Hyper pooling private trips into high occupancy transit like attractive shared rides

Rafał Kucharski & Oded Cats

The size of the solution space associated with the trip-matching problem has made the search for high-order ride-pooling prohibitive. We introduce hyper-pooled rides along with a method to identify them within urban demand patterns. Travellers of hyper-pooled rides walk to common pick-up points, travel with a shared vehicle along a sequence of stops and are dropped off at stops from which they walk to their destinations. While closely resembling classical mass transit, hyper-pooled rides are purely demand-driven, with itineraries (stop locations, sequences, timings) optimised for all co-travellers. For 2000 trips in Amsterdam the algorithm generated 40 hyper-pooled rides transporting 225 travellers. They would require 52.5 vehicle hours to travel solo, whereas in the hyper-pooled multi-stop rides, it is reduced sixfold to 9 vehicle hours only. This efficiency gain is made possible by achieving an average occupancy of 5.8 (and a maximum of 14) while remaining attractive for all co-travellers.

Ride-pooling offers a promising alternative in the context of urban mobility, being an intermediate mode between private rides and mass transit. However, its full potential has arguably not been yet realised, neither in practice nor in theory. Consequently, the scalability limits of pooled services in urban mobility remain unknown. In this study, we explore ways to increase the efficiency of ride-pooling systems, namely by means of increasing service occupancy and reducing vehicle mileage of ride-pooling systems while remaining attractive for travellers.

Ride-pooling companies primarily offer door-to-door services, picking up and dropping off off travellers along partially shared routes. Unlike public transit services, common pick-up and drop-off points, i.e. (virtual) stops, are rarely introduced. Furthermore, the majority of ride-pooling algorithms seek to maximise operational efficiency by focusing on real-time operations while disregarding the attractiveness of the yielded services, thereby implicitly assuming that users are captive to the service. Passenger-oriented solutions which focus on maximising users’ utility are scarce. Similarly, offline planning tools which could potentially unlock the full potential of bundling requests are limited.

We introduce a novel kind of pooled rides which is hereby denominated as hyper-pool rides. Travellers of hyper-pooled rides walk to common pick-up points, travel by means of a shared vehicle along a sequence of stops and are dropped off at stops from which they walk to their destinations. To identify attractive hyper-pooled rides, a sequence of bundling operations is conducted, resulting in moving successively from ‘private rides’ through ‘door-to-door pooled rides’ to arrive at ‘compact stop-to-stop rides’, and demonstrates the benefits of such rides in improving ride-pooling efficiency.

In particular, we are interested in rides which can be attractively composed using common pick-up and drop-off points. These compact stop-to-stop pooled rides are instrumental in introducing the ultimate level of pooling. Subsequently, multiple such stop-to-stop pooled rides may be bundled so as to compose a hyper-pooled ride, where travellers are picked up at common stop points, and the vehicle serves intermediate stops en-route. Our findings demonstrate that for such rides, occupancy increases substantially, improving thus the efficiency of pooled services and in some instances even starts resembling public transport operations while offering an on-demand service. Figure 1 illustrates ten selected hyper-pooled rides in Amsterdam which were yielded by employing the proposed algorithm.

The proposed bottom-up, hierarchical approach allows addressing the combinatorial increase of the associated search space consisting of all possible permutations of travel requests, which makes exact methods impractical. The central challenge to the ride-pooling problem pertains to its computational complexity - the number of feasible groups of travellers grows combinatorially, and even for relatively small problems, the search space quickly exceeds reasonable processing capabilities. An exact matching algorithm for solving the ride-sharing problem presented by Kucharski and Cats (2020) in is able to attain a sufficiently confined search space. This was achieved by (i) leveraging on user utility formulations so as to explicitly limit the consideration set to attractive pooling combinations only, and; (ii) exploiting properties of the so-called shareability graphs to effectively reduce the search space in the hierarchical searches. This study leverages the same principles to further explore higher degrees of pooled rides, and thereby devise ride-pooling solutions which contribute to improved efficiency and

---

1Faculty of Mathematics and Computer Science, Jagiellonian University, Krakow, Poland. 2Delft University of Technology, Delft, the Netherlands. e-mail: rafal.kucharski@uj.edu.pl

https://doi.org/10.1038/s44333-024-00006-4
sustainability objectives, i.e. by increasing the occupancy (reducing the vehicle mileage needed for transporting passengers). In the experiments reported by1, door-to-door pooled rides of degree (number of co-travellers) greater than five, which already yielded a search space of size 10^20, were seldom obtained. Conversely, with the two new levels of compact pooling proposed here, rides composed of up to 14 travellers, which would have otherwise required handling a search space of size 10^57, are identified. Hence, an exhaustive search is rendered impossible (calculated as a number of 14-element subsets of a 2000-element set with every possible sequence of 14 origins and destinations: \( \binom{2000}{14} \cdot 14! \)).

The potential of ride-pooling to improve urban mobility has triggered a bursting stream of research2–6, which led to an algorithmic breakthrough in the seminal works of Santi et al. (2014)7 and Alonso-Mora et al. (2017)8. First, the so-called shareability network was introduced along with a methodology to match travellers in pairs7. This network representation was then further exploited with a comprehensive algorithm to efficiently match incoming requests with available vehicles by sequentially adding new co-travellers into vehicles with empty seats and adjusting their routes8. By applying cut-offs on maximal detours and delays, the pooled rides are acceptable for travellers and by minimising the service costs (vehicle hours) when solving the matching problem, the solution becomes optimal for the supply side (platform and/or drivers).

Recent developments in ride-pooling research constitute a series of improvements to the original solutions, including: algorithmic—to reduce computational complexity9–11; heuristic—to narrow problem search space12,13; spatial decomposition—to reduce problem size14; or demand prediction—to better anticipate future requests15,16.

In parallel, behavioural research has made advancements in gaining a better understanding of users’ so-called willingness to share17–20. The discomfort associated with sharing has been quantified and can be now included in classical utility formulas within the random utility modelling framework. Either added as an alternative specific constant (fixed penalty21) or travel-time multiplier22. Furthermore, the extent of actual detours and delays of pooling were estimated from empirical data23–25, and the inherent variability of pooling travel and wait time has been incorporated into users’ choice models to capture the impacts of ride-pooling reliability26.

The ride-pooling problem suffers from combinatorial growth of search space, yet when rides are restricted to mutually attractive ones only, the search space implodes for practical problems and can be searched exhaustively without having to use heuristics. This is made possible by reformulating the fixed time-windows of a maximal acceptable detour and delay into utility-based compensation formulas as proposed and demonstrated in our previous study ‘Exact Matching of Attractive Shared Rides—ExMAS’, on which we build here. Such, utility-driven approaches, not only put the service user in the centre but also significantly reduce the search space. Other related user-centric works include accounting for individual customers’ pooling benefits and comparing it to existing provider-centred pooling mechanisms27, focus on the sensitivity of the results to user preferences in terms of the level of service (time to be served and excess trip time)28, or introduce willingness to share as a compensatory cost function at the individual passenger level to explicitly include the utility of pooled travellers29.

Fig. 1 | Ten selected hyper-pooled rides in Amsterdam. Each colour denotes a separate multi-stop pooled ride, small dots denote origins and destinations and larger dots denote (virtual) stops—linked with a walking path marked with lighter lines. The degree of rides (number of travellers) ranges between 4 (orange and brown rides) and 12 (grey and blue rides) and has been selected by all co-travellers (i.e. are mutually attractive). To our knowledge, identifying attractive pooled rides of such degree (number of travellers) has not been realised insofar. We can see direct short rides (e.g. brown in the central part) as well as rides spanning through the whole city (grey) and more curved yet attractive ones (orange). Occupancy levels often exceed four and vehicle hours are reduced five-fold when compared to serving all travellers by means of private rides, making hyper-pool rides effective and sustainable while remaining attractive for travellers. ©https://www.openstreetmap.org/copyright.

https://doi.org/10.1038/s44333-024-00006-4

Article
Ride-pooling can be considered as an intermediate mode of travel positioned between private ride-hailing and public transport, yet in practice—for example, in terms of its occupancy level—it still remains much closer to the former. There is considerable evidence to suggest that ride-hailing contributes to congestion. In contrast, a number of studies argue that ride-pooling can potentially become complementary to public transport. The critical differences between currently offered pooled services and public transport pertain to fares and travel times. Pooled services are typically more convenient: door-to-door, hence involving no transfers; yet remain significantly more expensive.

