

# Boosting Ride Sharing with Alternative Destinations

Vinicius Monteiro de Lira<sup>\*†‡</sup>, Raffaele Perego<sup>\*</sup>, Chiara Renso<sup>\*</sup>, Salvatore Rinzivillo<sup>\*</sup>, and Valeria Times<sup>†</sup>,  
<sup>\*</sup>ISTI-CNR (Italy) <sup>†</sup>Federal University of Pernambuco (Brazil) <sup>‡</sup>University of Pisa (Italy)

**Abstract**—People living in highly-populated cities increasingly suffer an impoverishment of their quality of life due to pollution and traffic congestion. With the objective of reducing the number of circulating vehicles, we investigate a novel approach to boost ride sharing opportunities based on the knowledge of the human activities behind individual mobility demands. We observe that in many cases the activity motivating the use of a private car (e.g., going to a shopping mall) can be performed in many different places. Therefore, when there is the possibility of sharing a ride, people having a pro-environment behavior or interested in saving money can accept to fulfill their needs at an alternative destination. We thus propose ACTIVITY-BASED RIDE MATCHING (ABRM), an algorithm aimed at matching ride requests with ride offers possibly reaching alternative destinations where the intended activity can be performed. By analyzing two large mobility datasets extracted from a popular social network, we show that our approach could largely impact urban mobility by resulting in an increase up to 54.69% of ride-sharing opportunities with respect to a traditional destination-oriented approach. Due to the high number of ride possibilities found by ABRM we introduce and assess a subsequent ranking step to provide the user with the top-k most relevant rides only. We discuss how ABRM parameters affect the fraction of car rides that can be saved and how the ranking function can be tuned to enforce pro-environment behaviors.

**Index Terms**—carpooling, ride sharing, flexibility, green mobility.

## I. INTRODUCTION

The increase of the fleet of vehicles in highly-populated cities brings detrimental consequences on the traffic congestion and the citizens' quality of life while the environment is harmed with the emission of dangerous pollutants. According to the International Business Times [1], large cities such as Beijing (China), Delhi (India) and Patna (India) are today smoggy places where pollution caused by industry and vehicular traffic has soared to hazardous levels around 20 times the limit recommended by the World Health Organization (WHO). A recent study by TomTom [2] shows how the traffic situation is often critical and heavily affects people mobility in cities like Istanbul (Turkey), Mexico City (Mexico), Rio de Janeiro (Brazil) and Moscow (Russia). Istanbul results to be the worst city in this unenviable ranking with a daily average delay of 29 minutes for a 30 minutes commute.

Large metropolitan areas have several problems and complexities but it is widely recognized that our individual behaviors have a substantial impact on the environment either in the positive and negative direction. Even a slight change in our mass mobility habits can lead to a large reduction of the number of circulating vehicles. The sharing of car journeys

in which a group of people commute together, commonly known as *Ride sharing* or *Carpooling* [3], [4] is one of the practices that can most benefit the reduction of urban traffic [5], [6], [7], [8], [9]. Worthily, carpooling recently started to attract the interest of car industries themselves<sup>1</sup>. Car manufacturers see, in fact, in the diffusion of this practice a challenging opportunity to make profit in modern urban transportation scenarios according to novel *Car-as-a-Service* business models.

Traffic congestion is increasingly becoming a crucial factor of decision in the people's life. When planning their daily activities people carefully consider *where to go*, *which time to go* and *how to go*, by considering how much delay they will incur given the expected traffic conditions. In addition, recent studies of human mobility highlight from one hand the significant potential for continued and future uptake of sustainable forms of urban mobility [10], [11], and, from the other hand, that the tendency to be regular or not in choosing the places where to perform some activities is an individual characteristic of people [12], [13].

Based on these observations, we hypothesize that changing habits can be rewarding under some aspect, and flexible people can accept to fulfill their mobility needs reaching alternative destinations when there is the possibility of sharing a ride thus saving time/money or enforcing their pro-environment behavior [11]. For example, a person may decide to go shopping in the morning after experiencing hassling traffic jams in the late afternoon, or even to go to a different shopping mall that is more easily reachable from her home or work place.

In this paper we investigate how to take advantage of a similar flexibility by proposing ACTIVITY-BASED RIDE MATCHING (ABRM), an algorithm aimed at matching ride requests with ride offers possibly reaching alternative destinations where the intended user activity can be performed. By exploiting the knowledge of the activity motivating ride requests, ABRM can boost at a high rate users' mobility demands by means of existing ride offers. The approach proposed is completely orthogonal and possibly complementary to popular ride sharing services like BlaBlaCar, UberPool and Lyft Line ([14], [15], [16]). Indeed, providing the user with activity-based ride options could enable novel business strategies to be incorporated in these services. For example, the service could support user's flexibility and increase her engagement by proposing the most convenient rides to alternative destinations where the intended activity can be performed.

Analyzing two large mobility datasets extracted from the FourSquare social network, we discovered that ABRM in-

This paper is co-funded by the MASTER project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie-Slodowska Curie grant agreement n. 777695.

<sup>1</sup>see for example: [goo.gl/d3tHdZ](http://goo.gl/d3tHdZ)

increases up to 78% the number of rides potentially satisfied with respect to traditional destination-oriented carpooling approaches. Note that the number of ride offers possibly matching an activity-based alternative destination varies on the basis of the generality and the characteristics of the activity performed. For example, the activity “shopping in a mall” may result in several different destinations, while daily activities such as “go to work place” or “go home” cannot benefit in any way from the proposed approach and are not considered as flexible ride requests in the experiments.

This paper extends our previous work on activity-based carpooling presented in [17], [18]. In [17] we firstly introduced the idea that carpooling can be based on the activity to be performed instead of a fixed destination, while in [18] we proposed a working demo and a mobile *app* providing an efficient implementation of the system. The relevant and original contributions of this paper, compared to these previous works and the state of art, are listed below:

- we propose a novel and improved version of the matching algorithm presented in [17]. While in previous work the algorithm used a approximate grid-based approach to model trips and temporal intervals, the current version precisely represents the information about the spatial location and temporal dimension.
- we propose a ranking model based on a linear combination of ride matching features and we assess it by investigating how the weighted combination of features compares with extreme scenarios where one of the features prevails over all the others.
- we design and run extensive experiments exploiting two new large-scale mobility datasets extracted from FourSquare in New York City and Tokyo<sup>2</sup>. We detail the methods used to build the datasets and enrich them with information coming from heterogeneous sources (FourSquare, Google maps, OpenStreetMap); Although the datasets used do not involve real ride sharing requests/offers, they are representative of real user activities and mobility demands for these activities at a large scale.
- we analyze and discuss in depth the potential environmental impact of our ABRM solution and we show how the ranking algorithm can be tuned to enforce the pro-environment behavior of users.

The rest of the paper is organized as follows. Section II discusses related works. Section III introduces some concepts used through the paper and define our activity-based alternative destinations ride problem. In Section IV we describe the algorithm for matching ride requests with alternative destinations and discuss the ranking model adopted. We discuss the experiments in Section V and draw conclusions envisioning future works in Section VI.

## II. RELATED WORKS

Ride sharing is a mode of transportation consisting of two (or more) persons sharing a vehicle to move along similar itineraries and time schedules [19], [20], [21]. A particular

kind of ride sharing is *carpooling*, where one of the user share her own car with other passengers.

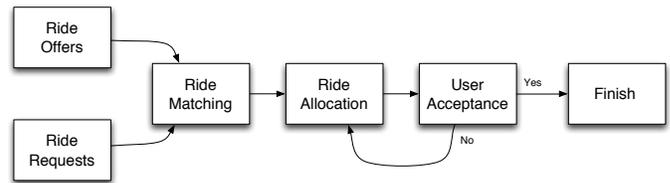


Fig. 1: A *Ride sharing* conceptual schema

A ride sharing system can be described by the schema showed in Figure 1. The system takes in input two sets: *ride offers* of users willing to share their trips, and *ride requests* of users searching for a lift to a destination. A *Ride Matching* algorithm is used to find a subset of ride offers that may supply a ride to given ride request. In this phase the system finds all matchings that are compatible with spatial and temporal constraints specified by the user. After the possible matchings have been identified, the *Ride Allocation* phase use some criteria to allocate ride requests to offers, limited by the vehicle capacity and based on the matchings.

