Topics in cooperative game theory and logistics

Ondrej Osicka

August 2020

Acknowledgements

I feel fortunate for being able to work on this thesis at the Department of Business and Management Science. The last four years in this incredibly friendly and supportive environment were amazing. Moreover, some of my colleagues have even become some of my closest friends.

If PhD is my first step into an academic career, I am proud to have a role model like my supervisor Mario Guajardo. His supervision has been nothing short of exemplary. Without his guidance, this thesis as well as the whole journey towards a PhD would have been much different. Even when running took away my focus and a global pandemic took away my productivity, Mario was always ready to provide the needed encouragement and for that I will be forever grateful. Besides Mario, I am also thankful to Kurt Jörnsten and Thibault van Oost which co-authored some of the articles included in this thesis. Unfortunately, none of the articles has been written with my co-advisor Stein W. Wallace, but to him I am grateful for training in stochastic programming utilized in the last chapter. Besides, I believe that our co-authored article is yet to come.

Lastly, I would like to express my great appreciation to my friends and family. Even though some of you live in a different country and we do not see each other as often as we used to and even though to my little niece I am just the uncle from the computer screen, I feel your continued support and I am immensely grateful for that.

Contents

Introduction

Logistics plays a crucial role in the global economy. In the U.S., over the last 10 years, business logistics costs accounted for over 7 % of the GDP (Kearney, 2020). In addition, the volume of freight movement is projected to grow by 40 % between 2015 and 2045 in the U.S. (Federal Highway Administration, 2017). In the EU, a growth of 1 to 4 % per year is expected (Ferrell et al., 2020).

The growth of the logistics sector has led to increased competition. This has been accompanied by introduction of products with shorter life cycles, rising labour prices, growing body of regulations, customers reducing order sizes and expecting shorter delivery times, etc. Thus, for logistics companies, operating in an economically efficient manner is becoming more and more challenging (Archetti et al., 2009, Cruijssen et al., 2007). Consequences of this inefficiency include among others, as presented by Ferrell et al. (2020), U.S. trucks' trailers being on average only 43 % full and 25 % of total travel being done with completely or nearly empty trailers. In the EU, the empty truck miles are estimated to range from 15 to 20 %.

To improve efficiency in supply chains, companies can join efforts and coordinate their activities. This can be referred to as collaborative logistics. Synergy effects associated with the cooperation often result in reduction in costs and increase in efficiency (Cruijssen et al., 2007). Moreover, collaborative logistics has also been identified as an opportunity to increase service levels, gain market shares, enhance capacities and reduce the negative impacts of the bullwhip effect (Audy et al., 2012).

In 2016, the transportation sector was responsible for 24.3 % of the total greenhouse gas emissions in the EU (European Union, 2018). By increasing efficiency in supply chains, cooperation may bring not only economical benefits, but can also contribute to reduction of the environmental impact. Fossil fuel combustion has significant impact on the environment. Thus, for instance, reduction of the empty miles would lead to decrease in the associated CO_2 emissions. Guajardo (2018) presented numerous case studies and reported substantial decrease in CO_2 emissions achieved by collaborative logistics.

Despite positive effects on several fronts, practical implementation of collaborative logistics is not always easy. Many collaboration efforts fail to meet the participants' expectations (Cao et al., 2010). According to Sabath and Fontanella (2002), collaboration has "the most disappointing track of the various supply chain management strategies" when it comes to its practical implementation. Audy et al. (2012) and Basso et al. (2019) described and classified common obstacles. Trust, fairness, conflicts of interest and choice of the right partners are just some of them. Nevertheless, this does not mean that there is no evidence of successful cooperation. Björnfot and Torjussen (2012) reported savings resulting from cooperation in Swedish timber industry. Cases of shared consolidation centers and cooperatively planned routing were described for example by Eyers (2010) and Paddeu (2017).

When companies begin to cooperate, exploiting the synergies and finding a new business strategy may often seem fairly straightforward. When it comes to inventory management, transportation planning or any other business aspect, companies usually have proper tools available before any potential cooperation occurs. Hence, with cooperation, the methodology of finding a new solution usually does not differ much as opposed to finding a solution prior to the cooperation. For example, with resources shared among partners, data can be aggregated and the problem to solve becomes larger, but the methodology remains the same. However, for a successful practical implementation, this may not suffice and more measures might need to be accounted for.

Ensuring that incentives to cooperate exist for all companies has been recognized as one of the main barriers for implementation of collaborative logistics (Basso et al., 2019). Cooperative game theory provides tools to recognize whether such a solution exists and, if it does, methods to find it. Additionally, the cooperative game-theoretic framework may help for example in answering questions revolving around fairness in collaborative efforts as well as with identification of an optimal partner or a coalition to join efforts with (whether it is to maximize economical benefits or likelihood of successful practical implementation).

This thesis is organized into four self-contained chapters. All of the chapters were written with an intention to publish in a scientific journal. Therefore, throughout the thesis, the terms chapter, article and paper are used interchangeably. With my co-authors, we address some of the questions revolving around cooperation within a scope of specific problems from transportation and logistics. In particular, chapter 1 focuses on optimal choice of strategies in a cooperative version of the traveling salesman problem with profits. Chapter 2 characterizes the cooperative variant of the location-routing problem in terms of various features such as the existence of core allocations, where players have no incentives to leave the collaboration. Chapter 3 focuses on finding a fairness-maximizing solution among a finite set of options in the traveling tournament problem. Lastly, chapter 4 introduces a methodology to determine the optimal coalition to join when the outcomes are uncertain. Besides cooperative game theory, techniques from various subfields of mathematical optimization are employed throughout the chapters, including integer, stochastic as well as multi-objective optimization. A more detailed description of each chapter follows.

Chapter 1. Cooperation of customers in traveling salesman problems with profits

with Mario Guajardo and Kurt Jörnsten

As opposed to the traveling salesman problem, in the traveling salesman problems with profits, not all customers need to be visited. The customers offer prizes and the carrier chooses which ones to visit based on the prize acquired upon their visit. The variant of the problem with an objective to maximize the difference between the total collected profit and the total traveling cost is called the profitable tour problem.

In this chapter, we focus on the prizes customers need to offer to ensure being visited by the carrier. This can be formulated as a cooperative game where customers may form coalitions and make decisions on the prize values cooperatively. We define the profitable tour game describing this situation and analyze the cost associated with each coalition of customers and prizes that help in achieving it. We derive properties of the optimal prizes to be offered when the grand coalition is formed. These properties describe relationship between the prizes and the underlying traveling salesman game to provide connection with extensive literature on core allocations in traveling salesman games. Our most compelling result states that the set of optimal prizes coincides with the core of the underlying traveling salesman game whenever this core is nonempty.

Chapter 2. Cooperative game-theoretic features of cost sharing in location-routing

with Mario Guajardo and Thibault van Oost

The location-routing problem deals with a question of locating facilities while simultaneously finding routes to serve customers from these facilities. The vast number of variants and extensions of this problem in the literature demonstrates its importance in logistics. Eyers (2010) and Paddeu (2017) report cases of companies sharing consolidation centers while simultaneously cooperating on transportation of their products to customers. With these sharing practices emerging as important mechanisms to improve operations, it is increasingly important for companies to understand the benefits and economic implications of cooperation in location-routing.

In this chapter, we formulate several variants of the collaborative version of the location-routing problem and classify them within a cooperative game-theoretic framework. We derive characteristics in terms of subadditivity, convexity, and non-emptiness of the core. Moreover, for some of the game variants, we show that for facility opening costs substantially larger than the costs associated with routing, the core is always non-empty. The theoretical results are supported by numerical experiments aimed at illustrating the properties and deriving insights. Among others, we observe that, while in general it is not possible to guarantee core allocations, in a huge majority of cases the core is non-empty.

Chapter 3. Fair travel distances in tournament schedules: A cooperative game theory approach

with Mario Guajardo

Generating fair schedules is an important aspect in the organization of sports competitions. The vast majority of the sports scheduling literature has focused on optimization problems where the tournament schedule is obtained by a minimization or a maximization of a single criterion. For example, the traveling tournament problem, the most studied problem in the literature, aims at finding a schedule that minimizes the total distance traveled by the teams. While minimizing the expenditure resulting from all traveling between games is efficient from the overall objective perspective, it overlooks the actual distribution of the travel among the teams. In consequence, some teams may end up better than others with respect to their single goals, an imbalance which may largely affect teams' often limited resources as well as preparedness for the games.

In this chapter, we adopt a cooperative game theory framework to deal with the question of fairness in sports scheduling. To obtain fair tournament schedules with respect to the travel distances of the teams, we develop the following approach. First, the scheduling problem is reformulated as a transferable utility game. Second, by means of well-established cost allocation methods, such as the egalitarian method, Shapley value and nucleolus, an ideal distance distribution among the teams is determined. Third, given the inherently discrete nature of the space of feasible solutions to the scheduling problem, we introduce fairness measures to produce a schedule which approximately resembles the ideal distribution. We also discuss how to obtain a solution in case of not pursuing only fairness, but rather a compromise between fairness and minimum total distance. To illustrate the approach, we compute numerical results in one of the classic data instances of the TTP.

Chapter 4. Player-centered approach to coalition formation in transferable utility games with uncertain payoffs

Partner selection has been identified by Basso et al. (2019) as one of the critical factors for successful practical implementation of collaboration. Despite the fact that transferable utility games (TU games) allow for reallocation of the worth of a coalition among its members, only a small number of studies has considered endogenously formed coalitions where the final allocation might actually affect the decision about which coalitions to establish. Moreover, to the best of my knowledge, there have been no studies considering a decentralized approach to endogenous coalition formation, i.e., finding coalitions optimal to form from a perspective of the players while taking into account the subsequent allocation.

In this chapter, I investigate endogenous coalition formation in TU games from a perspective of their players. In particular, the focus is on decision-making situations where coalitions need to be formed before their actual outcome is observable. Thus, the scope of this chapter is much broader than problems from transportation and logistics. Several approaches are formulated to determine which coalition is optimal for a given player to pursue while taking into account the subsequent payoff or cost allocation. The formulated models are divided into two main categories, those describing TU games where the subsequent allocation rules are known prior to the coalition formation and those describing TU games where negotiations within the formed coalitions are yet to take place after observing the uncertainty realization. Thus, in addition to a novel approach to the coalition formation, the models also take into account possible uncertainty in the TU games' properties and hence in their characteristic function values. The models are then addressed with a stochastic programming approach. Subsequently, the methodology is illustrated on an example of randomly winning coalitions and on an example of a collaborative transportation problem. The results support arguments against exogenous approaches to coalition formation and show that failing to take the uncertainty in parameter values into account might lead to suboptimal solutions and consequently to false conclusions.

References

Archetti, C., Feillet, D., Hertz, A., and Speranza, M. G. (2009). The capacitated team orienteering and profitable tour problems. *Journal of the Operational Research Society*, 60(6):831–842.

Audy, J. F., Lehoux, N., D'Amours, S., and Rönnqvist, M. (2012). A framework for an efficient implementation of logistics collaborations. *International Transactions in Operational Research*, 19(5):633–657.

Basso, F., D'Amours, S., Rönnqvist, M., and Weintraub, A. (2019). A survey on obstacles and difficulties of practical implementation of horizontal collaboration in logistics. *International Transactions in Operational Research*, 26(3):775–793.

Björnfot, A. and Torjussen, L. (2012). Extent and effect of horizontal supply chain collaboration among construction SME. *Journal of Engineering, Project, and Production Management*, 2(1):47–55.

Cao, M., Vonderembse, M. A., Zhang, Q., and Ragu-Nathan, T. S. (2010). Supply chain collaboration: conceptualisation and instrument development. *International Journal of Production Research*, 48(22):6613–6635.

Cruijssen, F., Cools, M., and Dullaert, W. (2007). Horizontal cooperation in logistics: Opportunities and impediments. *Transportation Research Part E: Logistics and Transportation Review*, 43(2):129–142.

European Union (2018). EU transport in figures: Statistical pocketbook 2018. Technical report.

Eyers, N. (2010). 'Pooling': a successful collaborative approach. Presented at *Horizontal Collaboration in the Supply Chain Summit*, Brussels, June 2010.

Federal Highway Administration (2017). Freight analysis framework. Technical report.

Ferrell, W., Ellis, K., Kaminsky, P., and Rainwater, C. (2020). Horizontal collaboration: opportunities for improved logistics planning. *International Journal of Production Research*, 58(14):4267–4284.

Guajardo, M. (2018). Environmental benefits of collaboration and allocation of emissions in road freight transportation. In Zeimpekis, V., Aktas, E., Bourlakis, M., and Minis, I., editors, *Sustainable Freight Transport*, pages 79–98. Springer, Cham.

Kearney (2020). CSCMP's 31st Annual State of Logistics Report. Technical report.

Paddeu, D. (2017). The Bristol-Bath urban freight consolidation centre from the perspective of its users. *Case Studies on Transport Policy*, 5(3):483–491.

Sabath, R. E. and Fontanella, J. (2002). The unfulfilled promise of supply chain collaboration. *Supply Chain Management Review*, 6(4):24–29.

Chapter 1

Cooperation of customers in traveling salesman problems with profits*

Ondrej Osicka[a], Mario Guajardo[a], Kurt Jörnsten[a]

[a]Department of Business and Management Science, NHH Norwegian School of Economics, Bergen, Norway

Abstract

In the profitable tour problem, the carrier decides whether to visit a particular customer with respect to the prize the customer offers for being visited and traveling cost associated with the visit, all in the context of other customers. Our focus is on the prizes customers need to offer to ensure being visited by the carrier. This can be formulated as a cooperative game where customers may form coalitions and make decisions on the prize values cooperatively. We define the profitable tour game describing this situation and analyze the cost associated with each coalition of customers and prizes that help to achieve it. We derive properties of the optimal prizes to be offered when the grand coalition is formed. These properties describe relationship between the prizes and the underlying traveling salesman game to provide connection with extensive literature on core allocations in traveling salesman games. The most important result states that the set of optimal prizes coincides with the core of the underlying traveling salesman game if this core is nonempty.

Keywords: Traveling salesman problem; Profitable tour problem; Prize-collecting TSP; Logistics; Cooperative game theory; Prize allocation

*A published version of this chapter appears in *Optimization Letters*, 14, 1219-1233.

1.1 Introduction

The traveling salesman problem (TSP) is one of the most studied problems in logistics (Laporte, 2007). It answers the question of how to visit several places within a single tour starting and finishing at a particular place while minimizing the total traveling cost. Throughout this paper, we will use words *customers* to refer to the places to be visited and *carrier* to refer to their visitor. There have been a lot of variants of the traveling salesman problem dealing with different aspects of the underlying situation. For example, the customers might offer prizes for being visited which introduces the traveling salesman problems with profits (Feillet et al., 2005). This offers alternations of the carrier's objective in terms of focusing on minimizing the total traveling cost, maximizing the total collected profit or any combination of these conflicting objectives. In contrast to TSP, in the traveling salesman problems with profits not all customers need to be visited and the carrier selects them based on a prize acquired when they are visited. The variant with an objective to maximize the difference between the total collected profit and the total traveling cost, is called the profitable tour problem (PTP) as described by Laporte and Martín (2007).

A cooperative formulation of the traveling salesman problem, the traveling salesman game, attracted scientific interest after a question proposed by Fishburn and Pollak (1983). After a road trip of a visitor visiting several sponsors, how should they be charged in a fair manner such that they cover the total cost of the trip? To find such allocation, Fishburn and Pollak (1983) stated several conditions on the allocated costs, then Potters et al. (1992) provided game-theoretic insights. Nowadays, there exist many studies focusing on such allocations (Engevall et al., 1998, Kimms and Kozeletskyi, 2016, Sun et al., 2015, Sun and Karwan, 2015, Tamir, 1989). In fact, traveling salesman games are the main topic of a great share of the atricles gathered by a recent survey on cost allocation methods used in collaborative transportation (Guajardo and Rönnqvist, 2016).

Estévez-Fernández et al. (2009) proposed a traveling salesman game alternative for the case of customers offering prizes, called the routing game with revenues. Its focus remains on the allocation of the total cost of a tour visiting all customers. However, the total cost incurred by the customers is in fact the sum of all offered prizes. As shown by the following example, sometimes it is needed to allocate more than the total traveling cost of the tour.

Example 1. Let $1, \ldots, 6$ be customers which desire to be visited by a carrier departing from and returning to depot 0 as illustrated in Figure 1.1, an example taken from Tamir (1989). All

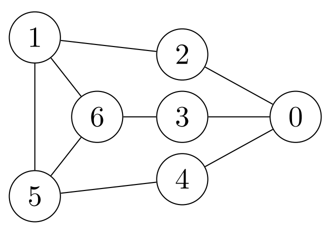


Figure 1.1: Problem in Example 1

edges of the graph represent unit traveling cost such that the minimal cost of traveling from 0 to 4 is of one unit (0-4) whereas the minimal cost of traveling from 2 to 3 is of two units (2-0-3). It is easy to see that the least costly tour would be of 8 units (0-4-5-6-3-6-1-2-0). Looking at the problem from a perspective of a prize-collecting carrier, let us assume the carrier is originally visiting all of the customers which implies the aforementioned cost of 8 units. If the carrier decides to drop customers 4 and 5 and only performs tour 0-3-6-1-2-0, the associated cost drops to 5 units. This means that customers 4 and 5 can make themselves worth visiting by offering a total prize of at least 3 units to cover the additional traveling cost associated with their visit. The same requirement of at least 3 units being offered could be derived for the pair of customers 3 and 6 and the pair of customers 1 and 2. Together the prizes of all customers must add up to at least 9 units. Otherwise, some customers do not get visited. In other words, at a cost of 8 units, the coalition of all customers fails to fulfill its purpose and both the traveling salesman game and the routing game with revenues do not describe such a situation properly as they assume all customers being visited at this cost.

The purpose of this paper is to define the profitable tour game, a cooperative version of the profitable tour problem, and to derive its properties. We are particularly interested in prize allocations that create incentives for the carrier to visit all relevant customers.

The traveling salesman game could be interpreted as a problem where the carrier is forced to visit all customers and suggest charges in a way that makes the customers willing to be visited. On the other hand, the profitable tour game allows the carrier to have a final word and introduces a problem where the customers need to compete or cooperate and make themselves worth being visited. The profitable tour game is not only of theoretical interest as it relates to many situations occurring in practice.

The most straightforward application is a cooperation of customers that can be served within a single route of a carrier. This might include for example less-than-truckload shipping where a carrier is able to serve demands of several customers with a single vehicle (Archetti et al., 2014). Whether it comes to delivery of goods or pickup of goods (such as waste collection (Šomplák et al., 2017)), the customers might need to induce the carrier to visit them by offering sufficient rewards. Subsequently, negotiation with other customers in the same position could lead to better prizes while the carrier's visit would remain guaranteed. This knowledge could also be utilized by the carrier by offering specifically tailored discounts on multiple orders from the same area.

The profitable tour game might also become relevant in evaluating and pricing of new customers. In Engevall et al. (1998), for example, the Logistics Department at Norsk Hydro Olje AB determines how gas stations across southern Sweden should be charged for distribution of gas utilizing the concept of the traveling salesman game. If another gas station would like to join the initial set of stations, a simple question of what charges should the gas station expect for being even considered interesting already requires a point of view as given by the profitable tour game.

The remainder of this paper is organized as follows. In section 1.2, we analyze the costs associated with different coalitions of customers and define the profitable tour game. Section 1.3 studies the optimal prizes to be offered by customers. Concluding remarks follow in section

1.4.

1.2 From the traveling salesman game to the profitable tour game

The problem outlined in the introduction is a two-stage conflict of $n+1$ decision-makers. First, n customers decide on prizes offered to a carrier for being visited and, after all prizes are observed, the carrier decides which customers to visit and how to perform the tour. We assume the carrier to be rational and profit-maximizing, which means, the strategy is to choose the tour with the largest difference between the total prize-based profit and the total traveling cost. With this in mind, the customers want to set the prizes in a least costly manner that still guarantees them to be visited. It might be useful to form coalitions with other customers. Such coalitions then aim to set prizes offered by their customers such that their sum is the least possible to guarantee being visited regardless of the other customers' offered prizes.

The whole situation is a zero-sum game, that is, whatever is paid out by the customers gets collected by the carrier. This does not offer opportunities for a meaningful cooperation of all players. On the contrary, leaving out the carrier and focusing on the conflict of customers only, the prizes can be set in a way that benefits other customers as well. This is the idea for defining the profitable tour game as a cooperative game of the customers.

For modeling purposes, we impose standard assumptions on the traveling costs c_{ij} among customers themselves and between them and the carrier's home depot. Denoting the set of all customers by N and the home depot by 0, these assumptions are

$$c_{ii} = 0 \qquad \forall i \in N \cup \{0\}, \tag{1.1}$$

$$c_{ij} \leq c_{ik} + c_{kj} \qquad \forall i, j, k \in N \cup \{0\}. \tag{1.2}$$

Assumptions (1.1) imply no traveling cost is associated with staying in the same place and the triangle inequalities (1.2) make sure that the costs always represent the lowest possible costs which cannot be beaten by going another way. These assumptions are common in the literature. There are studies of traveling salesman problems with symmetric as well as asymmetric traveling costs. We don't limit our focus by imposing assumptions on this symmetry.

1.2.1 The traveling salesman game

To define the profitable tour game and derive its properties, it comes in handy to utilize the definition of the traveling salesman game by Potters et al. (1992) which will become our starting point.

Let $N = \{1, \ldots, n\}$ be the set of all customers. For each group of customers $S \subseteq N$ which need to be visited from depot 0 using only one vehicle starting from and finishing at the depot, the least total traveling cost could be obtained by solving the *traveling salesman problem* (TSP) given by

$$\text{Cost}^{TSP}(S) = \min \sum_{i \in S \cup \{0\}} \sum_{j \in S \cup \{0\}} c_{ij} x_{ij} \tag{1.3}$$

$$\text{s.t.} \quad \sum_{i \in S \cup \{0\}} x_{ij} = 1 \qquad \forall j \in S \cup \{0\}, \tag{1.4}$$

$$\sum_{j \in S \cup \{0\}} x_{ij} = 1 \qquad \forall i \in S \cup \{0\}, \tag{1.5}$$

$$\sum_{i \in T} \sum_{j \in T} x_{ij} \leq |T| - 1 \qquad \forall T \subset S \cup \{0\} \colon T \neq \varnothing, \tag{1.6}$$

$$x_{ij} \in \{0,1\} \qquad \forall i, j \in S \cup \{0\}. \tag{1.7}$$

In this integer linear programming model, x_{ij} is a binary variable that indicates whether the vehicle travels directly from i to j in the final tour. Constraints (1.4) and (1.5) ensure that all customers in S and the depot are visited exactly once and constraints (1.6) eliminate any subtours to guarantee the solution to be a single tour. The binary nature of x_{ij} is prescribed by (1.7). A tour with the lowest traveling cost is then selected by (1.3).

The pair (N, Cost^{TSP}) is then the traveling salesman game as defined by Potters et al. (1992).

1.2.2 Cooperation in the profitable tour problem

The introduction of prizes offered by customers to the carrier for visiting them, denoted by Prize_j for the prize offered by customer $j \in N$, requires a slightly different view of the problem. Finding the optimal profit of the carrier is then the *profitable tour problem* (PTP) which can be formulated as

$$\text{Profit}^{PTP} = \max \left(\sum_{i \in N \cup \{0\}} \sum_{j \in N} \text{Prize}_j x_{ij} - \sum_{i \in N \cup \{0\}} \sum_{j \in N \cup \{0\}} c_{ij} x_{ij} \right) \tag{1.8}$$

$$\text{s.t.} \quad \sum_{i \in N \cup \{0\}} \sum_{j \in N \cup \{0\}} x_{ij} \leq M \sum_{j \in N} x_{0j}, \tag{1.9}$$

$$\sum_{i \in N \cup \{0\}} x_{ij} = \sum_{k \in N \cup \{0\}} x_{jk} \qquad \forall j \in N \cup \{0\}, \tag{1.10}$$

$$\sum_{i \in T} \sum_{j \in T} x_{ij} \leq |T| - 1 + \delta_T M \qquad \forall T \subset N \cup \{0\} \colon T \neq \varnothing, \tag{1.11}$$

$$\sum_{i \in (N \cup \{0\}) \setminus T} \sum_{j \in N \cup \{0\}} (x_{ij} + x_{ji}) \leq (1 - \delta_T) M \qquad \forall T \subset N \cup \{0\} \colon T \neq \varnothing, \tag{1.12}$$

$$x_{ij} \in \{0,1\} \qquad \forall i, j \in N \cup \{0\}, \tag{1.13}$$

$$\delta_T \in \{0,1\} \qquad \forall T \subset N \cup \{0\} \colon T \neq \varnothing, \tag{1.14}$$

where x_{ij} again indicates whether the vehicle travels directly from i to j, δ_T is a binary variable used for modeling subtour elimination, and M is a big enough number (for example, $M = |N \cup \{0\}|^2$ would be sufficient).

If there are any customers to be visited, constraint (1.9) ensures that the tour starts from the depot. Constraints (1.10) ensure that if the vehicle arrives at a certain customer it needs to continue its tour afterwards. For the depot, the intuition is that the tour needs to both start and finish there. Constraints (1.11) and (1.12) eliminate subtours by ensuring that, if there is a cycle over a set T, all customers not belonging to T remain unvisited. Constraints (1.13) and (1.14) state the binary nature of variables x_{ij} and δ_T. Overall, a tour maximizing the difference between the total prize-based profit and the total traveling cost is chosen by (1.8).

In the case of PTP, the carrier is not forced to visit all customers, but visits only the most profitable subset of them. On the other hand, if all customers in coalition $S \subseteq N$ (and only those) needed to be visited, new constraints could be introduced in the PTP model which would create what we refer to as the *profitable tour problem with all customers visited* (PTP-AV). This can be formulated generally for any group of customers S as

$$\text{Profit}^{AV}(S) = \max \left(\sum_{i \in S \cup \{0\}} \sum_{j \in S} \text{Prize}_j x_{ij} - \sum_{i \in S \cup \{0\}} \sum_{j \in S \cup \{0\}} c_{ij} x_{ij} \right) \tag{1.15}$$

$$\text{s.t.} \quad \sum_{i \in S \cup \{0\}} x_{ij} = \sum_{k \in S \cup \{0\}} x_{jk} \qquad \forall j \in S \cup \{0\}, \tag{1.16}$$

$$\sum_{i \in S \cup \{0\}} x_{ij} = 1 \qquad \forall j \in S \cup \{0\}, \tag{1.17}$$

$$\sum_{i \in S \cup \{0\}} \sum_{j \in S \cup \{0\}} x_{ij} \leq M \sum_{j \in S} x_{0j}, \tag{1.18}$$

$$\sum_{i \in T} \sum_{j \in T} x_{ij} \leq |T| - 1 + \delta_T M \qquad \forall T \subset S \cup \{0\} \colon T \neq \varnothing, \tag{1.19}$$

$$\sum_{i \in (S \cup \{0\}) \setminus T} \sum_{j \in S \cup \{0\}} (x_{ij} + x_{ji}) \leq (1 - \delta_T) M \qquad \forall T \subset S \cup \{0\} \colon T \neq \varnothing, \tag{1.20}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in S \cup \{0\}, \tag{1.21}$$

$$\delta_T \in \{0,1\} \qquad \forall T \subset S \cup \{0\} \colon T \neq \varnothing. \tag{1.22}$$

The model for $\text{Profit}^{AV}(N)$ indeed differs from PTP only by the inclusion of constraints (1.17), which ensure that all customers from S are visited (customers outside S are not visited as they are in fact not even part of the model). It is easy to see that constraint (1.18) is no more needed and the left-hand side of constraints (1.20) is never less than 2 and hence δ_T equals 0 for all nonempty $T \subset S \cup \{0\}$. The PTP-AV model could be then reformulated as

$$\text{Profit}^{AV}(S) = \max \left(\sum_{j \in S} \text{Prize}_j - \sum_{i \in S \cup \{0\}} \sum_{j \in S \cup \{0\}} c_{ij} x_{ij} \right) \tag{1.23}$$

$$\text{s.t.} \quad \sum_{j \in S \cup \{0\}} x_{ij} = 1 \qquad \forall i \in S \cup \{0\}, \tag{1.24}$$

$$\sum_{i \in S \cup \{0\}} x_{ij} = 1 \qquad \forall j \in S \cup \{0\}, \tag{1.25}$$

$$\sum_{i \in T} \sum_{j \in T} x_{ij} \leq |T| - 1 \qquad \forall T \subset S \cup \{0\} \colon T \neq \varnothing, \tag{1.26}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in S \cup \{0\}. \tag{1.27}$$

One of two additional changes done is the replacement of constraints (1.16) by constraints (1.24) as the combination of constraints (1.16)-(1.17) is obviously equivalent to the combination of constraints (1.24)-(1.25). The second change is the change in the first term of objective function (1.15) which indeed holds because of constraints (1.17). However, this term then becomes constant over the decision variables and hence, for each S, the optimal solution is the same as for the case of TSP and the objective value is

$$\text{Profit}^{AV}(S) = \sum_{j \in S} \text{Prize}_j - \text{Cost}^{TSP}(S) \tag{1.28}$$

for any S.

Under what conditions does PTP generate an optimal solution with all customers in a particular coalition S being visited? Clearly, there needs to exist set $T \subseteq N \setminus S$ such that

$$\text{Profit}^{PTP} = \text{Profit}^{AV}(S \cup T). \tag{1.29}$$

A relationship between the objective values of PTP and PTP-AV could be expressed as

$$\text{Profit}^{PTP} = \max_{R \subseteq N} \text{Profit}^{AV}(R) \tag{1.30}$$

and hence

$$\text{Profit}^{AV}(S \cup T) \geq \text{Profit}^{AV}(R) \qquad \forall R \subseteq N. \tag{1.31}$$

Thinking about how coalition S could achieve it by setting prizes offered by customers in S (and not by the others), it needs to be noted that any customer outside S could make themselves interesting for the carrier by setting the prize exceptionally high or uninteresting or at least indifferent by setting it to zero. Therefore, as coalition S has no control over the other prizes, instead of (1.31), S needs to set the prizes such that, for any $T \subseteq N \setminus S$, it is profitable for the carrier to visit all customers in S, that is

$$\text{Profit}^{AV}(S \cup T) \geq \text{Profit}^{AV}(R \cup T) \qquad \forall R \subseteq S,\ \forall T \subseteq N \setminus S \tag{1.32}$$

or, using relation (1.28),

$$\sum_{j \in S \cup T} \text{Prize}_j - \text{Cost}^{TSP}(S \cup T) \geq \sum_{j \in R \cup T} \text{Prize}_j - \text{Cost}^{TSP}(R \cup T) \qquad \forall R \subseteq S,\ \forall T \subseteq N \setminus S \tag{1.33}$$

which can be simplified as

$$\sum_{j \in S \setminus R} \text{Prize}_j \geq \text{Cost}^{TSP}(S \cup T) - \text{Cost}^{TSP}(R \cup T) \qquad \forall R \subseteq S,\ \forall T \subseteq N \setminus S. \tag{1.34}$$

1.2.3 The profitable tour game

As discovered in the previous subsection, the carrier would visit all customers in coalition S only if the prizes were offered in a way satisfying conditions (1.34). It is then easy to determine the minimal total cost associated with S as

$$\mathrm{Cost}^{PTP}(S) = \min \sum_{j \in S} \mathrm{Prize}_j \tag{1.35}$$

$$\text{s.t.} \sum_{j \in S \setminus R} \mathrm{Prize}_j \geq \mathrm{Cost}^{TSP}(S \cup T) - \mathrm{Cost}^{TSP}(R \cup T) \quad \forall R \subseteq S, \forall T \subseteq N \setminus S, \tag{1.36}$$

$$\mathrm{Prize}_j \geq 0 \quad \forall j \in S. \tag{1.37}$$

The pair (N, Cost^{PTP}) then defines a cooperative transferable-utility game of the customers which we name the *profitable tour game*.

