Carpooling with Transfers and Travel Time Uncertainty

Patrick Stokkink*[1], André de Palma[2], and Nikolas Geroliminis[1]

[1]'Ecole Polytechnique Fédérale de Lausanne (EPFL), Urban Transport Systems Laboratory (LUTS), Switzerland
[2]THEMA, CY Cergy Paris University, France

Short summary

Carpooling is known to have lower CO2 emissions compared to driving individually. One of the limitations of carpooling is that matched drivers and passengers need to have similar itineraries, or their generalized costs will be high. By allowing a single transfer at a designated transfer hub, their itineraries need to be only partially similar. We allow for transfers within the carpooling system and between carpooling and public transport. Thereby, we include travel time uncertainty to evaluate its effect on carpooling with transfers. We model the ride-matching problem with transfers and travel time uncertainty as a two-stage stochastic programming problem. The results indicate that a single transfer hub can already reduce the average generalized cost of passengers by 15%. When travel times are uncertain, commuters tend to find a match that performs relatively well in every traffic situation, rather than one that performs well for only one scenario.
Keywords: Carpooling, Ride-Matching, Ride-Sharing, Transfers, Uncertain Travel Time

1 Introduction

Transport accounts for a large share of global CO2 emissions. According to IEA (2022) cars and buses for passenger transport constitute 45.1% of transport emissions, including freight transport. Statistics gathered by the Center for Sustainable Systems, University of Michigan (2021) show that in 2019 the average car occupancy in the United States was 1.5 and in 2017 24% of the U.S. households had 3 or more vehicles. Carpooling as an alternative to traveling alone by car is known to reduce CO2 emissions directly and a well-functioning carpooling system is expected to reduce car ownership in the long term.

One of the main limitations of direct carpooling is that a pairing of drivers and passengers needs to be found with similar itineraries (in space and time). By allowing transfers, a larger set of potential matches is available for drivers and passengers since the itineraries only need to be partially similar, as they spend only a part of their trip together. Thereby, passengers are allowed to transfer to and from public transport. A graphic illustration is displayed in Figure 1. Carpooling or ride-matching models with transfers have been considered before by, among others, Herbawi & Weber (2012); Masoud & Jayakrishnan (2017); Huang et al. (2018); Lu et al. (2020).

Despite their benefit, transfers may impose additional difficulties when travel time is uncertain. Passengers and drivers may carry on their first-leg delays to the second leg, thereby influencing their match, or they may fully miss their connection. In the presence of uncertainty, transfers can make carpooling with transfers less appealing due to their effect on tardiness and uncomfortable and unanticipated waiting times. Long et al. (2018) consider a bi-objective ride-sharing-matching model under travel-time uncertainty. They consider delay and schedule delay penalties that may change according to this uncertainty.

We consider a matching framework for carpooling with transfers and uncertainty in travel time. We allow for transfers within the carpooling system and between carpooling and public transport. By considering uncertainty in travel time, schedule delay penalties of potential matches can depend on this uncertainty, similar to Long et al. (2018). In our problem, potential matches may be infeasible for some uncertain scenarios which can therefore affect the optimal matching. We model the matching problem with transfers and public transport as a deterministic integer programming problem. We extend this model to a two-stage stochastic programming problem where travel time uncertainty forms the division between first- and second-stage decisions. Matches on the first leg of the trip are made under uncertainty, whereas second-leg matches are made afterward when information on travel time is gathered.