One way to bridge this gap is the introduction of so-called pick-up and drop-off points (PUDOs) - which introduce access/egress walk into ride-pooling. Such services have been offered, for example, by Uber (Uber-ExpressPool), MOIA and ViaVan (see a relevant market review by Foljanty (2021) ). From the operator’s perspective, stops are introduced to minimise operating costs (5% vehicle-kilometres reduction in Munich) within acceptable access/egress walking distances. Experiments with various PUDO spacing and density revealed an average ride occupancy increase by 20% and mileage decrease by 27%. Solutions to the stop-location problem are inspired by pickup delivery locations or NP-hard multi-meeting-point route search. Notably, in this study, we formulate the problem so that it can be searched exhaustively.

Research on ride-pooling shares several key similarities with on-demand public transport (for a recent review of the latter, see ref. 43) that are worth highlighting here. The mathematical complexity is similarly challenging, with similar supply-side and demand-side constraints. In both public transport and ride-pooling a critical mass is needed to take off and economies of scale share similarities. A virtuous cycle results from a positive feedback loop between service improvements, reduced travel times and increased demand, equivalent to the Mohring Effect. In the case of ride-pooling this process is governed by the improved matching of users. At the same time, in the short-run, increased demand leads to higher vehicle loads and longer detours (similar to crowding in PT). The key difference is the bottom-up approach, while on-demand public transport typically departs from pre-defined fixed lines and timetables and dynamically adjusts them to the demand, here we start from the disaggregated demand and hierarchically build rides of increasing aggregation until we reach higher level occupancies.

Several recent studies have made initial steps in contributing toward bridging between these two approaches. Integrated public transport and flexible on-demand transport passenger assignment models developed using agent-based simulation models. Schlüter et al. (2021) demonstrated that DRT systems are able to reduce the economic and environmental costs of transportation between rural and urban areas. For medium cities, demand-responsive services minimise detours, waiting times and walking distances experienced by passengers while allowing for a higher shareability and efficiency of the service. The potential efficiency gains that can be made by bundling travel demand in the wide spectrum, which goes beyond forming shared rides on the one hand and without yet having to rely on introducing flexible variants to pre-defined service lines remains largely unknown.

Hyper-pool is an analytic, offline, utility-based algorithm. For a given disaggregated travel demand provided in the form of individual trip requests known in advance, attractive pooled rides of increasing levels of aggregation (as illustrated in Fig. 2) are successively identified. The main objective is to increase service occupancy levels while remaining attractive to users. The algorithm relies on explicit utility formulas that trade-off any detour, delay and discomfort induced by successive kinds of pooling against the discount offered for sharing the ride. The utility formulas are user-specific and assess the attractiveness of four pooling levels. Utility is calculated for each traveller-ride pair taking based on user-specific behavioural parameters (e.g. value of time, willingness to share, ride-specific attributes (ride, waiting and walking times) and service settings (pricing).

First, all feasible door-to-door ride-pooling rides, which can be composed for the input demand pattern provided in the form of individual trip requests are identified using ExMAS. In the next stage, these door-to-door ride-pooling rides are searched to identify candidates for compact stop-to-stop rides. In particular, the optimal pick-up point, departure time and drop-off point for each ride are identified to then assess whether the additional discounts compensate for the discomfort of walking for all travellers. Importantly, we note that those mutually attractive compact stop-to-stop rides have a structure similar to a single trip request. Thanks to this resemblance, a set of such rides can be subject to a new bundling operation. These rides are then provided as an input to ExMAS and with a few minor modifications can be pooled into the desired hyper-pool rides. In the last, crucial step, ExMAS finds pairs, triples and quadruples of stop-to-stop pooled rides and bundles them together. This second-degree pooling (pooling rides which have already been pooled) is critical to leap beyond the current limits of ride-pooling systems. The hierarchical, utility-based approach guarantees that all identified rides are attractive for all co-travellers (i.e. their perceived generalised cost does not exceed that of private ride-hailing). Notwithstanding, we do not claim for completeness, i.e. not all feasible rides (stop-to-stop and hyper-pooled) are identified using our method. The proposed method can be, therefore, considered a form of heuristic search in the very large solution space of high-degree pooling.

By following the above-mentioned step-wise approach, one obtains a complete set of rides from which, by solving the bi-partite matching (classical for ride-pooling), the final solution is obtained. In the solution, each of the travellers is assigned to a ride, which may be private, pooled or hyper-pooled, depending now not only on individual preferences but also on global optimality considerations (mileage reductions). Key performance indicators (KPIs) of interest at the system, user and ride levels can be then analysed. In the results, we report: travel times, walk times, fares, passenger utilities (decomposed into trip fare, in-vehicle travel time and walk times to access and egress to/from stop points) and their relative changes across pooling options, ride occupancy and ride fare-efficiency (vehicle travel time divided by the total fare).

The proposed hyper-pool method is replicable and reusable using a publicly available source code and parameterised through a set of design variables (e.g. prices and discounts) and behavioural parameters (value of time, willingness to share, etc.). The hyper-pool algorithm was tested for demand levels of up to 8000 trip requests per hour and maximal batches of 2000 travellers. Such levels are large enough for many real-world problems, and in the following, we demonstrate that they induce a critical mass of demand level, beyond which attractive hyper-pooled rides emerge.

Results

The hyper-pool method is applied to pool a batch of 2000 trips requested within half an hour in Amsterdam, The Netherlands. In the following, we first showcase how 10 travellers are hyper-pooled together into a single multi-stop shared ride, followed by a system-wide impact analysis of hyper-pooled rides for the Amsterdam case study.

Experimental setting

We sample 2000 Amsterdam PM peak trips from an updated version of the Albatross dataset. This demand is considered inelastic (no induced demand), i.e. travellers can only choose from the ride-hailing and ride-pooling alternatives offered. The actual Amsterdam fare for private ride-hailing rides of 1.5€/km is used and the following discounts are assumed: 25% for door-to-door ride-pooling, 60% for stop-to-stop and 75% for multi-stop pooling, i.e. the per-kilometre fares of 1.5, 1.11, 0.5 and 0.375€/km, respectively (corresponding to respective λs in the Methods section). Public transport fare in Amsterdam is ca. 0.3€/km, hence only 25% cheaper than the proposed hyper-pool rides. The in-vehicle time multiplier for pooling is set to βd = 1.3, and multipliers of 1.5 and 1.2 are specified for walking (βw) and waiting (βw) times, respectively. Travellers’ value-of-time is normally distributed around 126/h with a standard deviation of σ = 1.5€/km to represent travellers’ heterogeneity. Stop locations are searched over all nodes of the road network (we did not limit searches to PT stops) within a 7.5-min walk time radius centred around origins and destinations (εd). Fixed and network-wide flat speeds of 1.5 m/s for walking and 8 m/s (28.8 km/h) for vehicle movements (which can be easily refined with space- and time-varying speeds).
are specified. The detailed OSM graph of Amsterdam is used to obtain travel times and distances\(^2\). For public transport connection attributes, the actual GTFS public transport timetable and Open Trip Planner query engine are used (following the method described in ref. 34).

