COALITION FORMATION IN RIDESHARING WITH WALKING OPTIONS

Lucia Cipolina-Kun
School of Computer Science, Electrical and Electronic and Engineering Maths
University of Bristol. Bristol, UK
lucia.kun@bristol.ac.uk

Vahid Yazdanpanah, Enrico Gerding, Sebastian Stein
Electronics and Computer Science
University of Southampton. Southampton , UK
v.yazdanpanah@soton.ac.uk, {eg,ss2}@ecs.soton.ac.uk

ABSTRACT

In this work, we introduce a novel coalition formation mechanism for ridesharing services. Specifically, we extend the current literature to integrate walking options within the trip. This allows us to effectively account for a user's value of time when walking, which is not negligible. Additionally, we propose a cost allocation method that ensures proportionality for sharing the costs, i.e. those who walk more should pay less. We present a preliminary formal evaluation of the efficacy of our cost allocation method and discuss its desirable properties. Our study is a step towards developing smart mobility systems that recommend optimal ridesharing coalitions to users and suggest cost allocations that accounts for the marginal cost and contributions of each rider.

1 INTRODUCTION

Ridesharing services allow individuals with similar itineraries to share a car and split the cost of the trip. As a mobility option, ridesharing presents advantages on several fronts. From a social point of view, it contributes to lower congestion, bringing a positive externality to traffic Santos & Xavier (2015); Stiglic et al. (2015). From the user's standpoint, ridesharing is usually more affordable than travelling alone Li et al. (2021). For these reasons, ridesharing services have become a popular mobility option across big cities. In recent times, several commercial platforms have started to offer ridesharing services such as Uber[1], Lyft[2] and DiDi[3] as a sign of the changing landscape in mobility patterns. According to the *US Census Bureau* US Census Bureau (2020), by 2019, ridesharing represented approximately 8–11% of the transportation modality in Canada and the USA.

Inspired by the surge in ridesharing, we propose a mechanism to allocate riders in a cost-efficient way. We consider a setting where users meet at a common departure point and take a car to a single drop-off point, to where they walk to their own destinations. To account for the walking time, we introduce a mechanism to distribute the car cost, compensating riders for the inconvenience of walking. We believe it is important to acknowledge the walking time as a cost, since it is part of the user's *value of time* (VoT), and therefore, has a monetary value Mayet & Hansen (2000); Santos (2004).

The surge in ridesharing has motivated the production of academic studies on different related aspects. One of the main questions is how to combine different riders into the available set of cars. For this, there have been many approaches using game-theoretic solutions, for example the work by Bistaffa et al. (2015; 2017) uses cooperative game theory to find the optimal coalitions or riders. They propose the concept of *feasible coalitions* in which commuters are grouped according

[1] https://www.uber.com
[2] https://www.lyft.com
[3] http://www.didiglobal.com

to their proximity in social networks. This concept allows the reduction of the dimensionality of the search space; however, it requires a certain degree of information regarding the users' social network. Moreover, this work, like many others,Golpayegani & Clarke (2018); Ye et al. (2021); Lu & Quadrifoglio (2019) assumes that cars make multiple stops and riders are not required to walk, which is unrealistic in a ridesharing setting.

Focusing now on the cost allocation problem, the traditional literature borrows from game theoretic concepts to propose cost sharing mechanisms, for example, Yan et al. (2021) uses the VCG auction mechanism to achieve stable allocations, the work of Bistaffa et al. (2015) proposes the kernel as an allocation method and Lu & Quadrifoglio (2019) propose an algorithm to find the core. None of these methods include walking as a cost to consider in the allocation mechanism. Two works include some form of walking in the ridesharing problem. The paper by Stiglic et al. (2015) considers a ridesharing scheme where there is a single departure and drop-off point, with no additional stops or detours. However, their focus is on solving a bipartite matching problem between taxi cabs and riders. Their work does not factor in the walking time as a cost incurred by riders, and there is no monetary compensation for the walking time nor accounting for the different walking times that users might have. The second work by Kaan & Olinick (2013) considers a situation where employees drive from their homes to a common car park and board a company van to their workplace. The objective here is to determine the minimum van cost so that employees choose the company service over the private cars.

We expand the current literature by explicitly considering a more general setting in which riders are willing to walk: first from their home to a departure point, and then from the car drop-off point to their own destination. The main difference with the existing literature is in that we explicitly model the walking cost and compensate riders for their walking time. The example below depicts the main points of our methodology.

