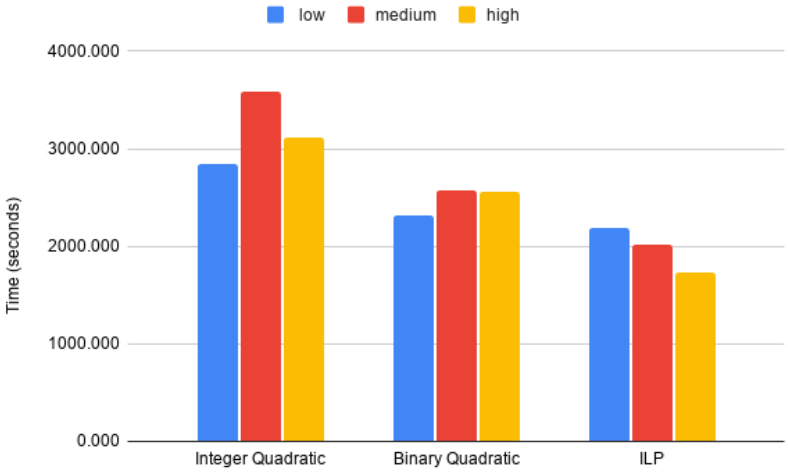
(b) ILP with inequalities 37,42,43,45,49

Figure 4 – Average computational times per time allocation fractions (α) for the synthetic instances with $n \in \{25, 50\}$. Scales are different in the two graphics.

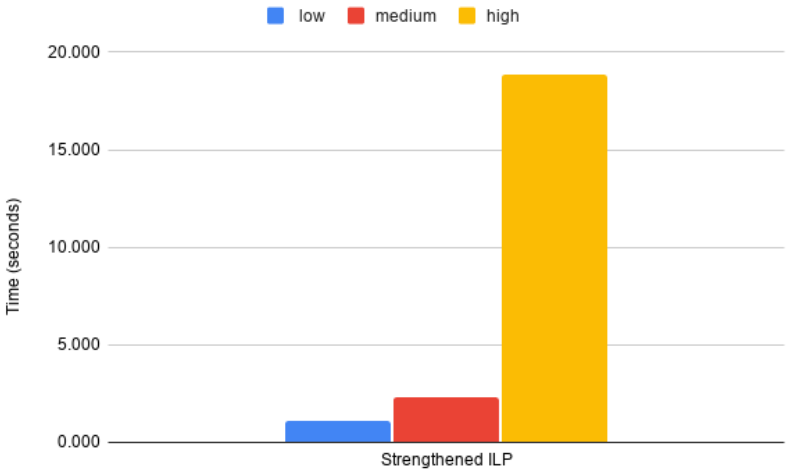
R, respectively. The boundary used to discriminate between sparse and dense depends on n. Figures 6(a)-6(b) reveal our observation. Another evidence appears when comparing the computational times for class 8 $(n = 50)$ and class 9 $(n = 100)$ in the same group II (see Tables 8 and 9). Although smaller in size, the former are defined by denser matrices R and take much time.

Interpersonal relationship matrix S

This matrix defines the second derivative of the objective functions of the quadratic formulations, thus determining their convexity/concavity. Therefore, it may have great impact on the quality of the bounds provided by the relaxations of these formulations and their linearized



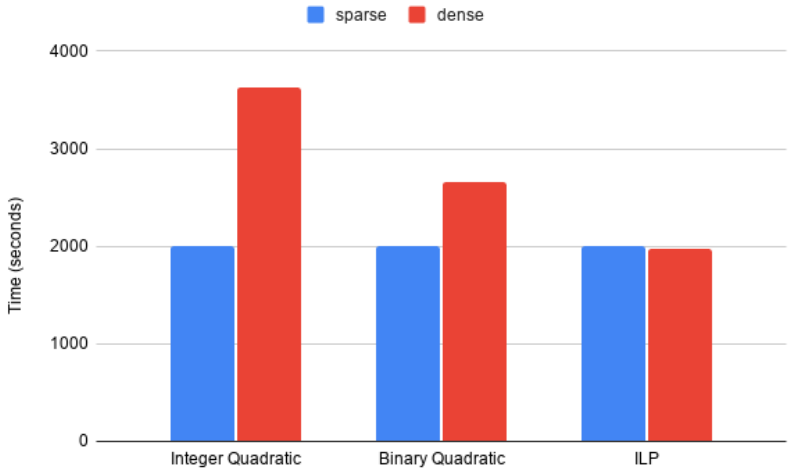
(a) Pure formulations



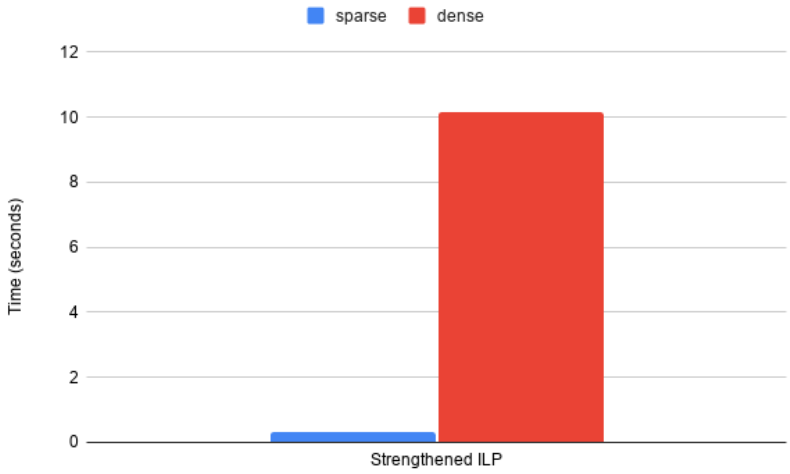
(b) ILP with inequalities 37,42,43,45,49

Figure 5 – Average computational times per level of requirements of individuals (low, medium, high) for the synthetic instances with $n \in \{25, 50\}$. Scales are different in the two graphics.

versions. We consider two attributes of matrix S that interfere with the quadratic form: (i) its density, and (ii) the difference between the number of positive and negative elements. In the case of the synthetic instances, both values are related to parameter p, which is the percentage of $+1$ entries and also defines the percentages $0.15(1-p)$ and $0.85(1-p)$ of -1 and null entries, respectively. For $p = 30\%, 50\%, 70\%$ (i.e. groups I, II and III), the density of S is 40.5%, 57.5% and 74.5%, respectively. Figure 7 shows the average computational times for the synthetic instances with $n \in \{25, 50\}$. For the pure formulations, the higher the value of p (and so the density of S), the longer the computational times. However, the use of the valid inequalities reverts such a tendency, as shows Figure 7(b). They are able to tighten the feasible region and cut non-optimal or symmetrical feasible solutions.



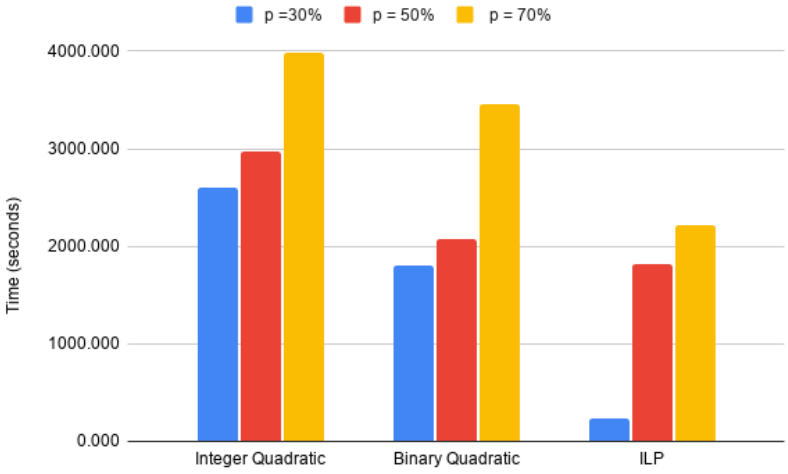
(a) Pure formulations



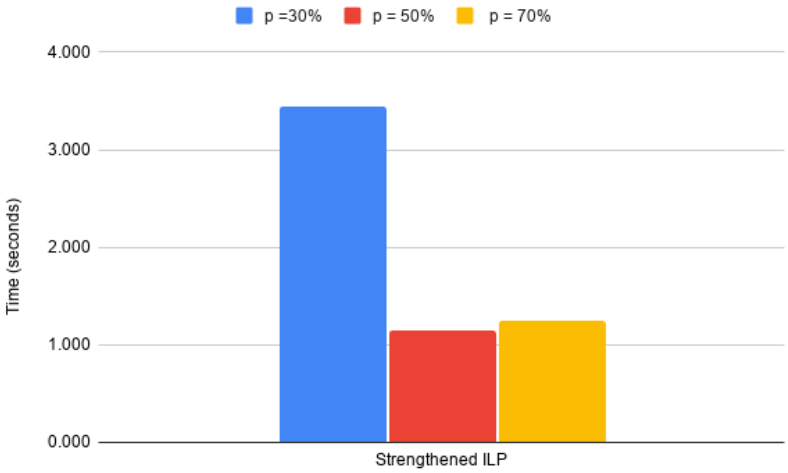
(b) ILP with inequalities 37,42,43,45,49

Figure 6 – Average computational times per density of the requirement matrices (sparse, dense) for the synthetic instances with $n \in \{25, 50\}$. Scales are different in the two graphics.

In the case of matrices S based on Epinions and BitCoin OCT networks, the number of positive and negative elements are not defined by a parameter. As this relation also influences the resolution time, we grouped the instances by density (L,M,H) and also by ρ, which is density weighted by the difference between the percentages of positive and negative elements. Tables 3 and 4 show that the increasing order in ρ corresponds to the increasing order in the density, except for the Epinions instances with $n = 25$, where L, M and H correspond to ρ equals to 27%, 45% and 31%, respectively. Figure 8 and 9 present the computational times for $n = 25$ and $n = 50$, respectively. In each group of 6 bars, the 3 leftmost bars relate to Epinions instances and the 3 rightmost ones are associated with BitCoin OCT instances. For the pure formulations, the charts essentially show that times usually increase as ρ increases. This means that times tend to



(a) Pure formulations



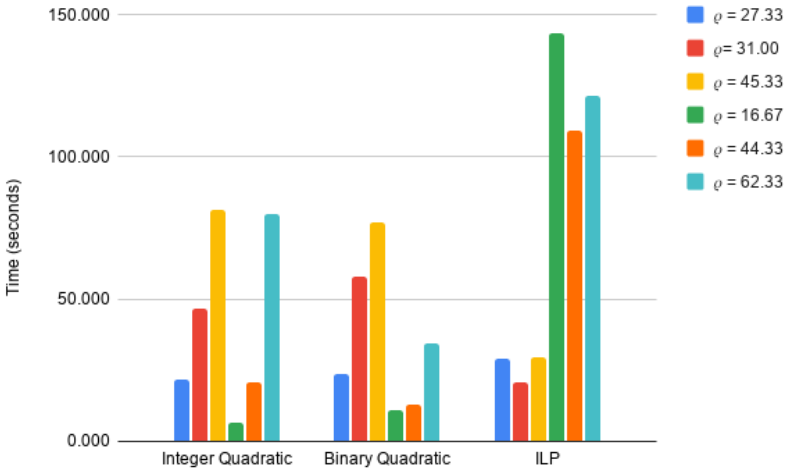
(b) ILP with inequalities 37,42,43,45,49

Figure 7 – Average computational times per percentage p of positive entries in S for the synthetic instances with $n \in \{25, 50\}$. Scales are different in the two graphics.

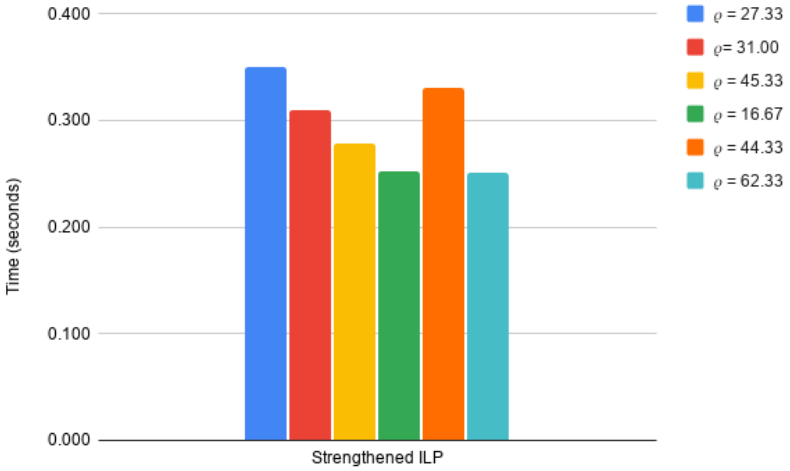
increase with density, but the number of positive and negative elements may alter such a tendency, as in the case of Epinions instances with $n = 25$. After using the valid inequalities to strengthen the ILP formulation, we could not draw a general conclusion on how the computational times depend on S.

Lists of individuals per skill $Q = \{Q_1, Q_2, ..., Q_f\}$

The possible worker-skill assignments are linked to the intersections among these lists, which are directly affected by parameter σ, the percentage of multi-skilled individuals. Recall that $\sigma = 30\%$, 50% and 70% for the MMTFP instances in categories A, B and C, respectively. Figure 10 presents the computational times for the MMTFP instances spent by



(a) Pure formulations

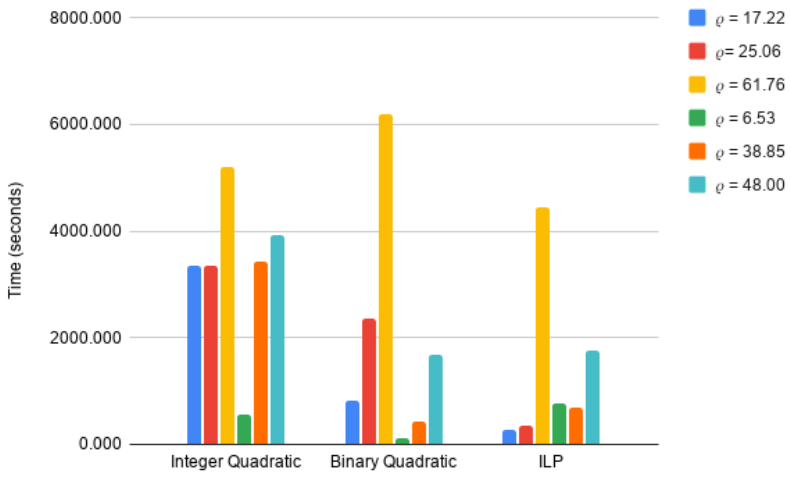


(b) ILP with inequalities 37,42,43,45,49

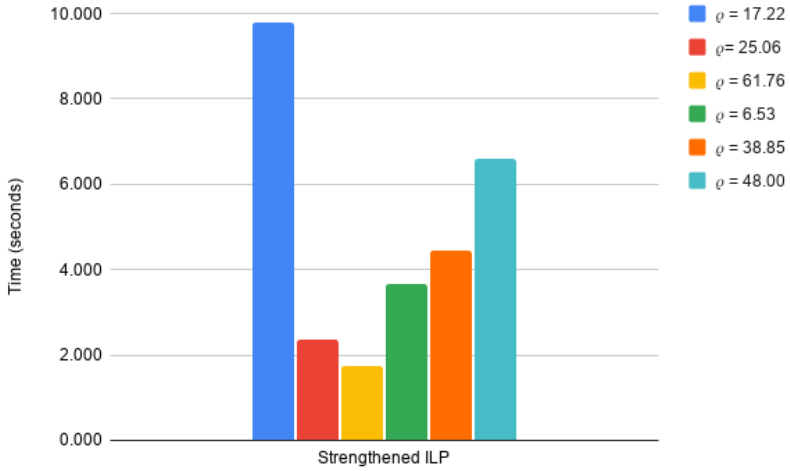
Figure 8 – Average computational times per ρ for Epinions and BitCoin OCT instances with $n = 25$. Scales are different in the two graphics.

formulations (YY') and (Y), both tightened by inequalities (2.37), (2.40), (2.42), (2.51), (2.53), (2.55), (2.56) and (2.57), as previously specified.

The charts for formulations (Y) and (YY') present similar patterns, and this usually holds regardless of the valid inequalities included. However, we have observed that, for the same formulation, the relation between the computational times and σ may significantly change in some cases depending on valid inequalities used. To illustrate this situation, Figure 11 presents the results when inequalities (2.40) and (2.42) are not discarded. In this figure, we can observe that times increase as σ increases in almost all of the cases. However, in Figure 10, the same does not hold for the synthetic instances of Class 5, for example. Therefore, the relative resolution times between instances depends on the their characteristics and also on how the formulation



(a) Pure formulations



(b) ILP with inequalities 37,42,43,45,49

Figure 9 – Average computational times per ρ for Epinions and BitCoin OCT instances with $n = 50$. Scales are different in the two graphics.

was strengthened.

2.11 Conclusion

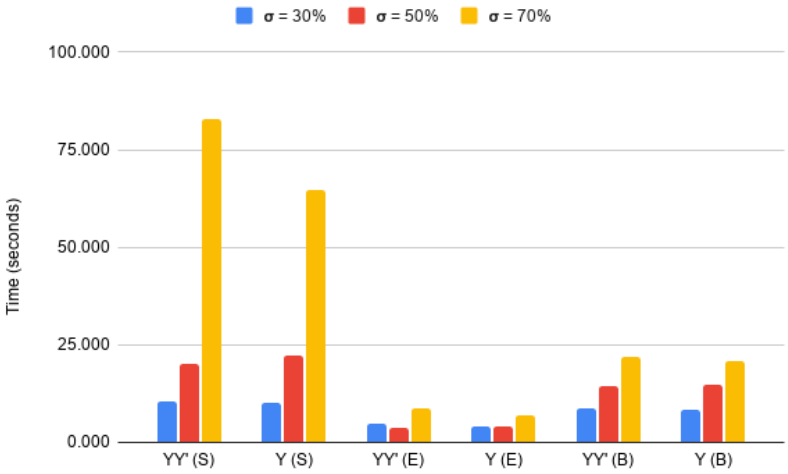
In this work, we introduced an ILP formulation for the MTFP, as defined by Gutiérrez *et al.* (2016). Since the problem is NP-Hard, instances of practical interest may not be solved to optimality in a reasonable time. In order to improve resolution time, we proposed some classes of valid inequalities. The performance of our ILP formulation together with the inequalities was compared with the unique model presented in the literature. The comparative analysis took into account a set of benchmark instances of size and characteristics compatible with real applications as well as instances generated from real-world social networks. It showed that our approach

consistently outperforms — in terms of effectiveness and efficiency — the one in the literature. Besides, our numerical experiments indicated that the derived valid inequalities are very useful to strengthen the formulation and to significantly reduce its resolution time.

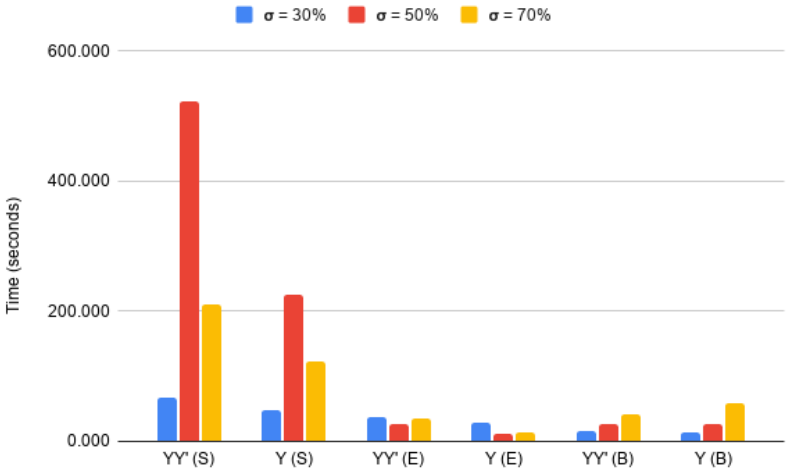
In addition, we introduced an extension of MTFP where individuals can have multiple skills. This scenario was suggested in the original work by Gutiérrez *et al.* (2016) and already considered by Campêlo *et al.* (2020) in the context of a related team formation problem. Differently from the model by Gutiérrez *et al.* (2016), our ILP model can be easily adapted to this generalized version at the expense of adding a polynomial number of variables and constraints. Alternatively, we were able to keep the same set of variables by using an exponential number of constraints, which we proved separable in polynomial time. The correctness of this second approach is grounded on a more ingenious proof that uses the max-flow/min-cut theorem.

We theoretically related the two resulting formulations for this extended version of the problem and presented valid inequalities that generalizes those proposed for the MTFP. Also, we computationally compared the two formulations on a set of instances adapted from the MTFP benchmark. The analysis of the computational times showed that formulation (Y) usually performs better than (YY') due to an effective separation procedure.

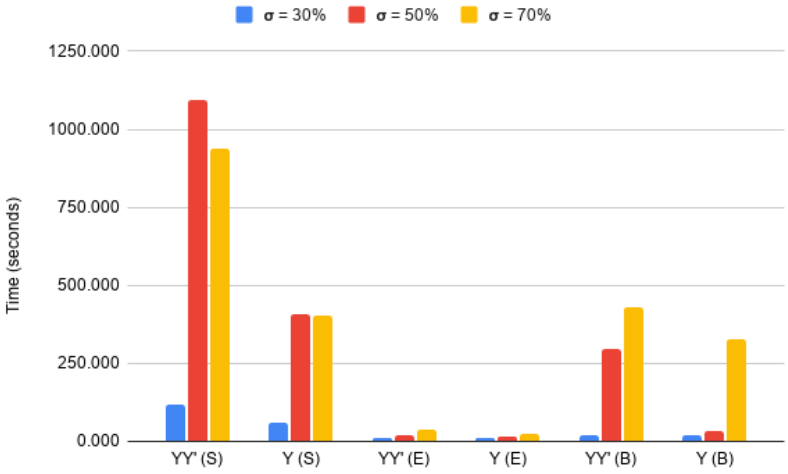
Since the used valid inequalities are hard to separate in general, we added them directly to the formulations in all tests. This was enough to solve the considered instances. As future work, the development of efficient (non-trivial) separation heuristics could be attempted to deal with larger and more complex instances. Another interesting track of research would be introducing a time component in the team formation problems. For instance, one could consider a time horizon for each project. Thus, an individual's dedication to a project would count only during (part of) that time horizon.



(a) Class 2

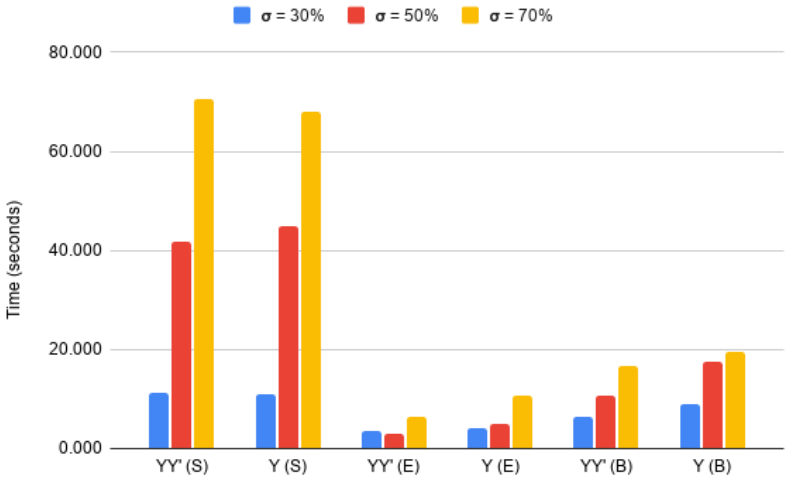


(b) Class 5

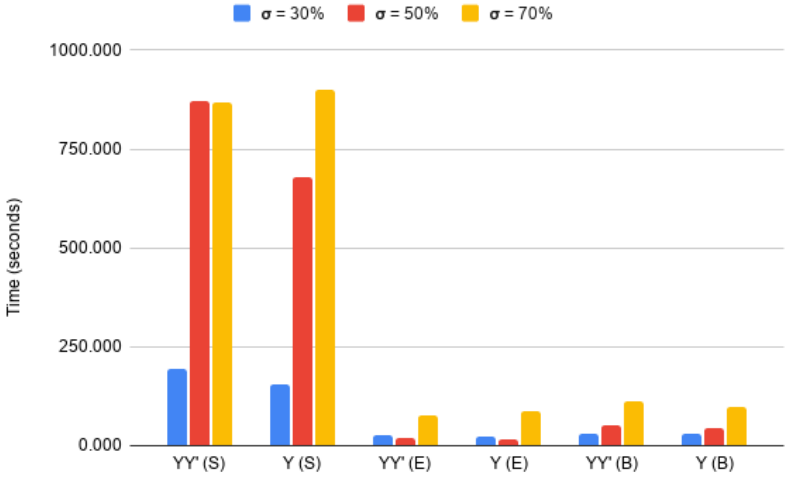


(c) Class 8

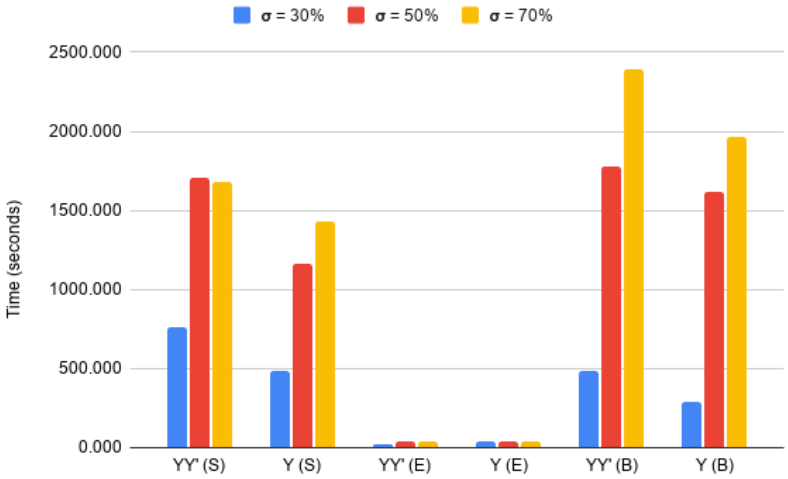
Figure 10 – Average computational times per σ for Synthetic (S), Epinions (E) and BitCoin OCT (B) instances with $n = 50$. Scales are different in the two graphics.



(a) Class 2



(b) Class 5



(c) Class 8

Figure 11 – Average computational times per σ for Synthetic (S), Epinions (E) and BitCoin OCT (B) instances with $n = 50$. Inequalities (2.40) and (2.42) are not used. Scales are different in the two graphics.

3 COMPETITIVE TEAMS FORMATION PROBLEM

Different scenarios where team formation problems appear give rise to many variations of the TFP. In Chapter 2, we have considered a version of the problem that maximizes affinities within teams. Another possibility would be to simultaneously maximize the intra-team affinities and the competitiveness between teams. Using positive and negative edge weights to represent affinity and competitiveness relationships, respectively, this would basically go back to considering the problem in Chapter 2 with a modified objective function. Alternatively, we could think of finding skilled teams as large as possible that meet a certain pattern of intra-team and inter-teams relationships. For instance, based on Heider's structural balance theory, we can prohibit negative relationships within a team as well as positive relationships between different teams. This strategy, to be better detailed in this chapter, leads us to define a new problem called Competitive Teams Formation Problem (CTFP). It uses Heider's structural balance theory to form work teams meeting some predefined competences. Although it is a novel problem, it can be related to other grouping problems that are grounded on this theory. We start by reviewing the literature about these related problems. Then, we formally define the CTFP and restate some properties on balance. We propose an ILP model for the problem and use the principle of structural balance to create valid inequalities that enhance the computational performance of the model. Several computational experiments are presented.

3.1 Introduction

Many times, organizations need to form collaborative work teams to design and implement decisions that require expertise from diverse areas. As examples, we can citeonline boards on budget planning and forecasting of companies, groups in political parliaments, scientific or technological innovations committees, and sports competition teams (FRANZ; JIN, 1995). A notable common element among these applications is that all of them occur in collaborative vs. conflicting environments, where competitiveness is constantly present and has great influence on the productivity and decision making of the teams (INOHARA, 2003; LI *et al.*, 2015b; SHI *et al.*, 2016; ARINIK *et al.*, 2017).

Goal complexity and diversity stemming from differences in requirements and specialization within these teams often induce conflict of opinions among participants. Agreement among work group leaders and labor negotiation teams, for example, requires collaboration with

a high potential for conflicts (FRANZ; JIN, 1995). Although conflicts can be viewed as negative at a first glance, they have important implications in increasing the effectiveness of a team's decision-making process and can be seen as a dynamic with a valuable effect on the outcome and productivity (ESQUIVEL; KLEINER, 1996).

Attempts to model collaborative vs. conflicting relationships have been the focus of researchers since the 1940s when Heider (1946) introduced the *structural balance theory*. According to this theory, social groups tend to organize themselves so as to reach a balanced state where conflict situations are avoided. Since then, efforts to understand the properties of structural balance in social networks continued for years and have led to the study of the so-called *signed graphs*, with seminal contributions by Harary (1953), Davis (1963) and Davis (1967). By considering a set of individuals as the vertices of a graph, each edge is labeled with a positive $(+)$ or negative $(-)$ sign, in order to indicate the nature of the relationship between the corresponding adjacent vertices. A signed graph can be used to model any system containing two types of antithetical relationships, such as like/dislike, for/against, etc (ARINIK *et al.*, 2017).

Although the initial studies about signed graphs started a long time ago, the advent and popularization of online social networks in the last few decades have greatly renewed the interest in this data structure. It has been shown to be an attractive representation for social networks when dealing with the challenge of analysing structural balance (YANG *et al.*, 2007; DOREIAN; MRVAR, 2009; LESKOVEC *et al.*, 2010; FACCHETTI *et al.*, 2011; TANG *et al.*, 2016).

As in these works, we also use signed graphs to represent relationships between workers in a social network. Using this representation, we define a new optimization problem to be called *Competitive Teams Formation Problem* (CTFP). The CTFP can be seen as a novel version of the problem proposed by Lappas *et al.* (2009) who, in a pioneering way, defined a team formation problem using social networks.

Given a project that requires a set of skills and a pool of single-skilled workers who are organized in a social network, the *Team Formation Problem* (TFP), as defined by Lappas *et al.* (2009), refers to finding a subset of the workers that cover the skill requirements and can communicate effectively with each other. Generalizations of the TFP that deal with multiple teams and/or multi-skilled individuals have also been considered (GUTIÉRREZ *et al.*, 2016; CAMPÊLO *et al.*, 2020). The concept of effective communication can be measured using several *communication cost* metrics.

Differently from Lappas *et al.* (2009), which compute the communication costs within a team based on the distances between its members in the social network, in this work we use the theory of social balance to define the intra-team and inter-teams communication costs. Using a signed graph to represent the social network, the teams must be formed in such a way that only positive relationships occur between individuals of the same team and only negative relationships happen between individuals of different teams. This kind of restrictions for people grouping is grounded on Heider's balance theory (HEIDER, 1946) and has been widely used, as detailed in the next section. They aim to ensure some certain affinity degree within each group while feeding the competitiveness among them.

