Cumulative constraint problem for assessing a ride sharing system

Vincent Armant, John Horan, Nahid Mabub, Kenneth N. Brown

Insight Centre for Data Analytics, Department of Computer Science,
University College Cork, Ireland

Abstract. Ride-sharing schemes attempt to reduce road traffic by matching prospective passengers to drivers with spare seats in their cars. To be successful, such schemes require a critical mass of drivers and passengers. Based on a history of advertised schedules and agreed shared trips, we observed an imbalance number of feasible matches between drivers and passengers. We assess the potential benefits of persuading existing drivers to become passengers by solving a cumulative constraint problem. We demonstrate that flexible participation has the potential to reduce the number of unmatched participants by up to 60%.

1 Introduction

Road traffic is one of the main generators of carbon emissions, and traffic congestion is a significant contributor to pollution around major cities and urban areas. Partly motivated by these issues, there has been a recently strong growth in ride-sharing schemes (e.g. Blabla car, Carma, Lyft, Sidecar, Uber), where participants post details of intended trips, and the system then proposes possible matches between drivers and prospective passengers. As more matches are agreed, the number of car journeys decreases, and the total driven distance also decreases, helping to reduce congestion, emissions and energy consumption. Increasing participation in such schemes is thus considered both as a benefit for society and a commercial objective for the system operators. However, users who receive few offers, or who are given offers that are a poor match for their travel plans, are unlikely to continue with such system. Therefore, there is a need to assess the performance of the current matching schemes, identify ways in which performance could be improved, and assess the improvements that could be gained. To do this, we employ data analytics to infer constraints on possible matches, and to assess current performance. We then use the inferred constraints to build optimisation models and to evaluate proposed improvements.

Specifically, (i) we use shortest path routing algorithms and time constraints on the departure and arrival time windows for drivers to infer the bipartite graph representing the feasible matches between potential riders and prospective drivers. (ii) From this graph we show an imbalance between drivers and passengers that may be hampering participation in the scheme; and (iii) we propose and evaluate the potential of persuading drivers to be flexible in their roles in the scheme, showing a reduction in unmatched participants of up to 60%.

2 Related work

The dial-a-ride problem has been long studied in the OR community [6]. Dial-a-ride typically assumes a single vehicle, picking up and dropping off riders at specified locations within time windows, although multiple vehicle problems have also been studied [5, 4]. The dial-a-ride drivers have no journey requirements of their own. For ride-sharing schemes [7], both the drivers and the riders have their own objectives. Specific schemes vary as to whether the drivers move to the riders locations or the riders move to and from the driver routes, and whether or not drivers take single or multiple riders on a trip. One extension includes participants known as *shifters*, who may either drive or ride as a rider [1]. Armant *et al.* [3] also include shifters, but also assume that each pure rider who is not served in the matching has a probability of driving on their own, included as a penalty in the objective function. Computing an optimal matching is hard [2], and the complexity increases as the number of shifters increases. Kamar and Horvitz [8] model the problem as one of collaborative planning, where agents must balance competing goals. Yousaf *et al.* [12] model the problem as multi source-destination path planning, with a wide range of competing objectives including privacy and incentives. Schilde *et al.* [10] and Manna and Prestwich [9] consider stochastic problems, in which trip requests arrive during the execution of the solution, using scenario-based methods to minimise expected delays or unserved requests. Simonin and O'Sullivan [11] focus on the matching problem, assuming an input graph of all feasible pairings, and establish the complexity of a number of two variations, showing that in some cases polynomial time solutions are possible.

3 Ride sharing optimisation model

To describe the trip schedules and the parameters inferred from the history of advertised trip schedules of four regions during a period of 6 months, we introduce the following notations. D denotes the set of possible drivers, R the set of possible riders, and $U = D \cup R$ the set of all users. A trip schedule is a tuple $ts_u = (et_u^{start}, lt_u^{dest}, l_d^{start}, l_d^{dest})$ describing user u's inferred earliest start time et_u^{start}, latest arrival time lt_u^{dest}, start location l_u^{start}, and destination l_u^{dest}. $TS = \{ts_{u_1}, \ldots, ts_{u_n}\}$ denotes the set of user trip schedules sent to the system.