Example 2. Looking back at Example 1, it is easy to compute the Cost^{PTP} values and compare them to those of Cost^{TSP}. These functions are defined for 64 different coalitions, but they differ for 10 of them only. These are reported in Table 1.1. Taking the first one, $S = \{1, 2, 3, 6\}$, as an example, suppose that customers 4 and 5 offer sufficiently high prizes such that the carrier would always visit them. Then, by the same logic as in Example 1, each pair of customers 1 and 2, and 3 and 6 would need to offer a total prize of at least 3 units, that adds up to at least 6 units in total. Hence, offering only 5 units does not guarantee all customers in S being visited. However, if they were to offer 6 units for instance in a way that customers 1 and 6 offer 1 unit each and 2 and 3 offer 2 units each, it would leave the carrier indifferent between visiting and not visiting all of them. Even a very small increase in the prizes would then create strong preference for visiting them. 6 units is hence indeed the minimal cost guaranteeing all customers in S being visited.

Table 1.1: Differences in values of cost functions Cost^{TSP} and Cost^{PTP} in Example 2

S	Cost^{TSP}	Cost^{PTP}	S	Cost^{TSP}	Cost^{PTP}
$\{1,2,3,6\}$	5	6	$\{1,2,3,5,6\}$	6	7
$\{1,2,4,5\}$	5	6	$\{1,2,4,5,6\}$	6	7
$\{3,4,5,6\}$	5	6	$\{1,3,4,5,6\}$	6	7
$\{1,2,3,4,5\}$	7	8	$\{2,3,4,5,6\}$	7	8
$\{1,2,3,4,6\}$	7	8	$\{1,2,3,4,5,6\}$	8	9

1.3 Properties of the profitable tour game

The definition of the profitable tour game provides information on costs associated with different coalitions, but prizes leading to such outcomes also deserve attention. At this point, a clear distinction between prize allocation and cost allocation needs to be made. Using game-theoretic terminology, prize allocation represents the strategies of the customers, whereas the cost allocation represents the outcome of the cooperation assigned to the customers.

When dealing with cost allocation, most studies utilize concept of the *core*. For a game (N, Cost) with $N = \{1, \ldots, n\}$, the core is defined as a set of all allocations $(x_1, \ldots, x_n)$, where x_i is the cost prescribed to be paid by customer $i \in N$, that satisfy constraints

$$\sum_{i \in N} x_i = \text{Cost}(N), \tag{1.38}$$

$$\sum_{i \in S} x_i \leq \text{Cost}(S) \qquad \forall S \subseteq N. \tag{1.39}$$

Constraint (1.38) guarantees that the total cost is allocated and constraints (1.39) ensure that no coalition can get better off by breaking the cooperation. It is important to note that for some games such allocations need not exist and then the core is empty.

1.3.1 Prize allocation

Many studies of traveling salesman games deal with conditions for nonemptiness of the core (Sun and Karwan, 2015, Tamir, 1989). It is then natural to study how the optimal prize allocations are affected by the emptiness of the core of the underlying traveling salesman game. In what follows, when referring to an *optimal prize allocation of the grand coalition*, we mean the prizes Prize_j that represent the optimal solution of model (1.35)-(1.37) when solved for $\text{Cost}^{PTP}(N)$.

Theorem 1. *If the core of the traveling salesman game* (N, Cost^{TSP}) *is nonempty, all allocations from the core, and no other, are optimal prize allocations of the grand coalition in the profitable tour game* (N, Cost^{PTP})*.*

Proof. For N, model (1.35)-(1.37) becomes

$$\text{Cost}^{PTP}(N) = \min \sum_{j \in N} \text{Prize}_j \tag{1.40}$$

$$\text{s.t.} \quad \sum_{j \in N \setminus R} \text{Prize}_j \geq \text{Cost}^{TSP}(N) - \text{Cost}^{TSP}(R) \qquad \forall R \subseteq N, \tag{1.41}$$

$$\text{Prize}_j \geq 0 \qquad \forall j \in N. \tag{1.42}$$

With optimal prizes, constraints (1.41) need to hold for all $R \subseteq N$ including $R = \varnothing$, that is

$$\sum_{j \in N} \text{Prize}_j \geq \text{Cost}^{TSP}(N). \tag{1.43}$$

The term on the left-hand side of this constraint is the same as in objective (1.40). Hence, the objective value must be greater than or equal to $\text{Cost}^{TSP}(N)$. The equality is achieved only when constraint (1.43) is binding, that is

$$\sum_{j \in N} \text{Prize}_j = \text{Cost}^{TSP}(N). \tag{1.44}$$

Assuming this to hold, constraints (1.41) can be rewritten as

$$\sum_{j \in R} \text{Prize}_j \leq \text{Cost}^{TSP}(R) \qquad \forall R \subseteq N. \tag{1.45}$$

If the core of the traveling salesman game (N, Cost^{TSP}) is nonempty, there exist prizes Prize_j that satisfy (1.44) and (1.45) and, hence, are optimal. Clearly, the set of all such prize allocations coincides with the core of (N, Cost^{TSP}). □

If the core contains more than just one allocation, to select a particular one, it might be useful to use allocation methods such as the nucleolus which by definition make as few constraints (1.45) binding as possible (Schmeidler, 1969, Guajardo and Jörnsten, 2015). This would contribute to lowering chances of leaving the carrier indifferent between visiting all and just some of the customers. A binding constraint (1.45) for a particular R leaves the carrier indifferent between visiting all customers and visiting only those in R. However, it still cannot rule out the indifference between visiting all customers and not performing a tour at all which is obvious from (1.44).

Example 3. Let 1, 2 and 3 be customers which desire to be visited by a carrier departing from and returning to depot 0 as illustrated in Figure 1.2. Each edge of the graph is accompanied

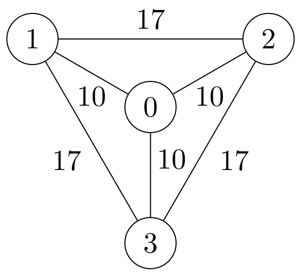


Figure 1.2: Problem in Example 3

by a number standing for the respective traveling cost. The traveling salesman problem can be solved to obtain costs for different coalitions. This results in $\text{Cost}^{TSP}(\varnothing) = 0$, $\text{Cost}^{TSP}(\{1\}) = \text{Cost}^{TSP}(\{2\}) = \text{Cost}^{TSP}(\{3\}) = 20$, $\text{Cost}^{TSP}(\{1,2\}) = \text{Cost}^{TSP}(\{1,3\}) = \text{Cost}^{TSP}(\{2,3\}) = 37$, $\text{Cost}^{TSP}(\{1,2,3\}) = 54$.

By Theorem 1, optimal prize allocation $(\text{Prize}_1, \text{Prize}_2, \text{Prize}_3)$ belongs to the core of game $(\{1,2,3\}, \text{Cost}^{TSP})$, that is

$$\text{Prize}_1 + \text{Prize}_2 + \text{Prize}_3 = 54, \tag{1.46}$$

$$\text{Prize}_1 + \text{Prize}_2 \leq 37, \tag{1.47}$$

$$\text{Prize}_1 + \text{Prize}_3 \leq 37, \tag{1.48}$$

$$\text{Prize}_2 + \text{Prize}_3 \leq 37, \tag{1.49}$$

$$\text{Prize}_1 \leq 20, \tag{1.50}$$

$$\text{Prize}_2 \leq 20, \tag{1.51}$$

$$\text{Prize}_3 \leq 20. \tag{1.52}$$

This is satisfied for example by $\text{Prize}_1 = 20$, $\text{Prize}_2 = 17$, $\text{Prize}_3 = 17$. However, one can observe that it makes the carrier indifferent between visiting all customers, visiting customers 1 and 2, visiting customers 1 and 3, visiting only customer 1, and visiting no customers. However, allocating the prizes as prescribed by the nucleolus, that is $\text{Prize}_1 = 18$, $\text{Prize}_2 = 18$, $\text{Prize}_3 = 18$, would leave the carrier indifferent only between visiting all customers and visiting no customers. Even a marginal increase in any of these prizes would then incentivize the carrier to perform the tour visiting all customers.[1]

Theorem 1 implies $\text{Cost}^{PTP}(N) = \text{Cost}^{TSP}(N)$. This was not the case for Examples 1 and 2 which necessarily means that the core of the underlying traveling salesman game is empty for this example. This is indeed shown by Tamir (1989).

Examples 1 and 2 might look strange as for many combinations of the traveling costs the triangle inequality becomes equality and the unit traveling costs are not represented in Figure 1.1 with edges of the same length. Different explanations of this might take place such as an underlying road network does not allow shorter ways between some customers and a short distance traveled on a bad quality road is as expensive as a long distance traveled on a good quality road. However, there are also examples of cases with Euclidean distance between all places which result in games with empty cores (Engevall et al., 1998).

Theorem 2. *If the core of the traveling salesman game* (N, Cost^{TSP}) *is empty, then prizes* Prize_j *stand for an optimal prize allocation of the grand coalition in the profitable tour game* (N, Cost^{PTP}) *if and only if they represent an optimal solution of the problem:*

$$\min \varepsilon \tag{1.53}$$

$$\text{s.t.} \quad \sum_{j \in N} \text{Prize}_j = \text{Cost}^{TSP}(N) + \varepsilon, \tag{1.54}$$

$$\sum_{j \in S} \text{Prize}_j \leq \text{Cost}^{TSP}(S) + \varepsilon \qquad \forall S \subset N, \tag{1.55}$$

$$\text{Prize}_j \geq 0 \qquad \forall j \in N, \tag{1.56}$$

$$\varepsilon \geq 0, \tag{1.57}$$

where the variable ε *stands for the cost to be allocated in form of prizes in addition to the total cost of a tour visiting all customers.*

Proof. Following the same path as in the proof of Theorem 1, it is clear that assumption (1.44) would not be correct in this case as there exist no prizes Prize_j that satisfy (1.44) and (1.45) when the core of the traveling salesman game (N, Cost^{TSP}) is empty.

[1] A question could be raised of which ones of the customers should increase the prize. This might open a long discussion since arguably every individual customer wants to minimize its own prize. However, in the cooperative game-theoretical framework we adopt, the increment could be already reflected in the cost function value $\text{Cost}^{TSP}(\{1, 2, 3\})$. Thus, the increment gets allocated in a manner that is fair according to the chosen allocation method (the nucleolus in this example).

Assuming instead

$$\sum_{j \in N} \text{Prize}_j = \text{Cost}^{TSP}(N) + \varepsilon \tag{1.58}$$

for an arbitrary value of ε, constraints (1.41) could be reformulated as

$$\sum_{j \in R} \text{Prize}_j \leq \text{Cost}^{TSP}(R) + \varepsilon \qquad \forall R \subseteq N. \tag{1.59}$$

Lastly, because the term on the left-hand side of constraint (1.58) is the same as in objective function (1.40) and $\text{Cost}^{TSP}(N)$ is constant over the decision variables, then, in terms of the optimal solution for prizes Prize_j, objective function (1.40) is equivalent to

$$\min \varepsilon. \tag{1.60}$$

Model (1.40)-(1.42) can then be reformulated as (1.53)-(1.57) while preserving the same optimal solution for prizes Prize_j. The optimal value of ε can be interpreted as the minimal cost that needs to be allocated in form of prizes in addition to the total cost of a tour visiting all customers. □

It is important to note that model (1.53)-(1.57) is always feasible. For example, prizes prescribed as $\text{Prize}_j = \frac{\text{Cost}^{TSP}(N)+\varepsilon}{|N|}$ for each $j \in N$ clearly satisfy constraint (1.54) and, since $\text{Cost}^{TSP}(S) \geq 0$ for any $S \subseteq N$, then $\varepsilon \geq |N| \, \text{Cost}^{TSP}(N)$ guarantees satisfaction of constraints (1.55)-(1.57) as well.

Whereas there might exist multiple optimal solutions for prizes Prize_j, optimal ε is clearly unique. Then, game $(N, \widehat{\text{Cost}}^{TSP})$ can be defined, where $\widehat{\text{Cost}}^{TSP}(S) = \text{Cost}^{TSP}(S) + \varepsilon$ for each $S \subseteq N$. Straightforwardly, all allocations from the core of $(N, \widehat{\text{Cost}}^{TSP})$, and no other, are optimal prize allocations of the grand coalition in profitable tour game (N, Cost^{PTP}). This allows for usage of allocation methods such as the nucleolus for problems with empty cores of the associated traveling salesman games with the same implications as in the case of nonempty cores.

Example 4. Solving model (1.53)-(1.57) for Examples 1 and 2 results in $\varepsilon = 1$. Thanks to Example 2, we actually already knew the value of ε because

$$\varepsilon = \text{Cost}^{PTP}(N) - \text{Cost}^{TSP}(N) = 1. \tag{1.61}$$

Defining $\widehat{\text{Cost}}^{TSP}$ such that $\widehat{\text{Cost}}^{TSP}(S) = \text{Cost}^{TSP}(S) + 1$ for each $S \subseteq N$ and analyzing the core of the respective game introduces a system of one equality and 62 inequalities describing the set of all optimal prize allocations. To choose just one of them, the nucleolus prescribes the prize allocation $\text{Prize}_1 = 1$, $\text{Prize}_2 = 2$, $\text{Prize}_3 = 2$, $\text{Prize}_4 = 2$, $\text{Prize}_5 = 1$, and $\text{Prize}_6 = 1$.

Note that Theorem 2 could be generalized to all profitable tour games regardless of the core emptiness of the respective traveling salesman game. In fact, if the core of the traveling salesman game (N, Cost^{TSP}) is nonempty, the optimal value of ε in problem (1.53)-(1.57) equals 0 and the optimal prizes Prize_j must satisfy (1.44) and (1.45) which define the core of (N, Cost^{TSP}) as in Theorem 1.

An interesting corollary of Theorems 1 and 2 appears when utilizing a concept of the *dual game* (Sudhölter, 1996, Tarashnina, 2011). For a game (N, Cost), the dual game is defined as a game (N, Cost^*) where

$$\text{Cost}^*(S) = \text{Cost}(N) - \text{Cost}(N \setminus S) \qquad \forall S \subseteq N. \tag{1.62}$$

The corollary can be stated without a proof as it follows directly from model (1.40)-(1.42).

Corollary 1. *Prizes* Prize_j *stand for an optimal prize allocation of the grand coalition in the profitable tour game* (N, Cost^{PTP}) *if and only if they represent an optimal solution of problem*

$$\min \sum_{j \in N} \text{Prize}_j \tag{1.63}$$

$$\text{s.t.} \quad \sum_{j \in R} \text{Prize}_j \geq \text{Cost}^{TSP*}(R) \qquad \forall R \subseteq N, \tag{1.64}$$

$$\text{Prize}_j \geq 0 \qquad \forall j \in N. \tag{1.65}$$

where Cost^{TSP*} *represents a cost function of the dual game to the traveling salesman game* (N, Cost^{TSP}).

1.4 Conclusion

We have studied the profitable tour problem where a profit-maximizing carrier decides whether to visit a particular customer with respect to the prize the customer offers for being visited and traveling cost associated with the visit, all in the context of other customers. Our focus has been on the prizes that customers need to offer to ensure being visited by the carrier. This can be formulated as a cooperative game where customers may form coalitions and make decisions on the prize values cooperatively. We have defined the profitable tour game describing the situation in which customers need to compete or cooperate and make themselves worth being visited. This problem relates to logistics applications such as the evaluation of new customers in traveling salesman problems or customer selection in less-than-truckload transportation.

We have found several properties of the optimal prizes to be offered if the coalition of all customers, the grand coalition, is formed. These properties describe a relationship between the prizes and the underlying traveling salesman game to provide another connection with an extensive literature on core allocations in traveling salesman games. Our most important result states that the set of optimal prizes to be offered coincides with the core of the underlying traveling salesman game if this core is nonempty. Other results include the optimal prizes description for the empty-core case and relation of the prizes to the dual game of the traveling salesman game.

Overall, we have analyzed the total cost associated with each coalition of customers and, in form of the prize allocation, the strategies to achieve it. Our analysis opens further research opportunities for studying cost allocation to divide the costs paid out in form of prizes among the cooperating customers. Analyzing the core of the profitable tour game and the performance of specific allocation methods are interesting avenues for future research.

References

Archetti, C., Speranza, M. G., and Vigo, D. (2014). Vehicle routing problems with profits. In Toth, P. and Vigo, D., editors, *Vehicle Routing: Problems, Methods, and Applications, 2nd ed.*, pages 273–297. SIAM, Philadelphia.

Engevall, S., Göthe-Lundgren, M., and Värbrand, P. (1998). The traveling salesman game: An application of cost allocation in a gas and oil company. *Ann. Oper. Res.*, 82:203–218.

Estévez-Fernández, A., Borm, P., Meertens, M., and Reijnierse, H. (2009). On the core of routing games with revenues. *Int. J. Game Theory*, 38(2):291–304.

Feillet, D., Dejax, P., and Gendreau, M. (2005). Traveling salesman problems with profits. *Transp. Sci.*, 39(2):188–205.

Fishburn, P. C. and Pollak, H. O. (1983). Fixed-route cost allocation. *Am. Math. Mon.*, 90(6):366–378.

Guajardo, M. and Jörnsten, K. (2015). Common mistakes in computing the nucleolus. *Eur. J. Oper. Res.*, 241(3):931–935.

Guajardo, M. and Rönnqvist, M. (2016). A review on cost allocation methods in collaborative transportation. *Int. Trans. Oper. Res.*, 23(3):371–392.

Kimms, A. and Kozeletskyi, I. (2016). Core-based cost allocation in the cooperative traveling salesman problem. *Eur. J. Oper. Res.*, 248(3):910–916.

Laporte, G. (2007). What you should know about the vehicle routing problem. *Nav. Res. Logist.*, 54(8):811–819.

Laporte, G. and Martín, I. R. (2007). Locating a cycle in a transportation or a telecommunications network. *Networks*, 50(1):92–108.

Potters, J. A. M., Curiel, I. J., and Tijs, S. H. (1992). Traveling salesman games. *Math. Program.*, 53(1–3):199–211.

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM J. Appl. Math.*, 17(6):1163–1170.

Šomplák, R., Nevrlý, V., Málek, M., Pavlas, M., and Klemeš, J. J. (2017). Network flow based model applied to sources, sinks and optimal transport of combustible waste. *Chem. Eng. Trans.*, 61:991–996.

Sudhölter, P. (1996). The modified nucleolus as canonical representation of weighted majority games. *Math. Oper. Res.*, 21(3):734–756.

Sun, L. and Karwan, M. H. (2015). On the core of traveling salesman games. *Oper. Res. Lett.*, 43(4):365–369.

Sun, L., Rangarajan, A., Karwan, M. H., and Pinto, J. M. (2015). Transportation cost allocation on a fixed route. *Comput. Ind. Eng.*, 83:61–73.

Tamir, A. (1989). On the core of a traveling salesman cost allocation game. *Oper. Res. Lett.*, 8(1):31–34.

Tarashnina, S. (2011). The simplified modified nucleolus of a cooperative TU-game. *TOP*, 19(1):150–166.

Chapter 2

Cooperative game-theoretic features of cost sharing in location-routing*

Ondrej Osicka[a], Mario Guajardo[a], Thibault van Oost[a,b]

[a]Department of Business and Management Science, NHH Norwegian School of Economics, Bergen, Norway
[b]Louvain School of Management, Université catholique de Louvain, Louvain-la-Neuve, Belgium

Abstract

While the interest in both collaborative logistics and location-routing has grown considerably, horizontal cooperation in location-routing problems remains fairly unattended. This article studies several variants of the location-routing problem using a cooperative game-theoretic framework. The authors derive characteristics in terms of subadditivity, convexity, and non-emptiness of the core. Moreover, for some of the game variants, it is shown that for facility opening costs substantially larger than the costs associated with routing, the core is always non-empty. The theoretical results are supported by numerical experiments aimed at illustrating the properties and deriving insights. Among others, it is observed that, while in general it is not possible to guarantee core allocations, in a huge majority of cases the core is non-empty.

Keywords: Collaborative logistics; Location-routing; Cooperative game theory; Cost allocation

*A published version of this chapter appears in *International Transactions in Operational Research*, 27(4), 2157-2183.

2.1 Introduction

The European Union describes horizontal cooperation as "an agreement or concerted practice between companies operating at the same level(s) in the market" (European Union, 2001). In transportation and logistics, horizontal cooperation has been recognized as an important instrument to reduce costs, increase productivity, improve customer service, stabilize or improve the market position, among other benefits (Cruijssen et al., 2007a,b). Recent literature surveys by Guajardo and Rönnqvist (2016), Gansterer and Hartl (2018), and Cleophas et al. (2019) reveal a considerable growth in the number of articles and applications of collaborative transportation. This literature positions collaborative transportation as an important mechanism to reduce cost and negative environmental effects, and to increase profits and service levels.

In parallel, the literature on location-routing problems has also grown considerably. The central problem in this literature is to locate facilities while simultaneously finding routes to serve customers from these facilities. It has early been recognized that tackling these decision problems separately may lead to suboptimal solutions (Perl, 1987, Salhi and Rand, 1989). Reviews by Prodhon and Prins (2014), Drexl and Schneider (2015), and Schneider and Drexl (2017) give account of a broad range of applications as well as a considerable progress in solution methods for location-routing problems. Recent emergence of dynamic on-demand warehousing (Sinha, 2016) emphasizes the potential in applicability even further as it introduces facility location decisions of an operational (short-term) nature.

Cleophas et al. (2019) list the cooperative planning within location-routing problems as one of the future challenges in operations research. Despite the growing interest in both collaborative transportation and location-routing, the integration of these two areas remains fairly unattended by researchers. The only exceptions, to our knowledge, are van Oost (2016), Quintero-Araujo et al. (2019), and Ouhader and Kyal (2017), which present promising studies on the benefits of collaboration in location-routing. Besides the academic interest, the intersection of these two areas is motivated from practical situations, such as the installation of urban consolidation centres for city logistics and the formation of strategic alliances. For example, Paddeu (2017) recounts the case of the Bristol-Bath freight consolidation centre from the perspective of its users (106 retailers) and points to pricing and cost coverage as important factors for success or failure. If a consolidation centre is located in the periphery of a city to serve retailers in the city centre by shared routes, they will naturally be concerned about the cost of the service and how this compares to the non-shared solution. The location of the centre may, therefore, play an important role in the willingness to adopt the shared solution. Another example in practice is given by the alliance between Colgate-Palmolive, GlaxoSmithKline, Henkel, and Sara Lee in France (Eyers, 2010). The alliance started in 2005 with cooperation on routing only, but subsequently led to a decrease in the number of facilities of the alliance from four to only one. With these sharing practices emerging as important mechanisms to improve operations, it is increasingly important for firms to understand the economic foundations for cooperation in location-routing.

In this article, we study several variants of the location-routing problem using a cooperative game-theoretic framework. After defining a transferable utility game for each of these variants, the main focus of this article turns to exploration of essential properties that characterize the behaviour of cooperation in location-routing from a cost sharing perspective such as the subadditivity, convexity and emptiness of the core. In particular, this framework is useful to answer whether incentives exist for all firms to align under a collaborative location-routing approach. This has been recognized as one of the main barriers for the implementation of collaboration in logistics (Audy et al., 2012, Basso et al., 2019). In this respect, we show that for some of the variants of the collaborative location-routing games, the core is guaranteed to be non-empty when the facility opening costs are substantially larger than the traveling costs and the costs of using vehicles.

The rest of the article is organized as follows. Section 2.2 introduces the game-theoretic concepts used throughout the article. In Section 2.3, we briefly overview related literature. Section 2.4 describes the main variants of the location-routing problem and defines the corresponding cooperative games. In Section 2.5, we derive and investigate theoretical properties of the location-routing games. Section 2.6 summarizes numerical results obtained from experiments. Section 2.7 presents our concluding remarks.

2.2 Preliminaries

Let $N = \{1, \ldots, n\}$ denote the set of all players and $\mathscr{S}$ the set of all subsets of N. A *transferable utility game* is a pair (N, C) where $C\colon \mathscr{S} \to \mathbb{R}$ is the *characteristic function* assigning to each coalition $S \in \mathscr{S}$ the optimal cost achievable by cooperation of players within this coalition.

A transferable utility game (N, C) is considered *subadditive* if its characteristic function is subadditive, i.e.,

$$C(S \cup T) \leq C(S) + C(T) \qquad \forall S, T \subseteq N\colon S \cap T = \varnothing. \tag{2.1}$$

If the subadditivity holds, no coalition is less profitable than some of its partitions. That is, there is no loss involved in merging coalitions with respect to the total costs incurred.

The *convexity* of a transferable utility game is a stronger property which requires submodularity of its characteristic function, i.e.,

$$C(S \cup T) + C(S \cap T) \leq C(S) + C(T) \qquad \forall S, T \subseteq N, \tag{2.2}$$

or equivalently (Schrijver, 2003),

$$C(S \cup \{i\}) - C(S) \geq C(T \cup \{i\}) - C(T) \qquad \forall i \in N,\, \forall S, T \subseteq N\colon S \subseteq T \subseteq N \setminus \{i\}. \tag{2.3}$$

Studies of transferable utility games naturally lead to studies of cost allocations which prescribe the costs to be paid by particular players within the cooperation. For a transferable utility game of n players, the cost allocation is a vector $(\pi_1, \ldots, \pi_n) \in \mathbb{R}^n$. Conditions of efficiency,

$$\sum_{j \in N} \pi_j = C(N), \tag{2.4}$$

and rationality,

$$\sum_{j \in S} \pi_j \leq C(S) \qquad \forall S \subseteq N, \tag{2.5}$$

define a set of cost allocations known as the *core* as first introduced by Shapley (1955). A vector in the core proposes a redistribution of total costs of the grand coalition N and is said to be *stable*, as there are no incentives for any subset of players to deviate from the collaboration. In the literature, this has been recognized as an essential condition to sustain the cooperation in practice. It is hence desirable for a transferable utility game to have a non-empty core and it is interesting to study whether a game admits or not allocations in the core.

2.3 Literature review

Cruijssen et al. (2007a,b) recognize horizontal cooperation in transportation and logistics as an important instrument to reduce costs, increase productivity, improve customer service, stabilize or improve the market position, among other benefits. This can be achieved for example with optimized vehicle capacity utilization, reduced empty mileage or minimized cost of non-core operations. A literature survey by Guajardo and Rönnqvist (2016) gathered articles and applications of collaborative transportation, with special focus on the benefits that the horizontal cooperation introduces and their redistribution among the cooperating parties using appropriate cost-sharing mechanisms. More recently, a survey by Gansterer and Hartl (2018) focuses specifically on collaboration in vehicle routing problems.

The vehicle routing problem tackles a question of how to perform tours to visit a group of customers from one or more facilities using one or more vehicles. For a current state of this literature, see for example Adewumi and Adeleke (2018). In the vehicle routing problems, the horizontal cooperation can be found on several different levels in form of cooperation of customers, carriers or shippers. For the customer level, the traveling salesman game and the basic vehicle routing game are among the most studied. For their definitions, see for example studies by Potters et al. (1992) and Göthe-Lundgren et al. (1996). These problems focus on allocating costs of realized tours among the visited customers. When it comes to the shippers and carriers, the collaborative vehicle routing problem (CoopVRP) has been studied. In the CoopVRP, the cooperating shippers and carriers pool their customers and allow for visiting customers of different shippers within the same tour. In their survey, Gansterer and Hartl (2018) acknowledge a usual lack of distinction between a cooperation of shippers and a cooperation of carriers. They suggest that it might be sometimes important to recognize a difference in the information they possess. However, since the focus of this article is on centralized collaborative planning, we assume that the shippers and carriers possess the same information and make decisions jointly. Thus, their distinction is not needed and we refer to them simply as shippers. When it comes to the game-theoretic properties, it is easy to see that in the CoopVRP, as defined for example by Zibaei et al. (2016), the subadditivity is generally satisfied as a combination of any non-cooperative solutions forms a feasible solution in the cooperative formulation as long as the facilities' capacities are unconstrained. For the case of traveling salesman games, Potters et al. (1992) show that the convexity is not generally satisfied and the core can be empty. Since the traveling salesman game can be formulated as a special case of the CoopVRP with vehicles of a large capacity and shippers each serving one customer and possessing one facility in the same location for all shippers, the results remain the same for the CoopVRP.

For situations where the shippers already know where their potential customers would be, but they do not have a facility from which to serve them, the facility location problem becomes useful. The facility location problem aims to find the optimal location of facilities such that each customer gets assigned to a facility. The optimality lies in minimizing the sum of facility opening costs and connection costs. Computation of the connection costs differs among various formulations of the facility location problem. For an overview of several variants of the problem, see Laporte et al. (2015). Goemans and Skutella (2004) introduced the cooperative facility location game to deal with a cooperative formulation of the problem in which several shippers aim to get their customers assigned to a facility. By allowing the customers to be connected to facilities of different shippers, this problem might allow for substantial savings. With no restriction on the facilities' capacities, the FLG is subadditive as a combination of any non-cooperative solutions forms a feasible solution in the cooperative formulation. On the other hand, Goemans and Skutella (2004) show that the convexity is not generally satisfied and the core can be empty.

Perl (1987) and Salhi and Rand (1989) pointed out that tackling the decisions on facility location and vehicle routing separately may lead to suboptimal solutions. With the aim to connect these problems, the location-routing problem was addressed by a large stream of literature since then, as documented in surveys by Nagy and Salhi (2007), Prodhon and Prins (2014), Drexl and Schneider (2015), and Schneider and Drexl (2017). Among others, these surveys discuss different variants of the location-routing problem such as, for example, the standard location-routing problem (LRP), the capacitated location-routing problem (LRP-C), and the location-routing problem with a limited number of facilities (LRP-L). The LRP is used in numerous applications as documented, for example, by Watson-Gandy and Dohrn (1973), Or and Pierskalla (1979), Bruns et al. (2000), and Ambrosino et al. (2009) for the cases of food and drink distribution, blood bank location, parcel delivery, and food distribution, respectively. As claimed by Prodhon and Prins (2014), the LRP-C is nowadays addressed to a larger extent than the LRP. For example, Nambiar et al. (1981, 1989), Gunnarsson et al. (2006), and Marinakis and Marinaki (2008) utilize the LRP-C in rubber plant location, shipping industry, and wood distribution, respectively. The LRP-L arises when some types of facilities cause nuisance and social rejection. A city taking this into account could decide to impose a limit on the number of a certain type of facilities. For instance, Caballero et al. (2007) deal with one such problem by using the LRP-L for location of incineration plants for disposal of solid animal waste.