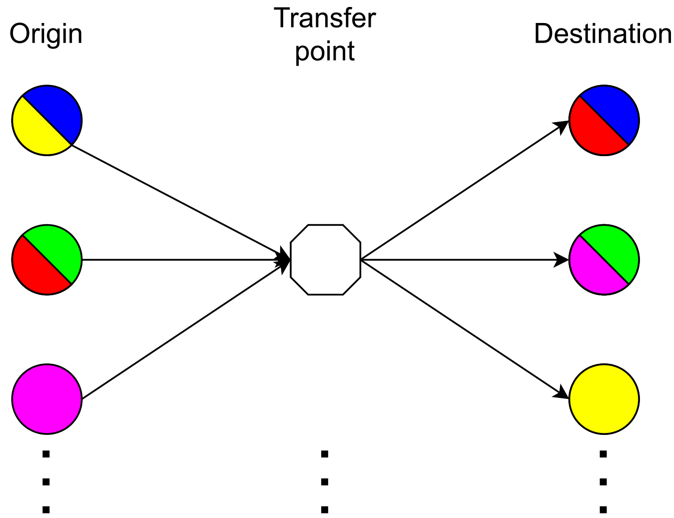

Figure 1: Graphic illustration of transfers

2 Methodology

We consider a matching problem of carpooling drivers and passengers. Every individual has a specific value of time, scheduling preferences, and desired arrival time. Unlike ride-hailing services, drivers are relatively inflexible in the system we consider. Drivers are only willing to perform pickups at their own origin or at a dedicated transfer hub along their route and are only willing to perform drop-offs at their own destination or at a dedicated transfer hub along their route. Thereby, drivers determine the departure time only based on their own schedule delay preferences and ignore those of the passengers. The reason for this is that in a complex system where drivers take multiple passengers and passengers take multiple drivers, coordinating jointly optimal departure times can be extremely difficult both theoretically and in practice. Contrary to the common carpooling approach where passengers spend their full trip with a single driver, we allow passengers to transfer at designated transfer hubs. These transfer hubs have connections to public transport services and allow for transfers between two carpooling drivers. We only allow one transfer to limit the incomfort of transfers and include the inconvenience of transferring and waiting in the generalized cost formulation.

Determinisitic Formulation

The deterministic matching approach is based on a set of predefined passenger paths. Let I be the set of passengers, J the set of drivers, and H the set of transfer hubs. Given that the possible number of matches for passengers is polynomial, we can generate all possible paths in advance. We let the drivers set the departure times, such that the costs of passenger paths are independent of each other. Then we only need to consider that drivers may take multiple passengers at the same time, but only on the same route, and that they can pick up new passengers at the transfer point.

Every individual has an origin o_i, a destination d_i and a desired arrival time t_i^*. Let K be the set of passenger paths and let $e_{ik} = 1$ if passenger path k corresponds to passenger i, and 0 otherwise. The cost of passenger path k is denoted by c_k, which only contains the cost for passengers. By definition, drivers do not incur any scheduling delay costs nor make a detour. Therefore, we assume they are fully compensated for the inconvenience of sharing their car and their costs are not included in the objective function. Let decision variable $x_k = 1$ if passenger path k is chosen and 0 otherwise. Let q_j be the capacity of driver j, that is, the number of passengers driver j is able to transport at the same time. We distinguish between the following three kinds of trips:

- Direct trip: $a_{jk}^0 = 1$ if driver j contributes to passenger path k through a direct trip.

- First leg of indirect trip: $a_{jk}^{1h} = 1$ if driver j contributes to passenger path k through a first-leg trip to transfer hub h.
- Second leg of indirect trip: $a_{jk}^{2h} = 1$ if driver j contributes to passenger path k through a second-leg trip from transfer hub h.

We use decision variable y_{jh} to define through which transfer hub driver j is going. This allows to formulate the deterministic matching problem as follows:

$$\textbf{(P1)} \; \textit{minimize} \sum_{k \in K} c_k x_k \tag{1a}$$

$$\sum_{k \in K} e_{ik} x_k = 1 \qquad \forall i \in I \tag{1b}$$

$$\sum_{k \in K} a_{jk}^{0} x_k \leq q_j \left(1 - \sum_{h \in H} y_{jh}\right) \qquad \forall j \in J \tag{1c}$$

$$\sum_{k \in K} a_{jk}^{1h} x_k \leq q_j y_{jh} \qquad \forall j \in J, h \in H \tag{1d}$$

$$\sum_{k \in K} a_{jk}^{2h} x_k \leq q_j y_{jh} \qquad \forall j \in J, h \in H \tag{1e}$$