To the best of our knowledge, Kouvatis *et al.* (2020) are the only authors to consider multiple teams formation problems with skill requirements that explicitly use signed graphs to define compatible/incompatible groups of individuals. Although they have studied different notions of compatibility, such as direct positive/negative edge compatibilities and compatibility based on the structural balance defined by shortest-path-structures, the case modeled by the CTFP was not considered. Thus, the present work introduces a novel version of the TFP.

The remainder of this text is structured as follows. In Section 3.2, we present a review of the literature about problems similar to CTFP in the sense that they are also based on the principle of structural balance. The CTFP is then formally defined in Section 3.3. An edge contraction operation on signed graphs is also defined in this section, which will be very useful along the text. For instance, in Section 3.4, it is used to restate some structural properties on balance which imply the NP-hardness of the CTFP. In the sequel, in Section 3.5, we revisit and extend some polyhedral results that will be the base to derive valid inequalities for the problem. In Section 3.6, we propose an ILP model for the CTFP. We use the principle of structural balance to create valid inequalities that enhance the computational performance of the model. In Section 3.7, we describe and analyze numerical experiments with the formulation and valid inequalities proposed in this work. Finally, in Section 3.8, we close the chapter with general conclusions and directions for future works.

3.2 Related works

The structural (or social) balance theory was first elaborated by Heider (1946) with the purpose of describing antagonistic relationships (like/dislike, love/hate, respect/disrespect, trust/distrust) between individuals pertaining to the same social group and then comprehending

the origin and structure of tensions and conflicts in the social network formed by them. While there may be some criticism of this theory, it states that there is a tendency for individuals to achieve a *balanced state*. Changes toward balance will occur since the state of imbalance produces tension. The outlines of this theory can be found in Wasserman e Faust (1994) and Easley e Kleinberg (2010), for instance.

Cartwright e Harary (1956) formalized Heider's theory via the concept of *signed graph*, introduced by Harary (1953), i.e. a graph where each edge is assigned either a positive or negative sign. They presented a rigorous generalization of Heider's concept of balance to mean the absence of cycle with an odd number of negative edges in the signed graph. The Structure Theorem (HARARY, 1953) then states that a balanced social group can be partitioned into two disjoint subgroups (one of them possibly empty) such that all relationships (the edges of the signed graph) within each subgroup are positive (internal solidarity) and all those between subgroups are negative (mutual antagonism). The term *balanced* is indistinctly used to qualify the social network and its corresponding signed graph.

Based on sociometric studies, Davis (1967) later introduced a generalized notion of structural balance (clustering in his terminology) by claiming that a balanced social group is one that can be partitioned into $k \geq 2$ mutually antagonistic and disjoint subgroups (or clusters), each having internal solidarity. Again, the intra-cluster relationships are all positive, and those inter-cluster are all negative. In this case, we refer to the group and the corresponding signed graph as *k-balanced* (partition in at most k clusters) or *clusterable* (k-balanced for some k or, equivalently, n-balanced for n equals the number of vertices). Davis (1967) showed that being clusterable is equivalent to containing no cycle with exactly one negative edge.

This generalized balance notion has been used in different scenarios (LEVORATO; FROTA, 2017). However, real-world signed social networks are not expected to be clusterable in general (DOREIAN; MRVAR, 2009). This may happen for different reasons, e.g. the network is still changing through time towards balance. Harary (1959) proposed some measures of balance/imbalance (HARARY *et al.*, 1965). One of them is the *edge-index for balance*, i.e. the minimum number of edges that, when removed (or have the sign reversed), leads to a balanced graph. Another one is the *vertex-index for balance*, i.e. the minimum number of vertices whose removal yields a balanced graph. Both indices can also be defined with respect to k-balance and clustering.

These measures of balance naturally give raise to optimization problems. However,

only decades after the works by Cartwright e Harary (1956) and Davis (1967), Doreian e Mrvar (1996) studied the following problem: given k as input, partition the signed graph into (at most) k clusters such that the number of negative intra-cluster edges and positive inter-cluster edges is minimized, i.e. the edge-index for k-balance is minimized. The authors proposed a local search heuristic for the problem. It was then revisited by Brusco e Steinley (2010) who developed a combinatorial branch-and-bound algorithm.

Motivated by a machine learning application and ignoring the structural balance theory, Bansal *et al.* (2004) formalized the variant of this problem where k is not fixed as the so-called *Correlation Clustering* (CC). Its goal is to maximize the number of positive edges within clusters plus negative edges between clusters, or equivalently to minimize the edge-index for clustering. Let us observe that CC corresponds to the previous problem with k equals the number of vertices, provided that empty clusters are allowed. Besides, it is worth noticing that, although equivalent at optimality, the maximization and minimization problems differ from the point of view of approximation. Bansal *et al.* (2004) showed that CC is NP-hard and presented polynomial time approximation schemes (PTAS) for the maximization version as well as a constant factor approximation for the minimization version, when the graph is complete. Demaine *et al.* (2006) presented an $O(\log n)$-approximation for the weighted version of the minimization problem on general signed graphs. The algorithm is based on an ILP formulation for graph clustering which was further studied by Figueiredo e Moura (2013). Metaheuristic solutions for CC have also been proposed (LEVORATO *et al.*, 2017). More results on CC and closely related problems can be found in Levorato e Frota (2017) and references therein.

Regarding the vertex-index for balance, the corresponding optimization problem is called *Maximum Balanced Subgraph Problem* (MBS), which consists in finding the largest subset of vertices inducing a balanced subgraph. It means deleting the minimum number of vertices so as to turn the graph balanced. Gülpinar *et al.* (2004) showed that MBS is equivalent to the problem of detecting a maximum embedded reflected network matrix (DMERN), and then it is an NP-hard problem due to a general result by Bartholdi (1982). Gülpinar *et al.* (2004) presented a heuristic approach for the problem that is able to find an optimal solution when the entire graph is balanced. Poljak e Turzík (1987) and Barahona e Mahjoub (1989) studied the balanced induced subgraph polytope, i.e. the convex hull of the incidence vectors of all vertex sets that induce balanced subgraphs. The relation between MBS and DMERN was further studied by Figueiredo *et al.* (2011) in terms of polyhedral results. The authors also presented the

first branch-and-cut algorithm for MBS which was improved by Figueiredo e Frota (2013) in the sequel. Figueiredo e Frota (2014) discussed pre-processing routines and heuristic approaches, including a GRASP metaheuristic, for MBS. The generalization of this problem to k-balance, to be denoted k-MBS, has been considered recently by Figueiredo *et al.* (2019). They propose an ILP representatives formulation along with a partial description of the associated polytope. The authors also present a branch-and-cut algorithm that uses an ILS metaheuristic for providing primal bounds.

Although problems such as CC and k-MBS provide solutions to creation of groups in collaborative vs. conflicting environments, they do not take into account individual technical competences, a factor recognized as fundamental in work teams design (FITZPATRICK *et al.*, 2001; TSENG *et al.*, 2004; HLAOITTINUN *et al.*, 2007). Actually, some previous works have addressed team formation problems that simultaneously consider skill requirements and social relationships. In particular, Gutiérrez *et al.* (2016) and Campêlo *et al.* (2020) represent the pairwise relationships between individuals by $0, \pm 1$-sociometric matrices, which can be seen as adjacency matrices of signed graphs. However, these works adopt the sum of the weights of the edges inside the groups as a measure of quality of the formed teams. They do not care about the links between different teams. This is the case of the problems studied in Chapter 2. Only in a very recent research, presented by Kouvatis *et al.* (2020), the structural balance theory was used to create work teams while considering skill requirements.

Following the pioneering work of Lappas *et al.* (2009), which introduced the team formation problem using social networks, Kouvatis *et al.* (2020) defined the *Team Formation in Signed Networks* (TFSN). Given a signed graph representing the social network of the considered skilled individuals, a compatibility (binary) relation based on this graph, and a set of skills to be covered, the TFSN consists in finding a subset of pairwise compatible individuals that collectively cover these skills and minimize a certain communication cost function. The authors considered several versions of TFSN by defining different compatibility relations. Basically, the parity of negative edges in the shortest paths between a pair of individuals is used to define their compatibility. Kouvatis *et al.* (2020) then provided heuristic algorithms for TFSN.

In the same vein as Kouvatis *et al.* (2020), we can think of introducing skill requirements in CC or k-MBS problems. In the case of CC, the resulting problem could be modeled in a similar way to the MMTFP, as studied in Chapter 2. It is essentially a matter of changing the objective function. On the other hand, the problem studied in this chapter can be seen as the

incluision of skill requirements in k-MBS.

3.3 Notation and problem definition

Let $G = (V, E, sign)$ be a *signed graph* where V is the set of n vertices, E is the set of edges (unordered pairs of vertices), and $sign : E \to \{+, -\}$ is a function that associates a positive or negative sign to each edge. For explicit reference to the graph, we may use $V(G)$, $E(G)$, $sign_G$ and $n(G)$. In this text, it is assumed that graph G is simple (has no loops or parallel edges). An edge $(u, v) \in E$ is called negative if $sign(u, v) = -$ and positive if $sign(u, v) = +$. The subsets of negative and positive edges in G are E^- and E^+, respectively. Therefore, $E = E^+ \cup E^-$ and $E^- \cap E^+ = \emptyset$. The spanning subgraphs with all negative edges and all positive edges respectively are $G^- = (V, E^-)$ and $G^+ = (V, E^+)$. For two subsets of vertices $W, W' \subseteq V$, let $E[W : W'] = \{(u, v) \in E \mid u \in W, v \in W'\}$ denote the subset of edges linking a vertex from W to a vertex from W'. In particular, we simply use $E[W] = E[W : W]$ to denote the subset of edges induced by W. The subgraph induced by W is $G[W] = (W, E[W])$. We similarly define $E^-[W, W']$ and $E^+[W, W']$.

The input of the *Competitive Teams Formation Problem* (CTFP) is a tuple $\vartheta = (G, S, s, T, t)$ where:

- $G = (V, E, sign)$ is a signed graph, whose vertices and edges represent individuals and their pairwise relationships, respectively;
- S is a skill set with cardinality $|S| = f$;
- s denotes a point-to-set function $s : V \to 2^S$ which returns the skill set $s(u) \subseteq S$ of each individual $u \in V$;
- T is a set of m projects (or teams) - each project is assigned to a work team and vice-versa;
- t is a demand function $t : T \times S \to \mathbb{N}$ that specifies the number $t(j, s)$ of individuals required for each team $j \in T$ and skill $s \in S$.

For simplicity of the notation, we will also use letter s to express a function $s : T \to S$ that informs the skill subset $s(j) \subseteq S$ demanded by team $j \in T$, that is, $s(j) = \{s \in S : t(j, s) > 0\}$. Besides, we denote by $s(u, j) = s(u) \cap s(j)$ the set of skills of individual $u \in V$ that are demanded by team $j \in T$.

We define the *Competitive Teams Formation Problem* (CTFP) as the problem of finding $|T|$ vertex-disjoint subgraphs in $G = (V, E, sign)$, each one representing a team, with the maximum number of vertices, under the following conditions. The subgraph associated with

team $j \in T$ must contain at least $t(j,s)$ individuals exercising skill s, for every $s \in S$. Besides being allocated to at most one team, each individual $u \in V$ may exercise at most one of his/her skills. The teams must be $|T|$-balanced, i.e. it is mandatory to exists only positive relationships between individuals of the same team and only negative relationships between individuals of different teams. In other terms, the CTFP consists in determining a maximum $|T|$-balanced induced subgraph of G where each cluster must satisfy skill requirements.

More formally, given a tuple $\vartheta = (G = (V,E,sign),S,s,T,t)$, the *CTFP* consists in finding a maximum induced subgraph H of G that has a *vertex-partition* $\mathscr{P} = \{\mathscr{P}_j : j \in T\}$, i.e. $\bigcup_{j\in T} \mathscr{P}_j = V(H)$ and $\mathscr{P}_i \cap \mathscr{P}_j = \emptyset$ for $i \neq j$, where:

1. for each $j \in T$, $\mathscr{P}_j$ can be partitioned into $|S|$ disjoint sets $\{\mathscr{P}_j^s : s \in S\}$ such that, for every $s \in S$, $|\mathscr{P}_j^s| \geq t(j,s)$ and $\mathscr{P}_j^s \subseteq \{u \in \mathscr{P}_j : s \in s(u)\}$;
2. H is $|T|$-balanced, i.e., $E[\mathscr{P}_j] \subseteq E^+$, for every $j \in T$, and $E[\mathscr{P}_j : \mathscr{P}_i] \subseteq E^-$, for every $i, j \in T$, $i \neq j$.

It is important to note that, by item (1), it is possible to have $\mathscr{P}_j^s \neq \emptyset$ even if $s \notin s(j)$.

In order to derive valid inequalities for the CTFP, we apply edge-contraction operations on graph G. Such an operation results in a signed multigraph. Since it will be applied iteratively, it will be convenient to define it over multigraphs. Thus, it also applies to G, which can be seen as a multigraph where every edge has unit multiplicity.

Let $\mathscr{G} = (\mathscr{V},\mathscr{E},\rho,\sigma)$ be a signed multigraph where $\mathscr{V}$ is the set of vertices, $\mathscr{E}$ is the multiset of edges, ρ is a function that relates each edge to a pair in $\mathscr{V} \times \mathscr{V}$, and $\sigma : \mathscr{E} \to \{+,-\}$ assigns to each edge a positive sign $(+)$ or negative sign $(-)$. For the sake of convenience, let us assume that each vertex is identified by a set, and $\mathscr{V}$ comprises pairwise disjoint sets, so that we can define, for each $(U,W) \in \mathscr{E}$, $\rho(U,W) = (u,w)$ for some $u \in U$ and $w \in W$. Besides, we can refer to the "union" $U \cup W$ of two vertices $U, W \in \mathscr{V}$.

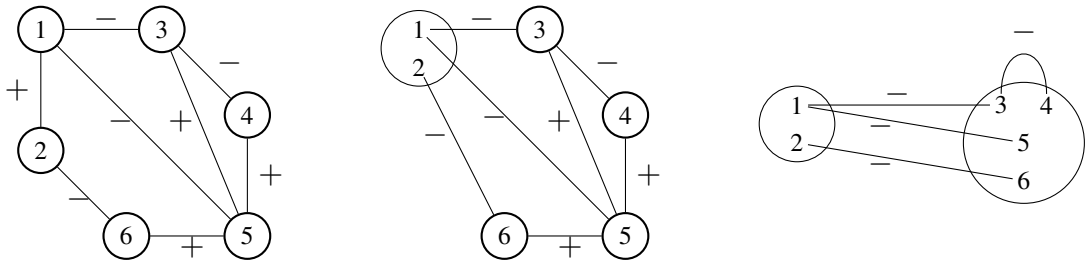
The edge-contraction of a positive edge $e = (U,W) \in \mathscr{E}$, with $U \neq W$, results in another signed multigraph $\mathscr{G} \parallel e = (\mathscr{V}',\mathscr{E}',\rho',\sigma')$, where vertices U and W are replaced by a single vertex $U \cup W$, which is made adjacent to all neighbors of U and W in $\mathscr{G}$, thus defining (possibly parallel) edges that keep their signs. Besides, positive edges between U and W are discarded whereas negative edges between U and W, if any, are kept as negative loops in $U \cup W$. See Figure 12. Precisely,

$$\mathscr{V}' \;=\; (\mathscr{V} \setminus \{U,W\}) \cup \{U \cup W\}$$

and $\mathscr{E}'$ comprises the following edges:

- $e' = e$ with $\rho'(e') = \rho(e)$ and $\sigma'(e') = \sigma(e)$, for all $e = (X,Y) \in \mathscr{E}$ such that $\{X,Y\} \cap \{U,W\} = \emptyset$;
- $e' = (U \cup W, X)$ with $\rho'(e') = \rho(e)$ and $\sigma'(e') = \sigma(e)$, for all $e = (X,Y) \in \mathscr{E}$ such that $\{X,Y\} \cap \{U,W\} = \{X\}$;
- $e' = (U \cup W, U \cup W)$ with $\rho'(e') = \rho(e)$ and $\sigma'(e') = \sigma(e)$, for all $e = (X,Y) \in \mathscr{E}$ such that $\{X,Y\} = \{U,W\}$ and $\sigma(e) = -$.

In particular, let us denote by $\mathscr{G}^- = (\mathscr{V}, \mathscr{E}, \rho, \sigma)$ the multigraph obtained from $G = (V,E,sign)$ by iteratively contracting all positive edges. It will be called *total positive contraction* of G. Note that every vertex in $\mathscr{G}^-$ is defined by a connected component of G^+ whereas every edge $e' = (U,W) \in \mathscr{E}$ corresponds to a negative edge $e = (u,w) \in E^-$ with $u \in U$, $w \in W$, and vice-versa. Besides, $\rho(e')$ keeps the original endpoints of e. Roughly speaking, $\mathscr{G}^-$ is a contraction of the connected components of G^+ into super-vertices.



a) Graph $\mathscr{G}$ *b*) Graph $\mathscr{G} \parallel (\{1\},\{2\})$ *c*) Graph $\mathscr{G}^-$

Figure 12 – Edge-contraction operation.

3.4 Structural properties on balance

Let $\mathscr{G}$ be a multigraph. An *independent set* of $\mathscr{G}$ is a subset of pairwise non-adjacent vertices. In particular, a vertex with a loop does not belong to any independent as it is adjacent to itself. Given a positive integer k, $\mathscr{G}$ is said to be *k-partite* if its vertex set can be partitioned into k disjoint independent sets, some of them possibly empty. Let $\alpha_k(\mathscr{G})$ be the maximum cardinality of a subset of vertices that induces a k-partite subgraph of $\mathscr{G}$. A 2-partite graph is also called *bipartite*. A *k-coloring* of $\mathscr{G}$ is an assignment of a label (color) from $\{1,\ldots,k\}$ to each vertex such that adjacent vertices receive different colors. If $\mathscr{G}$ has a k-coloring, it is said *k-colorable*. Since the vertices assigned a same color form an independent set, k-partite and k-colorable are equivalent concepts. Let us observe that a multigraph with a loop is not k-colorable for any k. On the other hand, the existence of parallel edges makes no difference in the above definitions,

which are directly applied to simple graphs. For more details see West (2000).

A *path* in $\mathscr{G}$ is a subgraph with vertices $U_1, U_1, \ldots, U_p$ and non-parallel edges U_iU_{i+1}, for all $i \in \{1, \ldots, p-1\}$, for some positive integer p. We may refer to it as a U_1U_p-path in order to make its extreme vertices explicit. Similarly, we define a *cycle* in $\mathscr{G}$ by additionally requiring an edge U_1U_p. The cycle is odd if its size p is odd. In particular, a *loop* is an odd cycle with unit size, that is, a vertex U and an edge UU.

Let G be a signed graph. An *odd negative cycle* in G is a cycle with an odd number of negative edges. Such a structure plays a fundamental role in the characterization of balance of G, as already anticipated in the previous section. It is important to observe the correspondence between negative odd cycles in G and odd cycles in $\mathscr{G}^-$, when $\mathscr{G}^-$ is loopless. Besides, a loop in $\mathscr{G}^-$ corresponds to a negative cycle in G with exactly one negative edge. Such correspondences allow us to restate important structural results on balance of G in terms of $\mathscr{G}^-$.

Proposição 3.4.1. *Let $G = (V, E, sign)$ be a signed graph and $\mathscr{G}^-$ be the corresponding total positive contraction. The following assertions are equivalent:*

1. *G is clusterable, i.e. V has a partition $V_1, \ldots, V_k$, for some k, such that $E[V_i] \subseteq E^+$ and $E[V_i : V_j] \subseteq E^-$ for all $i, j \in \{1, \ldots, k\}$, $i \neq j$;*
2. *G has no odd negative cycle with exactly one negative edge;*
3. *$\mathscr{G}^-$ has no loop.*

Proposição 3.4.2. *Let $G = (V, E, sign)$ be a signed graph and $\mathscr{G}^-$ be the corresponding total positive contraction. The following assertions are equivalent:*

1. *G is balanced, i.e. V has a bipartition V_1, V_2 such that $E[V_1] \cup E[V_2] \subseteq E^+$ and $E[V_1 : V_2] \subseteq E^-$;*
2. *G has no odd negative cycle;*
3. *$\mathscr{G}^-$ is bipartite.*

In Proposition 3.4.1 and 3.4.2, the equivalences between (1) and (2) were proved by Davis (1967) and Harary (1953), respectively. The equivalence with (3) is a direct consequence of the aforementioned correspondence between negative odd cycles in G and odd cycles in $\mathscr{G}^-$. Since the construction of $\mathscr{G}^-$ and checking the existence of loops can be done in polynomial time, Proposition 3.4.1 implies that the recognition of clusterable signed graphs is a polynomial problem. Similarly, Proposition 3.4.2 and the polynomial recognition of bipartite multigraphs

imply polynomial time complexity for deciding whether a signed graph is balanced. Actually, a linear time algorithm for this purpose was presented by Harary e Kabell (1980).

The characterization of balance of G in terms of $\mathscr{G}^-$ can be generalized for k-balance as follows.

Proposição 3.4.3. *Let $G = (V, E, sign)$ be a signed graph and $\mathscr{G}^-$ be the corresponding total positive contraction. Let $k \geq 2$ be an integer. The following assertions are equivalent:*

1. *G is k-balanced, i.e. V has a partition $V_1, \ldots, V_k$ such that $E[V_i] \subseteq E^+$ and $E[V_i : V_j] \subseteq E^-$ for all $i, j \in \{1, \ldots, k\}$, $i \neq j$;*
2. *$\mathscr{G}^-$ is k-partite.*

Proof. In the construction of $\mathscr{G}^-$, all positive edges are contracted whereas all negative edges are kept. Then, an independent set in $\mathscr{G}^-$ corresponds to a subset of vertices V' of G such that $E[V'] \subseteq E^+$, and there is only negative edges between any two of these sets. It follows that a k-partition of $\mathscr{G}^-$ into independent sets corresponds to a k-partition of G into $V_1, \ldots, V_k \subseteq V$ such that $E[V_i] \subseteq E^+$ and $E[V_i : V_j] \subseteq E^-$ for all $i, j \in \{1, \ldots, k\}$, $i \neq j$. □

Although the above characterization is immediate, we could not find it in the literature. Besides, it is worth observing that the graphs obtained from total positive contraction of signed graphs are not of a particular subclass (disregarding edge multiplicity which are irrelevant for k-partiteness). Indeed, $G = \mathscr{G}^-$ whenever $E^+ = \emptyset$. Thus, k-balance and k-partiteness carry the same issues in a certain way. In particular, it follows that deciding whether a signed graph is k-balanced, for any fixed $k \geq 3$, is an NP-complete problem, since checking whether a multigraph is k-colorable is NP-complete, for any $k \geq 3$ (STOCKMEYER, 1973; GAREY *et al.*, 1976).

Optimization versions of the decision problems mentioned above can be defined when searching for a k-balanced induced subgraph with maximum number of vertices. This problem has been studied for $k = 2$, as the Maximum Balanced Subgraph Problem (MBS), and for $k \geq 2$, as the k-MBS. In general, we can state the following complexity result.

Proposição 3.4.4. *For any $k \geq 2$, the Maximum k-Balanced Subgraph Problem (k-MBS), i.e. determining the k-balanced induced subgraph with the maximum number of vertices, is NP-hard.*

Proof. For $k = 2$, the problem is equivalent to DMERN (GÜLPINAR *et al.*, 2004) and so NP-hard (BARTHOLDI, 1982). For $k \geq 3$, it generalizes the problem of determining the maximum k-partite induced subgraph, which is an NP-hard problem (STOCKMEYER, 1973). Indeed, the latter problem is a special case where $G = \mathscr{G}^-$. □

The k-MBS can be viewed as a special case of our goal problem. This implies the complexity of the CTFP.

Corolário 3.4.1. *For any number of teams, the Competitive Teams Formation Problem (CTFP) is NP-hard.*

Proof. Consider the case of the CTFP where every individual has all skills and the function t is null. This makes condition (1) of the CFTP definition innocuous, and so the problem results in just determining a maximum k-balanced induced subgraph. □

3.5 Polyhedral results on balance

For any $U \subseteq V$, let x^U denote the incidence vector of U, i.e. the $|V|$-dimensional binary vector such that $x^U_v = 1$ if, and only if, $v \in U$. For each integer $k \geq 2$, we define the k-balanced subgraph polytope of a signed graph G:

$$\mathscr{P}_k(G) = \text{conv}\{x^U \in \{0,1\}^{|V|} : U \subseteq V, G[U] \text{ is } k\text{-balanced}\}.$$

Then, the maximum number of vertices of a k-balanced induced subgraph of G, to be denoted $r_k(G)$, is

$$r_k(G) = \max\left\{\sum_{v\in V} x_v : x \in \mathscr{P}_k(G)\right\}.$$

For any vector x indexed by V and any $U \subseteq V$, let $x(U) = \sum_{v\in U} x_v$. Possible valid inequalities for $\mathscr{P}_k(G)$ are

$$x(U) \leq r_k(G[U]) \quad \forall U \subseteq V.$$

This type of inequality is called *rank inequality* (BARAHONA; MAHJOUB, 1989). Let us observe that the inequality remains valid (although possibly weaker) if $G[U]$ is replaced by a non-induced subgraph of G with vertex set U.

For $k = 2$, polytope $\mathscr{P}_k(G)$ was studied by Poljak e Turzík (1987), Barahona e Mahjoub (1989), Figueiredo *et al.* (2011). They derived rank inequalities related to special subgraphs and presented facet-defining conditions. We could not find polyhedral results for $P_k(G)$, $k \geq 3$, in the literature. However, Figueiredo *et al.* (2019) have recently studied a closely related polytope, namely the polytope associated with the representatives formulation that they propose for k-MBS. As $P_k(G)$ can been seen as a relaxation of that one related to CTFP, in this

section we derive valid inequalities for $\mathscr{P}_k(G)$. They will be used later to get valid inequalities for the CTFP.

3.5.1 Rank inequalities

Here, we generalize for $\mathscr{P}_k(G)$ some of the inequalities proposed for $\mathscr{P}_2(G)$. They can be seen as counterparts of valid inequalities presented for the representatives formulation of the k-MBS (FIGUEIREDO *et al.*, 2019). We start by considering rank inequalities defined by odd negative cycles.

Proposição 3.5.1. *Let C be the set of vertices of an odd negative cycle of G. The following inequality is valid for $\mathscr{P}_2(G)$:*

$$x(C) \leq |C| - 1. \tag{3.1}$$

Besides, it is valid for $\mathscr{P}_k(G)$, for every $k \geq 2$, if the cycle has only one negative edge.

Proposition 3.5.1 is a direct consequence of Propositions 3.4.1 and 3.4.2. For $k = 2$, Barahona e Mahjoub (1989) showed necessary and sufficient conditions for inequality (3.1) to be facet-defining. See also Figueiredo *et al.* (2011).

A *clique* of a (multi)graph is a subset of vertices such that every two distinct vertices are adjacent. A *negative clique* of a signed graph G is a clique K such that $E[K] \subseteq E^-$. This structure also produces valid inequalities.

Proposição 3.5.2. *Let $k \geq 2$ be an integer and K be a negative clique of G with $|K| \geq k+1$. The following inequality is valid for $\mathscr{P}_k(G)$:*

$$x(K) \leq k. \tag{3.2}$$

The validity of inequality (3.2) follows by the fact that at most k vertices of any clique can belong to a k-partite graph. For $k = 2$, Figueiredo *et al.* (2011) showed necessary and sufficient conditions for (3.2) to be facet-defining. Besides, when $G = G^-$, Barahona e Mahjoub (1989) proved that inequalities (3.2) comprise all facet-defining inequalities of $\mathscr{P}_2(G)$ with integer coefficients and right-hand side 2.

Figueiredo *et al.* (2011) presented other valid inequalities for $\mathscr{P}_2(G)$ based on cliques with both positive and negative edges. Precisely, they consider cliques K where $E^-[K]$ induces either two disjoint subcliques, an odd hole or an odd anti-hole. These inequalities could

also be generalized for $\mathscr{P}_k(G)$. However, the last two types of structures do not seem to be usual in signed graphs representing social networks. So, we only deal with the first type here. For, we introduce the following concept.

Given an integer $p \geq 1$, a clique K in a signed graph G is *p-partitionable* if there exists a partition $K_1, K_2, \ldots, K_p$ of K such that $K_i \neq \emptyset$, $E[K_i] \subset E^-$ and $E[K_i : K_j] \subset E^+$, for all $i, j \in \{1, \ldots, p\}$, $i \neq j$. For $p \leq 2$, the concept was presented by Figueiredo *et al.* (2011), and the clique is simply called partitionable. In particular, note that a partitionable clique for $p = 1$ is a negative clique.

Proposição 3.5.3. *Let $k \geq 2$, $p \geq 1$, and K be a p-partitionable clique of G with $|K| > \kappa$, where $\kappa = \max\{p, \min\{\ell, k\}\}$ and $\ell \geq 1$ is the size of the largest part. The following inequality is valid for $\mathscr{P}_k(G)$:*

$$x(K) \leq \kappa. \tag{3.3}$$

Proof. Since K is p-partitionable, it is not possible to choose two vertices from one part and another vertex from a different part, otherwise they would induce an odd negative cycle. Therefore, the vertices chosen from K either belong to a same part or each one lies on a different part. In the first case, we can take at most $\min\{\ell, k\}$ vertices because of the negative edges within a part. In the second case, no more than p vertices can be taken. In any case, the number of vertices from K in any k-balanced graph is at most $\kappa = \max\{p, \min\{\ell, k\}\}$. □

Some remarks regarding inequality (3.3) are worthwhile. If $p = 1$, then $\ell = |K| > \kappa$ and so $\kappa = \min\{|K|, k\} = k$, which implies that (3.3) is exactly (3.2). If $p = k = 2$, we have $\kappa = \max\{2, \min\{\ell, 2\}\} = 2$. In this case, (3.3) becomes $x(K) \leq 2$, which is the inequality presented by Figueiredo *et al.* (2011). It is shown to be facet-defining for $\mathscr{P}_2(G)$ under some conditions. In general, if $p \leq k \leq \ell$, then $\kappa = k$, and (3.3) equals (3.2) (if $p = 1$) or dominates (3.2) (if $p \geq 2$). On the other hand, (3.3) is dominated by the same kind of inequality but related to a smaller clique in the following cases:

- $p > \min\{\ell, k\}$: in this case, $\kappa = p$. Given the partition $K_1, \ldots, K_p$ of K, let $C^i = K \setminus K_i$, for all $i \in \{1, \ldots, p\}$. Inequality (3.3) for C^i is $x(C^i) \leq \kappa_i$, where $\kappa_i = \max\{p-1, \min\{\ell_i, k\}\}$ and $\ell_i = \max\{|K_j| : j = 1, \ldots, p, j \neq i\} \leq \ell$. Since $p - 1 \geq \min\{\ell, k\}$, we have $\kappa_i = p - 1$. Summing up inequalities $x(C^i) \leq p - 1$, for all $i \in \{1, \ldots, p\}$, we get $\sum_{i=1}^{p} x(C^i) \leq \kappa(p-1)$. Since $\sum_{i=1}^{p} x(C^i) = (p-1)x(K)$, we conclude that $x(K) \leq \kappa$ is dominated.