While a large number of articles in this literature stream have been devoted to development of solution methods, very few have studied collaboration in location-routing. Among the exceptions, Quintero-Araujo et al. (2019) use numerical experiments to compare the non-cooperative and cooperative scenarios in the location-routing problem in terms of the total cost and the total traveling-related CO_2 emissions. For the location-routing problem variant with two-echelons, Ouhader and Kyal (2017) analyze the cooperation based on three different objectives, minimizing the total cost, minimizing the total amount of traveling-related CO_2 emissions, and maximizing the number of created job opportunities (social impact). Through numerical experiments, they observe how each objective separately affects the other measures. In literature on collaborative transportation problems, cooperative game theory is commonly used to derive theoretical properties and investigate implications of the collaboration. To our knowledge, our article is the first one studying cooperation in location-routing problems from a cooperative game-theoretic perspective.

2.4 Location-routing problems and game definitions

In this section, we briefly present some of the main location-routing problem variants and then formally introduce the definition of transferable utility games for these variants.

2.4.1 Location-routing problem variants

Standard location-routing problem

Our departing point is the standard location-routing problem (LRP). Its definition, as described for example by Prodhon and Prins (2014), assumes a set of potential facility locations and a set of customers (and their corresponding demands) to be given. The LRP then aims at finding locations and routes that minimize the total routing costs, costs of using vehicles and costs of opening facilities while assuring that all customers are visited and their demand is satisfied.

Let G be the set of feasible sites of candidate facilities, I the set of customers to be served and $V = G \cup I$ the set of all such nodes. Let K be a set of all vehicles available for routing from the facilities (no facility has a specific fleet). Let c_{ij} be the cost of traveling from node $i \in V$ to node $j \in V$, a the cost of acquiring a vehicle, and f_g the cost of establishing and operating a facility at site $g \in G$. The number of units demanded by customer $i \in I$ is d_i. The capacity of one vehicle is q. To make the values comparable, they need to be normalized with respect to a certain time period. This depends on a particular application in question. Sometimes, for example, an annual average may serve the purpose.

The LRP can be formulated as an integer linear programming model. Let X_{ijk} be a binary decision variable which takes value 1 if vehicle k travels directly from node i to node j, and zero otherwise (for $k \in K$, $i \in V$, $j \in V$, such that $i \neq j$ and at least one of these two nodes belongs to I), and Z_g a binary decision variable which equals 1 when facility at site $g \in G$ is open, and zero otherwise. Then, the model is formulated below.

$$\min_{X_{ijk}, Z_g} \left(\sum_{i \in V} \sum_{j \in V} \sum_{k \in K} c_{ij} X_{ijk} + \sum_{k \in K} \left(a \sum_{g \in G} \sum_{j \in I} X_{gjk} \right) + \sum_{g \in G} f_g Z_g \right) \tag{2.6}$$

$$\text{s.t.} \quad \sum_{k \in K} \sum_{i \in V} X_{ijk} = 1 \qquad \forall j \in I, \tag{2.7}$$

$$\sum_{j \in I} \sum_{i \in V} d_j X_{ijk} \leq q \qquad \forall k \in K, \tag{2.8}$$

$$\sum_{i \in V} X_{ipk} - \sum_{j \in V} X_{pjk} = 0 \qquad \forall p \in V, \forall k \in K, \tag{2.9}$$

$$\sum_{g \in G} \sum_{j \in I} X_{gjk} \leq 1 \qquad \forall k \in K, \tag{2.10}$$

$$\sum_{i \in S} \sum_{j \in S} X_{ijk} \leq |S| - 1 \qquad \forall S \subseteq I \colon S \neq \varnothing, \forall k \in K, \tag{2.11}$$

$$\sum_{j \in I} X_{gjk} - Z_g \leq 0 \qquad \forall k \in K, \forall g \in G, \tag{2.12}$$

$$X_{ijk} \in \{0, 1\} \qquad \forall k \in K, \forall i, j \in V \colon i \neq j, \{i, j\} \cap I \neq \varnothing, \tag{2.13}$$

$$Z_g \in \{0, 1\} \qquad \forall g \in G. \tag{2.14}$$

Objective function (2.6) minimizes the sum of routing costs, the costs of using vehicles and the costs of operating facilities. Constraints (2.7) state that each customer must be served by one vehicle. Constraints (2.8) state that the capacity of vehicles must be respected. Constraints (2.9) are flow conservation constraints. Constraints (2.10) express that no vehicle can depart from more than one facility. Constraints (2.9) and (2.10) impose that every used vehicle has to come back to the same facility it departed from. Constraints (2.11) eliminate sub-tours. Constraints (2.12) ensure that a facility can be used if and only if this facility is open. Constraints (2.13) and (2.14) state the binary nature of the variables.

Note that in this formulation, constraints (2.7) and (2.8) imply that each customer needs to be served by a single vehicle. In combination with constraints (2.12) this also means that the whole demand of a customer needs to be satisfied from only one facility. Consequently, the problem might become infeasible if there is not enough vehicles to satisfy the total demand or if a customer demands more than the capacity of the vehicles.

Capacitated location-routing problem

The capacitated location-routing problem (LRP-C) introduces an upper bound in the supply available at the facilities. The limited capacity of the facilities leads to the following modifications of the model (2.6)–(2.14).

Parameter w_g is added and defined as the capacity of facility $g \in G$. New binary decision variable F_{ig} takes value 1 if customer i is assigned to facility g, and 0 otherwise (for $i \in I$, $g \in G$). In addition to constraints (2.7)–(2.14), constraints

$$\sum_{u \in V} X_{guk} + \sum_{u \in V} X_{uik} \leq 1 + F_{ig} \qquad \forall g \in G, \forall i \in I, \forall k \in K \tag{2.15}$$

need to hold. To make constraints (2.12) account for the limited capacity of the candidate facilities, they are replaced by

$$\sum_{i \in I} d_i F_{ig} \leq w_g Z_g \qquad \forall g \in G. \tag{2.16}$$

Constraints (2.15) state that each customer served by a vehicle departing from a certain facility must be assigned to this facility. The term $\sum_{u \in V} X_{guk}$ is equal to 1 if the vehicle k departs from the facility g while $\sum_{u \in V} X_{uik}$ is equal to 1 if this vehicle supplies customer i. If both terms are equal to 1, it implies F_{ig} to equal 1. Constraints (2.16) ensure that the capacities of the facilities must be respected. At the same time, constraints (2.16) forbid customers to be assigned to a facility which is not open.

To avoid infeasibility of the LRP-C, there need to be enough facility candidates with large enough capacities such that each customer can be assigned to and fully supplied by only one facility. Besides, as for the LRP, there need to be enough vehicles to satisfy the total demand and customers cannot demand more than the capacity of the vehicles.

Location-routing problem with a limited number of facilities

The location-routing problem with a limited number of facilities (LRP-L) introduces an upper bound in the maximum number of facilities that can be opened. If a condition that only l facilities can be used takes place, the model (2.6)-(2.14) needs to be modified by introducing the parameter l and adding the constraint

$$\sum_{g \in G} Z_g \leq l \tag{2.17}$$

which ensures that the total number of opened facilities is less or equal to the limit l.

The feasibility of the LRP-L is subject to the same conditions as in the case of the LRP, that is, there need to be enough vehicles to satisfy the total demand and customers cannot demand more than the capacity of the vehicles.

2.4.2 Location-routing games

Our interest lies in a collaborative version of the location-routing problem. When shippers collaborate, the overall problem opens opportunities to combine their customers within the same tours and serve their demands from shared facilities. To model this situation, we use a cooperative game-theoretic framework.

Standard location-routing game

The standard location-routing game (LRG) is defined as a transferable utility game by the tuple (N, C) where $N = \{1, \ldots, n\}$ denotes the set of all players (shippers), $\mathscr{S}$ the set of all subsets of N, and $C : \mathscr{S} \to \mathbb{R}$ the characteristic function. The characteristic function C assigns to each coalition $S \in \mathscr{S}$ its optimal cost, that is, the optimal objective value of the corresponding LRP (by convention, $C(\varnothing) = 0$). Let I_n be the set of customers originally to be served by a shipper $n \in N$. The corresponding LRP for computing $C(S)$ is then the LRP in which the set of customers to be served is $I = \bigcup_{n \in S} I_n$. This definition requires each customer to be a client of only one shipper. However, customers with demand from several shippers may still be modeled as multiple customers with identical location.

This modeling approach assumes that the cooperation on location and routing decisions does not affect customers' choice of shippers and there is no competition for customers associated with the coalition formation. Particularly, in the model, the customers' demands from a shipper do not depend on the coalition that the shipper belongs to.

Location-routing game extensions

Similarly, for the two other variants of the location-routing problem, the LRP-C and the LRP-L, we define the capacitated location-routing game (LRG-C1) and the location-routing game with a limited number of facilities (LRG-L1), respectively. The LRG-C1 introduces a problem where each coalition solves a location-routing problem with capacitated facilities. The LRG-L1 extends the standard location-routing game by assuming a limit on the number of facilities that can be located by each coalition. These capacities and limits are independent of the coalitions' size or members and remain constant.

Unlike in other collaborative transportation problems (such as the collaborative vehicle routing problem or the facility location game), in both the LRG-C1 and the LRG-L1, when cooperation takes place, the original strategies before such cooperation are not necessarily feasible. When two shippers use the same facility and operate on its full capacity in their stand-alone strategies, this strategy is not possible once they cooperate. The same problem occurs for the LRG-L1 if the shippers already use the maximum allowed number of facilities in their stand-alone strategies. Therefore, we also formulate alternative models for the capacitated location-routing game and the location-routing game with a limited number of facilities as LRG-C2 and LRG-L2, respectively.

Let the LRG-C2 be defined as a transferable utility game in a similar way as the LRG-C1, that is, for computation of each characteristic function value, the LRP-C is solved. Now, however, these LRP-C's differ not only in the sets of customers $I = \bigcup_{n \in S} I_n$, but also in the capacities w_g. Here, each combination of a shipper $n \in N$ and a candidate facility $g \in G$ is associated with a parcial capacity w_{gn} which reflects the capacity of facility g available to shipper n in a non-cooperative case. For a coalition S, the capacity of facility g is determined as a sum of the partial capacities of all present shippers $w_g = \sum_{n \in S} w_{gn}$.

Similarly, let the LRG-L2 be defined as a transferable utility game in which the LRP-L is solved for each characteristic function value computation. As opposed to the LRG-L1, however, these LRP-L's differ not only in the sets of customers $I = \bigcup_{n \in S} I_n$, but also in the limits on the number of facilities l. Each shipper is now associated with a partial limit l_n, the limit on the number of facilities in a non-cooperative case. In the cooperative case, the limit on the number of facilities to be opened by a coalition S equals the sum of the partial limits of all present shippers $l = \sum_{n \in S} l_n$.

As opposed to the LRG-C1 and the LRG-L1, in the LRG-C2 and the LRG-L2, the capacities and limits are dependent on the coalitions' size and members. The shippers are associated with partial capacities or limits which may be utilized in any coalition by adding up the partial capacities or limits of all its members.

All the variants have their place in practice. The LRG-C1 can be used, for example, if shippers may for some reason be associated only with facilities up to a certain capacity per location. For situations where several facilities can be opened at each location, the LRG-C1 is an appropriate model too. This reflects a case in which facilities consist of blocks of a certain capacity and each shipper or coalition of shippers determines how many blocks to open. On the other hand, in a single-commodity situation, if shippers for example have a maximum supply available for each candidate facility, its value stands for the partial capacity of this facility in the LRG-C2 formulation. If the facilities cause social rejection and shippers do not want to be associated with more than a certain number of them, cooperation might not increase the number and the LRG-L1 becomes the appropriate model. If, on the other hand, there is an enforcement preventing a shipper from operating more than a certain number of facilities, cooperation might allow the total limit to be a sum of the limits of all involved shippers. Such situation requires the use of the LRG-L2.

2.5 Theoretical results

Cooperation does not necessarily guarantee beneficial outcomes for all parties. In the following, we investigate the satisfaction of subadditivity and convexity properties by the location-routing games as well as whether they admit or not allocations in the core.

2.5.1 Game-theoretic properties of the standard location-routing game

Proposition 1. *The standard location-routing game is subadditive.*

Proof. For the LRG, it is easy to see that for any disjoint sets of players S and T, the solution of a LRP where customers of shippers in S and T are served using the same facilities, routes and vehicles as in the separate problems for S and T is a feasible solution of the LRP for set $S \cup T$. The value of $C(S) + C(T)$ is therefore an upper bound on the optimal objective value $C(S \cup T)$ of the problem for $S \cup T$ and the LRG is hence subadditive. □

Proposition 2. *The core of the standard location-routing game can be empty.*

Proof. Göthe-Lundgren et al. (1996) present an example to prove that the core of the basic vehicle routing game can be empty. By adapting this example, we can prove that the core of the LRG can be empty as well. Figure 2.1 illustrates the location of three customers (1, 2 and 3) and feasible sites of candidate facilities (A, B and C). Each customer is a client of a different

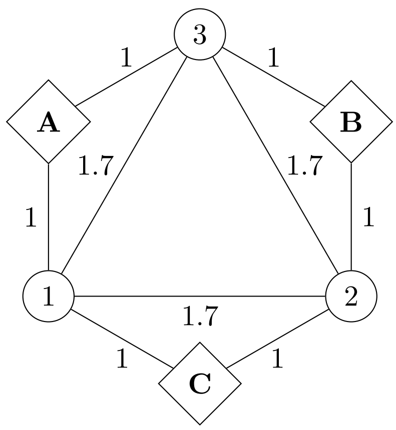


Figure 2.1: A standard collaborative location-routing problem with an empty core (the circles and diamonds represent customers and candidate facilities, respectively, and the arcs indicate the traveling costs)

shipper which can hence be referred to by 1, 2 and 3 as well. The figure also contains the transportation costs. The costs of establishing and operating a facility are equal to one unit for each of the sites. The demand of each customer is of one unit. The capacity of each vehicle is of two units. The cost of using a vehicle is set to zero. Calculating the characteristic cost function for the singletons, we obtain $C(\{1\}) = C(\{2\}) = C(\{3\}) = 3$. The routing cost is equal to 2 and the facility opening cost to 1. The facility opened is one of the two adjacent to the customer. For the two-player coalitions, $C(\{1,2\}) = C(\{1,3\}) = C(\{2,3\}) = 4.7$ (routing cost of 3.7 and facility opening cost of 1). For the three-player coalition, $C(\{1,2,3\}) = 7.7$ (routing cost of 5.7 and facility opening cost of 2). From the costs, we can notice there is an incentive for a two-player coalition. It is clearly more beneficial than a non-cooperative state. In case of the three-player coalition, whichever the cost allocation is, there will always be two shippers that could get better off by excluding the third one. Hence, the core is empty. □

Corollary 2. *The standard location-routing game is not necessarily convex.*

Proof. According to Shapley (1971), the core of a convex game is not empty. Equivalently, by contraposition, if the core of a game is empty, the game must be non-convex. Since the core of the LRG can be empty, it follows that this game cannot be convex in general.

Moreover, in the problem of Figure 2.1, taking coalitions $S = \{1\}$, $T = \{1,2\}$ and $i = 3$, the left-hand side of inequality (2.3) is $4.7 - 3 = 1.7$ while the right-hand side is $7.7 - 4.7 = 3$, thus the inequality is violated and this is an example of a non-convex game. □

2.5.2 Game-theoretic properties of the location-routing game extensions

Proposition 3. *The LRG-C1 is not necessarily subadditive or convex.*

Proof. A counter-example is illustrated in Figure 2.2. Again, we consider each customer being assigned to a different shipper. The costs of establishing and operating facilities are equal to one

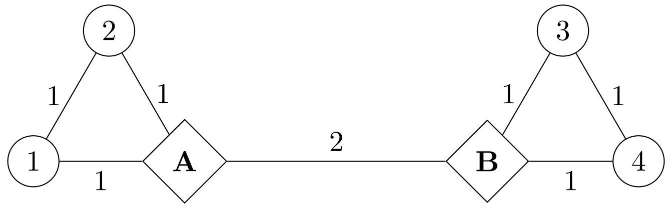


Figure 2.2: A non-subadditive capacitated location-routing game (the circles and diamonds represent customers and candidate facilities, respectively, and the arcs indicate the traveling costs)

unit for site A and six units for site B. Both candidate facilities have a maximum capacity equal to two units. The demand for each customer is of one unit. The capacity of each vehicle is of two units. The cost of using a vehicle is set to zero. For coalition $S = \{1,2\}$, facility A is opened and $C(S) = 4$. For coalition $T = \{3,4\}$, facility A is opened as well and $C(T) = 8$. However, for $S \cup T$, both facilities are opened and $C(S \cup T) = 13$ which causes that $C(S \cup T) \nleq C(S) + C(T)$ and the subadditivity (as well as convexity) does not hold. □

Proposition 4. *The LRG-L1 is not necessarily subadditive or convex.*

Proof. Considering the same example (Figure 2.2) for the location-routing game with a limited number of facilities, but now with omitting the maximum capacities, introducing a limit of $l = 1$ on the number of facilities and having the costs of establishing and operating facilities equal to 1 for both A and B, leads to $C(S) = 4$, $C(T) = 4$ and $C(S \cup T) = 11$. Hence, again $C(S \cup T) \nleq C(S) + C(T)$ and the subadditivity and convexity properties do not hold. □

Proposition 5. *The core of the LRG-C1 and the LRG-L1 can be empty.*

Proof. Clearly, if each facility offered a maximum capacity that could accommodate all customers in the game, the LRG-C1 would reduce to the LRG. Similarly, if the limit on the number of facilities l was equal or higher than the number of all potential facility sites, the LRG-L1 would reduce to the LRG. Therefore, the example in Figure 2.1 could be used to show a possibility of an empty core in the LRG-C1 and the LRG-L1 as well as for the case of the LRG. □

Proposition 6. *The LRG-C2 and the LRG-L2 are subadditive.*

Proof. For both the LRG-C2 and the LRG-L2, it is easy to see that for any disjoint sets S and T, the solution where customers of shippers in S and T are served using the same facilities, routes and vehicles as for the separate problems for S and T is a feasible solution for set $S \cup T$. The value of $C(S) + C(T)$ is therefore an upper bound on the optimal objective value $C(S \cup T)$ and subadditivity is hence satisfied. □

Proposition 7. *The core of the LRG-C2 and the LRG-L2 can be empty.*

Proof. For illustration of the possibility of an empty core, the same reasoning as in the case of the LRG-C1 and the LRG-L1 could be used. That is, large enough capacities and large enough limits on the number of facilities would reduce the LRG-C2 and the LRG-L2 to the LRG which, as shown in the example of Figure 2.1, can have an empty core. □

Corollary 3. *The LRG-C2 and the LRG-L2 are not necessarily convex.*

Proof. As in the proof of Corollary 2, the possible emptiness of the core consequently leads to non-convexity. □

All the aforementioned results are summarized in Table 2.1 along with the property satisfaction of some of the collaborative transportation problems outlined in Section 2.3, namely the collaborative vehicle routing problem (CoopVRP) and the facility location game (FLG). If a checkmark is missing, this property might be satisfied for particular instances, but does not hold in general. The results might suggest that all collaborative transportation problems allow for an empty core. To avoid such misinterpretation, without its definition, we also include the transportation game (TG) introduced by Samet et al. (1984). It is an example of a game which is always subadditive and has a non-empty core, but is not necessarily convex (Sánchez-Soriano et al., 2001).

Table 2.1: Properties of different variants of the location-routing games, the collaborative vehicle routing problem, the facility location game, and the transportation game

	Subadditive	Convex	Non-empty core
LRG	✓	-	-
LRG-C1	-	-	-
LRG-L1	-	-	-
LRG-C2	✓	-	-
LRG-L2	✓	-	-
CoopVRP	✓	-	-
FLG	✓	-	-
TG	✓	-	✓

2.5.3 Impact of the facility costs

The facility opening costs, f_g in the model (2.6)–(2.14), play an important role in the location-routing games. They stand for all costs necessary to establish and operate a facility. In what follows, we will show that, if the facility opening costs are substantially larger than the traveling costs and the costs of using vehicles, the core of the LRG is guaranteed to be non-empty.

Theorem 3. *For any traveling and vehicle-related costs, there exists $K \in \mathbb{R}$ such that, if $f_g \geq K$ for all $g \in G$, the core of the respective standard location-routing game is non-empty.*

Proof. Solving the model (2.6)–(2.14) for the grand coalition N results in an optimal objective value of $C(N)$. This cost could be allocated to the shippers for example such that each of them is allocated the same cost, that is,

$$\pi_j = \frac{C(N)}{|N|} \qquad \forall j \in N. \tag{2.18}$$

This allocation clearly satisfies the efficiency condition (2.4).

It would be feasible to open only the least costly facility and serve all customers from this facility. Denoting the opening cost of this facility by $f_{g^*} = \min_{g \in G} f_g$ and any (not necessarily minimal) routing costs and costs of using vehicles needed to serve all customers by utilizing only this facility by VRC_{g^*}, the total cost would be at least as large as $C(N)$,

$$C(N) \leq f_{g^*} + VRC_{g^*}, \tag{2.19}$$

and hence

$$\pi_j \leq \frac{1}{|N|}\left(f_{g^*} + VRC_{g^*}\right) \qquad \forall j \in N. \tag{2.20}$$

Then, for any proper subset $S \subset N$,

$$\sum_{j \in S} \pi_j \leq \frac{|S|}{|N|}\left(f_{g^*} + VRC_{g^*}\right) \leq \frac{|N|-1}{|N|}\left(f_{g^*} + VRC_{g^*}\right). \tag{2.21}$$

If the facility costs are substantially larger than the traveling costs and the costs of using vehicles, for example

$$VRC_{g^*} \leq \frac{f_{g^*}}{|N|-1}, \tag{2.22}$$

(2.21) then implies

$$\sum_{j \in S} \pi_j \leq f_{g^*} \qquad \forall S \subset N. \tag{2.23}$$

Since each coalition $S \subset N$ needs to locate at least one facility and f_{g^*} is the minimal cost associated with this, then

$$\sum_{j \in S} \pi_j \leq C(S) \qquad \forall S \subset N. \tag{2.24}$$

In combination with satisfaction of the efficiency condition, this means that all rationality conditions (2.5) are satisfied and $(\pi_1, \ldots, \pi_{|N|})$ belongs to the core. Hence, the core is non-empty.

Since VRC_{g^*} is not dependent on the value of f_{g^*}, it is easy to see that any K such that

$$VRC_g \leq \frac{K}{|N|-1} \qquad \forall g \in G \tag{2.25}$$

guarantees a non-empty core and proves the Theorem 3. □

This result for the LRG can be generalized to the LRG-L1 and the LRG-L2 as well. However, in the case of the LRG-C1 or the LRG-C2, due to their capacities, it is not always feasible to locate only one facility and the same argument cannot be used.

2.6 Numerical results

To illustrate the theoretical results and explore the satisfaction of the properties that do not hold in general, we have conducted a computational study using randomly generated instances as well as instances from the literature on location-routing problems. We also address here the problem of how to split the savings by applying cost allocation methods frequently used in the literature.

Unless stated otherwise, all presented models and methods are implemented and solved using AMPL/Gurobi 7.5.0.

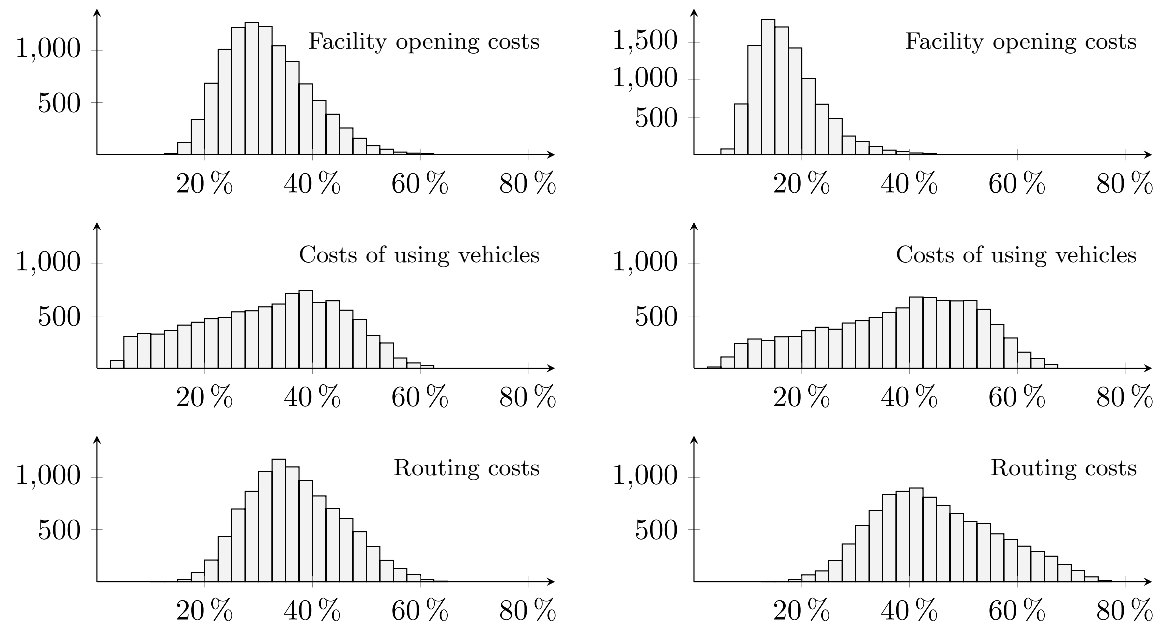
2.6.1 Experiment design

We have generated 10,000 instances, all of them containing nine sites of candidate facilities and three shippers, each of them having two or three customers.

Given a square 100×100, the coordinates x and y of the customers and facility candidates follow a uniform distribution between 0 and 100. The transportation cost c_{ij} between two nodes is the Euclidean distance between the two nodes. Each shipper is randomly assigned two or three customers. The demand d_j for each customer follows a uniform distribution between 10 and 100 and the vehicle capacity q follows a uniform distribution between 100 and 200. The fleet is homogeneous. The cost of using a vehicle a is the same for all vehicles and ranges between 10 and 200 and each facility opening cost f_g follows a uniform distribution between 100 and 300.

In the case of the LRG-L1, the limit on number of facilities l takes value 1, 2 or 3. The partial limits on number of facilities l_n in the LRG-L2 equal either 1 or 2. For the LRG-C1, the facility capacities w_g range from 100 to 500, whereas for the LRG-C2, the partial facility capacities w_{gn} range from 35 to 200. Generation of the values of w_{gn} in this way implies that for some coalitions there might be customers whose demand cannot be satisfied from only one facility. In such case, the LRG-C2 would not yield a feasible solution. We observed only two instances in which this occurred. For further analysis, we exclude those and take into account only the instances with feasible solutions.

The aim in choosing the parameters is to cover cases where the decisions on facility location are of a similar relevance as the decisions on routing. We attempt to achieve this by generating instances where the facility opening costs, the costs of using vehicles and the routing costs are of a similar magnitude. For illustration, in the non-cooperative case of the LRG, the facility opening costs range from 12 % to 67 % of the total costs, the costs of using a vehicle from 3 % to 64 % and the routing costs from 12 % to 65 %. Respective histograms are shown in Figure 2.3a.



(a) Non-cooperative case (b) Cooperative case

Figure 2.3: Histograms of facility opening costs, costs of using vehicles, and routing costs as a percentage of the total costs

In the cooperative case, as depicted in Figure 2.3b, we observe a slight shift of the facility opening costs to lower values as the cooperation is often accompanied by a reduction in the number of open facilities. Overall, however, we can see that the experiment covers a broad spectrum of how the costs could be distributed.

2.6.2 Game properties

Regarding the properties of subadditivity, convexity and core-emptiness, results of the experiment are shown in Table 2.2. All instances of the LRG, LRG-C2 and LRG-L2 are subadditive which confirms the theoretical results from Section 2.5. There is only a slight change in the results when it comes to the LRG-C1 and the LRG-L1 in which 2.7 % and 0.2 % of the instances respectively end up as non-subadditive. This is not surprising as it was not difficult to find examples violationg the subadditivity in the proofs of Propositions 3 and 4. We observe a huge majority of instances having a non-empty core. Nevertheless, only less than a third of the instances end up being convex. There are no substantial differences among the models. Only, in the case of the LRG-C1, satisfaction of the properties is generally slightly lower than in the other models.

Table 2.2: Satisfaction of properties in location-routing games

	Subadditive	Convex	Non-empty core
LRG	100 %	30.5 %	99.3 %
LRG-C1	97.3 %	22.7 %	92.6 %
LRG-L1	99.8 %	30.3 %	99.1 %
LRG-C2	100 %	26.1 %	99.7 %
LRG-L2	100 %	30.4 %	99.3 %

By definition, in the instances with an empty core, no cost allocation can satisfy the efficiency condition (2.4) and the rationality conditions (2.5) at the same time. By requiring only the efficiency condition to be satisfied, we can measure the amount of violation of the rationality constraints which is inevitable. For this, the strong ε-core can be utilized. The *strong ε-core*, as introduced by Shapley and Shubik (1966), is a set of all optimal solutions of $(\pi_1, \dots, \pi_{|N|})$ to problem

$$\min_{\pi_j, \varepsilon} \varepsilon \tag{2.26}$$

$$\text{s.t.} \quad \sum_{j \in N} \pi_j = C(N), \tag{2.27}$$

$$\sum_{j \in S} \pi_j \leq C(S) + \varepsilon \qquad \forall S \subseteq N, \tag{2.28}$$

$$\pi_j \in \mathbb{R} \qquad \forall j \in N, \tag{2.29}$$

$$\varepsilon \in \mathbb{R}, \tag{2.30}$$

where ε represents the maximal violation of the rationality constraints. For all variants of the location-routing game, the average maximal violation is rather low and ranges from 0.91 % to 1.70 % of the respective $C(N)$.

2.6.3 Impact of the facility costs

As discussed in Section 2.5.3, for the LRG, LRG-L1, and LRG-L2, the core is non-empty when the facility costs reach a certain size. We investigate this by generating new sets of instances with the only change being in the facility cost values. In the new instances we generate the facility opening costs as multiples of the original facility opening costs, that is, for a multiplier of 0, all facility opening costs equal zero, for a multiplier of 1, the results are the same as in Table 2.2, for a multiplier of 2, all facility opening costs are doubled, and so on. The results are reported in Table 2.3. For the LRG, we can see that the results confirm the theoretical findings

Table 2.3: Impact of the facility costs on properties of the location-routing games

	Facility cost multiplier	Subadditive	Convex	Non-empty core
LRG	0	100 %	41.4 %	97.1 %
	1	100 %	30.5 %	99.3 %
	2	100 %	32.0 %	99.9 %
	3	100 %	32.9 %	100 %
	4	100 %	33.5 %	100 %
	5	100 %	33.9 %	100 %
LRG-C1	0	95.5 %	38.4 %	93.7 %
	1	97.3 %	22.7 %	92.6 %
	3	93.9 %	19.2 %	75.1 %
	5	90.9 %	19.4 %	69.6 %
LRG-L1	0	57.3 %	17.6 %	67.1 %
	1	99.8 %	30.3 %	99.1 %
	3	100 %	32.9 %	100 %
	5	100 %	33.9 %	100 %
LRG-C2	0	100 %	23.9 %	97.3 %
	1	100 %	26.1 %	99.7 %
	3	100 %	23.3 %	99.5 %
	5	100 %	22.2 %	99.3 %
LRG-L2	0	100 %	25.7 %	97.5 %
	1	100 %	30.4 %	99.3 %
	3	100 %	32.9 %	100 %
	5	100 %	33.9 %	100 %

and the core becomes non-empty already for a multiplier of 3. For the other variants, we only report the results for multiplier values 0, 1, 3, and 5. As expected, the results are very similar for the case of the LRG-L1 and the LRG-L2. However, we do not observe similar behavior in the LRG-C1 and the LRG-C2.