To infer the time and geographical constraints, we use Open Street Map data to deduce minimal path distances and times between two locations. $L = \{l_1, \ldots, l_n\}$ denotes the set of road node locations identified by their GPS coordinates. A path $\pi = (l_i, \ldots, l_j)$ is an ordered list of locations, and $time(\pi)$ (resp. $dist(\pi)$) returns the driving path time (resp distance) for π. The path $\pi^*_{l_i,l_j}$ (resp. $\pi^\diamond_{l_i,l_j}$) denotes a minimal time (resp. distance) path from l_i to l_j. For a driver trip schedule ts_d, π_d denotes the inferred driver path from start l_d^{start} to destination l_d^{dest}. For a rider trip schedule ts_r, m_r^{pick} denotes the inferred maximal path distance r is willing to walk from his intended start l_r^{start} to a pick-up location l_r^{pick} on the driver path π_d. Similarly m_r^{drop} denotes the inferred maximal path distance the rider is willing to walk from a drop-off location l_r^{drop} to his destination l_r^{dest}.

Given the above notations, we define the feasible matches relaying both on the users' inferred path constraints and the users' inferred time constraints.

Definition 1 (inferred feasible ride match). *A driver's trip schedule and a rider's trip schedule, ts_d and ts_r, represent an inferred feasible ride match if:*

1. *their inferred time windows tw_d, tw_r are consistent with the rider's pick-up and drop-off time:*
   (a) $lt_d^{dest} - et_r^{start} > time(\pi^*_{l^{pick},l^{drop}})$*, the time interval between the driver latest arrival and the rider earliest start is greater than the fastest path from the rider's inferred pick-up to his inferred drop-off, or,*
   (b) $lt_r^{dest} - et_d^{start} > time(\pi^*_{l^{pick},l^{drop}})$*, the time interval between the earliest driver start and the latest rider arrival is greater than the fastest path from the rider inferred pick-up to the inferred drop-off.*
2. *The expected driving path intersects the rider's possible pick-up and drop-off points.*
   (a) $dist(\pi^{\diamond}_{l_r^{start},\pi_d}) < m_r^{pick}$*, the shortest path distance between the rider intended start and the expected driver path is lower than the maximal distance for the rider's pick-up.*
   (b) $dist(\pi^{\diamond}_{l_r^{start},\pi_d}) < m_r^{drop}$*, the shortest path distance between the rider intended destination and the expected driver path is lower than the maximal distance for the rider's drop-off.*

For each studied region, we iteratively check if each pair of trip schedules is a feasible ride match and build a bipartite graph $G = (TSD, TSR, E)$ s.t. $TSD \subseteq TS$ is a set of driver trip schedules, $TSR \subseteq TS$ is a set of rider trip schedules and every edge $(ts_d, ts_r) \in E$ is a feasible ride match. G is the input parameter of the constraint programming model we build to assess the potential of a ride-sharing scheme. Each feasible match (ts_d, ts_r) in G is associated to a ride share trip y_{ts_d,ts_r} encoded as a collection of decision variables s.t.:

$y_{ts_d,ts_r}.start$ represents the pick-up time of r,
$y_{ts_d,ts_r}.end$ denotes a the drop-off time of r,
$y_{ts_d,ts_r}.duration$ is the time duration of the rideshare trip,
$y_{ts_d,ts_r}.reqSeats$ represents the number of seats required by r,
$y_{ts_d,ts_r}.presence$ denotes presence of the ride share trip in the optimal solution.

We model a served rider using x_{ts_r} s.t. x_{ts_r} equal 1 when the rider is allocated to exactly one of the feasible share rides y_{ts_d,ts_r}. To assess the potential of a ride-sharing scheme, our objective is to maximize:

$$\sum_{(ts_d,ts_r)\in E} x_{ts_r} \tag{1}$$

subject to:

$$y_{ts_d,ts_r}.start \geq max(t_d^{early}, t_r^{early}),\ \ \forall (ts_d, ts_r) \in E \tag{2}$$

$$y_{ts_d,ts_r}.end \leq min(t_d^{latest}, t_r^{latest}),\ \ \forall (ts_d, ts_r) \in E \tag{3}$$

$$y_{ts_d,ts_r}.duration = y_{ts_d,ts_r}.end - y_{ts_d,ts_r}.start,\ \ \forall (ts_d, ts_r) \in E \tag{4}$$