Since it is desirable that both passengers and drivers find a ride matching that best meets individual preferences, the *User Acceptance* phase checks which suggestions have been accepted by the users and, in case of rejection, looks for an alternative allocation repeating the previous step of *Ride Allocation*. To maximize the probability of successfully matching, it is crucial to provide the *Ride Allocation* phase with a large set of candidates. A largely adopted solution for reaching this objective consists in relaxing spatio-temporal constraints of ride offers and requests [19], [22], [20], [23]. Our approach, as most of the works in the literature, focuses on the matching phase. A common approach to increase the number of matchings is to consider some **temporal flexibility** in the participants (driver and rider) requests: the participants specify an earliest possible departure time and latest possible arrival time. A recent work showed that even a small flexibility in terms of desired departure time or maximum detour time, can significantly impact the expected matching rate, especially when the number of ride offers in the system is not large [19].

Analogously, other approaches explore the **spatial flexibility** of users in terms of ride detours or walking distance to catch the ride. Detour ride sharing considers a scenario where the driver accept to make a detour from her original itinerary if this brings a satisfactory ride matching[22], [20], [19]. Geisberger et. al in [22] propose an innovative detour route planning algorithm able to find a reasonable match for more than 60% of the requests not satisfied by traditional matching strategies. Orthogonally, some works investigated the possibility that the passenger could walk to a meeting point to join a ride [21], [24], [25]. This approach is also known as *slugging*. In [21], the authors formally define the slugging problem and propose a close to optimal heuristic for the the matching problem considering vehicle capacity constraints. They performed experiments using real taxi cabs

<sup>2</sup>The datasets and the associated metadata will be made publicly available upon acceptance of the submission.

trips in Shanghai, and reported a saving up to 70% in the total distance traveled by vehicles. Minett and Pearce [26] investigate the impact of slugging in gasoline consumption. They estimate that slugging in San Francisco can save from 1.7 to 3.5 million liters of gasoline per year, much of which comes from the indirect impact on the rest of the traffic. In [27] a solution that exploits slugging in areas with High-Occupancy Vehicle lanes (HOV) is proposed. The authors also address some shortcomings associated with casual carpooling such as personal safety, the free-rider problem, and the maximization of the number of passengers sharing a ride. Other works investigate how to combine more than one ride offer to supply a single ride request. The results discussed in [28] show that such *multi-hop* ride sharing can significantly increase the number of requests satisfied. A *multi-modal* ride sharing approach considering a combination of multi-hop and slugging is considered in [24]. Experiments are conducted on real taxis trips in New York and report a reduction up to 40% in the number of trips. In this paper, compare to the ones presented above, we try to further increase the matching possibilities exploiting not only the spatial and temporal flexibility but also the alternative destinations based on the user desired activity.

Most of related works do not consider in their models the activity as a flexibility attribute, except the work by Cho et Al. [29] that first addresses activity-based carpooling. The authors propose the use of an ontology in an activity-based microsimulation. While no explicit evidence is presented, the focus of the paper is recognizing that the ontology is a useful and appropriate method for activity-based microsimulation research. Indeed, only a conceptual design and framework are suggested, and this study is a clearly preliminary step.

Social aspects are considered in [30], [31]. The authors of [31] propose a recommender system for carpooling services that leverages on learning-to-rank techniques to automatically derive a personalized ranking model for each user from the history of her choices (i.e., the type of accepted or rejected shared rides). The system builds the list of recommended rides by maximizing the estimated success rate of the offered matches extracted from Foursquare check-in information.

Finally, it is worth discussing the allocation problem aimed at maximizing the number of matchings by considering spatio-temporal and vehicle capacity constraints. Although the allocation problem is very important for the deployment of an actual ride sharing system (see for example [32], [33], [34], [7], [35]), it is not addressed in this paper. We note, in fact, that the allocation task asks for choosing a small subset of ride offers potentially satisfying a given request. The ABRM algorithm largely increases the number of candidate rides and it is not clear, in this new context, how to address ride allocation without making very strong and unrealistic assumptions on users' flexibility. Similar to [31] we preferred, instead, to adopt a ranking-based solution which orders the potential ride-offers on the basis of a weighted combination of a set of features modeling the different aspects of user flexibility. An advantage of this solution is that by sweeping these weights we can study and better understand the effect of user flexibility in accepting changes involving the three dimensions considered: the desired departure time, the distance of the pick up point and, more

importantly, the destination proposed.

### III. BASIC CONCEPTS AND PROBLEM FORMULATION

We introduce here the formulation of our ACTIVITY-BASED RIDE MATCHING problem and some basic concepts and notations used through the paper. A *trajectory* (or *trip*) represents the spatio-temporal movement evolution of a traveling object. It is generally detected by a tracking device into samples thus bringing a discrete representation of the movement. A more formal definition is given below.

*Definition 1 (trajectory):* A *trajectory* represents a finite sequence of spatio-temporal points assigned to a moving object and denoted by  $\langle objId, \langle x_1, y_1, t_1 \rangle, \dots, \langle x_n, y_n, t_n \rangle \rangle$ , where *objId* is the identifier of the moving object,  $x_i, y_i, t_i$  represent the spatial and temporal coordinates of the sample points.

Trajectories may have stops at venues or Points Of Interest (POIs), e.g., a shop, a bar, a restaurant or a gym [36]:

*Definition 2 (Point Of Interest (POI) or venue):* A POI is a geographical object, usually associated to a human activity, that is interesting for a specific application. We define a POI as a tuple  $\langle s, n, c \rangle$  where *s* is the representative spatial point, *n* is the name of the POI and *c* is its category drawn from a defined taxonomy *C*.

A trajectory can pass or stop close to a set of POIs where some *activity* (e.g., shopping, visiting, eating, working, exercising) can be performed.

*Definition 3 (Activity):* An *Activity* defines a task that can be performed at a POI. We assume that activities are related to specific POI categories and viceversa. Thus there is a mapping between a given activity *a* and a set of POI categories. For example, the activity *eating* is related to POI categories *Restaurants* and *Pizzeria*. Viceversa the POI category *Restaurants* is related to activities *eating* and *drinking*.

Given a POI *p*, we thus assume to be able to find a number of alternative venues where the activity performed in *p* can be performed as well. We call these POIs the *alternative destinations* for *p*.

*Definition 4 (Alternative Destination):* Given a POI *p*, the set of alternative destinations for *p* is a set of POIs  $\langle p_1, \dots, p_n \rangle$  where it is possible to perform the same activity as in *p*. These alternative destinations can be all the POIs belonging to the same category of *p* or a subset of them selected on the basis of some *criterium*, e.g., those most similar to *p*, or the most popular, or the ones preferred by the user.

The ABRM problem addressed assumes that a *ride request* for going to a POI *p* from a given spatial location *loc* at a given time *t* can be served with *ride offers* respecting the time and starting location constraints of the request and possibly dropping the user to one of the alternative destinations for *p*. We formally define *ride requests* and *ride offers* as follows.

*Definition 5 (Ride Request):* A Ride Request *q* is a tuple:  $\langle u, loc, venueDest, time, w\_dist, delay \rangle$ , where *u* identifies the requesting user, *venueDest* is the POI to be reached, *loc* and *time* are the starting location and the preferred departure time, while *w\\_dist* is the maximum walking distance the user is willing to walk to get the ride and *delay* the maximum time the user is willing to anticipate or delay the departure.

**Definition 6 (Ride Offer):** A ride offer  $rt$  is a tuple:  $\langle u, orig, dest, time, path \rangle$  where  $u$  is the driver,  $orig$  and  $dest$  are the fixed origin and destination of the ride offered by  $u$ ,  $time$  is the departure time, and  $path$  is the route followed by the vehicle offering available seats. These seats can be offered to passengers traveling from origins to destinations that are reachable along the vehicle route.

It is worth noticing that our definition of ride offer is intentionally generic to encompass different ride-sharing scenarios: the vehicle with available seats following a fixed route at a fixed time could be a private car (e.g., routinely going home from work at 5pm of working days), or a shared taxi serving a ride. We can now proceed in formulating our ACTIVITY-BASED RIDE MATCHING problem:

**Definition 7 (ACTIVITY-BASED RIDE MATCHING):** Given a set of ride offers  $RT$  and a ride request  $q = \langle u, loc, venueDest, time, w\_dist, delay \rangle$ , the ABRM problem asks to find all the matchings  $\langle m_1, \dots, m_k \rangle$  between  $q$  and the ride offers in  $RT$  that allow the passenger to reach  $p$  or an alternative destination for  $p$  within the maximum walking distance  $w\_dist$  and the maximum departure delay or anticipation  $delay$ .