Introductory example. Consider a group of friends who want to share a ride after a night out, or colleagues who live close by and want to share a ride. In both cases, the riders agree to meet at a departure point and ride together to a drop-off point. Given a set of riders, and assuming an infinite number of cars, there are many possible allocations of riders into cars. Our premise is that individuals choose with whom to share a ride based on the cost of the trip, including riding and walking costs. We illustrate the interplay between walking and riding costs below.

As an example, consider Figure 1 below. On the left panel there are 3 individuals $\{A, B, C\}$ who choose to share a car. They walk from their individual origins $\{A_o, B_o, C_o\}$ to a common departure point **o** and then drive to a single stop **d**, from where they walk to their individual destinations $\{A_d, B_d, C_d\}$. This is just one example of a possible arrangement of riders, but there can be many others. For example, the right panel on Figure 1 depicts an alternative arrangement. In this case, $\{A, B\}$ ride-share together while $\{C\}$ travels alone (and therefore has no need for walking). Indeed, there are five possible combinations in which three individuals can be arranged. The formation of these possible travel arrangements, their cost optimality, and the fair cost distribution is the subject of this paper.

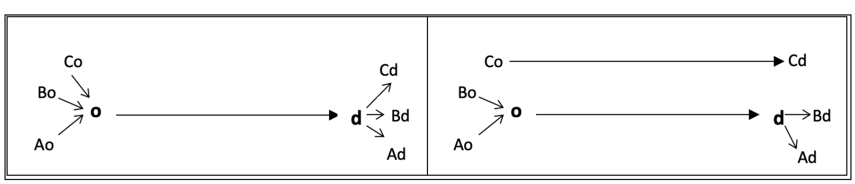

Figure 1: Example of two ridesharing formations for three individuals. The left panel shows three riders sharing a car and their respective walking distances, the right panel shows two individuals ridesharing and a third one opting to travel alone.

Our contributions are as follows: first, we propose a methodology based on game-theoretic guarantees for the cost-efficient formation of ridesharing groups. Secondly, we propose a cost-allocation method that compensates users for the walking time spent. To the best of our knowledge, we are the first to include the cost of walking into the ridesharing problem (in both source and destination) and factor this cost in when performing the cost allocation.

2 Formal Preliminaries

Following Rothe (2016) and Nisan et al. (2007), in cooperative games a set of $A = \{a_1, ..., a_n\}$ agents work together by forming coalitions, and take joint actions so as to maximize their utility. The goal is to find what is the best coalitional structure to form. A coalitional cost game can be described in terms of its *characteristic function*, which expresses the cost of each coalition. Intuitively, this is the cost that a subset of agents face and need to pay collectively.

Definition 1 *A coalitional cost game with transferable utilities (TU) is the tuple (A, C) where A is the set of agents and $C(S)$, is a characteristic function $C : 2^S \mapsto \mathbb{R}$ that returns the cost that a subset $S \subseteq A$ of agents face on their own regardless of what the remaining agents do (i.e., without any externality present).*

Given a coalitional cost game (A, C), the Coalition Structure Generation (CSG) problem focuses on generating a coalition structure (as a partitioning on the set of agents A) with desirable properties, e.g., those that yield a minimum cost.

Definition 2 *Given a coalitional cost game (A, C), a coalition structure $\mathbb{C} = \{S_1, \ldots, S_m\}$, is a partitioning on A, with $m \leq n$. That is, for arbitrary distinct $1 \leq k, l \leq m$, we have that $S_k \subset A$, $S_k \cap S_l = \varnothing (k \neq l)$, and $\cup_{k=1}^{m} S_k = A$. The set of all possible coalition structures is denoted by $\hat{\mathbb{C}}$.*

3 Ridesharing as a Coalitional Cost Game

A ridesharing game is a tuple (A, C, M) where $A = \{a_1, .., a_n\}$ is the set of riders, $C : 2^S \mapsto \mathbb{R}$ is the characteristic function, and $M : A \mapsto \mathbb{R}^2 \times \mathbb{R}^2$ is the *coordinate function* that yields the origin and destination coordinates for each rider e.g., $M(a_i) = \big((x_i^{origin}, y_i^{origin}), (x_i^{dest}, y_i^{dest})\big)$ are the origin and destination coordinates of a rider. For simplicity, we consider an unbounded supply of cars with up to four seats each. We formulate our problem in a game-theoretic setting where our goal is to construct partitions of A of up to four members (i.e., $|S_k| \leq 4, \forall k$) such that the *coalitional structure* obtained has the minimum cost across all the possible ones in $\hat{\mathbb{C}}$. The following sections cover how to calculate the coalitional cost and how the optimal coalition structure is obtained.