**Hyper-pool showcase**

To showcase the potential of hyper-pooling, we select ten trip requests from the Amsterdam PM peak travel demand, which the proposed algorithm managed to bundle into a single hyper-pooled ride. Figure 3 shows pooling levels resulting from consecutive bundling steps starting from private rides (a) through door-to-door pooled rides (b) and stop-to-stop rides (c) to ultimately yield a single hyper-pooled ride (d). The latter results in ten requests with a total length of 65 km being served by a single 9.6 km multi-stop ride. If these requests were served by private rides, the latter would have resulted in 2.25 vehicle-hours, compared to only 0.44 with the generated hyper-pooled ride (as detailed in Table 1). The total passenger-walking time is 78 min and the total passenger-in-vehicle time is 157 min. Remarkably, the total disutility for each of the travellers bundled in this ride is lower than for the alternative private ride (as our method guarantees that only travellers who willingly join a shared ride based on their utility considerations are mutually compatible). Moreover, for some travellers, the hyper-pooled ride is more attractive than the public transport (PT) alternative (depending also on their value-of-time). While offering an efficient bundling of travel requests, the total in-vehicle travel times are still 50% of those associated with PT and the total walking time is 30% shorter. In monetary terms, the total disutility drops from 127 € with private rides to only 95 € for the hyper-pool, which is also considerably lower than for PT (112 €). Notably, the total collected fare drops from 97 € for private rides to only 24. This reduction in revenues is partially compensated by the cost reductions resulting from fewer vehicle hours. In addition, service attractiveness may result in induced demand

---

**Fig. 2 | Study conceptual overview.** Each traveller issues a trip request (Input) and chooses among a variety of services. Each may choose a public transport (static alternative) and private ride (with own car or hailed) or opt for shared rides. First, in the input demand pattern, we identify all feasible door-to-door-ride pooling combinations with ExMAS. Then, we identify which may be attractively served with a single pair of stops (with exhaustive optimisation). Such stop-to-stop, thanks to their similarity to individual trip requests, can be pooled together (with modified ExMAS) to identify hyper-pooled rides. The disadvantages of increasing levels of pooling (waiting, walking and/or detour) are compensated with reduced trip fares. © https://www.openstreetmap.org/copyright.
which in turn would enlarge the overall revenues. Furthermore, service contribution to reducing externalities may justify subsidies. Note that the hyper-pool solution obtained by the proposed algorithm is identified from the gigantic search space of $3.86 \times 10^{40}$ combinations, otherwise practically impossible without employing the proposed hierarchical method.

**Half an hour hyper-pooling in Amsterdam**

To demonstrate the scalability of the proposed approach and its system-wide effects, it is next applied and investigated for a travel demand consisting of 2000 passenger trips requested within a 30-min batch (4000 trips/hour) in Amsterdam. The impact of consecutive levels of pooling is analysed starting with private rides, then adding door-to-door rides with ExMAS, to thereafter identify stop-to-stop rides and finally generate hyper-pooled rides. At each of these steps, the matching problem (detailed in the Methods section) is solved and the impact of the newly added rides on system performance is analysed.

The resulting rides’ compositions obtained at respective pooling levels are shown in Table 2. Each newly introduced level of pooling attracts a significant share of travellers. Door-to-door pooling is attractive for over 1300 travellers, and 171 of them can be successfully bundled into stop-to-stop rides. Eventually, over 225 travellers are successfully matched into attractive hyper-pooled rides. Notably, introducing new levels of pooling does not affect the number of private rides in the solution, which remains stable at around 30% of the travellers who are not matched with any other co-traveller. The impact of the newly introduced pooled rides is thus mainly on bundling more efficiently travellers who are already pooling. The number of feasible pooled rides (last column) identified by ExMAS exceeds 150,000. This number increases only slightly for stop-to-stop rides but then grows substantially into over a million feasible hyper-pooled rides.

Importantly, the demand of 225 travellers is successfully supplied using only 40 hyper-pooled rides composed, on average, of 5.6 travellers. If each one of the 225 travellers had travelled alone, they would require 52.5 vehicle-hours. This figure is reduced by the hyper-pooled multi-stop rides sixfold to 9 vehicle hours only, allowing to attain an average occupancy of 5.8. If we treat vehicle-hours reduction as a proxy for cost reductions (emission, congestion and other externalities), this may allow discounting of the fare by the assumed rate.

In general, introducing new levels of pooling improves all the KPIs of interest (Table 3). Vehicle hours are reduced, traveller’s costs (disutility) are lower, in-vehicle hours are reduced, and occupancy rates rise. Vehicle hours are reduced from 274 h for private rides to 171 for the solution obtained when offering all pooling services. The generalised total travel times of passengers, i.e. increase in disutility, is reduced from 7.56€ to 6.65€ per passenger. In-vehicle hours are lowest for private rides, then increase for door-to-door pooling, drop when stop-to-stop rides are introduced and drop again when hyper-pooling is introduced. This reduction is achieved by inducing modest average walking times. The mean fare collected by the service provider drops from 5.92€ to 4.02€ per passenger. The average occupancy of pooling increases only slightly when new pooling services are introduced (from 1.53 for door-to-door pooling to 1.60 for hyper-pooled). The greatest efficiency gain (fares per vehicle hours) is yielded for door-to-door pooling, that is, introducing hyper-pooling is not directly attractive for the service provider, whose efficiency drops from 50.8 to 47 €per vehicle hour, yet remains at a significantly higher level when compared to the private ride-hailing solution (47 versus 43€/veh-km).

A more detailed analysis of individual trips is presented in Figs. 4–6. This series of panels (distinguishing between rides/travellers levels) shows how introducing increasing levels of pooling impacts the solution. The different services within each solution are depicted using distinct colours.

First, we investigate how ride occupancy varies as a function of trip length in Fig. 4. One can see in the last panel the massive impact of hyper-pooled rides, reaching, to the best of our knowledge uncommon levels. The stop-to-stop rides (introduced in panel c) are not the main contributor to system efficiency, but rather serve as a mean to an end, i.e. being instrumental for composing hyper-pooled rides in the next step. The compact stop-to-stop rides have occupancy comparable to door-to-door rides, but are typically shorter. The following trend can be observed twice: first: longer private trips (blue) are more often pooled into door-to-door rides (green)—compare panel a with b; second: longer door-to-door rides (orange) are more often pooled into hyper-pooled rides (red)—compare panel c with d. Attractive stop-to-stop rides are typically short (presumably due to constraints of finding the optimal stop locations). However, this is not the case.
Table 1: An illustrative example of a set of 10 travellers bundled into a single hyper-pooled ride (depicted in Fig. 3)

| Rider | Utility | Fare | PT | d2d | s2s | Walk time (minutes) | Travel time (minutes) |
|-------|---------|------|----|-----|-----|---------------------|----------------------|
| 1     | 776     | 7.4  | 8.8| 10.2| 10.9| 3.6                 | 6.8                  |
| 2     | 1168    | 18.4 | 5.8| 15.1| 14.9| 2.9                 | 6.6                  |
| 3     | 1162    | 18.7 | 5.9| 15.0| 14.7| 3.4                 | 6.8                  |
| 4     | 1395    | 17.9 | 15.6| 15.6| 14.6| 3.6                 | 6.5                  |
| 5     | 1470    | 18.4 | 6.6| 15.7| 15.7| 3.7                 | 6.7                  |
| 6     | 1729    | 18.7 | 15.6| 15.6| 14.6| 3.7                 | 6.7                  |
| 7     | 1865    | 18.7 | 15.6| 15.6| 14.6| 3.7                 | 6.7                  |
| 8     | 1995    | 18.7 | 15.6| 15.6| 14.6| 3.7                 | 6.7                  |
| 9     | 2125    | 18.7 | 15.6| 15.6| 14.6| 3.7                 | 6.7                  |
| 10    | 2225    | 18.7 | 15.6| 15.6| 14.6| 3.7                 | 6.7                  |

Rows denote consecutive travellers. Values reported in columns correspond to (dis)utilities, travel times, walk times and fares paid for the four levels of pooling: private (p), door-to-door (d2d), stop-to-stop (s2s), hyper-pool (h), as well as value of time for the public transport (PT).

Table 1 includes an illustrative example of a set of 10 travellers bundled into a single hyper-pooled ride (depicted in Fig. 3). The table presents utility, fare, public transport (PT) values, door-to-door (d2d), stop-to-stop (s2s), and hyper-pooled (h) values for travel times and walk times.

Figure 5 shows how attractive are pooled rides of various types for travellers. It can be observed how travellers’ utility improves relative to a private ride. Here, each dot denotes one of our 2000000 riders and the x-axis distinguishes the riders’ based on their respective value-of-time ($\beta$). The grey background denotes the public transport alternative, which spans from non-attractive to highly competitive versus private rides. At first a clear pattern can be observed with a lower bound of attractiveness for door-to-door pooling (b)—the relative gains of pooling are greater for travellers with lower value-of-time. The impact of stop-to-stop pooling (c) on travellers’ utility is negligible. Stop-to-stop rides are bounded within the relative attractiveness of door-to-door ride-pooling. Hyper-pooling, however, breaks this trend and sharply reduces (dis)utilities, which can now be halved. While the main trend is still visible, it is not as evident anymore—effective hyper-pooled rides are generated and selected also by travellers with a relatively high value-of-time as well. Similarly to the impact on occupancy, the stop-to-stop rides are not attractive per se, despite being offers for a significant discount, i.e. 1.11€/km for a door-to-door ride is reduced in this experiment to 0.96€/km. This discount is apparently not sufficient to compensate for walking induced (which has a high perceived discomfort), yet the true benefits materialise once hyper-pooled services are offered. Notably, hyper-pooled rides are inclusive for both high and low value-of-time travellers. This is a surprising finding, since such services were expected to be attractive mainly to travellers with low value-of-time.