#### IV. ACTIVITY-BASED RIDE MATCHING

The example in Figure 2 illustrates a simple instance of our matching problem. The request  $q$  of user  $u$  is for POI  $v_1$  starting from location  $p$  at time  $t$  with a maximum walking distance of 500m and a temporal flexibility of 30 min. The alternative destinations for  $v_1$ , preserving the activity to be done, are POIs  $v_2, v_3, v_4, v_5$ . The circles around the POIs represent the area within the walking distance the user set in her request (e.g., 500 meters). The set  $RT$  does not offer any ride to  $v_1, v_2$  and  $v_3$  satisfying the time and starting location constraints of  $u$ . There is, however, a ride offer  $rt \in RT$  that intersects the circles around  $p, v_4$ , or  $v_5$  in the order.

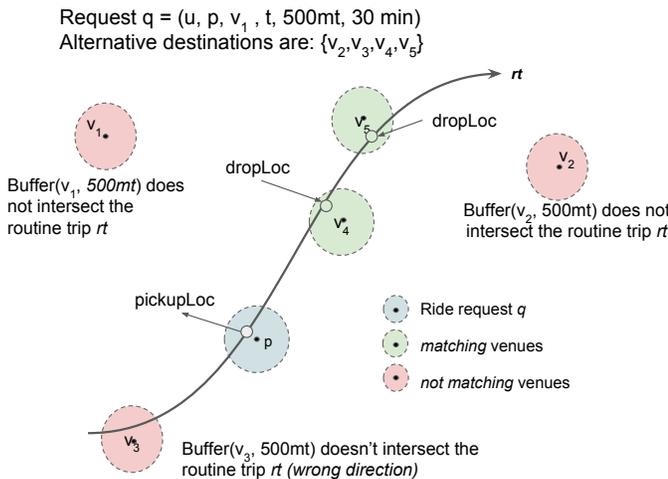


Fig. 2: A ride matching example

We have thus two possible matches: one possibility is to pick up the ride  $rt$  from  $pickupLoc$  to  $v_4$  and another

#### Algorithm 1: ACTIVITY-BASED RIDE MATCHING

```

Input :  $q = (u, loc, venueDest, time, w\_dist, delay)$  % ride request
          $RT$  % set of ride offers
          $V$  % set of POIs
Output:  $\mathcal{M}$  % set of matchings

1 begin
2    $\mathcal{M} \leftarrow \emptyset$ ;
3    $altVenues \leftarrow AlternativeDest(q.venueDest, V)$ ;
4   foreach  $rt \in RT$  do
5     if  $Distance(q.loc, rt.traj) \leq q.w\_dist$  then
6        $m.pickupLoc = closestPoint(q.loc, rt.traj)$ ;
7        $m.pickupTime = timeAt(rt.traj, m.pickupLoc)$ ;
8       if  $(|q.time - m.pickupTime| \leq q.delay)$  then
9         foreach  $vdest \in altVenues$  do
10          if  $Distance(vdest.s, rt.traj) \leq q.w\_dist$  then
11             $m.dropLoc = closestPoint(vdest.s, rt.traj)$ ;
12             $m.droptime = timeAt(rt.traj, m.dropLoc)$ ;
13          if  $m.droptime > m.pickupTime$  then
14             $m.dest = vdest$ ;
15             $\mathcal{M} \leftarrow \mathcal{M} + m$ ;
16 return  $\mathcal{M}$ 

```

possibility is take the same ride up to  $v_5$ . We call these ride possibilities the *matchings*. We note that a request may have different matchings, not only with different ride offers but also with the same ride to different alternative destinations. Each matching has a destination venue *altDestination*, a pickup location *pickupLoc* and time *pickupTime*, a drop-off location *dropLoc* and time *dropTime*. We specify that our approach is not based on the detour of the ride that remain fixed and the driver does not change her path to pickup the passenger. It is the passenger who moves from the location of the request to the *pickupLoc* point of the matching. The constraint is that *pickupLoc* and *dropLoc* have distances from *loc* and *altDestination* lower than the walk distance  $w\_dist$  specified in  $q$ . The pseudocode in Algorithm 1 illustrates the steps that find all the matchings  $\mathcal{M}$  for a user request  $q = (u, loc, venueDest, time, w\_dist, delay)$  in a set  $RT$  of ride offers.

The algorithm starts by finding the alternative destinations for *venueDest* with function *AlternativeDest()*. This function can be instantiated in several ways. A simple solution is to select all the POIs whose category is the same as the requested POI. For example, when the user requests a ride to a supermarket the alternative POIs are all the venues labeled with the "supermarket" category. Since the POIs in a given category can be many, we can restrict the number of alternative venues by choosing the  $k$  most popular ones or the  $k$  most similar to *venueDest* according to some similarity function.

The algorithm then iterates over the ride offers (line 4) checking if the starting location  $q.loc$  is within the walking distance from the path of a ride offer (line 5). In this case, the algorithm sets as pick up location the closest point between the offered ride path and the request location (line 6). Accordingly, the pick up time is computed as the time at which the offered trip passes in the pick up location (line 7). Then, if the candidate ride does not respect the time delay constraint for pick up it is discarded (line 8).

A similar computation is done for the drop point. The algorithm checks if the candidate ride passes within distance

$w\_dist$  from one of the alternative destinations (line 9-12) and checks that the direction of the ride is correct (line 13). In this case, the candidate ride is added to the matchings set (line 15).

Finally, the algorithm returns the set of matchings discovered (line 16).

Since the cardinality of this set can be high, we introduce a subsequent ranking step to order the matchings on the basis of an estimation of their relevance for the passenger (or on the basis of other criteria such as the saving of  $CO_2$ ). In the next section we discuss the ranking features and how we can combine them in a ranking function.

#### A. Ranking Model

In order to rank the matchings returned by Algorithm 1 we consider four features. The objective of these features is to measure how effectively a matching  $m$  fits a ride request  $q$ . The four features are:

**Time delay** ( $f_{dly}$ ). The *anticipation* or *delay* of the trip respect to the intended time specified in the request.

**Distance to walk** ( $f_{wld}$ ). The *distance* the passenger has to walk in order to get the ride and arrive at the proposed destination. It is computed as the sum of the distance between the passenger location and the pick-up point, plus the distance between the drop off point and the destination venue.

**Ride duration** ( $f_{dur}$ ). The estimated *duration* of the ride from the user location to the proposed destination.

**Ride length** ( $f_{len}$ ). The estimated *length* of the candidate ride.

These features are first rescaled in the range [0-1] on a ride-request basis by considering all the values occurring in the set of matchings  $\mathcal{M}$ .  $\mathcal{M}$  is then sorted by decreasing value of function  $Rank(m)$ :

$$Rank(m) = 1 - \sum_i w_i f_i$$

where  $w_i \in [0,1]$ ,  $\sum_i w_i = 1$  is the weight associated with feature  $f_i \in \{f_{dly}, f_{wld}, f_{dur}, f_{len}\}$ . By properly setting the weights  $w_i$  in the above linear combination of features, we can tune the importance of the different features. For example, we can reward rides with the shortest distance to reduce the emission of pollutants, or favor the passengers by ranking higher rides with the lowest walks and/or duration. Of course we are aware that more sophisticated ranking models considering different and complex aspects could be adopted (see for example [31]). However, the goal of this work is investigating the impact on carpooling of user flexibility and not studying application-specific rankers. We will see in the following how this simple ranking model allows us to sweep weights in order to understand their effect on the user and the environment.

## V. EXPERIMENTAL EVALUATION

In this section we present the experiments conducted to assess ABRM in terms of kilometers, liters of gasoline and  $CO_2$  potentially saved with respect to a traditional, destination-oriented, carpooling approach. Note that in the following we do not deal with the allocation problem neither consider the number of seats available in the cars offering the rides.