3.1 Cost Calculation

Our work considers two main costs, the individual walking cost and the car cost, which is shared by passengers sharing a car. A key input for the determination of both are the meeting points, which in turn, depend on the geometric median of the riders' coordinates. The main contribution of our cost calculation methodology is the addition of walking costs on top of the car cost. In a sense, the walking cost can be viewed as the *cost of entry* into a coalition, as it represents the time investment required to join a ridesharing coalition. The walking *distance* of each individual depends on the central origin and destination points, while the walking *cost* depends on the monetary value of time, as explained below.

Determination of the Central O-D Points. The first step in determining the walking time is to calculate the departure and destination points for all the the riders sharing a car. Our approach determines these points using the geometric median of the individuals' coordinates (as given by M) [4]. Once we have identified the departure and destination points, we assign a monetary value to the walking time.

Cost of Walking. In order to assign a monetary value to the time spent walking, we use the concept of value of time (VoT). We assume that the VoT can be expressed as a function proportional to the distance walked and this proportion increases at an increasing rate. Similar to Shrestha & Zolnik (2013), the justification for the *over-proportionality* of the walking cost follows from the idea of a threshold distance for reasonable walking, over which individuals see walking as a burden. To capture the marginal aversion to walking as a monetary cost, the VoT of walking from a to b for the individual a_i is modeled as a continuous and convex function as below.

[4]The geometric median of a set of points is defined as the point minimizing the sum of Euclidean distances to all points Cohen et al. (2016)

$$VoT_i(a,b) = dist(a,b) + \frac{1}{\alpha} \cdot dist(a,b) * \log_\alpha dist(a,b) \tag{1}$$

where α is an arbitrary parameter related to the function's convexity. Figure A.1 in the Appendix A illustrates the convexity of VoT_i against the simple Euclidean distance function.

Cost of Riding. The total cost of the car for a ridesharing coalition S is simply the product of the distance travelled from points a to b times a constant, accounting for the taxi fare: $P_{car}(S) = dist(a,b) \cdot K$.

Coalition Cost. For a set of riders $S \subseteq A$, we calculate the geometric median of their origin and destination coordinates [5], and derive the walking and car costs using the equations above. The total cost of the trip for this coalition (i.e. the characteristic function of the ridesharing game) is:

$$C(S) = \begin{cases} P_{car}(S), & |S| = 1 \\ P_{car}(S) + \sum_{i=1}^{|S|} VoT_i(S), & \text{otherwise} \end{cases} \tag{2}$$

where VoT_i is the total walking done by individual i (i.e. sum of walking at origin and at destination).

Lastly, from definition 2, the cost of a coalitional structure $\mathbb{C}$ is simply the total cost across its component coalitions:

$$\mathbb{C}_{cost}(\mathbb{C}) = \sum_{S \in \mathbb{C}} C(S) \tag{3}$$

Example. To illustrate the cost calculation, consider a given coalition S with only two individuals $\{a_1, a_2\}$. The rider's origin and destination coordinates are, respectively, $\left(x_i^{origin}, y_i^{origin}\right)$ and $\left(x_i^{dest}, y_i^{dest}\right)$ for $i = 1, 2$. The car's origin and destination points are calculated as the median of the individual's origin and destination coordinates.[6] In this example, the car departure point is $(x_o, y_o) = \left(\frac{1}{2}\sum_{i=1}^{2} x_i^{origin}, \frac{1}{2}\sum_{i=1}^{2} y_i^{origin}\right)$; and the car drop-off point is: $(x_d, y_d) = \left(\frac{1}{2}\sum_{i=1}^{2} x_i^{dest}, \frac{1}{2}\sum_{i=1}^{2} y_i^{dest}\right)$. The rideshare car travels: $dist\left((x_o, y_o), (x_d, y_d)\right)$, which is the input for P_{car}. Each rider walks at most two times, first from their individual origin to the meet-up point $dist\left((x_i^{origin}, y_i^{origin}), (x_o, y_o)\right)$ and then from the drop-off point to their destination $dist\left((x_i^{dest}, y_i^{dest}), (x_d, y_d)\right)$. Each of these walking distances are passed to equation 1[7].