Next, Fig. 6 focuses on the perspective of the service provider. Obviously, under fixed demand, offering discounts always yields lower total revenues. Discounts are offered on the premise that ride-pooling allows a potential cost reduction. To further investigate it, the total fares paid are divided by vehicle hours, which can be treated as a proxy of operating costs. Due to flat per-kilometre fares and fixed network speeds, offering private rides (panel a) always yields 44€/veh-hour (x-axis). On the contrary, this varies significantly for door-to-door rides (panel b). While for a small number of rides, door-to-door pooled rides are less efficient, for the majority, effectiveness increases and reaches up to 200€/veh-hour (e.g. when five highly compatible requests are pooled together). The introduction of stop-to-stop rides (panel c) proves to be ineffective and the discount is not compensated for by the reduced mileage. Again, similarly to the two previous figures, this trend is strongly reversed with hyper-pooled rides, with the majority of the mutually attractive rides generated being also more beneficial for the service provider than private rides. Although some of the identified hyper-pooled rides may be less profitable for operators than the private rides, the majority are more effective. This suggests that under such settings, subsidies for the service provider may not be necessary. These disaggregated results can be compared with the fare per vehicle hour reported in Table 3. Introducing hyper-pooling reduces effectiveness compared to door-to-door rides (from 50 to 47), yet it remains more cost-effective when compared to private rides only (47 to 43).

To understand how both the service design variables and the exogenous behaviour parameters can impact the results, we perform the following sensitivity analysis. We investigate the hyper-pooling results under various price settings and travellers’ value of time. For the 225 travellers in the hyper-pooled rides generated for the 2000 trips in Amsterdam, we test whether the proposed alternative remains attractive. The results are presented in Fig. 7, where attractiveness is compared to private rides (blue) and PT (green). We decrease and increase the per-kilometre fare of hyper-pooling ($\lambda_h$, in the first two panels) and the value of time ($\beta$, in the last two panels). First, we show how it impacts individual travellers (first and third panels) and then if all members comprising hyper-pooled rides still find it attractive (second and fourth panels).

If we increase the hyper-pooling price from 0.375 to 0.56€/km, only 30% of the 225 travellers who found hyper-pooling attractive will opt-out. However, for 70% of the hyper-pooled rides, at least one traveller becomes unsatisfied following this price increase, resulting in its disintegration. 
In comparison to PT for 80% of the travellers. This of time of 16
100% when the value of time drops to 8
pared to private alternative and 20% vis-a-vis public transport.

| Column | denotes the number of feasible rides in respective solutions. | by means of reducing vehicle-kilometres) to sustainability goals by means of reducing vehicle-kilometres)

When a default value of time (\( \beta = 12€/h \)) is used, hyper-pooling is attractive in comparison to PT for 80% of the travellers. This figure increases to almost 100% when the value of time drops to 8€/h. On the other hand, with a value of time of 16€/h only 40% of hyper-pooled rides are attractive when compared to private alternative and 20% vs-a-vis public transport.

### Discussion

The proposed hyper-pool method yields insofar unattainable high occupancy levels while ensuring that those are mutually attractive ride-pooling trips. Up to 14 travellers are pooled together in rides where the discomfort induced by pooling is compensated by reduced fares. The proposed approach hence offers great potential for travellers (whose discomfort induced by pooling is compensated by reduced fares. The cost-effectiveness of hyper-pooled rides exceeds the one of private ride-hailing; however, the cost of larger vehicles needs to be considered. The hyper-pooled rides were found to be selected by travellers with a variety of value-of-time, demonstrating that the obtained hyper-

![Table 2 | Rides composition obtained when pooling 2000 trip requests at various pooling levels](image)

| Solution with | Number of rides in the solution |
|---------------|---------------------------------|
| Private \( p \) | 2000 0 0 0 2000 |
| Door-to-door pooled \( p \cup d2d \) | 655 1345 0 0 159,702 |
| Stop-to-stop pooled \( p \cup d2d \cup s2s \) | 645 1184 171 0 160,239 |
| Hyper-pooled \( p \cup d2d \cup s2s \cup h \) | 651 1065 59 225 1.009,855 |

| Table 3 | Ride-pooling KPIs (formally defined in the Methods section) per traveller obtained for the case of a 30-min batch with 2000 trips in Amsterdam for the four consecutive levels of pooling |
| Solution | Vehicle hours | (dis)Utility [€] | Pax in-vehicle hours | Walk time [min] | #Fesible rides | Fare [€] | Revenue [€/veh-hour] | Average occupancy |
| Private | 0.137 | --7.56 | 0.137 | 0.00 | 2000 | 5.92 | 43.24 | 1.00 |
| Door-to-door pooled | 0.090 | --7.31 | 0.181 | 0.00 | 159702 | 4.55 | 50.81 | 1.53 |
| Stop-to-stop pooled | 0.088 | --7.18 | 0.169 | 0.01 | 160239 | 4.35 | 49.43 | 1.56 |
| Hyper-pooled | 0.086 | --6.65 | 0.146 | 0.02 | 1009855 | 4.02 | 47.03 | 1.60 |

Fig. 4 | Occupancy (y-axis) of rides obtained at respective levels of ride pooling, scattered against ride length (x-axis). Each panel denotes a solution where new levels of pooling are introduced, dots represent individual rides and colours denote the pooling service kind. Private rides (a) have a single-person occupancy. Door-to-door services of ExMAS (b) dominate for longer trips but are also present for shorter trips. Stop-to-stop trips (c) are identified mainly for the shorter trips and do not result in a significantly higher occupancy level. However, once they are bundled again into hyper-pooled rides (d) not only the rides are longer, but also attain higher occupancy level.
pooled solutions managed to cater for diverse travellers and needs by offering an affordable alternative while also effectively compensating travellers with greater values of time.

By employing behavioural constraints and capitalising on the hierarchical nature of the problem, the proposed hyper-pool algorithm is able to drastically reduce the size of the search space. However, this also implies that

**Fig. 5** Attractiveness of pooled rides of various kinds against a private ride utility: relative change in perceived costs (x-axis) for rides obtained at respective levels of ride-pooling, scattered against travellers value of time (x-axis). Relative attractiveness (disutility reduction $\Delta U$ on the y-axis as related to the private ride) obtained for private (a), door-to-door (b), stop-to-stop (c) and hyper-pooled (d) rides plotted against the value of time of the respective travellers (x-axis). Each dot here denotes a single traveller and is coloured according to the type of pooling service selected by the traveller, negative values are obtained when perceived costs of the ride were lower than the ones of private ride. The grey dots in the background represent the relative attractiveness of public transport (vs. private ride).

**Fig. 6** Trip efficiency (fare per vehicle-kilometre on the y-axis) as a function of ride length (x-axis) for various ride types (colours), each dot represents a single ride. Consecutive panels show results for solutions with private rides only (a), including door-to-door (b), stop-to-stop (c) and hyper-pooled (d) rides.

**Fig. 7** Sensitivity analysis of hyper-pooling. For the 225 travellers for whom our algorithm identified attractive hyper-pooled rides we test whether the hyper-pooling option remains attractive when we alter the fare (a, b) and value of time (c, d). Lines denote the share of 225 travellers who find hyper-pooling attractive in relation to either private ride-hailing (green) or public transport (blue) under a given fare or value of time. The analysis is done at a single traveller level (a, c) and at the level of rides, which remain attractive only if they are attractive for all co-travellers (b, d).
the algorithm is not exhaustive since we gradually increase the discount with the newly introduced pooling levels. Furthermore, the solution is generated offline. Transferring the proposed method and findings to a real-time setting remains an open challenge. In developing the proposed hyper-pool approach, the travellers’ perspective is deliberately adopted, treating the fleet operations as exogenous. This is deemed reasonable, since the focus here is on the potential of a given demand to be pooled while assuming that the available fleet is sufficiently large. In other words, it is assumed that the fleet size is such that there are vehicles available to serve the pooled rides without additional delays beyond those caused by the compromise among the co-travellers to minimise joint departure delays. Another consequence of the adopted perspective is that stops are set so as to minimise the walking time for the respective group of co-travellers. Even a minor shift in stop location and departure time may significantly shorten the overall vehicle route and avoid detours (or, in other words, shift the user equilibrium solution towards the system optimum one). This can improve the solution and further reduce mileage. On a related note, vehicle-hours as reported in this study (mileage) only constitute a rough approximation of the actual costs. First, deadheading and repositioning are not considered, focusing strictly on the distance travelled with a passenger onboard only. Second, no considerations are made regarding the potential deployment of different vehicle types. Rides composed of more than three travellers require special kinds of vehicles, resembling minibuses—presumably more expensive to operate, which may be taken into consideration in future research.