$$\sum_{h \in H} y_{jh} \leq 1 \qquad \forall j \in J \tag{1f}$$

$$x_k \in B \qquad \forall k \in K \tag{1g}$$

$$y_{jh} \in B \qquad \forall j \in J, h \in H \tag{1h}$$

The objective (1a) is to minimize the cost of all matches. Every passenger needs to be matched to exactly one driver, which is enforced by Constraints (1b). Feasibility of the solution from the perspective of a driver is enforced through Constraints (1c) - (1e). The feasibility of the solution from the perspective of a passenger is enforced directly on the set of paths K. That is, the set K only contains paths that are feasible for a passenger. On every leg, a driver $j \in J$ may have at most q_j passengers in their car, which is enforced jointly by Constraints (1c), (1d) and (1e). A driver may either serve passengers directly from their origin to their destination or through a transfer point, but not both. Constraints (1f) ensure that a driver only makes a stop at one transfer point.

In the remainder of this section, we discuss in detail the three types of paths that we consider and the corresponding parameter values in **P1**. For this, we let α be the value of time spent in a car, β the penalty for every unit of time an individual is early, and γ the penalty for every unit of time an individual is late. Waiting time is penalized by α^{wait} and the value of time spent in public transport is defined as α^{pt}. Travel time between o and d is defined as $tt(o, d)$. For the sake of notation, these parameters are all homogeneous, but the formulation allows for heterogeneous parameter values.

Public Transport Paths
Every passenger $i \in I$ has the option to take public transport instead of carpooling. Public transport has a fixed cost per unit of time such that $c_k = \alpha^{pt} tt(o_i, d_i)$.

Direct Carpool Paths
For every passenger $i \in I$, a direct match can be found with a driver $j \in J$ if $o_i = o_j$ and $d_i = d_j$. As the driver selects the departure time to minimize her own cost, the arrival time at the final destination is equal to the desired arrival time of the driver, possibly imposing schedule delay costs on the passenger. For a match between $i \in I$ and $j \in J$, $e_{ik} = 1$, $a_{jk}^{0} = 1$ and all other parameters are equal to 0. The cost of this direct match are as follows:

$$c_k = \alpha tt(o_i, d_i) + \beta(t_i^* - t_j^*)^+ + \gamma(t_j^* - t_i^*)^+ \tag{2}$$

Indirect Carpool Paths
For the sake of notation, we define $t_j^*(h)$ as the desired arrival time of driver j at transfer hub h if he travels through that hub. This is simply computed as $t_j^*(h) = t_j^* - tt(h, d_j)$.

We first consider the path where the passenger takes public transport on one of the two legs.

For a path where only the first leg is a carpooling leg, if passenger $i \in I$ and driver $j \in J$ are matched with a transfer at hub $h \in H$, they must share their origin $o_i = o_j$ and hub h must be on the path of driver j. For a path where only the second leg is a carpooling leg, if passenger $i \in I$ and driver $j \in J$ are matched with a transfer at hub $h \in H$, they must share their destination $d_i = d_j$ and hub h must be on the path of driver j. Given the fixed cost of public transport and the fact that we do not consider a schedule for public transport, the cost for first-leg or second-leg carpooling are highly similar and given as follows:

$$c_k = \alpha^{pt} tt(h, d_i) + \alpha tt(o_i, h) + \beta[t_i^* - t_j^*(h) - tt(h, d_i)]^+ + \gamma[t_j^*(h) + tt(h, d_i) - t_i^*]^+ \tag{3}$$

$$c_k = \alpha^{pt} tt(o_i, h) + \alpha tt(h, d_i) + \beta[t_i^* - t_j^*(h) - tt(h, d_i)]^+ + \gamma[t_j^*(h) + tt(h, d_i) - t_i^*]^+ \tag{4}$$