2.6.4 Impact of the vehicle cost

The cost of using a vehicle, a in model (2.6)–(2.14), stands for not only the usage of the vehicle as it might suggest. All kinds of costs regarding the tours, but not dependent on the distance traveled, should be included. This covers loading and unloading costs, driver-related costs, vehicle maintenance costs, and so on. The cost of using a vehicle might hence differ substantially across applications. Therefore, it is worthwhile to investigate how the properties of the location-routing games are affected by the magnitude of this cost.

For the LRG, when we compare instances resulting with an empty core with those resulting with a non-empty core, we can see a substantial difference between the average costs of using a vehicle. This average is 164 for the empty-core instances and 104 for the non-empty-core instances. We investigate this issue further by generating new sets of instances in the same way as for the facility costs analysis. Now, the new sets of instances differ only in the cost of using a vehicle. The results are provided in Table 2.4. For all the variants, we can notice that the

Table 2.4: Impact of the vehicle cost on properties of the location-routing games

	Vehicle cost multiplier	Subadditive	Convex	Non-empty core
LRG	0	100 %	28.1 %	100 %
	1	100 %	30.5 %	99.3 %
	2	100 %	30.7 %	97.3 %
	3	100 %	30.8 %	96.4 %
	4	100 %	30.7 %	95.8 %
	5	100 %	30.7 %	95.7 %
LRG-C1	0	97.1 %	21.4 %	92.7 %
	1	97.3 %	22.7 %	92.6 %
	3	97.3 %	23.8 %	90.8 %
	5	97.3 %	24.1 %	90.4 %
LRG-L1	0	99.7 %	27.6 %	99.8 %
	1	99.8 %	30.3 %	99.1 %
	3	99.8 %	30.6 %	96.2 %
	5	99.8 %	30.5 %	95.6 %
LRG-C2	0	100 %	23.6 %	99.9 %
	1	100 %	26.1 %	99.7 %
	3	100 %	26.4 %	97.0 %
	5	100 %	26.7 %	96.1 %
LRG-L2	0	100 %	27.9 %	100 %
	1	100 %	30.4 %	99.3 %
	3	100 %	30.7 %	96.3 %
	5	100 %	30.6 %	95.7 %

proportion of instances with a non-empty core decreases as the cost of using a vehicle increases. This indeed supports the observation of different average costs of using a vehicle.

2.6.5 Cost allocations

Regardless of the core emptiness or non-emptiness, in practice, a unique cost allocation often needs to be specified to assess the contribution of different cooperating parties. With a focus on collaborative transportation, Guajardo and Rönnqvist (2016) recognized the Shapley value (Shapley, 1953), the nucleolus (Schmeidler, 1969) and proportional methods to be some of the most preferred allocation methods used in literature.

Here, we investigate whether the allocation methods result in allocations that are rational, that is, satisfy constraints (2.5). This relates to suitability of different allocation methods for the location-routing game. Besides the Shapley value and the nucleolus, we investigate the lexicographic equal profit method known as EPML (Frisk et al., 2010, Dahlberg et al., 2017), and two proportional methods, the first of them proportional to the stand-alone costs of each shipper and the second one to the total demand of each shipper's customers. All these allocation methods by definition satisfy the efficiency condition (2.4). Additional satisfaction of the rationality conditions hence means that the respective cost allocation belongs to the core. Such an analysis is therefore meaningless for the instances with an empty core in which all allocations would end up as non-rational. We investigate how often the particular allocation methods satisfy the rationality conditions in the instances with a non-empty core. The results are presented in Table 2.5.

Table 2.5: Rationality satisfaction by various cost allocation methods

	Shapley value	Nucleolus	EPML	Cost proportional method	Demand proportional method
LRG	97.0 %	100 %	100 %	79.7 %	67.7 %
LRG-C1	84.4 %	100 %	100 %	60.5 %	53.1 %
LRG-L1	96.9 %	100 %	100 %	79.3 %	67.4 %
LRG-C2	95.9 %	100 %	100 %	83.7 %	80.3 %
LRG-L2	97.0 %	100 %	100 %	79.7 %	67.7 %

The nucleolus and the EPML excel in the rationality satisfaction, which is not surprising as they are defined to belong to the core if it is not empty. In a huge majority of the instances, the rationality is also satisfied by the Shapley value. The proportional methods do not perform that well. Nevertheless, in the computation of the Shapley value, the nucleolus, and the EPML, the cost associated with each possible coalition needs to be determined. This becomes an obstacle when the number of shippers increases as the number of coalitions grows exponentially. The proportional methods, on the other hand, do not face this problem and might therefore be a preferred option. In case of seeking a proportional method, from Table 2.5 we can suggest the allocation proportional to the stand-alone costs to be the one to choose. In fact, we observe that the cost proportional method performs better in all the location-routing game variants.

As already mentioned, in the instances with an empty core, no allocation method can satisfy all the rationality conditions. To propose a similar allocation quality measure, a modification of the strong ε-core model (2.26)–(2.30) becomes useful. If, for a given allocation $(\pi_1, \ldots, \pi_{|N|})$, the model is formulated as

$$\min_{\varepsilon} \varepsilon \tag{2.31}$$

$$\text{s.t.} \quad \sum_{j \in S} \pi_j \leq C(S) + \varepsilon \qquad \forall S \subseteq N, \tag{2.32}$$

$$\varepsilon \in \mathbb{R}, \tag{2.33}$$

the optimal solution of ε represents the maximal violation of the rationality constraints (2.5). Table 2.6 shows the average maximal violation as a percentage of the respective $C(N)$ for various cost allocation methods (the EPML is not included because it is defined only for games with a non-empty core). The results support the previous findings with the nucleolus having the lowest average maximal violation of the rationality constraints. This, ranging from 0.9 % to 1.7 %, outperforms the Shapley value, the cost proportional method, and the demand proportional method in all the location-routing game variants. Among the proportional methods, the cost proportional method again performs better in all the variants.

Table 2.6: Average maximal violation of rationality constraints by various cost allocation methods in instances with an empty core

	Shapley value	Nucleolus	Cost proportional method	Demand proportional method
LRG	2.5 %	1.1 %	5.3 %	7.3 %
LRG-C1	3.5 %	1.7 %	5.5 %	7.7 %
LRG-L1	2.7 %	1.2 %	5.6 %	7.3 %
LRG-C2	3.2 %	0.9 %	5.3 %	7.3 %
LRG-L2	2.5 %	1.1 %	5.3 %	7.3 %

2.6.6 Savings

From the results of the LRG, it can be seen that the savings of the shippers can be substantial, as shown in Figure 2.4. The histogram shows percentage savings when the grand coalition is

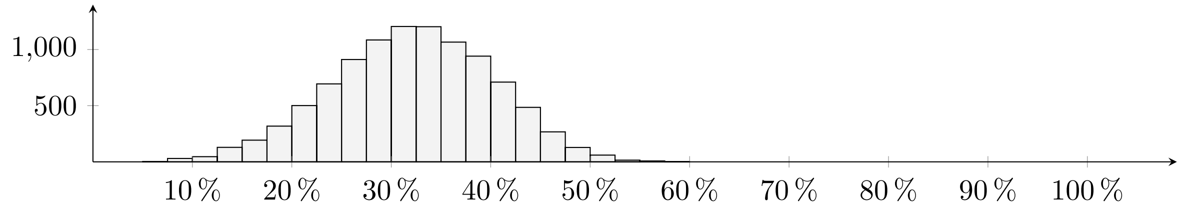


Figure 2.4: Histogram of percentage savings of the grand coalition with respect to the sum of the stand-alone costs

formed as opposed to a non-cooperative case represented as a sum of the stand-alone costs. These cost savings range from 6 % to 62 % with an average of 32 %. The main source of the savings is that less facilities are often needed when the grand coalition forms. The facility opening costs are on average reduced by 63 %.

Although it might seem counter-intuitive, we observe 14 % of instances where the total routing cost increases in comparison to the non-cooperative solution as seen in Figure 2.5. The intuition for this is that, since collaboration aims at reducing the overall costs (which include routing costs, facility opening costs and costs of using vehicles), it is possible that in some situations the selection of a reduced number of facilities might offset the increase in costs due to larger distance traveled. On average, however, the routing costs are reduced by 16 %.

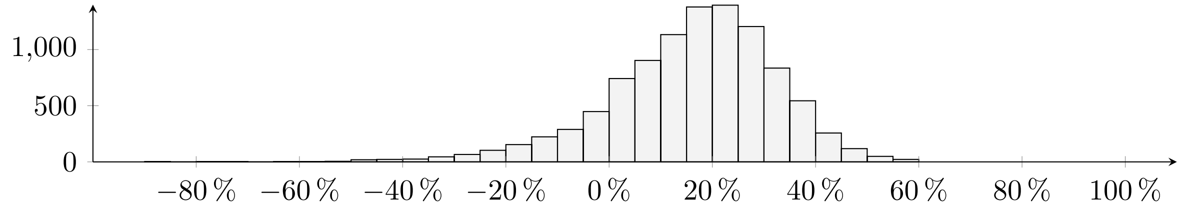


Figure 2.5: Histogram of percentage difference of the grand coalition's routing cost with respect to the sum of the stand-alone routing costs

The transportation is commonly recognized as one of the main contributors of CO_2 emissions (Ballot and Fontane, 2010). The increase in the traveled distance might thus lead to negative environmental effects. However, as this is accompanied by a change in facility selection, the emissions from transportation might be outbalanced by a reduction in emissions produced by the facilities due to a lower number of them needed to be open. Due to the wide range of applications of the location-routing problems, it is difficult to draw general conclusions. In fact, most of literature on location-routing problems only considers emissions stemming from transportation (Koç et al., 2016, Toro et al., 2017), whereas consideration of emissions produced by the facilities is more recent (Palacio et al., 2018).

2.6.7 Other instances

In addition to the results obtained for the experiment above, we also conduct numerical computations using some instances from the literature on location-routing problems. In order to find optimal solutions to a large enough number of instances, we have selected three relatively small instances. The first two, coord20-5-1 and coord20-5-2, have been used by Prins et al. (2006a,b, 2007) and are available from Prodhon (2010). Both of these instances contain 20 customers and 5 feasible sites of candidate facilities and the latter consists of customers located within two clusters. The third instance, Gaskell67-21x5, comes from a Ph.D. thesis by Barreto (2004) and is available from Barreto (nd). This instance contains 21 customers and 5 feasible sites of candidate facilities.

Since these instances have been designed for the location-routing problem assuming a single shipper, we use them as a basis to generate new instances with three, four and five shippers with randomly assigned customers. For example, for the instance coord20-5-1, we assume three shippers and generate 100 new instances by randomly assigning the customers such that the first and second shipper serve 7 customers each and the third shipper serves the remaining 6 of them. For an overview of all cases, see Table 2.7 where each row represents 100 generated instances. The mentioned example corresponds to the row for the instance coord20-5-1 with customers per shipper being 7, 7, 6. The row for the instance Gaskell67-21x5 with customers per shipper being 5, 4, 4, 4, 4 would then stand for five shippers with 5, 4, 4, 4, and 4 customers respectively. To cover a wide range of cases, some of the instances also assume shippers with substantially more or substantially less customers than the other shippers.

To find the optimal solutions to each of the instances, we use the following two-step exact approach. First, for each of the original instances, based on the customers' demands and the vehicle capacity, we generate all feasible tours and, with the Concorde TSP Solver (Applegate et al., 2007), we compute their costs. Second, for each of the new instances, we formulate a set partitioning model similar to the one used by Akca et al. (2008) and find the characteristic function values using Python/CPLEX 12.8.0.0.

For the LRG, the results on satisfaction of the properties of subadditivity, convexity and emptiness of the core are provided in Table 2.7. All 2,100 instances lead to a non-empty core. Moreover, the results on convexity instances suggest that there might be an effect of the number of shippers on the convexity satisfaction. In fact, as the number of shippers increases from 3 to 5, the share of instances satisfying the convexity decreases to 0 %.

Table 2.7: Satisfaction of properties in LRG and LRG-C1 for selected instances from literature

		LRG			LRG-C1		
	Customers per shipper	Subadditive	Convex	Non-empty core	Subadditive	Convex	Non-empty core
coord20-5-1	7, 7, 6	100 %	33 %	100 %	100 %	38 %	100 %
	8, 8, 4	100 %	80 %	100 %	100 %	77 %	100 %
	12, 4, 4	100 %	75 %	100 %	100 %	0 %	94 %
	5, 5, 5, 5	100 %	6 %	100 %	98 %	0 %	100 %
	6, 6, 6, 2	100 %	0 %	100 %	100 %	0 %	100 %
	8, 4, 4, 4	100 %	9 %	100 %	99 %	0 %	90 %
	4, 4, 4, 4, 4	100 %	0 %	100 %	96 %	0 %	16 %
coord20-5-2	7, 7, 6	100 %	73 %	100 %	100 %	1 %	100 %
	8, 8, 4	100 %	65 %	100 %	100 %	3 %	100 %
	12, 4, 4	100 %	38 %	100 %	100 %	0 %	64 %
	5, 5, 5, 5	100 %	0 %	100 %	100 %	0 %	48 %
	6, 6, 6, 2	100 %	0 %	100 %	100 %	0 %	69 %
	8, 4, 4, 4	100 %	0 %	100 %	100 %	0 %	21 %
	4, 4, 4, 4, 4	100 %	0 %	100 %	100 %	0 %	0 %
Gaskell67-21x5	7, 7, 7	100 %	61 %	100 %	100 %	71 %	100 %
	9, 9, 3	100 %	54 %	100 %	100 %	54 %	100 %
	13, 4, 4	100 %	56 %	100 %	100 %	63 %	100 %
	6, 5, 5, 5	100 %	0 %	100 %	100 %	0 %	100 %
	6, 6, 6, 3	100 %	0 %	100 %	100 %	0 %	100 %
	9, 4, 4, 4	100 %	0 %	100 %	100 %	1 %	100 %
	5, 4, 4, 4, 4	100 %	0 %	100 %	100 %	0 %	100 %

For the LRG-C1, we set the facility capacities w_g equal to the capacities in the original instances. The results in Table 2.7 show that in contrast to the LRG, there exist few instances where the subadditivity property is violated. Satisfaction of the convexity property seems to follow a similar trend as in the LRG. In the LRG-C1, also the core non-emptiness seems to follow this trend for the instance coord20-5-2 and partly for the instance coord20-5-1. Nevertheless, since this is not consistent over all of the instances, it is difficult to draw a general conclusion.

To define values of the partial facility capacities in the LRG-C2, we have followed two different approaches. In the first, for each shipper $n \in N$, we set the partial facility capacity w_{gn} equal to the capacity of facility $g \in G$ in the original instance. In the second approach, we divide the capacity in the original instances among the partial capacities proportionally to the numbers of customers served by the shippers. For example, if a capacity of facility $g \in G$ equals 50 in the original instance, for a shipper $n \in N$ serving 4 out of 20 customers, the partial capacity w_{gn} equals 50 with the first approach and 10 with the second approach. It should be noted that, since the demand of each customer needs to be satisfied from only one facility, the second approach might lead to infeasible solutions in computation of the characteristic function for some coalitions. In fact, out of the total of 2,100 instances, 305 result in an infeasibility. In Table 2.8, we denote the first approach by LRG-C2a and the second by LRG-C2b. The results marked with * are based on less than 100 instances as we have removed those with infeasible solutions. All instances produce a non-empty core and, again, show a declining trend in the

Table 2.8: Satisfaction of properties in LRG-C2a and LRG-C2b for selected instances from literature

	Customers per shipper	LRG-C2a Subadditive	LRG-C2a Convex	LRG-C2a Non-empty core	LRG-C2b Subadditive	LRG-C2b Convex	LRG-C2b Non-empty core
coord20-5-1	7, 7, 6	100 %	33 %	100 %	100 %	0 %	100 %
	8, 8, 4	100 %	80 %	100 %	100 %	3 %	100 %
	12, 4, 4	100 %	0 %	100 %	100 %	3 %	100 %
	5, 5, 5, 5	100 %	6 %	100 %	100 %	0 %	100 %
	6, 6, 6, 2	100 %	0 %	100 %	100 %*	0 %*	100 %*
	8, 4, 4, 4	100 %	9 %	100 %	100 %	0 %	100 %
	4, 4, 4, 4, 4	100 %	0 %	100 %	100 %	0 %	100 %
coord20-5-2	7, 7, 6	100 %	73 %	100 %	100 %	6 %	100 %
	8, 8, 4	100 %	65 %	100 %	100 %*	11 %*	100 %*
	12, 4, 4	100 %	2 %	100 %	100 %*	13 %*	100 %*
	5, 5, 5, 5	100 %	0 %	100 %	100 %*	0 %*	100 %*
	6, 6, 6, 2	100 %	0 %	100 %	100 %*	0 %*	100 %*
	8, 4, 4, 4	100 %	0 %	100 %	100 %*	0 %*	100 %*
	4, 4, 4, 4, 4	100 %	0 %	100 %	100 %*	0 %*	100 %*
Gaskell67-21x5	7, 7, 7	100 %	61 %	100 %	100 %	19 %	100 %
	9, 9, 3	100 %	54 %	100 %	100 %*	24 %*	100 %*
	13, 4, 4	100 %	55 %	100 %	100 %	23 %	100 %
	6, 5, 5, 5	100 %	0 %	100 %	100 %	0 %	100 %
	6, 6, 6, 3	100 %	0 %	100 %	100 %*	0 %*	100 %*
	9, 4, 4, 4	100 %	0 %	100 %	100 %	1 %	100 %
	5, 4, 4, 4, 4	100 %	0 %	100 %	100 %	0 %	100 %

convexity satisfaction.

In the LRG, in the optimal solutions for all coalitions of all instances generated from coord20-5-1 and coord20-5-2, we observe that only one facility is to be opened. Therefore, in the case of the LRG-L1 and the LRG-L2, there would be no change in the solutions by introducing a positive integer limit on the number of facilities and we skip these instances. For the case of Gaskell67-21x5, the highest number of opened facilities is 2. To observe any changes in the optimal solutions, we set the limit on number of facilities in the LRG-L1 as well as all the partial limits on number of facilities in the LRG-L2 equal to 1. The results provided in Table 2.9 show only few instances violating the subadditivity and even less instances with an empty core in the LRG-L1. Again, we observe the same trend in the convexity satisfaction.

Overall, the results to a large extent confirm the findings from the numerical experiment in Section 2.6.1. Additionally, we do not observe any effect when one of the shippers serves substantially more or substantially less customers than the other shippers. Also, we do not observe any effect of customers being clustered. However, since we only included a case with two clusters, the results might not be representative enough. This should certainly be tested on larger instances with more clusters for which exact methods might not suffice anymore. The same holds for the drop in the number of instances with a non-empty core in the LRG-C1 which has not been observed for all of the instances. This might, nevertheless, reflect the somewhat

Table 2.9: Satisfaction of properties in LRG-L1 and LRG-L2 for the Gaskell67-21x5 instance

		LRG-L1			LRG-L2		
	Customers per shipper	Subadditive	Convex	Non-empty core	Subadditive	Convex	Non-empty core
Gaskell67-21x5	7, 7, 7	99 %	38 %	100 %	100 %	57 %	100 %
	9, 9, 3	99 %	28 %	100 %	100 %	46 %	100 %
	13, 4, 4	100 %	34 %	100 %	100 %	42 %	100 %
	6, 5, 5, 5	99 %	0 %	99 %	100 %	0 %	100 %
	6, 6, 6, 3	99 %	0 %	99 %	100 %	0 %	100 %
	9, 4, 4, 4	99 %	0 %	100 %	100 %	0 %	100 %
	5, 4, 4, 4, 4	100 %	0 %	100 %	100 %	0 %	100 %

lower number of instances with a non-empty core in Table 2.2.

2.7 Concluding remarks

Horizontal cooperation is receiving more and more attention across transportation and logistics processes. The location-routing problem is no exception with companies cooperating on both locating their facilities and serving their customers. While there exists evidence of successful cooperation in practice (Eyers, 2010, Paddeu, 2017), the literature lacks general assessment of benefits coming from horizontal cooperation in location-routing problems.

In this article, we have introduced the standard location-routing game, a collaborative formulation of the standard location-routing problem. For both the capacitated location-routing problem and the location-routing problem with limited number of facilities, we have defined two alternative formulations of their collaborative versions. In three of these problems, we have shown the subadditivity property to hold in general. However, in the other two, the subadditivity is not always satisfied. Moreover, none of the problems guarantees the convexity or a non-empty core. Nevertheless, for the standard location-routing game and the location-routing game with limited number of facilities, we have shown that, when the facility opening costs are substantially larger than the traveling costs and the costs of using vehicles, the core is guaranteed to be non-empty.

Although it is not possible to guarantee core allocations in general, with a numerical experiment, we have shown that the core allocations exist in a huge majority of our instances. This has also been confirmed by using selected instances from the location-routing literature. The numerical results have also supported the findings of the effect of facility opening costs on the core emptiness. As the facility opening costs increase, the likelihood of a non-empty core increases. On the other hand, with the costs of using vehicles, we have observed the opposite effect. As the vehicle costs increase, the likelihood of a non-empty core decreases. It is important to note that the core emptiness does not necessarily outrule the cooperation. Often, regardless of the emptiness or non-emptiness of the core, it is preferred to pursue the cooperation and choose a unique cost allocation. We have tested the performance of various cost allocation methods. With respect to the stability, the results have shown dominance of the nucleolus and the lexicographical equal profit method. The latter is however not defined for the cases with an empty core and cannot thus be used under any circumstances. The Shapley value has shown a fairly good performance as well, followed by the cost proportional method and the demand proportional method.

The focus of this article was on exploration of the properties of location-routing games. While the numerical experiment was conducted in small instances that can be solved to optimality, an interesting avenue for future research is to explore location-routing games where the approximate solutions, instead of the optimal ones, are taken into account. In fact, large-scale instances often occurring in practice are commonly solved with heuristic approaches (Schneider and Drexl, 2017). It is worthwhile to investigate whether different solution approaches preserve the same properties.

Although most of the research on collaborative logistics focuses on cost reduction, literature has also reported other benefits such as shorter delivery times (Yang et al., 2016) and reduction of greenhouse gas emissions (Pérez-Bernabeu et al., 2015, Guajardo, 2018). Studying these features in collaborative location-routing problems is also a relevant direction for future research.

References

Adewumi, A. O. and Adeleke, O. J. (2018). A survey of recent advances in vehicle routing problems. *International Journal of System Assurance Engineering and Management*, 9(1):155–172.

Akca, Z., Berger, R. T., and Ralphs, T. K. (2008). Modeling and Solving Location, Routing, and Scheduling Problems. Technical report, COR@L Lab. Report 08T-009, Lehigh University. Available at http://coral.ie.lehigh.edu/~ted/files/papers/LRSP08.pdf [Accessed 30 March 2019].

Ambrosino, D., Sciomachen, A., and Scutellà, M. G. (2009). A heuristic based on multi-exchange techniques for a regional fleet assignment location-routing problem. *Computers & Operations Research*, 36(2):442–460.

Applegate, D. L., Bixby, R. E., Chvátal, V., and Cook, W. J. (2007). *The Traveling Salesman Problem: A Computational Study*. Princeton University Press, Princeton.

Audy, J.-F., Lehoux, N., D'Amours, S., and Rönnqvist, M. (2012). A framework for an efficient implementation of logistics collaborations. *International Transactions in Operational Research*, 19(5):633–657.

Ballot, E. and Fontane, F. (2010). Reducing transportation co2 emissions through pooling of supply networks: perspectives from a case study in french retail chains. *Production Planning & Control*, 21(6):640–650.

Barreto, S. S. (2004). *Análise e Modelização de Problemas de localização-distribuição*. PhD thesis, University of Aveiro.

Barreto, S. S. (n.d.). Location-routing problems (LRP). [online] Available at http://sweet.ua.pt/sbarreto/_private/SergioBarretoHomePage.htm [Accessed 28 June 2018].

Basso, F., D'Amours, S., Rönnqvist, M., and Weintraub, A. (2019). A survey on obstacles and difficulties of practical implementation of horizontal collaboration in logistics. *International Transactions in Operational Research*, 26(3):775–793.

Bruns, A., Klose, A., and Stähly, P. (2000). Restructuring of swiss parcel delivery services. *OR Spectrum*, 22(2):285–302.

Caballero, R., González, M., Guerrero, F. M., Molina, J., and Paralera, C. (2007). Solving a multiobjective location routing problem with a metaheuristic based on tabu search. application to a real case in andalusia. *European Journal of Operational Research*, 177(3):1751–1763.

Cleophas, C., Cottrill, C., Ehmke, J. F., and Tierney, K. (2019). Collaborative urban transportation: Recent advances in theory and practice. *European Journal of Operational Research*, 273(3):801–816.

Cruijssen, F., Cools, M., and Dullaert, W. (2007a). Horizontal cooperation in logistics: Opportunities and impediments. *Transportation Research Part E: Logistics and Transportation Review*, 43(2):129–142.

Cruijssen, F., Dullaert, W., and Fleuren, H. (2007b). Horizontal cooperation in transport and logistics: A literature review. *Transportation Journal*, 46(3):22–39.

Dahlberg, J., Göthe-Lundgren, M., and Engevall, S. (2017). A note on the nonuniqueness of the equal profit method. *Applied Mathematics and Computation*, 308:84–89.

Drexl, M. and Schneider, M. (2015). A survey of variants and extensions of the location-routing problem. *European Journal of Operational Research*, 241(2):283–308.

European Union (2001). Commission notice: Guidelines on the applicability of Article 81 of the EC Treaty to horizontal cooperation agreements (2001/C 3/02).

Eyers, N. (2010). 'Pooling': a successful collaborative approach. Presented at *Horizontal Collaboration in the Supply Chain Summit*, Brussels, June 2010.

Frisk, M., Göthe-Lundgren, M., Jörnsten, K., and Rönnqvist, M. (2010). Cost allocation in collaborative forest transportation. *European Journal of Operational Research*, 205(2):448–458.

Gansterer, M. and Hartl, R. F. (2018). Collaborative vehicle routing: a survey. *European Journal of Operational Research*, 268(1):1–12.

Goemans, M. X. and Skutella, M. (2004). Cooperative facility location games. *Journal of Algorithms*, 50(2):194–214.

Göthe-Lundgren, M., Jörnsten, K., and Värbrand, P. (1996). On the nucleolus of the basic vehicle routing game. *Mathematical Programming*, 72(1):83–100.

Guajardo, M. (2018). Environmental benefits of collaboration and allocation of emissions in road freight transportation. In Zeimpekis, V., Aktas, E., Bourlakis, M., and Minis, I., editors, *Sustainable Freight Transport: Theory, Models, and Case Studies*, pages 79–98. Springer, Cham.

Guajardo, M. and Rönnqvist, M. (2016). A review on cost allocation methods in collaborative transportation. *International Transactions in Operational Research*, 23(3):371–392.

Gunnarsson, H., Rönnqvist, M., and Carlsson, D. (2006). A combined terminal location and ship routing problem. *The Journal of the Operational Research Society*, 57(8):928–938.

Koç, Ç., Bektaş, T., Jabali, O., and Laporte, G. (2016). The impact of depot location, fleet composition and routing on emissions in city logistics. *Transportation Research Part B: Methodological*, 84:81–102.

Laporte, G., Nickel, S., and Saldanha da Gama, F., editors (2015). *Location Science*. Springer International Publishing, Berlin.

Marinakis, Y. and Marinaki, M. (2008). A bilevel genetic algorithm for a real life location routing problem. *International Journal of Logistics Research and Applications*, 11(1):49–65.

Nagy, G. and Salhi, S. (2007). Location-routing: Issues, models and methods. *European Journal of Operational Research*, 177(2):649–672.

Nambiar, J. M., Gelders, L. F., and van Wassenhove, L. N. (1981). A large scale location-allocation problem in the natural rubber industry. *European Journal of Operational Research*, 6(2):183–189.

Nambiar, J. M., Gelders, L. F., and van Wassenhove, L. N. (1989). Plant location and vehicle routing in the malaysian rubber smallholder sector: A case study. *European Journal of Operational Research*, 38(1):14–26.

Or, I. and Pierskalla, W. P. (1979). A transportation location-allocation model for regional blood banking. *A I I E Transactions*, 11(2):86–95.

Ouhader, H. and Kyal, M. E. (2017). Combining facility location and routing decisions in sustainable urban freight distribution under horizontal collaboration: How can shippers be benefited? *Mathematical Problems in Engineering*, 2017:18 pages. Article ID 8687515.

Paddeu, D. (2017). The Bristol-Bath urban freight consolidation centre from the perspective of its users. *Case Studies on Transport Policy*, 5(3):483–491.

Palacio, A., Adenso-Díaz, B., and Lozano, S. (2018). Analysing the factors that influence the Pareto frontier of a bi-objective supply chain design problem. *International Transactions in Operational Research*, 25(6):1717–1738.

Pérez-Bernabeu, E., Juan, A. A., Faulin, J., and Barrios, B. B. (2015). Horizontal cooperation in road transportation: a case illustrating savings in distances and greenhouse gas emissions. *International Transactions in Operational Research*, 22(3):585–606.

Perl, J. (1987). The multi-depot routing allocation problem. *American Journal of Mathematical and Management Sciences*, 7(1–2):7–34.

Potters, J. A. M., Curiel, I. J., and Tijs, S. H. (1992). Traveling salesman games. *Mathematical Programming*, 53(1–3):199–211.

Prins, C., Prodhon, C., and Calvo, R. W. (2006a). A memetic algorithm with population management (MAjPM) for the capacitated location-routing problem. In Gottlieb, J. and Raidl, G. R., editors, *Lecture Notes in Computer Science*, volume 3906, pages 183–194. Springer, Budapest.

Prins, C., Prodhon, C., and Calvo, R. W. (2006b). Solving the capacitated location-routing problem by a GRASP complemented by a learning process and a path relinking. *4OR*, 4(3):221–238.

Prins, C., Prodhon, C., Soriano, P., Ruiz, A., and Calvo, R. W. (2007). Solving the capacitated LRP by a cooperative Lagrangean relaxation-granular tabu search heuristic. *Transportation Science*, 41(4):470–483.

Prodhon, C. (2010). Classical instances for LRP. [online] Available at http://prodhonc.free.fr/Instances/instances_us.htm [Accessed 28 June 2018].

Prodhon, C. and Prins, C. (2014). A survey of recent research on location-routing problems. *European Journal of Operational Research*, 238(1):1–17.

Quintero-Araujo, C. L., Gruler, A., Juan, A. A., and Faulin, J. (2019). Using horizontal cooperation concepts in integrated routing and facility-location decisions. *International Transactions in Operational Research*, 26(2):551–576.

Salhi, S. and Rand, G. (1989). The effect of ignoring routes when locating depots. *European Journal of Operational Research*, 39(2):150–156.

Samet, D., Tauman, Y., and Zang, I. (1984). An application of the Aumann-Shapley prices for cost allocation in transportation problems. *Mathematics of Operations Research*, 9(1):25–42.

Sánchez-Soriano, J., López, M. A., and García-Jurado, I. (2001). On the core of transportation games. *Mathematical Social Sciences*, 41(2):215–225.