We believe that the proposed approach is not only beneficial for improving the efficiency of ride-pooling but may also inspire the development of novel solutions to the transit network design problem, which is commonly solved by means of (meta)heuristics (see ref. 37). While the resulting hyper-pool rides do not correspond to public transport lines (i.e. recurrent service configuration patterns), they may be used as spatial patterns driving the network design solutions towards a new optimum. Broader experiments with larger travel demand sets may reveal recurrent patterns, and clustering methods may be used to spatially and temporally aggregate hyper-pooled rides into potential service lines. An additional direction for further research pertains to the development of personalised pricing strategies and related revenue management techniques where discounts are determined for each traveller-ride pair. Furthermore, whether discounts offered for hyper-pool services can also be made attractive for the system provider remains an open question. Traditionally, ride-pooling services are considered a commercial product. However, arguably, once they reach occupancy levels comparable to transit, one may consider treating them as public services, e.g. subject to tendering and subsidies. Related policy decisions depend on the extent to which these services contribute to sustainability goals and their added value to urban accessibility, which largely depends on whether they compete with or complement public transport and active modes of transport.

**Methods**

The proposed method revolves around the Exact Matching of Attractive Shared rides (ExMAS) algorithm. Door-to-door attractive pooled rides computed with ExMAS are compressed into attractive stop-to-stop rides. Subsequently, identified attractive stop-to-stop pooled rides are themselves pooled once again using a modified version of the ExMAS algorithm, as detailed below.

**Travel demand data**

The proposed method computes the hyper-pooling for a given demand pattern known in advance \( Q = \{Q_1, Q_2, \ldots, Q_n\} \), with a single trip \( i \) defined in terms of its origin, destination and departure time:

\[
Q_i = (s_i, d_i, t_p)
\]

The origins and destinations are the nodes \( N \) in the road network graph \( (o_i, d_i \in N) \). We propose an offline method, suitable for processing travel demand data that are readily and publicly available for multiple cities, e.g. New York or Chicago.

**Door-to-door ride-pooling with ExMAS**

We start by computing attractive pooled door-to-door rides using our ExMAS algorithm. (Full methodological details of the ExMAS algorithm are available in ref. 1 while its software implementation is publicly available at https://github.com/RafalKucharskiPK/ExMAS. Here we provide only a brief introduction.) ExMAS first identifies the pairwise shareable trip requests and gradually exploits the search space of pooled rides of increasing degree (number of travellers). First, the algorithm aims at extending identified pairwise shareable trip request pairs into triplets, then extending triplets into quadruples, and so forth. ExMAS exhaustively exploits the search space of attractive pooled rides and terminates when no ride can be extended further. The feasible pooled rides resulting from ExMAS serve as input for the new pooling services proposed here.

In ExMAS, pooled rides are considered attractive if the pooled-ride utility is greater than the utility associated with the alternative private ride for all respective co-travellers. In a door-to-door private (non-shared) ride, the utility \( U_p \) is composed of a direct travel time \( t_p \) weighted by \( \beta_p \) (corresponding to the value-of-time) and a distance-based fare \( \lambda_p \)

\[
U_p = \beta_p t_p - \lambda_p l + \epsilon
\]

To identify attractive pooling, the above utility term is compared with the door-to-door pooled ride utility \( U_{dtd} \):

\[
U_{dtd} = \beta_p t_{dtd} l_{dtd} + \lambda_p d_{dtd} - \lambda_{dtd} l + \epsilon
\]

For a traveller to select the pooled alternative, the lower distance-based fare \( \lambda_{dtd} < \lambda_p \) needs to at least compensate for the longer travel time \( t_{dtd} \geq t_p \), delay induced by pooling \( d_{dtd} > 0 \) and additional discomfort (expressed as \( \beta_{dtd} > 1 \) to represent the so-called willingness-to-share). Mind that the utility here is, in fact, a disutility and is negative, due to the negative sign in \( \lambda_p \) and negative values of \( \beta_p \). Thus, the computed utility values correspond to generalised travel costs. The delay in offline ride-pooling is understood as the absolute difference between the desired and actual departure time (as detailed in ref. 1). The experiments performed in this study examine the deterministic part of the utility function \( \epsilon = 0 \). To account for heterogeneity among travellers, i.e. taste variations, traveller-specific value-of-time \( \beta_p \) drawn from a normal distribution are incorporated into the utility calculation.

ExMAS outputs the set of attractive stop-to-stop rides \( R_{dtd} = \{r_{dtd}\} \), with a ride defined generically as:

\[
r_{dtd} = (Q_r, O_r, D_r, t_r^o),
\]

where \( Q_r \) is a set of served trips, \( O_r \) and \( D_r \) are the ordered sequences of served trips’ origins and destinations, respectively, and \( t_r^o \) denotes traveller’s desired departure time.

Note that ExMAS generates so-called FIFO trips only, i.e. where the sequence of pick-ups precedes the sequence of drop-offs.

**Optimal pickup point for stop-to-stop pooled rides**

Next, we identify which door-to-door pooled rides generated by ExMAS can be used using common pick-up and drop-off points. For each pooled ride, the algorithm examines whether travellers can walk to a single common pick-up point and arrive at a common drop-off point while yielding a superior travel alternative, using the following method.

The following utility formula is used for assessing the attractiveness of a stop-to-stop pooled ride, where travellers walk to a pick-up point, travel directly to the drop-off point, and walk to their destination:

\[
U_{dpl} = \beta_p (\beta_{dpl} l_{dpl} + \beta_d d_{dpl} + \beta_w w_{dpl}) - \lambda_{dpl} l + \epsilon
\]

Similarly to ExMAS, only attractive stop-to-stop pooled rides are considered, i.e. such that \( U_{dpl} \geq U_{dtd} \) for each co-traveller \( i \in Q_r \). This means that the reduced price \( \lambda_{dpl} < \lambda_{dtd} \) and possibly shorter travel time \( l_{dpl} \leq l_{dtd} \) must at least compensate for the induced walking discomfort.
(\beta \cdot w_{\lambda,i}). Mind that no additional in-vehicle discomfort ($\beta \cdot g_{\Lambda,i} = \beta \cdot h_{\Lambda,i}$) is applied, since the ride is anyhow pooled. The discomfort associated with walking time ($\beta > 1$) is explicitly accounted for. Moreover, in-vehicle travel time is likely to become shorter since the number of pick-up and drop-off locations is reduced, and their locations are determined so as to reduce detours. This effect gives an additional slack, which makes it easier to compensate for the negative impact of walking.

The yielded stop-to-stop pooled rides are denoted by $r_{\lambda,i}$ and are defined by means of a set of served travellers ($Q_i$) and a triplet of origin (pick-up), destination (drop-off) and departure time:

$$r_{\lambda,i} = (Q_i, o, d, t)$$

(6)

The following step identifies the optimal pick-up and drop-off locations and respective departure times for a given stop-to-stop ride. Instead of finding the optimal triplet in three-dimensional search space decomposes the problem into three independent steps solved sequentially. The algorithm first loops over common pick-up points, then determines the optimal departure time, and finally loops over common drop-off points. This exhaustive search did not result in a computational bottleneck since the search space solutions are sufficiently small and can be easily performed in parallel. This approach was undertaken rather than applying an optimisation programme for a non-linear objective function on a multidimensional discrete search space, which is considered out of scope for this study:

1. Identifying access points (i.e. virtual stops) accessible from all the origins within a fixed walking time threshold $t_w^{\text{max}}$, either based on the full network graph or only among a predetermined set of locations (e.g. existing public transport stops).