For paths that consist of two carpooling legs, we consider a passenger $i \in I$ and two drivers $j_1, j_2 \in J$ where j_1 takes i on the first leg and j_2 takes i on the second leg with a transfer at transfer hub h. Similar to before, this is only feasible if $o_i = o_{j_1}$, $d_i = d_{j_2}$ and h is both on the path of j_1 and j_2. Thereby, $t_{j_1}^*(h) \leq t_{j_2}^*(h)$ to ensure that the passenger is dropped off at the transfer hub before the scheduled pickup. The cost for the passenger is then defined as follows:

$$c_k = \alpha[tt(o_i, h) + tt(h, d_i)] + \alpha^{wait}[t_{j_2}^*(h) - t_{j_1}^*(h)] + \beta[t_i^* - t_{j_2}^*(h) - tt(h, d_i)]^+ + \gamma[t_{j_2}^*(h) + tt(h, d_i) - t_i^*]^+ \tag{5}$$

Stochastic Formulation

To allow for uncertainty in travel times, we adapt our formulation to a two-stage stochastic programming problem. The matching is determined a-priori but may be adapted based on the observed state of the system (i.e. the travel times). The first leg of every driver and passenger is fixed and cannot be altered after observing the state. This can be seen as a contract between the driver and the passenger. The second leg, however, may be changed after observing the state of the system. We assume the state is observed after the first leg has been fixed (i.e., the contract has been negotiated) but before the second leg commenced. With modern technologies, commuters are aware of traffic conditions during or shortly before their trip. Let Ω be the uncertainty set and $\omega \in \Omega$ a realization of the uncertain travel times. We also refer to such a realization of the uncertain travel times as a scenario.

All variables and parameters are altered to be dependent on the scenario ω. This means that in stead of x_k we use $x_k(\omega)$ and in stead of y_{jh} we use $y_{jh}(\omega)$. In addition to this, the cost of path k also depends on the scenario as it influences travel time and may even make paths infeasible. Therefore, we change c_k to $c_k(\omega)$, where $c_k(\omega) = \infty$ if it path k is infeasible for scenario ω. This may happen, for example, when the passenger arrives at the transfer point after their driver has already departed because of a delay. We denote $p(\omega)$ the probability of scenario ω occuring, such that $p(\omega) \geq 0$, $\sum_{\omega \in \Omega} p(\omega) = 1$.

We enforce that all first-stage decisions are the same for all scenarios. Specifically, if driver j is involved in path k for scenario ω that is a direct match from origin to destination, he must commit to the same direct path in any other scenario. Therefore, as direct paths only have one leg, the chosen paths are identical for every scenario and are used to minimize the expected cost. This is enforced through

$$a_{jk}^0 x_k(\omega) = a_{jk}^0 x_k(\omega') \qquad \forall j \in J, k \in K, \omega, \omega' \in \Omega \tag{6}$$

For an indirect path, only the first leg is fixed. In this case, the full path need not be the same, as long as the same driver goes to the same hub in both scenarios. In addition to this, we impose that both paths need to correspond to the same passenger. To enforce this, we use the following set of constraints. By enforcing the matched driver-passenger pair as well as the hub at which the passenger is dropped to be equal across scenarios, we guarantee that first-leg matches are fixed in advance.

$$\sum_{k \in K} e_{ik} a_{jk}^{1h} x_k(\omega) = \sum_{k \in K} e_{ik} a_{jk}^{1h} x_k(\omega') \qquad \forall i \in I, j \in J, h \in H, \omega, \omega' \in \Omega \tag{7}$$

The full formulation of the stochastic programming problem, to which we refer as **P2**, is given as follows:

$$\textbf{(P2)}\ \mathit{minimize} \sum_{\omega \in \Omega} \sum_{k \in K} p(\omega) c_k(\omega) x_k(\omega) \tag{8a}$$

$$\sum_{k \in K} e_{ik} x_k(\omega) = 1 \qquad \forall i \in I, \omega \in \Omega \tag{8b}$$

$$\sum_{k \in K} a^0_{jk} x_k(\omega) \le q_j \left(1 - \sum_{h \in H} y_{jh}(\omega)\right) \qquad \forall j \in J, \omega \in \Omega \tag{8c}$$