TABLE I: Datasets Statistics

Dataset	Checkins	Users	Venues	Categories
New York (NYC)	227,428	1,083	38,333	251
Tokyo (TKY)	573,703	2,293	61,858	247

Ride request-offer allocation is a well-known optimization problem, orthogonal to this proposal. Any scheduling solution addressing this problem for destination-oriented carpooling fits also our activity-based approach. Since the same assumptions hold for the destination-oriented solution used as baseline, we believe that the choice of not considering allocation does not constitute a limitation of the work. Thus, below we investigate the **potential impact** of ABRM and of the settings of the ranking function in the reduction of the number of circulating vehicles, of pollutant emissions and consequent improvement of quality of the urban environment. The experiments conducted aim to answer comprehensively the following research questions:

**RQ1:** *To what extent can ABRM increase ride sharing opportunities?*

**RQ2:** *How well the matchings discovered by ABRM fit the constraints in the ride request?*

**RQ3:** *What is the impact of tuning the weights used by the ranking function on the number of requests potentially supplied?*

**RQ4:** *Which are the most favorable activities for exploiting the alternative destinations approach?*

#### A. Experimental setup

Our experiments are conducted on two semi-synthetic datasets of ride requests and ride offers, obtained by processing and enriching two publicly-available *Foursquare* datasets [37]. These datasets record the check-ins of *Foursquare* users in New York City (NYC) and Tokyo (TKY) for about 10 months (from 12 April 2012 to 16 February 2013). Each check-in is associated with a time stamp, the GPS coordinates of the POI and a fine-grained venue-category. Table I summarizes the total number of check-ins, the number of distinct users and distinct venues, and the number of possible categories for venues. For both datasets we consider only the users with at least 100 check-ins.

We enriched the above datasets by gathering the POI information using the *Venue*<sup>3</sup> and *Similar*<sup>4</sup> *Foursquare* APIs. Specifically, we gathered for each POI in our datasets the number of check-ins performed at the POI (popularity), the number of “likes” received (favorite), and the *top-5* most similar venues according to an unknown *Foursquare* similarity measure.

We emphasize that the datasets used for the experiments provide only a simulation of a traffic scenario and are not representative of a general mobility graph. In these datasets the urban traffic flow is surely under-represented as most actual trips are likely not to be between two *FourSquare* destinations checked-in by the drivers. Nevertheless, this dataset has the

<sup>3</sup><https://developer.foursquare.com/docs/venues/venues>

<sup>4</sup><https://developer.foursquare.com/docs/venues/similar>

important advantage that the activity performed by the users is explicitly reported as check-ins and this is crucial for our activity-based ride sharing approach. Rather than representing a general urban traffic flow, we simulate using these datasets the activity-based ride requests and offers. In the previous preliminary work [17], we experimented the use of actual GPS traces of cars and we faced the non trivial problem of associating the raw GPS points to the performed activity at stops. Also, we had strong privacy problems and we could not make the dataset public. Therefore we privileged here the use of datasets that, although semi-synthetic and with clear limitations, are public and representative of a large-scale activity-based scenario.

We exploit the above datasets of Foursquare check-ins to build two semi-synthetic datasets representing disjoint sets of *Ride Requests* and *Ride Offers*. By matching the ride requests with the ride offers by means of the ABRM algorithm and the baseline, we assess our proposal by considering the set of requests potentially satisfied by ABRM but not by the baseline algorithm.

We identify the ride offers as the trips of each user between the two most frequently visited venues  $v_a$  and  $v_b$ . The intuition behind this choice is that the rides between the most frequently visited locations constitute a reasonable surrogate of routine trips a user could offer as driver. This is also supported by a manual inspection of the data that shows that very often people have their most frequent check-ins at venues such as home, work place, university, school [12].

The exact procedure followed for each user  $u$  to populate the set  $RT$  of ride offers is detailed below:

- 1) Let  $v_a$  and  $v_b$  be the two POIs most frequently visited by user  $u$ , and  $rt_{ab}$  and  $rt_{ba}$  the candidate routine trips from  $v_a$  to  $v_b$  and from  $v_b$  to  $v_a$ , respectively.
- 2)  $rt_{ab}$  and  $rt_{ba}$  are added to set  $RT$  for all the days of the week in which there is at least a check-in of  $u$  in both the places. The arrival time in  $v_a$  and  $v_b$  for the above two ride offers are computed as the median among the timestamps associated with the check-ins in  $v_a$  and  $v_b$ .
- 3) For each ride offer  $rt$  obtained with steps 1-2 we compute a representative trajectory of the fastest car route from the departure to the arrival locations by using *Google Maps*<sup>5</sup>. In addition, the arrival time  $t_{arrival}$  of  $rt$  is used to estimate the duration of the ride, its length and the time of departure  $t_{departure}$ .

By following the above procedure we obtained 11,426 and 25,306 ride offers for NYC and TKY respectively. Figure 3 shows a geospatial visualization of a sample of ride offers from the Tokyo dataset. Each color represents a single trajectory. As we can see, the ride offers cover main streets and avenues of the city, particularly in the city center characterized by a higher density.

For the extraction of the set  $Q$  of ride requests, we focus instead on the venues least frequently visited by each user. The insight is that occasionally visited venues are the ones for which a user is most likely to be open to accept a ride for an alternative destination. For example, venues like bars,

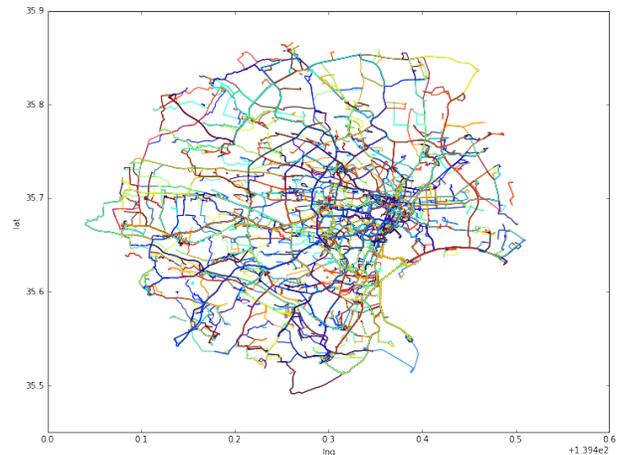


Fig. 3: Samples of ride offers extracted from the Tokyo dataset.

restaurants, pubs, markets, cinemas are in general places not routinely visited for which we expect that users can be more flexible [12]. Based on this idea, we extract the ride requests for user  $u$  as follows. We first remove from the check-ins of  $u$  the check-ins in the two most frequently visited venues (see the previous procedure) and all the check-ins in venues belonging to the following categories: 'Home (private)', 'Office', 'Airport', 'Subway', 'Neighborhood', 'Road', 'Building', 'Residential Building (Apartment / Condo)', 'Government Building', 'Train Station', 'Road', 'Bus Station', 'Hotel', 'City' and 'Bridge'. We assume in fact that the activities associated to these categories of POIs can difficultly be performed in alternative places. For example, a passenger could not be dropped to a different airport from the one she has the flight, or to a different hotel. Then, for each remaining check-in  $c$  we create a ride request  $q$  considering as starting location the place most frequently visited by the user. The destination is obviously the same of the check-in, while the departure time is inferred from the check-in time and the travel time needed to reach the destination from the starting location (as estimated by Google map). Resulting ride request corresponding to rides shorter than 1 km are discarded.

In addition, unless differently specified, we set in all the ride requests 60 minutes and 500 meters as the maximum delay and the maximum walk distance, respectively. We also consider 5.0 km/h as average walking speed [38]. The resulting datasets contain 98,008 and 160,271 ride requests in NYC and TKY, respectively.

## B. Evaluation metrics

The acceptance of a ride offer is an absolutely subjective decision whose modeling is out of the scope of the present work. Thus, given the set of ride requests  $Q$  and the set of ride offers  $RT$  in the NYC and TKY datasets, we assess our proposal by simply considering the potential matchings returned by ABRM and the destination-oriented baseline, if any. In order to minimize the effect of differences in the implementations, the destination-oriented solution used as baseline is a simple modification of our implementation where

<sup>5</sup>developers.google.com/maps/documentation/directions/

the requested destination in the ride request is considered fixed in the matchings. All the other parameters (including the maximum walking distance and time delay) are set exactly to the same values in order to directly measure the boost in the number of *requests potentially supplied* by our activity-based approach versus the destination-based counterpart.