4 Optimal Coalition Structure

Across all the possible coalitions of up to four individuals, our aim is to find the coalitional structure with minimal travel cost. Below we explain the steps performed.

4.1 Dimensionality reduction

The problem of searching for optimal structures is NP-hard because the size of different possible coalition structures is exponential in the number of coalition members, since there are 2^n possible coalitions for a set of n individuals. The common practice in the ridesharing literature is to reduce the dimensionality of the problem in a sensible way. Here, a popular technique is the use of DBSCAN Ester et al. (1996) to find proximity clusters among riders with similar origin and destination (see for example, the works of Samy & Elkorany (2018) and Syafira et al. (2021)). Following this approach, we use DBSCAN to reduce the partitions space by first grouping together individuals whose origin and destination points are closer (i.e., clustered structures). This allows us to narrow down the space of possible coalitions into a selection of *feasible coalitions*.

[5]for the geometric median, we used the Weiszfeld (1937) algorithm

[6]In this example, since $n = 2$, we use the fact that the geometric median matches the mean.

[7]In the particular case of two riders, both walk the same distance.

Definition 3 (Feasible Coalition) *In a ridesharing game (A, C, M), we say a coalition $S \subseteq A$ is ϵ-feasible if all of its members are interior points of a closed ball $N_\epsilon(p) = \{q \in \mathbb{R}^4 \mid dist(p, q) \leq \epsilon\}$, where the $dist(\cdot)$ function is the Euclidean distance of points in a four-dimensional space composed by (x_o, y_o, x_d, y_d) (as sources and destinations of any two riders in S), ϵ is an arbitrary radius in 4D, and p is an arbitrary point in this space.*

In Section 7 we develop a numerical example exemplifying the implementation of DBSCAN and the selection of ϵ.

4.2 Optimal Coalitional Structure Algorithm

After partitioning the space with DBSCAN, we calculate the set of all feasible coalitions structures of up to four riders on each cluster. We calculate their cost of riding following equation 2, then we aggregate them to obtain the cost of each of the possible coalitional structures as in equation 3. The objective is to find the combination of coalitions per cluster (i.e., the coalition structure) that minimizes this cost. Formally, we want the specific $\mathbb{C} \in \hat{\mathbb{C}}$ such that $\arg\min_{\mathbb{C}}\{\mathbb{C}_{cost}(\mathbb{C})\}$. If there exist more than one coalitional structures with the minimum cost, we (randomly) take one with the least number of coalitions.

Algorithm 1, summarizes the steps required to find the coalition structure with the minimum cost on each cluster. We refer to such a coalition structure as the *optimal coalition structure*. The algorithm starts by calculating all the possible partitions (i.e. coalition structures) within the elements of a cluster. For this, the function func_partitions takes in the list of individuals in a cluster and returns a list of all possible coalition structures with coalitions up to size four. Then the algorithm loops over all coalitions within a coalition structure to calculate equations 2 and 3. Lastly, it returns the minimum cost and its associated structure per cluster.

Algorithm 1: Determination of the socially optimal coalition structure on each cluster

input: set of all clusters, game components (A, C, M)
output: socially optimal coalition structure per cluster and its cost
foreach *cluster in set of clusters* **do**
    $\hat{\mathbb{C}}_{lst}$ ←func_partitions($a_i \in cluster$, 4) ; `// list of coalition structures`
    foreach *coalition structure* $\mathbb{C}$*, in* $\hat{\mathbb{C}}_{lst}$ **do**
        $\mathrm{C}_{cost}(C) \leftarrow 0$
        foreach *coalition* S *in* $\mathbb{C}$ **do**
            $\mathrm{C}_{cost}(C) \leftarrow \mathbb{C}_{cost}(C) + C(S)$
        end
    end
    return $\arg\min_{\mathbb{C}}\{\mathbb{C}_{cost}(\mathbb{C})\}$
end

5 Cost Allocation Problem

After obtaining the socially optimal coalition structure, the next question is how to split the price of the car across members of a coalition (i.e., riders sharing a car). This aspect is important, as it will determine whether individuals would prefer ridesharing rather than traveling alone, and thus a coalition can be formed.