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics*, 17(6):1163–1170.

Schneider, M. and Drexl, M. (2017). A survey of the standard location-routing problem. *Annals of Operations Research*, 259(1–2):389–414.

Schrijver, A. (2003). *Combinatorial Optimization: Polyhedra and Efficiency*. Springer, Berlin.

Shapley, L. S. (1953). A value for n-person games. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games*, volume 2, pages 307–317. Princeton University Press, Princeton.

Shapley, L. S. (1955). Markets as cooperative games. *Rand Corporation Papers*, P-629:1–5.

Shapley, L. S. (1971). Cores of convex games. *International Journal of Game Theory*, 1(1):11–26.

Shapley, L. S. and Shubik, M. (1966). Quasi-cores in a monetary economy with nonconvex preferences. *Econometrica*, 34(4):805–827.

Sinha, A. (2016). The emerging "uberization" of dynamic on-demand warehousing for ecommerce. [online] Supply Chain 24/7. Available at http://www.supplychain247.com/article/the_uberization_of_dynamic_ondemand_warehousing [Accessed 3 August 2018].

Toro, E. M., Franco, J. F., Echeverri, M. G., and Guimarães, F. G. (2017). A multi-objective model for the green capacitated location-routing problem considering environmental impact. *Computers & Industrial Engineering*, 110:114–125.

van Oost, T. (2016). Cooperative game theory in location-routing. Master's thesis, Université Catholique de Louvain and NHH Norwegian School of Economics.

Watson-Gandy, C. D. T. and Dohrn, P. J. (1973). Depot location with van salesmen – a practical approach. *Omega*, 1(3):321–329.

Yang, F., Yang, M., Xia, Q., and Liang, L. (2016). Collaborative distribution between two logistics service providers. *International Transactions in Operational Research*, 23(6):1025–1050.

Zibaei, S., Hafezalkotob, A., and Ghashami, S. S. (2016). Cooperative vehicle routing problem: an opportunity for cost saving. *Journal of Industrial Engineering International*, 12(3):271–286.

Chapter 3

Fair travel distances in tournament schedules: A cooperative game theory approach

Ondrej Osicka[a], Mario Guajardo[a]

[a]Department of Business and Management Science, NHH Norwegian School of Economics, Bergen, Norway

Abstract

Generating fair schedules is an important aspect in the organization of sports competitions. The vast majority of the sports scheduling literature has focused on optimization problems where the performance of alternative solutions is measured by an overall goal aggregating all teams of the competition. For example, the most studied problem in the literature, so-called traveling tournament problem (TTP), aims at finding a schedule that minimizes the total distance traveled by the teams. While minimizing the expenditure resulting from all traveling between games is efficient from the overall objective perspective, it overlooks the actual distribution of the travel among the teams. In consequence, some teams may end up better than others with respect to their single goals, an imbalance which may largely affect teams' often limited resources as well as preparedness for the games. In this article, we adopt a cooperative game theory framework to deal with the question of fairness in sports scheduling. To obtain fair tournament schedules, we develop the following approach. First, the scheduling problem is reformulated as a transferable utility game. Second, by means of well-established cost allocation methods, such as the egalitarian method, Shapley value and nucleolus, an ideal distance distribution among the teams is determined. Third, given the inherently discrete nature of the space of feasible solutions to the scheduling problem, we introduce fairness measures to produce a schedule which approximately resembles the ideal distribution. We also discuss how to obtain a solution in case of not pursuing only fairness, but rather a compromise between fairness and minimum total distance. To illustrate the approach, we compute numerical results in one of the classic data instances of the TTP.

Keywords: Sports scheduling; Fairness; Cooperative game theory; Cost allocation; Traveling tournament problem

3.1 Introduction

Due to its importance in practice and the methodological challenges it involves, scheduling sport tournaments has received a lot of attention from researchers and practitioners. The application of operations research techniques and related fields has allowed for the incorporation of many criteria to improve the schedules of sport tournaments. These criteria often derive from logical and operational conditions (such as the number of matches between each pair of opponents and the number of time slots available to schedule them), and from other league-specific requirements. Then, these criteria are put together into an optimization problem, where some of the conditions are modeled as constraints and the others as part of the objective function (see Van Bulck et al. (2020) for a recent survey documenting typical constraints and objective functions in the literature). Whether it is a minimization or a maximization problem, the performance of alternative solutions is usually measured by an overall goal aggregating all teams (e.g. minimization of total travel costs or maximization of total revenues from attendance and TV). In consequence, some teams may end up better than others with respect to their single goals. It is therefore interesting to address how the benefits of the overall best solution could be fairly split among the different teams. Cooperative game theory offers a framework to deal with that question. This article focuses on the intersection of such cooperative game theory framework and fairness in sports scheduling, with particular emphasis on the traveling tournament problem (TTP).

Introduced in the seminal work by Easton et al. (2001), the TTP is the most studied problem in sports scheduling literature. The objective of the TTP is to find a schedule that minimizes the total traveled distances incurred by the teams, while satisfying some essential constraints of the tournament. While most of the related work on the TTP is methodological and focuses on solving it, there are also real-world applications reported in the literature, such as Bonomo et al. (2012) and Durán et al. (2019). Despite the numerous works on the TTP and other related problems, the discussion of fairness issues in sports scheduling has not received so much attention. Recently, Van Bulck and Goossens (2020) and Durán et al. (in press) have produced pioneer efforts, but focused on rather specific tournament designs where the fairness criteria are *ad hoc* to such designs.

In this article, we propose an approach to obtain a fair tournament schedule with respect to the travel distances of the teams and show that the optimal solution of the TTP is not necessarily the fairest solution. Furthermore, by analyzing the Pareto efficiency, we suggest a way to find a compromise solution between the schedule minimizing the total distance and the schedule maximizing fairness.

To the best of our knowledge, our article is the first one adopting a cooperative game theory framework to address fairness in sports scheduling. This is a well-established framework founded on general principles of fairness which have been studied in a broad range of applications. Also, by focusing on the flagship problem of the area, namely the TTP, our article brings the fairness discussion further to a common venue of research. Moreover, there is not much work on cooperative game theory in sports tournaments, where most of the game theory work has focused on its non-cooperative branch. This is perhaps because a sport tournament is essentially a competitive situation. However, we may argue that not even the top teams of the top leagues would be what they are without *smaller* teams taking part in the same tournaments. The success of the tournaments is subject to the participation of all teams, which also motivates the introduction of a cooperative game theory perspective. In this regard, recent contributions include for example Bergantiños and Moreno-Ternero (2020) focusing on sharing the revenues from broadcasting sport league events among participating teams.

In the next section, we present the TTP as well as some relevant concepts from cooperative game theory. In Section 3.3, we propose a methodology to find a tournament schedule by maximizing fairness which, in Section 3.4, we illustrate on the NL6 instance, one of the classic data instances of the TTP. Finally, Section 3.5 summarizes and concludes the article.

3.2 Preliminaries

In this section, we describe the traveling tournament problem (TTP) and present definitions of the relevant concepts from the cooperative game theory.

3.2.1 Traveling tournament problem

In some sports leagues, the travel associated with visiting other teams' venues plays an important role due to the associated financial burden, time spent on road or its plain inconvenience. In order to find schedules reducing the amount of travel, the TTP was introduced by Easton et al. (2001). Given n teams, n even, the teams' locations, and values L and U, $L \leq U$, the TTP generates a tournament schedule with $2(n-1)$ rounds where every pair of teams plays twice, once at home and once away for each team, the number of consecutive games at home and consecutive games away is between L and U inclusive, and the total distance traveled by the teams is minimized. The values of L and U are often set to 1 and 3, respectively.

A tournament in which every team faces each other team exactly twice, once at home and once away, is often referred to as a double round-robin tournament. Additionally, with n even and a total of $2(n-1)$ rounds, the solution of the TTP corresponds with a *compact* double round-robin tournament, i.e., the number of rounds is minimum and every team plays exactly once in every round. This is a common requirement in sports scheduling competitions (see e.g. Rasmussen and Trick (2008)). Besides, the TTP definition by Easton et al. (2001) introduces a *no-repeaters* condition.

We adopt the above specifications for the rest of this study. These summarize as follows:

(C1) Every team faces each other team exactly twice, once at home and once away for each team.

(C2) The number of rounds is minimum.

(C3) No team can play more than three consecutive games at home or three consecutive games away.

(C4) No two teams can play against each other in two consecutive rounds (no repeaters).[1]

In order to determine a schedule satisfying conditions (C1)–(C4) while minimizing the total distance, we use a model based on the *formulation with* $O(n^4)$ *variables* described by Melo et al.

[1]Note that this condition requires the number of teams to be greater than 2.

(2009). With the set of teams N, the set of rounds $R = \{1, \ldots, 2(n-1)\}$, and parameters $d_{i,j}$ stating the distance between venues of teams i and j for each combination $i, j \in N$, the model can be formulated as

$$\min \sum_{t\in N}\sum_{i\in N}\sum_{j\in N}\sum_{k\in R} d_{i,j}\, x_{t,i,j,k} + \sum_{t\in N}\sum_{i\in N}\sum_{j\in N} d_{j,t}\, x_{t,i,j,|R|} \tag{3.1}$$

$$\text{s.t.} \quad \sum_{\substack{j\in N\\ j\neq t}}\sum_{k\in R} x_{i,j,t,k} = 1 \qquad \forall t \in N, i \in N : i \neq t, \tag{3.2}$$

$$\sum_{i\in N}\sum_{j\in N} x_{t,i,j,k} = 1 \qquad \forall t \in N, k \in R : k > 1, \tag{3.3}$$

$$\sum_{j\in N} x_{t,t,j,1} = 1 \qquad \forall t \in N, \tag{3.4}$$

$$\sum_{j\in N} x_{t,i,j,k} = \sum_{j\in N} x_{t,j,i,k-1} \qquad \forall t \in N, i \in N, k \in R : k > 1, \tag{3.5}$$

$$\sum_{j\in N} x_{t,j,t,k} = \sum_{\substack{i\in N\\ i\neq t}}\sum_{\substack{j\in N\\ j\neq t}} x_{i,j,t,k} \qquad \forall t \in N, k \in R, \tag{3.6}$$

$$\sum_{j\in N}\sum_{\substack{i\in N\\ i\neq t}}\sum_{l=0}^{3} x_{t,j,i,k+l} \leq 3 \qquad \forall t \in N, k \in R : k \leq |R| - 3, \tag{3.7}$$

$$\sum_{\substack{j\in N\\ j\neq t}}\sum_{i\in N}\sum_{l=0}^{3} x_{j,i,t,k+l} \leq 3 \qquad \forall t \in N, k \in R : k \leq |R| - 3, \tag{3.8}$$

$$\sum_{j\in N}\sum_{l=0}^{1}(x_{t,j,i,k+l} + x_{i,j,t,k+l}) \leq 1 \qquad \forall t \in N, i \in N, k \in R : i \neq t, k < |R|, \tag{3.9}$$

$$x_{t,i,j,k} \in \{0, 1\} \qquad \forall t \in N, i \in N, j \in N, k \in R. \tag{3.10}$$

The binary variable $x_{t,i,j,k}$ equals 1 if team $t \in N$ travels from venue of team $i \in N$ to venue of team $j \in N$ in round $k \in R$. Hence, the objective function (3.1) minimizes the total distance traveled by all teams over all rounds (including the return home after the last game). Constraints (3.2), (3.3), (3.4), (3.5), and (3.6) ensure satisfaction of conditions (C1) and (C2) by prescribing each team to play at home against each other team once, play one game each round, start at home before the first round, travel in each round from the venue of the previous round's game, and meet another team for a game when staying home, respectively. Condition (C3) is enforced by constraints (3.7) and (3.8). For each team and for each sequence of four consecutive rounds, constraints (3.7) restrict the number of games away to be at most three. Constraints (3.8) work analogously for the home games. Constraints (3.9) make sure that no pair of teams plays against each other in two consecutive rounds, i.e., condition (C4). Lastly, (3.10) states the binary nature of the variables.

After solving model (3.1)–(3.10), the optimal objective value indicates the total distance traveled by the teams. The associated optimal schedule can be derived from the values of $x_{t,i,j,k}$.

3.2.2 Cooperative game theory

Let N denote the set of all players and $\mathscr{S}$ the set of all subsets of N. A transferable utility game (TU game) is a pair (N, v) where $v: \mathscr{S} \to \mathbb{R}$ is the characteristic function assigning to each coalition $S \in \mathscr{S}$ the optimal cost achievable by cooperation of players within this coalition. Depending on the context, the characteristic function may be defined to represent either costs or payoffs. Although some of the following definitions are universally applicable for both variants, we limit the attention only to the case of v describing costs.

TU games allow for side payments, and thus for redistribution of the incurred costs among cooperating players. One of the central questions in the cooperative game theory revolves around a fair distribution of the costs. Assuming all players from N to cooperate and thus incurring a total cost $v(N)$, it is often not trivial to decide how big share of the cost should each player bear. In practice, there is no consensus on what fairness actually means. In the same vein, the game-theoretic literature recognizes multiple cost allocation methods based on different views of fairness. Here, we describe three of the most commonly used, the egalitarian method, the Shapley value and the nucleolus.

Proportional methods assign each player $p \in N$ a share α_p of the cost $v(N)$, i.e.,

$$x_p = \alpha_p \cdot v(N) \qquad \forall p \in S \tag{3.11}$$

where $\sum_{p \in N} \alpha_p = 1$. These cost allocation methods are very straightforward and easy to communicate in practice.

Different strategies are employed in the literature as how to specify the shares α_p. The simplest option is the egalitarian method assigning to each player the same share of $v(N)$, i.e.,

$$x_p = \frac{v(N)}{|N|} \qquad \forall p \in S \tag{3.12}$$

where $|N|$ denotes the cardinality of N, i.e., the total number of players.

The Shapley value is another commonly used cost allocation method. Algaba et al. (2019) shows its importance across many different applications. To achieve a fair distribution, Shapley (1953) derives the share of $v(N)$ assigned to player $p \in N$ from this player's marginal contributions to any possible coalition as

$$x_p = \sum_{S \subseteq N:\, p \in S} \frac{(|S|-1)!\,(|N|-|S|)!}{|N|!} \cdot \left(v(S) - v(S \backslash \{p\})\right). \tag{3.13}$$

The nucleolus, as defined by Schmeidler (1969), takes a different approach to determine a fair distribution of $v(N)$. In this case, the shares are derived from excesses of the coalitions in order to minimize players' incentives to leave N and form another coalition. For $N = \{p_1, \dots, p_{|N|}\}$ and a cost allocation $x = (x_{p_1}, \dots, x_{p_{|N|}})$, the excess of coalition $S \subseteq N$ at x is defined as $\varepsilon(x, S) = v(S) - \sum_{p \in S} x_p$. Furthermore, with $S_1, \dots, S_m$ denoting all subsets of N, $e(x) = (\varepsilon(x, S_1), \dots, \varepsilon(x, S_m))$ denotes the excess vector at x and $\theta(e(x))$ is a vector resulting from arranging the components of $e(x)$ in nondecreasing order. The nucleolus is then defined as the cost allocation x which lexicographically maximizes $\theta(e(x))$.

3.3 Fair scheduling in the TTP

The TTP generates a schedule guaranteeing minimum total distance traveled by the teams. This does not imply minimum distance traveled by each particular team. In fact, some teams might need to travel more if it keeps the collective distance low. It suggests a question: What a fair schedule considering the collective as well as all individual preferences would look like? In the following, we determine fair schedules by, first, finding an *ideal* distance distribution and, then, determining the *nearest* schedule.

To find the ideal distance distribution, we employ the cost allocations introduced in the previous section. As most of the cost allocations require the characteristic function, in Section 3.3.1, we define models to compute distances associated with each set of teams. With determined ideal distance distribution, we proceed by finding the nearest solution, a schedule in which the distances traveled by the teams resemble the distribution the most. Sections 3.3.2 and 3.3.3 focus on measuring of the resemblance and finding the nearest schedule, respectively. However, just as the distance-minimizing TTP does not guarantee a fair solution, the fairness maximizing solution does not guarantee the lowest distance. Therefore, in Section 3.3.4 we investigate the Pareto efficiency and discuss the choice of the best schedule.

One could argue that the cost allocations introduced in Section 3.2.2 are defined for transferable utility games while the utility in TTP, the distance traveled, is not transferable. A monetary value could be put on some aspects of the traveling distance which may or may not allow for this property. Nonetheless, some more or less important aspects such as affected players' recovery while on road clearly remains non-transferable.

The literature on non-transferable utility games (NTU games) covers several concepts dealing with payoff or cost allocation (McLean, 2002). However, they are often accompanied by assumptions not suitable for our setting. The standard definition of an NTU game with set of players N, as formulated for example by Hart (2004), requires feasible utility combinations (corresponding to the distance distributions) for each coalition $S \subseteq N$ to form a nonempty, strict, closed, convex and comprehensive subset of Euclidean space $\mathbb{R}^{|N|}$. In the TTP, this is often unachievable due to the finite number of feasible schedules. This makes the NTU-game-based approach unfitting and, therefore, we substitute it by formulating the two-step approach described above.

3.3.1 Characteristic function

With N standing for the set of all teams in a tournament, we can define a TU game (N, v). The characteristic function v assigns to each coalition $S \subseteq N$ the optimal traveling distance a tournament of teams within S could achieve. The computation of $v(S)$ varies depending on the properties of S.

First, we look at coalitions of two or less teams. For any $S \subseteq N$ such that $|S| = 0$ or $|S| = 1$, there are no games to be played. Hence, $v(S) = 0$. For any $S \subseteq N$ such that $|S| = 2$, exactly two games need to be played. The teams in S need to face each other at each team's venue. However, a tournament schedule satisfying conditions (C1)–(C4) cannot be constructed. With only two teams in a tournament, a feasible schedule needs to allow for the teams to play against each other in two consecutive rounds, i.e., violate condition (C4). Therefore, the value of $v(S)$ can be determined by replacing N in model (3.1)–(3.10) by S and removing constraints (3.9) corresponding with condition (C4).

For any $S \subseteq N$ such that $|S| > 2$ and $|S|$ is even, the value of $v(S)$ can be determined as the optimal objective value of model (3.1)–(3.10) with N replaced by S. For $S \subseteq N$ such that $|S| > 2$ and $|S|$ is odd, the computation of $v(S)$ is not as straightforward. With an odd number of teams, in each round there has to be at least one team unmatched with an opponent. We refer to this situation as a team having a rest or having a bye as described for example by Bao and Trick (2010) in the case of the relaxed traveling tournament problem. With teams having byes, not only the minimum number of rounds no longer equals $2(|S| - 1)$, but even with the correct number, model (3.1)–(3.10) cannot be applied. For instance, the combination of constraints (3.4) and (3.6) explicitly requires each team to play one game each round. The model thus needs to be modified.

In modeling of the byes, we assume that, when having a bye, the teams stay on road, i.e., the teams do not travel home during a round with no game. We still require the model to follow conditions (C1)–(C4), but to avoid long periods without a game we introduce an additional one as follows:

(C5) No team can have two or more consecutive byes.

To comply with condition (C2), the number of rounds needs to be minimum. With n teams in a tournament, each team plays $2(n - 1)$ games. This means a total of $n(n - 1)$ games in the tournament. With n even, a maximum number of games in any round equals $\frac{n}{2}$. With n odd, a maximum number of games in any round equals $\frac{n-1}{2}$. Therefore, the least number of rounds equals $2(n - 1)$ for n even and $2n$ for n odd. Hence, with an odd number of teams, each team needs to have a rest for at least two rounds.

Using a general formulation for any $S \subseteq N$, a model computing $v(S)$ while allowing for byes can be stated as

$$v(S) = \min \sum_{t\in S} \sum_{i\in S} \sum_{j\in S} \sum_{k\in R} d_{i,j}\, x_{t,i,j,k} + \sum_{t\in S} \sum_{i\in S} \sum_{j\in S} d_{j,t}\, x_{t,i,j,|R|} \tag{3.14}$$

$$\text{s.t.} \quad \sum_{\substack{j\in S\\ j\neq t}} \sum_{k\in R} x_{i,j,t,k} = 1 \qquad \forall t \in S, i \in S : i \neq t, \tag{3.15}$$

$$\sum_{i\in S} \sum_{j\in S} x_{t,i,j,k} = 1 \qquad \forall t \in S, k \in R : k > 1, \tag{3.16}$$

$$\sum_{j\in S} x_{t,t,j,1} = 1 \qquad \forall t \in S, \tag{3.17}$$

$$\sum_{j\in S} x_{t,i,j,k} = \sum_{j\in S} x_{t,j,i,k-1} \qquad \forall t \in S, i \in S, k \in R : k > 1, \tag{3.18}$$

$$\sum_{j\in S} x_{t,j,t,k} \geq \sum_{\substack{i\in S\\ i\neq t}} \sum_{\substack{j\in S\\ j\neq t}} x_{i,j,t,k} \qquad \forall t \in S, k \in R : |S| > 1, \tag{3.19}$$

$$\sum_{j\in S}\sum_{\substack{i\in S\\ i\neq t, i\neq j}}\sum_{\substack{m\in R\\ k\leq m<k+l}} x_{t,j,i,m} \leq 3+\sum_{\substack{j\in S\\ j\neq t}}\sum_{\substack{i\in S\\ i\neq t}}\sum_{\substack{m\in R\\ k\leq m<k+l}} M\, x_{j,i,t,m}$$
$$\forall t \in S, k \in R, l \in R: \quad k \leq |R|-3, 4 \leq l \leq |R|-k+1, \quad (3.20)$$

$$\sum_{\substack{j\in S\\ j\neq t}}\sum_{\substack{i\in S\\ i\neq t}}\sum_{\substack{m\in R\\ k\leq m<k+l}} x_{j,i,t,m} \leq 3+\sum_{j\in S}\sum_{\substack{i\in S\\ i\neq t, i\neq j}}\sum_{\substack{m\in R\\ k\leq m<k+l}} M\, x_{t,j,i,m}$$
$$\forall t \in S, k \in R, l \in R: \quad k \leq |R|-3, 4 \leq l \leq |R|-k+1, \quad (3.21)$$

$$\sum_{\substack{j\in S\\ j\neq i}}\sum_{l=0}^{1} x_{t,j,i,k+l} + \sum_{\substack{j\in S\\ j\neq t}}\sum_{l=0}^{1} x_{i,j,t,k+l} \leq 1$$
$$\forall t \in S, i \in S, k \in R: \quad i \neq t, k < |R|, |R| > 2, \quad (3.22)$$

$$\sum_{l=0}^{1} x_{t,i,i,k+l} + \sum_{\substack{j\in S\\ j\neq t}}\sum_{l=0}^{1} x_{t,j,t,k+l} - \sum_{\substack{j\in S\\ j\neq t}}\sum_{\substack{h\in S\\ h\neq t}}\sum_{l=0}^{1} x_{j,h,t,k+l} \leq 1$$
$$\forall t \in S, i \in S, k \in R: \quad k < |R|, |S| > 1, \quad (3.23)$$

$$x_{t,i,j,k} \in \{0,1\} \qquad \forall t \in S, i \in S, j \in S, k \in R, \quad (3.24)$$

where $R = \{1, \ldots, 2(n-1)\}$ for $|S|$ even or $R = \{1, \ldots, 2n\}$ for $|S|$ odd and M is a sufficiently large number, for example $M = |R|$.

The binary variables $x_{t,i,j,k}$ in model (3.14)–(3.24) work in the same way as in model (3.1)–(3.10). If team $t \in N$ travels from a venue of team $i \in N$ to play a game at a venue of team $j \in N$ in round $k \in R$, $x_{t,i,j,k}$ equals 1. Additionally, if team $t \in N$ is to have a rest after being at a venue of team $i \in N$, $x_{t,i,i,k}$ equals 1. The objective function (3.14) and constraints (3.15)–(3.18), (3.24) correspond with the objective function (3.1) and constraints (3.2)–(3.5), (3.10), respectively, with the only difference being in replacing the set of teams N by S. Constraints (3.19) state that a team may travel to another team's venue for a game only if the other team is going to be there as well. Constraints (3.20) and (3.21) guarantee compliance with condition (C3). If any sequence of at least four consecutive rounds includes no game at home, constraints (3.20) require the number of games away in this sequence to be at most three. Constraints (3.21) work analogously to restrict the number of consecutive games at home. Constraints (3.22) prevent teams from playing against the same opponent in two consecutive rounds, i.e., condition (C4). Additionally, condition $|R| > 2$ in constraints (3.22) avoids infeasibility of the model for tournaments of less than 3 teams. Lastly, for each team, constraints (3.23) enforce condition (C5) by restricting the number of rests in any two consecutive rounds to be at most one.

Model (3.14)–(3.24) can be solved for any coalition $S \subseteq N$ regardless of its cardinality or parity of this cardinality. With fully specified characteristic function v, it is then easy to determine cost allocations such as the egalitarian method, the Shapley value or the nucleolus. Since these allocations follow certain concepts of fairness, they may serve as the ideal distance distribution.

3.3.2 Measuring fairness

With the knowledge of the ideal distance distribution, we want to determine a schedule which resembles this distribution the most. Since we consider resemblance regarding only the distances traveled by the teams, we can represent each schedule by a vector of the distances. Hence, when we compare the ideal distance distribution with a schedule or a schedule with another schedule, we essentially compare two vectors of length $|N|$.

For a similar approach aiming for fairness in kidney exchange programmes, Biró et al. (2020) apply a metric based on the sum of absolute deviations. For vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, this equals

$$\sum_{j=1}^{n} |x_j - y_j|. \tag{3.25}$$

For example, comparing vectors $(100, 100, 100, 100)$ and $(90, 110, 90, 110)$ results in a value of 40. However, the same result can be obtained when comparing vectors $(100, 100, 100, 100)$ and $(100, 100, 100, 140)$. In our application, each element of the vectors stands for a different team. With the objective to achieve fairness, it is then only natural to prefer the first case and try to avoid the second one. Therefore, we do not consider this metric suitable for our purpose.

Perea and Puerto (2019) proposed a heuristic procedure for computing the nucleolus. To assess its quality and compare the obtained result with the actual nucleolus, they proposed two metrics, RD_a and RD_e. For a given allocation $x = (x_1, \ldots, x_n)$ and the nucleolus $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_n)$, they are defined as

$$RD_a = \frac{\sqrt{\sum_{j=1}^{n} (\tilde{x}_j - x_j)^2}}{\sqrt{\sum_{j=1}^{n} \tilde{x}_j^2}} \tag{3.26}$$

and

$$RD_e = \frac{\sqrt{\sum_{k=1}^{k_{\max}} (\theta(e(\tilde{x}))_k - \theta(e(x))_k)^2}}{\sqrt{\sum_{k=1}^{k_{\max}} \theta(e(\tilde{x}))_k^2}} \tag{3.27}$$

where $\theta(e(x))$ is the excess vector at x arranged in nondecreasing order as described in the definition of the nucleolus and $k_{\max}$ is the number of the largest excesses to account for.

Furthermore, we introduce a new metric by adapting RD_a for the use in fairness maximization. For the RD_a metric, the signs of the differences in $\sum_{j=1}^{n} (\tilde{x}_j - x_j)^2$ do not matter. However, due to the inclusion of the positive differences (corresponding with the prescribed distances lower than the ideal ones), using this metric for schedule comparison could turn out to be counterproductive. For illustration, let us assume a four-team tournament with ideal distance distribution $(150, 100, 200, 80)$ and only two feasible schedules prescribing distances

$(300, 80, 210, 75)$ and $(300, 95, 210, 75)$. In both solutions, the first, third, and fourth team face the same distances. On the other hand, the second team could benefit from the first solution even though the distance prescribed by the second solution is closer to the ideal one. Nevertheless, due to the closer distance in the second solution, this would be the preferred schedule from the perspective of the RD_a metric. Hence, pursuing the RD_a metric could make a particular team travel a longer distance even if no other team would benefit from it. This can be avoided by focusing only on the negative differences. Therefore, we propose metric RD_a^- as a modification of the RD_a metric accounting only for the negative differences:

$$RD_a^- = \frac{\sqrt{\sum_{j=1:\tilde{x}_j - x_j < 0}^{n} (\tilde{x}_j - x_j)^2}}{\sqrt{\sum_{j=1}^{n} \tilde{x}_j^2}} \tag{3.28}$$

3.3.3 Maximizing fairness

With a set of teams $N = \{t_1, \dots, t_{|N|}\}$, a fully specified characteristic function v and a chosen allocation method, it is easy to obtain the ideal distance distribution $\tilde{x} = (\tilde{x}_{t_1}, \dots, \tilde{x}_{t_{|N|}})$. Then, with a chosen metric of fairness, to find the nearest solution (schedule maximizing the fairness), one can find a schedule minimizing this metric.

With the RD_a metric, the nearest solution can be determined by a model with objective function

$$\min \frac{\sqrt{\sum\limits_{t \in N} \left(\tilde{x}_t - \sum\limits_{i \in S} \sum\limits_{j \in S} \sum\limits_{k \in R} d_{i,j}\, x_{t,i,j,k} - \sum\limits_{i \in S} \sum\limits_{j \in S} d_{j,t}\, x_{t,i,j,|R|} \right)^2}}{\sqrt{\sum\limits_{t \in N} \tilde{x}_t^2}} \tag{3.29}$$

while satisfying constraints (3.15)–(3.24) in which all occurrences of set S are replaced by N. We refer to this model as the RD_a-minimizing TTP.

Similarly, this means objective function

$$\min \frac{\sqrt{\sum\limits_{k=1}^{k_{\max}} \left(\theta(e(\tilde{x}))_k - \theta\left(e\left(\left(\sum\limits_{i \in S} \sum\limits_{j \in S} \sum\limits_{k \in R} d_{i,j}\, x_{t,i,j,k} + \sum\limits_{i \in S} \sum\limits_{j \in S} d_{j,t}\, x_{t,i,j,|R|} \right)_{t \in N} \right) \right)_k \right)^2}}{\sqrt{\sum\limits_{k=1}^{k_{\max}} \theta(e(\tilde{x}))_k^2}} \tag{3.30}$$

and the RD_e-minimizing TTP in case of the RD_e metric, and objective function

$$\min \frac{\sqrt{\sum_{t \in N : D_t < 0} D_t^2}}{\sqrt{\sum_{t \in N} \tilde{x}_t^2}} \tag{3.31}$$

where

$$D_t = \tilde{x}_t - \sum_{i \in S} \sum_{j \in S} \sum_{k \in R} d_{i,j}\, x_{t,i,j,k} - \sum_{i \in S} \sum_{j \in S} d_{j,t}\, x_{t,i,j,|R|} \qquad \forall t \in N \tag{3.32}$$

and the RD_a^--minimizing TTP in case of the RD_a^- metric.

In all three objective functions, one can notice that the denominators do not depend on the variables of the models and, thus, remain constant. Therefore, and because the square root function is a strictly increasing function, the objective functions (3.29), (3.30), (3.31) may be replaced by

$$\min \sum_{t \in N} \left(\tilde{x}_t - \sum_{i \in S} \sum_{j \in S} \sum_{k \in R} d_{i,j}\, x_{t,i,j,k} - \sum_{i \in S} \sum_{j \in S} d_{j,t}\, x_{t,i,j,|R|} \right)^2, \tag{3.33}$$

$$\min \sum_{k=1}^{k_{\max}} \left(\theta(e(\tilde{x}))_k - \theta \left(e \left(\left(\sum_{i \in S} \sum_{j \in S} \sum_{k \in R} d_{i,j}\, x_{t,i,j,k} + \sum_{i \in S} \sum_{j \in S} d_{j,t}\, x_{t,i,j,|R|} \right)_{t \in N} \right) \right)_k \right)^2, \tag{3.34}$$

$$\min \sum_{t \in N : D_t < 0} D_t^2, \tag{3.35}$$

respectively, with no change in the optimal solutions. Despite the simplification, all the objective functions remain nonlinear and nonlinear integer programming solvers are required to solve the respective models.

