Technological Forecasting & Social Change 149 (2019) 119733

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Improving ridesplitting services using optimization procedures on a shareability network: A case study of Chengdu

Meiting Tu[a,b], Ye Li[a,*], Wenxiang Li[a,*], Minchao Tu[c], Olivier Orfila[b], Dominique Gruyer[b]

[a] *College of Transportation Engineering, Tongji University, Key Laboratory of Road and Traffic Engineering of the Ministry of Education, 4800 Cao'an Road, Shanghai 201804, PR China*
[b] *Laboratory LIVIC, IFSTTAR, 25 allée des Marronniers, 78000 Versailles, France*
[c] *School of Automobile, Chang'an University, Xi'an, Shanxi Province, PR China*

ARTICLE INFO

Accepted by Xiao Luo

Keywords:
Ridesplitting
Ridesourcing
Shareability network
Policy implications

ABSTRACT

Ridesourcing services play a crucial role in metropolitan transportation systems and aggravate urban traffic congestion and air pollution. Ridesplitting is one possible way to reduce these adverse effects and improve the transport efficiency, especially during rush hours. This paper aims to explore the potential of ridesplitting during peak hours using empirical ridesourcing data provided by DiDi Chuxing, which contains complete datasets of ridesourcing orders in the city of Chengdu, China. A ridesplitting trip identification algorithm based on a shareability network is developed to quantify the potential of ridesplitting. Then, we evaluate the gap between the potential and actual scales of ridesplitting. The results show that the percentage of potential cost savings can reach 18.47% with an average delay of 4.76 min, whereas the actual percentage is 1.22% with an average delay of 9.86 min. The percentage of shared trips can be increased from 7.85% to 90.69%, and the percentage of time savings can reach 25.75% from 2.38%. This is the first investigation of the gap between the actual scale and the potential of ridesplitting on a city scale. The proposed ridesplitting algorithm can not only bring benefits on a city level but also take passenger delays into consideration. The quantitative benefits could encourage transportation management agencies and transportation network companies to develop sensible policies to improve the existing ridesplitting services.

1. Introduction

Vehicular traffic congestion and air pollution during rush hours are two difficult problems for urban cities worldwide. The cost of congestion in the United States alone is approximately $121 billion per year, representing 1% of the GDP, which includes 5.5 billion h of time lost and an extra 2.9 billion gal of fuel burned (Schrank et al., 2012). The CO_2 emissions of Shanghai have increased greatly during these years, so it is necessary to advocate the low-carbon transport (Luo et al., 2017). Additionally, the road transport sector is now the leading cause of air-pollution-related deaths (Roskilly et al., 2015). Therefore, it is imperative and urgent to mitigate congestion and air pollution. Ridesplitting is one possible and effective way to reduce the adverse effects of on-demand ridesourcing services with respect to pollution, energy consumption, and congestion in urban cities.

Shaheen and Chan (2016) offer an explicit classification of shared mobility; see Fig. 1. There are some confusing terms relating to shared mobility for academic transportation researchers. In summary, ridesourcing refers to transportation services that connect drivers who drive private cars (instead of commercial vehicles) with passengers via smartphone applications (Rayle et al., 2016). Ridesplitting is a form of ridesourcing wherein riders with similar routes are matched to the same driver and vehicle in real time (Shaheen et al., 2016). Ridesourcing is different from ridesharing. Drivers who provide ridesourcing services are usually not travelers, and they drive for the purpose of making money. In contrast, in ridesharing, including carpooling and vanpooling, drivers are travelers who share similar origins/destinations with their riders for the purpose of conserving resources, saving money, or saving time (Jin et al., 2018). E-hailing is a kind of on-demand taxi service.

The shared mobility platforms represent a new mobility paradigm, a win-win practice for both drivers and passengers, capable of reducing each traveler's costs, offering a new way to decrease environmental impact (Casprini et al., 2019). Prearranged and on-demand ride services

* Corresponding authors.
E-mail addresses: 1710917@tongji.edu.cn (M. Tu), jamesli@tongji.edu.cn (Y. Li), lwxffff@gmail.com (W. Li), 2016903246@chd.edu.cn (M. Tu), olivier.orfila@ifsttar.fr (O. Orfila), gruyer@ifsttar.fr (D. Gruyer).

https://doi.org/10.1016/j.techfore.2019.119733
Received 31 March 2019; Received in revised form 17 July 2019; Accepted 29 August 2019
Available online 22 October 2019

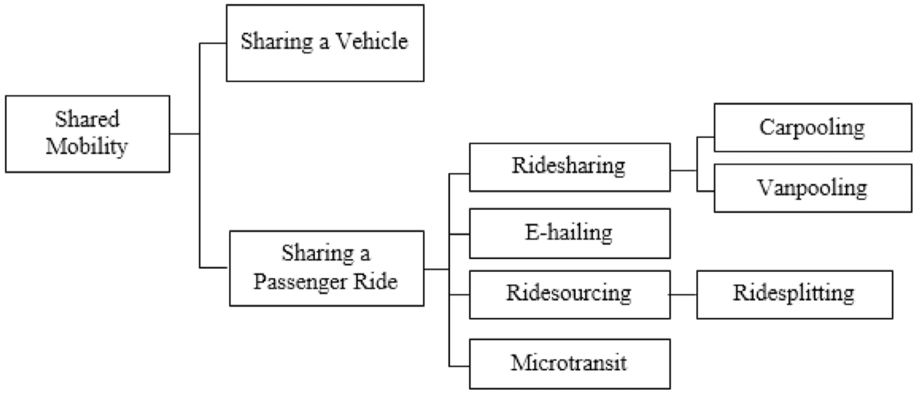

Fig. 1. Classification of shared mobility.
[Source: modified from Shaheen and Chan, 2016].

that match drivers and passengers using smartphone applications have rapidly spread worldwide due to their convenience. As of 2017, Uber operated in 84 countries in over 760 cities (Uber, 2018) with approximately 7 million drivers (Smith, 2018), while Lyft operated in 200 cities in the US with over 315,000 drivers (Lyft, 2018). DiDi Chuxing has served 450 million users in over 400 cities, with the number of daily rides reaching 25 million (Didichuxing, 2017a). Many ridesourcing companies have launched ridesplitting services to improve transportation efficiency, thus reducing carbon emissions. Lyft Line and UberPOOL are leading companies that allow passengers whose routes overlap to split rides and fares (Shaheen et al., 2015). As of April 2018, more than 2.4 million users travel through DiDi ExpressPool (including ridesplitting services) per day over 60 cities (Didichuxing, 2018).

The primary objective of this study is to explore the potential opportunity and evaluate the benefits of ridesplitting using an empirical ridesourcing dataset during rush hours in the city of Chengdu, China. More specifically, this study includes the following three tasks: First, we calculate the actual scale of ridesplitting and analyze the current characteristics of ridesplitting. Second, we compare the potential of ridesplitting under three different objectives to evaluate the benefits of ridesplitting, namely, reduced fleet size, time savings, and cost savings. Third, we analyze the reasons for the gap between the potential of ridesplitting and the actual scale of ridesplitting in the real world.