Formally, the goal is to calculate a payoff vector (x) that is the share of the coalition cost allocated to rider a_i after joining coalition S as one of the components of the optimal coalition structure. For the price sharing phase,the challenge is that the walking part of the cost is an NTU (non-transferable utility) component and the only part of the cost that has transferable utilities is the cost of the car. According to Nisan et al. (2007), two common desirable properties are required over the payoff redistribution:

1. *Efficiency*. The entire cost of coalition S is split among the coalition members.
2. *Individual rationality*. The cost paid by a coalition member (walking plus car share) is at most equal to its singleton cost (i.e., riding alone).

Next we develop our methodology for solving the cost allocation problem.

Inversely Proportional Cost Allocation

The idea of the inversely-proportional cost allocation method is that those who walk proportionally less, pay a greater share of the price. Let $P_{car}(S)$ be the cost of the car to be distributed among members of coalition S. To allocate this cost we define the share of the payment as inversely proportional to the *proportion* of walking done by each member.

Let $w_i = \frac{VoT_i}{\sum_{i\in S} VoT_i}$ be the proportion of walking done by individual $a_i \in S$. Let w_i' be the proportion of the price paid by individual a_i. The goal is to find a proportional split such that the usual condition holds:

$$\sum_{i\in S} w_i' = 1 \tag{4}$$

As an additional condition, we want the w_i' to be inversely proportional to walking, i.e.:

$$w_i' = \frac{1}{w_i} \cdot \kappa \tag{5}$$

where the κ term is needed for equation 4 to hold. To determine κ we plug equation 5 into equation 4, obtaining $\kappa = \frac{1}{\sum \frac{1}{w_i}}$.

Finally, the share of the car paid by each member is: $c_i = P_{car}(S) \cdot w_i'$.

6 Properties

Below, we analyze the properties of the cost-allocation algorithm based on the inverse proportion of walking. We start by showing that the formula for the distribution of cost is efficient.

6.1 Efficiency and Fairness

Equation 4 guarantees that the inverse-walking payment allocation is efficient. Additionally, it is easy to see that the distribution in 5 has the following properties: (a) if two individuals have walked the same proportion, their payment allocation is equal (i.e. if $w_i = w_j$ then $w_i' = w_j'$) and also (b) if individual i has walked p times more (for example) than individual j, then $w_i' = \frac{1}{p} w_j'$.

This cost allocation method is fair in the sense that it compensates individuals for the cost for entering a coalition. Fairness in this sense would not be captured by an equal allocation or traditional metrics such as the Shapley value Roth (1988) . The formal results on desired properties of our approach are further described in Section 6.2.

6.2 Study on the Conditions for Individual Rationality

There are certain conditions that have to be met for the socially-optimal coalition to be individually rational. In short, the requirement is that the distance travelled by car in ridesharing should be less than the distance travelled by individual ride, to compensate for the cost of walking. This is obtained when the origin and destination of ridesharing individuals are close enough. Below we formalize this idea.

Let the cost of traveling alone be $c(\{a_i\}) = P_{car}(o,d)$, where $P_{car} = dist(o,d) \cdot K$ is the cost of the taxi for travelling from coordinate points o to d. Without loss of generality, assume the taxi fare $K = 1$. Let the cost of riding in coalition S for individual a_i be $c_i(S) = VoT_i(S) + \Psi_i(S)$, with: $\Psi_i(S) = P_{car}(o', d') \cdot w_i'(S)$, where $w_i'(\cdot)$ is as defined in equation 5 and (o', d') are the common meetup and drop-off points. We want to analyze under which conditions $c_i(S) \leq c(\{a_i\})$.

We first focus on comparing the distance travelled by the singleton car and the ride share, which is basically asking that $dist(o,d) \leq dist(o', d')$. This inequality depends on the calculation of the points (o', d'), these points are the result of taking the geometric median across the origin and destination coordinates of the individuals on the coalition. If individuals are too far apart the ridesharing scheme is no longer convenient. This distance is controlled by the ϵ parameter of the DBSCAN algorithm. In the same line, if individuals are close enough to each other, that also helps in reducing the walking distance term $VoT_i(S)$ and thus the coalitional cost.