$$\sum_{k \in K} a^{1h}_{jk} x_k(\omega) \le q_j y_{jh}(\omega) \qquad \forall j \in J, h \in H, \omega \in \Omega \tag{8d}$$

$$\sum_{k \in K} a^{2h}_{jk} x_k(\omega) \le q_j y_{jh}(\omega) \qquad \forall j \in J, h \in H, \omega \in \Omega \tag{8e}$$

$$\sum_{h \in H} y_{jh}(\omega) \le 1 \qquad \forall j \in J, \omega \in \Omega \tag{8f}$$

$$a^0_{jk} x_k(\omega) = a^0_{jk} x_k(\omega') \qquad \forall j \in J, k \in K, \omega, \omega' \in \Omega \tag{8g}$$

$$\sum_{k \in K} e_{ik} a^{1h}_{jk} x_k(\omega) = \sum_{k \in K} e_{ik} a^{1h}_{jk} x_k(\omega') \qquad \forall i \in I, j \in J, h \in H, \omega, \omega' \in \Omega \tag{8h}$$

$$x_k \in B \qquad \forall k \in K \tag{8i}$$

$$y_{jh} \in B \qquad \forall j \in J, h \in H \tag{8j}$$

The Objective (8a) is to minimize the expected costs, which is a linear function weighted by the probability of each scenario occurring. Constraints (8b) to (8f) are the same as in **(P1)**, but extended them with the scenario dependency ω. Constraints (8g) ensure that the same direct paths are chosen for every scenario. Constraints (8h) enforce the first leg of drivers to be the same on indirect paths and that they carry the same passenger.

3 Results and discussion

We evaluate our model on a circular city consisting of 33 nodes, as depicted in Figure 2. Every passenger and driver has an origin and destination at one of the 33 nodes. Origins are more likely to be in the suburbs (the outer rings) whereas destinations are more likely to be in the city center. Transfer hubs can be at any of the nodes in the network. Drivers can perform a pick-up or a drop-off at one of the transfer hubs, but only if the hub is on their shortest path. We consider 1000 drivers and 500 passengers. Desired arrival times are drawn from a truncated normal distribution with a mean at 8:00 and a standard deviation of 1 hour. The distribution is truncated such that we only allow desired arrival times between 6:30 and 9:30.

The parameter settings are homogeneous among the entire population and are defined as follows. The value of time spent in a car α is equal to 6.4[$\$/h$]. Earliness and lateness are penalized with β and γ equal to 3.9[$\$/h$] and 15.21[$\$/h$] respectively. The value of time in public transport α^{pt} is higher and is set equal to 11.0[$\$/h$]. In addition to this, public transport has a fixed cost of 2.0$ per trip to compensate for waiting times. Waiting time is penalized by α^{wait} which is equal to 13.5[$\$/h$] such that $\beta < \alpha < \alpha^{pt} < \alpha^{wait} < \gamma$, consistent with the literature.

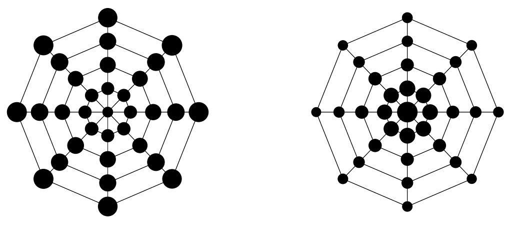

Figure 2: Circular city with the distribution of origins (left) and destinations (right)

Influence of transfers on cost and mode choice

We consider the influence transfers make on the average cost per individual in the deterministic system. The results are displayed in Figure 3 where the left-hand panel displays the modal split of passengers and the right-hand panel displays the composition of the average cost of passengers. Clearly, when there are no transfer hubs, the only possible mode choices are direct carpooling and public transport. By opening transfer hubs, a modal shift to the two modes that use transfers is observed. Especially the number of passengers carpooling on two separate legs increases. The reason for this is that by using a transfer, more options exist for matching to someone with the same destination and a similar desired arrival time, at the cost of waiting at the transfer point. The number of direct matches may be limited as the origin and destination of the passenger and driver need to be identical and the desired arrival time needs to be somewhat similar.