To our knowledge, most studies have replaced taxi services with ridesplitting services to explore the potential of ridesplitting services. However, taxi services are regarded as premium services. Passengers using taxi services are more concerned about the value of their time and privacy. Therefore, most taxi passengers are not willing to use ridesplitting services. To fill this gap, this study aims to evaluate the potential of ridesplitting services on a city scale using empirical ridesourcing data. The proposed ridesplitting service can not only mitigate air pollution but also take passenger delays into consideration to guarantee the ridesplitting service quality. The findings could help ridesourcing companies improve their existing ridesplitting services. Furthermore, this study can provide enlightening insights to policy makers to provide a better understanding of the potential of ridesplitting and its related environmental implications to support better planning of low-carbon transport and urban transport management.

The rest of this paper is organized as follows. Section 2 summarizes previous relevant studies. Section 3 presents the trip data and map data used in this study. A potential ridesplitting identification algorithm based on a shareability network is described in Section 4. Section 5 analyzes the results in detail and further discusses the insights and recommendations on how to increase the adoption of ridesplitting services. The final section summarizes the study's contributions and limitations and proposes future research directions.

2. Literature review

2.1. Data sources of shared mobility-related studies

Four data sources have typically been used to research the problem of sharing a passenger ride: (a) survey data, (b) mobile phone data, (c) empirical data on taxis, and (d) empirical data on ridesourcing services collected directly from transportation network companies.

Ridesourcing surveys mainly collect data through questionnaires, including intercept surveys (Handke and Jonuschat, 2012) and telephone surveys (Hara Associates Inc, 2013); this method has several weaknesses, including the limited number of respondents who can be surveyed, the incomplete representation of the ridesourcing market, and the huge investment of time and resources. Mobile phone data refer to average daily origin-destination (OD) trips from mobile phone call detail records (Alexander and González, 2015); but the observations are recorded only when an individual interacts with his or her device, resulting in heterogeneous sampling frequencies and an incomplete picture of daily behavior across users, therefore the data lack full authenticity to assess the impacts of sharing a passenger ride. Some studies used empirical data on taxis to quantify the benefits of ridesplitting in New York (Alonso-Mora et al., 2017; Santi et al., 2014; Simonetto et al., 2019). These articles assume that all passengers using taxis are willing to use the ridesplitting service and that there is only one passenger in each ride. However, passengers using taxis are distinct from passengers using ridesourcing services. Taxi services are regarded as premium services. Passengers using taxis probably are not willing to use ridesplitting services because of the issues of time value and privacy. Furthermore, the average number of passengers per ride for taxis in New York was 1.6 in 2014 (King and Saldarriaga, 2017). Therefore, the actual passenger capacity is much larger than what is assumed in the two studies.

Few studies use empirical data on ridesourcing services to evaluate the impacts of ridesplitting because transportation network companies are reluctant to share their ridesourcing data with the public due to privacy protection. To our knowledge, the two ridesplitting studies (Chen et al., 2017; Chen et al., 2018) are the only published papers using empirical data from transportation network companies. Chen et al. (2017) used empirical data from DiDi Chuxing to predict the ridesplitting choices of individual passengers. Chen et al. (2018) analyzed ridesplitting behaviors and the associated impact on multimodal mobility. However, these two papers do not investigate the potential of ridesplitting to evaluate the benefits. This paper uses empirical data on ridesourcing services collected directly from DiDi Chuxing to explore the potential shared trips, time savings and cost savings for ridesplitting services. It is meaningful for transportation network companies, governments and passengers to quantify the gap between the actual scale and the potential of ridesplitting services to improve the present ridesplitting services.

2.2. Objectives of shared mobility-related models

The objectives of shared mobility-related models can be summarized as follows: (a) maximizing the number of ridesplitting trips, (b) minimizing the total travel time or minimizing the vehicle miles of all trips, and (c) maximizing the value saved by ridesplitting.

The maximum number of ridesplitting trips can ideally satisfy passengers' trip demand with a minimum fleet size. Two shared trips can be served by a single vehicle instead of two. Thus, shared trips can be used to reduce the required fleet size to mitigate traffic congestion for society. This also benefits passengers by enhancing their mobility.

Generally, the cost is proportional to the travel time or vehicle miles. Therefore, an optimization objective is proposed for the purpose of minimizing the cost to society.

The Senseable City Laboratory of MIT uses empirical taxi data to quantify the benefits of ridesplitting in the future based on the above two objectives (Santi et al., 2014; Vazifeh et al., 2018). However, the delay experienced by passengers is not included in the objective function.

Maximizing the cost saved by ridesplitting is considered from two aspects: on the one hand, we want to minimize the total time and travel distance of all trips from the perspective of the ridesplitting system. On the other hand, the optimal solution for the whole society may not necessarily be optimal for individual passengers, since the travel time is the time actually spent in the vehicle and does not include delays caused by detours. This objective takes the value of the delay experienced by individual passengers into consideration to guarantee the rights of passengers when using the ridesplitting service. Kleiner et al. (2011) presented an objective of maximizing saved value by carpooling while simultaneously considering the tradeoff of a tolerable delay for commuters due to the detour distance. However, as demonstrated in the definitions we mentioned above, they focus on the ridesharing for commuters rather than ridesplitting.

This paper explores the potential of ridesplitting in terms of the three above-outlined objectives. From these three objectives, interesting findings about ridesplitting services in the future can be found.

2.3. Solution algorithms of shared mobility-related models

Problems related to sharing a passenger ride can be regarded as "pick-up and delivery" problems with desired time windows in which multi-vehicles and multi-passengers should be matched efficiently (Cortés et al., 2010; Ropke et al., 2007). Traditionally, such problems are solved by complex mathematical programming with a large number of variables and constraints. This method can be used to address problems with only a few thousand vehicles and trips at most (Baker and Ayechew, 2003; Laporte, 1992); thus, it is not applicable for the hundreds of thousands of trips and vehicles on a large city scale. Ma et al. (2013) take the first step toward solving the taxi ridesharing problem with dynamic large scales. The Senseable City Laboratory of MIT (Santi et al., 2014) proposes an efficient "shareability network" method that translates spatiotemporal sharing problems into a graph-theoretic framework to realize shared taxi services. Furthermore, these researchers address the minimum fleet problem using a "vehicle sharing network" (Vazifeh et al., 2018).

The problem of maximum matching in ridesharing is NP hard. There are two main algorithms for solving NP hard problems: exact algorithms and approximation algorithms. Exact algorithms, such as dynamic programming algorithms (Psaraftis, 1980), branch-cut methods (Cortés et al., 2010; Ropke et al., 2007), and Lagrangian relaxation algorithms (Baldacci et al., 2004), can be used to solve the ridesharing problem only for small-scale trip requests due to computational challenges. There are many heuristic approximation algorithms, such as tabu search algorithms (Braekers et al., 2014; Kirchler and Calvo, 2013), genetic algorithms (Baker and Ayechew, 2003), adaptive greedy algorithms (Santos and Xavier, 2015), and variable neighborhood search algorithms (Li et al., 2016; Parragh et al., 2010). A hybrid iterative local search algorithm based on a variable neighborhood depth search algorithm (ILS-VND), which has the characteristics of fast calculation, good solution quality and high stability, outperforms other heuristic algorithms in relation to the maximum weight independent set problem (Nogueira et al., 2018). This paper transforms the ridesplitting assignment problem into a maximum weight independent set problem and solves it with the ILS-VND algorithm. This method is effective in addressing the large-scale dynamic ridesplitting assignment problem through real-time computation.