7 Numerical Example

This section provides a numerical implementation of the theory previously described. Our methodology consists of three steps: *(a)* dimensionality reduction, *(b)* determination of the socially optimal coalitional value and *(c)* cost sharing among ridesharing participants. An implementation example is shown below.

Dimensionality reduction step. We start by simulating an initial set of 30 random points on a plane. Then, we perform DBSCAN clustering, for the choice of the $epsilon$ parameter we set it to be proportional to the dimensions of the space, as 0.1 of the total area. Results are shown below. Note that the clustering zones depicted in Figure 7 below are only indicative, as we actually performed a DBSCAN clustering in four dimensions (corresponding to the origin and destination coordinates combined).

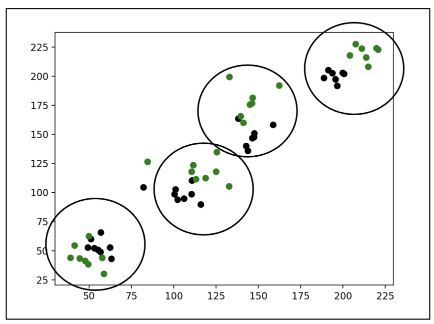

Figure 2: Origin and destination of commuters on the plane with clustering zones. Simplification of clusters in 4D space. The black dots correspond to individuals' origin and green dots to their destination.

Optimal Coalitional Structure Step. The next step is to calculate the walking and riding cost for each coalition structure, as in Equation 2, taking Equation 1 as input. For the α parameter it was selected according to the dimensions of the threshold value of 800 m suggested by Shrestha & Zolnik (2013). We used α = 10, which makes the convexity more pronounced at the chosen threshold value of 800 m.

The next step is to determine which is the socially optimal coalition structure among all feasible ones. This is achieved by following Algorithm 1, Table 1 shows an example composition of the socially optimal coalition structures per cluster. Individuals are arranged in cars of up to four riders such that the total cost of travel (walking time plus trip cost) is minimized. The column **Cost** represents the total cost of the coalitional structure (i.e., the sum of the cost of the car for each coalition plus the total walking cost for each individual on the cluster). The Figure A.2 in Appendix A shows a visual representation of the socially optimal coalition structure within each cluster.

7.1 Ensuring Efficiency and Fairness

Achieving the socially optimal coalitional structure is a desirable property at the aggregate level; however, at the individual level, coalitional costs should be allocated in a way that is individually rational. After the socially optimal coalition is found, we calculated the allocation of the car cost among its members. The graph below shows how much each agent saves by ridesharing compared to riding alone. In no case the individual rationality is violated. For this example, the average cost savings across all individuals is 63%, which means that, ridesharing produces a saving of more than 60 percent compared to traveling in a taxi alone. The median saving is 81%.

Table 1: Optimal coalition structure per cluster

Cluster	Optimal Coalition Structures	Cost
1	$\{\{a_0, a_6, a_{13}, a_{14}\}, \{a_7, a_{11}, a_{24}, a_{27}\}\}$	351.03
2	$\{\{a_3, a_8, a_{17}\}, \{a_1, a_{18}, a_{19}, a_{22}\}\}$	685.34
3	$\{\{a_4, a_5, a_{10}, a_{16}\}, \{a_2, a_{12}, a_{21}\}\}$	555.5
4	$\{\{a_9, a_{20}, a_{25}\}, \{a_{23}, a_{26}, a_{28}, a_{29}\}\}$	613.43
5	$\{\{a_{15}\}\}$	220.85

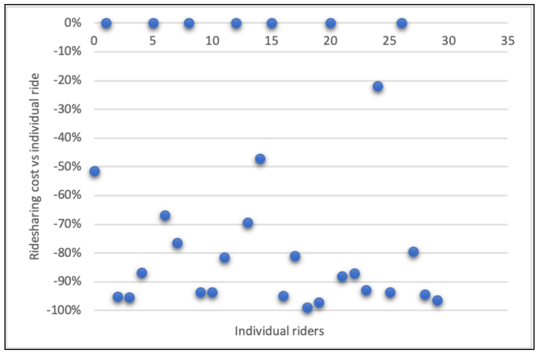

Figure 3: Cost savings of ridesharing versus travelling alone, using the inversely proportional cost allocation method. The horizontal axis represent the agent's a_i index and the vertical axis represents the savings percentage when ridesharing vs traveling alone.