By using a single transfer hub in the center of the network, the average cost decreases from 14\$ to 12\$ ($\approx$ 15%). Increasing the number of transfer hubs allows a further decrease in the average cost, but not nearly as substantial as for the first hub in the center. When all 9 hubs are opened, the average cost decreases to 11\$ ($\approx$ 22%). We emphasize that carpooling benefits from economies of scale when the number of commuters increases. The reason for this is that the matching opportunities increase, which decreases the expected cost of a match.

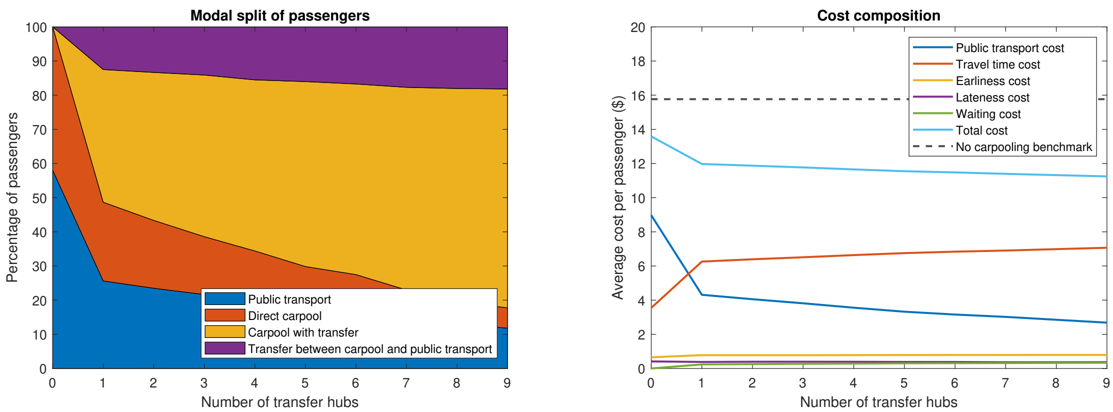

Figure 3: Statistics for a varying number of transfer hubs

Distribution of passengers by desired arrival time

We evaluate the distribution of passengers by desired arrival time and the mode they use to commute. We consider the deterministic system with only one hub in the center of the network. The results are displayed in Figure 4 where the left-hand panel displays the number of passengers using a mode and the right-hand panel displays the proportion of passengers with a specific desired arrival time using a mode.

The proportion of passengers using public transport is the highest in the tails. The reason for this is that the number of potential matches with identical origins and destinations and similar desired arrival times is low since the number of individuals here is rather low. This effect is more apparent for passengers with an early desired arrival time. When these passengers match to a driver, it is highly likely that the desired arrival time of the driver is later than that of the passenger, and therefore the passenger will suffer from lateness. As lateness is penalized heavier than earliness, the effect is more apparent at the start of the morning commute than it is at the end. At the peak of the rush hour, the number of carpoolers is the highest. We see a slight skewness towards later desired arrival times, which follows the same reasoning as stated before.

Stochastic Programming Results

We analyze the results of the stochastic programming problem, using three scenarios ($|\Omega| = 3$), with all travel times at 100%, 125%, and 150% of the free-flow travel times, respectively. A selection of 50 passengers is made to display the matching, where a bar represents a match to a driver. The color of the bar, as well as the index displayed on the bar, identifies the driver. The first bar