In summary, previous studies on the potential of ridesourcing that used observed data are quite limited, and empirical research on ridesplitting is even sparser because of the lack of data. To our knowledge, no publicly available reference exists to quantify the actual scale and potential of ridesplitting on a city scale. This paper explores the potential of ridesplitting from the three objectives mentioned above. The findings of this study can help reveal the potential benefits of ridesplitting, which will be crucial for transportation management agencies and ridesourcing companies in developing sensible policies to promote the use of ridesplitting services.

3. Data description

3.1. Trip data

The trip data used in this study are from the DiDi GAIA Initiative, DiDi's open data project (DiDi Chuxing, 2017b). The dataset contains the complete order data of DiDi Express and DiDi Premier, two of DiDi Chuxing's core ridesourcing services, in the city of Chengdu, China, from November 1–30, 2016.

According to a previous study (Li et al., 2019) on this dataset, 13:00–15:00 is the rush hour for using ridesourcing services. This research only concentrates on rush hours. The extracted order dataset contains over 1.04 million trips, including fields such as OrderID, start and end timestamps, and pick-up and drop-off locations (sample data are shown in Table 1).

3.2. Map data

We use data from www.openstreetmap.org to create the street

Table 1
Order data structure (37).

Field	Type	Sample	Comment
OrderID	String	mjiwdgkqmonDFvCk3ntBpron5mwfrqvI	Anonymized
Ride start timestamp	String	1501581031	Unix timestamp, in seconds
Ride end timestamp	String	1501582195	Unix timestamp, in seconds
Pick-up longitude	String	104.11225	GCJ-02 Coordinate System
Pick-up latitude	String	30.66703	GCJ-02 Coordinate System
Drop-off longitude	String	104.07403	GCJ-02 Coordinate System
Drop-off latitude	String	30.6863	GCJ-02 Coordinate System

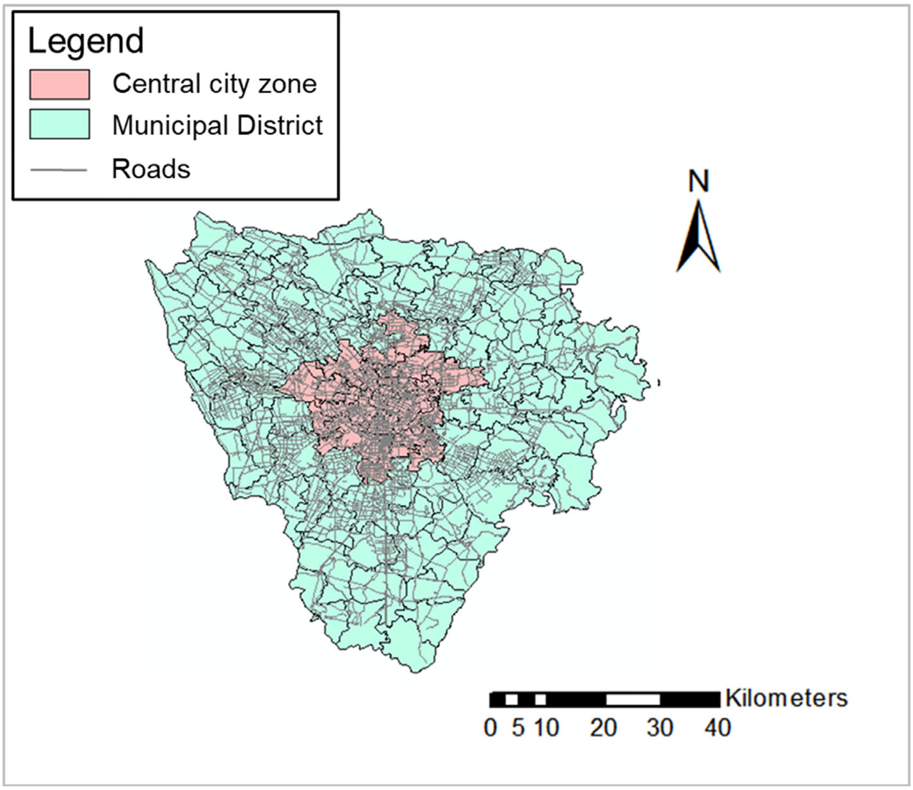

Fig. 2. Research area in Chengdu.

network of Chengdu and choose the central city zone of Chengdu as the research area; see Fig. 2. We build a network in which nodes are intersections and links are roads connecting those intersections. The network contains 7783 nodes and 13,087 links. We retain trips with durations of 1 min to 120 min, thus discarding outliers. After the preprocessing and filtering phase, more than 97% of the trips are kept for further analysis. We match the pick-up and drop-off locations with the intersection nodes from the trip dataset. We compute the travel time between all origins and destinations using the average speed on road segments retrieved from Baidu Map.

4. Methodology

We build the model based on the shareability network derived from The Senseable City Laboratory of MIT [Santi et al., 2014]. We define three performance indicators, which are the percentages of shared trips, time savings and cost savings. A hybrid iterative local search algorithm based on the variable neighborhood depth search algorithm (ILS-VND) is used to find the optimal potential ridesplitting solutions under each of the three objectives. We find the best potential ridesplitting solution of the three and quantify the gap between the actual scale and potential opportunity. Policy implications are proposed to improve ridesplitting services.

4.1. Definitions and notations

The proposed model has the following assumptions, and the related notations are shown in Table 2.

- All passengers in the ridesourcing data set are willing to choose a ridesplitting service.
- More than 90% of ridesplitting trips consist of only two shared rides because of DiDi Chuxing's rule that a driver is only allowed to accept at most two ride requests simultaneously. In addition, the number of passengers for each shared ride is limited to no more than two. Therefore, it is reasonable to assume that at most two trips can be combined in modeling.

Table 2
Notations of all variables.

Notation	Comment
$T = \{T_1, T_2, \ldots, T_n\}$	Trip set
$T_{i,\ j}$	Shared trips T_i and T_j
$T.o$, $T.d$	Pick-up, drop-off locations of the trip
t_i^s, t_i^e	Desired pick-up, drop-off times of the trip
t_i^o, t_i^d	Actual pick-up, drop-off times for the trip after ridesplitting
$t\ (i,j)$	Travel time between node i and node j
δ	Time horizon of the search time for ridesplitting
Δ	Maximum tolerable delay for passengers
$L = {}_{i,\ j\in\{1,2\ldots,n\}}L(T_{i,\ j})$	Link set that represents feasible ridesplitting
S	Shareability network
M	Matching of the shareability network S
M_{max}	Maximum weighted matching of the shareability network S
S'	Intersection graph
$V = \{V_1, V_2, \ldots, V_n\}$	Node set of intersection graph S'
R	Independent set of intersection graph S'
R_{max}	Maximum weighted independent set of intersection graph S'
Delay (T_i)	Delay for T_i (min)
$V_{delay\text{-}Ti}$	Value of the delay (T_i)

4.2. Model description

Let $S = (T, L)$ be the shareability network defined as follows. Node set $T = \{T_1, \ldots, T_n\}$ corresponds to the set of all trips. If trip nodes T_i and T_j can be combined, the shared trip can be denoted as $T_{i,j}$, and we can add link $L(T_{i,\ j})$ to connect the two nodes. We define link set $L = {}_{i,\ j\in\{1,2\ldots,n\}}L(T_{i,\ j})$ as all feasible trips for ridesplitting.