8 Conclusions and Future Work

In this work we show that the socially-optimal ridesharing problem can be modelled as a coalition formation problem and solved within the framework of game theory. We proposed a methodology to include the cost of walking into the coalitional cost structure. Additionally, we proposed a methodology for the cost allocation among coalition members that is fair in the sense that it compensates members for the cost of walking, which is the cost of participation in each coalition. Our preliminary exploration of the results shows that the proposed method is computationally tractable and leads to results that are socially optimal.

Our proposed methodology is a centralized approach, having the following benefits. First, it preserves privacy for users of the system. Users communicate location and travel preferences to a trusted centralized system, without having to share personal information with others. Additionally, each agent does not need to have observability of all other agents, and communicates only with the central system.

For future work we would like to explore and compare additional cost allocation methods, and test stability concepts like the core or Shapley value. Additionally, we could explore different cost functions where, for example, the cost of the car is dependent on the occupancy level. One extension could also be to include multiple stops in the car route.

Data Access Statement

Data sharing is not applicable to this article because no new data were created or analyzed in this study.

Acknowledgments

We thank the anonymous reviewers for their incisive comments that were most useful in revising this paper. Additionally, the authors would like to thank Ayalvadi Ganesh and Ian Gemp for their comments. This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) through a Turing AI Fellowship (EP/V022067/1) on Citizen-Centric AI Systems (`https://ccais.soton.ac.uk/`). Lucia Cipolina-Kun is funded by British Telecom.

REFERENCES

F. Bistaffa, A. Farinelli, and Ramchurn S. D. Sharing rides with friends: A coalition formation algorithm for ridesharing. *Proceedings of the National Conference on Artificial Intelligence*, 1: 608–614, 2015.

F. Bistaffa, A. Farinelli, G. Chalkiadakis, and S. D. Ramchurn. A cooperative game-theoretic approach to the social ridesharing problem. *Artificial Intelligence*, 246:86–117, 2017. ISSN 00043702. doi: 10.1016/j.artint.2017.02.004. URL `http://dx.doi.org/10.1016/j.artint.2017.02.004`.

M.B. Cohen, Y. Lee, T. Yin, G. Miller, J. Pachocki, and Sidford A. Geometric median in nearly linear time. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450341325. doi: 10.1145/2897518.2897647. URL `https://doi.org/10.1145/2897518.2897647`.

M. Ester, H.P Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, Oregon, USA*. AAAI Press, 1996.

F. Golpayegani and S. Clarke. Co-ride: Collaborative preference-based taxi-sharing and taxi-dispatch. *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*, 2018-Novem:864–871, 2018. ISSN 10823409. doi: 10.1109/ICTAI.2018.00135.

L. Kaan and E. V. Olinick. The Vanpool Assignment Problem: Optimization models and solution algorithms. *Computers and Industrial Engineering*, 66(1):24–40, 2013. ISSN 03608352. doi: 10.1016/j.cie.2013.05.020. URL `http://dx.doi.org/10.1016/j.cie.2013.05.020`.

Y. Li, J. Wan, R. Chen, J. Xu, X. Fu, H. Gu, P. Lv, and M. Xu. Top-k Vehicle Matching in Social Ridesharing: A Price-Aware Approach. *IEEE Transactions on Knowledge and Data Engineering*, 33(3):1251–1263, 2021. ISSN 15582191. doi: 10.1109/TKDE.2019.2937031.

Wei Lu and Luca Quadrifoglio. Fair cost allocation for ridesharing services – modeling, mathematical programming and an algorithm to find the nucleolus. *Transportation Research Part B: Methodological*, 121:41–55, 2019. ISSN 01912615. doi: 10.1016/j.trb.2019.01.001. URL `https://doi.org/10.1016/j.trb.2019.01.001`.

J. Mayet and M. Hansen. Congestion pricing with continuously distributed value of time. *Journal of Transport Economics and Policy Journal of Transport Economics and Policy*, 34(3):359–369, 2000.

N. Nisan, T. Roughgarden, T. Tardos, and Vazirani V. *Algorithmic Game Theory*. Cambridge University Press, 2007.

ed. Roth, A. E. *The Shapley Value: Essays in Honor of Lloyd S. Shapley*. Cambridge University Press, 1988.

J. Rothe. *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*. Springer, 2016.

M. Samy and A. M. Elkorany. Using semantic features for enhancing car pooling system. *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018*, pp. 702–708, 2018. doi: 10.1109/ASONAM.2018.8508713.