– $p < \ell \leq k$: now, $\kappa = \ell \geq 2$. Let $K_1, \ldots, K_p$ be the partition of K. W.l.o.g. suppose that $|K_j| = \ell$ if and only if $j \in \{1, \ldots, r\}$. Let $C^1, C^2, \ldots, C^{\ell^r}$ be all possible cliques obtained from K by removing exactly one vertex from K_j, for every $j \in \{1, \ldots, r\}$. Each clique C^i is still p-partitionable with partition $\{K'_j = K_j \cap C^i : j = 1, \ldots, p\}$ where $|K'_j| = \ell - 1$ for all $j \in \{1, \ldots, r\}$ and $|K'_j| \leq \ell - 1$ for all $j \in \{r+1, \ldots, p\}$. Then, inequality (3.3) for C^i is $x(C^i) \leq \kappa_i$, where $\kappa_i = \max\{p, \min\{\ell - 1, k\}\}$. Since $p \leq \ell - 1 < k$, we have $\kappa_i = \ell - 1$. Summing up inequalities $x(C^i) \leq \ell - 1$, for all $i \in \{1, \ldots, \ell^r\}$, we get $\sum_{i=1}^{\ell^r} x(C^i) \leq \ell^r(\ell - 1)$. Notice a vertex from K_j appears in all these inequalities, if $j > r$, and in exactly $(\ell - 1)\ell^{r-1}$ of them, if $j \leq r$. Thus, $\sum_{i=1}^{\ell^r} x(C^i) \geq (\ell - 1)\ell^{r-1} x(K)$. Therefore, the obtained inequality dominates $(\ell - 1)\ell^{r-1} x(K) \leq \ell^r(\ell - 1)$ or equivalently dominates $x(K) \leq \ell = \kappa$.

3.5.2 Valid inequalities based on edge-contractions

Instead of looking at particular structures in G and probably get a few strong inequalities, we could think of generating a bunch of (possibly weaker) valid inequalities for $\mathscr{P}_k(G)$ that could be translated into constraints for the CTFP. Besides, we could search for interesting structures in a smaller but related graph, for instance the graphs obtained after contraction of positive edges. The following results ground this strategy.

Let $U \subseteq V$. A set of vertices $C \subseteq V \setminus U$ is a *U-connecting set* if $G^+[U \cup C]$ is connected. Besides, it is a *minimal U-connecting set* if $G^+[U \cup (C \setminus v)]$ is not connected for every $v \in C$. Now, assume that U is a vertex of a multigraph obtained from G by a sequence of positive edge contractions. It follows that $G^+[U]$ is connected. Therefore, for any $U' \subseteq U$, there is always $C \subseteq U \setminus U'$ such that C is a minimal U'-connecting set. Actually, a minimal U'-connecting set C comprises the Steiner vertices of a Steiner tree in $G^+[U]$ with U' being the terminals. In particular, when $U' = \{u, v\}$, C is an induced uv-path in $G^+[U]$.

Lemma 2. *Let $\mathscr{G} = (V(\mathscr{G}), E(\mathscr{G}), \rho, \sigma)$ be a multigraph obtained after a sequence of positive edge contractions in $G = (V(G), E(G), sign)$ and the removal of the remaining positive edges. Let $\mathscr{H}$ be a subgraph of $\mathscr{G}$ with $V(\mathscr{H}) = \{U_1, U_2, \ldots, U_p\}$. Let $T_{\mathscr{H}} = \bigcup_{e \in E(\mathscr{H})} \rho(e)$, and $T_i = T_{\mathscr{H}} \cap U_i$ for all $i = 1, \ldots, p$. Let $S_i \subset U_i \setminus T_i$ be a T_i-connecting set, and $S_{\mathscr{H}} = \bigcup_{i=1}^{p} S_i$. Let $k \geq 2$. The following inequality is valid for $\mathscr{P}_k(G)$:*

$$x(T_{\mathscr{H}} \cup S_{\mathscr{H}}) \leq \alpha_k(\mathscr{H}) - |V(\mathscr{H})| + |T_{\mathscr{H}} \cup S_{\mathscr{H}}|. \tag{3.4}$$

Recall that each vertex U_i of $\mathscr{H}$ represents a subset of vertices of G (endpoints

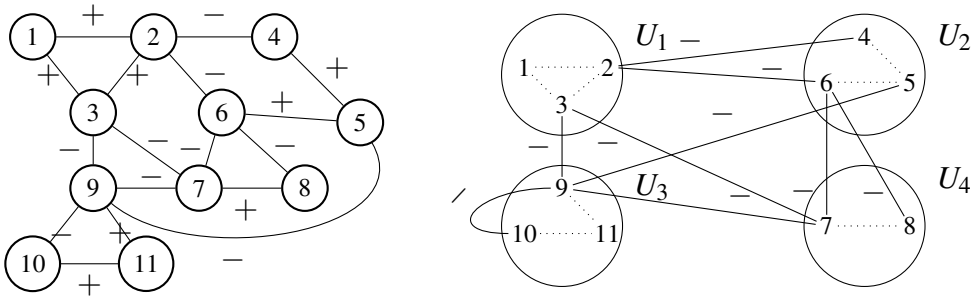
of contracted positive edges). Thus, $T_{\mathscr{H}}$ is exactly the endpoints of the negative edges of G containing in $\mathscr{H}$, and T_i are those endpoints in U_i. Besides, $S_i \cup T_i \subseteq U_i$ are endpoints of positive edges whose contraction contributes to form U_i. See Figures 13 and 14 for an illustration of Lemma 2. Figure 13(a) and 13(b) present graph G and a possible contracted graph $\mathscr{G}$, respectively. Just to make the visualization easier, in Figure 13(b) we keep the contracted positive edges as dashed lines. Figure 14 presents two valid inequalities based on two different subgraphs $\mathscr{H}$ of $\mathscr{G}$. Actually, both subgraphs have the same vertex set as $\mathscr{G}$ but different edge sets. It is worth observing that $\mathscr{H}$ does not need to be an induced subgraph of $\mathscr{G}$ in the statement of the lemma. In particular, we can get stronger inequalities by removing parallel edges in $\mathscr{H}$, since it does not modify $\alpha_k(\mathscr{H})$ but may reduce $|T_{\mathscr{H}} \cup S_{\mathscr{H}}|$. Similarly, we should prefer minimal T_i-connecting sets in order to reduce $|S_{\mathscr{H}}|$. The inequality in Figure 14(a) dominates that one in Figure 14(b). Let us observe that (3.4) can only be useful if $\alpha_k(\mathscr{H}) < |V(\mathscr{H})|$.

Proof. First, notice that the family $\{T_1, \ldots, T_p, S_1, \ldots, S_p\}$ is formed by disjoint subsets of vertices of G. Besides, $\{T_i : i = 1, \ldots, p\}$ and $\{S_i : i = 1, \ldots, p\}$ are partitions of $T_{\mathscr{H}}$ and $S_{\mathscr{H}}$, respectively. Let $R_i = T_i \cup S_i$, for all $i \in \{1, \ldots, p\}$, and $R = T_{\mathscr{H}} \cup S_{\mathscr{H}}$.

Let $\mathscr{G}'$ be the multigraph obtained from G by contracting all edges in $E^+[R]$ that were also contracted to get $\mathscr{G}$ and then by the removal of the remaining positive edges. Since $G^+[R_i]$ is connected, R_i is a vertex of $\mathscr{G}'$. Let $\mathscr{H}'$ be the subgraph of $\mathscr{G}'$ with vertex set $V(\mathscr{H}') = \{R_1, \ldots, R_p\}$ and edge set corresponding to the (negative) edges of $\mathscr{H}$, i.e. $E(\mathscr{H}') = \{e' \in E(\mathscr{G}') : \sigma(e') = \sigma(e), e \in E(\mathscr{H})\}$. Observe that $\mathscr{H}'$ is indeed a subgraph of $\mathscr{G}'$ because $R_i \supset T_i$ and every (negative) edge in $E(\mathscr{H})$ links two vertices of T_i in G. Let $T_{\mathscr{H}'} = \bigcup_{e \in E(\mathscr{H}')} \sigma(e)$. The definition of $V(\mathscr{H}')$ and $E(\mathscr{H}')$ respectively imply $|V(\mathscr{H}')| = |V(\mathscr{H})| = p$ and $T_{\mathscr{H}'} = T_{\mathscr{H}}$. It follows that $\alpha_k(\mathscr{H}) = \alpha_k(\mathscr{H}')$.

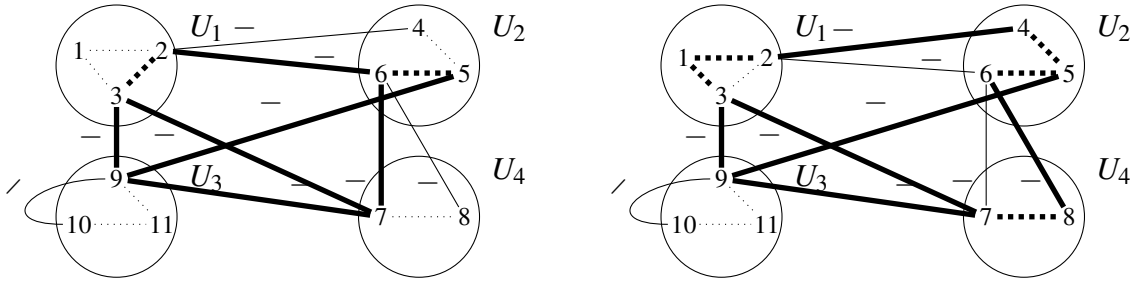
Now, let $\bar{x}$ be an integer point in $\mathscr{P}_k(G)$. Suppose by contradiction that $\bar{x}$ violates (3.4). Therefore, $\bar{x}$ chooses at least $\alpha_k(\mathscr{H}') - p + |R| + 1$ vertices in $R = \bigcup_{i=1}^{p} R_i$. Since $|R| = \sum_{i=1}^{p} |R_i|$, it means at least $\alpha_k(\mathscr{H}') + 1 + \sum_{i=1}^{p} (|R_i| - 1)$ vertices. In other terms, all vertices of at least $\alpha_k(\mathscr{H}') + 1$ parts among $R_1, \ldots, R_p$ are chosen by $\bar{x}$. Since each R_i represents a vertex of $\mathscr{H}'$, the corresponding subgraph taken in $\mathscr{H}'$ is not k-partite. Since $G^+[R_i]$ is connected, for every $i \in \{1, \ldots, p\}$, this means that the subgraph of G induced by the chosen vertices is not k-balanced: a contradiction. □

Inequalities (3.4) generalize (3.1) and (3.2). The contraction of the positive edges of an odd negative cycle C with $|C|$ vertices leads to a negative cycle $\mathscr{C}$ with p vertices, for



a) Graph G *b*) Graph $\mathscr{G} = \mathscr{G}^-$

Figure 13 – Illustration of Lemma 4.



a) $x_2 + x_3 + x_5 + x_6 + x_7 + x_9 \leq 2 - 4 + 6$ *b*) $\sum_{i=1}^{9} x_i \leq 2 - 4 + 9$

Figure 14 – Illustration of Lemma 2. Valid inequalities for $\mathscr{P}_2(G)$ derived from different subgraphs $\mathscr{H}$, whose edges are shown in bold lines. Connecting sets are identified by bold dashed lines.

some odd integer p. Taking $\mathscr{H} = \mathscr{C}$ in Lemma 2, we have $|V(\mathscr{C})| = p$, $|T_{\mathscr{C}} \cup S_{\mathscr{C}}| = |C|$, and $\alpha_k(\mathscr{C}) = p - 1$, for any $k \geq 2$ (if $\mathscr{C}$ is a loop) or for $k = 2$ (if $\mathscr{C}$ is not a loop). In any case, the righthand side of (3.4) is $|C| - 1$. When K is a negative clique of G, then we can take $\mathscr{H} = K$ as a subgraph of $\mathscr{G} = G$. Then, $V(\mathscr{H}) = T_{\mathscr{H}} = K$, $S_{\mathscr{H}} = \emptyset$, and $\alpha_k(\mathscr{H}) = k$. Now, the righthand side of (3.4) is k. Regarding, partitionable clique inequalities (3.3), they are not generally obtainable after contraction operations.

It is worth highlighting that inequalities (3.4) are not always rank inequalities because $\alpha_k(\mathscr{H}) - |V(\mathscr{H})| + |T_{\mathscr{H}} \cup S_{\mathscr{H}}|$ is only an upper bound for $r_k(H)$, where H is the subgraph of G corresponding to the contracted multigraph $\mathscr{H}$. In this sense, using Lemma 2 rather than directly searching for structures of G may lead to weaker inequalities. On the other hand, the contracted graphs can be made smaller and smaller by successive contraction operations, which may reveal interesting structures that could not be found easily otherwise. Since $\alpha_k(\mathscr{H})$ is hard to compute in general, one may favor positive edge contractions in G that lead to structures for which this parameter is known a priori. For instance, $\alpha_k(\mathscr{H}) = k$ if $\mathscr{H}$ is a clique. Even if we restrict ourselves to cliques, the original structures in G may be quite diverse, depending on the connecting sets. In Figures 14(a)-(b), for example, $\mathscr{H}$ is clique that leads to different subgraphs

in G.

3.6 An ILP formulation for the CTFP

Consider the following binary variable:

$$x_{us}^{j} = \begin{cases} 1, \text{if the individual } u \text{ is in team } j \text{ with skill } s, \\ 0, \text{otherwise.} \end{cases} \quad \forall u \in V, \forall j \in T, \forall s \in s(u).$$

An ILP formulation for the CTFP can be defined as follows:

$$(\text{CTF}_x)\ \max \quad \sum_{j \in T} \sum_{u \in V} \sum_{s \in s(u)} x_{us}^{j} \tag{3.5}$$

$$\text{s.t.} \quad \sum_{s \in s(u)} x_{us}^{j} + \sum_{s \in s(v)} x_{vs}^{j} \leq 1, \quad \forall (u,v) \in E^{-}, \forall j \in T \tag{3.6}$$

$$\sum_{s \in s(u)} x_{us}^{i} + \sum_{s \in s(v)} x_{vs}^{j} \leq 1, \quad \forall (u,v) \in E^{+}, \forall i, j \in T, i \neq j \tag{3.7}$$

$$\sum_{j \in T} \sum_{s \in s(u)} x_{us}^{j} \leq 1, \quad \forall u \in V \tag{3.8}$$

$$\sum_{u \in V : s \in s(u)} x_{us}^{j} \geq t(j,s), \quad \forall j \in T, \forall s \in s(j) \tag{3.9}$$

$$x_{us}^{j} \in \{0,1\}, \quad \forall j \in T, \forall u \in V, \forall s \in s(u) \tag{3.10}$$

Constraints (3.6) and (3.7) ensure that teams are balanced. Indeed, constraints (3.6) forbid any two vertices (individuals) linked by a negative edge to be in a same subgraph (team). Similarly, constraints (3.7) ensure that, if the edge between a pair of vertices is positive, then they cannot belong to different subgraphs. Constraints (3.8) state that each individual can be part of at most one team with at most one skill. Constraints (3.9) guarantee that every skill demand is met for each team. Finally, constraints (3.10) determine that the variables are binary. The objective function (3.5) maximizes the total size of the teams.

We remark that constraints (3.7) can be replaced by the following stronger and smaller set of inequalities:

$$\sum_{i \in T : i \neq j} \sum_{s \in s(v)} x_{vs}^{i} + \sum_{s \in s(u)} x_{us}^{j} \leq 1, \quad \forall (u,v) \in E^{+}, \forall j \in T. \tag{3.11}$$

To see the validity of (3.11), first observe that each term in the left handside is at most 1 due to (3.8). Besides, given a pair (u,v) linked by a positive edge, if v is allocated to a team $j \in T$, then u cannot be assigned to a team other than j.

3.6.1 Valid inequalities

The ILP formulation presented above can be strengthened by valid inequalities derived from those presented for $\mathscr{P}_{|T|}(G)$ in Section 3.5. We start with a general statement.

Lemma 3. *If $\sum_{u\in V(G)} \pi_u x_u \leq \pi_0$ is a valid inequality for $\mathscr{P}_{|T|}(G)$ then*

$$\sum_{u\in V(G)} \sum_{j\in T} \sum_{s\in s(u)} \pi_u x^j_{us} \leq \pi_0 \tag{3.12}$$

is valid for (CTF$_x$).

Proof. Let $\bar{x}$ be feasible solution of (CTF$_x$). Let $u \in V(G)$. Define $\bar{y}_u = \sum_{j\in T} \sum_{s\in s(u)} \bar{x}^j_{us}$. By (3.8) and (3.10), $\bar{y}_u \in \{0,1\}$. Actually, $\bar{y}_u$ indicates whether vertex u is chosen or not by $\bar{x}$ (to be allocated to some team with some skill). Besides, $\bar{y}_u = \sum_{j\in T} \bar{y}^j_u$, where $\bar{y}^j_u = \sum_{s\in s(u)} \bar{x}^j_{us} \in \{0.1\}$ indicates whether u belongs to cluster $j \in T$ or not. By (3.6) and (3.7), the subsets $\{u \in V(G) : \bar{y}^j_u = 1\}$, for $j \in T$, define the clusters of a $|T|$-balanced induced subgraph. In other terms, $\bar{y} = [\bar{y}_u]_{u\in V}$ is the characteristic vector of a $|T|$-balanced induced subgraph, and so $\bar{y} \in \mathscr{P}_{|T|}(G)$. Then, $\sum_{u\in V(G)} \pi_u \bar{y}_u \leq \pi_0$ or still $\sum_{u\in V(G)} \sum_{j\in T} \sum_{s\in s(u)} \pi_u \bar{x}^j_{us} \leq \pi_0$. Therefore, (3.12) is valid. □

As a direct consequence of Lemma 3 and Proposition 3.5.3, we get the following valid inequalities for (CTF$_x$).

Corolário 3.6.1. *Let K be a p-partitionable clique of G with $|K| > \max\{p, \min\{\ell, |T|\}\}$, where $\ell \geq 1$ is the size of the largest part. The following p-*partitionable clique inequality *is valid for (CTF$_x$):*

$$\sum_{u\in K} \sum_{j\in T} \sum_{s\in s(u)} x^j_{us} \leq \max\{p, \min\{\ell, |T|\}\}. \tag{3.13}$$

Similarly, Lemmas 2 and 3 lead to valid inequalities based on contraction operations on G.

Lemma 4. *Let $\mathscr{G} = (V(\mathscr{G}), E(\mathscr{G}), \rho, \sigma)$ be a multigraph obtained after a sequence of positive edge contractions in $G = (V(G), E(G), sign)$ and the removal of the remaining positive edges. Let $\mathscr{H}$ be a subgraph of $\mathscr{G}$ with $V(\mathscr{H}) = \{U_1, U_2, \ldots, U_p\}$. Let $T_{\mathscr{H}} = \bigcup_{e\in E(\mathscr{H})} \rho(e)$, and $T_i = T_{\mathscr{H}} \cap U_i$ for all $i = 1, \ldots, p$. Let $S_i \subset U_i \setminus T_i$ be a T_i-connecting set, and $S_{\mathscr{H}} = \bigcup_{i=1}^{p} S_i$. The following inequality is valid for (CTF$_x$):*

$$\sum_{u\in T_{\mathscr{H}} \cup S_{\mathscr{H}}} \sum_{j\in T} \sum_{s\in s(u)} x^j_{us} \leq \alpha_{|T|}(\mathscr{H}) - |V(\mathscr{H})| + |T_{\mathscr{H}} \cup S_{\mathscr{H}}|. \tag{3.14}$$

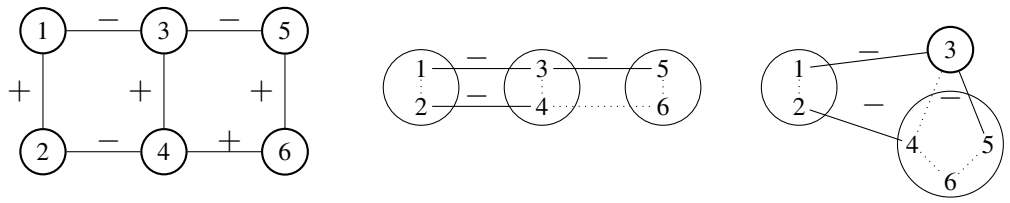
We specialize Lemma 4 for odd cycles and cliques. We start with loops.

Corolário 3.6.2. *Let e be a loop in $\mathscr{G}^-$ with $\rho(e) = \{u,u'\} \subseteq V(G)$. Let P be a uu'-path of G^+. Then, the following* loop inequality *is valid for (CTF_x):*

$$\sum_{u\in V(P)} \sum_{j\in T} \sum_{s\in s(u)} x^j_{us} \leq |V(P)| - 1. \tag{3.15}$$

Proof. In Lemma 4, take $\mathscr{G} = \mathscr{G}^-$ and $\mathscr{H}$ as loop e. Then, $T_{\mathscr{H}} = \{u,u'\}$ and the internal vertices of P are a $T_{\mathscr{H}}$-connecting set. Since $\alpha_{|T|}(\mathscr{H}) = 0$, $|V(\mathscr{H})| = 1$ and $T_{\mathscr{H}} \cup S_{\mathscr{H}} = V(P)$, inequality (3.14) becomes (3.15). □

Inequality (3.15) can be generalized for any odd cycle when $|T| = 2$. It can be expressed in terms of $\mathscr{G}^-$ or any other contracted graph. In this case, we prefer to use $\mathscr{G}^\oslash = (\mathscr{V}^\oslash, \mathscr{E}^\oslash, \rho^\oslash, \sigma^\oslash)$, which stands for a multigraph obtained from $G = (V,E,sign)$ by iteratively contracting all positive edges except for those that would generate a loop, and then just removing the remaining positive edges. Note that $\mathscr{G}^\oslash$ is a loopless multigraph that has all negative edges of G. Let us note that $\mathscr{G}^\oslash$ is not unique and depends on the sequence of edge-contractions. See Figure 15.



a) Graph G *b*) Graph $\mathscr{G}^\oslash$ *c*) Graph $\mathscr{G}^\oslash$

Figure 15 – Illustration of graph $\mathscr{G}^\oslash$ obtained by different sequences of edge-contractions. Dashed lines indicate contracted and removed edges.

Proposição 3.6.1. *Let $\{U_0, U_1, \ldots, U_{p-1}\}$ and $\{e_i = U_iU_{(i+1) \bmod p} : i = 0, \ldots, p-1\}$ be the vertices and edges of an odd cycle in $\mathscr{G}^\oslash$. For every $i \in \{0, \ldots, p-1\}$, let $u_i = \rho(e_i) \cap U_i$, $u'_i = \rho(e_{(i-1) \bmod p}) \cap U_i$, and P_i be a $u_iu'_i$-path of $G^+[U_i]$. The following* odd cycle inequality *is valid for (CTF_x) when $|T| = 2$:*

$$\sum_{i=0}^{p-1} \sum_{u\in V(P_i)} \sum_{j\in T} \sum_{s\in s(u)} x^j_{us} \leq \sum_{i=0}^{p-1} |V(P_i)| - 1. \tag{3.16}$$

Proof. In Lemma 4, take $\mathscr{G} = \mathscr{G}^\oslash$ and $\mathscr{H}$ as the chosen odd cycle. Note that the two edges of the cycle incident to U_i are e_i and $e_{(i-1) \bmod p}$. So, u_i and u'_i are well-defined (vertices in

U_i). Let us remark that it may happen that $u_i = u'_i$, and P_i is simply u_i in this case. Since all vertices of P_i belong to U_i, they do not intercept the vertices of P_j, for $j \neq i$. It follows that $T_{\mathscr{H}} = \bigcup_{i=0}^{p-1} \{u_i, u'_i\}$, $T_{\mathscr{H}} \cup S_{\mathscr{H}} = \bigcup_{i=0}^{p-1} V(P_i)$, and $|T_{\mathscr{H}} \cup S_{\mathscr{H}}| = \sum_{i=0}^{p-1} |V(P_i)|$. Since $|V(\mathscr{H})| = p$ and $\alpha_2(\mathscr{H}) = p-1$, inequality (3.14) becomes (3.16). □

We now specialize inequality (3.14) for cliques in $\mathscr{G}^{\oslash}$.

Proposição 3.6.2. *Let $\mathscr{K} = \{U_1, \ldots, U_p\}$ be a clique of $\mathscr{G}^{\oslash}$ with $p > |T|$, and $\mathscr{E}'$ be a maximal subset of edges of $\mathscr{G}^{\oslash}[\mathscr{K}]$ with no parallel edges. Let $T_{\mathscr{K}} = \bigcup_{e \in \mathscr{E}'} \rho(e)$, and $T_i = T_{\mathscr{K}} \cap U_i$ for all $i = 1, \ldots, p$. Let H_i be a tree of $G^+[U_i]$ that spans T_i, i.e. T_i is contained in $V(H_i)$ and contains every leaf of H_i. The following* contracted clique inequality *is valid for (CTF$_x$):*

$$\sum_{i=1}^{p} \sum_{u \in V(H_i)} \sum_{j \in T} \sum_{s \in s(u)} x_{us}^j \leq |T| + \sum_{i=1}^{p} |E(H_i)|. \tag{3.17}$$

Proof. In Lemma 4, take $\mathscr{G} = \mathscr{G}^{\oslash}$ and $\mathscr{H} = \mathscr{G}^{\oslash}[\mathscr{K}]$. Note that $\mathscr{H}$ is a clique, and so $\alpha_{|T|}(\mathscr{H}) = |T|$. Besides, $T_{\mathscr{H}} = T_{\mathscr{K}}$ and $V(H_i) \setminus T_i$ is a T_i-connecting set, for all $i \in \{1, \ldots, p\}$. Then, $T_{\mathscr{H}} \cup S_{\mathscr{H}} = \bigcup_{i=1}^{p} V(H_i)$. Since the trees H_i are vertex-disjoint, it follows that $|T_{\mathscr{H}} \cup S_{\mathscr{H}}| - |V(\mathscr{H})| = \sum_{i=1}^{p} |V(H_i)| - p = \sum_{i=1}^{p} |E(H_i)|$. Therefore, inequality (3.14) becomes (3.17). □

3.6.2 Generation of inequalities

Each group of valid inequalities presented in section 3.6.1 is potentially exponential in size. So, it is not practicable to include all of them in the ILP model (CTF$_x$). In this work, we chose to include only a subset of them. Below we present the algorithms used to generate each subset of inequalities.

p-Partitionable clique inequalities

First, let us remark that the separation of the p-partitionable clique inequalities is an NP-hard problem, even when $G = G^-$. In this case, we must have $p = 1$. So, condition $|K| > \max\{p, \min\{\ell, |T|\}\}$ in Corollary 3.6.1 translates into $|K| > \min\{|K|, |T|\}$, which leads the righthand side of (3.13) to be $|T|$. Therefore, separating these inequalities from a point $\bar{x} = [\bar{x}_{us}^j]$ is equivalent to finding a maximum vertex-weighted clique in G, where the weight of each vertex $u \in V(G)$ is $\sum_{j \in T} \sum_{s \in s(u)} \bar{x}_{us}^j$.

To generate a subset of inequalities (3.13), we apply the procedure presented in Algorithm 1. We determine (at least) one p-partitionable clique containing each negative edge

and then generate the corresponding inequality, if it satisfies the condition of Corollary 3.6.1. List $\mathscr{L}$, initialized at line 1, keeps the edges not yet covered. To find each part (negative subclique) of the p-partitionable clique, we use the CQL2 heuristic proposed by Nemhauser e Sigismondi (1992).

Algorithm 1: Generation of p-partitionable clique inequalities (3.13)

Result: Subset of p-partitionable clique inequalities
Input : $\vartheta = (G = (V, E, sign), S, s, T, t)$

1 $\mathscr{L} \leftarrow$ list of all edges in $E^-(G)$ (lexicographically ordered);
2 **foreach** $(u, u') \in \mathscr{L}$ **do**
3 $U_1 \leftarrow$ maximal clique containing u, u' in G^- ; `// apply the greedy procedure CLQ2 in` G^-
4 $p \leftarrow 1$;
5 $\ell \leftarrow |U_1|$;
6 $C \leftarrow N^+(U_1)$;
7 **while** $p < |T|$ *and* $C \neq \emptyset$ **do**
8 $p \leftarrow p + 1$;
9 **if** $E^-[C] = \emptyset$ **then**
10 Randomly choose $v \in C$;
11 $U_p \leftarrow \{v\}$;
12 $C \leftarrow \emptyset$;
13 **else**
14 $(v, v') \leftarrow$ lexicographically minimum edge in $E^-[C]$;
15 $U_p \leftarrow$ maximal clique containing v, v' in $G^-[C]$; `// apply the greedy procedure CLQ2 in` $G^-[C]$
16 $C \leftarrow C \cap N^+(U_p)$;
17 **end**
18 $\ell \leftarrow \max\{\ell, |U_P|\}$;
19 **end**
20 **if** $\ell < |T|$ **then**
21 **if** $\ell < p$ **then**
22 Randomly remove $p - \ell$ sets from $\{U_2, \ldots, U_p\}$;
23 **end**
24 **if** $\ell > p$ **then**
25 Randomly remove $\ell - p$ vertices from U_i for all $i \in \{1, \ldots, p\}$ such that $|U_i| > p$;
26 **end**
27 **end**
28 $K \leftarrow \bigcup_{i=1}^{p} U_i$;
29 **if** $|K| > \max\{p, \min\{\ell, |T|\}\}$ **then**
30 Generate inequality (3.13) for K;
31 $\mathscr{L} \leftarrow \mathscr{L} \setminus E^-[K]$;
32 **else**
33 $\mathscr{L} \leftarrow \mathscr{L} \setminus \{(u, u')\}$;
34 **end**
35 **end**

First, we enumerate a maximal (negative) clique U_1 on G^- to be one part of the

p-partitionable clique K (line 3). We initialize the number of negative subcliques in K ($p = 1$) as well as the size of the largest part ($\ell = |U_1|$) (lines 4 and 5). The set C of candidate vertices for the other parts of K is then $N^+(U_1) = \bigcap_{u \in U_1} N^+(u)$, where $N^+(u) = \{v \in V(G) : uv \in E^+(G)\}$ is the set of positive neighbors of u. Similarly, we determine a maximal negative clique in $G^-[C]$ (lines 9-17). Then, we update p (line 8), ℓ (line 18), and the candidate set C to keep only the common positive neighbors of all vertices in all determined subcliques (line 12 or 16). This process is carried out until $p = |T|$ or $C = \emptyset$. The first itemized remark presented after Proposition 3.5.3 alerts that having $p > |T|$ leads to a dominated inequality.