$$y_{ts_d,ts_r}.duration \geq \pi *_{l^{start},l^{dest}}\ ,\ \ \forall (ts_d, ts_r) \in E \tag{5}$$

$$y_{ts_d,ts_r}.reqSeats = \begin{cases} 1 \text{ if } y_{ts_d,ts_r}.presence = 1 \text{ and} \\ \qquad y_{ts_d,ts_r}.start \leq i \leq y_{ts_d,ts_r}.end \\ 0 \text{ otherwise}, \forall i \in \mathbb{N}, \forall (ts_d, ts_r) \in E \end{cases} \tag{6}$$

$$CUMULATIVE(\{y_{ts_d,ts_r}\}, nbSeats_d, \leq),\ \ \forall ts_d \in TSD \tag{7}$$

$$ALTERNATIVE(x_{ts_r}, \{y_{ts_d,ts_r} | (ts_d, ts_r) \in E\}),\ \ \forall ts_r \in TSR \tag{8}$$

$$(x_{ts_c}.presence \Rightarrow y_{tr_c,ts_r}.presence),\ \ \forall (ts_d, ts_r) \in E \tag{9}$$

The aim is to maximize the total number of served riders (1). The constraints (2) force each rideshare trip to start after the earliest rider start and the earliest driver start. Similarly, the constraints (3) force each rideshare trip to end before the latest rider arrival and the latest driver arrival. The duration of the rideshare trip is the difference between the end and the start (4) and it is greater than the rider shortest path (5). For a ride share trip, $reqSeats$ corresponds to a number of occupied seats in a driver's car for the decided trip duration if the ride share trip is chosen in a solution and is equal to 0 otherwise (6). The cumulative constraints (7) restrict each driver car occupancy to not exceed the number of available seats at any moment of the trip. The alternative constraints (8) enforce that at most one y_{ts_d,ts_r} rideshare trip is chosen. In the successful case of the rider rideshare trip x_{ts_r} is equal to the chosen rideshare y_{ts_d,ts_r} otherwise the rider is not chosen. The constraints (9) state that a shifter assigned to be a rider does not drive.

4 Assessing Ride Sharing Scheme

To assess the potential of the ride-sharing scheme, we use the inferred CP model to compute the maximum number of assignments of riders to drivers' cars. The analysis and experiments have been run using on a set of trip schedules collected during a period of one year. We solved the constraint problems on a machine having 2 processors of 2.5GHz, 12 cores, 64 GB of memory, using the CP Optimizer solver of IBM and either stopped the search at an optimal gap of 5% or retained the best solution found after a time limit of 1 hour. In the next table, we compare the number of matched users found in G (heuristic FM), with number of matched users found when persuading some drivers to become passenger (heuristic FMDS). This flexible heuristic has been envisaged after noticing a significant imbalance among the participants. In the table below, the low ratio R/D for FM means many drivers will be unmatched, and thus will drive with empty seats. In addition, frequent failed attempts to find a match are likely to deter those users from participating. The current optimisation model prioritises riders, and thus some drivers may have multiple passengers. The results of running these flexible models FMDS allows us to match significantly more participants. The number of unmatched participants drops by a factor of 1.33 in the worst case (region 2) and by a factor of 2.5 in the best case (region 3). We conclude that, where there is a participant imbalance, the focus of the ride sharing scheme operators should be to persuade drivers to be flexible in their roles.

region 1	users	ratio R/D	matched riders	matched drivers	% unmatch	region 2	users	ratio R/D	matched riders	matched drivers	% unmatch
FM	838	0.54	94	111	75.54	FM	512	0.66	40	66	79.30
FMDS		1.54	281	148	48.81	FMDS		1.66	127	75	60.55
region 3						region 4					
FM	1578	0.72	284	282	64.13	FM	4250	0.83	712	680	67.25
FMDS		1.72	930	257	24.78	FMDS		1.83	1890	736	38.20

In the next table, we compare the CPU time for building the feasible match graph G, rewriting the graph into a cumulative constraint problem and solving the problem per region and per hypothesis. Both the building time and the solving increase rapidly and reach 6 hours or hit the time limit when the number of trip schedules growth. Even if the modeling time remains below 3 secs, and the FM heauristic is solved fast, the global time performance is too slow to envisage real-time answer.