Whether trips T_i and T_j can be combined or not depends on the spatial/temporal properties of the two trips and on an upper bound Δ. If

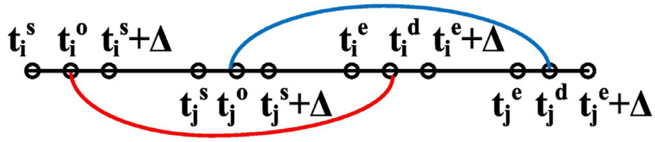

Fig. 3. Ridesplitting trip example.

the waiting time and delay time after ridesplitting are both less than the upper bound Δ for passengers, this means that the pick-up times and drop-off times of the 2 trips after ridesplitting are within the given time window, as shown in Fig. 3 below, and the 2 trips can be regarded as feasible for ridesplitting and be added to the link set.

As shown in Fig. 4, there are 4 situations for ridesplitting with two arbitrary trips, T_i and T_j. If any of the above 4 situations satisfies the time window constraint, T_i and T_j can be combined into $T_{i,j}$. Specifically, the constraints are as follows:

$$\boldsymbol{t_i^s} \le \boldsymbol{t_i^o} \le \boldsymbol{t_i^s} + \Delta \tag{1}$$

$$\boldsymbol{t_i^e} \le \boldsymbol{t_i^d} \le \boldsymbol{t_i^e} + \Delta \tag{2}$$

$$\boldsymbol{t_j^s} \le \boldsymbol{t_j^o} \le \boldsymbol{t_j^s} + \Delta \tag{3}$$

$$\boldsymbol{t_j^e} \le \boldsymbol{t_j^d} \le \boldsymbol{t_j^e} + \Delta \tag{4}$$

After deriving the link set L of combinable trips, the shareability network $S = (T, L)$ forms. Given a shareability network $S = (T, L)$, M is a subset of S. If there is no repeated node on either link of M, then M is said to be a match of S.

For example, given trip $T = \{T_1, T_2, T_3, T_{4,} T_5, T_6\}$, known (T_1, T_2), (T_2, T_3), (T_1, T_5), (T_2, T_6), (T_2, T_4), and (T_5, T_6) are feasible for ridesplitting. Then, shareability network S consists of 6 links: $L = \{L_1, L_2, L_{3,} L_4\ L_{5,}\ L_6\} = \{L(T_{1,2}), L(T_{2,3})\ L(T_{1,5})\ L(T_{2,6})\ L(T_{2,4})\ L(T_{5,6})\}$, as shown in Fig. 5.

We can give relevant weights to the links according to the optimization objective. Therefore,$\omega(\mathrm{L}_i)$, the weight of link $\{T_i, T_j\} \in L$, can be defined as follows:

(1) For the objective of maximizing the number of ridesplitting trips, the weight $\omega(\mathrm{L}_i)$ of each edge L_i is defined as the number of trips, which is 2 in all cases in this paper.

$$\omega(\mathrm{L}_i) = 2 \tag{5}$$

(2) For the objective of maximizing the total time of all trips, the weight $\omega(\mathrm{L}_i)$ of each edge L_i is defined as the time saved by ridesplitting.

$$\omega(L_i) =_{i\varepsilon\{1,2,\ldots,n\}} t\ (T_i) + t\ (T_j) - t\ (T_{ij}) \tag{6}$$

(3) For the objective of maximizing the value saved by ridesplitting, the weight $\omega(\mathrm{L}_i)$ of each edge L_i is defined as the value saved by ridesplitting.

$$\omega(\mathrm{L}_i) = P_1{*}saved\ time + P_2{*}saved\ distance - P_3{*}delay\ time \tag{7}$$

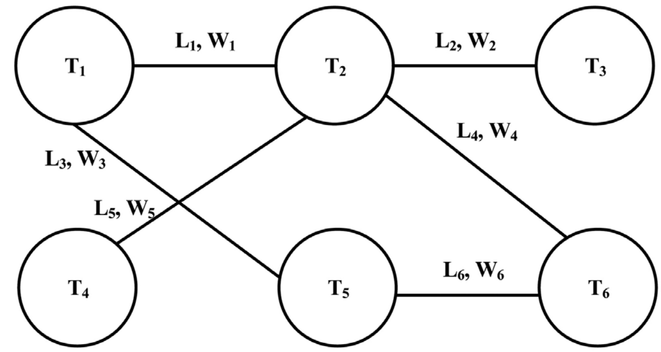

Fig. 5. Shareability network.

$$Saved\ \ \mathrm{time}\ (\mathrm{T_{i,j}}) = \mathrm{t}\ (\mathrm{T_i}) + \mathrm{t}\ (\mathrm{T_j}) - \mathrm{t}\ (\mathrm{T_{i,j}}) \tag{8}$$

$$Saved\ \ \mathrm{distance}\ (\mathrm{T_{i,j}}) = \mathrm{S}\ (\mathrm{T_i}) + \mathrm{S}\ (\mathrm{T_j}) - \mathrm{S}\ (\mathrm{T_{i,j}}) \tag{9}$$

$$Delay(T_{i,j}) = Delay\ (T_i) + Delay\ (T_j) = (t_i^d - t_i^e) + (t_j^d - t_j^e) \tag{10}$$

$$\omega(L_\mathrm{i}) = P_1{*}[\mathrm{t}\ (\mathrm{T_i}) + \mathrm{t}\ (\mathrm{T_j}) - \mathrm{t}\ (\mathrm{T_{i,j}})] + P_2{*}[\mathrm{S(T_i)} + \mathrm{S(T_j)} - \mathrm{S(T_{i,j})}] - P_3{*}(t_i^d - t_i^e + t_\mathrm{j}^d - t_\mathrm{j}^e) \tag{11}$$

where P_1 represents the average price of time, P_2 represents the average price of distance, P_3 represents the value of travel time, and Delay ($\mathrm{T_i}$) represents the delay time for T_i.

Specifically, the average price of time P_1 and the average price of distance P_2 can be calculated by the price of the DiDi ridesourcing service in Chengdu, as shown in Table 3. Thus, for the research time interval, $\mathrm{P_1} = 0.3$ RMB/min and $\mathrm{P_2} = 1.7$ RMB/km.

Another step is the calculation of the value of the delay time experienced by passengers in Chengdu. Becker (1965) uses the time allocation theory to estimate the value of time and proposes that the value of travel time equals the wage rate. Wardman (1998) states that the hourly wage rate can be used to determine the time value. When using the hourly wage rate method to estimate the value of travel time, it should be multiplied by the reduction factor α for revision since the travel time saved is usually not fully used for work. Miller (1989) proposes a reduction factor close to 60%. Small et al. (2012) proposes that the reduction factor can be taken as the interval value of 20% to 100%.

$$P_3 = \alpha{*}\frac{Income}{T} \tag{12}$$

where *Income* represents the per capita annual income and T represents the per capita working hours.