Let us indicate with  $Q_s$  the subset of all ride requests  $Q$  satisfied by at least one ride offer.  $|Q_s|$  is the number of ride requests potentially satisfied and  $|Q_s|/|Q|$  the ratio (measured in percentage) between the number of requests supplied and the total number of requests in  $Q$ .

In order to estimate the potential impact of carpooling solutions on the reduction of kilometers traveled by cars, we assume ride requests are satisfied by the ride offer most highly ranked retrieved by ABRM and by the baseline, if any. We clearly assume a user can avoid to take its own vehicle when a ride possibility is offered. Every satisfied request thus corresponds to a vehicle less in circulation. The number of kilometers potentially saved is thus:

$$Km_{saved}(Q_s) = \sum_{q \in Q_s} D(q.pickupLoc, q.dropLoc)$$

where  $D(q.pickupLoc, q.dropLoc)$  is the length of the route connecting the pickup and the (alternative) drop off locations estimated by the Google maps service.

### C. RQ1: boosting ride sharing opportunities

In this section we address the first research question by comparing the number of ride sharing requests potentially satisfied by ABRM and the baseline. We varied the method used by ABRM to choose the set of alternative venues. Specifically we experimented the following variations: (a) *all alternative destinations*, the passenger can accept to go to any alternative destination where she could perform the desired activity (all the venues in the same category of the requested POI); (b / c) *liked / popular*, the passenger can accept to go to one of the  $k$  most *liked / visited* POIs in the same category; (d) *preferred*, the passenger can accept to go to one among her  $k$  most preferred destinations. Since we do not have user-level preference information in our datasets, we simulate this case by randomly selecting  $k$  POIs in the same category for each user; (e) *Foursquare similarity*, the passenger can accept to go to one of the venues most *similar* to the requested destination according to the FourSquare black-box similarity function. We remark that for the liked, popular and preferred criteria we experimented values of  $k$  equal to 5, 10 or 20. For the *Foursquare similarity* case, given the implementation of the APIs, at most 5 alternative POIs are returned even if in many cases a lower number of similar venues is suggested.

Table II reports the number ( $|Q_s|$ ) and percentage ( $|Q_s|/|Q|$ ) of ride requests potentially supplied, the improvement of ABRM compared to the destination-oriented baseline (gain), and the impact for the environment in terms of potentially saved kilometers ( $Km_{saved}$ ).

As expected, we observe that the *all alternative destinations* method, due to the higher number of possible destinations,

reaches the best results with a gain of 44.46% and 54.69%, compared to the baseline, for NYC and TKY, respectively. In general, we measured a higher performance in the Tokyo dataset compared to New York, probably due to the higher number of ride offers and alternative destinations. Another general trend we can note is that the number of requests potentially satisfied increases when more flexibility is assumed, i.e., when the number  $k$  of possible alternative destinations increases. On the other side, we observed similarly that ABRM remarkably increases also the number of ride offers matched to some ride requests. For example, in the NYC case, considering the most flexible criteria (i.e. all alternative destinations), 63% of the ride offers are matched to at least one ride request against 30% measured with the baseline.

TABLE II: Ride requests matched with the baseline and ABRM on the NYC and TKY datasets.

	NYC	$ Q_s $	$ Q_s / Q $	gain	$Km_{saved}$
baseline	9,288	9.48%	-	-	35,901
all alt. destinations	52,864	53.94%	44.46%	44.46%	388,293
likes (k=5)	18,378	18.75%	9.27%	9.27%	117,027
likes (k=10)	24,534	25.03%	15.56%	15.56%	157,840
likes (k=20)	32,483	33.14%	23.67%	23.67%	213,254
popular (k=5)	20,268	20.68%	11.20%	11.20%	129,702
popular (k=10)	25,011	25.52%	16.04%	16.04%	160,634
popular (k=20)	32,614	33.28%	23.80%	23.80%	213,908
preferred (k=5)	20,157	20.57%	11.09%	11.09%	130,523
preferred (k=10)	28,627	29.21%	19.73%	19.73%	190,249
preferred (k=20)	36,628	37.37%	27.90%	27.90%	250,761
Foursquare similarity	12,642	12.90%	3.42%	3.42%	55,783
	TKY	$ Q_s $	$ Q_s / Q $	gain	$Km_{saved}$
baseline	36,695	22.90%	-	-	213,838
all alt. destinations	124,356	77.59%	54.69%	54.69%	1,014,079
likes (k=5)	74,674	46.59%	23.70%	23.70%	566,748
likes (k=10)	89,326	55.73%	32.84%	32.84%	690,604
likes (k=20)	98,782	61.63%	38.74%	38.74%	773,103
popular (k=5)	74,329	46.38%	23.48%	23.48%	565,492
popular (k=10)	87,908	54.85%	31.95%	31.95%	678,237
popular (k=20)	98,839	61.67%	38.77%	38.77%	775,156
preferred (k=5)	67,044	41.83%	18.94%	18.94%	513,269
preferred (k=10)	85,236	53.18%	30.29%	30.29%	662,167
preferred (k=20)	97,223	60.66%	37.77%	37.77%	764,810
Foursquare similarity	48,844	30.48%	7.58%	7.58%	317,694

Table II also shows the estimated amount of kilometers traveled potentially saved by the corresponding carpooling solution. Once more we recall that the figures reported here are upper-bound estimates computed on the basis of the assumptions made on the flexibility and willingness of users to accept shared rides possibly at alternative destinations. From these values we can easily approximate the liter of fuel (e.g. gasoline) saved and consequently the saved amount of  $CO_2$ . For the sake of simplicity, in this computation we assume each car consumes in average 1 liter of gasoline for 7.449 Km as reported in the official statistics of the Bureau of Transportation<sup>6</sup>. By assuming the passengers are flexible enough to change their requested location, the estimate for the number of kilometers saved in our test cities amount to 388,293 in NYC and 1,014,079 in Tokyo. These values in turn correspond to 52,120 and 136,117 less liters of gasoline and to 99.4 tons and 259.60 tons less  $CO_2$  emission for New York and Tokyo, respectively. The baseline based on fixed

<sup>6</sup>goo.gl/CIDSfL

destinations may save in NYC 35,786 Km only (corresponding to 4,803 gasoline liters and 1.22  $CO_2$  tons). For TKY we estimate instead that the baseline can save 213,838 Km (28,703 gasoline liters and 7.34  $CO_2$  tons).

On both the datasets, the method based on all the alternative destinations significantly outperforms the other methods. On the TKY dataset the average number of matches for each ride request is 690 for ABRM and only 3.02 for the baseline. Slightly similar figures are measured on the NYC datasets. The ride request in TKY with the highest number of alternative destinations counts about 23k matches against the 45 achieved by the baseline. Such high numbers motivates the need of a ranking model later discussed in the sections V-D and V-E.

In conclusion, in relation to *RQ1*, we observe that the results reported prove the potential boost of ride sharing services involving the offers of rides to alternative destinations. The next step is to investigate how much these possibilities fit the passenger requests and the public good. These aspects are discussed in the next sections.

#### D. *RQ2: How well the user requests meet the ride offers to alternative destinations?*

In this section, we address our second research question, related to the contribution of the ride features to the ranking method defined in Section IV-A. With these features we intend to model how much the rides to alternative destinations can meet the requirements specified in the user request. Figure 4 supports this study. Each plot shows eight curves reporting the cumulative distribution of  $|Q_s|/|Q|$  for the *baseline*, the *all alternative destination* criterion and the *popular* ( $k=5,10,20$ ) and *preferred* ( $k=5,10,20$ ) ones. The four plots for NYC (TKY) report each the effect on  $|Q_s|/|Q|$  of varying the value of one of the features  $f_{dly}$ ,  $f_{wld}$ ,  $f_{dur}$ , and  $f_{len}$  by keeping all the other fixed.

We computed the cumulative distributions by considering the fraction of supplied ride requests having a value for the feature considered lower than the one reported in the x axis. In this way we can see how the fraction of supplied requests changes when the feature value increases. Looking at the plots, specifically for the TKY dataset and the temporal shift feature  $f_{dly}$ , we see that even when we consider matches with only 10 minutes shift, more than half of the requests could be potentially satisfied with a ride offer to an alternative POI with the *all alternative destinations* method. A similar consideration can be done for the plot analyzing the variation on the walk distance feature  $f_{wld}$ .

