Transportation Research Part C 80 (2017) 20–36

Contents lists available at ScienceDirect

Transportation Research Part C

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

The GRAAL of carpooling: GReen And sociAL optimization from crowd-sourced data ☆

Michele Berlingerio [a,*], Bissan Ghaddar [a,b], Riccardo Guidotti [a,c], Alessandra Pascale [a], Andrea Sassi [a,d]

[a] IBM Research Ireland, Ireland
[b] University of Waterloo, Canada
[c] University of Pisa, Italy
[d] University of Modena and Reggio Emilia, Italy

A R T I C L E I N F O

Article history:
Received 14 September 2015
Received in revised form 14 February 2017
Accepted 28 February 2017

Keywords:
Carpooling
Social networks
Multi-objective optimization
User study

A B S T R A C T

Carpooling, i.e. the sharing of vehicles to reach common destinations, is often performed to reduce costs and pollution. Recent work on carpooling takes into account, besides mobility matches, also social aspects and, more generally, non-monetary incentives. In line with this, we present GRAAL, a data-driven methodology for GReen And sociAL carpooling. GRAAL optimizes a carpooling system not only by minimizing the number of cars needed at the city level, but also by maximizing the *enjoyability* of people sharing a trip. We introduce a measure of enjoyability based on people's interests, social links, and tendency to connect to people with similar or dissimilar interests. GRAAL computes the enjoyability within a set of users from crowd-sourced data, and then uses it on real world datasets to optimize a weighted linear combination of number of cars and enjoyability. To tune this weight, and to investigate the users' interest on the social aspects of carpooling, we conducted an online survey on potential carpooling users. We present the results of applying GRAAL on real world crowd-sourced data from the cities of Rome and San Francisco. Computational results are presented from both the city and the user perspective. Using the crowd-sourced weight, GRAAL is able to significantly reduce the number of cars needed, while keeping a high level of enjoyability on the tested data-set. From the user perspective, we show how the entire per-car distribution of enjoyability is increased with respect to the baselines.

1. Introduction

Carpooling is a scheme in which people share a vehicle in order to reach common or nearby destinations. Despite its clear advantages in reducing costs, pollution, and time spent in finding a car park, there are still a few obstacles that prevents it from being the preferred way to move: safety of passengers, sub-optimal mobility matches, and time flexibility, among others.

☆ This article belongs to the Virtual Special Issue on "Smart City".

* Corresponding author.
E-mail addresses: michele.berlingerio@gmail.com (M. Berlingerio), bghaddar@uwaterloo.ca (B. Ghaddar), riccardo.guidotti@di.unipi.it (R. Guidotti), apascale@ie.ibm.com (A. Pascale), andrea.sassi@unimore.it (A. Sassi).

http://dx.doi.org/10.1016/j.trc.2017.02.025
0968-090X/

A common underlying aspect across many such obstacles is a hidden psychological barrier that makes carpooling less attractive. However, due to the increasing popularity of online social networks in the last few years, there are some social aspects that people intentionally decide to share with the outside world, including strangers. In fact, sharing interests, pictures, and visited locations, are the basis of the success of services such as Facebook, Twitter, and Foursquare. The availability of such information allows external services and people to use this data for third party applications. As a result, such social aspects can be now *measured*, and *exploited* to overcome this invisible psychological barrier in the context of carpooling.

Inspired by the literature on carpooling (Teal, 1987; Yang and Huang, 1999; Correia and Viegas, 2010; de Almeida Correia et al., 2013; Maniezzo et al., 2004), and by the recent work on data-driven analysis in urban networks (Pascale et al., 2015) and data-driven optimization of urban transit networks (Naoum-Sawaya et al., 2015; Berlingerio et al., 2013), we present a mathematical formulation of the carpooling problem taking into account the above factors, and a data-driven methodology to automatically derive mobility and social matches to be used as recommendations for the carpooling system. The main goal of our work is to present a "what-if analysis" in which we measure, from sources available online, how users would *enjoy* sharing a trip with other people, and to devise a new methodology for carpooling driven by these measurements. Our contribution is mainly methodological, rather than a carpooling system tested on the field. Thus, we focus in this paper on the theoretical core of carpooling, i.e. the data-driven multi-objective optimization problem.

In contrast with on-demand carpooling setting, where the user typically opens a mobile application to select origin, destination and departure time, and find matching drivers, we process data in temporal batches and focus on recurring trips. In turn, well known results on human behavior analytics (Gonzalez et al., 2008; Song et al., 2010) show that our mobility is largely predictable, i.e. processing data in batches, rather than in an on-demand basis, covers a large portion of our demand. Moreover, this allows us to gain more room for optimization, as we treat space, time, and interest patterns of users all at once.

Based on all the above, we build GRAAL, a methodology for GReen And sociAL carpooling. GRAAL optimizes a carpooling system, at the city level, not only by minimizing the number of cars needed, but also by maximizing the *enjoyability* of people traveling together. Starting from the concept presented by the authors in Guidotti et al. (2015) we introduce a measure of enjoyability based on people's interests, social links, and tendency to connect to people with similar or dissimilar interests. Specifically, our enjoyability measure takes into account two factors: (i) what we call *like-mindness*, i.e. a topic similarity between any two users; and (ii) what we define as *homophily*, i.e. the tendency of a person to group with similar ones. Previous attempts to use social context in carpooling include putting together in a car people who are friends (Cici et al., 2014). However, by looking only at the direct (or even the two-hop) friends, we may loose other good matches from the optimization model, as the set of potential drivers (or co-passengers) is usually much larger than the typical number of friends pairs in a social network.

In GRAAL, we introduce a multi-objective optimization based on a weighted linear combination of two components: (i) number of cars (which is minimized) and (ii) total enjoyability of the users in the system (which is maximized). In our experiments, we vary this weight, which we refer to as α, between 0 and 1. Moreover, we learn a crowd-sourced value for α by means of an online survey. The survey has the double effect of both confirming the interest of potential carpooling users to a more social solution, and providing us with a realistic estimation of α to use in the optimization model. We present the results of applying GRAAL on real world crowd-sourced data from Twitter, geo-located in the cities of Rome and San Francisco. Results are presented from both the city-wide and the user perspective, and we compare them with different baselines: a random model; a heuristic model aimed at maximizing the user enjoyability; GRAAL model with a value of α equals to one that makes GRAAL minimize only the number of cars (which is derived from the state of art of carpooling); GRAAL model with α set to zero, such that only the maximization of enjoyability is performed. Results show that with the crowd-sourced α, GRAAL is able to reduce the number of cars needed compared to using private vehicles (i.e., each user driving his/her own car), while keeping a high level of total enjoyability. From the user perspective, we show how the per-car distribution of enjoyability is increased with respect to our baselines. We also compare our algorithm with the methodology described in Cici et al. (2014). Although the computational results are based on real-data, however the outcome and the analysis are theoretical as they depend on the assumption that all users in the system would accept our recommendations. To summarize, the main contributions of this paper are:

- we formulate the carpooling problem as a multi-objective optimization of number of cars and enjoyability;
- we learn the weight for the multi-objective optimization by means of a user study;
- we build a methodology extracting enjoyability and mobility demand from data like Twitter;
- we show the results of the application of our methodology on real world data, from the perspectives of the city, and the users, and we evaluate against different baselines.

The paper is organized as follows: related work is presented in Section 2; we define the carpooling problem and the formulation in Section 3; the methodology is described in Section 4; the user study is presented in Section 5; experiments on Rome and San Francisco data are presented in Section 6; limitations and future work are presented in Section 7; the conclusion is summarized in Section 8.