Still based on the remarks on Proposition 3.5.3, we perform a post-processing to adjust the number of parts and the size of the largest one in clique K (lines 29-34). If $\ell \geq |T|$, then we are done because $p \leq |T|$, which ensures that none of the cases enumerated after the proposition occurs. In case $\ell < |T|$, we remove $p - \ell$ parts (if $p > \ell$) in order to reduce p to ℓ, or we remove $\ell - p$ elements (if $\ell > p$) of each part larger that p so as to reduce ℓ to p. In both subcases, ℓ and p turn to be equal. Finally, if the resulting clique K satisfies $|K| > \max\{p, \min\{\ell, |T|\}\}$, the corresponding inequality is generated. Otherwise, the inequality is useless.

Loop inequalities

Instead of using graph $\mathscr{G}^-$, we generate loop inequalities (3.15) as follows. We initially construct subgraph $G^+ = (V, E^+)$ by removing all negative edges from the graph $G = (V, E, sign)$. Then, for each pair of vertices $u, u' \in V(G)$, such that $uu' \in E^-(G)$, we apply Dijkstra's algorithm to calculate a shortest path in G^+ between u and u'. If it exists, this path ensures a loop e in $\mathscr{G}^-$ with $\sigma(e) = \{u, u'\}$. In this case, inequality (3.15) for such a path is added to (CTF$_x$).

It is worth remarking that the separation problem related to the loop inequalities is polynomially solvable. Indeed, we can rewrite (3.15) as

$$\sum_{u \in V(P)} \left(1 - \sum_{j \in T} \sum_{s \in s(u)} x^j_{us} \right) \geq 1.$$

Thus, given a point $\bar{x}$, feasible to the linear relaxation of (CTF$_x$), let

$$w_u = 1 - \sum_{j \in T} \sum_{s \in s(u)} \bar{x}^j_{us}$$

be a weight assigned to each vertex of G. Note that $w_u \geq 0$ due to (3.8). The separation problem consists in solving minimum vertex-weighted paths in G^+ between all pairs (u, u') such that

$uu' \in E^-(G)$. If the minimum weight of all these paths is greater than or equal to 1, all loop inequalities are satisfied by $\bar{x}$. Otherwise, any path with weight less than 1 defines a violated loop inequality.

Multigraph $\mathscr{G}^{\oslash}$

To generate subsets of inequalities (3.16) and (3.17), we first construct multigraph $\mathscr{G}^{\oslash} = (\mathscr{V}^{\oslash}, \mathscr{E}^{\oslash}, \rho^{\oslash}, \sigma^{\oslash})$ by contracting all positive edges, expect for those whose contraction would generate a loop, which are then simply discarded. Recall that, depending on the order of the contractions, we may get different multigraphs. The applied procedure is presented in the Algorithm 2.

Initially, we create an empty multigraph $\mathscr{G}^{\oslash}$ and an ordered list $\mathscr{Q}$ with all vertices of G (lines 1 and 2). List $\mathscr{Q}$ will guide the order of the contractions. At each iteration of the main loop (lines 3–23), a vertex $w \in \mathscr{Q}$ is chosen to form a new vertex U_w in $\mathscr{V}^{\oslash}$. Starting from w, the other vertices still in $\mathscr{Q}$ to be contracted into U_w are taken so as to ensure that there is no negative edge between the vertices belonging to U_w and that the set of vertices in U_w are connected, i.e., there is at least one positive path between any pair of vertices belonging to U_w (condition guaranteed by lines 10 to 21). The vertices in U_w are then removed from $\mathscr{Q}$, and U_w is included in the vertex set $\mathscr{V}^{\oslash}$ of $\mathscr{G}^{\oslash}$. As long as $\mathscr{Q}$ remains non-empty, a new set of vertices is contracted in a similar manner. Lastly, the edges in $\mathscr{E}^{\oslash}$ are created. Note that $\mathscr{G}^{\oslash}$ is a loopless multigraph that has all negative edges of G.

Contracted clique inequalities

As the contracted clique inequalities (3.17) coincide with the p-partitionable clique inequalities (3.13) when $G = G^-$, the corresponding separation problem is also NP-hard.

The subset of contracted clique inequalities (3.17) included in (CTF$_x$) is guven by the procedure presented in Algorithm 3. First, we generate the multigraph $\mathscr{G}^{\oslash}$ with Algorithm 2 (line 1). Then, for each negative edge of G still not covered by a generated contracted clique inequality, we use the greedy procedure CLQ2, proposed by Nemhauser e Sigismondi (1992), to find a maximal (negative) clique on $\mathscr{G}^{\oslash}$ containing this edge (loop 3–12). Let $\mathscr{K} = \{U_1, \ldots, U_p\}$ be the found clique (line 5). If $p \leq |T|$, we discard the clique, since the corresponding inequality is useless. Otherwise, to generate a contracted clique inequality (3.17) related to $\mathscr{K}$, we have to determine the sets of terminal and connecting vertices (line 8). The negative edges covered by the inequality are no longer considered as starting edges for generating other cliques (line 9).

Algorithm 2: Construction of the multigraph $\mathscr{G}^{\oslash} = (\mathscr{V}^{\oslash}, \mathscr{E}^{\oslash}, \rho^{\oslash}, \sigma^{\oslash})$

Result: Multigraph $\mathscr{G}^{\oslash} = (\mathscr{V}^{\oslash}, \mathscr{E}^{\oslash}, \rho^{\oslash}, \sigma^{\oslash})$
Input : Signed graph $G = (V, E, sign)$

1 $\mathscr{Q} \leftarrow$ ordered list with all vertices in $V(G)$;
2 $\mathscr{G}^{\oslash} = (\mathscr{V}^{\oslash}, \mathscr{E}^{\oslash}, \rho^{\oslash}, \sigma^{\oslash}) \leftarrow \emptyset$;
3 **while** $\mathscr{Q} \neq \emptyset$ **do**
4 remove a vertex $w \in \mathscr{Q}$;
5 $U_w \leftarrow U_w \cup \{w\}$;
6 count_neg $\leftarrow 0$;
7 count_pos $\leftarrow 0$;
8 **foreach** $u \in \mathscr{Q}$ **do**
9 **foreach** $v \in U_w$ **do**
10 **if** $uv \in E^-(G)$ **then**
11 count_neg++;
12 **end**
13 **if** $uv \in E^+(G)$ **then**
14 count_pos++;
15 **end**
16 **end**
17 **if** *count_neg* $= 0$ *&& count_pos* > 0 **then**
18 remove vertex u from $\mathscr{Q}$;
19 $U_w \leftarrow U_w \cup \{u\}$;
20 **end**
21 **end**
22 $\mathscr{V}^{\oslash} \leftarrow \mathscr{V}^{\oslash} \cup \{U_w\}$;
23 **end**
24 **foreach** $U_{w_1}, U_{w_2} \in \mathscr{V}^{\oslash} : w_1 \neq w_2$ **do**
25 **foreach** $(u,v) \in U_{w_1} \times U_{w_2}$ **do**
26 **if** $uv \in E^-(G)$ **then**
27 $\mathscr{E}^{\oslash} \leftarrow \mathscr{E}^{\oslash} \cup \{U_{w_1}U_{w_2}\}$; // $\mathscr{E}^{\oslash}$ is a multiset
28 $\rho^{\oslash}(U_{w_1}U_{w_2}) \leftarrow (u,v)$;
29 $\sigma^{\oslash}(U_{w_1}U_{w_2}) \leftarrow -$;
30 **end**
31 **end**
32 **end**

As commented just after Lemma 2, it is desired to have these sets as small as possible. This is the goal of Algorithm 4. For each $i \in \{1, \ldots, p\}$, the terminal vertices in U_i will be stored in U_i^*. First, we initialize $U_i^* = \emptyset$, for every $i \in \{1, \ldots, p\}$ (line 1). Then, for each pair (U_i, U_j), $i < j$, we determine an edge $e \in E^-[\mathscr{K}]$ such that $\rho(e) = (u,v) \in U_i \times U_j$. Whenever possible, we choose vertex u (resp. v) that already belongs to U_i^* (resp. U_j^*) at line 3. Precisely, (u,v) is the lexicographically minimum pair in one of the sets $U_i^* \times U_j^*$, $U_i^* \times (U_j \setminus U_j^*)$, $(U_i \setminus U_i^*) \times U_j^*$

or $(U_i \setminus U_i^*) \times (U_j \setminus U_j^*)$ (considered in this order of priority). We then add u to U_i^* and v to U_j^* (lines 4-5). Edge (u,v) is stored to be considered as covered by the contracted clique inequality (line 6).

At the end of this first loop, we need to connect all the vertices included in each set U_i^*. Thus, we apply a heuristic to find a Steiner tree in $G^+[U_i]$ that contains all vertices in U_i^* (lines 11 to 25). In other words, we find a tree $H_i = (U_i', E_i)$ such that $U_i^* \subseteq U_i' \subseteq U_i$ and $E_i \subseteq E^+[U_i']$. We start with $U_i' = \{u\}$, for some $u \in U_i^*$, $E_i = \emptyset$, and iterate as follows. Every vertex $v \in U_i^* \setminus U_i'$ that directly connects to some vertex $u \in U_i'$ is added to U_i', and the edge uv is added to E_i. If no vertex in $U_i^* \setminus U_i'$ has a neighbor in U_i', we add to U_i' any vertex $v \in U_i \setminus U_i'$ that is adjacent to some vertex $w \in U_i'$, whereas vw is added to E_i. This process is repeated until we get $U_i^* \subseteq U_i'$. At the end, an iterative post-processing phase is performed to remove all vertices in $U_i' \setminus U_i^*$ with degree 1 in H_i. This post-processing is repeated in the reduced tree until there is no more vertex to be removed (lines 21 to 23).

With these trees H_i at hand, the corresponding inequality (3.17) is added to the model if $|\mathscr{K}| > |T|$ (line 10 of Algorithm 3).

Odd cycle inequalities

For the case where $|T| = 2$, we also add a subset of the odd cycle inequalities (3.16). Differently from the other inequalities, these ones are included through a separation routine.

Let $\bar{x} = [\bar{x}_{us}^j] \in \mathbb{R}^{m \times n \times f}$ be a feasible (fractional) solution of linear relaxation of (CTF$_x$). Define a weight $w_u = 1 - \sum_{j \in T} \sum_{s \in s(u)} \bar{x}_{us}^j$ for each vertex $u \in V$. Note that $w_u \geq 0$ due to (3.8). Similarly to the case of the loop inequalities, we can conclude that the separation problem for the odd cycle inequalities consists in checking whether or not there is an *odd negative cycle* in G with vertex-weight less then 1. This problem can be solved in polynomial time (FIGUEIREDO *et al.*, 2011).

Algorithm 3: Generation of contracted clique inequalities (3.17)

Result: Subset of contracted clique inequalities

Input : $\vartheta = (G = (V, E, sign), S, s, T, t)$

$\mathscr{G}^{\oslash} = (\mathscr{V}^{\oslash}, \mathscr{E}^{\oslash}, \rho^{\oslash}, \sigma^{\oslash}) \leftarrow$ apply Algorithm 2 on G;
$\mathscr{L} \leftarrow$ list of all edges in $E^{-}(G)$ (lexicographically ordered);
foreach $(u, u') \in \mathscr{L}$ **do**
    Let $UU' \in \mathscr{E}^{\oslash}$ such that $u \in U$, $u' \in U'$;
    $\mathscr{K} \leftarrow$ clique $\{U_1, \ldots, U_p\} \supseteq \{U, U'\}$ in $\mathscr{G}^{\oslash}$; `// apply the greedy procedure CLQ2 in` $\mathscr{G}^{\oslash}$
    $\mathscr{L} \leftarrow \mathscr{L} \setminus \{(u, u')\}$;
    if $p > |T|$ **then**
        $(E', H_1, \ldots, H_p) \leftarrow$ apply Algorithm 4 on $\mathscr{K}$;
        $\mathscr{L} \leftarrow \mathscr{L} \setminus E'$;
        Generate inequality (3.17) for $(H_1, \ldots, H_p)$;
    end
end

Algorithm 4: Generation of connecting sets for contracted clique

Result: Set E' of negative edges, and set $H_1, \ldots, H_p$ of connecting trees

Input : clique $\mathscr{K} = \{U_1, \ldots, U_p\}$ in $\mathscr{G}^{\oslash}$

$U_1^*, U_2^*, \ldots, U_p^* \leftarrow \emptyset$; `//` U_i^* `stores the terminal vertices in` U_i
$E' \leftarrow \emptyset$;
foreach $U_i, U_j \in \mathscr{K} : i < j$ **do**
    Determine an edge $e \in E^{-}[\mathscr{K}]$ such that $\rho(e) = (u, v) \in U_i \times U_j$ and minimizes $|U_i^* \cup \{u\}| + |U_j^* \cup \{v\}|$;
    $U_i^* \leftarrow U_i^* \cup \{u\}$;
    $U_j^* \leftarrow U_j^* \cup \{v\}$;
    $E' \leftarrow E' \cup \{(u, v)\}$;
end
for $i \leftarrow 1$ *to* p **do**
    if $|U_i^*| > 1$ **then**
        `/* create a tree` $H_i = (U_i', E_i)$ `such that` $U_i^* \subseteq U_i' \subseteq U_i$ `and` $E_i \subseteq E^{+}[U_i']$; `*/`
        $U_i' \leftarrow \{u\}$, for some $u \in U_i^*$; $E_i \leftarrow \emptyset$;
        while $U_i^* \not\subseteq U_i'$ **do**
            if $\exists v \in U_i^* \setminus U_i'$ *such that* $uv \in E^{+}[U_i^*]$ *for some* $u \in U_i'$ **then**
                $U_i' \leftarrow U_i' \cup \{v\}$;
                $U_i^* \leftarrow U_i^* \setminus \{v\}$;
                $E_i \leftarrow E_i \cup \{uv\}$;
            else
                choose any vertex $v \in U_i \setminus U_i'$ such that $vw \in E^{+}[U_i]$ for some $w \in U_i'$;
                $U_i' \leftarrow U_i' \cup \{v\}$;
                $E_i \leftarrow E_i \cup \{vw\}$;
            end
        end
        foreach $u \in U_i' \setminus U_i^*$ *such that* $|\{uv \in E_i : v \in U_i'\}| = 1$ **do**
            $U_i' \leftarrow U_i' \setminus \{u\}$;
        end
    end
end

Instead of exactly solving this problem on G, we apply a separation heuristic on a smaller (contracted) graph. Such a graph, to be denoted $\hat{\mathscr{G}}^{\oslash}$, is obtained by applying Algorithm 2 on the subgraph of G induced by the vertices with strictly positive weight. An odd cycle in $\hat{\mathscr{G}}^{\oslash}$ gives an odd negative cycle in G. To find odd cycles in $\mathscr{G}^{\oslash}$, we create a bipartite graph B as follows. For each vertex $U \in V(\hat{\mathscr{G}}^{\oslash})$, we create to copies U' and U'' in B. Besides, if U and W are adjacent in $\hat{\mathscr{G}}^{\oslash}$, we make U' adjacent to W'' and W' adjacent to U'' in B. For each pair (U',U''), we determine a $U'U''$-shortest path in B, which corresponds to an odd cycle in $\hat{\mathscr{G}}^{\oslash}$. Then, given an odd cycle $\mathscr{C}$ in $\hat{\mathscr{G}}^{\oslash}$ with vertex set $\{U_0,U_1,\ldots,U_{p-1}\}$, we build an odd negative cycle C of G as follows. For each $i \in \{0,\ldots,p-1\}$, we order the vertices in U_i and then search for the edge $e = U_iU_{(i+1) \mod p} \in E(\mathscr{C})$ such that $\rho(e) = (u_i, u'_{(i+1) \mod p}) \in U_i \times U_{(i+1) \mod p}$ is lexicographically minimum, i.e. $u_i = \min\{u \in U_i : uv \in E^+(G) \text{ for some } v \in U_{(i+1) \mod p}\}$ and $u'_{(i+1) \mod p} = \min\{v \in U_{(i+1) \mod p} : u_iv \in E^+(G)\}$. The vertices $\{u_i, u'_i : i = 0,\ldots,p-1\}$ and the negative edges $\{u_iu'_{(i+1) \mod p} : i = 0,\ldots,p-1\}$ will be part of C. Note that u_i and u'_i, which may be distinct or not, are the vertices taken from U_i. If they are distinct vertices, they must be connected through positive edges belonging to $E^+[U_i]$. To find them, we apply Dijkstra's algorithm to calculate the shortest $u_iu'_i$-path in $G^+[U_i]$. The internal vertices of these $u_iu'_i$-paths, $i \in \{0,\ldots,p-1\}$, and their edges will complete cycle C. If violated by $\bar{x}$, the odd cycle inequality (3.16) for C is added to the linear relaxation at the root node. This is done for every vertex U of $\hat{\mathscr{G}}^{\oslash}$.

3.7 Computational experiments and results

We report on computational results for the formulation presented in Section 3.6 with and without the inclusion of valid inequalities presented in section 3.6.1. The algorithms were coded in C language using the CPLEX 12.8 callable library with parallelism enabled and four threads. The experiments were run on a machine equipped with Intel Core i5 Dual Core, 4 $\times$ 1.8Ghz, 8GB of RAM under macOS Catalina, version 10.15.5.

A set of 522 instances were used in the computational experiments. They were obtained from 6 synthetic randomly generated graphs and 23 signed graphs used in Gülpinar *et al.* (2004), available at http://www.cs.rhul.ac.uk/~zvero/thesis/sga_ref/all_models.tar.gz. As we have mentioned before, Gülpinar *et al.* (2004) proposed a heuristic approach based on signed graphs to the solution of the DMERN problem. GÜLPINAR *et al.*'s graphs were also used by Figueiredo *et al.* (2011) as MBS instances.

The synthetic signed graphs were randomly generated by using the triple (n,d,p) as input parameters, where n is the number of vertices (set to 50), the $d \in \{20\%, 30\%\}$ is the probability that an edge exists, and $p \in \{40\%, 50\%, 50\%\}$ is the probability of an existing edge to be positive. Table 19 presents this triple as well as the resulting edge-density, percentage of positive edges (E^+) and percentage of negative edges (E^-) for each generated signed graph. Besides, it shows the number of loops, contracted cliques larger than $|T|$ and p-partitionable cliques that generated inequalities for the model, all found by the algorithms presented in the Section 3.6.2. Similar information is presented in Table 20 for the GÜLPINAR *et al.*'s graphs.

Table 19 – Parameter values for each synthetic instance graph.

Synthetic Signed Graphs											
Name	(n,d,p)	Density (%)	E+ (%)	E- (%)	Loops	Cliques			p-partitionable Cliques		
						|T| = 4	|T| = 3	|T| = 2	|T| = 4	|T| = 3	|T| = 2
S1	(50, 20, 60)	22.36	59.48	40.52	111	42	54	54	9	9	10
S2	(50, 20, 40)	20.08	42.68	57.32	141	49	61	63	7	7	7
S3	(50, 20, 50)	20.80	49.41	50.59	129	31	50	60	4	4	8
S4	(50, 30, 50)	30.28	47.44	52.56	195	71	76	76	5	5	5
S5	(50, 30, 60)	29.22	53.91	46.09	165	56	78	81	9	12	12
S6	(50, 30, 40)	30.69	42.02	57.98	218	73	75	84	8	8	9

Table 20 – Parameter values for each instance graph presented by Gülpinar *et al.* (2004).

Name	Instance	Vertices (n)	Density (%)	E^+ (%)	E^- (%)	Loops	Cliques			p-Partitionable Cliques		
							|T| = 4	|T| = 3	|T| = 2	|T| = 4	|T| = 3	|T| = 2
G1	25fv47	224	0.01	0.85	0.15	25	0	0	8	0	0	0
G2	agg2	141	0.05	0.99	0.01	4	0	0	0	0	0	0
G3	agg3	141	0.05	0.99	0.01	4	0	0	0	0	0	0
G4	bnl1	275	0.01	0.67	0.33	23	0	0	24	0	0	0
G5	fffff800	157	0.05	0.50	0.50	154	0	0	155	0	0	0
G6	maros	305	0.01	0.57	0.43	50	0	29	29	0	1	1
G7	modszk1	148	0.02	0.49	0.51	11	0	0	15	0	0	0
G8	nesm	190	0.01	0.84	0.16	24	0	0	0	0	0	0
G9	perold	150	0.02	0.70	0.30	21	0	7	9	0	0	0
G10	pilot	275	0.01	0.68	0.32	18	0	0	17	0	0	0
G11	pilot.ja	205	0.01	0.63	0.37	32	0	0	29	0	0	0
G12	pilot.we	202	0.01	0.47	0.53	6	0	0	34	0	0	0
G13	pilot87	339	0.01	0.56	0.44	21	0	0	70	0	0	0
G14	pilotnov	209	0.01	0.65	0.35	13	0	0	27	0	0	0
G15	scfxm2	282	0.01	0.64	0.36	72	0	0	80	0	0	0
G16	woodw	301	0.01	1.00	0.00	0	0	0	0	0	0	0
G17	cycle	505	0.00	0.70	0.30	75	0	0	41	0	0	1
G18	ganges	631	0.01	0.54	0.46	353	14	153	244	4	4	4
G19	gfrd-pnc	590	0.01	0.50	0.50	554	53	162	283	4	4	4
G20	sctap2	483	0.01	0.50	0.50	247	142	201	231	0	0	0
G21	seba	456	0.01	0.66	0.34	455	0	209	372	0	6	6
G22	shell	483	0.00	0.94	0.06	20	0	0	0	0	0	0
G23	ship12s	456	0.03	0.86	0.14	247	0	0	94	0	0	12

Let us recall that a CTFP instance is represented by a tuple $\vartheta = (G,S,s,T,t)$. Besides the signed graph $G = (V,E,sign)$, the other data required to create each instance was also randomly generated according to the following patterns. First, we choose the number of skills

($|S| = f$) and teams ($|T| = m$), where $|S| = 5$ and $|T| \in \{2,3,4\}$. Then, each individual $u \in V$ is assigned a set $s(u) \in S = \{1,2,...,f\}$ of skills. We predefined a percentage $\sigma \in \{30\%, 50\%, 70\%\}$ of individuals to have multiple skills. Thus, we classify the instances into three categories: A ($\sigma = 30\%$), B ($\sigma = 50\%$) and C ($\sigma = 70\%$). The multi-skilled individuals were randomly selected from set V. Then, the number of multiple skills (2, 3 or 4 skills) as well as the skills themselves were also randomly chosen. Finally, the values $t(j,s)$ for each $j \in T = \{1,2,\ldots,m\}$ and $s \in S$ were also chosen randomly while ensuring that $t(j,s) \leq \sum_{u \in V : s \in s(u)} 1/m$. Two different functions t were generated.

Based on each graph G, we generated 18 instances, each of them associated with a possible combination of $|T|$, category and function t. Each instance is identified by a sequence of letters and numbers corresponding to: the base graph presented in Table 19 or Table 20; a letter (A, B or C) that matches one of the three categories; an arabic number (1 or 2), referring to one of the two functions t generated; and a roman number (II, III, IV) that matches the value of $|T| \in \{2,3,4\}$. Thus, S3A2II is the instance $\vartheta = (G,S,s,T,t)$, where G is the third graph in Table 19, 30% of the individuals have multiple skills from S, the second demand function t was used, and the number of required teams is $|T| = 2$.

The next 3 tables show the results of computational experiments with the 108 instances based on the synthetic graphs presented in Table 19, grouped by category (A, B and C) and separated by number of requested teams ($|T| \in \{2,3,4\}$). Table 21 compares the basic ILP formulation (3.5)-(3.10) and this formulation with inequality (3.7) replaced by (3.11), identified as ILP-(3.11). We present the average computational time (Time) in seconds and the average number of generated nodes (Nodes) to solve 12 instances of a same category (A, B and C) with a fixed number of requested teams ($|T| \in \{2,3,4\}$).

We can remark that formulation ILP-(3.11) outperforms or is comparable to formulation ILP (3.5)-(3.10) in all test cases. See Table 26 in Appendix B for detailed results. It is also important to note that there is a gradual increase both in spent time and number of nodes as the percentage of multi-skilled individuals increases. Even more expressive is the increase in these figures as the number of requested teams increases.

Table 22 compares similar computational results for formulation ILP-(3.11) and for this formulation with the inclusion of inequalities (3.13), (3.15) together with (3.16) in case $|T| = 2$, and (3.17). We evaluate the inclusion of each set of inequalities separately as well as all of them together. Each line presents the average computational time (Time) in seconds and

Table 21 – Average computational time and number of nodes to solve instances based on synthetic graphs for formulations ILP (3.5) - (3.10) and ILP-(3.11).

Cat.	|T|	ILP (3.5)-(3.10)		ILP-(3.11)	
		Time	Nodes	Time	Nodes
A	2	1.14	1370.17	0.94	964.42
B		1.50	2324.50	1.46	2252.25
C		2.13	3813.42	2.04	3739.75
A	3	16.49	14060.50	14.18	11494.00
B		18.11	17211.17	17.00	14528.25
C		23.21	25279.92	20.80	20172.25
A	4	85.24	60659.50	58.57	31436.08
B		141.34	118959.83	82.91	54885.00
C		320.50	273204.33	138.53	126015.75

average number of generated nodes (Nodes) to solve 12 instances of a same category (A, B and C) with a fixed number of requested teams ($|T| = \{2,3,4\}$). The best values are shown in bold. For more details, see Table 27 in Appendix B.

Table 22 – Average computational time and number of nodes to solve instances generated with synthetic graphs for formulation ILP-(3.11) without and with inclusion of valid inequalities.

Cat.	|T|	ILP-(3.11)		ILP-(3.13)		ILP-(3.15)*		ILP-(3.17)		ILP-ALL	
		Time	Nodes	Time	Nodes	Time	Nodes	Time	Nodes	Time	Nodes
A	2	0.94	964.42	**0.89**	**900.75**	0.90	913.33	0.89	905.25	0.90	1195.67
B		1.88	2835.58	1.64	2697.83	1.47	2355.58	1.77	2591.08	**1.42**	**2342.50**
C		2.20	3864.75	1.96	3385.67	2.14	3525.17	2.04	3507.17	**1.86**	**3134.33**
A	3	15.29	12044.00	14.17	11437.50	13.15	9715.00	14.91	10104.42	**13.11**	**9659.83**
B		17.00	14528.25	16.53	14729.08	16.71	12770.67	17.90	14098.75	**16.48**	**13402.83**
C		20.80	20172.25	17.91	17815.58	18.77	17668.08	20.49	18941.17	**17.51**	**17362.92**
A	4	58.57	31436.08	55.92	29467.08	48.43	24170.67	53.02	30249.58	**48.17**	**23916.75**
B		82.91	54885.00	73.87	44002.17	65.53	32876.25	77.19	53850.25	**64.52**	**32645.08**
C		138.53	126015.75	136.42	101624.25	120.12	83252.83	131.01	103495.00	**114.82**	**64906.67**

Note*. For the instances with $|T| = 2$, inequalities (3.16) were also added.

We can observe that formulation ILP-(3.11) with the addition all sets of inequalities (3.13), (3.15), (3.16) and (3.17) (i.e. ILP-ALL) presents the best performance when compared with formulation ILP-(3.11) restricted by each set inequalities separately. ILP-ALL got the lowest values for both the elapsed time and the number of nodes.

Table 23 presents the percentage integrality gaps between the optimal solution and the solution of the linear relaxation of each ILP formulation. Besides the average percentage gap for the 12 instances related to the same category and number of teams, each line also shows the elapsed time (Time) in seconds to solve the linear relaxation of each formulation ILP. We can observe that formulations ILP-(3.15) and ILP-ALL (in bold) present the smallest gaps when compared to the other formulations, thus showing that inequalities (3.15) provide the main contribution to narrow the gap. The times spent to solve the relaxations are negligible in all cases.

To close this section, we present the results of the computational experiments with

Table 23 – Average computational time and gap for the linear relaxation of the formulation ILP -(3.11) without and with inclusion of valid inequalities

Cat.	|T|	ILP-(3.11)		ILP-(3.13)		ILP-(3.15*)		ILP-(3.17)		ILP-ALL	
		Time	gap (%)	Time	gap (%)	Time	gap (%)	Time	gap (%)	Time	gap (%)
A	2	0.00	96.48	0.01	80.30	**0.05**	**39.68**	0.01	80.85	**0.04**	**39.68**
B		0.00	98.15	0.01	81.26	**0.07**	**38.29**	0.01	80.49	**0.04**	**38.29**
C		0.00	95.03	0.01	81.74	**0.06**	**36.94**	0.01	81.87	**0.03**	**36.94**
A	3	0.05	79.56	0.03	79.42	**0.15**	**30.52**	0.05	72.40	**0.09**	**30.52**
B		0.05	80.04	0.03	74.37	**0.16**	**27.37**	0.05	71.40	**0.10**	**27.37**
C		0.04	77.94	0.03	75.27	**0.15**	**25.67**	0.04	71.93	**0.09**	**25.67**
A	4	0.11	79.14	0.04	79.47	**0.28**	**27.22**	0.10	75.39	**0.19**	**27.22**
B		0.11	76.63	0.06	74.34	**0.25**	**23.56**	0.09	72.16	**0.19**	**23.56**
C		0.09	75.50	0.05	74.10	**0.29**	**22.21**	0.10	72.96	**0.18**	**22.21**

Note*. For the instances with $|T| = 2$, inequalities (3.16) were also added.

the 414 instances based on gulpinar2004extracting's graphs described in Table 20. Table 24 shows the results for these instances separated in two groups, according to the number of vertices: medium-size instances (between 141 and 339 vertices) and large-size instances (between 456 and 631 vertices). As we have 16 and 7 instances in the former and the latter group, there are 288 and 126 medium-size and large-size instances, respectively. We present the average computational time (Time) in seconds and the average number of generated nodes (Nodes) to solve the instances within the same category (A, B and C) and requesting equal number of teams ($|T| \in \{2,3,4\}$), i.e. 32 instances (medium-size) or 14 instances (large-size). Besides, we calculate the average percentage integrality gap of the linear relaxation for these instances. We compare the performance of formulations ILP -(3.11) and ILP-ALL (where all sets of inequalities (3.13), (3.15), (3.16) and (3.17) are included). Best results are shown in bold. See Table 29 in Appendix B for detailed results.

When using the valid inequalities, we can note a significant reduction in the average resolution time for all test sets. It also leads to a decease in the number of generated nodes. We can also observe that the spent time and the number of nodes usually increase with the number of vertices, the number of requested teams and the percentage of multi-skilled individuals. Regarding the gaps, we can observe that they are already small for ILP-(3.11), and formulation ILP-ALL can still reduce them in almost all groups. Similarly to the synthetic instances, the time spent to solve the relaxations are insignificant (below 1 second for all instances) and are not presented in table. See Table 30 in Appendix B for more details.