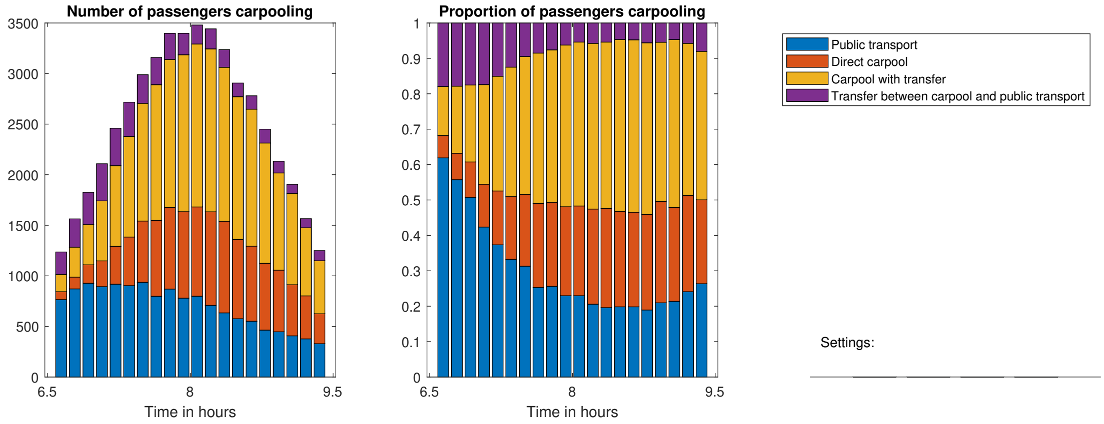

Figure 4: Distribution of passengers by desired arrival time and mode

displays a first-leg match, the second bar displays a second-leg match and the full bar represents a direct match. When no bar is given on a leg, the passenger uses public transport on this leg. The results are given in Figure 5.

We observe that in line with Constraints (8g) and (8h) the direct matches and first-leg matches are the same for all scenarios. Approximately 75% of the matches are identical across all three scenarios, whereas for the remaining 25% the passenger either changes mode or driver on the second part of the trip. We compare the result to the wait-and-see benchmark, where Constraints (8g) and (8h) are relaxed. This is displayed in Figure 6. For the wait-and-see benchmark only 50% of the matches are identical. The wait-and-see benchmark shows that the higher the travel times (for scenarios 2 and 3) the fewer matches are optimal. For the recourse problem, however, a middle ground needs to be found since the matching cannot be fully adapted to the exact traffic situation. Therefore, we observe slightly fewer matches in scenario 1 for the recourse problem, whereas we observe significantly more matches for scenario 3 compared to the wait-and-see problem.

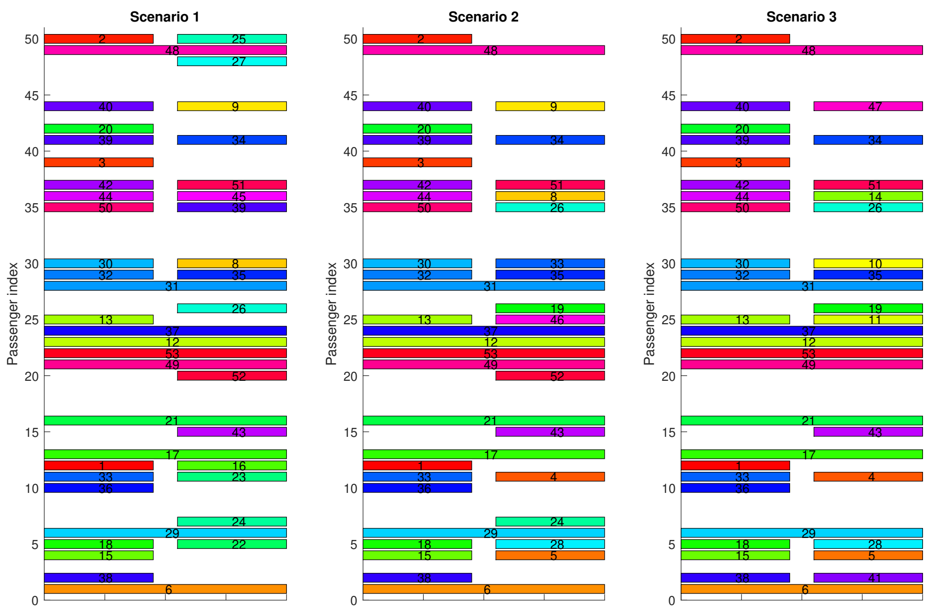

Figure 5: Gantt chart of matching for the recourse problem.