For the trip length feature  $f_{len}$ , we report on the x axis the ratio between the matched ride offer having the shortest length and the shortest car distance to the requested destination computed by the Google Maps service. By looking to the plot corresponding to the NYC dataset we observe that more than 20% of the ride requests can be supplied to an alternative destination with a trip which is half in duration or shorter than the trip originally requested. We observe even better results on the TKY dataset. Similarly, for the duration feature  $f_{dur}$  we report on the x axis the ratio between the estimated duration to reach the alternative POI and the travel time to arrive at the

requested destination computed by the Google Maps service. On both the NYC and TKY datasets we can see that ABRM proposes to passengers rides to alternative destinations that are in most of the cases shorter in duration than the ones to the requested destination offered by the destination-oriented baseline. We can conclude this section by considering that independently from the feature considered, and thus from the subjective importance given by the user to the carpooling aspect modeled by the specific feature (time shift, walking distance, ride duration and ride length) ABRM is likely to provide a better carpooling service to flexible users accepting to reach alternative destinations for performing the intended activities.

#### E. *RQ3: Effectiveness of the Ranking Method*

ABRM returns the set of candidate Ride offers matching a given Ride Request. The ranking model defined in IV-A order the candidate offers on the basis of a linear combination of weighted. The purpose of the ranking step is helping the users to choose the most relevant rides among a possibly large set of offers. Since no golden standard recording user preferences is available to optimize the weights, we consider here a uniform weighting schema giving the same importance to all the features and we compare such setting with four extreme scenarios where we prefer one features over all the others by setting, in turn, the corresponding weight to 1 and the others to zero. Given the order of features previously used, we identify the uniform weighting schema with the vector  $w = [1/4, 1/4, 1/4, 1/4]$  while the scenario giving, for example, only importance to ride duration (feature  $f_{dur}$ ) corresponds to vector  $w = [0, 0, 1, 0]$ . We chose the *all alternative destinations* as alternative destination criteria since it is the one providing the highest numbers of matchings, not biased by other parameters (like number of likes, popularity, etc.). As in the previous section,  $f_{len}$  and  $f_{dur}$  are normalized based on the ratio between the estimated length and temporal duration to reach the alternative POI and the requested destination. For both datasets the top-1 offer returned by the ranking model was compared with the best values of the extreme approach giving importance to only that feature. Figure 5 shows the cumulative distribution of value  $|Q_s|/|Q|$  considering the top-ranked result for each of these five configurations of vector  $w$ .

As expected, the highest values for  $|Q_s|/|Q|$  are achieved in all the plots for the extreme weighting schema considering the associated feature only. However, the curves corresponding to the uniform weighting schema are in all the plots but the two in the most right hand side the closest to the highest curves, thus showing in general a very good performance. Uniform weighting performs as third solution under the considered metric only when the ride length aspect is analyzed, and the curves result to be very close to the second one. We observe in fact that the third and fourth features, ride duration and length, are highly correlated. On the other hand, the other two features, time delay and walking distance, are more selective on the generation of candidates. This suggest us that delay and walking distance are strong constraints to the matching

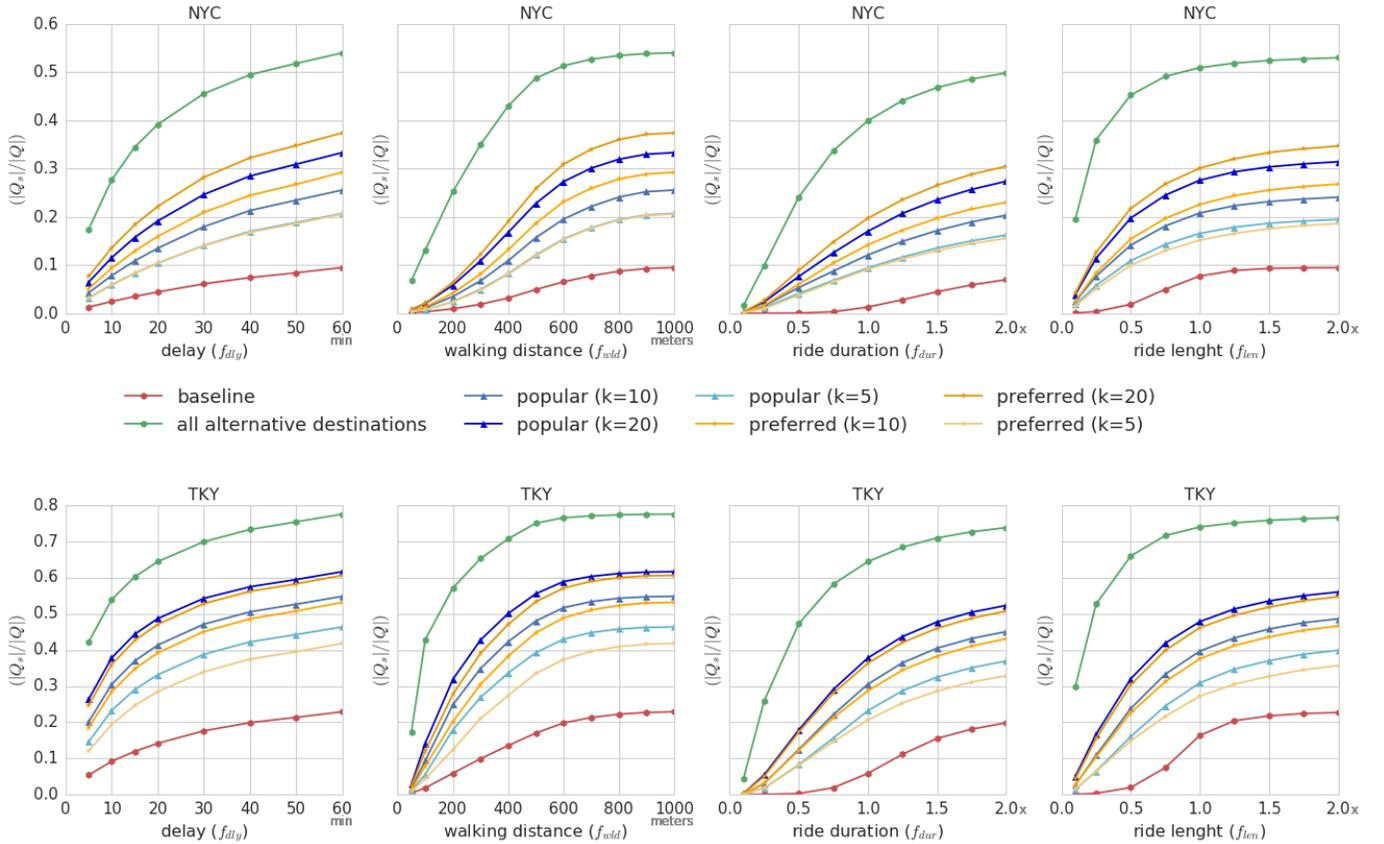


Fig. 4: Cumulative distribution of  $|Q_s|/|Q|$  by varying each feature in isolation.

of candidate offers, whereas the other two features may be weighted differently to try to optimize user acceptance on one hand and to improve public goodness by minimizing the total distance traveled by cars.

#### F. RQ4: Activities mostly favored by ABRM

We conclude the experiments by addressing RQ4, namely discussing which activities most favor the ride sharing in the alternative destination scenario. In other words, we analyze our ride matching results in the two datasets for understanding which activities have the largest boost in terms of number of ride matches.

We identified the most frequent ride requests by intended activity and, from them, we analyze the possible matches. Table III reports, for both datasets the top-10 activities that result in the largest number of ABRM ride matchings. We report the number of requests  $|Q|$ , the percentage  $(|Q_s|/|Q|)$  of ride requests potentially supplied with the baseline and the improvement in percentage obtained with ABRM. We use the *popular* with  $k = 5$  (more restrictive) and *all alternative destinations* (less restrictive) as alternative destination criteria. We observe for NYC a boost on ride-sharing possibilities mainly for activities related to “Italian Restaurant” and “Bar” with an improvement of +58.02% and +57.86% respectively. In turn, for TKY, activity “Food & Drink Shop” achieves the highest boost with +64.28%. This insight confirm our intuition

that entertainment or eating are in general the activities that can most benefit from the proposed approach. The specific results reported in the table are however also correlated to a combination of factors such as the number of venues for each categories in the cities and their location.

TABLE III: Activities favoring ABRM boosting.