3.8 Conclusion

In this chapter, we have introduced a new variant of the TFP problem denominated Competitive Teams Formation Problem (CTFP). Using the theory of social balance, we represent

Table 24 – Average computational time and number of nodes to solve instances based on GÜLPINAR *et al.*'s graphs for formulations ILP -(3.11) with and without valid inequalities

Cat.	|T|	Medium-size instances (141-339 vertices)						Large-size instances (456-631 vertices)					
		ILP-(3.11)			ILP-ALL			ILP-(3.11)			ILP-ALL		
		Time	Nodes	Gap (%)	Time	Nodes	Gap (%)	Time	Nodes	Gap (%)	Time	Nodes	Gap (%)
A	2	4.04	1070.97	0.02	**2.98**	**1050.97**	**0.01**	48.88	4042.86	0.06	**45.64**	**2839.50**	**0.02**
B		4.17	1164.28	0.02	**3.10**	**904.78**	**0.01**	159.15	9704.00	0.06	**101.13**	**6103.14**	**0.02**
C		4.23	902.44	0.02	**3.36**	**800.00**	**0.01**	658.57	6287.36	0.07	**118.07**	**5202.50**	**0.02**
A	3	3.88	655.75	0.03	**2.85**	**571.97**	**0.01**	54.03	2972.71	0.06	**36.24**	**1752.50**	**0.02**
B		4.28	809.16	0.03	**3.67**	**786.16**	**0.01**	137.65	7963.29	0.06	**111.46**	**7583.64**	**0.02**
C		5.96	1515.47	0.03	**4.47**	**818.13**	**0.01**	661.16	27153.00	0.06	**403.04**	**16572.64**	**0.02**
A	4	34.70	8031.66	0.02	**26.54**	**7462.38**	**0.02**	81.07	4344.57	0.05	**51.69**	**3378.79**	**0.02**
B		63.11	17049.13	0.02	**50.16**	**14161.06**	**0.02**	83.23	5053.57	0.05	**52.01**	**4006.86**	**0.02**
C		61.33	16393.44	0.02	**45.93**	**10179.53**	**0.02**	299.60	20303.57	0.05	**138.98**	**11710.21**	**0.02**

Note*. For the instances with $|T| = 2$, inequalities (3.16) were also added.

the social network that connects the involved individuals as a signed graph and consider both intra-team and inter-teams communication costs by asking to have only positive relationships between individuals of a same team and only negative relationships between individuals of different teams. To solve it, we propose an ILP formulation and study structural properties on balance applied to signed graphs for generation of valid inequalities. We also show that it is possible to characterize the presented valid inequalities via a signed (smaller) multigraph, generated from edge-contracting operations. Since each group of valid inequalities presented is potentially exponential in size so that it is impracticable to include all of them in the ILP model, we present some polynomial-time algorithms to generate a subset of each family of valid inequalities.

The performance of the ILP formulation was compared with and without the use of inequalities presented. The comparative analysis took into account a set of synthetic instances of size and characteristics compatible with real applications as well a set instance used by Gülpinar *et al.* (2004). It has shown that the ILP with inclusion of inequalities consistently outperforms — in terms of effectiveness and efficiency — the ILP model without inclusion of inequalities. This indicates that the derived valid inequalities are very useful to strengthen the formulation and to significantly reduce its resolution time.

4 FINAL REMARKS AND FUTURE WORK

In this thesis, we study two generalizations for the Team Formation Problem. The first problem, called Multiple Team Formation Problem (MTFP), defined by Gutiérrez *et al.* (2016), allows distinct demands of workers per ability as well as multiple work teams and fractions of dedication time per team for each individual. Additionally, we consider a variant of the problem where individuals may have multiple skills, as in the extension of the TFP proposed by Campêlo *et al.* (2020). We model the MTFP as an integer linear program and present valid inequalities that are derived via reformulation-linearization techniques - RLT (SHERALI; ADAMS, 1990). Computational experiments demonstrate that the ILP formulation strengthened with valid inequalities consistently outperforms the quadratic model by Gutiérrez *et al.* (2016). The tests were carried out on a large variety of MTFP instances. We use extra variables to straightforwardly adapt our MTFP model to deal with the case where individuals may have multiple skills. Besides, we project out some of the variables to get a second ILP model with fewer variables but an exponential number of constraints. We show how to separate these constraints in polynomial time. The correctness of this model and the separation procedure is demonstrated via max-flow/min-cut arguments. We also present valid inequalities for both formulations, again using RLT techniques. We apply the separation routine to solve the second program with the branch-and-cut method. We computationally compare the two formulations and show their potential to solve instances generated from the original MTFP instances.

We could think of other generalizations for the MTFP. An interesting one would be the addition of the temporal dimension, as commented in Section 2.11. It could be related to the individuals or the projects. This can model other real scenarios at the expense of making the problem more complex. Since in MTFP an individual can partition his/her working time and, therefore, participate in more than one project/team, the possibility of having different human resources along the time horizon together with schedules for the projects can greatly increase the number of feasible solutions, especially in cases of large instances. For these cases, exact approaches could be possibly prohibitive so that the development of efficient heuristics and metaheuristics would be an alternative.

Variants of the MFTP that consider different objective functions or changes in the desired characteristics of the teams are also worth to mention as possible extensions of our work. For instance, one could adopt a multicriteria approach by taking both the communication costs and the project skill requirements as objectives. Other metrics for the communication

costs could be considered. Still, one could desire teams with a certain degree of heterogeneity among their members (in order to avoid "bubles" and favor innovation and originality) or a certain equilibrium among the efficiency of the formed teams (instead of focusing on the global efficiency). Another possibility would be the inclusion of disjunctive or dependency constraints (VIANA *et al.*, 2021), for instance to force some individuals of a subgroup to participate on some team or to avoid them to be simultaneously chosen for the teams.

The second problem studied in this thesis is a new variant of the TFP to be called Competitive Teams Formation Problem (CTFP). It simultaneously considers skill requirements and structural balance constraints. For its resolution, we propose an ILP formulation together with valid inequalities which are derived from structural properties on balance applied to signed graphs. We also define edge-contraction operations on the input signed graph so as to characterize the proposed valid inequalities. This favors the description and generation of the inequalities. Since each group of valid inequalities presented is potentially exponential in size and the separation problems are usually NP-hard, we propose a set of polynomial-time algorithms to generate the inequalities used to strengthen the model. We carry out some computational experiments to evaluate the performance of the ILP formulation and the inequalities. The benchmark instances comprise a set of synthetic instances of size and characteristics compatible with real applications as well a set instance adapted from Gülpinar *et al.* (2004). The computational results indicate that the derived valid inequalities are very useful to strengthen the formulation and to significantly reduce its resolution time.

As a future work, we would suggest the design of efficient separation heuristics, which would allow the use of the valid inequalities as cuts. A deeper study of the polytope associated with the CTFP would be another proposal for the continuation of this work. Results on these two points could be used to improve the efficiency of the branch-and-cut.

Simililarly to the MTFP, we could also think of several possible variants of the CTFP. For instance, the time dimension could be introduced as well. Particularly, one could consider other communication cost metrics in signed graphs, such as those presented by Kouvatis *et al.* (2020), where negative edge parity in the shortest paths between a pair of individuals is used to define their compatibility. As the authors provide only heuristic algorithms for the generation of the teams, an interesting track of research would be the generation of ILP models using such metrics. It is worth noting the valid inequalities presented here would also be valid for these new models.

BIBLIOGRAPHY

ADAMS, W. P.; SHERALI, H. D. A tight linearization and an algorithm for zero-one quadratic programming problems. **Management Science**, v. 32, n. 10, p. 1274–1290, 1986.

ANAGNOSTOPOULOS, A.; BECCHETTI, L.; CASTILLO, C.; GIONIS, A.; LEONARDI, S. Online team formation in social networks. In: **Proceedings of the 21st International Conference on World Wide Web**. [*S. l.*: *s. n.*], 2012. p. 839–848.

ANDREWS, M. C.; KACMAR, K. M.; BLAKELY, G. L.; BUCKLEW, N. S. Group cohesion as an enhancement to the justice—affective commitment relationship. **Group & Organization Management**, Sage Publications Sage CA: Los Angeles, CA, v. 33, n. 6, p. 736–755, 2008.

ARINGHIERI, R. Composing medical crews with equity and efficiency. **Central European Journal of Operations Research**, Elsevier, v. 17, n. 3, p. 343–357, 2009.

ARINIK, N.; FIGUEIREDO, R.; LABATUT, V. Signed graph analysis for the interpretation of voting behavior. **International Workshop on Recommender Systems and Social Network Analysis**, 2017.

ASKIN, R. G.; HUANG, Y. Forming effective worker teams for cellular manufacturing. **International Journal of Production Research**, Taylor & Francis, v. 39, n. 11, p. 2431–2451, 2001.

AWAL, G. K.; BHARADWAJ, K. K. Team formation in social networks based on collective intelligence – an evolutionary approach. **Applied Intelligence**, Springer, v. 41, n. 2, p. 627–648, 2014.

BAGHEL, V. S.; BHAVANI, S. D. Multiple team formation using an evolutionary approach. In: IEEE. **2018 Eleventh International Conference on Contemporary Computing (IC3)**. [*S. l.*], 2018. p. 1–6.

BALLESTEROS-PÉREZ, P.; GONZÁLEZ-CRUZ, M. C.; FERNÁNDEZ-DIEGO, M. Human resource allocation management in multiple projects using sociometric techniques. **International Journal of Project Management**, Elsevier, v. 30, n. 8, p. 901–913, 2012.

BANSAL, N.; BLUM, A.; CHAWLA, S. Correlation clustering. **Machine learning**, Springer, v. 56, n. 1-3, p. 89–113, 2004.

BARAHONA, F.; MAHJOUB, A. R. Facets of the balanced (acyclic) induced subgraph polytope. **Mathematical Programming**, v. 45, n. 1, p. 21–33, 1989.

BARRICK, M. R.; STEWART, G. L.; NEUBERT, M. J.; MOUNT, M. K. Relating member ability and personality to work-team processes and team effectiveness. **Journal of Applied Psychology**, American Psychological Association, v. 83, n. 3, p. 377–391, 1998.

BARTHOLDI, J. J. A good submatrix is hard to find. **Operations Research Letters**, v. 1, n. 5, p. 190–193, 1982.

BRUSCO, M.; STEINLEY, D. K-balance partitioning: An exact method with applications to generalized structural balance and other psychological contexts. **Psychological Methods**, v. 15, n. 2, p. 145–157, 2010.

BRUSCO, M.; STEINLEY, D. Exact and approximate algorithms for variable selection in linear discriminant analysis. **Computational Statistics & Data Analysis**, Elsevier, v. 55, n. 1, p. 123–131, 2011.

CAMPÊLO, M.; FIGUEIREDO, T. Integer programming approaches to the multiple team formation problem. **Computers & Operations Research**, 2021. To appear.

CAMPÊLO, M.; FIGUEIREDO, T.; SILVA, A. The sociotechnical teams formation problem: a mathematical optimization approach. **Annals of Operations Research**, v. 286, p. 201–216, 2020.

CAMPION, M. A.; MEDSKER, G. J.; HIGGS, A. C. Relations between work group characteristics and effectiveness: Implications for designing effective work groups. **Personnel Psychology**, Wiley Online Library, v. 46, n. 4, p. 823–847, 1993.

CARTWRIGHT, D.; HARARY, F. Structural balance: a generalization of heider's theory. **Psychological Review**, v. 63, n. 5, p. 277–293, 1956.

CHEN, R.-C.; LI, J.-Y.; MA, N.-J.; CHANG, Y.-T. Application of sociometry and genetic algorithm to selection of class officers. **Information – An International Interdisciplinary Journal**, International Information Institute, v. 16, n. 2, p. 1233–1241, 2013.

CHEN, S.-J.; LIN, L. Modeling team member characteristics for the formation of a multifunctional team in concurrent engineering. **IEEE Transactions on Engineering Management**, IEEE, v. 51, n. 2, p. 111–124, 2004.

CHHABRA, M.; DAS, S.; SZYMANSKI, B. Team formation in social networks. In: GELENBE, E.; LENT, R. (ed.). **Computer and Information Sciences III**. [*S. l.*]: Springer, 2013. p. 291–299.

COHEN, S. G.; BAILEY, D. E. What makes teams work: Group effectiveness research from the shop floor to the executive suite. **Journal of Management**, Sage Publications Sage CA: Thousand Oaks, CA, v. 23, n. 3, p. 239–290, 1997.

COLARELLI, S. M.; BOOS, A. L. Sociometric and ability-based assignment to work groups: Some implications for personnel selection. **Journal of Organizational Behavior**, Wiley Online Library, v. 13, n. 2, p. 187–196, 1992.

DAVIS, J. A. Structural balance, mechanical solidarity, and interpersonal relations. **American Journal of Sociology**, University of Chicago Press, v. 68, n. 4, p. 444–462, 1963.

DAVIS, J. A. Clustering and structural balance in graphs. **Human Relations**, Sage Publications Sage CA: Thousand Oaks, CA, v. 20, n. 2, p. 181–187, 1967.

DEMAINE, E. D.; EMANUEL, D.; FIAT, A.; IMMORLICA, N. Correlation clustering in general weighted graphs. **Theoretical Computer Science**, v. 361, n. 2, p. 172–187, 2006.

DOREIAN, P.; MRVAR, A. A partitioning approach to structural balance. **Social Networks**, Elsevier, v. 18, n. 2, p. 149–168, 1996.

DOREIAN, P.; MRVAR, A. Partitioning signed social networks. **Social Networks**, Elsevier, v. 31, n. 1, p. 1–11, 2009.

EASLEY, D.; KLEINBERG, J. **Networks, Crowds, and Markets: Reasoning About a Highly Connected World**. [*S. l.*]: Cambridge University Press, 2010.

ESQUIVEL, M. A.; KLEINER, B. H. The importance of conflict in work team effectiveness. **Team Performance Management: An International Journal**, MCB UP Ltd, v. 2, p. 42–48, 1996.

FACCHETTI, G.; IACONO, G.; ALTAFINI, C. Computing global structural balance in large-scale signed social networks. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 108, n. 52, p. 20953–20958, 2011.

FALKENAUER, E. **Genetic algorithms and grouping problems**. [*S. l.*]: John Wiley & Sons, Inc., 1998.

Farhadi, F.; Sorkhi, M.; Hashemi, S.; Hamzeh, A. An effective expert team formation in social networks based on skill grading. In: **2011 IEEE 11th International Conference on Data Mining Workshops**. [*S. l.*: *s. n.*], 2011. p. 366–372.

FERSHTMAN, M. Cohesive group detection in a social network by the segregation matrix index. **Social Networks**, Elsevier, v. 19, n. 3, p. 193–207, 1997.

FIGUEIREDO, R.; FROTA, Y. **An improved Branch-and-cut code for the maximum balanced subgraph of a signed graph**. [*S. l.*], 2013.

FIGUEIREDO, R.; FROTA, Y. The maximum balanced subgraph of a signed graph: Applications and solution approaches. **European Journal of Operational Research**, Elsevier, v. 236, n. 2, p. 473–487, 2014.

FIGUEIREDO, R.; FROTA, Y.; LABBÉ, M. A branch-and-cut algorithm for the maximum k-balanced subgraph of a signed graph. **Discrete Applied Mathematics**, Elsevier, v. 261, p. 164–185, 2019.

FIGUEIREDO, R.; MOURA, G. Mixed integer programming formulations for clustering problems related to structural balance. **Social Networks**, v. 35, n. 4, p. 639 – 651, 2013.

FIGUEIREDO, R. M. V.; LABBÉ, M.; SOUZA, C. C. d. An exact approach to the problem of extracting an embedded network matrix. **Computers & Operations Research**, Elsevier, v. 38, n. 11, p. 1483–1492, 2011.

FITSILIS, P.; GEROGIANNIS, V.; ANTHOPOULOS, L. Software project team selection based on enterprise social networks. In: GEN, M.; KIM, K.; HUANG, X.; HIROSHI, Y. (ed.). **Industrial Engineering, Management Science and Applications 2015**. [*S. l.*]: Springer, 2015. p. 375–384.

FITZPATRICK, E.; ASKIN, R.; GOLDBERG, J. Using student conative behaviors and technical skills to form effective project teams. In: IEEE. **31st Annual Frontiers in Education Conference. Impact on Engineering and Science Education. Conference Proceedings**. [*S. l.*], 2001. v. 3, p. S2G–8.

FITZPATRICK, E. L.; ASKIN, R. G. Forming effective worker teams with multi-functional skill requirements. **Computers & Industrial Engineering**, Elsevier, v. 48, n. 3, p. 593–608, 2005.

FORD, L. R.; FULKERSON, D. R. Maximal flow through a network. **Canadian Journal of Mathematics**, Cambridge University Press, v. 8, p. 399–404, 1956.

FORSYTH, E.; KATZ, L. A matrix approach to the analysis of sociometric data: preliminary report. **Sociometry**, JSTOR, v. 9, n. 4, p. 340–347, 1946.

FRANZ, C. R.; JIN, K. G. The structure of group conflict in a collaborative work group during information systems development. **Journal of Applied Communication Research**, v. 23, n. 2, p. 108–127, 1995.

GAREY, M.; JOHNSON, D.; STOCKMEYER, L. Some simplified np-complete graph problems. **Theoretical Computer Science**, v. 1, n. 3, p. 237–267, 1976.

GASTON, M.; SIMMONS, J.; DESJARDINS, M. Adapting network structure for efficient team formation. In: **Proceedings of the AAAI Fall Symposium on Artificial Multi-agent Learning**. [*S. l.*: *s. n.*], 2004. p. 1–8.

GLADSTEIN, D. L. Groups in context: A model of task group effectiveness. **Administrative Science Quarterly**, JSTOR, v. 29, n. 4, p. 499–517, 1984.

GOODMAN, P. S.; DEVADAS, R.; HUGHSON, T. Groups and productivity: Analyzing the effectiveness of self-managing teams. In: **Productivity in Organizations: New Perspectives From Industrial and Organizational Psychology**. [*S. l.*]: Jossey-Bass San Francisco, 1988. p. 295–327.

GÜLPINAR, N.; GUTIN, G.; MITRA, G.; ZVEROVITCH, A. Extracting pure network submatrices in linear programs using signed graphs. **Discrete Applied Mathematics**, Elsevier, v. 137, n. 3, p. 359–372, 2004.

GUNDLACH, M.; ZIVNUSKA, S.; STONER, J. Understanding the relationship between individualism–collectivism and team performance through an integration of social identity theory and the social relations model. **Human Relations**, Sage Publications London, Thousand Oaks CA, New Delhi, v. 59, n. 12, p. 1603–1632, 2006.

GUTIÉRREZ, J. H.; ASTUDILLO, C. A.; BALLESTEROS-PÉREZ, P.; MORA-MELIÀ, D.; CANDIA-VÉJAR, A. The multiple team formation problem using sociometry. **Computers & Operations Research**, Elsevier, v. 75, p. 150–162, 2016.

GUZZO, R. A.; SHEA, G. P. Group performance and intergroup relations in organizations. In: **Handbook of Industrial and Organizational Psychology**. [*S. l.*]: Consulting Psychologists Press, 1992. p. 269–313.

HARARY, F. On the notion of balance of a signed graph. **The Michigan Mathematical Journal**, The University of Michigan, v. 2, n. 2, p. 143–146, 1953.

HARARY, F. On the measurement of structural balance. **Behavioral Science**, v. 4, n. 4, p. 316–323, 1959.

HARARY, F.; KABELL, J. A. A simple algorithm to detect balance in signed graphs. **Mathematical Social Sciences**, v. 1, n. 1, p. 131 – 136, 1980.

HARARY, F.; NORMAN, R.; CARTWRIGHT, D. **Structural Models: An Introduction to the Theory of Directed Graphs**. [*S. l.*]: Wiley, 1965.

HARDING, J. Review of sociometry, experimental method and the science of society: An approach to a new political orientation. **Psychological Bulletin**, American Psychological Association, v. 49, n. 3, p. 283–284, 1952.

HEIDER, F. Attitudes and cognitive organization. **The Journal of psychology**, Taylor & Francis, v. 21, n. 1, p. 107–112, 1946.

HLAOITTINUN, O.; BONJOUR, E.; DULMET, M. A team building approach for competency development. In: IEEE. **2007 IEEE International Conference on Industrial Engineering and Engineering Management**. [*S. l.*], 2007. p. 1004–1008.

INOHARA, T. Clusterability of groups and information exchange in group decision making with approval voting system. **Applied Mathematics and Computation**, Elsevier, v. 136, n. 1, p. 1–15, 2003.

KARGAR, M.; AN, A. Discovering top-k teams of experts with/without a leader in social networks. In: **Proceedings of the 20th ACM international conference on Information and knowledge management**. [*S. l.*: *s. n.*], 2011. p. 985–994.

KARGAR, M.; AN, A.; ZIHAYAT, M. Efficient bi-objective team formation in social networks. In: SPRINGER. **Joint European Conference on Machine Learning and Knowledge Discovery in Databases**. [*S. l.*], 2012. p. 483–498.

KOLBE, K. **Pure instinct**. New York: Random House, 1993.

KOUVATIS, I.; SEMERTZIDIS, K.; ZERVA, M.; PITOURA, E.; TSAPARAS, P. Forming compatible teams in signed networks. In: **Proceedings of the 23rd International Conference on Extending Database Technology (EDBT)**. [*S. l.*]: OpenProceedings.org, 2020. p. 363–366.

KUIPERS, B. S.; HIGGS, M. J.; TOLKACHEVA, N. V.; WITTE, M. C. de. The influence of Myers-Briggs type indicator profiles on team development processes: An empirical study in the manufacturing industry. **Small Group Research**, Sage Publications Sage CA: Los Angeles, CA, v. 40, n. 4, p. 436–464, 2009.

LAPPAS, T.; LIU, K.; TERZI, E. Finding a team of experts in social networks. In: ACM. **Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**. [*S. l.*], 2009. p. 467–476.

LESKOVEC, J.; HUTTENLOCHER, D.; KLEINBERG, J. Signed networks in social media. In: **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. [*S. l.*: *s. n.*], 2010. p. 1361–1370.

LEVORATO, M.; FIGUEIREDO, R.; FROTA, Y.; DRUMMOND, L. Evaluating balancing on social networks through the efficient solution of correlation clustering problems. **EURO Journal on Computational Optimization**, Springer, v. 5, n. 4, p. 467–498, 2017.

LEVORATO, M.; FROTA, Y. Brazilian congress structural balance analysis. **Journal of Interdisciplinary Methodologies and Issues in Science**, v. 2, p. 1–27, 2017.

LI, C.-T.; SHAN, M.-K. Team formation for generalized tasks in expertise social networks. In: IEEE. **2010 IEEE Second International Conference on Social Computing**. [*S. l.*], 2010. p. 9–16.

LI, C.-T.; SHAN, M.-K.; LIN, S.-D. On team formation with expertise query in collaborative social networks. **Knowledge and Information Systems**, Springer, v. 42, n. 2, p. 441–463, 2015.

LI, Y.; CHEN, W.; WANG, Y.; ZHANG, Z.-L. Voter model on signed social networks. **Internet Mathematics**, Taylor & Francis, v. 11, n. 2, p. 93–133, 2015.

LIND, M. R. The effectiveness of electronic work groups for student cases. **Proceedings of the First Americas Conference on Information Systems**, v. 155, p. 1–7, 1995.

LITTLE, B. L.; MADIGAN, R. M. The relationship between collective efficacy and performance in manufacturing work teams. **Small Group Research**, SAGE PUBLICATIONS, INC. 2455 Teller Road, Thousand Oaks, CA 91320, v. 28, n. 4, p. 517–534, 1997.

LODAHL, T. M.; PORTER, L. W. Psychometric score patterns, social characteristics, and productivity of small industrial work groups. **Journal of Applied Psychology**, American Psychological Association, v. 45, n. 2, p. 73–79, 1961.

MAJUMDER, A.; DATTA, S.; NAIDU, K. Capacitated team formation problem on social networks. In: **Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining**. [*S. l.*: *s. n.*], 2012. p. 1005–1013.

MORENO, J. L. Foundations of sociometry: An introduction. **Sociometry**, JSTOR, v. 4, n. 1, p. 15–35, 1941.

MYERS, I. B. **The Myers-Briggs type indicator manual**. Princeton, NJ: Educational Testing Service, 1962.

MYERS, I. B. **Introduction to type: A description of the theory and applications of the Myers-Briggs Type Indicator**. [*S. l.*]: Vision Australia Student Support, 1997.

NEMHAUSER, G. L.; SIGISMONDI, G. A strong cutting plane/branch-and-bound algorithm for node packing. **Journal of the Operational Research Society**, Taylor & Francis, v. 43, n. 5, p. 443–457, 1992.

NEUMAN, G. A.; WAGNER, S. H.; CHRISTIANSEN, N. D. The relationship between work-team personality composition and the job performance of teams. **Group & Organization Management**, Sage Publications Sage CA: Thousand Oaks, CA, v. 24, n. 1, p. 28–45, 1999.

POLJAK, S.; TURZíK, D. On a facet of the balanced subgraph polytope (english). **Časopis pro Pěstování Matematiky**, v. 112, n. 4, p. 373–380, 1987.

RANGAPURAM, S. S.; BÜHLER, T.; HEIN, M. Towards realistic team formation in social networks based on densest subgraphs. In: **Proceedings of the 22nd international conference on World Wide Web**. [*S. l.*: *s. n.*], 2013. p. 1077–1088.

ROISTACHER, R. C. A review of mathematical methods in sociometry. **Sociological Methods & Research**, Elsevier, v. 3, n. 2, p. 123–171, 1974.

SHERALI, H.; ADAMS, W. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. **SIAM Journal on Discrete Mathematics**, v. 3, n. 3, p. 411–430, 1990.

SHI, G.; PROUTIERE, A.; JOHANSSON, M.; BARAS, J. S.; JOHANSSON, K. H. The evolution of beliefs over signed social networks. **Operations Research**, INFORMS, v. 64, n. 3, p. 585–604, 2016.

SILVA, I. E. da; KROHLING, R. A. A fuzzy sociometric approach to human resource allocation. In: IEEE. **2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)**. [*S. l.*], 2018. p. 1–8.

STOCKMEYER, L. Planar 3-colorability is polynomial complete. **SIGACT News**, Association for Computing Machinery, v. 5, n. 3, p. 19–25, 1973.

TANG, J.; CHANG, Y.; AGGARWAL, C.; LIU, H. A survey of signed network mining in social media. **ACM Computing Surveys (CSUR)**, ACM New York, NY, USA, v. 49, n. 3, p. 1–37, 2016.

TSENG, T.-L. B.; HUANG, C.-C.; CHU, H.-W.; GUNG, R. R. Novel approach to multi-functional project team formation. **International Journal of Project Management**, Elsevier, v. 22, n. 2, p. 147–159, 2004.

VIANA, L.; CAMPÊLO, M.; SAU, I.; SILVA, A. A unifying model for locally constrained spanning tree problems. **Journal of Combinatorial Optimization**, Springer, p. 1–26, 2021.

WASSERMAN, S.; FAUST, K. **Social network analysis: Methods and applications**. Princeton, NJ: Cambridge University Press, 1994.

WEST, D. B. **Introduction to Graph Theory**. 2nd. ed. [*S. l.*]: Pearson, 2000.

WHEELAN, S. A. Group size, group development, and group productivity. **Small Group Research**, Sage Publications Sage CA: Los Angeles, CA, v. 40, n. 2, p. 247–262, 2009.

WI, H.; OH, S.; MUN, J.; JUNG, M. A team formation model based on knowledge and collaboration. **Expert Systems with Applications**, Elsevier, v. 36, n. 5, p. 9121–9134, 2009.

YANG, B.; CHEUNG, W.; LIU, J. Community mining from signed social networks. **IEEE Transactions on Knowledge and Data Engineering**, IEEE, v. 19, n. 10, p. 1333–1348, 2007.

APPENDIX A – DETAILED COMPUTATIONAL TIMES - CTFP

Table 25 – Average computational times for Class 1, Class 4 and Class 7 instances with the quadratic model, the linear model and the linear model with inclusion of the all possible combinations of valid inequalities.