In the case of the RD_e-minimizing TTP, the nonlinearity stems not only from the second power, but also from the ordering function θ. In our attempts to find the nearest schedule using solvers Bonmin and Couenne, even for the four-team NL4 instance, we failed to obtain an optimal solution within 72 hours. On the other hand, the RD_a-minimizing TTP and the RD_a^--minimizing TTP both generate an integer quadratic programming problem solvable by solvers such as Gurobi or CPLEX. Therefore, in Section 3.4, we disregard the RD_e-minimizing TTP and focus only on the two remaining problems.

3.3.4 Pareto efficiency

The TTP selects a schedule associated with the shortest total distance. Maximizing fairness, on the other hand, results in a schedule minimizing a given metric. Although the solutions might coincide, it is not always the case. Since both criteria are often desirable, it is not trivial to decide which schedule is the best. To aid this decision, one may employ a method from multi-objective optimization called the Pareto front.

The RD_a-minimizing TTP, RD_e-minimizing TTP, RD_a^--minimizing TTP, and model (3.14)–(3.24) for $S = N$ contain the same set of constraints and differ only in their objective functions. Hence, their feasible regions are identical and, for a given ideal distance distribution, values of all the objective functions can be evaluated at each feasible solution.

As described for example by Hwang and Masud (1979), for a problem with two objective functions, a solution is considered Pareto efficient if none of the objective functions can be improved without worsening the other one. Pareto front is then a set of all Pareto efficient solutions.

For example, let us assume selection of the RD_a^- metric in a problem with 8 feasible schedules depicted in Figure 3.1. Only the solutions denoted by a full circle belong to the Pareto front. For each of the other solutions, there exists another solution improving at least one of the objectives

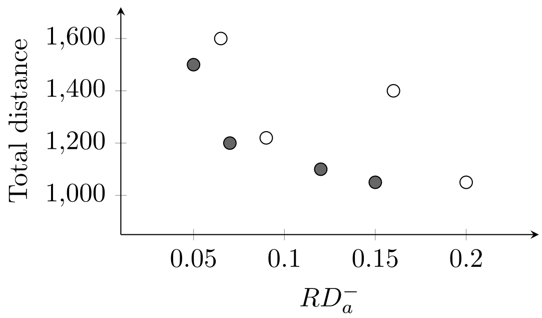


Figure 3.1: Example of a Pareto front

without worsening the other one. Therefore, it is reasonable to choose a schedule belonging to the Pareto front. The left-most solution on the Pareto front is the schedule minimizing RD_a^-, while the right-most solution corresponds with minimizing the total distance. The two remaining solutions offer some compromise between the two criteria.

The Pareto front may offer some options to choose from and help understand the trade-offs. Nonetheless, there is no general answer for what is the right choice from the schedules on the Pareto front. It very much depends on the tournament itself as well as on available resources, inclination towards fairness, etc.

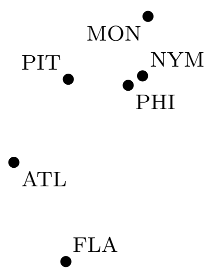
3.4 Numerical results for the NL6 instance

In this section, to illustrate the proposed methodology, we attempt to find a fair schedule in the NL6 instance by Easton et al. (2001). All our models were implemented in AMPL and solved with the Gurobi solver.

The NL6 instance contains six teams from the North American baseball league known as the National League. Figure 3.2 describes the teams and location of their venues. In the rest

Handle	Name	Location
ATL	Atlanta Braves	Atlanta, Georgia
NYM	New York Mets	New York, New York
PHI	Philadelphia Phillies	Philadelphia, Pennsylvania
MON	Montreal Expos	Montreal, Canada
FLA	Florida Marlins	Miami, Florida
PIT	Pittsburgh Pirates	Pittsburgh, Pennsylvania

(a) Names and locations



(b) Geographical distribution

Figure 3.2: Description of teams from NL6 instance

of this section, we refer to the teams only by their handles.

In Figure 3.2b, one can notice that teams NYM and PHI are based fairly close to each other. Teams FLA and ATL are on the other hand quite distanced from all other teams. The notion of fairness might vary across different leagues and tournaments. For instance, one could see team FLA in an unfortunate position given its location and seek balance in traveled distances by minimizing this team's distance on the expense of other teams. This viewpoint would correspond with a choice of the egalitarian method or a similar allocation. Nonetheless, one could also look at centrally located teams like NYM and PHI and claim that they should not be punished for other teams such as team FLA, i.e., the teams should travel in line with marginal distances they bring to the tournament. Such an approach would lead to a choice of the Shapley value, nucleolus or a similar allocation.

We proceed by solving the TTP model (3.1)–(3.10). The optimal schedule prescribes distances shown in the first part of Table 3.1. Afterwards, model (3.14)–(3.24) is solved for each $S \subseteq N$ in order to evaluate the characteristic function which is then used to compute the egalitarian method, Shapley value and nucleolus. These are listed in the second part of Table 3.1. The egalitarian method prescribes each team the same distance of 3,986. The Shapley value and the nucleolus, as expected from the geographical distribution of the teams' venues, prescribe the lowest distances to teams NYM and PHI with no big difference while the distance determined for team FLA is more than twice as large. To find schedules with distances similar

Table 3.1: NL6. DRR. Allocations.

	ATL	NYM	PHI	MON	FLA	PIT	Total distance
TTP	4,414	3,328	3,724	3,996	5,135	3,319	23,916
Egalitarian method	3,986	3,986	3,986	3,986	3,986	3,986	23,916
Shapley value	3,781.1	3,141.3	3,008.3	4,488.8	6,336	3,161.5	23,916
Nucleolus	4,176.8	2,938.5	2,893.5	4,117.5	6,715.3	3,074.5	23,916
RD_a-min. TTP (egal. m.)	4,147	3,857	4,174	4,090	4,953	4,058	25,279
RD_a^--min. TTP (egal. m.)	4,147	3,857	4,174	4,090	4,953	4,058	25,279
RD_a-min. TTP (Shapley v.)	4,263	3,472	3,372	4,424	6,407	3,597	25,535
RD_a^--min. TTP (Shapley v.)	4,147	3,365	3,372	4,537	5,596	3,604	24,621
RD_a-min. TTP (nucleolus)	4,551	3,509	3,200	3,994	6,455	3,218	24,927
RD_a^--min. TTP (nucleolus)	4,551	3,509	3,200	3,994	6,455	3,218	24,927

to the three computed allocations, we can now solve the RD_a-minimizing TTP and the RD_a^--minimizing TTP. The results are listed in the last part of Table 3.1. Although the schedules do not match the allocations exactly, they tend to be more similar to the respective allocations than the solution of the original TTP.

In Section 3.3.2, we have discussed the main advantage of using the RD_a^- metric over the RD_a metric. In this example, for the egalitarian method and the nucleolus, the models actually result in the same solutions for both of the metrics. This is not the case with the Shapley value. The distances determined by the models differ for all teams except for PHI. In case of ATL, NYM, MON, and PIT the distances are still fairly similar. The issue occurs in the case of team FLA. While the RD_a^- metric leads to a distance of 5,596, the RD_a metric determines a solution with a distance of 6,407. Although this is much closer to the value of 6,336 determined by the Shapley value, team FLA would certainly prefer the first option. It hence seems like an unnecessary increase in the distance for team FLA which is also reflected in the total distance. This supports the discussion in Section 3.3.2 and, therefore, for a detailed analysis of the results, we proceed only with the models based on the RD_a^- metric.

Figure 3.3 reports the optimal schedules obtained by solving the TTP and the RD_a^--

Team	Round 1	2	3	4	5	6	7	8	9	10	Distance
ATL	FLA	NYM	PIT	@PHI	@MON	@PIT	PHI	MON	@NYM	@FLA	4,414
NYM	@PIT	@ATL	@FLA	MON	FLA	@PHI	@MON	PIT	ATL	PHI	3,328
PHI	@MON	FLA	MON	ATL	@PIT	NYM	@ATL	@FLA	PIT	@NYM	3,724
MON	PHI	@PIT	@PHI	@NYM	ATL	FLA	NYM	@ATL	@FLA	PIT	3,996
FLA	@ATL	@PHI	NYM	PIT	@NYM	@MON	@PIT	PHI	MON	ATL	5,135
PIT	NYM	MON	@ATL	@FLA	PHI	ATL	FLA	@NYM	@PHI	@MON	3,319
										Total distance	23,916

(a) Optimal solution of the TTP

Team	Round 1	2	3	4	5	6	7	8	9	10	Distance
ATL	FLA	MON	@PHI	@NYM	@MON	PIT	PHI	NYM	@PIT	@FLA	4,147
NYM	PIT	PHI	@PIT	ATL	@PHI	MON	FLA	@ATL	@FLA	@MON	3,857
PHI	MON	@NYM	ATL	@FLA	NYM	FLA	@ATL	@PIT	@MON	PIT	4,174
MON	@PHI	@ATL	@FLA	PIT	ATL	@NYM	@PIT	FLA	PHI	NYM	4,090
FLA	@ATL	@PIT	MON	PHI	PIT	@PHI	@NYM	@MON	NYM	ATL	4,953
PIT	@NYM	FLA	NYM	@MON	@FLA	@ATL	MON	PHI	ATL	@PHI	4,058
										Total distance	25,279

(b) Optimal solution of the RD_a^--minimizing TTP (egalitarian method)

Team	Round 1	2	3	4	5	6	7	8	9	10	Distance
ATL	FLA	NYM	PIT	@MON	@NYM	@PHI	MON	PHI	@PIT	@FLA	4,147
NYM	@PIT	@ATL	@FLA	PHI	ATL	@MON	FLA	PIT	MON	@PHI	3,365
PHI	MON	@PIT	@MON	@NYM	FLA	ATL	PIT	@ATL	@FLA	NYM	3,372
MON	@PHI	FLA	PHI	ATL	@PIT	NYM	@ATL	@FLA	@NYM	PIT	4,537
FLA	@ATL	@MON	NYM	PIT	@PHI	@PIT	@NYM	MON	PHI	ATL	5,596
PIT	NYM	PHI	@ATL	@FLA	MON	FLA	@PHI	@NYM	ATL	@MON	3,604
										Total distance	24,621

(c) Optimal solution of the RD_a^--minimizing TTP (Shapley value)

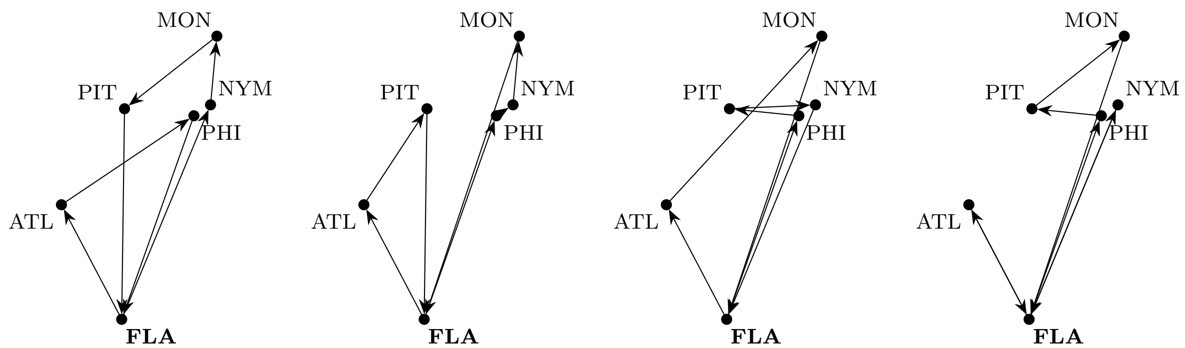
Team	Round 1	2	3	4	5	6	7	8	9	10	Distance
ATL	@FLA	PHI	MON	NYM	@PIT	@MON	@NYM	PIT	@PHI	FLA	4,551
NYM	@MON	FLA	PIT	@ATL	@FLA	@PIT	ATL	PHI	MON	@PHI	3,509
PHI	@PIT	@ATL	@FLA	PIT	MON	FLA	@MON	@NYM	ATL	NYM	3,200
MON	NYM	PIT	@ATL	@FLA	@PHI	ATL	PHI	FLA	@NYM	@PIT	3,994
FLA	ATL	@NYM	PHI	MON	NYM	@PHI	@PIT	@MON	PIT	@ATL	6,455
PIT	PHI	@MON	@NYM	@PHI	ATL	NYM	FLA	@ATL	@FLA	MON	3,218
										Total distance	24,927

(d) Optimal solution of the RD_a^--minimizing TTP (nucleolus)

Figure 3.3: Optimal schedules

minimizing TTP for all three allocations (character @ preceding a team handle indicates an away game). Although the schedules, as well as the distances, differ significantly, one can notice some similar patterns. For instance, the schedules of team ATL in Figures 3.3b and 3.3c imply the same distance despite coinciding only in rounds 1, 9 and 10. This demonstrates the flexibility in scheduling and shows that accommodating one team does not necessarily have a negative effect on all other teams.

To better understand how the different schedules result in different distances, in Figure 3.4, we display the actual travel of team FLA whose distances vary the most across the optimal schedules. This makes the travel of team FLA ideal for illustrative purposes. For the interested



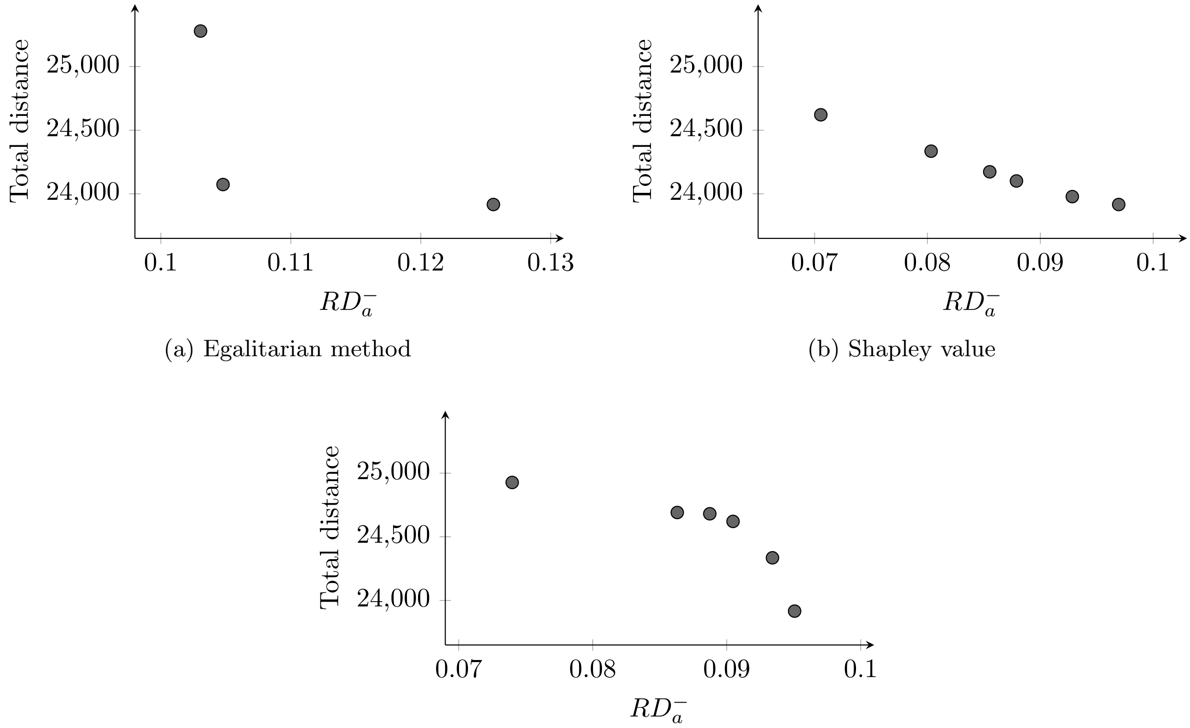
(a) TTP (b) RD_a^--minim. TTP (egalitarian method) (c) RD_a^--minim. TTP (Shapley value) (d) RD_a^--minim. TTP (nucleolus)

Figure 3.4: Travel of team FLA

reader, figures depicting the travel of all other teams are provided in Appendix 3.A. The TTP results with FLA covering a distance of 5,135. This is achieved by two trips, first visiting teams ATL and PHI and second visiting teams NYM, MON and PIT. Although this model minimizes the total distance, looking at Figure 3.4a, it might seem unreasonable that teams PHI and NYM are not visited right after each other. With the RD_a^--minimizing TTP and the egalitarian method, this is resolved and the distance drops to 4,953. This is in fact the shortest possible option for team FLA. Hence, even though the distance is still much larger than the one prescribed by the egalitarian method, it has got as close as it could. The Shapley value and the nucleolus, on the other hand, prescribe team FLA to cover longer distance than 5,135. The RD_a^--minimizing TTP's with respect to these allocations indeed increase the distance. Figures 3.4c and 3.4d show the associated travel. One could argue that the travel seems very inefficient. However, since we are using the RD_a^- metric, this inefficiency and increase in distance is a result of other teams' decreasing distances to achieve a more fair solution.

In Table 3.1, we can see that, for all allocations, the RD_a^--minimizing TTP is associated with an increase in the total distance. Therefore, as a last step, we can draw the Pareto fronts. These are displayed in Figure 3.5 where each circle represents two schedules. This is due to the fact that, in our problem, each feasible schedule can be reversed and run from the last round to the first round without violating any constraints while resulting in the same distances.

In the case of the egalitarian method, one can see in Figure 3.5a that, although the upper-left circle minimizes the RD_a^- metric and the lower-right circle minimizes the total distance, the



(a) Egalitarian method

(b) Shapley value

(c) Nucleolus

Figure 3.5: Pareto fronts

lower-left circle performs very well on both of the criteria. Therefore, a schedule corresponding with the lower-left circle seems to be a reasonable choice. In Figures 3.5b and 3.5c, it is not as simple and the tournament organizer would need to decide based on the properties of the tournament and importance of fairness.

3.5 Conclusion

In this article, we have developed a cooperative game theory approach to produce fair tournament schedules in sports competitions, with particular focus on the celebrated TTP. To our knowledge, this is the first research effort combining sports scheduling and cooperative game theory concepts. Given the importance of scheduling design in sports competitions and the importance of fairness in the outcomes and in the perception of the stakeholders around sport competitions, the study of such well-established concepts of fairness in well-established problems is relevant for the progress of the area.

To determine an ideal distribution of travel distances, we have tested three different methods. As illustrated by our numerical results, the choice of a specific method is crucial for the final schedule. Since the different methods are based on different notions of fairness, this choice should reflect tournament planners' perception of what fairness in a given tournament means.

Next, to find the nearest schedule to a particular distance distribution, we have discussed several metrics to measure the resemblance. Out of the options, we have found that our proposed

RD_a^- metric allows for fairly low computation time while not having some of the other metrics' negative side effects.

Overall, we have shown how to find a fair schedule for a given tournament as well as how to arrive at a solution representing a compromise between fairness and minimum total distance.

Although our focus has been on the TTP, the framework of finding an ideal distance distribution and a nearest schedule is general enough to adjust to different tournament settings. In this regard, we remark that the input to define the cooperative game is a result of the underlying scheduling problem, regardless of whether it is the TTP or another problem. In turn, the computation of this input is often hard in larger-size instances and the computation time increases exponentially with the number of teams, thus developing approximate solution methods remains of interest. On the applied side, it is particularly interesting to analyze how scheduling considering fairness could help teams in low-tier leagues, where resources are likely to be very limited and travel costs account for a large part of total expenses.

Finally, although transferable utility games have gained popularity to address collaborative problems in logistics and transportation, this article has studied the rather unexplored area of how these TU games can be used to compute ideally fair solutions to problems of discrete nature where non-transferable utility issues are in place. In our case, these non-transferable utility issues arose from the burden associated to traveling fatigue and its potential effects on the performance of players. We argue that similar issues arise in collaborative logistics problems and thus our approach can be extended further. For example, implementing a cooperative solution for the scheduling problem of two logistics companies based merely on an economic metric (e.g. cost minimization or revenue maximization) may imply drivers of one company having a more comfortable workload than drivers of the other company. This motivates finding solutions that incorporate driver fatigue into the discussion of the collaboration, where extending the approach developed in this article may be of interest.

References

Algaba, E., Fragnelli, V., and Sánchez-Soriano, J., editors (2019). *Handbook of the Shapley Value*. Chapman and Hall/CRC, New York.

Bao, R. and Trick, M. A. (2010). The relaxed traveling tournament problem. In *Proceedings of the 8th International Conference on the Practice and Theory of Automated Timetabling, PATAT 2010*, pages 167–178.

Bergantiños, G. and Moreno-Ternero, J. D. (2020). Sharing the revenues from broadcasting sport events. *Management Science*, 66(6):2291–2799.

Biró, P., Gyetvai, M., Klimentova, X., Pedroso, J. P., Pettersson, W., and Viana, A. (2020). Compensation scheme with shapley value for multi-country kidney exchange programmes. In *Proceedings of the 34th International ECMS Conference on Modelling and Simulation, ECMS 2020*, pages 129–136.

Bonomo, F., Cardemil, A., Durán, G., Marenco, J., and Sabán, D. (2012). An application of the traveling tournament problem: The Argentine volleyball league. *Interfaces*, 42(3):245–259.

Durán, G., Durán, S., Marenco, J., Mascialino, F., and Rey, P. A. (2019). Scheduling Argentina's professional basketball leagues: A variation on the travelling tournament problem. *European Journal of Operational Research*, 275(3):1126–1138.

Durán, G., Guajardo, M., Lopez, A., Marenco, J., and Zamorano, G. (in press). Scheduling multiple sports leagues with travel distance fairness: An application to Argentinean youth football. *INFORMS Journal on Applied Analytics*.

Easton, K., Nemhauser, G., and Trick, M. (2001). The travelling tournament problem description and benchmarks. In Walsh, T., editor, *Principles and Practice of Constraint Programming — CP 2001*, volume 2239, pages 580–584. Springer, Berlin, Heidelberg.

Hart, S. (2004). A comparison of non-transferable utility values. *Theory and Decision*, 56(1–2):35–46.

Hwang, C.-L. and Masud, A. S. M. (1979). *Multiple Objective Decision Making - Methods and Applications*, volume 164. Springer, Berlin, Heidelberg.

McLean, R. P. (2002). Values of non-transferable utility games. In Aumann, R. and Hart, S., editors, *Handbook of Game Theory with Economic Applications*, volume 3, pages 2077–2120. Elsevier, Amsterdam.

Melo, R. A., Urrutia, S., and Ribeiro, C. C. (2009). The traveling tournament problem with predefined venues. *Journal of Scheduling*, 12:607–622.

Perea, F. and Puerto, J. (2019). A heuristic procedure for computing the nucleolus. *Computers & Operations Research*, 112: Article 104764.

Rasmussen, R. V. and Trick, M. A. (2008). Round robin scheduling – a survey. *European Journal of Operational Research*, 188(3):617–636.

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics*, 17(6):1163–1170.

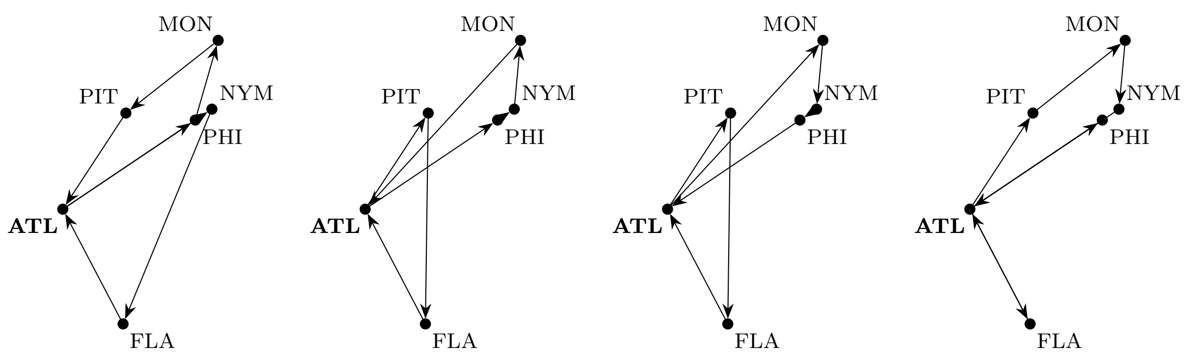
Shapley, L. S. (1953). A value for n-person games. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games*, volume 2, pages 307–317. Princeton University Press, Princeton.

Van Bulck, D. and Goossens, D. (2020). Handling fairness issues in time-relaxed tournaments with availability constraints. *Computers & Operations Research*, 115: Article 104856.

Van Bulck, D., Goossens, D., Schönberger, J., and Guajardo, M. (2020). RobinX: A three-field classification and unified data format for round-robin sports timetabling. *European Journal of Operational Research*, 280(2):568–580.

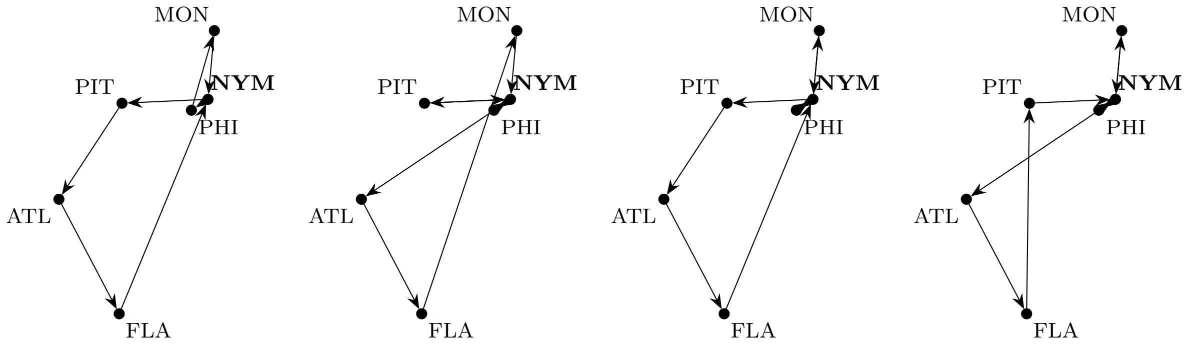
Appendix 3.A Teams' travel for different schedules

For the different schedules, Figures 3.6, 3.7, 3.8, 3.9, and 3.10 depict the actual travel of teams ATL, NYM, PHI, MON, and PIT, respectively. Similar conclusions could be drawn from these figures as in the case of team FLA in section 3.4.



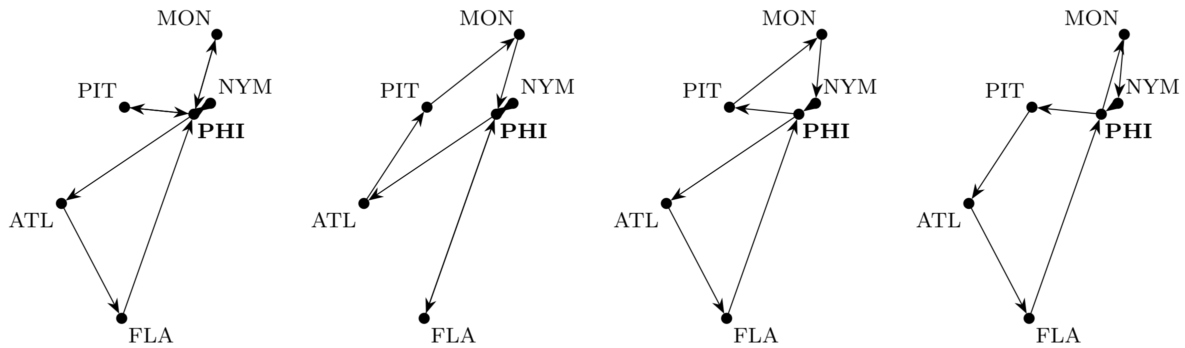
(a) TTP (b) RD_a^--minim. TTP (egalitarian method) (c) RD_a^--minim. TTP (Shapley value) (d) RD_a^--minim. TTP (nucleolus)

Figure 3.6: Travel of team ATL



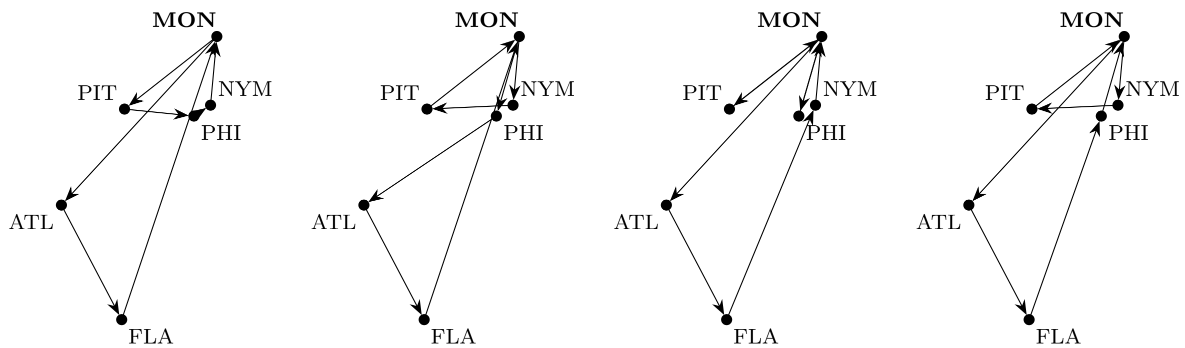
(a) TTP (b) RD_a^--minim. TTP (egalitarian method) (c) RD_a^--minim. TTP (Shapley value) (d) RD_a^--minim. TTP (nucleolus)

Figure 3.7: Travel of team NYM



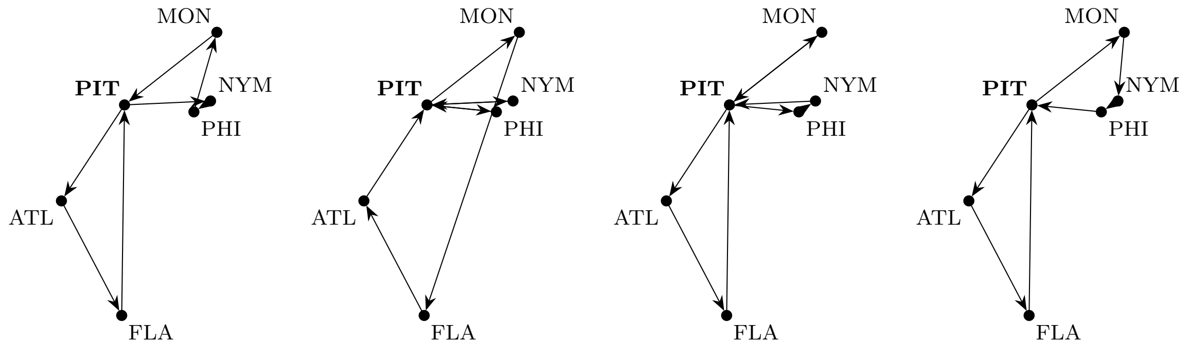
(a) TTP (b) RD_a^--minim. TTP (egalitarian method) (c) RD_a^--minim. TTP (Shapley value) (d) RD_a^--minim. TTP (nucleolus)

Figure 3.8: Travel of team PHI



(a) TTP (b) RD_a^--minim. TTP (egalitarian method) (c) RD_a^--minim. TTP (Shapley value) (d) RD_a^--minim. TTP (nucleolus)

Figure 3.9: Travel of team MON



(a) TTP | (b) RD_a^--minim. TTP (egalitarian method) | (c) RD_a^--minim. TTP (Shapley value) | (d) RD_a^--minim. TTP (nucleolus)

Figure 3.10: Travel of team PIT

Chapter 4

Player-centered approach to coalition formation in transferable utility games with uncertain payoffs

Ondrej Osicka[a]

[a]Department of Business and Management Science, NHH Norwegian School of Economics, Bergen, Norway

Abstract

This study considers cooperative games with transferable utility (TU games) and investigates endogenous coalition formation from a perspective of their players. Several approaches are formulated to determine which coalition is optimal for a given player to pursue while taking into account the subsequent payoff or cost allocation. In particular, the focus is on decision-making situations where coalitions need to be formed before their actual outcome is observable. The formulated models are divided into two main categories, those describing TU games where the subsequent allocation rules are known prior to the coalition formation and those describing TU games where negotiations within the formed coalitions are yet to take place after observing the uncertainty realization. Thus, in addition to a novel approach to the coalition formation, the models also take into account possible uncertainty in the TU games' properties and hence in their characteristic function values. The models are then addressed with a stochastic programming approach. Subsequently, the methodology is illustrated on an example of randomly winning coalitions and on an example of a collaborative transportation problem. The results support arguments against exogenous approaches to coalition formation and show that failing to take the uncertainty in parameter values into account might lead to suboptimal solutions and consequently to false conclusions.