	$ Q_s $	Bas. $ Q_s / Q $	ABRM <i>pop.</i> , $k = 5$	ABRM <i>all alt. dest.</i>
<b>NYC</b>				
Italian Restaurant	1,315	13.08%	+12.97%	+58.02%
Bar	11,242	10.60%	+13.64%	+57.86%
Music Venue	1,027	9.45%	+25.61%	+55.50%
Coffee Place	918	8.06%	+20.59%	+54.68%
Chinese Restaurant	932	10.41%	+10.73%	+53.33%
Deli / Bodega	1,621	5.31%	+15.48%	+52.81%
Bakery	899	10.23%	+20.91%	+51.72%
Park	3,015	11.21%	+16.80%	+50.95%
American Restaurant	2,681	14.58%	+21.04%	+50.09%
Mexican Restaurant	1,468	13.42%	+19.69%	+49.11%
<b>TKY</b>				
Food & Drink Shop	6,766	15.05%	+29.83%	+64.28%
Convenience Store	7,360	16.28%	+23.63%	+63.93%
Park	4,026	14.83%	+27.81%	+63.88%
Fast Food Restaurant	3,698	18.33%	+27.96%	+60.28%
Chinese Restaurant	2,804	20.97%	+29.81%	+60.06%
Japanese Restaurant	9,365	25.46%	+15.07%	+59.62%
Ramen / Noodle House	10,618	22.29%	+32.59%	+59.12%
Mall	6,185	20.03%	+26.89%	+57.83%
Coffee Shop	4,756	23.91%	+27.02%	+56.79%
Bar	8,051	26.89%	+33.21%	+55.21%

## VI. CONCLUSIONS

In this paper we proposed and discussed the ABRM algorithm aimed at matching users’ carpooling requests with ride offers reaching alternative destinations where the intended user activity can be performed. Experiments conducted on two large

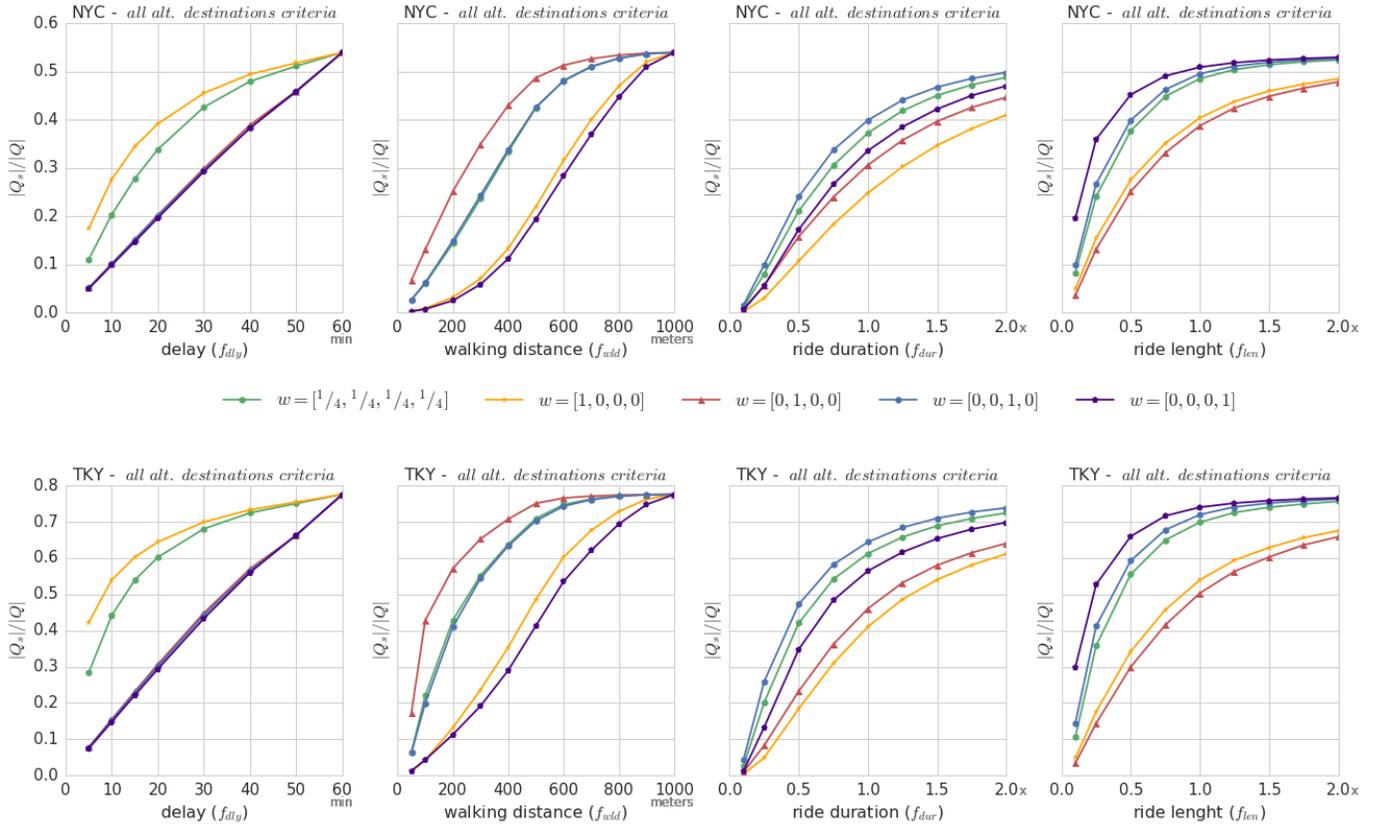


Fig. 5: Comparison by ride feature between the Top-1 ranked ride and the best result by feature

semi-synthetic datasets recording mobility demands and the categories of POIs visited (extracted from Foursquare checkins and made publicly available to favor the reproducibility of our results) showed that ABRM can boost to up 54.69% the percentage of carpooling request satisfied with compatible ride offers with respect to traditional destination-oriented carpooling. Since the number of carpooling opportunities provided by ABRM can be very large we proposed and analyzed in detail how the candidate ride offers can be ranked in order to better meet user expectations or to enforce their pro-environment behaviors in order to maximize the beneficial impact of carpooling on the environment in terms of  $CO_2$  emissions saved. Our analysis is based on the assumption that users might be flexible in their mobility habits and provided interesting insights on the extent to which a ride-sharing service could take advantage of this spontaneous attitude. We notice however that an ad-hoc study with real users would be necessary to have a reliable measure of the actual acceptance of ride offers towards alternative destinations. We leave such assessment as future work together with the study of the ride allocation problem in the ABRM context.

## REFERENCES

- [1] “The 10 most polluted cities in the world,” <http://www.ibtimes.co.uk/world-environment-day-10-most-polluted-cities-world-1504260>, last access April 18, 2017.
- [2] “TomTom traffic index,” [https://www.tomtom.com/en\\_gb/trafficindex/](https://www.tomtom.com/en_gb/trafficindex/), last access April 18, 2017.
- [3] B. Cici, A. Markopoulou, E. Frías-Martínez, and N. Laoutaris, “Quantifying the potential of ride-sharing using call description records,” in *Proceedings of the 14th Workshop on Mobile Computing Systems and Applications*, ser. HotMobile ’13. New York, NY, USA: ACM, 2013.
- [4] N. Perterer, P. Sundström, A. Meschtscherjakov, D. Wilfinger, and M. Tscheligi, “Come drive with me: An ethnographic study of driver-passenger pairs to inform future in-car assistance,” in *Proc. of the 2013 Conference on Computer Supported Cooperative Work*. New York, NY, USA: ACM, pp. 1539–1548.
- [5] R. Gilbert and A. Perl, *Transport Revolutions Moving People and Freight without Oil*. Earthscan, London, UK, 2008.
- [6] T. Vanoutrive, E. Van De Vijver, L. Van Malderen, B. Jourquin, I. Thomas, A. Verhetsel, and F. Witlox, “What determines carpooling to workplaces in Belgium: Location, organisation, or promotion?” *Journal of Transport Geography*, vol. 22, pp. 77–86, 2012.
- [7] A. Dixit, S. Bora, S. Chemate, and N. Kolpekwar, “Real-Time Carpooling System for Android Platform,” *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 2, no. 6, pp. 436–437, 2012.
- [8] M. Stiglic, N. Agatz, M. Savelsbergh, and M. Gradisar, “Making dynamic ride-sharing work: The impact of driver and rider flexibility,” *Transportation Research Part E*, vol. 91, p. 190207, 2016.
- [9] N. Agatz, A. Erera, M. Savelsbergh, and X. Wang, “Optimization for dynamic ride-sharing: A review,” *European Journal of Operational Research*, vol. 223, p. 295303, 2012.
- [10] P. Rode, C. Hoffmann, J. Kandt, A. Graff, and D. Smith, “Towards new urban mobility: the case of London and Berlin,” 2015, the London School of Economics and Political Science, LSE Cities, London, UK.
- [11] L. Steg and C. Vlek, “Encouraging pro-environmental behaviour: An integrative review and research agenda,” *Journal of Environmental Psychology*, vol. 29, p. 309317, 2009.
- [12] V. M. de Lira, S. Rinzivillo, C. Renso, V. C. Times, and P. C. A. R. Tedesco, “Investigating semantic regularity of human mobility lifestyle,” in *Proc. of IDEAS 2014*.
- [13] F. Wu and Z. Li, “Where did you go: Personalized annotation of mobility records,” in *Proc. of ACM CIKM 2016*. New York, NY, USA: ACM, pp. 589–598.