According to the Chengdu Statistical Yearbook 2017 compiled by the Chengdu Statistic Bureau, National Bureau of Standards Survey Office in Chengdu, the average wage of fully employed persons in urban units in 2016 was 61,330 RMB. The average time spent working per year is 250 days/year * 8 h/day = 2000 h/year. The wage rate of people in Chengdu is 61,330/2000 = 30.67 RMB/h. We set reduction

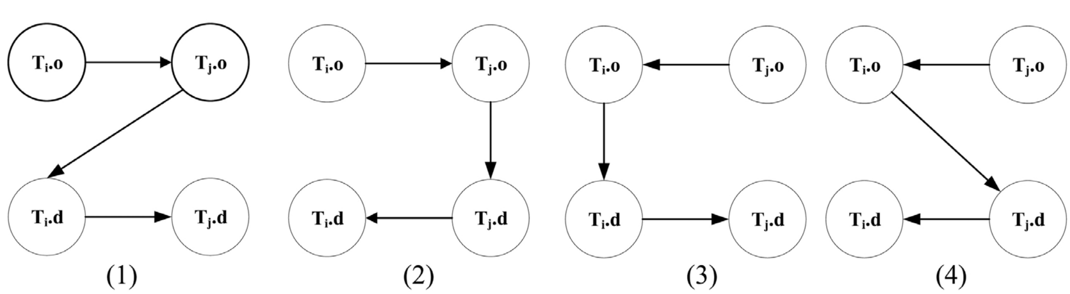

Fig. 4. Four situations of ridesplitting with two rides.

Table 3
Price of the DiDi ridesourcing service in Chengdu.

Time	Fee per kilometer (RMB)
23:00–07:00	2.4
Other time	1.7

Time	Fee per minute (RMB)
7:00–10:00 & 17:00–19:00	0.38
Other time	0.3

factor α as 0.6. Thus, $P_3 = 0.307$ RMB/min.

$$\omega(L_i) = 0.3*[t(T_i) + t(T_j) - t(T_{i,j})] + 1.7*[S(T_i) + S(T_j) - S(T_{i,j})] - 0.307*(t_i^d - t_i^e + t_j^d - t_j^e) \tag{13}$$

Ridesourcing companies such as DiDi and Uber must respond to trip requests in real time. Trip requests in a short time window can be known, and we set this window as δ. Passengers need to know whether they will get a shared ride within time window δ. For trips T_i and T_j, only when $|t_i^s - t_j^s| \le \delta$ can we connect the two trips as a feasible ridesplitting trip. Obviously, if we set a larger δ, there are more sharing opportunities for trips. However, δ should be reasonably small to offer passengers timely service. Therefore, we set $\delta = 1$ min in the model.

4.3. Algorithm of the shareability network

M with the $Max\{\Sigma\omega(L_i)\}$ is the maximum weighted matching of the shareability network S, denoted as M_{max}, which corresponds to the optimal ridesplitting solution. The maximum weight matching problem of the shareability network can be solved efficiently by converting the shareability network into the intersection graph. Node V_i of the intersection graph corresponds to link L_i of the shareability network. We add a link if two nodes contain the same trip. The weight of each node $w(V_i)$ in the intersection graph equals the weight of the corresponding link $w(L_i)$. R is a subset of S'. If there is no link between two arbitrary nodes of R, R is said to be an independent set of the intersection graph. The R with the $Max\{\Sigma\omega(V_i)\}$ is the maximum weight independent set of the intersection graph S', denoted as R_{max}. Obviously, R_{max} equals the corresponding M_{max}. We can transform the shareability network into an intersection graph to solve the maximum weighted matching problem efficiently, as shown in Fig. 6.

In this paper, a hybrid iterative local search algorithm (ILS-VND) based on a variable neighborhood depth search is used to solve the large-scale ridesplitting problem. The specific steps of the algorithm are shown in Table 4. The input of the algorithm is the intersection graph G(V, E, W) and the maximum number of iterations, MaxIter. The algorithm follows the standard flow of the local iterative search

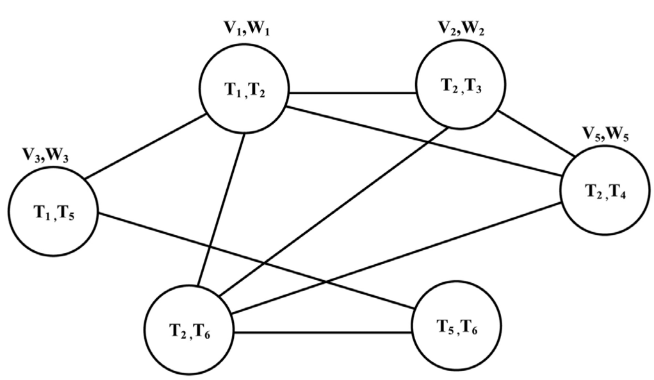

Fig. 6. Transformation into an intersection graph.

Table 4
Pseudocode under three objectives.

Objective: Maximizing shared trips, time savings or cost savings
Input: G(V, E, W), maximum number of iterations MaxIter
Output: The set of selected matching M_{max}, and the total weight of the set ω_{max}
1: $\omega(L_i)$
2: M_0 = Initialize(M)
3: **While** iter ≤ MaxIter **do**
Perturb(M)
LocalSearch(M)
4: **End while**
5: ω_{max} = Weight(M_{max})
6: Return M_{max},ω_{max}

algorithm. We add disturbance to the local optimal solution obtained by local search to avoid falling into the local optima. The current optimal solution is determined by the total weight of the selected set until the maximum number of iterations is reached.

5. Results and discussion

5.1. Results analysis

5.1.1. Actual scale of Ridesplitting

Based on the ridesplitting trip identification method of a previous study (Li et al., 2019) using this dataset, single and shared rides can be distinguished to obtain the actual current scale of DiDi ridesplitting services. As shown in Table 5, it can be found that only 7.85% of total ride orders are shared orders. The hours saved by ridesplitting account for 2.4% of the total ride hours. The percentage of cost savings is 1.22%. All of these values are fairly low. In addition, the average delay of actual ridesplitting is 9.86 min.

5.1.2. Potential of ridesplitting

Under the three different objectives, as the maximum tolerable delay increases, the optimal shared trips, time savings and cost savings continue to increase due to more opportunities for ridesplitting. We evaluate the performances of the three different objectives based on three indicators: shared trips, time savings and cost savings.

For the indicator of shared trips, there is no doubt that the objective of maximizing the number of shared trips has the best performance. As shown in Fig. 7, the potential opportunity can reach 91.37% with an average delay of 7.27 min, while the actual figure is only 7.85%. There is almost no gap between the objectives of maximizing the shared trips and maximizing the cost savings. Under the objective of maximizing cost savings, the percentage of shared trips can reach 90.69%. Under

Table 5
Actual scale of ridesplitting.