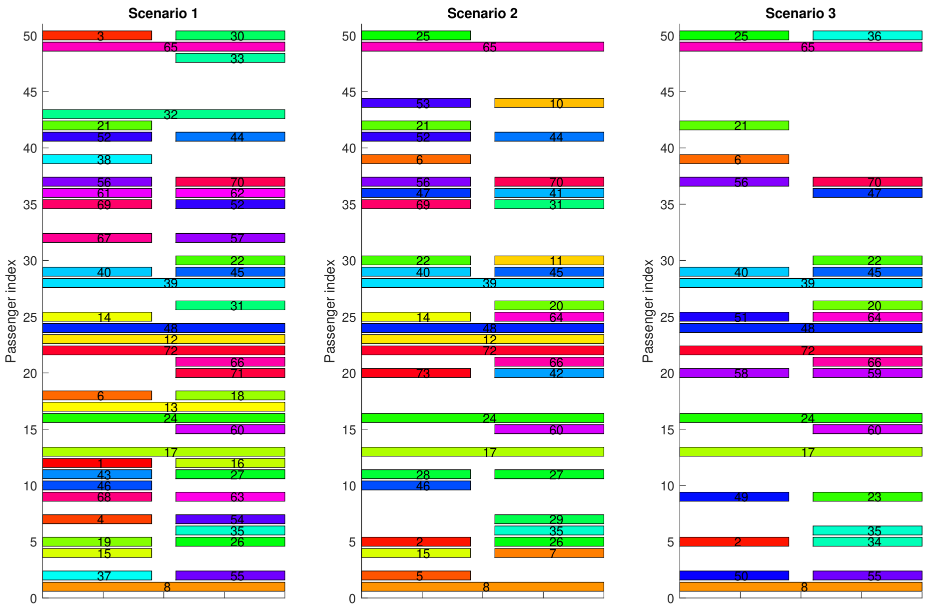

Figure 6: Gantt chart of matching for the wait-and-see benchmark problem.

4 Conclusions

We conclude that transfers can significantly reduce the average costs of commuters. Carpooling with transfers partially replaces direct carpooling as well as public transport as a mode of transport. It is especially beneficial for commuters during the peak of the commute, whereas commuters in the tails are usually better off taking public transport. In the case of stochastic travel times, we observe a large share of the commuters stick to the same mode and match for each scenario. However, 25% of the commuters may change their mode or match. We also observe that due to uncertainty, the number of carpoolers generally decreases.

References

Center for Sustainable Systems, University of Michigan. (2021). *Center for Sustainable Systems, University of Michigan. 2021. "Personal Transportation Factsheet." Pub. No. CSS01-07.*

Herbawi, W., & Weber, M. (2012). The ridematching problem with time windows in dynamic ridesharing: A model and a genetic algorithm. In *2012 ieee congress on evolutionary computation* (pp. 1–8).

Huang, H., Bucher, D., Kissling, J., Weibel, R., & Raubal, M. (2018). Multimodal route planning with public transport and carpooling. *IEEE Transactions on Intelligent Transportation Systems*, *20*(9), 3513–3525.

IEA. (2022). *Transport sector CO2 emissions by mode in the Sustainable Development Scenario, 2000-2030, IEA, Paris https://www.iea.org/data-and-statistics/charts/transport-sector-co2-emissions-by-mode-in-the-sustainable-development-scenario-2000-2030, IEA. Licence: CC BY 4.0.*

Long, J., Tan, W., Szeto, W., & Li, Y. (2018). Ride-sharing with travel time uncertainty. *Transportation Research Part B: Methodological*, *118*, 143–171.

Lu, W., Liu, L., Wang, F., Zhou, X., & Hu, G. (2020). Two-phase optimization model for ride-sharing with transfers in short-notice evacuations. *Transportation research part C: emerging technologies*, *114*, 272–296.

Masoud, N., & Jayakrishnan, R. (2017). A decomposition algorithm to solve the multi-hop peer-to-peer ride-matching problem. *Transportation Research Part B: Methodological*, *99*, 1–29.