Model	Class 1			Class 4			Class 7		
	I	II	III	I	II	III	I	II	III
Quad. Integer	0.177	0.297	0.549	24.359	78.058	564.577	470.557	591.871	1333.226
Quad. Binary	0.169	0.355	0.559	11.614	34.155	22.324	377.069	439.082	1083.731
ILP	0.348	0.407	0.376	21.039	35.034	14.414	569.736	260.326	202.214
37	0.465	0.445	0.446	25.736	27.904	12.965	65.959	44.675	36.534
40	0.645	0.640	0.571	21.132	25.655	17.343	111.838	61.993	59.369
42	0.481	0.467	0.462	20.162	16.756	7.679	238.279	98.061	124.271
43	0.197	0.155	0.154	3.531	3.861	2.976	27.064	24.676	56.789
45	0.487	0.475	0.485	23.602	16.875	13.032	84.441	48.887	39.283
49	0.432	0.403	0.413	26.226	23.373	19.295	83.111	46.343	46.803
50	0.470	0.751	0.529	23.577	21.756	14.915	102.345	38.741	33.587
37-40	0.701	0.653	0.687	25.057	24.366	17.224	142.649	86.176	101.655
37-42	0.528	0.509	0.588	25.707	30.049	18.126	129.466	73.356	55.479
37-43	0.128	0.205	0.155	4.004	4.841	3.891	8.704	6.234	6.543
37-45	0.427	0.462	0.423	39.856	40.720	18.976	72.539	49.820	42.024
37-49	0.395	0.433	0.405	46.884	29.343	19.262	152.419	60.856	54.126
37-50	0.457	0.502	0.505	34.339	38.828	22.399	107.232	58.535	58.549
37-40-42	0.579	0.649	0.485	20.494	30.665	24.838	161.624	99.161	80.546
37-40-43	0.142	0.127	0.116	3.513	3.589	3.316	5.201	4.320	3.422
37-40-45	0.483	0.485	0.453	26.985	24.558	28.155	117.386	66.591	36.820
37-40-49	0.703	0.639	0.600	21.535	25.143	21.089	149.019	85.089	60.629
37-40-50	0.679	0.626	0.589	25.332	26.656	15.622	118.260	94.110	76.409
37-42-43	0.135	0.194	0.174	2.088	2.149	1.221	4.348	3.884	2.556
37-42-45	0.442	0.476	0.427	22.086	37.553	17.058	122.071	79.605	55.583
37-42-49	0.573	0.620	0.532	20.797	19.976	12.899	201.954	118.979	97.371
37-42-50	0.647	0.611	0.661	24.555	21.917	22.692	137.393	85.430	53.011
37-43-45	**0.171**	**0.151**	**0.143**	**0.634**	**0.926**	**0.812**	**0.858**	**1.253**	**0.635**
37-43-49	0.125	0.122	0.117	5.199	5.145	3.667	7.291	6.782	6.501
37-43-50	0.115	0.123	0.119	6.454	7.618	3.862	7.477	8.028	5.903
37-45-49	0.502	0.524	0.519	26.821	32.376	29.127	99.283	66.911	50.804
37-45-50	0.576	0.578	0.593	29.282	24.643	22.035	57.706	42.883	27.278

37-49-50	0.436	0.493	0.480	36.228	33.831	23.446	174.060	126.665	59.716
37-40-42-43	0.189	0.165	0.163	1.904	2.183	1.474	3.370	2.942	1.754
37-40-42-45	0.607	0.537	0.572	23.602	29.416	27.154	154.381	83.023	71.430
37-40-42-49	0.719	0.863	0.626	18.188	18.099	17.102	207.493	105.435	115.565
37-40-42-50	0.683	0.735	0.670	26.844	35.893	13.195	218.685	90.397	91.947
37-40-43-45	**0.171**	**0.153**	**0.136**	**0.667**	**0.866**	**0.669**	**0.806**	**0.767**	**0.595**
37-40-43-49	0.171	0.149	0.165	3.616	4.461	6.417	8.586	6.679	5.463
37-40-43-50	0.112	0.116	0.126	5.027	4.209	3.729	5.433	5.053	3.670
37-40-45-49	1.055	0.972	0.883	20.790	29.643	18.128	194.744	104.931	59.816
37-40-45-50	0.983	0.828	0.811	32.936	26.827	19.759	114.138	97.741	60.540
37-40-49-50	0.787	0.954	0.783	30.398	27.962	19.721	237.438	101.577	94.773
37-40-42-43-45	**0.183**	**0.204**	**0.196**	**0.667**	**0.926**	**0.739**	**1.206**	**1.382**	**0.977**
37-40-42-43-49	0.194	0.208	0.192	2.654	2.233	2.860	3.673	3.993	2.461
37-40-42-43-50	0.193	0.179	0.167	2.126	1.982	2.503	2.807	3.200	1.779
37-40-42-45-49	1.080	0.838	0.730	30.523	23.125	12.914	223.866	128.326	84.760
37-40-42-45-50	0.968	0.812	1.017	22.730	49.014	15.938	132.508	92.687	91.463
37-40-42-49-50	0.867	0.979	0.727	19.034	25.237	18.577	225.287	141.647	138.198
37-40-43-45-49	**0.199**	**0.170**	**0.162**	**0.927**	**1.325**	**0.700**	**1.214**	**1.137**	**0.806**
37-40-43-45-50	**0.123**	**0.130**	**0.118**	**0.664**	**1.192**	**0.618**	**0.947**	**0.813**	**0.643**
37-40-43-49-50	0.123	0.123	0.139	4.125	3.925	3.473	9.282	8.942	7.420
37-40-45-49-50	0.695	0.638	0.675	32.868	54.711	16.331	152.488	92.376	50.750
37-42-43-45	0.187	0.209	0.176	3.378	3.372	2.392	4.767	4.683	2.770
37-42-43-49	0.143	0.149	0.149	2.508	2.396	2.382	3.098	3.292	2.112
37-42-43-50	0.163	0.134	0.115	2.893	3.066	2.913	3.792	3.270	2.109
37-42-45-49	0.523	0.508	0.442	31.221	35.608	14.293	229.375	118.975	137.096
37-42-45-50	0.577	0.554	0.484	24.945	26.802	10.073	121.677	59.000	48.746
37-42-49-50	0.487	0.554	0.470	22.631	21.202	11.983	367.680	167.411	121.204
37-43-45-49	**0.116**	**0.166**	**0.137**	**0.554**	**0.938**	**0.591**	**0.993**	**1.857**	**0.670**
37-43-45-50	**0.135**	**0.143**	**0.112**	**0.648**	**0.772**	**0.870**	**0.729**	**1.448**	**0.679**
37-43-49-50	0.181	0.181	0.130	4.653	4.485	9.297	8.610	9.100	7.592
37-45-49-50	0.538	0.562	0.602	34.946	34.710	21.032	96.553	72.064	49.632
37-40-42-43-45-49	**0.275**	**0.183**	**0.230**	**0.618**	**0.811**	**0.738**	**1.231**	**1.027**	**0.827**
37-40-42-43-45-50	**0.119**	**0.156**	**0.121**	**0.612**	**1.197**	**0.782**	**0.818**	**0.989**	**0.677**
37-40-42-43-49-50	0.152	0.210	0.189	3.202	1.916	2.554	2.987	5.487	2.416
37-40-42-45-49-50	0.724	0.630	0.663	21.734	26.379	9.553	374.141	90.394	119.958
37-40-42-43-45-49-50	**0.279**	**0.268**	**0.208**	**0.553**	**0.861**	**0.667**	**1.277**	**1.216**	**0.954**
37-40-43-45-49-50	**0.201**	**0.147**	**0.139**	**0.695**	**1.212**	**0.538**	**1.166**	**0.956**	**0.652**

37-42-43-45-49	**0.214**	**0.176**	**0.147**	**0.619**	**0.749**	**0.698**	**1.382**	**1.114**	**0.717**
37-42-43-45-50	**0.103**	**0.139**	**0.123**	**0.530**	**0.841**	**0.575**	**0.822**	**0.843**	**0.517**
37-42-43-49-50	0.111	0.134	0.151	2.553	2.948	2.138	3.891	4.454	2.764
37-42-45-49-50	0.666	0.650	0.570	38.014	47.427	28.057	282.767	170.171	111.286
37-43-45-49-50	0.156	0.175	0.149	0.768	1.109	0.693	2.028	3.312	2.701
37-42-43-45-49-50	0.135	0.148	0.180	0.618	0.947	0.624	1.828	2.981	3.654
40-42	0.736	0.816	0.607	16.829	42.588	24.961	154.214	100.063	80.958
40-43	0.136	0.149	0.135	4.302	4.784	3.539	7.179	4.940	3.329
40-45	0.666	0.600	0.626	27.887	35.272	16.467	148.383	86.179	64.297
40-49	0.698	0.675	0.634	24.398	25.102	15.758	173.586	122.070	75.766
40-50	0.703	0.587	0.726	23.466	28.183	18.745	100.426	64.061	67.134
40-42-43	0.163	0.145	0.156	2.905	2.828	2.464	3.291	2.568	1.992
40-42-45	0.698	0.653	0.545	20.334	30.346	11.574	134.308	89.504	74.603
40-42-49	0.759	0.673	0.608	22.524	24.439	14.696	300.488	92.096	133.105
40-42-50	1.316	1.342	0.971	31.493	32.341	25.262	217.612	104.772	144.809
40-43-45	0.171	0.214	0.150	0.615	0.887	0.555	2.725	2.752	3.443
40-43-49	0.178	0.151	0.153	5.573	4.609	3.688	6.011	6.448	5.397
40-43-50	0.126	0.120	0.112	4.839	4.446	3.691	6.767	6.131	4.049
40-45-49	0.708	0.759	0.627	20.262	30.569	11.270	153.209	89.550	58.933
40-45-50	0.812	0.870	0.725	20.782	22.872	20.005	158.838	68.692	45.694
40-49-50	0.817	0.743	0.685	31.348	32.627	23.796	268.015	95.861	99.017
40-42-43-45	0.206	0.182	0.158	0.759	0.865	0.835	3.803	2.810	3.676
40-42-43-49	0.181	0.141	0.144	2.416	1.964	2.604	3.576	3.690	2.337
40-42-43-50	0.131	0.136	0.143	2.816	2.295	2.854	2.402	3.337	1.573
40-42-45-49	0.912	0.694	0.765	28.135	34.732	15.397	273.492	100.818	122.724
40-42-45-50	0.957	0.930	0.996	22.052	38.532	15.482	129.243	71.824	75.997
40-42-49-50	0.921	1.062	0.909	21.831	23.674	14.762	196.253	111.242	133.021
40-43-45-49	0.231	0.197	0.217	0.653	0.785	0.786	3.271	4.161	2.939
40-43-45-50	0.227	0.217	0.236	0.818	1.174	0.773	1.904	3.748	2.607
40-43-49-50	0.114	0.129	0.187	6.485	3.965	4.483	6.393	7.573	6.450
40-45-49-50	1.063	0.818	0.896	26.444	27.941	19.192	184.050	96.560	72.106
40-42-43-45-49	0.244	0.210	0.223	0.720	0.853	0.646	3.095	3.405	2.855
40-42-43-45-50	0.191	0.187	0.204	0.576	0.886	0.828	3.048	3.090	2.705
40-42-43-49-50	0.143	0.194	0.162	2.208	4.204	2.700	3.913	3.178	2.489
40-42-45-49-50	1.031	0.992	0.723	17.029	39.950	14.341	319.569	131.639	141.285
40-43-45-49-50	0.181	0.188	0.173	0.773	1.359	0.898	3.128	1.125	2.760
42-43	0.170	0.184	0.152	3.376	2.314	3.532	3.629	5.703	2.329

42-45	0.679	0.737	0.552	21.664	28.606	21.157	103.620	113.864	59.841
42-49	0.586	0.577	0.555	23.128	18.803	9.493	184.300	123.343	79.124
42-50	0.600	0.598	0.648	21.589	23.613	14.557	117.747	65.942	40.588
42-43-45	0.145	0.231	0.177	0.592	0.905	0.716	3.359	3.233	0.939
42-43-49	0.223	0.221	0.189	3.033	3.399	2.209	4.014	4.490	2.108
42-43-50	0.129	0.127	0.130	1.652	1.949	1.326	2.418	3.713	1.349
42-45-49	0.638	0.663	0.464	31.782	36.313	14.820	183.619	143.474	84.053
42-45-50	0.631	0.660	0.492	20.035	36.445	14.798	109.336	66.187	53.771
42-49-50	0.623	0.562	0.576	19.387	18.417	12.832	292.273	195.491	82.110
42-43-45-49	0.244	0.172	0.232	0.748	0.988	0.724	3.161	1.963	1.764
42-43-45-50	0.137	0.173	0.186	0.552	0.566	0.512	1.837	3.870	1.554
42-43-49-50	0.126	0.160	0.129	2.404	3.240	1.951	4.218	4.025	2.439
42-45-49-50	0.553	0.667	0.527	28.970	33.080	11.909	211.029	123.555	102.734
42-43-45-49-50	0.216	0.232	0.193	0.828	1.198	0.846	2.287	3.453	1.931
43-45	0.350	0.296	0.327	1.085	1.938	1.476	1.035	3.306	1.066
43-49	0.324	0.292	0.248	8.464	7.446	6.781	12.372	23.455	10.804
43-50	0.246	0.228	0.214	7.931	7.656	5.865	11.639	14.520	12.383
43-45-49	0.167	0.156	0.162	0.815	1.222	1.110	2.989	3.744	1.796
43-45-50	0.190	0.154	0.176	0.687	1.127	0.983	2.609	3.082	1.486
43-49-50	0.153	0.119	0.113	6.080	4.374	8.819	10.172	10.905	8.797
43-45-49-50	0.241	0.273	0.253	0.959	1.218	1.024	2.114	3.721	1.882
45-49	0.687	0.758	0.736	25.846	29.160	17.806	94.761	58.423	43.918
45-50	0.420	0.503	0.462	21.805	26.893	12.069	84.589	39.032	29.208
45-49-50	0.706	0.725	0.737	47.065	41.478	22.330	118.668	80.330	65.892
49-50	0.525	0.516	0.484	36.455	52.164	23.222	192.581	95.649	82.996
AGV (all comb.)	0.425	0.422	0.391	14.279	16.299	9.915	85.972	48.155	40.092
AGV (less than 3 sec.)	0.176	0.168	0.152	0.645	0.970	0.692	1.035	1.139	0.719

APPENDIX B – DETAILED COMPUTATIONAL TIMES - CTFP

Table 26 – Computational time and number of nodes used to solve instances generated with synthetic graphs for formulations ILP (3.5)-(3.10) and ILP-(3.11).

Instance	OTM	ILP (3.5)-(3.10)		ILP-(3.11)	
		Time	Nodes	Time	Nodes
|T| = 2					
S1A1II	27	1.057	2592	1.09	1718
S1A2II	25	0.859	1213	0.95	793
S2A1II	27	1.947	4586	1.36	2888
S2A2II	27	0.356	0	0.38	0
S3A1II	29	0.978	459	0.76	405
S3A2II	27	0.859	435	0.61	0
S4A1II	23	1.758	1655	1.24	1553
S4A2II	23	0.461	0	0.32	0
S5A1II	23	1.579	1780	1.41	1533
S5A2II	22	0.880	796	1.06	698
S6A1II	22	2.048	2618	1.37	1985
S6A2II	21	0.936	308	0.71	0
Avg.		**1.14**	**1370.17**	**0.94**	**964.42**
S1B1II	27	1.32	1852	2.36	4320
S1B2II	27	1.07	1096	0.97	1506
S2B1II	27	3.00	5457	4.39	5578
S2B2II	27	1.17	2245	1.83	3086
S3B1II	29	0.98	900	0.94	885
S3B2II	28	1.38	1097	0.82	1440
S4B1II	23	1.68	2596	1.63	3236
S4B2II	23	1.29	996	1.17	566
S5B1II	23	1.69	3679	3.24	5011
S5B2II	23	1.10	2195	1.44	1696
S6B1II	22	1.70	3646	2.22	4223
S6B2II	21	1.66	2135	1.49	2480
Avg.		**1.50**	**2324.50**	**1.88**	**2835.58**
S1C1II	26	2.934	4069	2.23	4380
S1C2II	26	1.564	2856	1.61	3836
S2C1II	26	4.518	12405	3.68	12563
S2C2II	26	2.430	4780	2.59	5038
S3C1II	29	0.981	2094	0.85	834
S3C2II	27	1.353	2695	0.86	2033
S4C1II	23	1.284	1892	1.92	3393
S4C2II	23	1.079	666	1.43	1571
S5C1II	23	2.850	5551	2.93	4085
S5C2II	23	1.182	1613	1.58	2133
S6C1II	22	3.224	4490	4.19	4513

S6C2II	22	2.209	2650	2.55	1998
Avg.		**2.13**	**3813.42**	**2.20**	**3864.75**
|T| = 3					
S1A1III	27	27.413	18031	31.95	37773
S1A2III	26	8.255	9442	3.65	3704
S2A1III	30	22.713	29017	13.75	10064
S2A2III	29	1.858	1124	1.81	1209
S3A1III	30	21.268	19544	16.16	12495
S3A2III	29	2.322	2152	2.72	2909
S4A1III	25	31.995	25280	31.43	21255
S4A2III	24	2.655	1556	2.39	1551
S5A1III	25	29.085	26450	30.37	21216
S5A2III	24	2.053	1471	2.90	1462
S6A1III	25	45.768	33099	30.70	23123
S6A2III	24	2.436	1560	15.67	7767
Avg.		**16.49**	**14060.50**	**15.29**	**12044**
S1B1III	28	24.20	18465	23.31	17570
S1B2III	28	4.07	2833	3.70	3384
S2B1III	30	19.82	26507	22.55	23058
S2B2III	30	13.31	12976	6.30	5626
S3B1III	30	23.92	28898	18.41	20252
S3B2III	30	9.33	8735	7.42	5504
S4B1III	25	38.63	46590	40.01	37022
S4B2III	23	16.75	15581	19.38	17252
S5B1III	26	21.46	12030	21.64	19304
S5B2III	25	10.04	8103	5.07	5741
S6B1III	26	28.94	20712	20.55	11859
S6B2III	25	6.84	5104	15.67	7767
Avg.		**18.11**	**17211.17**	**17.00**	**14528.25**
S1C1III	28	23.12	21700	20.98	15413
S1C2III	28	7.20	4073	5.26	3541
S2C1III	30	40.14	52415	24.37	23362
S2C2III	29	9.95	11906	11.52	13283
S3C1III	30	24.69	24078	15.07	12891
S3C2III	30	8.34	6438	9.71	7478
S4C1III	24	53.21	82776	59.91	78469
S4C2III	23	28.90	31197	28.07	30840
S5C1III	26	21.33	17806	27.28	20826
S5C2III	25	12.00	9120	11.77	7507
S6C1III	26	30.01	28096	20.40	14008
S6C2III	25	19.65	13754	15.24	14449
Avg.		**23.21**	**25279.92**	**20.80**	**20172.25**
|T| = 4					
S1A1IV	27	44.74	25260	55.36	27282

S1A2IV	25	22.35	13202	23.94	13374
S2A1IV	31	109.30	101141	31.24	14951
S2A2IV	30	37.58	27122	33.21	22439
S3A1IV	31	143.69	125475	43.15	20206
S3A2IV	31	24.97	18968	21.15	13441
S4A1IV	25	146.09	91603	106.82	53017
S4A2IV	24	33.85	16482	31.05	13561
S5A1IV	25	351.20	241053	273.38	162011
S5A2IV	24	37.38	17543	34.05	17144
S6A1IV	26	67.38	48773	43.01	17618
S6A2IV	28	4.41	1292	6.51	2189
Avg.		**85.24**	**60659.50**	**58.57**	**31436.08**
S1B1IV	28	145.66	106984	169.36	112006
S1B2IV	28	60.64	35636	48.33	27033
S2B1IV	31	227.77	167211	48.87	26559
S2B2IV	31	96.92	89371	84.89	48423
S3B1IV	31	96.92	117003	39.17	21495
S3B2IV	31	58.87	74012	57.42	31575
S4B1IV	25	475.82	430792	217.78	158263
S4B2IV	24	134.23	104742	90.13	66629
S5B1IV	26	164.76	123858	98.43	86175
S5B2IV	26	44.24	28267	48.44	30347
S6B1IV	28	128.56	110452	29.94	17129
S6B2IV	27	61.73	39190	62.13	32986
Avg.		**141.34**	**118959.83**	**82.91**	**54885**
S1C1IV	28	259.62	189090	186.73	174828
S1C2IV	28	67.53	48690	40.24	26267
S2C1IV	31	131.91	119925	44.86	25107
S2C2IV	31	61.79	41758	66.26	53341
S3C1IV	30	743.39	658157	110.13	112725
S3C2IV	31	78.04	63316	48.75	39039
S4C1IV	25	918.24	781789	387.16	425617
S4C2IV	25	164.61	126851	125.35	105828
S5C1IV	25	793.63	686496	448.52	380460
S5C2IV	25	120.43	104658	95.77	92562
S6C1IV	27	417.94	392753	55.69	44158
S6C2IV	27	88.83	64969	52.91	32257
Avg.		**320.50**	**273204.33**	**138.53**	**126015.75**

Table 27 – Computational time and number of nodes used to solve instances generated with synthetic graphs for formulation ILP-(3.11) without and with inclusion of all valid inequalities (ILP-ALL).

Instance	ILP-(3.11)		ILP-(3.13)		ILP-(3.15)*		ILP-(3.17)		ILP-ALL	
	Time	Nodes	Time	Nodes	Time	Nodes	Time	Nodes	Time	Nodes

|T| = 2										
S1A1II	1.09	1718	1.02	1724	0.73	1216	1.07	1715	0.951	1731
S1A2II	0.95	793	0.79	629	0.69	768	0.75	671	0.633	863
S2A1II	1.36	2888	1.24	2488	1.40	2895	1.45	2532	1.373	3540
S2A2II	0.38	0	0.37	0	0.44	0	0.29	0	0.463	0
S3A1II	0.76	405	0.75	404	0.85	403	0.48	451	0.706	386
S3A2II	0.61	0	0.54	0	0.56	0	0.67	0	0.499	456
S4A1II	1.24	1553	1.16	1553	1.47	1513	1.33	1405	1.297	1830
S4A2II	0.32	0	0.53	0	0.49	0	0.38	0	0.501	0
S5A1II	1.41	1533	1.46	1530	1.45	1528	1.34	1530	1.422	2538
S5A2II	1.06	698	0.62	548	0.71	640	0.53	638	0.622	534
S6A1II	1.37	1985	1.41	1933	1.30	1997	1.72	1921	1.283	1993
S6A2II	0.71	0	0.83	0	0.75	0	0.72	0	1.018	477
Avg.	**0.94**	**964.42**	**0.89**	**900.75**	**0.90**	**913.33**	**0.89**	**905.25**	**0.90**	**1195.67**
S1B1II	2.36	4320	1.89	3712	1.29	1734	1.39	2056	0.909	1811
S1B2II	0.97	1506	1.24	1004	0.73	607	0.82	875	1.071	974
S2B1II	4.39	5578	3.96	6591	1.92	5002	2.89	5950	1.726	5220
S2B2II	1.83	3086	1.34	2326	1.79	2775	1.49	2234	1.481	2531
S3B1II	0.94	885	0.70	1168	0.79	861	1.12	1298	0.528	861
S3B2II	0.82	1440	1.02	1083	0.82	857	1.30	1583	0.594	766
S4B1II	1.63	3236	1.74	3236	1.12	3243	1.88	3273	1.631	3077
S4B2II	1.17	566	1.06	833	1.18	503	1.16	680	1.679	503
S5B1II	3.24	5011	1.73	3855	2.20	5019	3.62	4313	2.226	3942
S5B2II	1.44	1696	1.26	994	1.60	1214	1.46	1434	1.458	1250
S6B1II	2.22	4223	1.86	4784	2.40	4220	2.59	4365	2.262	4929
S6B2II	1.49	2480	1.90	2788	1.84	2232	1.49	3032	1.460	2246
Avg.	**1.88**	**2835.58**	**1.64**	**2697.83**	**1.47**	**2355.58**	**1.77**	**2591.08**	**1.42**	**2342.50**
S1C1II	2.23	4380	2.88	3262	1.41	3568	1.82	3853	1.855	3276
S1C2II	1.61	3836	1.54	2462	1.86	3853	1.62	3488	1.475	2480
S2C1II	3.68	12563	4.07	12449	3.61	10879	3.92	12173	3.763	9602
S2C2II	2.59	5038	1.92	4861	2.86	5013	2.46	4666	1.600	4721
S3C1II	0.85	834	0.85	818	0.91	813	1.15	806	0.815	813
S3C2II	0.86	2033	0.71	2054	1.30	2111	0.85	2120	0.942	1499
S4C1II	1.92	3393	1.87	2317	1.40	2222	1.91	2721	1.650	2380
S4C2II	1.43	1571	0.97	707	1.43	1008	1.29	824	1.475	707
S5C1II	2.93	4085	2.13	4083	2.63	4066	2.60	4087	2.148	4004
S5C2II	1.58	2133	1.18	1352	1.90	2542	1.28	1287	1.361	1285
S6C1II	4.19	4513	3.46	4387	4.08	4524	4.07	4342	3.703	5063
S6C2II	2.55	1998	1.97	1876	2.27	1703	1.55	1719	1.561	1782
Avg.	**2.20**	**3864.75**	**1.96**	**3385.67**	**2.14**	**3525.17**	**2.04**	**3507.17**	**1.86**	**3134.33**
|T| = 3										
S1A1III	31.95	37773	31.17	30122	22.38	19994	31.80	24710	25.226	20862
S1A2III	3.65	3704	3.00	2628	3.14	2686	4.30	5254	2.953	2583
S2A1III	13.75	10064	14.93	12606	13.43	10293	13.50	11434	13.124	13316
S2A2III	1.81	1209	2.24	2117	2.00	1276	1.71	1270	1.035	1131
S3A1III	16.16	12495	15.37	11984	9.63	8019	14.38	10423	13.452	11210
S3A2III	2.72	2909	2.60	2479	2.42	2598	2.67	2764	2.854	2260

S4A1III	31.43	21255	28.72	26069	29.78	21941	34.28	21157	32.498	20842
S4A2III	2.39	1551	2.90	1768	2.93	1485	2.41	1195	2.987	1260
S5A1III	30.37	21216	29.67	28619	27.90	18007	32.84	21068	26.768	21752
S5A2III	2.90	1462	2.34	1221	2.93	1563	2.64	1580	3.006	1409
S6A1III	30.70	23123	34.00	16509	27.27	21484	35.20	18798	29.537	17956
S6B2III	15.67	7767	3.07	1128	14.03	7234	3.15	1600	3.865	1337
Avg.	**15.29**	**12044**	**14.17**	**11437.5**	**13.15**	**9715**	**14.91**	**10104.42**	**13.11**	**9659.83**
S1B1III	23.31	17570	24.62	16745	22.39	18552	20.90	16857	23.232	17933
S1B2III	3.70	3384	4.14	2852	4.27	3419	3.73	2686	3.956	3533
S2B1III	22.55	23058	23.41	20794	19.70	23436	19.52	20825	19.321	22393
S2B2III	6.30	5626	10.53	8258	6.22	5686	12.31	8394	6.928	6777
S3B1III	18.41	20252	24.30	22948	11.81	13570	18.27	19160	12.427	16339
S3B2III	7.42	5504	8.59	8040	6.35	5063	14.05	10719	7.559	4421
S4B1III	40.01	37022	27.79	32109	46.10	36362	46.81	36945	38.156	35426
S4B2III	19.38	17252	15.89	15342	20.87	15309	20.47	16353	23.658	17194
S5B1III	21.64	19304	17.47	14487	21.22	11448	22.53	11016	20.630	11076
S5B2III	5.07	5741	10.13	8583	6.02	3519	6.39	4073	4.816	5843
S6B1III	20.55	11859	18.91	15303	21.56	9650	22.23	16484	20.015	12008
S6B2III	15.67	7767	12.63	11288	14.03	7234	7.56	5673	17.073	7891
Avg.	**17.00**	**14528.25**	**16.53**	**14729.08**	**16.71**	**12770.67**	**17.90**	**14098.75**	**16.48**	**13402.83**
S1C1III	20.98	15413	22.94	11877	23.31	14536	20.52	19806	14.802	10525
S1C2III	5.26	3541	3.45	3048	3.70	2585	4.06	2693	4.766	3021
S2C1III	24.37	23362	17.95	23413	15.50	9601	18.20	21734	15.465	11774
S2C2III	11.52	13283	8.51	10136	11.61	13189	14.97	13225	12.779	11230
S3C1III	15.07	12891	14.73	12216	14.05	12870	17.88	12641	11.981	12261
S3C2III	9.71	7478	7.22	5226	6.35	5341	10.27	5524	6.408	5946
S4C1III	59.91	78469	49.63	71298	59.97	74492	60.55	77001	54.321	79918
S4C2III	28.07	30840	28.65	30536	25.35	30722	30.32	30614	28.854	30742
S5C1III	27.28	20826	19.36	10893	19.63	13132	20.56	10845	20.316	8023
S5C2III	11.77	7507	10.45	7302	9.74	6516	12.62	7241	11.810	6588
S6C1III	20.40	14008	16.70	14016	21.23	14693	20.97	13163	15.387	14737
S6C2III	15.24	14449	15.35	13826	14.79	14340	14.90	12807	13.272	13590
Avg.	**20.80**	**20172.25**	**17.91**	**17815.58**	**18.77**	**17668.08**	**20.49**	**18941.17**	**17.51**	**17362.92**
|T| = 4										
S1A1IV	55.36	27282	79.45	29033	50.86	24105	67.76	27103	52.24	24093
S1A2IV	23.94	13374	25.71	13260	11.88	4963	31.41	14168	26.09	13921
S2A1IV	31.24	14951	30.59	14714	29.14	11467	36.26	14082	24.07	9508
S2A2IV	33.21	22439	32.67	18411	23.89	14738	26.42	22374	25.99	14439
S3A1IV	43.15	20206	31.76	14403	23.09	8684	28.29	12905	16.02	5307
S3A2IV	21.15	13441	30.48	22961	17.39	16300	20.59	15194	21.70	12437
S4A1IV	106.82	53017	139.90	64083	108.97	49149	79.10	37012	108.70	41275
S4A2IV	31.05	13561	30.00	15460	30.38	13153	39.60	16946	30.77	13937
S5A1IV	273.38	162011	178.98	127699	231.25	126929	220.41	165436	177.97	120960
S5A2IV	34.05	17144	32.57	17266	30.54	13409	35.35	17410	37.61	12879
S6A1IV	43.01	17618	49.30	14549	17.32	6333	42.99	17448	48.33	17083
S6A2IV	6.51	2189	9.62	1766	6.43	818	8.07	2917	8.61	1162
Avg.	**58.57**	**31436.08**	**55.92**	**29467.08**	**48.43**	**24170.67**	**53.02**	**30249.58**	**48.17**	**23916.75**

S1B1IV	169.36	112006	90.17	33564	102.01	51102	115.24	73227	58.91	24998
S1B2IV	48.33	27033	38.83	17586	41.45	10777	37.09	19718	41.13	14184
S2B1IV	48.87	26559	40.01	20216	42.97	18159	41.41	22000	39.86	18632
S2B2IV	84.89	48423	96.30	41680	49.33	29263	66.35	42140	46.58	18737
S3B1IV	39.17	21495	44.63	21145	22.64	11038	32.40	18019	24.42	15331
S3B2IV	57.42	31575	35.68	30552	26.55	17338	55.76	17727	30.19	19417
S4B1IV	217.78	158263	222.59	123876	214.33	150472	250.46	165060	217.88	157565
S4B2IV	90.13	66629	95.61	57570	96.12	51662	97.22	56372	103.26	43504
S5B1IV	98.43	86175	97.36	112902	90.11	30975	95.83	162000	92.75	42709
S5B2IV	48.44	30347	43.76	23980	25.45	5934	41.83	20816	22.58	5858
S6B1IV	29.94	17129	25.79	12741	22.51	5740	22.03	27714	33.43	6407
S6B2IV	62.13	32986	55.67	32214	52.89	12055	70.62	21410	63.29	24399
Avg.	**82.91**	**54885**	**73.87**	**44002.17**	**65.53**	**32876.25**	**77.19**	**53850.25**	**64.52**	**32645.08**
S1C1IV	186.73	174828	205.83	164737	103.52	136320	174.60	152018	92.16	52519
S1C2IV	40.24	26267	37.24	25910	51.67	17674	44.23	26073	53.57	19792
S2C1IV	44.86	25107	32.07	29107	25.96	10494	49.38	28773	24.97	8838
S2C2IV	66.26	53341	30.36	32195	33.22	11176	53.05	31427	31.67	13959
S3C1IV	110.13	112725	125.73	214124	95.21	102006	111.16	202919	74.29	44883
S3C2IV	48.75	39039	44.17	13468	39.64	10287	52.87	35311	31.41	13705
S4C1IV	387.16	425617	427.66	157746	352.13	163895	323.84	156100	338.55	206554
S4C2IV	125.35	105828	94.91	73653	125.04	36229	106.75	82590	105.69	90893
S5C1IV	448.52	380460	447.27	362992	422.18	367431	426.26	366885	427.95	211908
S5C2IV	95.77	92562	85.64	66332	80.98	84050	116.88	74854	86.43	45368
S6C1IV	55.69	44158	52.07	54119	60.41	24873	60.74	54065	56.16	37542
S6C2IV	52.91	32257	54.09	25108	51.51	34599	52.35	30925	54.97	32919
Avg.	**138.53**	**126015.75**	**136.42**	**101624.25**	**120.12**	**83252.83**	**131.01**	**103495.00**	**114.82**	**64906.67**

Table 28 – Computational time and gap for the linear relaxation of the formulation ILP-(3.11) without and with inclusion of valid inequalities for instances generated with synthetic graphs.