Keywords: Cooperative game theory; Stochastic cooperative games; Endogenous coalition formation; Decentralized coalition formation

4.1 Introduction

In cooperative game theory and more specifically in studies of cooperative games with transferable utility (TU games), authors usually focus on the payoff or cost allocation. The allocation is a mechanism that splits a certain amount (the worth of the coalition) among the players while adhering to certain rules, complaints, or claims of these players. For example, a set of players might claim a certain worth as otherwise, they would be better off leaving the current coalition. Albeit it is important to account for fairness and overall compliance to the allocation, from a perspective of a particular player it might be preferable to completely avoid some players' complaints or claims by simply choosing a coalition without these players in the first place.

In the literature, this option is often dismissed as the majority of studies assumes that a coalition of all players (the grand coalition) is established. However, there are exceptions to this and one can observe a growing academic interest in formation of coalition structures (the set of players partitioned into disjoint coalitions; see Aumann and Maschler (1964), Aumann and Dreze (1974), Guajardo and Rönnqvist (2015) for examples or Rahwan et al. (2015) for a literature review) or coalition configurations (all players assigned to one or more coalitions which may even overlap; see Shehory and Kraus (1996), Chalkiadakis et al. (2010), Guajardo et al. (2018) for examples). Even there, nonetheless, most studies of the allocations consider exogenously given coalitions, i.e., they compute the allocations for coalitions determined beforehand.

Only a small number of studies considers endogenously formed coalitions where the final allocation might actually affect the decision about which coalitions to establish. For example, the model developed by Hart and Kurz (1983) evaluates various coalition structures and, based on this evaluation, determines the stable ones. Similar approaches can be found in articles by Ray and Vohra (1999) or Belleflamme (2000). These studies address the coalition formation from a perspective of a central planner and aim for an outcome advantageous to all players. Nonetheless, it is not difficult to imagine a situation where a small group of players is satisfied with cooperation while some others might never reach mutual agreement. Even if these two groups of players may never affect each other, the central-planner approach would fail while trying to achieve each player's satisfaction.

In this article, we propose a decentralized approach to endogenous coalition formation. In other words, we investigate which coalitions are optimal to form from a perspective of the players while taking into account the subsequent allocation. Additionally, our framework is not limited to TU games with deterministic characteristic functions, but is broad enough to capture decision-making situations where coalitions need to be formed before their actual outcome is observable.

There have been studies considering TU games with uncertain payoffs. Granot (1977) proposed a two-stage procedure to determine an allocation. First, an allocation likely to be realizable is promised. Then, after the actual payoffs are observed, this allocation may be adjusted according to circumstances. Suijs et al. (1998, 1999) and Suijs and Borm (1999) studied an allocation method with application in insurance. Their allocation prescribes to each player a combination of a fixed amount and a proportion of a random loss. Timmer et al. (2004, 2005) proposed several allocations based on the Shapley value prescribed as fractions of the observed payoffs. Fernández et al. (2002) applied methods based on stochastic ordering to define two different notions of core. Other studies such as Benati et al. (2019) have acknowledged the difficult task of computing allocations in TU games like the Shapley value and focused on finding their statistic estimates through stochastic approximations of these games.

In this article, we follow a different approach to the TU games with uncertain payoffs altogether. Instead of deciding on the allocations before the actual outcomes are observed, we postpone the decision after the observation. We provide two examples to motivate this choice.

1. What sensitive information to share with other companies is often an essential question when establishing coalitions and assessing the contribution of different players. With uncertainty, this becomes even more problematic. For example, a company might choose to report an inaccurate probability distribution of their resource availability to seem more interesting to cooperate with and hence secure themselves a better share of the coalition's worth. Afterwards, even when less compelling amounts of the resources tend to occur, it might be difficult to hold this company accountable for their promise.[1]

2. Even when the intentions of a company are honest, the reported probability distribution might not reflect the actual distribution when cooperating. For example, a company might report an accurate probability distribution of their customers' demand. However, after the coalition forms and the company gets promised a certain share of the coalition's worth, the company might become less concerned for example about their marketing efforts and affect the demand's probability distribution. This might be for instance due to feeling more secure in the coalition or observing other companies' behavior and not willing to contribute to the common wealth more than others do.

The methodology proposed in this article aims to find the optimal coalition to form assuming that the worth of the coalition is allocated based on the actual observations. With respect to the aforementioned examples, this should help reduce misreporting and provide players with incentives to operate as efficiently as possible. Note that allocating the actual outcome does not reduce the problem to a deterministic one as the coalitions still need to be established before the actual payoffs or costs can be observed.

The whole decision-making situation is structured in three steps as illustrated by the sequence in Figure 4.1. From the perspective of a particular player, the decisions to be made and

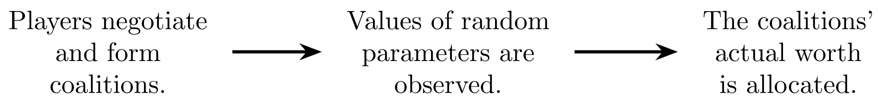


Figure 4.1: Sequence of the coalition formation process

the outcomes are the following. First, the player negotiates with others about which coalition to form. As soon as an agreement is reached, the chosen coalition forms. Second, all the random variables and thus also the coalitions' outcomes are observed. Lastly, the player obtains

[1]Basso et al. (2019) address building of a trustful partnership as one of the main challenges cooperating partners often face.

the final payoff or pays the final cost as prescribed by a given allocation method. Naturally, a rational decision maker would make all decisions in order to maximize this payoff or minimize this cost. In the whole process, for the player, there is only one decision to be made. Which coalition should the player pursue? Our methodology can serve each player as a tool to answer this very question.

The rest of the article is organized as follows. Section 4.2 defines TU games and outlines the aspects of coalition formation with uncertainty in payoffs or costs. In the process outlined in Figure 4.1, one detail still remains unanswered. How and when does one reach an agreement on which allocation method to use? We recognize and investigate two options. In section 4.3, we assume that the allocation method is decided upon at the time of coalition formation, i.e., before the observation of the uncertainty. In section 4.4, we explore the option that the allocation method is to be decided upon only after all coalitions are formed and all parameters' actual values are revealed. In section 4.5, the methodology is illustrated on two examples. Lastly, section 4.6 summarizes the observations and concludes the article.

4.2 Characteristics of the player-centered coalition formation

In this section, we introduce the game-theoretic concepts used throughout the article and discuss options of handling uncertainty in the context of coalition formation.

4.2.1 Preliminaries

In game theory, a cooperative game with transferable utility (TU game) is defined as a pair (N, v). Here, N represents the set of all players (decision makers) and v stands for the characteristic function. Classically, the characteristic function is defined as v: $\mathcal{P}(N) \to \mathbb{R}$ where $\mathcal{P}(N)$ denotes the power set of N. For each nonempty subset $S \subseteq N$, the value $v(S)$ measures the worth of a coalition formed by players included in S. This is usually the highest payoff or the lowest cost this coalition is able to achieve. Furthermore, $v(\varnothing) = 0$. In the rest of this article, all definitions assume characteristic functions representing payoffs. The cost variants would be developed analogously.

To account for uncertainty in the coalitions' worth, we assume a univariate or multivariate random variable $\boldsymbol{\xi}$ to be the sole driver of this uncertainty. This allows us to define the stochastic characteristic function as v: $\mathcal{P}(N) \times Y_{\boldsymbol{\xi}} \to \mathbb{R}$ where $Y_{\boldsymbol{\xi}}$ is the range of $\boldsymbol{\xi}$. For any realization ξ, the worth of coalition S then equals $v(S, \xi)$. In other words, for each coalition, the stochastic characteristic function does not prescribe a single value as is the case in the "classical" cooperative game theory, but rather a random variable $v(S, \boldsymbol{\xi})$.

In the following, when referring to an allocation within coalition $S = \{p_1, \ldots, p_m\}$, we understand a vector $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ satisfying $\sum_{i=1}^{m} x_i = v(S, \xi)$. Here, ξ is the actual realization of $\boldsymbol{\xi}$ and x_i stands for the share of the worth of coalition S allocated to player p_i.

Throughout this study, we assume that the player seeking the best coalition knows payoffs of all potential coalitions. Although it is a strong assumption, it is an essential one as it allows the player to evaluate their prospects in different coalitions. Some games are regulated by rules which automatically guarantee that this information is public to all players, while, in other games, avoiding its disclosure might lead to an important strategic advantage. Sometimes, the data sharing may even be restricted due to privacy or security issues (National Academies of Sciences, Engineering, and Medicine, 2020). Nonetheless, since we implement the stochastic variant of the characteristic function, our methodology also opens opportunities for capturing one's beliefs and expectations of competitors' payoffs.

4.2.2 Uncertainty and the optimal solution

With the uncertainty in the coalitions' worth, it is not always clear what characterizes the optimal solution, i.e., the optimal coalition to pursue.

If a coalition is to be formed once and sustain over a long period of time with a variety of realizations of $\boldsymbol{\xi}$, players might prefer to form coalitions that maximize their expected payoffs. When it comes to long-term cooperation, the lower outcomes for some realizations and higher outcomes for other realizations balance each other and might pose no complications. At the same time, there might be cases where some degree of consistency in the outcomes is needed. If for example the realization of $\boldsymbol{\xi}$ affects the number of employees that need to be on site, the expected value might not be the only important criterion to consider. In such a case, it might be preferable to minimize the variance of the outcomes, maybe while requiring a certain level of the expected value as for example in the portfolio optimization model by Markowitz (1959).

In the opposite extreme, when the coalition is formed with a purpose to sustain for only one realization of $\boldsymbol{\xi}$, the expected value might or might not be desirable. This becomes more of a question of the players' risk aversion. Some might remain inclined to optimize with respect to the expected value or the variance, whereas others might prefer another optimization criterion such as a certain quantile in the positive or negative spectrum of possible outcomes or a certain risk measure such as the value at risk (VaR) or the conditional value at risk (CVaR). For an overview of alternative risk measures, see for example McNeil et al. (2005).

In the following sections, we limit our attention to optimization with respect to the expected value and the variance.

4.3 Player-centered coalition formation with prior agreement on allocation method

At the time when a coalition forms, its worth might not be known and, as argued in the introduction, not even the exact allocation of it should be decided upon. However, it is natural to assume that the coalition formation is accompanied by an agreement of all involved players. To join a coalition, the players might require some prior knowledge of how their share of the benefits is going to be determined. In this section, we assume the shares to be prescribed in accordance to a given allocation method. In other words, we assume a common knowledge that, after forming a coalition and observing $\boldsymbol{\xi}$, the coalition's worth is going to be allocated by a certain allocation method.

There have been numerous allocation methods proposed and used in the literature. With a focus on collaborative transportation, a survey by Guajardo and Rönnqvist (2016) recognized more than 40 different methods. Among these, the Shapley value and various proportional methods were the prevalent choice. Therefore, here we focus especially on these methods. However, the approach is rather straightforward and, for different methods, the formulation should not pose any complications.

4.3.1 Proportional allocation methods

For a coalition $S \subseteq N$, a proportional method assigns each player $p \in S$ a share α_p of the worth $v(S)$, i.e.,

$$x_p = \alpha_p \cdot v(S) \qquad \forall p \in S \tag{4.1}$$

where $\sum_{p \in N} \alpha_p = 1$. Proportional methods are very straightforward and easy to communicate in practice. However, due to their simplicity, they often do not satisfy important game-theoretic properties (Özener and Ergun, 2008).

Different strategies are employed in the literature as how to specify the shares α_p. The simplest option is the egalitarian method assigning to each player the same share, i.e.,

$$x_p = \frac{v(S)}{|S|} \qquad \forall p \in S \tag{4.2}$$

where $|S|$ denotes the cardinality of S.

Another frequently used proportional method is one where the shares are determined in line with the stand-alone payoffs, i.e., the payoffs each player is able to achieve with no cooperation. According to this method,

$$x_p = \frac{v(\{p\})}{\sum\limits_{q \in S} v(\{q\})} \cdot v(S) \qquad \forall p \in S. \tag{4.3}$$

To find the optimal coalitions to pursue, we formulate an approach based on the egalitarian method. Nevertheless, the implementation would be analogous for a different proportional method.

If a random variable $\boldsymbol{\xi}$ is observed as ξ, each player in coalition S receives a payoff equal to

$$\frac{v(S,\xi)}{|S|}. \tag{4.4}$$

Therefore, a risk-neutral player would naturally aim to be part of a coalition which maximizes the expected value of this payoff. To find such a coalition, player p can simply evaluate the value of

$$\mathbf{E}\left[\frac{v(S,\boldsymbol{\xi})}{|S|}\right] \tag{4.5}$$

for each coalition S such that $p \in S$ and choose the coalition producing the highest value. This is the optimal coalition for player p to pursue.

As mentioned in section 4.2, maximizing the expected value might not always be the preferred approach. In case player p would be more interested in minimizing the variance in their allocated share of the coalition's worth, the expected value operator in (4.5) could be replaced by the variance operator. The coalition S such that $p \in S$ associated with the lowest value of

$$\mathbf{Var}\left[\frac{v(S,\boldsymbol{\xi})}{|S|}\right] \tag{4.6}$$

would then indicate the optimal coalition to aim for. Such an approach would then result in a coalition with player p allocated the most consistent share across different realizations of $\boldsymbol{\xi}$. For example, in case of uncertainty in production, this approach could suggest partnership between companies with productions to a certain extent complementary to each other with respect to this uncertainty.

4.3.2 Shapley value

According to Thomson (2019), the Shapley value (Shapley, 1953) is "a centerpiece of the branch of game theory known as 'cooperative game theory'". Its numerous applications covered by Algaba et al. (2019) show its importance across many fields.

For a coalition $S \subseteq N$, the Shapley value assigns each player $p \in S$ a share of the worth $v(S)$ as

$$x_p = \sum_{T \subseteq S \colon p \in T} \frac{(|T|-1)!\,(|S|-|T|)!}{|S|!} \cdot \left(v(T) - v(T \setminus \{p\})\right). \tag{4.7}$$

In a way corresponding to the approach discussed for the proportional methods, a coalition maximizing the expectation of a share allocated to player $p \in N$ according to the Shapley value can be determined as the coalition S associated with the highest value of

$$\mathbf{E}\left[\sum_{T \subseteq S \colon p \in T} \frac{(|T|-1)!\,(|S|-|T|)!}{|S|!} \cdot \left(v(T,\boldsymbol{\xi}) - v(T \setminus \{p\},\boldsymbol{\xi})\right)\right]. \tag{4.8}$$

One may notice that for $S = N$ expression (4.8) prescribes each player a value which coincides with all three Shapley-like solutions proposed by Timmer et al. (2004) assuming all players having "expectation preferences". Nonetheless, whereas they use it to determine the players' shares in the final allocation, we utilize it only as a criterion to find the optimal coalition.

Again, in the case of a player seeking the most consistent outcome of the cooperation, the expected value would be replaced by the variance and the coalition generating the lowest value would be selected.

4.3.3 From preferences to coalitions

When the player's only goal is to be part of a certain coalition, it is not automatically guaranteed that such a coalition will actually be established. There are other players that might still affect the result. If the goal coalition in fact involves other players, these have to be somehow persuaded. If this does not succeed, maybe it is time for some alternative options.

The aforementioned approaches can determine the optimal coalition to pursue for a particular player. It is also easy to obtain the preference order over all coalitions by simply sorting the coalitions this player could be part of from the best one to the worst one. The results may then serve the player as a guidance on which coalitions to prioritize in negotiations with other players.

Additionally, when it is possible to determine which preference order applies to the other players and reasonable to assume their rationality in subsequent negotiations, one could also predict which coalitions will actually form. This is useful as it may provide insight on the real value of cooperation. A player could for example benefit substantially from being in a particular coalition but, if this player is very likely to end up alone anyway, taking part in the negotiations might not be worth the effort.

Nevertheless, before making an assumption on other players' preference orders, one should be very cautious. Naturally, rational decision makers would evaluate their odds in different coalitions to determine the preference orders, but their attitudes towards risk as well as other important aspects of the evaluation might be very difficult to observe by other players.

Literature describes situations where all players' preference orders over all coalitions are revealed as the coalition formation games. For a survey of this literature, see Hajduková (2006). Here, we shortly outline the core stable coalition structures as a potential outcome of the coalition formation.

According to Hajduková (2006), a coalition $T \subseteq N$ blocks a coalition structure P_N (i.e., a partition of N), if each player $p \in T$ strictly prefers the new coalition T to their current coalition $S \in P_N$ where $p \in S$. A coalition structure which admits no blocking coalition is said to be core stable.

Example 5. Consider a game with a set of players $N = \{1, 2, 3\}$ and preference orders

$$\{1,2\} \succ_1 \{1,3\} \succ_1 \{1,2,3\} \sim_1 \{1\},$$
$$\{1,2,3\} \succ_2 \{2,3\} \sim_2 \{1,2\} \succ_2 \{2\},$$
$$\{2,3\} \succ_3 \{1,3\} \succ_3 \{1,2,3\} \sim_3 \{3\},$$

where $\succ_p$ stands for strict preference of player $p \in N$ and $\sim_p$ for their indifference. This means that for example player 1 strictly prefers coalition $\{1, 2\}$ to $\{1, 3\}$ and is indifferent between coalitions $\{1, 2, 3\}$ and $\{1\}$.

One could observe that coalition structures $\{\{1, 2\}, \{3\}\}$ and $\{\{1\}, \{2, 3\}\}$ are core stable as they admit no blocking coalitions. In other words, no group of players would be better off in a different coalition without making at least one of them worse off. Hence, resulting in one of these coalition structures is a likely outcome of the negotiation. On the other hand, for instance the coalition structure $\{\{1, 2, 3\}\}$, i.e., formation of the grand coalition, does not imply the core stability despite it being the best option for player 2. In fact, for both players 1 and 3, forming the two-player coalition $\{1, 3\}$ would be a strictly preferred alternative.

4.4 Player-centered coalition formation without prior agreement on allocation method

In this section, we investigate the situation where players form coalitions before their actual payoffs can be observed, but postpone the decision on the allocation method after the observation. In other words, after the actual worth of a coalition is revealed, players in this coalition negotiate on the allocation. When a player chooses which coalition to pursue, it is hence natural to take into account both the coalition's expected worth and the odds of being able to successfully bargain for a large share of it.

4.4.1 The core and stability

One of the most important concepts in the cooperative theory is the core. This was first introduced by Shapley (1955) and represents a set of allocations within the coalition of all players, the grand coalition. For an allocation $x = (x_1, \ldots, x_{|N|})$ to belong to the core of a TU game (N, v), the conditions of efficiency,

$$\sum_{q \in N} x_q = v(N), \tag{4.9}$$

and rationality,

$$\sum_{q \in T} x_q \geq v(T) \qquad \forall T \subseteq N, \tag{4.10}$$

need to be satisfied. These conditions guarantee that, with any allocation in the core, all profits are divided among the players and there are no incentives for any subset of players to deviate from the collaboration. Such an allocation is then said to be stable. Note that for some TU games the core may also be empty.

For this study, we need to extend the definition of the core so that we can evaluate stability of an allocation within any coalition, not only the grand coalition.

Assuming a coalition S to be established, the efficiency condition can be extended in a straightforward manner as

$$\sum_{q \in S} x_q = v(S). \tag{4.11}$$

Players would have incentives to break out of coalition S and form a new coalition if everyone in the new coalition could be better off. The rationality conditions are supposed to prevent this behavior. For the case of players breaking out of S and forming smaller coalitions within S, conditions based on (4.10) can be formulated as

$$\sum_{q \in T} x_q \geq v(T) \qquad \forall T \subseteq S. \tag{4.12}$$

We refer to these conditions as the conditions of internal rationality.

The players can nonetheless also threaten to break out of S and form a coalition with players which are not part of S. Aumann and Dreze (1974) approached such behavior by simply extending $\forall T \subseteq S$ in conditions (4.12) to $\forall T \subseteq N$. This can be applicable in the case of a central planner with control over all coalitions that are formed as well as all respective allocations. From the players' perspective the situation is more complicated since the players in S have no prior information on how the players outside of S are organized and what their current profit is. We formulate conditions of external rationality to suppress such deviations as

$$\sum_{q \in S \cap T} x_q + \sum_{q \in T \setminus S} y_q \geq v(T) \qquad \forall T \subseteq N \colon T \not\subseteq S, T \cap S \neq \varnothing \tag{4.13}$$

where y_q stands for the current payoff received by player q.

We say that an allocation x within coalition S is stable, if it satisfies conditions (4.11), (4.12) and (4.13). As opposed to the model by Aumann and Dreze (1974), players in coalition S have no control over the values of y_q in (4.13). Moreover, these values may be unknown to them. Therefore, we proceed by formulating sufficient conditions of external rationality.

Let us assume coalitions S and T for which (4.13) is violated. It could be interpreted as a threat to stability of S by forming coalition T instead. However, in the case coalition T would actually establish, the players from $T \setminus S$ would immediately break out unless, in total, they are allocated something better than what they can achieve without the players from S. In other words, coalition T poses no threat to stability of S if, in T, players from $T \setminus S$ cannot be allocated a total payoff of more than

$$v^*(T \setminus S) = \max \sum_{\mathscr{S} \in P_{T \setminus S}} v(\mathscr{S}) \tag{4.14}$$

where $P_{T \setminus S}$ is a set of all partitions of $T \setminus S$.[2] At the same time, it is easy to see that, due to rationality of the remaining players, the players from $T \setminus S$ may be allocated at most a value of

$$v(T) - \sum_{q \in S \cap T} x_q. \tag{4.15}$$

Therefore, if

$$v(T) - \sum_{q \in S \cap T} x_q \leq v^*(T \setminus S), \tag{4.16}$$

coalition T poses no threat to stability of S.

We refer to constraints

$$\sum_{q \in S \cap T} x_q + v^*(T \setminus S) \geq v(T) \qquad \forall T \subseteq N \colon T \not\subseteq S, T \cap S \neq \varnothing \tag{4.17}$$

as the sufficient conditions of external rationality. Overall, if an allocation x within coalition S satisfies conditions (4.11), (4.12) and (4.17), it is stable.

4.4.2 Stability as a proxy for players' best options

We utilize the idea behind the core and stability when assessing the prospects of a player in each coalition.

Let us assume that a coalition S has been formed, the random variable $\boldsymbol{\xi}$ has been observed as ξ and the allocation is now to be negotiated upon. Clearly, the final allocation must satisfy

[2]Note that for a superadditive characteristic function, $v^*(T \setminus S) = v(T \setminus S)$.

efficiency as the payoff $v(S, \xi)$ needs to be fully distributed among the members of S. Additionally, during the negotiation, each member as well as each group of members can claim a share of the payoff which they could accomplish without the rest of the members. This corresponds to the conditions of internal rationality (4.12). Hence, it is reasonable to expect the final allocation to be efficient and internally rational.

For any player p, staying alone would lead to payoff $v(\{p\}, \xi)$. Although this is an efficient and internally rational allocation within $\{p\}$, it is not necessarily the optimal coalition for player p. To select the best coalition to be in, one needs to also account for the conditions of external rationality with respect to the deviations player p could be part of. If there is a coalition S including player p and satisfying conditions

$$\sum_{q \in S \cap T} x_q + \sum_{q \in T \setminus S} y_q \geq v(T, \xi) \qquad \forall T \subseteq N \colon T \not\subseteq S, p \in T, \tag{4.18}$$

it guarantees that player p has no incentives to deviate from coalition S. This essentially makes S the best coalition for player p. In the following, instead of (4.18), we use a sufficient condition corresponding to (4.17) with $v^*(T \setminus S, \xi)$ defined as

$$v^*(T \setminus S, \xi) = \max \sum_{\mathscr{S} \in P_{T \setminus S}} v(\mathscr{S}, \xi) \tag{4.19}$$

where $P_{T \setminus S}$ is a set including all partitions of $T \setminus S$.

In summary, for a player p and $\boldsymbol{\xi}$ observed as ξ, if there exists a coalition S including player p and an allocation x within S satisfying conditions

$$\sum_{q \in S} x_q = v(S, \xi), \tag{4.20}$$

$$\sum_{q \in T} x_q \geq v(T, \xi) \qquad \forall T \subseteq S, \tag{4.21}$$

$$\sum_{q \in S \cap T} x_q + v^*(T \setminus S, \xi) \geq v(T, \xi) \qquad \forall T \subseteq N \colon T \not\subseteq S, p \in T, \tag{4.22}$$

it is optimal to be in this coalition.

Nonetheless, such a coalition does not always need to exist. In such a case, it might be impossible to predict the outcome of the negotiation and find the optimal coalition. The vast amount of different allocation methods in the literature demonstrates that there is no consensus on bargaining power and fairness which could help make the prediction. Nevertheless, it is easy to see that for instance for TU games with nonempty cores, the optimal coalition always exists in form of the grand coalition.

Stability likelihood maximization

For the player-centered coalition formation in TU games with uncertain payoffs, we propose an approach we refer to as the stability likelihood maximization (SLM). According to this approach, coalition S is optimal for player p, if $p \in S$ and S maximizes the probability that an allocation

x within S satisfying conditions (4.20)–(4.22) exists. In other words, for player p, this approach determines the coalition associated with the highest probability of being the best coalition to be in.

The SLM approach comes with a few drawbacks. For example, there might be a situation when it is better to be in the second best coalition with a probability of 100 % than to be in the best coalition with a probability of only 60 %. Moreover, the model does not control for what happens when the stability cannot even be achieved. However, without the knowledge of the actual outcomes, the efficiency and rationality are the only concepts to work with. This is all due to the fact that the allocation method is not determined at time of the coalition formation. In fact, if one had a prediction for how the worth will be distributed, the methodology introduced in section 4.3 should always be the preferred option.

The conditions (4.22) might pose another limitation. Since (4.22) stand only for sufficient conditions of external rationality, the probability associated with the optimal coalition may actually underestimate the real probability of a stable outcome.

For a discrete random variable $\boldsymbol{\xi}$, a mixed-integer linear programming model to determine the optimal coalition can be formulated. Denoting the set of all possible realizations of $\boldsymbol{\xi}$ by Ξ and the probability of realization ξ by π_ξ, the model for player p can be stated as

$$\max \sum_{S\subseteq N:\, p\in S} \sum_{\xi\in\Xi} \pi_\xi \kappa_{S,\xi} \tag{4.23}$$

$$\text{s.t.} \quad \sum_{q\in T} x_{q,\xi} \geq v(T,\xi) - (1-\kappa_{S,\xi})\cdot M \qquad \forall S,T\subseteq N, \xi\in\Xi\colon p\in S, T\subseteq S, \tag{4.24}$$

$$\sum_{q\in S\cap T} x_{q,\xi} + v^*(T\setminus S,\xi) \geq v(T,\xi) - (1-\kappa_{S,\xi})\cdot M \qquad \forall S,T\subseteq N, \xi\in\Xi\colon p\in S, T\nsubseteq S, p\in T, \tag{4.25}$$

$$\sum_{q\in N} x_{q,\xi} = \sum_{S\subseteq N:\, p\in S} \delta_S \cdot v(S,\xi) \qquad \forall \xi\in\Xi, \tag{4.26}$$

$$\sum_{q\in N\setminus S} x_{q,\xi} \leq (1-\delta_S)\cdot M \qquad \forall S\subseteq N, \xi\in\Xi\colon p\in S, \tag{4.27}$$

$$\sum_{S\subseteq N:\, p\in S} \delta_S = 1, \tag{4.28}$$

$$\kappa_{S,\xi} \leq \delta_S \qquad \forall S\subseteq N, \xi\in\Xi\colon p\in S, \tag{4.29}$$

$$\delta_S \in \{0,1\} \qquad \forall S\subseteq N\colon p\in S, \tag{4.30}$$

$$\kappa_{S,\xi} \in \{0,1\} \qquad \forall S\subseteq N, \xi\in\Xi\colon p\in S, \tag{4.31}$$

$$x_{q,\xi} \in \mathbb{R} \qquad \forall q\in N, \xi\in\Xi \tag{4.32}$$

where M is a big enough number. For example, $M = |N|\cdot \max_{S\subseteq N,\xi\in\Xi}(v(S,\xi))$ would suffice.

In the model, binary variable δ_S equals 1 when coalition S is the optimal coalition for player p to join, 0 otherwise. Binary variable $\kappa_{S,\xi}$ equals 1 when δ_S equals 1 and a stable allocation can be achieved within coalition S under realization ξ, 0 otherwise. Continuous variable $x_{q,\xi}$ expresses the share allocated to player q under realization ξ. The objective function (4.23) maximizes the probability of a stable allocation. Constraints (4.24) and (4.25) correspond to (4.21) and (4.22), respectively, and additionally allow for violation via forcing the respective variable $\kappa_{S,\xi}$ to equal 0 and thus having no contribution to the objective function. The combination of constraints (4.26) and (4.27) enforces the efficiency condition (4.20) by dividing the worth of coalition S among its players while prescribing $x_{q,\xi} = 0$ to all players outside of S. Lastly, constraint (4.28) guarantees that only one coalition is selected while constraints (4.29) make sure that only allocations within this coalition can contribute to the objective function.

Example 6. Let us assume a TU game of three players 1, 2 and 3 in which each coalition $S \subseteq \{1,2,3\}$ with at least two members receives a payoff of 1, i.e., $v(S) = 1$ when $|S| \geq 2$, whereas staying alone leads to no payoff, i.e., $v(\{1\}) = v(\{2\}) = v(\{3\}) = 0$. This example hence assumes a deterministic characteristic function. Focusing on the prospects of player 1 in various coalitions, and thus solving model (4.23)–(4.32) for $p = 1$, would lead to one of two optimal coalitions. Coalitions $\{1,2\}$ and $\{1,3\}$ both result in an objective value of 100 %. Let us assume that the coalition $\{1,2\}$ subsequently forms. This coalition receives a payoff of 1 which can be split among its members. As long as both players 1 and 2 are allocated a nonnegative share of the payoff, they would not be better off alone.