2. For each candidate pickup point, determine the optimal departure time. To this end, the delay experienced by each traveller is calculated as the absolute difference between the desired departure time $t_i$ and the departure time to walk to the pick-up point ($t_w(o_i, o)$) and reach it at the time of common departure $t$:

$$d_{\lambda,i,o,t} = |t - t - t_w(o_i, o)|$$

(7)

Subsequently, the optimal departure time $t$ from stop point $o$ is determined as the minimum over the squared sum of delays:

$$t = \arg \min_{t \in \mathbb{R}} \sum_{o} (d_{\lambda,i,o,t})^2$$

(8)

This step is independent of the drop-off point $d$ and it is thus possible to execute it once for each origin to thereafter use it for all candidate destinations in the next step.

3. Finally, all potential drop-off locations from which destinations are reachable within a fixed walking time threshold $t_w^{\text{max}}$ are explored. We can now identify the optimal ride by tracing how the total utility changes for a given triplet (pick-up location, destination time, and drop-off location). Note that most of the utility components change as a function of the triplet specifications: the pick-up location and departure time determine access walk time and delays, respectively, and the drop-off location determines the egress walk time. In addition, the combination of pick-up and drop-off locations determines in-vehicle time changes. Similarly to eq. (8), to avoid utility imbalances between co-travellers, the so-called log sum formula is applied as follows:

$$(o, d, t) = \arg\max_{o \in N \times t \in \mathbb{R} \times d \in N} \ln \left( \sum_{\lambda,i} \exp(U_{\lambda,i}(o, t, d)) \right)$$

(9)

Since the goal is to identify attractive rides, we further consider only stop-to-stop pooled rides, which are found to be more attractive than their door-to-door pooled service counterparts for all respective co-travellers:

$$R_{\lambda,i} = \{(Q_i, o, d, t) : U_{\lambda,i} \geq U_{\lambda,i}^{\text{hyper}}, \forall i \in Q, \forall r \in r_{\lambda,i}\}$$

(10)

This can be further relaxed to identify stop-to-stop rides which are more attractive than private rides: $U_{\lambda,i} \geq U_{p,i} - \varepsilon$ we did not exploit this direction in the illustrative experiment. The outcome of this step is used as input to the subsequent hyper-pool algorithm.

### Pooling stop-to-stop rides to the hyper-pooled rides

Stop-to-stop pooled rides are, just like private rides, defined using a triplet: origin, destination and departure time (compare eqs. (1) and (9)). They can, therefore, be treated as input to the ExMAS algorithm and thereby bundle pooled rides again into so-called hyper-pooled rides. In short, we aim to pool stop-to-stop rides, which are already in themselves composed of several rides bundled by means of single pick-up and drop-off locations.

To this end, the utility of a stop-to-stop pooled ride is expressed, in a fashion similar to a private ride in eq. (2), using the following formula:

$$U_{\lambda,i} = \beta \cdot t_{\lambda,i} - \lambda h_{\lambda,i} + \varepsilon$$

(11)

Here, $h_{\lambda,i}$ is the fare paid by the travellers (fares are equal among the co-travellers in the stop-to-stop pooled ride, since the travel distance is identical for all fellow travellers) and $t_{\lambda,i}$ is the direct in-vehicle time (between the pick-up location and the drop-off location). The previously identified pick-up and drop-off locations are hereby treated as the origin $o$ and destination $d$ of such a pooled ride, respectively. Similarly, the optimal departure time computed earlier is used as the desired departure $t$ for all the respective pooled travellers. The maximal value-of-time $\beta$ among co-travellers is employed in subsequent calculations since it serves as an upper bound in reducing the available slack for pooling while ensuring that the hyper-pooled ride is attractive for the most demanding co-traveller included in the pool.

Next, the algorithm identifies which of such rides may be pooled again so as to form hyper-pooled rides $h$. A formula similar to eq. (3) for door-to-door pooled rides is used:

$$U_{h} = \beta \cdot (\beta^h t_{h} + \beta^h d_{h}) - \lambda h l + \varepsilon$$

(12)

An additional discount is applied ($\lambda_h < \lambda_{\lambda,i}$) as well as a penalty ($\beta^h$) which is lower than ($\beta_{\lambda,i}^{\text{hyper}}$), since travellers are already pooled.

Similarly to ExMAS, the hyper-pool algorithm first identifies pairwise shareable rides and matches them together. From such pairs, a new shareability graph is constructed, in which we explore pooled rides of a gradually increasing degree. Note that each stop-to-stop ride is composed of at least two individual trip requests; thus, any pair of shareable stop-to-stop rides pool together at least four individual trip requests.

Several adjustments to the original ExMAS algorithm have to be made when applied to the stop-to-stop pooled rides. First, while single trips are pairwise independent, stop-to-stop rides are composed of individual trips which may be overlapping. For instance, a stop-to-stop ride composed of trips $\{1, 2\}$ cannot be pooled with another stop-to-stop ride composed of trips $\{2, 3\}$ - since traveller 2 cannot be pooled with him- or herself. We, therefore, add a constraint: pooling candidates need to be mutually exclusive ($Q_i \cap Q_j = \emptyset$). Second, it is essential to ensure that each ride remains attractive for all co-travellers included in the bundle, i.e. $U_{h,i} > U_{p,i}$, for each $i \in Q$. It should be noted that this condition is tested at the stop-to-stop ride level only, which does not guarantee attractiveness at the individual level (for each single trip request). While for some co-travellers, stop-to-stop pooling is very attractive and close to both origin and destination, for others, it may lie at the limit of attractiveness. Similarly, ExMAS outputs the delay relative to the desired departure time, which for the hyper-pooled rides becomes an optimal departure time computed with (eq. (8)), thus the individual delay evaluated with eq. (7) needs to be updated. For hyper-pooling travellers, similar to stop-to-stop rides, the delay is the absolute difference between the actual and desired departure from the origin to arrive at the pick-up point in time to
enter the shared vehicle. For some travellers, shifting the departure due to hyper-pooling is positive (moving closer to the desired departure), while for others, it is negative. Based on a series of experiments, it can be concluded that using aggregated utility formulas of eq. (11) and eq. (12) well approximates the individual utilities encountered by co-travellers, and only a small share of identified hyper-pool rides does not meet the attractiveness criteria (ca. 5% travellers), which can be easily post-processed (as we do in the experiments).

Employing the above adjustments, the original ExMAS algorithm successfully pools stop-to-stop rides into hyper-pool rides. The output of the original ExMAS $R_h$ is now enriched with information concerning the access and egress times of individual travellers:

$$r_h = (Q_{st}, O_i, D_r, t^p_r),$$

$Q_{st}$ denotes a set of stop-to-stop rides pooled together, $O_i$ and $D_r$ are the sequences of pickup and drop-off locations, respectively. Access and egress walk times of individual travellers can be obtained from their stop-to-stop rides, while the departure from the first stop $t^p$ and the route along the intermediate stops $O_i$ and $D_r$ allows calculating their in-vehicle travel times and delays.

**Solution**

Applying the previous steps results in a rich set of feasible rides of various kinds, within which it is now possible to search for the optimal solution. Since typically, each traveller has multiple travel (at the very least a private ride is available and typically multiple pooled rides of various kinds are feasible), it is necessary to find the ride that will be performed by each traveller. The global choice set is now composed of rides of four kinds:

$$R = R_p \cup R_{d2d} \cup R_{s2s} \cup R_h$$

To address this, a so-called trip-ride coverage problem is formulated as an assignment problem. Hereby, each trip $i$ is unilaterally matched with a ride $r \in R$. The typical formulation where the assignment is formulated as the problem of identifying a binary vector $x_r$ of length equal to the number of rides $||R||$ is adopted, i.e. an assignment variable which takes the value one if a ride is selected and zero otherwise (second constrain of eq. (15)). The costs $C_r$ are then expressed as the cost of each ride $c_r$, multiplied by the assignment variable $x_r$ and aggregated for all rides (eq. (15)).