CPU Time	region 1 FM	region 1 FMDS	region 2 FM	region 2 FMDS	region 3 FM	region 3 FMDS	region 4 FM	region 4 FMDS
building G	2min30s	5min31s	3min	6min56s	1h4min4s	2h2min55s	3h44min16s	6h40min45s
CP modeling	157ms	287ms	79ms	214ms	1s204ms	2s770ms	2s410ms	3s226ms
CP solving	1s235ms	> 1h	587ms	3min18s527ms	13s448ms	> 1h	14s179ms	> 1h

5 Conclusion

Ride-sharing is a rapidly growing practice for reducing the number of cars on the road in urban regions. Successful ride sharing schemes require committed users, and they in turn require the scheme to provide them with feasible ride matches in real-time. In current systems, the emphasis has been on the real-time requirement rather than the feasibility of the matches. We have developed a model which uses route planning and time windows to describe feasible matches as a constraint satisfaction problem, and the ultimate goal of the ride-matching scheme as constraint optimisation. By applying the model to the data sets of advertised trips, we identify the errors in the current heuristics, and find an imbalance among participants in the ride sharing schemes. We consider the benefits that might be obtained if drivers can be persuade to switch roles and act as passengers, and by re-running the optimisation model we show that there is potential to reduce the number of unmatched participants by up to 60%. Such flexible switching would have a societal benefit, and encourage sustained user participation. Future work will focus on validating the hypothesis through field trial with users in the scheme, and on developing real-time responses respecting the constraints on feasible matches.

Bibliography

[1] Agatz, N., Erera, A., Savelsbergh, M., Wang, X.: Optimization for dynamic ride-sharing: A review. European Journal of Operational Research 223(2), 295 – 303 (2012)

[2] Agatz, N.A., Erera, A.L., Savelsbergh, M.W., Wang, X.: Dynamic ride-sharing: A simulation study in metro atlanta. Transportation Research Part B: Methodological 45(9), 1450 – 1464 (2011)

[3] Armant, V., Brown, K.N.: Minimizing the driving distance in ride sharing systems. In: ICTAI 2014, Proc 26th IEEE International Conference on Tools with Artificial Intelligence (2014)

[4] Attanasio, A., Cordeau, J.F., Ghiani, G., Laporte, G.: Parallel tabu search heuristics for the dynamic multi-vehicle dial-a-ride problem. Parallel Comput. 30(3), 377–387 (Mar 2004)

[5] Berbeglia, G., Cordeau, J.F., Laporte, G.: Dynamic pickup and delivery problems. European journal of operational research 202(1), 8–15 (2010)

[6] Cordeau, J.F., Laporte, G.: The dial-a-ride problem: models and algorithms. Annals of Operations Research 153(1), 29–46 (2007)

[7] Furuhata, M., Dessouky, M., Ordóñez, F., Brunet, M.E., Wang, X., Koenig, S.: Ridesharing: The state-of-the-art and future directions. Transportation Research Part B: Methodological 57(C), 28–46 (2013), http://ideas.repec.org/a/eee/transb/v57y2013icp28-46.html

[8] Kamar, E., Horvitz, E.: Collaboration and shared plans in the open world: Studies of ridesharing. In: Proceedings of the 21st International Jont Conference on Artifical Intelligence. pp. 187–194. IJCAI'09, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2009)

[9] Manna, C., Prestwich, S.: Online stochastic planning for taxi and ridesharing. In: 26th IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2014, Limassol, Cyprus, November 10-12, 2014. pp. 906–913 (2014)

[10] Schilde, M., Doerner, K.F., Hartl, R.F.: Metaheuristics for the dynamic stochastic dial-a-ride problem with expected return transports. Comput. Oper. Res. 38(12), 1719–1730 (Dec 2011)

[11] Simonin, G., O'Sullivan, B.: Optimisation for the ride-sharing problem: a complexity-based approach. In: ECAI 2014 - 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic - Including Prestigious Applications of Intelligent Systems (PAIS 2014). pp. 831–836 (2014)

[12] Yousaf, J., Li, J., Chen, L., Tang, J., Dai, X., Du, J.: Ride-sharing: A multi source-destination path planning approach. In: Thielscher, M., Zhang, D. (eds.) Australasian Conference on Artificial Intelligence. Lecture Notes in Computer Science, vol. 7691, pp. 815–826. Springer (2012)