One could oppose that the result of the example is not stable as the player 3 could still interfere. In fact, if player 2 does not receive the full payoff $v(\{1,2\})$, they could threaten to deviate and form coalition $\{2,3\}$ with an allocation from which both players 2 and 3 would benefit. This may or may not be a valid threat to the stability. In some settings, when a coalition forms, claims involving external players might be allowed. In other settings, they might be restricted for instance by the contract design. The SLM approach corresponds to the latter situation. To account for the former situation instead, i.e., claims involving external players are valid in the allocation negotiations, the model can be easily modified as follows.

Stability likelihood maximization with negotiations involving external players

Conditions (4.21) describe the negotiations in a formed coalition with respect to claims involving players within the coalition. This can be extended to involve the external players by additionally accounting for conditions

$$\sum_{q \in S \cap T} x_q + v^*(T \setminus S, \xi) \geq v(T, \xi) \qquad \forall T \subseteq N \colon T \not\subseteq S, T \cap S \neq \varnothing \tag{4.33}$$

corresponding to the sufficient conditions of external rationality (4.17). Note that the set of conditions (4.33) includes also conditions (4.22).

In line with the SLM approach, for a player p, coalition S, such that $p \in S$, is optimal if it maximizes the probability that an allocation x within S satisfying conditions (4.20), (4.21) and (4.33) exists. We refer to this approach as the stability likelihood maximization with negotiations involving external players (SLMext).

Looking at Example 6, one would now find out that for all players, all coalitions are associated with a stability likelihood of 0 %. For this example, this approach hence prescribes no preferable coalitions.

4.4.3 From optimal coalitions to coalition structures

Although a player may prefer a certain coalition, this coalition is not yet guaranteed to actually form. As in section 4.3, it is interesting to investigate which coalition structures could be the result of all players using the SLM or the SLMext approach.

As opposed to the analysis in section 4.3, these models provide players only with the optimal coalitions and not with the full preference orders. Even though, it is often possible to determine the likely outcome of the negotiation. If, in a three-player TU game, players 1 and 2 both pursue coalition $\{1, 2\}$, no matter what player 3's preference is, the coalition $\{1, 2\}$ will form and player 3 will be left alone. Generally speaking, a coalition S can be expected to form if it is an optimal coalition for all players $p \in S$. On the other hand, when player 1 aims for coalition $\{1, 2\}$, player 2 for $\{2, 3\}$ and player 3 for $\{1, 3\}$, the situation gets much more complicated and the outcome is unclear.

Assuming a TU game (N, v) and all its players using the SLM approach, one would need to determine their optimal coalitions and see if there exists a coalition $T_1 \subseteq N$ optimal for all its members. If there indeed is such a coalition, it will form. Then, since the players of T_1 have achieved their optimal solution, the remaining players cannot affect them anymore. Hence, the remaining players can apply the SLM approach on a TU game $(N \setminus T_1, v)$. Again, if there exists a coalition $T_2 \subseteq N$ optimal for all its members, it will form. This can be repeated until all the players are assigned to a coalition resulting in a coalition structure $\{T_1, T_2, \dots\}$. If, at any point, there is no coalition optimal to all its members, no conclusion can be made about the remaining players. If, on the other hand, there is more than one coalition optimal to all its members, the process can be broken into separate branches possibly resulting in multiple coalition structures.

With the SLMext approach, the process is easier. One can notice that conditions (4.20), (4.21) and (4.33) do not explicitly include the player p. Hence, a coalition S maximizing the probability that an allocation x within S satisfying conditions (4.20), (4.21) and (4.33) exists is also an optimal coalition for all players $p \in S$ according to the SLMext approach. This can be again repeated for a TU game of all the remaining players until the complete coalition structure is revealed. In this case, the final coalition structure can always be determined and, again, in case of multiple optimal coalitions, the process can lead to multiple coalition structures.

4.5 Player-centered coalition formation on examples

To illustrate the proposed framework for endogenous coalition formation from the players' perspective, we apply the described methods on two examples. First, the concepts are illustrated on a simple example of a three-player TU game where coalitions receive positive payoffs only with a given probability. Then, the methods are applied for coalition formation in an example of a collaborative transportation problem.

To see the added value the stochastic programming approaches bring as compared to deterministic ones and to highlight the importance of proper specification of the characteristic functions, we also present results of the methods' deterministic reformulations.

4.5.1 Problem of randomly winning coalitions

Let us assume a TU game of three players 1, 2 and 3 similar to the deterministic example from the previous section. Staying alone again leads to no payoff, i.e., $v(\{1\}) = v(\{2\}) = v(\{3\}) = 0$. On the other hand, cooperation with other players may or may not lead to a positive payoff. Particularly, for each S such that $|S| \geq 2$, $v(S) = 1$ (S is a winning coalition) with probability of 50 % and $v(S) = 0$ (S is not a winning coalition) otherwise. The payoffs of different coalitions are mutually independent. Hence, there might be multiple winning coalitions at the same time as well as there might be no such coalition.

There are four coalitions which may or may not be winning. This implies 16 distinct combinations which can be interpreted as scenarios with probability of realization equal to $0.5^4 = 0.0625$. For instance, one scenario may depict coalitions $\{1, 2\}$ and $\{2, 3\}$ as winning while all others as not winning. With such a specification of the uncertainty, it is easy to apply the methods proposed in this article.

We also present results of the methods' deterministic reformulations. These are based on the expected values of all parameters. In other words, we solve the models as if there was only one scenario with probability of 100 % with $v(\{1\}) = v(\{2\}) = v(\{3\}) = 0$ and $v(\{1, 2\}) = v(\{1, 3\}) = v(\{2, 3\}) = v(\{1, 2, 3\}) = 0.5$.

Prior agreement on the egalitarian method

Given that the players from the example would like to form a coalition with a prior knowledge that the egalitarian allocation method will be applied, they could determine the best coalition with the approach described in section 4.3.1. Taking player 1 for instance, the potential coalitions would lead to expected allocated payoffs and expected variance of the allocated payoffs shown in Table 4.1.

Table 4.1: Properties of the share allocated to player 1 by the egalitarian allocation method

Coalition	Expected value	Variance
$\{1\}$	0	0
$\{1, 2\}$	0.25	0.06
$\{1, 3\}$	0.25	0.06
$\{1, 2, 3\}$	0.17	0.03

From the expected-value viewpoint, pursuing a two-player coalition is the optimal solution as it generates the highest expected value. It is rather straightforward as the stand-alone option yields no payoff and the three-player option is associated with the same probability of winning, but the payoff needs to be allocated among more players. Interestingly, the results are exactly the same with the deterministic reformulation.

The variance minimizing approach, on contrary, favors staying alone followed by forming the grand coalition. The two-player coalitions yield the highest variance and are therefore the least preferred option. Nonetheless, in this simple example, the variance minimizing approach might be quite misleading as, the higher the variance is, the further away from 0 the allocated share is in case of winning.

To find a compromise between the highest expected value and the lowest variance, one could also choose a coalition from a set of coalitions for which there exists no coalition that would yield lower expected payoff with lower variance, i.e., a set corresponding to the so-called Markowitz efficient frontier. For player 1, this includes all the coalitions since the stand-alone option dominates in terms of variance, the two-player coalitions in terms of the expected value and the grand coalition stands for an intermediate option on both measures.

As described in section 4.3.3, with preference orders of all players, one can evaluate the stability of different coalition structures to predict which coalitions are likely to actually form. Assuming all players minimizing the expectation of their allocated share, one would obtain preference orders

$$\{1,2\} \sim_1 \{1,3\} \succ_1 \{1,2,3\} \succ_1 \{1\},$$
$$\{1,2\} \sim_2 \{2,3\} \succ_2 \{1,2,3\} \succ_2 \{2\},$$
$$\{2,3\} \sim_3 \{1,3\} \succ_3 \{1,2,3\} \succ_3 \{3\}.$$

Clearly, it is not possible to achieve all players' optimal coalitions and at least one player needs to step back. In fact, three core stable coalition structures can be found and all involve one player remaining alone, namely, $\{\{1,2\},\{3\}\}$, $\{\{1,3\},\{2\}\}$ and $\{\{1\},\{2,3\}\}$.

Prior agreement on the Shapley value

With players' prior agreement on employing the Shapley value to allocate the final payoff, the approach described in section 4.3.2 results for player 1 in values reported in Table 4.2. Compared to the agreement on the egalitarian method, there is only one difference. In case of

Table 4.2: Properties of the share allocated to player 1 by the Shapley value

Coalition	Expected value	Variance
$\{1\}$	0	0
$\{1,2\}$	0.25	0.06
$\{1,3\}$	0.25	0.06
$\{1,2,3\}$	0.17	0.07

the grand coalition, the variance is slightly higher at value 0.07. Altogether, with the exception of the grand coalition not being the least preferred option from the variance viewpoint and the Markowitz efficient frontier now consisting only of coalitions $\{1\}$, $\{1,2\}$ and $\{1,3\}$, the results remain the same.

Stability likelihood maximization

In the case the allocation method is not agreed upon beforehand, one should resort to one of the approaches from section 4.4 based on stability likelihood maximization. Due to the finite number of scenarios, the SLM approach can be carried out with model (4.23)–(4.32). For the SLMext approach, an analogous model can be formulated.

For player 1, the SLM approach determines coalitions $\{1,2\}$ and $\{1,3\}$ as the optimal ones, both associated with an objective value of 62.5 %. This means that with probability of 62.5 % the players in the coalition are able to find an allocation such that there is no better coalition for player 1 and there is no better coalition without involving external players for the other player.

When the stability may be affected by external players, i.e., applying the SLMext approach, $\{1,2,3\}$ is the optimal coalition for player 1. The optimal objective value equals 50 %. This is less than the 62.5 % associated with coalitions $\{1,2\}$ and $\{1,3\}$ using the SLM approach. However, since the SLMext approach would evaluate the two-player coalitions at only 43.75 %, the grand coalition is a strictly preferred option when external players have a say in the negotiations.

Following the approaches from section 4.4.3, one can investigate which coalitions are actually likely to establish with all players following the SLM or the SLMext approach. With the SLM approach, coalition structures $\{\{1,2\},\{3\}\}$, $\{\{1,3\},\{2\}\}$ and $\{\{1\},\{2,3\}\}$ represent possible outcomes with two players achieving their optimal coalition and one player forced to stay alone. On contrary, with the SLMext approach, formation of the grand coalition $\{1,2,3\}$ can be expected as it is the optimal coalition for all players.

Looking at the deterministic reformulations for both the SLM and the SLMext approach, it is clear that the optimal objective value equals either 100 % or 0 % since the expected values essentially generate a single scenario. This weakness of the deterministic approaches is apparent from the results. With the SLM approach for player 1, coalitions $\{1,2\}$ and $\{1,3\}$ result in 100 % while all other coalitions result in 0 %. Although in this case the optimal solution coincides with the original optimal solution, the stability likelihood is largely overestimated. The situation is different with the SLMext approach for player 1, where the deterministic reformulation results in 0 % for all coalitions. Hence, such a result provides the player with no insight. This indicates the importance of proper specification of the uncertainty in the characteristic function.

4.5.2 Collaborative transportation problem

In this part, we illustrate the proposed methods for coalition formation on a more complex and arguably more practical example of a TU game, the collaborative transportation problem. We use a similar formulation as used by Frisk et al. (2010). It is a TU game with a set of players N transporting a certain commodity. As opposed to Frisk et al. (2010), we assume a single commodity instead of multiple commodities. Each player $p \in N$ is associated with a distinct set of supply points (origins) I_p and a distinct set of demand points (destinations) J_p. When cooperation among the players takes place, it allows satisfying some of the demand from partners' supply points with an objective to minimize the total transportation cost. We assume the supply to be random. Hence, with the formulation by Frisk et al. (2010), we could face infeasibility issues if the supply did not suffice to satisfy the total demand. Therefore, we modify the model by introducing a penalty for any unsatisfied demand. The model to determine the optimal cost associated with each coalition S and each realization ξ of random variable $\boldsymbol{\xi}$ can then be formulated as

$$v(S,\xi) = \min \sum_{i\in I_S}\sum_{j\in J_S} c_{ij}y_{ij} + \sum_{j\in J_S} c_j^e z_j \tag{4.34}$$

$$\text{s.t.} \quad \sum_{j\in J_S} y_{ij} \le s_i(\xi) \qquad \forall i \in I_S, \tag{4.35}$$

$$\sum_{i\in I_S} y_{ij} + z_j = d_j \qquad \forall j \in J_S, \tag{4.36}$$

$$y_{ij} \ge 0 \qquad \forall i \in I_S, j \in J_S, \tag{4.37}$$

$$z_j \ge 0 \qquad \forall j \in J_S. \tag{4.38}$$

Here, sets I_S and J_S contain all supply points and all demand points of players in S, respectively, i.e., $I_S = \bigcup_{p\in S} I_p$ and $J_S = \bigcup_{p\in S} J_p$. Parameter c_{ij} stands for the cost of transporting one unit of the commodity from supply point i to demand point j. Parameter c_j^e represents the cost of unsatisfied demand at demand point j per unit of the commodity. The supply at supply point i and the demand at demand point j are expressed by $s_i(\xi)$ and d_j, respectively. In this case, since the supply is subject to uncertainty, it is formulated as a function of the realization of random variable $\boldsymbol{\xi}$.

Variables y_{ij} and z_j equal the amount to be transported from supply point i to demand point j and the amount of unsatisfied demand at demand point j, respectively. They are determined by minimizing the total cost in (4.34) while not exceeding the available supply at each supply point as prescribed by (4.35). Lastly, (4.36) are the demand constraints allowing for unsatisfied demand and constraints (4.37) and (4.38) enforce the variables to be nonnegative. Note that in this example, the characteristic function does not stand for payoffs, but for costs of the coalitions.

Let us assume a situation depicted in Figure 4.2. Here, five players own one supply point

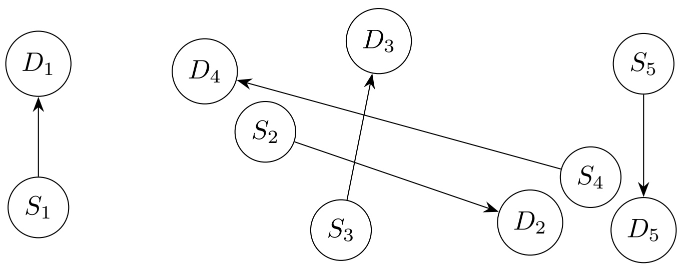


Figure 4.2: Example of a collaborative transportation problem

and one demand point each. For example, S_1 and D_1 are a supply point and a demand point of player 1, respectively. The values of c_{ij} are generated as the euclidean distance between the respective points whose position is described in Table 4.3. Furthermore, we assign each demand point a demand of 250 and a cost per unit of unsatisfied demand of 1000. For the supply points, the supply values are generated as realizations of i.i.d. random variables following normal distribution $N(300, 1000)$. We discretize these variables by generating 100 scenarios with uniform probabilities of occurrence.

Table 4.3: Position of the supply and demand points

	S_1	S_2	S_3	S_4	S_5	D_1	D_2	D_3	D_4	D_5
x-coordinate	10	40	50	83	90	10	75	55	32	90
y-coordinate	3	13	0	7	22	22	1	25	21	0

In the following, we again apply all the proposed methods of finding the best coalition. As for the previous example, we also include results of their deterministic reformulations. These are again based on the expected values of the parameters. In other words, we solve the models as if there was only one scenario with probability of 100 % with the supply at each supply point equal to 300.

Prior agreement on the egalitarian method

In the case the players from the example would like to form a coalition with a knowledge that the egalitarian allocation method will be applied, the approach described in section 4.3.1 can be used. Player 1, for instance, could evaluate all coalitions in terms of the expected allocated cost and expected variance of the allocated costs. These values are reported in Table 4.4.

Table 4.4: Properties of the share allocated to player 1 by the egalitarian allocation method

Coalition	Expected value	Variance	Coalition	Expected value	Variance
$\{1\}$	5995.4	31740745	$\{1,3,4\}$	6541.2	1230991
$\{1,2\}$	7210.0	3715158	$\{1,3,5\}$	5559.7	4846
$\{1,3\}$	5792.5	3418771	$\{1,4,5\}$	6697.9	126341
$\{1,4\}$	8608.5	4480485	$\{1,2,3,4\}$	4044.7	4272
$\{1,5\}$	5422.6	6109464	$\{1,2,3,5\}$	5336.4	1150
$\{1,2,3\}$	5339.6	355733	$\{1,2,4,5\}$	3755.9	10238
$\{1,2,4\}$	3378.0	5368	$\{1,3,4,5\}$	5498.0	17561
$\{1,2,5\}$	6346.5	119722	$\{1,2,3,4,5\}$	4214.3	7028

Prioritizing the lowest expected cost would lead player 1 to pursuing coalition $\{1,2,4\}$. On the other hand, a variance minimizing objective would suggest coalition $\{1,2,3,5\}$. The Markowitz efficient frontier, in addition to these two, contains also coalition $\{1,2,3,4\}$ representing an intermediate option on both measures.

Assuming all players aiming to minimize their expected cost, their preference orders over all coalitions can be determined. With them, we find that there is only one core stable coalition structure, namely the coalition structure $\{\{3\},\{2,4\},\{1,5\}\}$. This is hence a possible outcome of the coalition formation.

The deterministic reformulation results for all players with the same optimal coalitions. It can be however observed that the deterministic formulation slightly underestimates the actual expected cost. For example, for player 1, the optimal objective equals 3359.5 instead of the previously obtained cost of 3378.0. Although we do not see differences in the optimal coalitions, the preference orders are affected and generate a different core stable coalition structure, specifically, $\{\{1\},\{2,4\},\{3\},\{5\}\}$. Hence, omitting the proper specification of the parameters and simplifying the problem into a deterministic one may in fact lead to suboptimal solutions and inefficient managerial decisions.

Prior agreement on the Shapley value

With players' agreement on the Shapley value, the approach described in section 4.3.2 results for player 1 in values reported in Table 4.5.

Table 4.5: Properties of the share allocated to player 1 by the Shapley value

Coalition	Expected value	Variance	Coalition	Expected value	Variance
$\{1\}$	5995.4	5008463	$\{1,3,4\}$	4926.3	3672493
$\{1,2\}$	5177.6	6647699	$\{1,3,5\}$	5133.8	756194
$\{1,3\}$	5221.9	2082219	$\{1,4,5\}$	4533.0	1991516
$\{1,4\}$	4811.4	7694333	$\{1,2,3,4\}$	4910.3	1344385
$\{1,5\}$	5418.5	1620829	$\{1,2,3,5\}$	4959.5	1952171
$\{1,2,3\}$	5045.3	2610719	$\{1,2,4,5\}$	4818.9	1306237
$\{1,2,4\}$	4893.8	4061667	$\{1,3,4,5\}$	4736.8	1054453
$\{1,2,5\}$	5091.7	3509091	$\{1,2,3,4,5\}$	4840.3	722815

The optimal coalitions for player 1 are $\{1,4,5\}$ in terms of expected value and $\{1,2,3,4,5\}$ in terms of variance. The Markowitz efficient frontier then includes also coalitions $\{1,4,5\}$ and $\{1,3,4,5\}$.

Again, when all players follow the same approach, their preference orders over all coalitions can be determined. These then lead to five core stable coalition structures, namely,

$$\{\{1,3\},\{2,4\},\{5\}\},$$
$$\{\{3\},\{1,2,4,5\}\},$$
$$\{\{1,3\},\{2,4,5\}\},$$
$$\{\{2,4\},\{1,3,5\}\},$$
$$\{\{1,2,4\},\{3,5\}\}.$$

The deterministic reformulation again leads to the same optimal coalitions with an exception of player 5. In that case, coalition $\{1,3,4,5\}$ becomes the most desirable option instead of $\{3,4,5\}$. Moreover, the preference orders of all players are affected as well as the core stable coalition structures. With the deterministic approach, the coalition structure $\{\{1,2,4\},\{3,5\}\}$ is no more core stable while four new coalition structures become core stable.

Stability likelihood maximization

For the case of players with no prior agreement on the allocation method, we apply the approaches based on stability likelihood maximization.

As mentioned in section 4.4, for TU games admitting core allocations, the grand coalition is always an optimal coalition. For a problem similar to the collaborative transportation problem (4.34)–(4.38), Sánchez-Soriano et al. (2001) proved that the core is always nonempty. In fact, we observe the same for all 100 scenarios. On the other hand, even in logistics there are many problems which might result in an empty core. This has been shown for example for cooperative vehicle routing problems or location-routing games by Potters et al. (1992) and Osicka et al. (2020), respectively.

To keep the example simple and easy to follow as well as to demonstrate another useful application of the proposed methodology, we proceed with the collaborative transportation problem. However, as done for example by Guajardo and Rönnqvist (2015), we impose a condition on maximum cardinality of the established coalition. We set the value to 3 which means that only coalitions including up to 3 players may be formed. Hence, this excludes the five-player grand coalition as well as all four-player coalitions in the example.

The constrained cardinality affects not only the set of potential coalitions but also the set of coalitions which may serve as threats to deviate. Therefore, the SLM and SLMext approaches need to be modified by adding condition $|T| \leq 3$ into (4.21), (4.22) and (4.33). With such modified models, optimal coalitions can be determined for each player as shown in Table 4.6.

Table 4.6: Optimal coalitions with approaches based on stability likelihood maximization

	SLM		SLMext	
Player	Optimal coalition	Stability likelihood	Optimal coalition	Stability likelihood
1	$\{1,2,4\}$	89 %	$\{1,2,4\}$	1 %
2	$\{2,4,5\}$	72 %	$\{2,4,5\}$	71 %
3	$\{2,3,4\}$	89 %	$\{2,3,4\}$	21 %
4	$\{2,4,5\}$	71 %	$\{2,4,5\}$	71 %
5	$\{2,4,5\}$	86 %	$\{2,4,5\}$	71 %

It is interesting to see that both approaches prescribe the same coalitions to all players. It is however also easy to observe the fundamental difference between the approaches. The coalition $\{1,2,4\}$ promises player 1 likelihood of 89 % with the use of the SLM approach, but as soon as external players enter the negotiations, and more specifically player 5, it changes radically. This is obvious from the fact that coalition $\{2,4,5\}$ is optimal for all involved players with both approaches. Therefore, the SLM approach might be useful for player 1 as long as players 2 and 4 are unaware of the added value player 5 could bring.

Following the procedures from section 4.4.3, one can find a potential outcome of all players using any of these approaches to be a coalition structure $\{\{1,3\},\{2,4,5\}\}$. This is not surprising based on the prevalent preference for coalition $\{2,4,5\}$.

Results of the deterministic reformulations of both approaches again to some extent correspond with the results, but due to their binary nature (100 % or 0 %), they might appear very misleading. Both approaches would for example prescribe the coalition $\{2,4,5\}$ with 100 % for all involved players. If such a coalition formed, it is easy to imagine that a realization allowing no stable allocation could make the players question their decisions. From the stochastic formulations, we already know that such a realization may actually occur. The inconsistency between the predicted and the actual results could even lead to distrust among the partners and affect their future relations. Proper specification of the uncertain nature of the TU games is hence very important and might be essential for a successful cooperation.

4.6 Conclusion

We have studied endogenous coalition formation in TU games from a perspective of their players. For this, we have proposed several approaches to determine which coalition is optimal for a given player to pursue while taking into account the subsequent payoff or cost allocation. This makes the methodology different from the fairly common exogenous coalition formation which to an extent disregards the allocation when deciding on the optimal coalitions. This methodology also differs from the traditional central-planner approach as it maximizes the welfare of a particular player instead of the aggregated welfare of all players.

Our framework also captures possible uncertainty in the TU games' properties and hence in their characteristic function values. In particular, the focus is on decision-making situations where coalitions need to be formed before their actual outcome is observable. The introduced methods are divided into two main categories, those describing TU games where the subsequent allocation rules are known prior to the coalition formation and those describing TU games where negotiations within the formed coalitions are yet to take place after observing the uncertainty realization. We illustrate the framework on an example of randomly winning coalitions and on an example of a collaborative transportation problem.

For the case when players agree on using a particular allocation method prior to the coalition formation, we demonstrate the methodology for the egalitarian method and for the Shapley value. The results from the collaborative transportation problem show that different allocation methods might lead to different optimal coalition and subsequently to different partitions of the players. This implies that the allocation method choice might substantially affect the optimal coalitions. Unless the coalition formation in a TU game is inherently exogenous, this undermines the reliability of the traditional methods and strengthens the argument for using endogenous methods as those described in this article.

Because for some TU games the prior agreement on an allocation method might not be possible, we have also introduced two approaches based on stability likelihood maximization as alternatives to the exogenous coalition formation. As shown on the examples, they can be easily used to determine the optimal coalitions along with the probability of it being the best coalition. Additionally, as illustrated on the collaborative transportation problem, these approaches can be also used in TU games with restrictions on the coalitions' cardinality.

Furthermore, comparing all the approaches with their deterministic reformulations, we have observed slightly different results. Although the results might sometimes coincide, we have seen weaknesses of the deterministic formulations in several aspects. This highlights the importance of taking the uncertainty in the TU games' parameters into account as its omission might lead to suboptimal decisions.

To conclude, we suggest few directions for further research. One natural avenue would be to extend the SLM and SLMext approaches by an approximation of what happens when the stability cannot be achieved. Another interesting direction could be evaluation of different negotiation strategies. With knowledge of the optimal coalitions as well as likely final coalition structures, from a player's point of view, there might be a coalition structure more beneficial than others and there might be a negotiation strategy which is more likely to reach it. Lastly, as suggested, the stochastic characteristic function might not only capture the real probabilistic distributions but may be based on players' beliefs and expectations. In that case, sensitivity analysis of the proposed approaches providing insights into how potential errors in the distributions affect the optimal solutions might be useful for the decision makers.

References

Algaba, E., Fragnelli, V., and Sánchez-Soriano, J., editors (2019). *Handbook of the Shapley Value*. Chapman and Hall/CRC, New York.

Aumann, R. J. and Dreze, J. H. (1974). Cooperative games with coalition structures. *International Journal of Game Theory*, 3(4):217–237.

Aumann, R. J. and Maschler, M. (1964). The bargaining set for cooperative games. In Dresher, M., Shapley, L. S., and Tucker, A. W., editors, *Advances in Game Theory*, volume 52, pages 443–476. Princeton University Press, Princeton.

Basso, F., D'Amours, S., Rönnqvist, M., and Weintraub, A. (2019). A survey on obstacles and difficulties of practical implementation of horizontal collaboration in logistics. *International Transactions in Operational Research*, 26(3):775–793.

Belleflamme, P. (2000). Stable coalition structures with open membership and asymmetric firms. *Games and Economic Behavior*, 30(1):1–21.

Benati, S., López-Blázquez, F., and Puerto, J. (2019). A stochastic approach to approximate values in cooperative games. *European Journal of Operational Research*, 279(1):93–106.

Chalkiadakis, G., Elkind, E., Markakis, E., Polukarov, M., and Jennings, N. R. (2010). Cooperative games with overlapping coalitions. *Journal of Artificial Intelligence Research*, 39(1):179–216.

Fernández, F. R., Puerto, J., and Zafra, M. J. (2002). Cores of stochastic cooperative games with stochastic orders. *International Game Theory Review*, 4(3):265–280.

Frisk, M., Göthe-Lundgren, M., Jörnsten, K., and Rönnqvist, M. (2010). Cost allocation in collaborative forest transportation. *European Journal of Operational Research*, 205(2):448–458.

Granot, D. (1977). Cooperative games in stochastic characteristic function form. *Management Science*, 23(6):621–630.

Guajardo, M. and Rönnqvist, M. (2015). Operations research models for coalition structure in collaborative logistics. *European Journal of Operational Research*, 240(1):147–159.

Guajardo, M. and Rönnqvist, M. (2016). A review on cost allocation methods in collaborative transportation. *International Transactions in Operational Research*, 23(3):371–392.

Guajardo, M., Rönnqvist, M., Flisberg, P., and Frisk, M. (2018). Collaborative transportation with overlapping coalitions. *European Journal of Operational Research*, 271(1):238–249.

Hajduková, J. (2006). Coalition formation games: A survey. *International Game Theory Review*, 8(4):613–641.

Hart, S. and Kurz, M. (1983). Endogenous formation of coalitions. *Econometrica*, 51(4):1047–1064.

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press.

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton.

National Academies of Sciences, Engineering, and Medicine (2020). *Data Sharing Guidance for Public Transit Agencies - Now and in the Future*. The National Academies Press, Washington, DC.

Osicka, O., Guajardo, M., and van Oost, T. (2020). Cooperative game-theoretic features of cost sharing in location-routing. *International Transactions in Operational Research*, 27(4):2157–2183.

Özener, O. Ö. and Ergun, Ö. (2008). Allocating costs in a collaborative transportation procurement network. *Transportation Science*, 42(2):146–165.

Potters, J. A. M., Curiel, I. J., and Tijs, S. H. (1992). Traveling salesman games. *Mathematical Programming*, 53(1–3):199–211.

Rahwan, T., Michalak, T. P., Wooldridge, M., and Jennings, N. R. (2015). Coalition structure generation: A survey. *Artificial Intelligence*, 229:139–174.

Ray, D. and Vohra, R. (1999). A theory of endogenous coalition structures. *Games and Economic Behavior*, 26(2):286–336.

Sánchez-Soriano, J., López, M. A., and García-Jurado, I. (2001). On the core of transportation games. *Mathematical Social Sciences*, 41(2):215–225.

Shapley, L. S. (1953). A value for n-person games. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games*, volume 2, pages 307–317. Princeton University Press, Princeton.

Shapley, L. S. (1955). Markets as cooperative games. Technical Report P-629, The RAND Corporation.

Shehory, O. and Kraus, S. (1996). Formation of overlapping coalitions for precedence-ordered task-execution among autonomous agents. In *Proceedings of International Conference on Multi-Agent Systems, ICMAS-96*, pages 330–337.

Suijs, J. and Borm, P. (1999). Stochastic cooperative games: Superadditivity, convexity, and certainty equivalents. *Games and Economic Behavior*, 27(2):331–345.

Suijs, J., Borm, P., De Waegenaere, A., and Tijs, S. (1999). Cooperative games with stochastic payoffs. *European Journal of Operational Research*, 113(1):193–205.

Suijs, J., De Waegenaere, A., and Borm, P. (1998). Stochastic cooperative games in insurance. *Insurance: Mathematics and Economics*, 22(3):209–228.

Thomson, W. (2019). The Shapley value, a crown jewel of cooperative game theory. In Algaba, E., Fragnelli, V., and Sánchez-Soriano, J., editors, *Handbook of the Shapley Value*, pages 1–15. Chapman and Hall/CRC, New York.

Timmer, J., Borm, P., and Tijs, S. (2004). On three Shapley-like solutions for cooperative games with random payoffs. *International Journal of Game Theory*, 32(4):595–613.

Timmer, J., Borm, P., and Tijs, S. (2005). Convexity in stochastic cooperative situations. *International Game Theory Review*, 7(1):25–42.