The assignment must assign each trip request to exactly one ride, obtained through the row-wise sum of assignment variable $x_r$ in the tript-ride incident matrix $I_{r}$. The latter is a binary matrix, in which each entry is one if ride $r$ serves trip $i$ and zero otherwise (first constrain of eq. (15)). Eventually, the solution to the problem (eq. (15)) is the set of rides $R^* \subseteq R$ such that $x_r = 1 \forall r \in R^*$. The matching problem is formulated using the following binary programme:

$$\min \quad C_R(x_r) = \sum_{r \in R} c_r x_r$$

subject to

$$\sum_{r \in R} I_{r} x_r = 1,$$

$$x_r \in [0, 1].$$

Typically, two perspectives on cost are considered: the one taken by the service provider (minimising the sum of vehicle hours) and the one taken by travellers (maximising the sum of utilities $U_r$). Despite taking a strictly demand-orientated stance in this study, operator’s costs, i.e. vehicle hours, are minimised in solving this matching problem. This choice is made since the algorithm has already ensured that all the pooled rides in the search space are attractive for all travellers, thus, when optimising for the service operator, one does not run the risk of including unattractive rides in the solution. Moreover, we are explicitly interested in identifying the potential for high-occupation pooled rides, i.e. those for which vehicle hours are minimal.

Optionally, one may filter the solution $R$ in eq. (15) to consist only of a subset of services. For instance, in the experiments performed in this study, the impact of successively introducing new kinds of services on system-wide KPIs is investigated.

**Implementation**

The hyper-pool algorithm solves a service planning problem in an acceptable time. It is parameterised via a custom configuration file and can be easily deployed on external computational servers. The open-source Python code is publicly available, allowing for further developments and experiments. The pseudocode algorithm is presented in Algorithm 1 below. It starts with trip requests as input data to travel demand and gradually searches for attractive rides of four kinds, finally matching travellers with rides by solving the coverage problem.

The hyper-pool algorithm outputs contain information concerning the ride serving each trip request (traveller). Rides in the solution can be of four kinds: private door-to-door-ride (for those who couldn’t be successfully matched with anyone), pooled door-to-door rides, pooled stop-to-stop rides or multi-stop rides. The fare paid by each traveller depends on the distance travelled and the type of service. Other components of travellers’ utility: detour, delay, and walk time are stored in the solution for further analysis.
Despite the radical reduction in the search space size, it may (similarly to the ExMAS algorithm) still become prohibitive under certain settings (beyond some demand levels and discount rates). Nevertheless, 10-minute batches with up to 2000 trip requests are computed within less than two hours on a standard laptop (MacBook Pro 2021 M1 chip, 16GB RAM), sufficient for offline strategic analyses of most practical real-world ride-pooling problems. Furthermore, the stop-to-stop rides algorithm can be solved in parallel. The bottleneck of the current implementation seems to be in the second run of ExMAS when stop-to-stop rides are bundled into hyper-pooled rides. Since at this step, sharing is not penalised by applying yet another multiplier ($\beta_i$) and with a considerable discount being offered, one often ends up with a very large number of identified hyper-pooled rides (over a million in our half-hour batch, which includes only 250 out of 2000 individual trip requests).

**Key performance indicators**
The following set of key performance indicators is generated to measure the performance of the obtained hyper-pool rides:

- **vehicle hours:** i.e. total time spent by vehicles to serve the demand (empty kilometres, like deadheading and repositioning, are not included here)

- **passenger utility:** i.e. total (dis)utility of travellers, both total as well as decomposed into:
  - trip fare (direct monetary costs)
  - in-vehicle travel time
  - walk time (to access and egress to/from stop points)

- **efficiency:** i.e. vehicle travel time divided by the total fare. Expressed in €per vehicle-kilometre and used as a proxy of service efficiency for the provider.

- **and ride occupancy:**

  \[
  \left( \frac{\sum \limits_{i \in Q} t_p}{\sum \limits_{r \in R} t_r} \right)
  \]

  The above formulation of the ‘time-averaged effective occupancy’ is arguably the most meaningful KPI. While in public transit, the occupancy is calculated for the actual passenger-hours, the shortest-path passenger-hours are specified here, i.e. the total passenger-hours for the private rides ($\sum \limits_{i \in Q} t_p$) as the nominator. In the denominator, the proposed metric takes the actual vehicle-hours of the selected rides ($\sum \limits_{r \in R} t_r$). We propose this formulation for the ride-occupancy metric, since it is sensitive to detours (which are included in the denominator only) and directly reflects the compactness of pooling.

**Data Availability**
The input data to reproduce the experiment is available on the public repository https://doi.org/10.5281/zenodo.12806397. The road graph data used in experiments was downloaded from OpenStreetMap public dataset on July 2024. The raw data with the results of the conducted experiments is available on the public data repository: https://doi.org/10.5281/zenodo.12721529.

**Code availability**
Hyper-pool Python code, along with scripts to reproduce the experiment is available on a public repository: https://github.com/RafalKucharskiPK/ExMAS/releases/tag/hyperpool and under https://doi.org/10.5281/zenodo.12806397 without use restrictions.

Received: 9 January 2024; Accepted: 12 August 2024; Published online: 13 September 2024

**References**
1. Kucharski, R. & Cats, O. Exact matching of attractive shared rides (exmas) for system-wide strategic evaluations. Transp. Res. B Methodol. 139, 285–310 (2020).
2. Agatz, N., Erera, A. L., Savelsbergh, M. W. & Wang, X. Dynamic ride-sharing: a simulation study in metro Atlanta. Procedia Soc. Behav. Sci. 17, 532–550 (2011).
3. Chan, N. D. & Shaheen, S. A. Ridesharing in North America: past, present, and future. Transp. Rev. 32, 93–112 (2012).
4. Furuhata, M. et al. Ridesharing: the state-of-the-art and future directions. Transp. Res. B Methodol. 57, 28–46 (2013).
5. Tachtet, R. et al. Scaling law of urban ride sharing. Sci. Rep. 7, 42688 (2017).
6. Bischoff, J., Maciejewski, M. & Nagel, K. City-wide shared taxis: a simulation study in Berlin. In 2017 IEEE 20th international conference on intelligent transportation systems (ITSC) 275–280 (IEEE, 2017).
7. Santi, P. et al. Quantifying the benefits of vehicle pooling with shareability networks. Proc Natl Acad. Sci. USA 111, 12390–12394 (2014).
8. Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E. & Rus, D. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. Proc. Natl Acad. Sci. USA 114, 462–467 (2017).
9. Engelhardt, R., Dandl, F. & Bogenberger, K. Speed-up heuristic for an on-demand ride-pooling algorithm. Preprint at arxiv https://arxiv.org/abs/2007.14877 (2020).
10. Simonetto, A., Monteil, J. & Gambella, C. Real-time city-scale ridesharing via linear assignment problems. Transp. Res. C Emerg. Technol. 101, 208–232 (2019).
11. Yao, R. & Bekhor, S. A dynamic tree algorithm for peer-to-peer ridesharing matching. Netw. Spatial Econ. 21, 801–837 (2021).
12. Liu, Y. & Samaranayake, S. Proactive rebalancing and speed-up techniques for on-demand high capacity ridesourcing services. In IEEE Transactions on Intelligent Transportation Systems 819–826 (IEEE, 2020).
13. Shah, S., Lowalekar, M. & Varakantham, P. Neural approximate dynamic programming for on-demand ride-pooling. In Proc. AAAI Conference on Artificial Intelligence 507–515 (2020).
14. Bischoff, J., Kaddoura, I., Maciejewski, M. & Nagel, K. Simulation-based optimization of service areas for pooled ride-hailing operators. Procedia Comput. Sci. 130, 816–823 (2018).
15. Kontou, E., Gankapati, V. & Hou, Y. Reducing ridesourcing empty vehicle travel with future travel demand prediction. Transp. Res. C Emerg. Technol. 121, 102826 (2020).
16. Grahn, R., Qian, S. & Hendrickson, C. Improving the performance of first-and last-mile mobility services through transit coordination, real-time demand prediction, advanced reservations, and trip prioritization. Transp. Res. C Emerg. Technol. 133, 103430 (2021).
17. Kang, S., Mondal, A., Bhat, A. C. & Bhat, C. R. Pooled versus private ride-hailing: a joint revealed and stated preference analysis recognizing psycho-social factors. Transp. Res. C Emerg. Technol. 124, 102906 (2021).
18. Lavieri, P. S. & Bhat, C. R. Modeling individuals’ willingness to share trips with strangers in an autonomous vehicle future. Transp. Res. A Policy Pract. 124, 242–261 (2019).
19. Alonso-González, M. J., Cats, O., van Oort, N., Hoogendoorn-Lanser, S. & Hoogendoorn, S. What are the determinants of the willingness to share rides in pooled on-demand services? Transportation 48, 1733–1765 (2021).
20. Lazarus, J. R., Caicedo, J. D., Bayen, A. M. & Shaheen, S. A. To pool or not to pool? understanding opportunities, challenges, and equity considerations to expanding the market for pooling. Transp. Res. A Policy Pract. 148, 199–222 (2021).
21. Geržinčič, N., van Oort, N., Hoogendoorn-Lanser, S., Cats, O. & Hoogendoorn, S. Potential of on-demand services for urban travel. Transportation 1–33 (2022).
22. Li, M., Jiang, G. & Lo, H. K. Pricing strategy of ride-sourcing services under travel time variability. Transp. Res. E Logist. Transp. Rev. 159, 102631 (2022).
23. Young, M., Farber, S. & Palm, M. The true cost of sharing: a detour penalty analysis between uberpools and uberox trips in Toronto. Transp. Res. D Transp. Environ. 87, 102540 (2020).
24. Zwick, F. & Axhausen, K. W. Ride-pooling demand prediction: a spatiotemporal assessment in Germany. J. Transp. Geogr. 100, 103307 (2022).