Item	Number	Percentage
Single ride orders	959,513	92.15%
Shared ride orders	81,694	**7.85**%
Total orders	1,041,207	100%
Single ride hours	365,336.4 h	89.19%
Shared ride hours	44,293.0 h	10.81%
Total ride hours before ridesplitting	409,629.4 h	100%
Single trip hours	365,336.4 h	91.36%
Ridesplitting trip hours	34,548.5 h	8.64%
Total trip hours after ridesplitting	399,884.9 h	100%
Hours saved by ridesplitting[a]	9744.5 h	**2.38**%
Total trip cost before ridesplitting	21,300,728.8 RMB	–
Single trip cost after ridesplitting	18,997,492.8 RMB	90.82%
Ridesplitting trip cost after ridesplitting	2,043,811.4 RMB	9.18%
Total trip cost after ridesplitting	21,041,304.2 RMB	100%
Cost saved by ridesplitting[b]	259,424.6 RMB	**1.22**%

[a] Percentage of saved hours within the total ride hours before ridesplitting.
[b] percentage of cost savings within the total cost before ridesplitting.

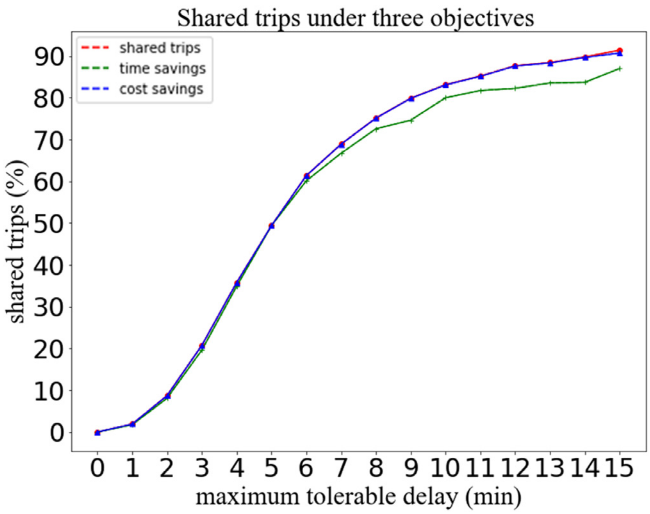

Fig. 7. Shared trips under the three objectives.

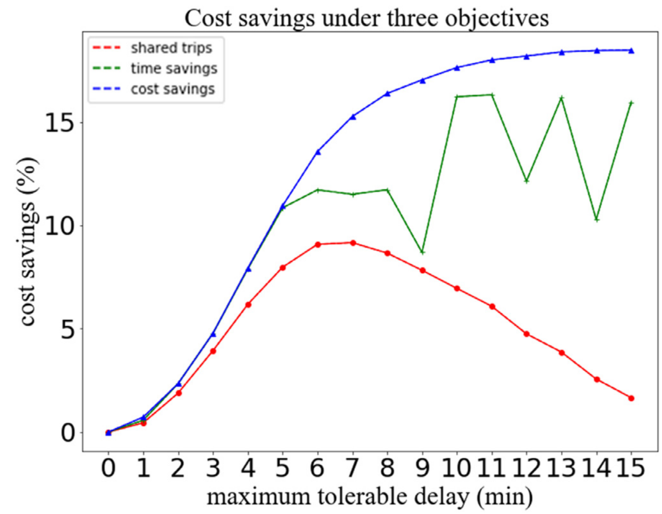

Fig. 9. Cost savings under the three objectives.

the objective of maximizing the time savings, the percentage of shared trips is a little less than that under the other two objectives, which can reach 86.99%.

For the indicator of time savings, the percentage can reach 25.77% with an average delay of 5.62 min under the objective of maximizing the time savings, see Fig. 8. By contrast, the actual figure is only 2.38%. The percentages of time savings under the objectives of maximizing the time savings and maximizing the cost savings are basically the same. This is because the travel time is also taken into consideration under the objective of maximizing the cost savings, where the percentage of time savings can reach 25.75%. Under the objective of maximizing shared trips, the percentage of time savings is 13.5%, much less than that under the other two objectives. In addition, the percentage of time savings goes down slightly when the maximum tolerable delay is more than 11 min. This is because it only considers the number of shared trips rather than the time savings, which can result in longer detours.

For the indicator of cost savings, the percentage of time savings can reach 18.47% with an average delay of 4.76 min under the objective of maximizing the cost savings, see Fig. 9. By contrast, the actual figure is only 1.22%. Under the objective of maximizing the time savings, the percentage of cost savings can reach 16.32% with an average delay of 4.98 min. The percentage of cost savings varies between 10% and 15% under the objective of maximizing the time savings. This is because there is no consideration of the delay time of passengers, which can also impact the cost savings. Under the objective of maximizing shared trips, the percentage of cost savings can reach 9.16%. Furthermore, the percentage of cost savings goes down gradually when the maximum tolerable delay is more than 7 min. This is because it only considers the number of shared trips rather than the time savings and delay time of passengers, which can bring about larger costs.

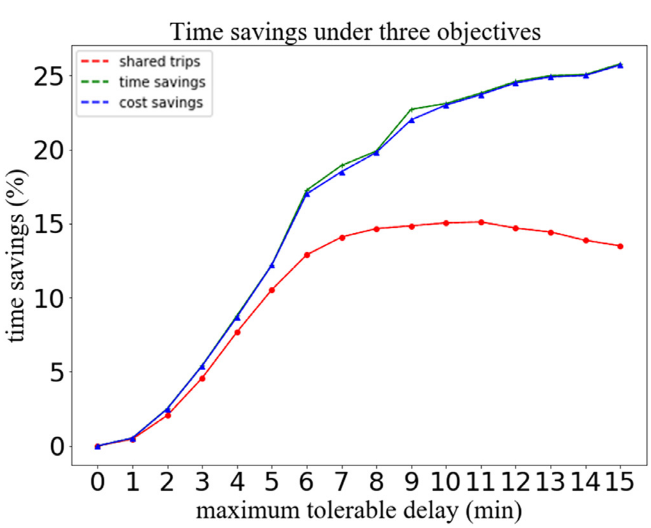

Fig. 8. Time savings under three objectives.

As shown in Table 6, the average delay increases with the increase of the maximum tolerable delay. When the maximum tolerable delay increases to 15 min, the average delay is 4.76 min under the objective of maximizing the cost savings, while the average delay reaches 7.27 min and 5.62 min under the objectives of maximizing the shared trips and the time savings, respectively. Obviously, the average delays of the potential ridesplitting solutions under the three objectives are all lower than the delay of 9.86 min in the real world. Therefore, it can be found that passengers can wait for less time on average when using the ridesplitting algorithm proposed in this paper. Furthermore, the average delay under the objective of maximizing the cost savings is the lowest of the three objectives.

5.2. Discussion

5.2.1. Benefits

To evaluate the potential benefits of ridesplitting, we compare the potential of ridesplitting under three objectives with the actual scale. We summarize and discuss the main findings in this section.

Table 6
Average delay under the three objectives.