Instance	ILP-(3.11)		ILP-(3.13)		ILP-(3.15)*		ILP-(3.17)		ILP-ALL	
	Time	GAP	Time	GAP	Time	GAP	Time	GAP	Time	GAP
|T|=2										
S1A1II	0.00	85.19	0.01	70.37	0.04	28.46	0.01	70.37	0.02	28.46
S1A2II	0.00	84.00	0.01	72.00	0.03	35.54	0.01	74.00	0.02	35.54
S2A1II	0.00	85.19	0.01	68.52	0.04	31.65	0.01	66.67	0.03	31.65
S2A2II	0.00	70.37	0.01	57.87	0.04	29.63	0.01	59.26	0.02	29.63
S3A1II	0.00	72.41	0.01	57.76	0.03	22.76	0.01	55.17	0.03	22.76
S3A2II	0.00	70.37	0.01	54.63	0.02	30.81	0.01	55.56	0.02	30.81
S4A1II	0.01	117.39	0.02	95.65	0.08	45.39	0.02	95.65	0.06	45.39
S4A2II	0.01	100.00	0.02	89.13	0.06	43.87	0.01	89.13	0.04	43.87
S5A1II	0.01	117.39	0.01	100.00	0.06	45.25	0.01	104.35	0.05	45.25
S5A2II	0.00	109.09	0.01	93.18	0.05	50.76	0.01	95.45	0.03	50.76
S6A1II	0.01	127.27	0.01	104.55	0.09	53.65	0.01	104.55	0.07	53.65
S6A2II	0.00	119.05	0.01	100.00	0.05	58.42	0.01	100.00	0.03	58.42
Avg.	**0.00**	**96.48**	**0.01**	**80.30**	**0.05**	**39.68**	**0.01**	**80.85**	**0.04**	**39.68**

S1B1II	0.00	85.19	0.01	70.37	0.04	28.46	0.01	70.37	0.03	28.46
S1B2II	0.01	77.78	0.01	66.67	0.05	27.58	0.01	62.96	0.02	27.58
S2B1II	0.00	85.19	0.01	66.67	0.11	31.65	0.01	66.67	0.03	31.65
S2B2II	0.00	77.78	0.01	62.96	0.04	29.63	0.01	62.96	0.03	29.63
S3B1II	0.01	72.41	0.01	51.72	0.04	22.76	0.01	53.45	0.03	22.76
S3B2II	0.01	71.43	0.01	53.57	0.04	25.58	0.01	55.36	0.03	25.58
S4B1II	0.00	117.39	0.01	97.83	0.08	45.39	0.03	95.65	0.07	45.39
S4B2II	0.01	108.70	0.01	91.30	0.06	44.48	0.01	89.13	0.05	44.48
S5B1II	0.00	117.39	0.01	104.35	0.09	45.25	0.01	104.35	0.05	45.25
S5B2II	0.01	108.70	0.01	95.65	0.05	44.66	0.01	95.65	0.04	44.66
S6B1II	0.00	127.27	0.01	104.55	0.07	53.65	0.01	104.55	0.08	53.65
S6B2II	0.00	128.57	0.01	109.52	0.11	60.36	0.01	104.76	0.05	60.36
Avg.	**0.00**	**98.15**	**0.01**	**81.26**	**0.07**	**38.29**	**0.01**	**80.49**	**0.04**	**38.29**
S1C1II	0.00	88.46	0.01	75.00	0.04	30.77	0.01	73.08	0.03	30.77
S1C2II	0.01	78.85	0.01	69.23	0.05	28.21	0.01	69.23	0.02	28.21
S2C1II	0.00	88.46	0.01	73.08	0.05	33.58	0.01	73.08	0.03	33.58
S2C2II	0.00	78.85	0.01	65.38	0.03	33.33	0.01	67.31	0.02	33.33
S3C1II	0.01	68.97	0.01	51.72	0.05	20.94	0.01	52.59	0.03	20.94
S3C2II	0.00	72.22	0.01	59.26	0.04	27.35	0.01	61.11	0.02	27.35
S4C1II	0.01	113.04	0.01	100.00	0.10	42.03	0.02	95.65	0.05	42.03
S4C2II	0.00	102.17	0.02	89.13	0.06	42.03	0.02	88.04	0.04	42.03
S5C1II	0.01	113.04	0.01	100.00	0.06	42.03	0.01	104.35	0.04	42.03
S5C2II	0.00	102.17	0.01	93.48	0.08	40.58	0.01	93.48	0.04	40.58
S6C1II	0.00	122.73	0.01	104.55	0.08	51.47	0.01	104.55	0.05	51.47
S6C2II	0.01	111.36	0.01	100.00	0.06	51.02	0.01	100.00	0.05	51.02
Avg.	**0.00**	**95.03**	**0.01**	**81.74**	**0.06**	**36.94**	**0.01**	**81.87**	**0.03**	**36.94**
|T| = 3										
S1A1III	0.01	85.19	0.03	77.78	0.11	28.46	0.02	77.78	0.07	28.46
S1A2III	0.01	71.35	0.02	84.62	0.08	33.32	0.02	69.90	0.05	33.32
S2A1III	0.04	66.67	0.03	58.33	0.11	18.48	0.06	56.67	0.09	18.48
S2A2III	0.02	54.63	0.02	62.07	0.07	22.54	0.02	49.54	0.04	22.54
S3A1III	0.04	66.67	0.02	56.67	0.07	18.67	0.05	57.78	0.06	18.67
S3A2III	0.02	55.17	0.03	62.07	0.06	22.76	0.02	52.80	0.05	22.76
S4A1III	0.14	100.00	0.03	88.00	0.41	33.76	0.10	88.00	0.23	33.76
S4A2III	0.06	85.00	0.03	95.83	0.23	39.32	0.04	81.94	0.10	39.32
S5A1III	0.09	100.00	0.02	92.00	0.19	33.63	0.07	92.00	0.12	33.63
S5A2III	0.04	85.00	0.03	100.00	0.09	39.20	0.05	81.06	0.09	39.20
S6A1III	0.09	100.00	0.03	84.00	0.19	35.21	0.05	84.00	0.12	35.21
S6A2III	0.03	85.00	0.04	91.67	0.17	40.84	0.05	77.39	0.09	40.84
Avg.	**0.05**	**79.56**	**0.03**	**79.42**	**0.15**	**30.52**	**0.05**	**72.40**	**0.09**	**30.52**
S1B1III	0.07	78.57	0.02	71.43	0.15	23.87	0.03	71.43	0.08	23.87
S1B2III	0.02	68.13	0.02	71.43	0.09	23.86	0.02	62.79	0.11	23.86
S2B1III	0.02	66.67	0.03	57.50	0.11	18.48	0.03	56.67	0.07	18.48
S2B2III	0.02	57.31	0.03	56.67	0.08	18.48	0.02	52.09	0.05	18.48

S3B1III	0.02	66.67	0.03	56.67	0.10	18.67	0.04	56.67	0.07	18.67
S3B2III	0.02	56.92	0.04	56.67	0.09	18.67	0.03	52.47	0.05	18.67
S4B1III	0.08	100.00	0.03	84.00	0.18	33.76	0.09	88.00	0.17	33.76
S4B2III	0.07	104.68	0.04	104.35	0.18	45.39	0.06	97.36	0.14	45.39
S5B1III	0.07	92.31	0.03	80.77	0.24	28.49	0.04	80.77	0.13	28.49
S5B2III	0.09	88.31	0.03	92.00	0.20	33.63	0.05	82.77	0.09	33.63
S6B1III	0.09	92.31	0.04	76.92	0.28	30.01	0.05	76.92	0.14	30.01
S6B2III	0.05	88.62	0.03	84.00	0.17	35.16	0.08	78.86	0.13	35.16
Avg.	**0.05**	**80.04**	**0.03**	**74.37**	**0.16**	**27.37**	**0.05**	**71.40**	**0.10**	**27.37**
S1C1III	0.03	75.00	0.04	67.86	0.15	21.43	0.02	67.86	0.10	21.43
S1C2III	0.01	65.60	0.02	67.86	0.11	21.43	0.02	62.07	0.11	21.43
S2C1III	0.06	63.33	0.05	56.67	0.10	15.77	0.02	56.67	0.06	15.77
S2C2III	0.02	59.89	0.03	59.77	0.07	19.76	0.03	55.46	0.06	19.76
S3C1III	0.04	63.33	0.04	56.67	0.12	16.90	0.05	56.67	0.05	16.90
S3C2III	0.03	54.53	0.04	56.67	0.09	16.90	0.04	52.94	0.06	16.90
S4C1III	0.09	104.17	0.04	95.83	0.25	36.11	0.07	95.83	0.13	36.11
S4C2III	0.04	101.57	0.02	104.35	0.21	42.03	0.07	93.48	0.14	42.03
S5C1III	0.03	88.46	0.04	84.62	0.14	25.64	0.03	84.62	0.12	25.64
S5C2III	0.02	85.44	0.04	92.00	0.14	30.67	0.04	82.00	0.07	30.67
S6C1III	0.10	88.46	0.03	76.92	0.23	28.17	0.06	76.92	0.11	28.17
S6C2III	0.07	85.44	0.03	84.00	0.21	33.29	0.08	78.67	0.09	33.29
Avg.	**0.04**	**77.94**	**0.03**	**75.27**	**0.15**	**25.67**	**0.04**	**71.93**	**0.09**	**25.67**
|T| = 4										
S1A1IV	0.02	85.19	0.03	81.48	0.16	28.46	0.05	81.48	0.17	28.46
S1A2IV	0.06	86.85	0.02	96.00	0.40	38.73	0.09	86.44	0.19	38.73
S2A1IV	0.06	61.29	0.03	56.99	0.11	14.66	0.09	56.99	0.13	14.66
S2A2IV	0.07	57.53	0.04	62.22	0.12	18.48	0.07	56.45	0.09	18.48
S3A1IV	0.04	61.29	0.03	58.06	0.22	14.84	0.06	58.06	0.11	14.84
S3A2IV	0.08	52.62	0.03	58.06	0.18	14.84	0.07	52.18	0.09	14.84
S4A1IV	0.22	100.00	0.08	92.00	0.51	33.76	0.10	88.00	0.27	33.76
S4A2IV	0.15	93.36	0.03	100.00	0.29	39.33	0.16	92.82	0.21	39.33
S5A1IV	0.19	100.00	0.05	96.00	0.29	33.63	0.09	96.00	0.31	33.63
S5A2IV	0.13	93.09	0.04	104.17	0.38	39.20	0.11	92.95	0.19	39.20
S6A1IV	0.13	92.31	0.06	80.77	0.37	30.01	0.13	80.77	0.34	30.01
S6A2IV	0.14	66.13	0.04	67.86	0.35	20.73	0.13	62.49	0.17	20.73
Avg.	**0.11**	**79.14**	**0.04**	**79.47**	**0.28**	**27.22**	**0.10**	**75.39**	**0.19**	**27.22**
S1B1IV	0.03	78.57	0.03	75.00	0.27	23.87	0.04	75.00	0.15	23.87
S1B2IV	0.05	71.72	0.03	75.00	0.19	23.87	0.06	69.88	0.16	23.87
S2B1IV	0.05	61.29	0.08	56.99	0.15	14.66	0.05	54.84	0.11	14.66
S2B2IV	0.10	55.72	0.04	58.06	0.16	14.66	0.07	54.31	0.09	14.66
S3B1IV	0.07	61.29	0.08	58.06	0.24	14.84	0.05	58.06	0.14	14.84
S3B2IV	0.11	55.26	0.04	58.06	0.22	14.84	0.10	53.40	0.10	14.84
S4B1IV	0.15	100.00	0.10	92.00	0.34	33.76	0.12	92.00	0.32	33.76
S4B2IV	0.18	101.11	0.04	100.00	0.39	39.33	0.14	97.45	0.24	39.33

S5B1IV	0.10	92.31	0.06	88.46	0.19	28.49	0.08	88.46	0.28	28.49
S5B2IV	0.12	85.34	0.03	88.46	0.20	28.49	0.16	83.88	0.17	28.49
S6B1IV	0.17	78.57	0.08	67.86	0.29	20.73	0.11	67.86	0.29	20.73
S6B2IV	0.13	78.39	0.05	74.07	0.34	25.20	0.16	70.75	0.23	25.20
Avg.	**0.11**	**76.63**	**0.06**	**74.34**	**0.25**	**23.56**	**0.09**	**72.16**	**0.19**	**23.56**
S1C1IV	0.02	75.00	0.04	71.43	0.29	21.43	0.03	75.00	0.17	21.43
S1C2IV	0.04	70.37	0.03	71.43	0.19	21.43	0.12	69.08	0.19	21.43
S2C1IV	0.05	58.06	0.05	58.06	0.16	12.03	0.05	54.84	0.12	12.03
S2C2IV	0.07	53.91	0.05	54.84	0.19	12.03	0.06	51.96	0.10	12.03
S3C1IV	0.06	63.33	0.08	61.67	0.15	16.90	0.07	61.67	0.15	16.90
S3C2IV	0.11	53.76	0.05	51.61	0.18	13.13	0.07	52.93	0.11	13.13
S4C1IV	0.16	96.00	0.07	92.00	0.41	30.67	0.18	92.00	0.27	30.67
S4C2IV	0.14	91.09	0.04	88.00	0.30	30.67	0.15	88.02	0.20	30.67
S5C1IV	0.05	96.00	0.06	96.00	0.36	30.67	0.06	96.00	0.18	30.67
S5C2IV	0.09	90.36	0.04	96.00	0.33	30.67	0.11	88.50	0.19	30.67
S6C1IV	0.18	81.48	0.06	74.07	0.44	23.42	0.12	74.07	0.27	23.42
S6C2IV	0.15	76.60	0.05	74.07	0.49	23.42	0.14	71.39	0.26	23.42
Avg.	**0.09**	**75.50**	**0.05**	**74.10**	**0.29**	**22.21**	**0.10**	**72.96**	**0.18**	**22.21**

Table 29 – Computational time and number of nodes used to solve instances generated with GÜLPINAR *et al.*'s graphs for formulation ILP - (3.11) without and with inclusion of all valid inequalities (ILP-ALL).

Category A																
T=|2|						T=|3|						T=|4|				
Instance	OTM	ILP-(3.11)		ILP-ALL		Instance	OTM	ILP-(3.11)		ILP-ALL		Instance	OTM	ILP-(3.11)		ILP-ALL	
		Time	Nodes	Time	Nodes			Time	Nodes	Time	Nodes			Time	Nodes	Time	Nodes
G1A1II	221	2.97	22	0.47	0	G1A1III	218	5.80	2314	5.54	2250	G1A1IV	219	110.19	7573	50.65	6574
G1A2II	219	0.29	0	0.05	0	G1A2III	219	0.30	0	0.05	0	G1A2IV	219	1.78	51	1.12	0
G2A1II	128	36.87	19577	36.55	19237	G2A1III	117	38.81	4387	31.54	4273	G2A1IV	125	380.71	93418	327.79	88680
G2A2II	121	0.23	0	0.27	0	G2A2III	134	0.76	353	0.83	361	G2A2IV	120	2.98	281	1.49	0
G3A1II	133	17.16	4632	16.29	4412	G3A1III	131	33.87	5049	30.63	4593	G3A1IV	129	335.48	133299	328.57	132602
G3A2II	123	0.06	0	0.05	0	G3A2III	135	0.96	580	0.54	158	G3A2IV	134	1.09	19	0.73	0
G4A1II	270	7.97	13	0.71	0	G4A1III	270	8.10	7	1.07	0	G4A1IV	269	28.13	532	5.09	0
G4A2II	255	0.08	0	0.09	0	G4A2III	268	0.81	0	0.30	0	G4A2IV	268	3.62	43	1.47	0
G5A1II	146	29.20	6710	23.39	7357	G5A1III	136	9.11	1285	8.58	1183	G5A1IV	144	103.43	7234	57.84	5364
G5A2II	138	0.32	0	0.19	0	G5A2III	148	7.17	6050	2.68	4526	G5A2IV	137	1.24	116	0.96	107
G6A1II	298	5.38	333	1.03	0	G6A1III	296	0.34	0	0.22	0	G6A1IV	298	24.67	5198	3.71	690
G6A2II	269	0.04	0	0.04	0	G6A2III	296	0.61	0	0.19	0	G6A2IV	292	0.09	0	0.09	0
G7A1II	144	0.78	1	0.04	0	G7A1III	143	0.13	0	0.05	0	G7A1IV	144	0.69	0	0.04	0
G7A2II	123	0.05	0	0.04	0	G7A2III	142	0.31	0	0.04	0	G7A2IV	141	0.04	0	0.02	0
G8A1II	189	0.05	0	0.15	0	G8A1III	188	0.11	0	0.04	0	G8A1IV	189	0.26	0	0.06	0
G8A2II	167	0.03	0	0.05	0	G8A2III	189	0.03	0	0.05	0	G8A2IV	188	0.02	0	0.05	0
G9A1II	147	0.70	0	0.04	0	G9A1III	147	0.29	0	0.04	0	G9A1IV	147	1.67	5	0.10	0
G9A2II	126	0.03	0	0.05	0	G9A2III	145	0.21	0	0.09	0	G9A2IV	144	0.04	0	0.05	0

G10A1II	270	2.84	90	0.23	0	G10A1III	267	0.30	0	0.13	0	G10A1IV	270	3.71	226	0.32	0
G10A2II	239	0.02	0	0.01	0	G10A2III	267	0.75	0	0.24	0	G10A2IV	262	0.05	0	0.05	0
G11A1II	201	1.03	18	0.22	0	G11A1III	201	0.10	0	0.04	0	G11A1IV	202	1.81	0	0.19	0
G11A2II	177	0.13	0	0.04	0	G11A2III	198	0.61	0	0.06	0	G11A2IV	195	0.21	0	0.16	0
G12A1II	201	0.19	0	0.13	0	G12A1III	198	0.17	0	0.14	0	G12A1IV	201	0.33	0	0.05	0
G12A2II	177	0.02	0	0.04	0	G12A2III	197	0.28	0	0.04	0	G12A2IV	195	0.04	0	0.07	0
G13A1II	337	0.89	0	0.12	0	G13A1III	332	0.50	0	0.23	0	G13A1IV	337	0.77	0	0.32	0
G13A2II	309	0.05	0	0.04	0	G13A2III	335	0.20	0	0.16	0	G13A2IV	330	0.09	0	0.08	0
G14A1II	206	0.24	0	0.19	0	G14A1III	203	0.10	0	0.05	0	G14A1IV	207	0.22	0	0.05	0
G14A2II	178	0.03	0	0.04	0	G14A2III	202	0.19	0	0.06	0	G14A2IV	199	0.06	0	0.05	0
G15A1II	266	7.48	250	1.14	0	G15A1III	262	4.33	21	1.05	0	G15A1IV	266	42.37	4294	2.46	55
G15A2II	233	0.05	0	0.05	0	G15A2III	263	2.69	5	1.10	26	G15A2IV	257	0.25	0	0.22	0
G16A1II	299	14.09	2625	13.45	2625	G16A1III	298	5.72	933	5.24	933	G16A1IV	299	64.04	4724	65.14	4724
G16A2II	262	0.02	0	0.03	0	G16A2III	300	0.38	0	0.29	0	G16A2IV	299	0.39	0	0.39	0
Avg.		**4.04**	**1070.97**	**2.98**	**1050.97**	**Avg.**		**3.88**	**655.75**	**2.85**	**571.97**	**Avg.**		**34.70**	**8031.66**	**26.54**	**7462.38**
G17A1II	495	4.08	36	2.33	0	G17A1III	501	3.84	327	1.09	0	G17A1IV	501	17.79	38	3.75	0
G17A2II	494	4.41	68	0.49	0	G17A2III	494	1.82	0	1.42	0	G17A2IV	493	2.15	0	1.96	0
G18A1II	499	5.87	32	8.54	0	G18A1III	523	5.96	31	5.09	0	G18A1IV	544	94.80	5391	16.99	305
G18A2II	497	94.98	7198	146.86	7406	G18A2III	493	8.05	5	8.10	5	G18A2IV	492	2.64	0	5.12	42
G19A1II	531	223.58	10573	213.89	8431	G19A1III	537	293.02	15725	90.98	5914	G19A1IV	535	198.94	6403	123.27	5204
G19A2II	461	10.79	1917	34.78	3918	G19A2III	486	138.67	7397	257.02	10441	G19A2IV	513	218.70	10134	131.09	9318
G20A1II	281	56.38	7214	51.21	5781	G20A1III	281	50.81	6313	44.38	4328	G20A1IV	289	117.76	9991	65.02	6979
G20A2II	273	74.83	19138	54.55	8232	G20A2III	265	3.64	447	2.63	370	G20A2IV	275	124.76	14780	92.69	11795
G21A1II	276	58.80	2060	56.65	2010	G21A1III	276	66.31	4383	48.20	2381	G21A1IV	277	156.97	5148	118.64	5031
G21A2II	275	10.10	915	6.68	514	G21A2III	265	9.93	429	7.63	275	G21A2IV	276	72.26	3668	70.20	3644
G22A1II	468	14.45	1610	15.95	1610	G22A1III	469	7.12	31	7.81	31	G22A1IV	467	64.60	3779	64.32	3779
G22A2II	468	1.12	45	1.18	45	G22A2III	468	1.14	49	1.25	49	G22A2IV	455	9.92	1141	10.95	1141
G23A1II	382	25.25	330	10.13	5	G23A1III	382	67.32	967	17.35	80	G23A1IV	382	30.54	119	10.52	0
G23A2II	368	99.74	5464	35.77	1801	G23A2III	368	98.71	5514	14.38	661	G23A2IV	382	23.20	232	9.09	65
Avg.		**48.88**	**4042.86**	**45.64**	**2839.50**	**Avg.**		**54.03**	**2972.71**	**36.24**	**1752.50**	**Avg.**		**81.07**	**4344.57**	**51.69**	**3378.79**

Category B

T=|2|						T=|3|						T=|4|					
Instance	OTM	ILP-(3.11)		ILP-ALL		Instance	OTM	ILP-(3.11)		ILP-ALL		Instance	OTM	ILP-(3.11)		ILP-ALL	
		Time	Nodes	Time	Nodes			Time	Nodes	Time	Nodes			Time	Nodes	Time	Nodes
G1B1II	221	2.64	21	0.50	0	G1B1III	218	5.56	1943	6.58	2097	G1B1IV	219	111.74	8328	61.07	6882
G1B2II	219	0.29	0	0.05	0	G1B2III	219	0.33	0	0.10	0	G1B2IV	218	2.15	206	1.63	167
G2B1II	130	39.45	19600	29.01	14736	G2B1III	120	38.77	5135	36.34	5016	G2B1IV	125	1056.97	246965	768.20	170445
G2B2II	121	0.24	0	0.17	0	G2B2III	135	0.59	102	0.53	148	G2B2IV	123	5.37	2311	4.32	1122
G3B1II	133	17.86	5445	14.79	3385	G3B1III	131	46.47	9964	38.78	9345	G3B1IV	129	591.12	257578	584.67	257299
G3B2II	120	0.47	0	0.37	0	G3B2III	134	0.89	470	0.83	424	G3B2IV	135	0.42	0	0.38	0
G4B1II	270	5.31	3	0.92	0	G4B1III	270	4.60	0	0.88	0	G4B1IV	269	21.43	353	6.25	5
G4B2II	256	0.04	0	0.11	0	G4B2III	268	0.80	0	0.61	0	G4B2IV	268	2.79	9	0.85	0
G5B1II	146	19.58	4689	19.03	4680	G5B1III	139	21.95	3431	21.20	3198	G5B1IV	143	89.97	8628	94.70	8740
G5B2II	139	0.21	0	0.17	0	G5B2III	148	4.22	4778	4.45	4929	G5B2IV	141	8.35	5171	8.28	2530
G6B1II	298	3.26	77	0.79	0	G6B1III	296	0.49	0	0.65	0	G6B1IV	298	6.99	1364	1.94	486
G6B2II	271	0.02	0	0.05	0	G6B2III	296	0.56	0	0.18	0	G6B2IV	294	0.04	0	0.05	0
G7B1II	144	0.92	7	0.06	0	G7B1III	143	0.18	0	0.04	0	G7B1IV	144	1.22	17	0.08	0

G7B2II	129	0.01	0	0.02	0	G7B2III	142	0.16	0	0.06	0	G7B2IV	142	0.12	0	0.07	0
G8B1II	189	0.09	0	0.03	0	G8B1III	188	0.07	0	0.09	0	G8B1IV	189	0.23	0	0.25	0
G8B2II	174	0.01	0	0.01	0	G8B2III	189	0.02	0	0.02	0	G8B2IV	188	0.09	0	0.02	0
G9B1II	147	0.96	0	0.05	0	G9B1III	147	0.17	0	0.03	0	G9B1IV	147	1.51	11	0.05	0
G9B2II	128	0.03	0	0.04	0	G9B2III	145	0.27	0	0.03	0	G9B2IV	144	0.06	0	0.07	0
G10B1II	270	2.45	535	0.28	0	G10B1III	267	0.31	0	0.22	0	G10B1IV	270	6.89	4886	0.59	0
G10B2II	247	0.02	0	0.02	0	G10B2III	267	1.10	0	0.50	0	G10B2IV	264	0.07	0	0.08	0
G11B1II	201	1.13	23	0.27	0	G11B1III	200	0.12	0	0.08	0	G11B1IV	202	1.02	0	0.27	0
G11B2II	178	0.02	0	0.02	0	G11B2III	198	0.27	0	0.19	0	G11B2IV	197	0.06	0	0.09	0
G12B1II	201	0.18	0	0.13	0	G12B1III	200	0.04	0	0.04	0	G12B1IV	201	0.26	0	0.06	0
G12B2II	183	0.01	0	0.02	0	G12B2III	197	0.27	0	0.07	0	G12B2IV	193	0.05	0	0.04	0
G13B1II	337	0.23	0	0.07	0	G13B1III	334	0.22	0	0.18	0	G13B1IV	337	0.83	0	0.25	0
G13B2II	317	0.02	0	0.04	0	G13B2III	335	0.19	0	0.15	0	G13B2IV	331	0.05	0	0.06	0
G14B1II	206	0.47	0	0.26	0	G14B1III	205	0.14	0	0.06	0	G14B1IV	207	0.26	0	0.27	0
G14B2II	186	0.02	0	0.04	0	G14B2III	202	0.22	0	0.16	0	G14B2IV	201	0.06	0	0.05	0
G15B1II	266	7.65	705	1.92	0	G15B1III	263	3.52	61	0.88	0	G15B1IV	266	38.91	4438	3.92	171
G15B2II	234	0.03	0	0.06	0	G15B2III	263	2.62	9	1.63	0	G15B2IV	258	0.21	0	0.22	0
G16B1II	299	29.78	6152	29.75	6152	G16B1III	299	1.65	0	1.66	0	G16B1IV	299	68.86	4622	65.01	4622
G16B2II	264	0.01	0	0.01	0	G16B2III	300	0.23	0	0.23	0	G16B2IV	298	1.46	685	1.46	685
Avg.		**4.17**	**1164.28**	**3.10**	**904.78**	**Avg.**		**4.28**	**809.16**	**3.67**	**786.16**	**Avg.**		**63.11**	**17049.13**	**50.16**	**14161.06**
G17B1II	500	3.82	0	3.52	0	G17B1III	501	5.37	36	1.44	0	G17B1IV	501	15.93	14	4.59	0
G17B2II	494	3.69	16	2.74	0	G17B2III	494	2.24	36	1.01	0	G17B2IV	498	3.76	0	1.61	0
G18B1II	513	39.74	1530	31.46	1590	G18B1III	528	52.10	2566	30.28	1755	G18B1IV	546	66.27	2590	11.43	270
G18B2II	503	294.61	15392	259.28	10267	G18B2III	499	148.72	9113	384.56	14777	G18B2IV	504	8.72	41	7.06	25
G19B1II	529	1456.61	86221	754.14	43691	G19B1III	537	551.82	37691	402.71	37317	G19B1IV	535	111.41	2227	19.20	1146
G19B2II	473	9.73	107	12.87	327	G19B2III	485	786.82	41434	502.81	35581	G19B2IV	514	68.96	3203	17.37	1568
G20B1II	282	122.54	13954	116.02	13579	G20B1III	281	62.18	6661	103.56	8641	G20B1IV	289	270.99	16566	60.73	9750
G20B2II	273	107.63	8481	101.10	8866	G20B2III	269	3.45	138	4.60	41	G20B2IV	274	264.26	33071	258.79	29676
G21B1II	277	27.45	1371	25.61	1103	G21B1III	276	71.87	4213	63.14	4892	G21B1IV	277	166.65	5768	110.00	3704
G21B2II	275	26.52	3571	10.37	868	G21B2III	275	19.41	2734	18.27	2511	G21B2IV	276	64.78	2886	148.57	5629
G22B1II	469	3.74	30	4.50	30	G22B1III	469	10.10	59	10.11	59	G22B1IV	468	55.29	2381	54.15	2381
G22B2II	468	1.46	70	1.51	70	G22B2III	468	1.46	31	1.39	31	G22B2IV	466	16.72	1947	16.58	1947
G23B1II	382	21.96	79	12.89	25	G23B1III	382	110.55	516	17.46	158	G23B1IV	382	33.44	0	13.00	0
G23B2II	368	108.56	5034	79.87	5028	G23B2III	368	101.00	6258	19.11	408	G23B2IV	382	18.07	56	5.10	0
Avg.		**159.15**	**9704**	**101.13**	**6103.14**	**Avg.**		**137.65**	**7963.29**	**111.46**	**7583.64**	**Avg.**		**83.23**	**5053.57**	**52.01**	**4006.86**

Category C

T=|2|						T=|3|						T=|4|				
Instance	OTM	ILP-(3.11)		ILP-ALL		Instance	OTM	ILP-(3.11)		ILP-ALL		Instance	OTM	ILP-(3.11)		ILP-ALL	
		Time	Nodes	Time	Nodes			Time	Nodes	Time	Nodes			Time	Nodes	Time	Nodes
G1C1II	221	2.72	18	0.63	0	G1C1III	218	4.90	373	4.49	318	G1C1IV	219	115.825783	7408	67.31	6065
G1C2II	219	0.34	0	0.17	0	G1C2III	219	0.40	0	0.06	0	G1C2IV	218	3.096699	499	2.65	123
G2C1II	130	26.11	9730	22.55	9565	G2C1III	122	56.45	10391	50.70	8508	G2C1IV	125	931.842102	359835	809.86	212014
G2C2II	124	0.11	0	0.08	0	G2C2III	135	0.70	453	0.44	66	G2C2IV	123	1.510156	232	1.31	95
G3C1II	133	28.30	9404	26.10	6734	G3C1III	132	73.86	23743	52.81	9079	G3C1IV	131	167.414536	47751	104.59	41924
G3C2II	124	0.28	106	0.28	0	G3C2III	135	0.80	482	0.84	597	G3C2IV	135	1.138822	25	1.68	70
G4C1II	270	8.63	51	0.62	0	G4C1III	270	6.25	6	1.47	0	G4C1IV	269	28.567719	951	8.03	0
G4C2II	255	0.19	0	0.16	0	G4C2III	268	1.13	0	0.51	0	G4C2IV	268	2.46696	64	1.31	0