- [14] “BlaBLaCar, trusted ridesharing,” <https://www.blablacar.com/>, last access 18 October 2017.
- [15] “Lyft Line, turn miles into money,” <https://www.lyft.com/line>, last access 18 October 2017.
- [16] “Uber Pool,” <https://www.uber.com/uberpool/>, last access 18 October 2017.
- [17] V. Monteiro De Lira, V. C. Times, C. Renso, and S. Rinzivillo, “ComeWithMe: An activity-oriented carpooling approach,” in *Proc. of IEEE ITSC 2105*.
- [18] V. Monteiro De Lira, R. Perego, C. Renso, and S. Rinzivillo, “The ComeWithMe system for searching and ranking activity-based carpooling rides,” in *proc. of ACM SIGIR 2016*.
- [19] M. Stiglic, N. Agatz, M. Savelsbergh, and M. Gradisar, “Making dynamic ride-sharing work: The impact of driver and rider flexibility,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 91, no. Supplement C, pp. 190 – 207, 2016.
- [20] M. Furuhashi, M. Dessouky, F. Ordez, M.-E. Brunet, X. Wang, and S. Koenig, “Ridesharing: The state-of-the-art and future directions,” *Transportation Research Part B: Methodological*, vol. 57, pp. 28 – 46, 2013.
- [21] S. Ma and O. Wolfson, “Analysis and evaluation of the slugging form of ridesharing,” in *Proc. of ACM SIGSPATIAL 2013*, pp. 64–73.
- [22] R. Geisberger, D. Luxen, S. Neubauer, P. Sanders, and L. Völker, “Fast detour computation for ride sharing,” *CoRR*, vol. abs/0907.5269, 2009.
- [23] “Optimization for dynamic ride-sharing: A review,” *European Journal of Operational Research*, vol. 223, no. 2, pp. 295 – 303, 2012.
- [24] J. Lin, S. Sasidharan, S. Ma, and O. Wolfson, “A model of multimodal ridesharing and its analysis,” in *Proc. of IEEE MDM 2016*, pp. 164–173.
- [25] M. Stiglic, N. Agatz, M. Savelsbergh, and M. Gradisar, “The benefits of meeting points in ride-sharing systems,” *Transportation Research Part B: Methodological*, vol. 82, no. Supplement C, pp. 36 – 53, 2015.
- [26] P. Minett and J. Pearce, “Estimating the Energy Consumption Impact of Casual Carpooling,” pp. 126–139, 2011.
- [27] K. L. Kelley, “Casual Carpooling Enhanced,” *Journal of Public Transportation*, vol. 10, no. 4, pp. 119–130, 2007.
- [28] W. Herbawi and M. Weber, “Modeling the multihop ridematching problem with time windows and solving it using genetic algorithms,” in *Proc. of IEEE International Conference on Tools with Artificial Intelligence*, 2012, pp. 89–96.
- [29] S. Cho, J.-Y. Kang, A.-U.-H. Yasar, L. Knapen, T. Bellemans, D. Janssens, G. Wets, and C.-S. Hwang, “An Activity-based Carpooling Microsimulation Using Ontology,” *Procedia Computer Science*, no. Ant, pp. 48–55, Jan.
- [30] R. Guidotti, A. Sassi, M. Berlingerio, A. Pascale, and B. Ghaddar, “Social or green? a data-driven approach for more enjoyable carpooling,” in *Proc. of IEEE ITSC 2015*.
- [31] M. G. Campana, F. Delmastro, and R. Bruno, “A machine-learned ranking algorithm for dynamic and personalised car pooling services,” in *Proc. of IEEE ITSC 2016*.
- [32] M. Sghaier, H. Zgaya, S. Hammadi, and C. Tahon, “A novel approach based on a distributed dynamic graph modeling set up over a subdivision process to deal with distributed optimized real time carpooling requests,” in *Proc. of IEEE ITSC 2011*, pp. 1311–1316.
- [33] M. Sghaier, H. Zgaya, S. Hammadi, C. Tahon, “A distributed Dijkstra’s algorithm for the implementation of a real time carpooling service with an optimized aspect on siblings,” in *Proc. of IEEE ITSC 2010*, pp. 795–800.
- [34] V. Dimitriesski, “Real-Time Carpooling and Ride-Sharing: Position Paper on Design Concepts, Distribution and Cloud Computing Strategies,” in *Proc. of FedCSIS, 2013*, pp. 781–786.
- [35] N. V. Pukhovskiy and R. E. Lepshokov, “Real-time carpooling system,” in *Proc. of IEEE ICCTS*, pp. 648–649.
- [36] C. Parent, S. Spaccapietra, C. Renso, G. Andrienko, N. Andrienko, V. Bogorny, M. L. Damiani, A. Gkoulalas-Divanis, J. A. Macedo, N. Pelekis, Y. Theodoridis, and Z. Yan, “Semantic trajectories modeling and analysis,” *ACM Computing Surveys*, vol. 45, no. 4, 2013.
- [37] D. Yang, D. Zhang, V. W. Zheng, and Z. Yu, “Modeling user activity preference by leveraging user spatial temporal characteristics in lbsns,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 1, pp. 129–142, Jan 2015.
- [38] B. J. Mohler, W. B. Thompson, S. H. Creem-Regehr, H. L. Pick, and W. H. Warren, “Visual flow influences gait transition speed and preferred walking speed,” *Experimental Brain Research*, vol. 181, no. 2, pp. 221–228, 2007.



**Vinicius Monteiro de Lira** holds a MSc degree in computer science from the Federal University of Pernambuco, Brazil, in 2014. He currently is PhD Student at HPC Lab of ISTI-CNR with a joint supervision from the University of Pisa, Italy and the Federal University of Pernambuco, Brazil. His main research interests include mobility data mining, machine learning and big data analytics.



**Raffaele Perego** (<http://hpc.isti.cnr.it/~raffaele>) is a senior researcher at ISTI-CNR, where he leads the HPC Lab (<http://hpc.isti.cnr.it/>). His main research interests include high performance computing, Web information retrieval, and data mining. He coauthored more than 150 papers on these topics published in journals and proceedings of international conferences.



**Chiara Renso** (<http://hpc.isti.cnr.it/~renso>) holds a PhD and M.Sc. degree in Computer Science from the University of Pisa (1992, 1998). She is researcher at HPC Lab of ISTI-CNR (Italy) working mainly on trajectory data mining and semantic trajectories. She authored more than 100 peer-reviewed publications. She is co-editor of the book “Mobility Data: Modelling, Management, and Understanding” Cambridge Press 2013.



**Salvatore Rinzivillo** is a researcher at the KDD Lab at ISTI-CNR (Italy). His main research interests include Data Mining and Visual Analytics for spatial, spatiotemporal, and mobility data. He has published several papers in international peer-reviewed conferences and journals.



**Valeria Cesario Times** received a MSc degree in computer science from the Federal University of Pernambuco, Brazil, in 1994, and a PhD in computer science in 1999 from University of Leeds, United Kingdom. She is currently an associate professor in the Informatics Center at the Federal University of Pernambuco, Brazil. Her research interests include GIS, data privacy, trajectory analysis, advanced database applications.