Maximum tolerable delay (min)	Shared trips	Time savings	Cost savings
1	0.48	0.48	0.47
2	1.14	1.12	1.11
3	1.72	1.70	1.68
4	2.34	2.27	2.21
5	2.9	2.78	2.69
6	3.45	3.36	3.12
7	3.98	3.85	3.46
8	4.44	4.32	3.78
9	4.93	4.91	4.03
10	5.4	4.73	4.23
11	5.83	4.98	4.39
12	6.3	5.59	4.53
13	6.66	5.35	4.64
14	6.91	6.13	4.73
15	7.27	5.62	4.76

(i) Ridesplitting has great potential to mitigate congestion and air pollution: under the objective of maximizing cost savings, the percentage of potential shared trips is 90.69%, while the actual scale is 7.85%. The required fleet size can also be reduced considerably by ridesplitting because two shared trips can be served by a single vehicle instead of two vehicles. Therefore, shared trips can be used to reduce the required fleet size to mitigate traffic congestion for society. This also benefits passengers by enhancing their mobility. Zheng et al. (2018) found that ridesplitting in on-demand services can lead to a reduction of 12,327 vehicles in Hangzhou in the future, representing nearly 0.53% of all vehicle ownership in Hangzhou. Furthermore, the potential time savings afforded by ridesplitting can reach 25.75%, whereas this figure is 2.38% in the real world. Sperling (2018) stated that pooling to fill the empty seats in all vehicles is the most important strategy and innovation for achieving sustainable transportation.

(ii) Maximum tolerable delay is a significantly important indicator to maximize benefits. With the increase of the maximum tolerable delay, the percentages of shared trips, time savings and cost savings increase because of there being more opportunities for ridesplitting. However, the average delays also increase due to the increase of maximum tolerable delay. Therefore, trade–off solutions should be proposed to avoid overly long delays and guarantee the travel time reliability for passengers. Specifically, under the objective of maximizing cost savings, the cost savings rise significantly before the maximum tolerable delay increases to 12 min. The percentage of cost savings basically maintains the same value when the maximum tolerable delay is more than 12 min. Therefore, policy makers and on-demand ridesplitting companies can consider the marginal benefit when setting the maximum delay in matching and routing algorithms.

(iii) According to the results of the above three indicators, it can be found that the objective of maximizing the cost savings has better outcomes than the objectives of maximizing the shared trips or time savings. This is because maximizing the cost savings takes into account not only the value of time savings and saved vehicle miles but also the value of passenger delays. From the perspective of the whole system, we want to minimize the total time and travel distance of all trips. However, from the perspective of individual passengers, the optimal solution for the whole society may not necessarily be optimal for themselves, since the travel time does not include delays caused by detours. This objective also takes the value of the delay experienced by individual passengers into consideration to guarantee the rights of passengers when using the ridesplitting service. Therefore, the objective is more reasonable than the other two objectives.

5.2.2. Limitations

There are some limitations to the current study, which can motivate a few future research directions. First, it assumes that all the passengers using the ridesourcing service are willing to use a ridesplitting service. In fact, the percentage of passengers who choose ridesplitting services is related to the quality of the ridesplitting service such as the delay time. Second, a constant speed is used to compute the travel time. The speed should be dynamic according to the road segments and time sequences in a follow-up study. Finally, the density of vehicles can also impact the results. The proposed ridesplitting algorithm can be used in different cities to make a comparison.

5.2.3. Recommendations

Upon exploring the actual scale and potential of ridesplitting during rush hours, it can be found that there is a large gap between the actual scale and potential opportunity. Specific strategies are recommended for ridesplitting companies and policy makers to improve the present ridesplitting service, as follows:

- Ridesplitting companies (e.g., Didi Chuxing) should improve matching and routing algorithms to reduce the extra delays and improve the travel time reliability of the ridesplitting service.
- Ridesplitting companies should establish credit rating systems for both passengers and drivers to discourage tardy behavior. Then the ridesplitting system can operate more efficiently and the travel time reliability can also be improved.
- Ridesplitting companies should subsidize drivers and passengers proportionately to improve their enthusiasm for and participation in ridesplitting. Additionally, transportation management agencies should provide more policy incentives to promote the use of ridesplitting, such as high-occupancy vehicle lanes and waivers for road fees and parking fees.

6. Conclusions and future research

This paper aims to improve the ridesplitting service during rush hours using empirical ridesourcing data provided by DiDi Chuxing, which contains complete datasets on ridesourcing orders in the city of Chengdu, China. A ridesplitting trip identification algorithm based on a shareability network is developed to quantify the potential of ridesplitting. We analyze shared trips, time savings and cost savings under three objectives. Furthermore, we evaluate the gap between the potential of ridesplitting and the actual scale of ridesplitting. The main findings are as follows:

The percentage of potential cost savings can reach 18.47% with an average delay of 4.76 min, whereas the actual percentage is 1.22% with an average delay of 9.86 min. Passengers can wait less time on average when using the proposed ridesplitting algorithm. The percentage of shared trips can be increased from 7.85% to 90.69%, and the percentage of time savings can reach 25.75% from 2.38%.

To our knowledge, this is the first investigation of the gap between the actual and the potential of ridesplitting on a city scale. The findings of this study can help reveal the great potential of ridesplitting. The quantitative benefits could encourage governments to take more action to improve the present ridesplitting services such as by providing high-occupancy vehicle lanes and waivers for road fees and parking fees. The comparative evaluation between the potential and actual scale of ridesplitting provides policy implications for improving existing ridesplitting services. The ridesplitting companies should establish credit rating systems to improve the efficiency of ridesplitting services. Furthermore, the maximum tolerable delay should be set properly for policy makers to obtain more benefits for low-carbon urban transport and guarantee the ridesplitting service quality at the same time.

The “shareability network” method can solve the dynamic ridesplitting problem on a large city scale efficiently. However, there are some limitations to the current study such as the willingness to use the ridesplitting service, the constant speed and simulation hypotheses. The authors recommend that future studies focus on these issues.

Acknowledgments

This research was funded by the National Natural Science Foundation of China [grant numbers: 71774118, 71734004, 51508484], the Science and Technology Commission of Shanghai Municipality [grant numbers: 16511105204, 17511105200, 17511105203], and the Guangzhou Major Innovation Project [grant number: 201704020201].

References

Alexander, L.P., González, M.C., 2015. Assessing the impact of real-time ridesharing on urban traffic using mobile phone data. In: Proc. UrbComp, pp. 1–9.
Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., Rus, D., 2017. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. Proc. Natl. Acad. Sci. 114, 462–467.
Baker, B.M., Ayechew, M., 2003. A genetic algorithm for the vehicle routing problem.