D. O. Santos and E. C. Xavier. Taxi and ride sharing: A dynamic dial-a-ride problem with money as an incentive. *Expert Systems with Applications*, 42(19):6728–6737, 2015.

G. Santos. Road pricing theory and evidence. *Research in Transportation Economics*, 9(3):2–308, 2004.

R. M. Shrestha and E. J. Zolnik. Eliminating bus stops: Evaluating changes in operations, emissions and coverage. *Journal of Public Transportation*, 16:153–175, 2013. doi: http://doi.org/10.5038/2375-0901.16.2.8.

M. Stiglic, N. Agatz, M. Savelsbergh, and M. Gradisar. The benefits of meeting points in ride-sharing systems. *Transportation Research Part B: Methodological*, 82:36–53, 2015. ISSN 01912615. doi: 10.1016/j.trb.2015.07.025. URL `http://dx.doi.org/10.1016/j.trb.2015.07.025`.

I. G. Syafira, Y. Satria, and H. Burhan. Application of Density-Based Spatial Clustering of Application with Noise to Optimize Matching Problems in Ridesharing for Maximize Total Distance Proximity Index. *Journal of Physics: Conference Series*, 1821(1), 2021. ISSN 17426596. doi: 10.1088/1742-6596/1821/1/012015.

US Census Bureau. American community survey, 2019. `https://www.census.gov/programs-surveys/acs/data.html`, 2020. Online; accessed 25 February 2022.

E. Weiszfeld. Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tohoku Mathematical Journal*, 1937.

P. Yan, C. Lee, C. Chu, C. Chen, and Z. Luo. Matching and pricing in ride-sharing: Optimality, stability, and financial sustainability. *Omega (United Kingdom)*, 102:102351, 2021. ISSN 03050483. doi: 10.1016/j.omega.2020.102351. URL `https://doi.org/10.1016/j.omega.2020.102351`.

Y. Ye, L. Zheng, Y. Chen, and Longquan L. *Discovering Stable Ride-Sharing Groups for Commuting Private Car Using Spatio-Temporal Semantic Similarity*, volume 12816 LNAI. Springer International Publishing, 2021. ISBN 9783030821463. doi: 10.1007/978-3-030-82147-0{_}52. URL `http://dx.doi.org/10.1007/978-3-030-82147-0_52`.

A Appendix

This section provides additional content to improve the clarity on some of the concepts presented in the main body.

A.1 Additional graph for the walking cost

The graph below provides a visualization of the weighted walking function against a linear distance function. The graph depicts the convexity of the walking cost function as described by Equation 1.

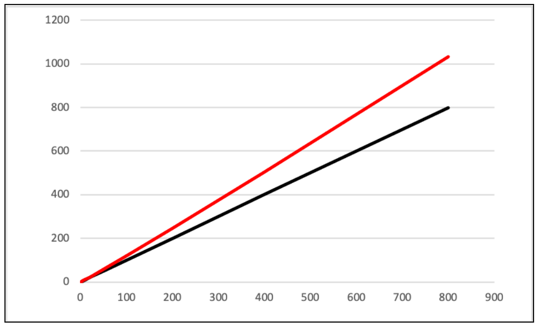

Figure 4: Comparison between the weighted walking (in red) function and linear distance (in black)

A.2 Additional graph for the Optimal Coalitions

The Figure below shows an example of the socially optimal coalitions within each cluster. The different colors show how individuals are sub divided into different ridesharing structures (i.e. coalitional structures).

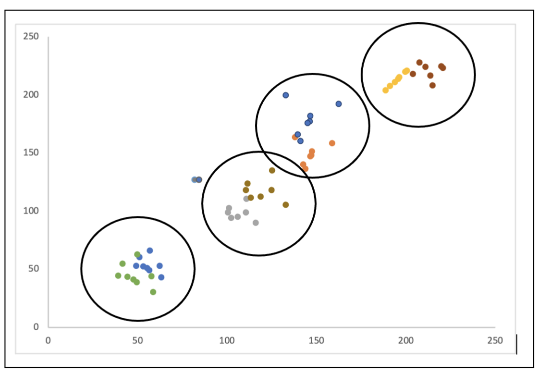

Figure 5: Socially optimal coalition structures per clusters using a visual simplification of clusters in 4D.