25. Shulika, O., Bujak, M., Ghasemi, F. & Kucharski, R. Spatiotemporal variability of ride-pooling potential—half a year New York City experiment. J. Transp. Geogr. 114, 103767 (2024).

26. Albo, González, M. J., van Oort, N., Cats, O., Hoogendoorn-Lanser, S. & Hoogendoorn, S. Value of time and reliability for urban pooled on-demand services. Transp. Res. C Emerg. Technol. 115, 102621 (2020).

27. Karaneku, P., Schiffer, M. & Waldherr, S. The customer is always right: customer-centered pooling for ride-sharing systems. Preprint at arXiv:2107.01161 (2021).

28. Ma, Z. & Koutsopoulos, H. N. Near-on-demand mobility, the benefits of user flexibility for ride-sharing services. Transp. Res. C Emerg. Technol. 135, 103530 (2022).

29. de Ruijter, A., Cats, O., Alonso-Mora, J. & Hoogendoorn, S. Ride-pooling adoption, efficiency and level of service under alternative demand, behavioural and pricing settings. Transp. Plan. Technol. 46, 407–436 (2023).

30. Erhardt, G. D. et al. Do transportation network companies decrease or increase congestion? Sci. Adv. 5, eaau2670 (2019).

31. Tarduno, M. The congestion costs of Uber and Lyft. J. Urban Econ. 122, 103318 (2021).

32. Bilali, A., Fastenrath, U. & Bogenberger, K. Analytical model to estimate ride-pooling traffic impacts by using the macroscopic fundamental diagram. Transp. Res. Rec. 2676, 03611981211064892 (2022).

33. Hall, J. D., Palsson, C. & Price, J. Is Uber a substitute or complement for public transit? J. Urban Econ. 108, 36–50 (2018).

34. Cats, O., Kucharski, R., Danda, S. R. & Yap, M. Beyond the dichotomy: how ride-hailing competes with and complements public transport. PLoS ONE 17, e0262496 (2022).

35. Erhardt, G. D. et al. Do transportation network companies increase or decrease transit ridership? empirical evidence from San Francisco. Transportation 49, 313–342 (2022).

36. Stiglic, M., Agat, N., Savelbergh, M. & Gradisar, M. The benefits of meeting points in ride-sharing systems. Transp. Res. B Methodol. 82, 36–53 (2015).

37. Foljanty, L. On-demand ridepooling market size. https://lukas-foljanty.medium.com/on-demand-ridepooling-market-size-f3f93845c5c (2021).

38. Engelhardt, R. & Bogenberger, K. Benefits of flexible boarding locations in on-demand ride-pooling systems. In 2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS) 1–6 (2021).

39. Fielbaum, A., Bai, X. & Alonso-Mora, J. On-demand ridesharing with optimized pick-up and drop-off walking locations. Transp. Rest C Emerg. Technol. 126, 103061 (2021).

40. Gurnumurthy, K. M. & Kockelman, K. M. Dynamic ride-sharing impacts of greater trip demand and aggregation at stops in shared autonomous vehicle systems. Transp. Res. A Policy Pract. 160, 114–125 (2022).

41. Zhao, M., Yin, J., An, S., Wang, J. & Feng, D. Ridesharing problem with flexible pickup and delivery locations for app-based transportation service: mathematical modeling and decomposition methods. J. Adv. Transp. 2018 (2018).

42. Li, R.-H., Qin, L., Yu, J. X. & Mao, R. Optimal multi-meeting-point route search. IEEE Trans. Knowl. Data Eng. 28, 770–784 (2015).

43. Vansteenkiste, L. et al. A survey on demand-responsive public bus systems. Transp. Res. C Emerg. Technol. 137, 103573 (2022).

44. Yoon, G., Chow, I. Y. & Rath, S. A simulation sandbox to compare fixed-route, semi-flexible transit, and on-demand microtransit system designs. KSCE J. Civil Eng. 26, 1–20 (2022).

45. Papanikolau, A. & Babas, A. Analytical models for comparing demand responsive transport with bus services in low demand interurban areas. Transp. Lett. 13, 255–262 (2021).

46. Fielbaum, A., Tirachini, A. & Alonso-Mora, J. Economies and diseconomies of scale in on-demand ridepooling systems. Econ. Transp. 34, 100313 (2023).

47. Mohring, H. Optimization and scale economies in urban bus transportation. Am. Econ. Rev. 62, 591–604 (1972).

48. de Ruijter, A., Cats, O. & van Lint, H. Emerging dynamics in ridesourcing markets. SSRN 4258151 (2022).

49. Atasoy, B., Ikeda, T., Song, X. & Ben-Akiva, M. E. The concept and impact analysis of a flexible mobility on demand system. Transp. Res. C Emerg. Technol. 56, 373–392 (2015).

50. Leffler, D., Burghout, W., Cats, O. & Jenelius, E. An adaptive route choice model for integrated fixed and flexible transit systems. Transp. B Transp. Dyn. 12, 2303047 (2024).

51. Calabrò, G. et al. Fixed-route vs. demand-responsive transport feeder services: An exploratory study using an agent-based model. J. Adv. Transp. 2022, 1–20 (2022).

52. Narayan, J., Cats, O., van Oort, N. & Hoogendoorn, S. Integrated route choice and assignment model for fixed and flexible public transport systems. Transp. Res. C Emerg. Technol. 115, 102631 (2020).

53. Schlüter, J., Bossert, A., Rössy, P. & Kersting, M. Impact assessment of autonomous demand responsive transport as a link between urban and rural areas. Res. Transp. Business Manag. 39, 100613 (2021).

54. Calabrò, G. et al. Designing demand responsive transport services in small-sized cities using an agent-based model. Transp. Res. Procedia 69, 759–766 (2020).

55. Arends, T. A. & Timmermans, H. J. A learning-based transportation oriented simulation model. Transp. Res. B Methodol. 38, 613–633 (2004).

56. Boeing, G. Osmnx: new methods for acquiring, constructing, analyzing, and visualizing complex street networks. Comput. Environ. Urban Syst. 65, 126–139 (2017).

57. Farahani, R. Z., Miaoabachi, E., Szeto, W. Y. & Rashidi, H. Review of urban transportation network design problems. Eur. J. Operation. Res. 229, 281–302 (2013).

58. Dey, B. K., Tirtha, S. D., Eruru, N. & Konduri, K. C. Transformation of ridesharing in New York City: a quantitative assessment. Transp. Res. C Emerg. Technol. 129, 102325 (2021).

59. Dean, M. D. & Kockelman, K. M. Spatial variation in shared ride-hail trip demand and factors contributing to sharing: lessons from Chicago. J. Transp. Geogr. 91, 102944 (2021).

Acknowledgements

This research was supported by the CriticalMaaS project (Grant Number 804489), which is financed by the European Research Council and Amsterdam Institute for Advanced Metropolitan Solutions, as well as by the National Science Centre in Poland programme OPUS 19 (Grant Number 2020/37/B/HS4/01847).

Author contributions

R.K.: conceptualisation; methodology; data curation; formal analysis; writing—original; writing—editing and review; funding acquisition. O.C.: conceptualisation; methodology; resources; validation; writing—original; writing—editing and review; supervision, funding acquisition.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Rafał Kucharski.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