Comput. Oper. Res. 30, 787–800.
Baldacci, R., Maniezzo, V., Mingozzi, A., 2004. An exact method for the car pooling problem based on Lagrangean column generation. Oper. Res. 52, 422–439.
Becker, G.S., 1965. A theory of the allocation of time. Econ. J. 493–517.
Braekers, K., Caris, A., Janssens, G.K., 2014. Exact and meta-heuristic approach for a general heterogeneous dial-a-ride problem with multiple depots. Transp. Res. B Methodol. 67, 166–186.
Casprini, E., Di Minin, A., Paraboschi, A., 2019. How do companies organize nascent markets? The BlaBlaCar case in the inter-city shared mobility market. Technol. Forecast. Soc. Chang. 144, 270–281.
Chen, X.M., Zahiri, M., Zhang, S., 2017. Understanding ridesplitting behavior of on demand ride services: an ensemble learning approach. Transp. Res. C Emerg. Technol. 76, 51–70.
Chen, X., Zheng, H., Wang, Z., Chen, X., 2018. Exploring impacts of on-demand ridesplitting on mobility via real-world ridesourcing data and questionnaires. Transportation 1–21.
Cortés, C.E., Matamala, M., Contardo, C., 2010. The pickup and delivery problem with transfers: Formulation and a branch-and-cut solution method. Eur. J. Oper. Res. 200, 711–724.
DiDi Chuxing, 2017a. DiDi Chuxing Corporate Citizenship Report 2017. https://www.didiglobal.com/news/newsDetail?id=384&type=news, Accessed date: 16 May 2018.
DiDi Chuxing, 2017b. What does the GAIA Initiative provide in the first phase? https://outreach.didichuxing.com/research/opendata/en/, Accessed date: 5 February 2018.
DiDi Chuxing, 2018. Daily rides of DiDi ExpressPool exceed 2.4 million. http://www.DiDichuxing.com/en/press-news/j3km2785.html, Accessed date: 16 May 2018.
Handke, V., Jonuschat, H., 2012. Flexible Ridesharing: New Opportunities and Service Concepts for Sustainable Mobility. Springer Science & Business Media.
Hara Associates, 2013. Best Practices Studies of Taxi Regulation: Taxi User Surveys. SFMTA, San Francisco. https://www.sfmta.com/sites/default/files/Draft%20SF%20UserSurvey%2055%20WEB%20version04042013.pdf.
Jin, S.T., Kong, H., Wu, R., Sui, D.Z., 2018. Ridesourcing, the sharing economy, and the future of cities. Cities 76, 96–104.
King, D.A., Saldarriaga, J.F., 2017. Access to taxicabs for unbanked households: an exploratory analysis in New York City. J. Public Transp. 20 (1).
Kirchler, D., Calvo, R.W., 2013. A granular tabu search algorithm for the dial-a-ride problem. Transp. Res. B Methodol. 56, 120–135.
Kleiner, A., Nebel, B., Ziparo, V.A., 2011. A mechanism for dynamic ride sharing based on parallel auctions. In: Twenty-second International Joint Conference on Artificial Intelligence.
Laporte, G., 1992. The vehicle routing problem: an overview of exact and approximate algorithms. Eur. J. Oper. Res. 59, 345–358.
Li, B., Krushinsky, D., Van Woensel, T., Reijers, H.A., 2016. An adaptive large neighborhood search heuristic for the share-a-ride problem. Comput. Oper. Res. 66, 170–180.
Li, W., Pu, Z., Li, Y., Ban, X.J., 2019. Characterization of ridesplitting based on observed data: A case study of Chengdu, China. Transp. Res. C Emerg. Technol. 100, 330–353.
Luo, X., Dong, L., Dou, Y., Li, Y., Liu, K., Ren, J., Liang, H., Mai, X., 2017. Factor decomposition analysis and causal mechanism investigation on urban transport CO2 emissions: comparative study on Shanghai and Tokyo. Energy Policy 107, 658–668. https://doi.org/10.1016/j.enpol.2017.02.049.
Lyft, 2018. Economic Impact Report. https://www.lyftimpact.com/impact, Accessed date: 16 May 2018.
Ma, S., Zheng, Y., Wolfson, O., 2013. T-share: a large-scale dynamic taxi ridesharing service. In: IEEE 29th International Conference on Data Engineering (ICDE), pp. 410–421.
Miller, T.R., 1989. The Value of Time and the Benefit of Time Saving [M]. Urban Institute.
Nogueira, B., Pinheiro, R.G., Subramanian, A., 2018. A hybrid iterated local search heuristic for the maximum weight independent set problem. Optim. Lett. 12, 567–583.
Parragh, S.N., Doerner, K.F., Hartl, R.F., 2010. Variable neighborhood search for the dial-a-ride problem. Comput. Oper. Res. 37 (6), 1129–1138.
Psaraftis, H.N., 1980. A dynamic programming solution to the single vehicle many-to-many immediate request dial-a-ride problem. Transp. Sci. 14, 130–154.
Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S., 2016. Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transp. Policy 45, 168–178.
Ropke, S., Cordeau, J.F., Laporte, G., 2007. Models and branch-and-cut algorithms for pickup and delivery problems with time windows. Networks 49 (4), 258–272.
Roskilly, A.P., Palacin, R., Yan, J., 2015. Novel technologies and strategies for clean transport systems. Appl. Energy 100 (157), 563–566.
Santi, P., Resta, G., Szell, M., Sobolevsky, S., Strogatz, S.H., Ratti, C., 2014. Quantifying the benefits of vehicle pooling with shareability networks. Proc. Natl. Acad. Sci. U. S. A. 111, 13290–13294.
Santos, D.O., Xavier, E.C., 2015. Taxi and ride sharing: a dynamic dial-a-ride problem with money as an incentive. Expert Syst. Appl. 42, 67286737.
Schrank, D., Eisele, B., Lomax, T., 2012. TTI' 2012 urban mobility report[J]. Texas A&M Transportation Institute. The Texas A&M University System.
Shaheen, S., Chan, N., 2016. Mobility and the sharing economy: potential to facilitate the first- and last-mile public transit connections. Built Environ. 42 (5), 573588.
Shaheen, S., Chan, N., Bansal, A., Cohen, A., 2015. Shared mobility. In: Definitions, Industry Developments, and Early Understanding. University of California Berkeley Transportation Sustainability Research Center, Berkeley.
Shaheen, S., Cohen, A., Zohdy, I., 2016. Shared Mobility: Current Practices and Guiding Principles[R].
Simonetto, A., Monteil, J., Gambella, C., 2019. Real-time city-scale ridesharing via linear assignment problems. Transp. Res. C Emerg. Technol. 101, 208–232.
Small, K.A., Winston, C., Evans, C.A., 2012. Road work: a new highway pricing and investment policy [M]. Brookings Institution Press.
Smith, C., 2018. Amazing Uber Statistics, Demographics and Facts. https://expandedramblings.com/index.php/uber-49 statistics/, Accessed date: 10 February 2018.
Sperling, D., 2018. Three Revolutions: Steering Automated, Shared, and Electric Vehicles to a Better Future. Island Press.
Uber Estimator, 2018. Uber Cities. https://uberestimator.com/cities, Accessed date: 10 February 2018.
Vazifeh, M., Santi, P., Resta, G., Strogatz, S., Ratti, C., 2018. Addressing the minimum fleet problem in on-demand urban mobility. Nature 557 (7706), 534.
Wardman, M., 1998. The value of travel time a review of British evidence. JTEP 32, 285–316.
Zheng, H., Chen, X., Chen, X.M., 2018. How does on-demand ridesplitting influence vehicle use and ownership? A case study in Hangzhou, China. In: Transportation Research Board 97th Annual Meeting.