G5C1II	146	19.75	4452	20.77	4510	G5C1III	133	13.81	2474	13.69	2424	G5C1IV	141	475.80249	79170	351.35	55693
G5C2II	135	0.13	0	0.13	0	G5C2III	148	5.44	7912	1.80	2541	G5C2IV	135	0.578477	0	0.26	0
G6C1II	298	2.60	64	0.80	0	G6C1III	296	2.17	0	0.77	0	G6C1IV	298	19.114811	3898	1.38	786
G6C2II	283	0.04	0	0.04	0	G6C2III	296	1.09	0	0.15	0	G6C2IV	291	0.22728	0	0.20	0
G7C1II	144	0.46	0	0.04	0	G7C1III	143	0.55	0	0.06	0	G7C1IV	144	0.829522	104	0.05	0
G7C2II	132	0.06	0	0.06	0	G7C2III	142	0.30	0	0.05	0	G7C2IV	141	0.25944	0	0.05	0
G8C1II	189	0.09	0	0.28	0	G8C1III	189	0.16	0	0.06	0	G8C1IV	189	0.348166	0	0.36	0
G8C2II	174	0.04	0	0.04	0	G8C2III	189	0.02	0	0.04	0	G8C2IV	188	0.02726	0	0.04	0
G9C1II	147	0.95	5	0.05	0	G9C1III	147	0.11	0	0.04	0	G9C1IV	147	1.263358	5	0.07	0
G9C2II	132	0.02	0	0.05	0	G9C2III	145	0.34	0	0.07	0	G9C2IV	144	0.072396	0	0.05	0
G10C1II	270	3.16	7	0.25	0	G10C1III	268	0.67	0	0.38	0	G10C1IV	270	5.957949	3670	0.53	0
G10C2II	243	0.03	0	0.04	0	G10C2III	267	0.77	0	0.25	0	G10C2IV	262	0.06343	0	0.05	0
G11C1II	201	1.39	70	0.22	0	G11C1III	200	0.49	0	0.21	0	G11C1IV	202	1.542753	237	0.22	0
G11C2II	183	0.11	0	0.08	0	G11C2III	198	0.89	0	0.17	0	G11C2IV	198	0.104456	0	0.08	0
G12C1II	201	0.20	0	0.04	0	G12C1III	200	0.16	0	0.04	0	G12C1IV	201	0.231713	0	0.06	0
G12C2II	183	0.05	0	0.04	0	G12C2III	197	0.27	0	0.15	0	G12C2IV	197	0.031756	0	0.04	0
G13C1II	337	0.28	0	0.08	0	G13C1III	335	0.14	0	0.16	0	G13C1IV	337	0.540603	0	0.27	0
G13C2II	318	0.04	0	0.03	0	G13C2III	335	0.20	0	0.09	0	G13C2IV	331	0.109995	0	0.07	0
G14C1II	206	0.30	0	0.22	0	G14C1III	203	0.22	0	0.18	0	G14C1IV	207	0.235714	0	0.17	0
G14C2II	188	0.02	0	0.04	0	G14C2III	202	0.32	0	0.05	0	G14C2IV	200	0.05984	0	0.09	0
G15C1II	266	6.62	189	1.49	9	G15C1III	266	4.23	14	0.73	0	G15C1IV	266	87.49646	11766	2.18	0
G15C2II	239	0.06	0	0.10	0	G15C2III	263	2.08	15	1.01	15	G15C2IV	261	0.428919	0	0.24	0
G16C1II	299	32.14	4782	32.05	4782	G16C1III	299	11.58	2632	11.39	2632	G16C1IV	299	115.182243	8975	115.04	8975
G16C2II	283	0.02	0	0.02	0	G16C2III	300	0.28	0	0.26	0	G16C2IV	300	0.283495	0	0.27	0
Avg.			**902.44**	**3.36**	**800.00**	**Avg.**		**5.96**	**1515.47**	**4.47**	**818.13**	**Avg.**		**61.33**	**16393.44**	**45.93**	**10179.53**
G17C1II	500	5.71	0	4.56	0	G17C1III	501	4.95	47	1.21	0	G17C1IV	501	20.460629	27	3.37	0
G17C2II	494	4.61	29	3.29	26	G17C2III	494	3.27	25	2.54	24	G17C2IV	500	7.461777	19	2.09	0
G18C1II	525	120.59	4646	62.51	3470	G18C1III	542	3410.04	150184	117.14	2473	G18C1IV	556	809.607666	25643	30.21	453
G18C2II	522	484.40	8293	459.87	7399	G18C2III	511	196.04	6911	168.26	5751	G18C2IV	517	8.946139	264	7.35	0
G19C1II	531	8043.48	35943	711.02	30948	G19C1III	538	269.99	20319	85.41	4856	G19C1IV	538	114.011116	2850	8.84	5
G19C2II	483	148.42	6928	114.53	6088	G19C2III	493	5000.53	176682	5002.13	197390	G19C2IV	522	2081.410645	176084	697.45	48508
G20C1II	283	48.78	5091	40.58	3566	G20C1III	283	81.15	11973	110.95	12318	G20C1IV	289	234.032639	16305	74.59	8914
G20C2II	273	154.19	16716	151.03	16295	G20C2III	268	3.89	689	9.83	2565	G20C2IV	277	303.855499	39880	499.36	82327
G21C1II	276	63.63	3272	47.73	1572	G21C1III	277	63.53	2423	63.46	2890	G21C1IV	277	174.454361	6468	162.06	6096
G21C2II	275	17.84	2896	17.17	2782	G21C2III	275	37.25	4860	25.15	1514	G21C2IV	276	97.143639	4121	206.96	6545
G22C1II	469	3.10	0	3.19	0	G22C1III	469	6.03	0	6.05	0	G22C1IV	468	152.500549	6200	153.48	6200
G22C2II	469	0.69	0	0.61	0	G22C2III	469	0.64	0	0.61	0	G22C2IV	467	75.31797	4742	75.58	4742
G23C1II	382	32.10	9	14.65	0	G23C1III	382	87.20	1790	19.13	13	G23C1IV	382	83.868576	1348	12.73	0
G23C2II	368	92.42	4200	22.19	689	G23C2III	368	91.69	4239	30.68	2223	G23C2IV	382	31.333338	299	11.61	153
Avg.		**658.57**	**6287.36**	**118.07**	**5202.50**	**Avg.**		**661.16**	**27153.00**	**403.04**	**16572.64**	**Avg.**		**299.60**	**20303.57**	**138.98**	**11710.21**

Table 30 – Computational time and gap for the linear relaxation on instances generated with GÜLPINAR *et al.*'s graphs for formulation ILP - (68) without and with inclusion of all valid inequalities.

Category A

T=|2|					T=|3|				T=|4|				
Instance	ILP-(3.11)		ILP-ALL		Instance	ILP-(3.11)		ILP-ALL		Instance	ILP-(3.11)		ILP-ALL	
	Time	Gap (%)	Time	Gap (%)		Time	Gap (%)	Time	Gap (%)		Time	Gap (%)	Time	Gap (%)
G1A1II	0.04	0.01	0.04	0.00	G1A1III	0.11	0.01	0.11	0.01	G1A1IV	0.12	0.02	0.15	0.01
G1A2II	0.01	0.02	0.01	0.00	G1A2III	0.01	0.02	0.01	0.00	G1A2IV	0.04	0.01	0.05	0.01
G2A1II	0.03	0.10	0.04	0.09	G2A1III	0.15	0.09	0.16	0.09	G2A1IV	0.08	0.13	0.09	0.12
G2A2II	0.01	0.02	0.01	0.02	G2A2III	0.01	0.05	0.01	0.04	G2A2IV	0.05	0.04	0.10	0.04
G3A1II	0.06	0.06	0.04	0.05	G3A1III	0.26	0.05	0.28	0.05	G3A1IV	0.12	0.09	0.14	0.09
G3A2II	0.01	0.01	0.01	0.01	G3A2III	0.02	0.04	0.01	0.04	G3A2IV	0.12	0.02	0.09	0.02
G4A1II	0.07	0.02	0.09	0.00	G4A1III	0.09	0.02	0.12	0.00	G4A1IV	0.13	0.02	0.09	0.01
G4A2II	0.01	0.00	0.01	0.00	G4A2III	0.02	0.02	0.03	0.00	G4A2IV	0.05	0.02	0.05	0.01
G5A1II	0.05	0.08	0.06	0.07	G5A1III	0.24	0.04	0.26	0.03	G5A1IV	0.16	0.09	0.15	0.08
G5A2II	0.02	0.02	0.02	0.02	G5A2III	0.01	0.06	0.01	0.05	G5A2IV	0.09	0.02	0.07	0.02
G6A1II	0.04	0.02	0.05	0.01	G6A1III	0.04	0.01	0.04	0.00	G6A1IV	0.05	0.02	0.06	0.01
G6A2II	0.00	0.00	0.00	0.00	G6A2III	0.01	0.03	0.02	0.00	G6A2IV	0.02	0.01	0.02	0.01
G7A1II	0.01	0.03	0.01	0.00	G7A1III	0.01	0.02	0.01	0.01	G7A1IV	0.01	0.03	0.01	0.00
G7A2II	0.01	0.01	0.00	0.01	G7A2III	0.00	0.04	0.00	0.00	G7A2IV	0.01	0.00	0.01	0.00
G8A1II	0.01	0.01	0.01	0.00	G8A1III	0.01	0.00	0.01	0.00	G8A1IV	0.03	0.01	0.03	0.00
G8A2II	0.00	0.00	0.00	0.00	G8A2III	0.00	0.01	0.00	0.00	G8A2IV	0.01	0.00	0.01	0.00
G9A1II	0.03	0.02	0.01	0.00	G9A1III	0.01	0.00	0.01	0.00	G9A1IV	0.03	0.02	0.02	0.00
G9A2II	0.00	0.00	0.00	0.00	G9A2III	0.00	0.03	0.01	0.00	G9A2IV	0.01	0.00	0.01	0.00
G10A1II	0.02	0.02	0.03	0.00	G10A1III	0.04	0.01	0.02	0.01	G10A1IV	0.04	0.02	0.06	0.00
G10A2II	0.00	0.00	0.00	0.00	G10A2III	0.01	0.03	0.01	0.01	G10A2IV	0.01	0.01	0.01	0.01
G11A1II	0.02	0.02	0.02	0.00	G11A1III	0.01	0.00	0.02	0.00	G11A1IV	0.03	0.01	0.03	0.00
G11A2II	0.00	0.00	0.00	0.00	G11A2III	0.01	0.04	0.01	0.01	G11A2IV	0.01	0.01	0.01	0.01
G12A1II	0.02	0.00	0.02	0.00	G12A1III	0.02	0.01	0.01	0.01	G12A1IV	0.02	0.00	0.01	0.00
G12A2II	0.00	0.00	0.00	0.00	G12A2III	0.01	0.03	0.01	0.00	G12A2IV	0.01	0.00	0.01	0.00
G13A1II	0.03	0.01	0.03	0.00	G13A1III	0.04	0.01	0.03	0.01	G13A1IV	0.07	0.01	0.06	0.00
G13A2II	0.00	0.00	0.00	0.00	G13A2III	0.02	0.01	0.01	0.00	G13A2IV	0.02	0.01	0.01	0.01
G14A1II	0.02	0.01	0.02	0.00	G14A1III	0.01	0.01	0.01	0.00	G14A1IV	0.03	0.01	0.02	0.00
G14A2II	0.00	0.00	0.00	0.00	G14A2III	0.01	0.03	0.01	0.00	G14A2IV	0.01	0.00	0.01	0.00
G15A1II	0.05	0.06	0.06	0.01	G15A1III	0.08	0.04	0.04	0.01	G15A1IV	0.08	0.06	0.09	0.01
G15A2II	0.00	0.00	0.01	0.00	G15A2III	0.01	0.07	0.02	0.00	G15A2IV	0.02	0.02	0.02	0.01
G16A1II	0.02	0.01	0.02	0.01	G16A1III	0.09	0.01	0.08	0.01	G16A1IV	0.06	0.01	0.06	0.01
G16A2II	0.00	0.00	0.00	0.00	G16A2III	0.01	0.00	0.01	0.00	G16A2IV	0.02	0.00	0.02	0.00
Avg.	**0.02**	**0.02**	**0.02**	**0.01**	**Avg.**	**0.04**	**0.03**	**0.04**	**0.01**	**Avg.**	**0.05**	**0.02**	**0.05**	**0.02**
G17A1II	0.23	0.00	0.25	0.00	G17A1III	0.21	0.01	0.21	0.00	G17A1IV	0.28	0.01	0.30	0.00
G17A2II	0.05	0.02	0.07	0.00	G17A2III	0.03	0.02	0.05	0.00	G17A2IV	0.22	0.01	0.24	0.00
G18A1II	0.06	0.06	0.35	0.03	G18A1III	0.22	0.08	0.43	0.04	G18A1IV	0.30	0.08	0.75	0.04
G18A2II	0.02	0.08	0.21	0.04	G18A2III	0.02	0.07	0.21	0.03	G18A2IV	0.04	0.05	0.19	0.03
G19A1II	0.32	0.11	0.70	0.02	G19A1III	0.41	0.10	0.45	0.01	G19A1IV	0.27	0.10	0.44	0.01
G19A2II	0.02	0.08	0.17	0.03	G19A2III	0.02	0.11	0.43	0.04	G19A2IV	0.15	0.08	0.48	0.02
G20A1II	0.06	0.06	0.12	0.04	G20A1III	0.09	0.06	0.16	0.04	G20A1IV	0.21	0.04	0.30	0.01
G20A2II	0.04	0.09	0.18	0.05	G20A2III	0.01	0.07	0.07	0.04	G20A2IV	0.03	0.09	0.13	0.06
G21A1II	0.66	0.07	1.22	0.03	G21A1III	0.52	0.08	0.58	0.03	G21A1IV	1.23	0.07	1.01	0.03
G21A2II	0.08	0.08	0.41	0.03	G21A2III	0.09	0.12	0.44	0.07	G21A2IV	0.34	0.06	0.85	0.03
G22A1II	0.12	0.00	0.11	0.00	G22A1III	0.52	0.00	0.54	0.00	G22A1IV	0.63	0.01	0.57	0.01

G22A2II	0.05	0.00	0.05	0.00	G22A2III	0.06	0.00	0.07	0.00	G22A2IV	0.24	0.01	0.25	0.01
G23A1II	0.36	0.05	0.33	0.01	G23A1III	1.23	0.05	2.33	0.01	G23A1IV	1.02	0.02	1.82	0.01
G23A2II	0.04	0.09	0.13	0.02	G23A2III	0.03	0.09	0.13	0.02	G23A2IV	0.35	0.05	0.45	0.01
Avg.	**0.15**	**0.06**	**0.31**	**0.02**	**Avg.**	**0.25**	**0.06**	**0.44**	**0.02**	**Avg.**	**0.38**	**0.05**	**0.56**	**0.02**

Category B

T=|2|					T=|3|				T=|4|				
Instance	ILP-(3.11)		ILP-ALL		Instance	ILP-(3.11)		ILP-ALL		Instance	ILP-(3.11)		ILP-ALL	
	Time	Gap (%)	Time	Gap (%)		Time	Gap (%)	Time	Gap (%)		Time	Gap (%)	Time	Gap (%)
G1B1II	0.04	0.01	0.04	0.00	G1B1III	0.13	0.02	0.11	0.01	G1B1IV	0.12	0.02	0.11	0.01
G1B2II	0.01	0.02	0.01	0.00	G1B2III	0.01	0.02	0.01	0.00	G1B2IV	0.04	0.02	0.06	0.01
G2B1II	0.02	0.08	0.03	0.08	G2B1III	0.14	0.09	0.16	0.09	G2B1IV	0.07	0.13	0.07	0.12
G2B2II	0.01	0.01	0.01	0.01	G2B2III	0.01	0.04	0.01	0.04	G2B2IV	0.05	0.04	0.05	0.04
G3B1II	0.04	0.06	0.09	0.05	G3B1III	0.33	0.05	0.40	0.05	G3B1IV	0.07	0.09	0.08	0.09
G3B2II	0.01	0.02	0.01	0.02	G3B2III	0.01	0.05	0.02	0.04	G3B2IV	0.09	0.02	0.11	0.02
G4B1II	0.06	0.02	0.09	0.00	G4B1III	0.11	0.02	0.07	0.00	G4B1IV	0.15	0.02	0.16	0.01
G4B2II	0.01	0.00	0.01	0.00	G4B2III	0.03	0.02	0.02	0.00	G4B2IV	0.04	0.02	0.04	0.01
G5B1II	0.06	0.08	0.05	0.07	G5B1III	0.52	0.05	0.39	0.04	G5B1IV	0.12	0.10	0.21	0.09
G5B2II	0.01	0.02	0.02	0.02	G5B2III	0.01	0.06	0.01	0.05	G5B2IV	0.14	0.04	0.11	0.03
G6B1II	0.02	0.02	0.04	0.00	G6B1III	0.03	0.02	0.04	0.01	G6B1IV	0.04	0.02	0.04	0.01
G6B2II	0.00	0.00	0.01	0.00	G6B2III	0.01	0.03	0.02	0.00	G6B2IV	0.01	0.00	0.02	0.00
G7B1II	0.01	0.03	0.01	0.00	G7B1III	0.01	0.02	0.01	0.01	G7B1IV	0.01	0.03	0.01	0.00
G7B2II	0.00	0.00	0.00	0.00	G7B2III	0.00	0.04	0.01	0.00	G7B2IV	0.01	0.02	0.01	0.01
G8B1II	0.01	0.01	0.01	0.00	G8B1III	0.01	0.00	0.01	0.00	G8B1IV	0.03	0.01	0.04	0.00
G8B2II	0.00	0.00	0.00	0.00	G8B2III	0.01	0.01	0.00	0.00	G8B2IV	0.01	0.00	0.00	0.00
G9B1II	0.01	0.02	0.01	0.00	G9B1III	0.01	0.00	0.01	0.00	G9B1IV	0.02	0.02	0.01	0.00
G9B2II	0.00	0.00	0.00	0.00	G9B2III	0.00	0.03	0.01	0.00	G9B2IV	0.01	0.00	0.01	0.00
G10B1II	0.02	0.02	0.02	0.00	G10B1III	0.03	0.01	0.02	0.01	G10B1IV	0.03	0.02	0.03	0.00
G10B2II	0.00	0.00	0.00	0.00	G10B2III	0.01	0.03	0.01	0.01	G10B2IV	0.01	0.01	0.01	0.00
G11B1II	0.02	0.02	0.02	0.00	G11B1III	0.01	0.01	0.02	0.00	G11B1IV	0.02	0.01	0.02	0.00
G11B2II	0.00	0.00	0.00	0.00	G11B2III	0.01	0.04	0.01	0.01	G11B2IV	0.01	0.01	0.01	0.01
G12B1II	0.01	0.00	0.01	0.00	G12B1III	0.01	0.00	0.01	0.00	G12B1IV	0.02	0.00	0.02	0.00
G12B2II	0.00	0.00	0.00	0.00	G12B2III	0.01	0.03	0.01	0.00	G12B2IV	0.01	0.01	0.01	0.01
G13B1II	0.02	0.01	0.02	0.00	G13B1III	0.03	0.01	0.02	0.01	G13B1IV	0.05	0.01	0.04	0.00
G13B2II	0.00	0.00	0.01	0.00	G13B2III	0.02	0.01	0.02	0.00	G13B2IV	0.03	0.01	0.02	0.01
G14B1II	0.01	0.01	0.01	0.00	G14B1III	0.01	0.01	0.01	0.00	G14B1IV	0.02	0.01	0.02	0.00
G14B2II	0.00	0.00	0.00	0.00	G14B2III	0.01	0.03	0.01	0.00	G14B2IV	0.01	0.01	0.01	0.01
G15B1II	0.05	0.06	0.05	0.01	G15B1III	0.05	0.04	0.05	0.01	G15B1IV	0.10	0.06	0.10	0.01
G15B2II	0.01	0.00	0.01	0.00	G15B2III	0.01	0.07	0.03	0.00	G15B2IV	0.02	0.02	0.02	0.01
G16B1II	0.03	0.01	0.03	0.01	G16B1III	0.08	0.01	0.09	0.01	G16B1IV	0.07	0.01	0.07	0.01
G16B2II	0.00	0.00	0.01	0.00	G16B2III	0.02	0.00	0.02	0.00	G16B2IV	0.02	0.01	0.02	0.01
Avg.	**0.02**	**0.02**	**0.02**	**0.01**	**Avg.**	**0.05**	**0.03**	**0.05**	**0.01**	**Avg.**	**0.04**	**0.02**	**0.05**	**0.02**

Instance	ILP-(3.11)		ILP-ALL		Instance	ILP-(3.11)		ILP-ALL		Instance	ILP-(3.11)		ILP-ALL	
	Time	Gap (%)	Time	Gap (%)		Time	Gap (%)	Time	Gap (%)		Time	Gap (%)	Time	Gap (%)
G17B1II	0.27	0.00	0.25	0.00	G17B1III	0.22	0.01	0.17	0.00	G17B1IV	0.25	0.01	0.35	0.00
G17B2II	0.03	0.02	0.07	0.00	G17B2III	0.02	0.02	0.04	0.00	G17B2IV	0.28	0.00	0.31	0.00
G18B1II	0.07	0.07	0.62	0.03	G18B1III	0.24	0.08	0.52	0.04	G18B1IV	0.36	0.08	0.81	0.03

G18B2II	0.02	0.10	0.32	0.05	G18B2III	0.02	0.09	0.25	0.05	G18B2IV	0.10	0.06	0.70	0.04
G19B1II	0.29	0.11	0.49	0.02	G19B1III	0.48	0.10	0.73	0.01	G19B1IV	0.29	0.10	0.44	0.01
G19B2II	0.02	0.09	0.27	0.03	G19B2III	0.02	0.12	0.45	0.05	G19B2IV	0.17	0.08	0.49	0.01
G20B1II	0.05	0.07	0.11	0.03	G20B1III	0.10	0.06	0.15	0.04	G20B1IV	0.11	0.04	0.37	0.01
G20B2II	0.04	0.09	0.20	0.06	G20B2III	0.01	0.06	0.11	0.03	G20B2IV	0.04	0.09	0.16	0.06
G21B1II	0.69	0.07	1.34	0.03	G21B1III	0.54	0.08	0.43	0.03	G21B1IV	0.60	0.07	0.50	0.03
G21B2II	0.08	0.08	0.43	0.03	G21B2III	0.08	0.09	0.39	0.03	G21B2IV	0.30	0.06	0.85	0.03
G22B1II	0.12	0.00	0.12	0.00	G22B1III	0.25	0.00	0.22	0.00	G22B1IV	0.21	0.00	0.22	0.00
G22B2II	0.05	0.00	0.04	0.00	G22B2III	0.05	0.00	0.04	0.00	G22B2IV	0.25	0.00	0.25	0.00
G23B1II	0.42	0.05	0.52	0.01	G23B1III	0.99	0.05	2.64	0.01	G23B1IV	3.25	0.02	1.89	0.01
G23B2II	0.07	0.09	0.28	0.02	G23B2III	0.07	0.09	0.21	0.02	G23B2IV	0.57	0.05	0.99	0.01
Avg.	**0.16**	**0.06**	**0.36**	**0.02**	**Avg.**	**0.22**	**0.06**	**0.45**	**0.02**	**Avg.**	**0.48**	**0.05**	**0.60**	**0.02**
Category C														
T=|2|					T=|3|				T=|4|				
G1C1II	0.03	0.01	0.04	0.00	G1C1III	0.14	0.02	0.10	0.01	G1C1IV	0.10	0.02	0.16	0.01
G1C2II	0.01	0.02	0.02	0.00	G1C2III	0.01	0.02	0.02	0.00	G1C2IV	0.05	0.02	0.06	0.01
G2C1II	0.03	0.08	0.03	0.08	G2C1III	0.11	0.08	0.11	0.08	G2C1IV	0.06	0.13	0.06	0.12
G2C2II	0.01	0.01	0.01	0.01	G2C2III	0.01	0.04	0.01	0.04	G2C2IV	0.10	0.04	0.06	0.04
G3C1II	0.03	0.06	0.05	0.05	G3C1III	0.17	0.05	0.13	0.05	G3C1IV	0.06	0.08	0.08	0.07
G3C2II	0.01	0.01	0.07	0.01	G3C2III	0.01	0.04	0.03	0.04	G3C2IV	0.05	0.02	0.08	0.02
G4C1II	0.07	0.02	0.08	0.00	G4C1III	0.10	0.02	0.10	0.00	G4C1IV	0.21	0.02	0.14	0.01
G4C2II	0.01	0.00	0.01	0.00	G4C2III	0.05	0.02	0.04	0.00	G4C2IV	0.05	0.02	0.05	0.01
G5C1II	0.07	0.07	0.06	0.07	G5C1III	0.14	0.04	0.12	0.04	G5C1IV	0.15	0.11	0.12	0.11
G5C2II	0.02	0.01	0.01	0.01	G5C2III	0.01	0.06	0.01	0.05	G5C2IV	0.06	0.02	0.07	0.01
G6C1II	0.02	0.02	0.03	0.00	G6C1III	0.04	0.02	0.04	0.01	G6C1IV	0.04	0.02	0.05	0.01
G6C2II	0.00	0.00	0.01	0.00	G6C2III	0.01	0.03	0.02	0.00	G6C2IV	0.02	0.02	0.03	0.01
G7C1II	0.01	0.03	0.01	0.00	G7C1III	0.01	0.03	0.01	0.01	G7C1IV	0.01	0.03	0.01	0.00
G7C2II	0.00	0.01	0.00	0.00	G7C2III	0.00	0.04	0.01	0.00	G7C2IV	0.01	0.02	0.01	0.01
G8C1II	0.01	0.01	0.02	0.00	G8C1III	0.02	0.00	0.01	0.00	G8C1IV	0.03	0.01	0.02	0.00
G8C2II	0.00	0.00	0.00	0.00	G8C2III	0.01	0.01	0.00	0.00	G8C2IV	0.01	0.00	0.01	0.00
G9C1II	0.01	0.02	0.01	0.00	G9C1III	0.01	0.00	0.01	0.00	G9C1IV	0.02	0.02	0.02	0.00
G9C2II	0.00	0.00	0.00	0.00	G9C2III	0.01	0.03	0.01	0.00	G9C2IV	0.01	0.00	0.01	0.00
G10C1II	0.02	0.02	0.02	0.00	G10C1III	0.03	0.01	0.02	0.01	G10C1IV	0.03	0.02	0.05	0.00
G10C2II	0.00	0.00	0.00	0.00	G10C2III	0.01	0.03	0.02	0.01	G10C2IV	0.01	0.01	0.01	0.01
G11C1II	0.01	0.02	0.02	0.00	G11C1III	0.03	0.02	0.01	0.01	G11C1IV	0.02	0.01	0.03	0.00
G11C2II	0.00	0.01	0.01	0.00	G11C2III	0.01	0.04	0.01	0.01	G11C2IV	0.01	0.01	0.01	0.01
G12C1II	0.01	0.00	0.01	0.00	G12C1III	0.01	0.00	0.02	0.00	G12C1IV	0.02	0.00	0.02	0.00
G12C2II	0.00	0.00	0.01	0.00	G12C2III	0.01	0.03	0.01	0.00	G12C2IV	0.01	0.00	0.01	0.00
G13C1II	0.02	0.01	0.01	0.00	G13C1III	0.03	0.01	0.03	0.00	G13C1IV	0.05	0.01	0.02	0.00
G13C2II	0.01	0.00	0.01	0.00	G13C2III	0.01	0.01	0.02	0.00	G13C2IV	0.02	0.01	0.01	0.01
G14C1II	0.01	0.01	0.01	0.00	G14C1III	0.02	0.02	0.01	0.01	G14C1IV	0.02	0.01	0.02	0.00
G14C2II	0.00	0.00	0.00	0.00	G14C2III	0.01	0.03	0.01	0.00	G14C2IV	0.01	0.01	0.01	0.01
G15C1II	0.04	0.06	0.05	0.01	G15C1III	0.08	0.04	0.06	0.00	G15C1IV	0.08	0.06	0.07	0.01
G15C2II	0.01	0.01	0.01	0.00	G15C2III	0.01	0.07	0.03	0.00	G15C2IV	0.03	0.02	0.03	0.01
G16C1II	0.05	0.01	0.06	0.01	G16C1III	0.13	0.01	0.10	0.01	G16C1IV	0.14	0.01	0.13	0.01
G16C2II	0.00	0.00	0.00	0.00	G16C2III	0.03	0.00	0.03	0.00	G16C2IV	0.04	0.00	0.04	0.00
Avg.	**0.02**	**0.02**	**0.02**	**0.01**	**Avg.**	**0.04**	**0.03**	**0.04**	**0.01**	**Avg.**	**0.05**	**0.02**	**0.05**	**0.02**

G17C1II	0.25	0.01	0.26	0.00	G17C1III	0.22	0.01	0.27	0.00	G17C1IV	0.27	0.01	0.39	0.00
G17C2II	0.05	0.02	0.07	0.00	G17C2III	0.03	0.02	0.05	0.00	G17C2IV	0.30	0.01	0.42	0.00
G18C1II	0.13	0.08	0.78	0.03	G18C1III	0.32	0.08	0.71	0.04	G18C1IV	0.45	0.07	0.99	0.03
G18C2II	0.03	0.10	0.71	0.04	G18C2III	0.02	0.11	0.60	0.05	G18C2IV	0.12	0.06	0.69	0.03
G19C1II	0.37	0.11	0.65	0.02	G19C1III	0.44	0.10	0.51	0.00	G19C1IV	0.37	0.10	0.59	0.00
G19C2II	0.03	0.12	0.40	0.04	G19C2III	0.03	0.15	0.46	0.04	G19C2IV	0.26	0.09	0.74	0.02
G20C1II	0.09	0.06	0.21	0.03	G20C1III	0.12	0.06	0.13	0.03	G20C1IV	0.18	0.04	0.50	0.01
G20C2II	0.03	0.10	0.15	0.06	G20C2III	0.01	0.06	0.10	0.04	G20C2IV	0.05	0.08	0.15	0.05
G21C1II	0.32	0.07	0.85	0.03	G21C1III	0.53	0.08	0.43	0.03	G21C1IV	0.65	0.08	0.68	0.03
G21C2II	0.04	0.09	0.27	0.03	G21C2III	0.06	0.09	0.29	0.03	G21C2IV	0.25	0.07	0.84	0.03
G22C1II	0.13	0.00	0.11	0.00	G22C1III	0.28	0.00	0.26	0.00	G22C1IV	0.23	0.00	0.22	0.00
G22C2II	0.04	0.00	0.04	0.00	G22C2III	0.05	0.00	0.04	0.00	G22C2IV	0.18	0.01	0.20	0.01
G23C1II	0.57	0.05	0.45	0.01	G23C1III	1.18	0.05	1.31	0.01	G23C1IV	1.81	0.03	1.57	0.01
G23C2II	0.09	0.09	0.28	0.02	G23C2III	0.09	0.09	0.21	0.02	G23C2IV	0.42	0.05	1.22	0.01
Avg.	**0.15**	**0.07**	**0.37**	**0.02**	**Avg.**	**0.24**	**0.06**	**0.38**	**0.02**	**Avg.**	**0.40**	**0.05**	**0.66**	**0.02**