Transportation Research Part C 85 (2017) 573–590

Contents lists available at ScienceDirect

Transportation Research Part C

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

# Commuter ride-sharing using topology-based vehicle trajectory clustering: Methodology, application and impact evaluation

Zihan Hong[a], Ying Chen[b], Hani S. Mahmassani[a,*], Shuang Xu[c]

[a] *Transportation Center, Northwestern University, 600 Foster Street, Evanston, IL 60208, United States*
[b] *Department of Civil and Environmental Engineering, Northwestern University, 600 Foster Street, Evanston, IL 60208, United States*
[c] *Department of Geographical and Sustainability Sciences, The University of Iowa, 223 Jessup Hall, Iowa City, IA 52242, United States*

ARTICLE INFO

*Keywords:*
Ride-sharing
Data mining
Travel demand management
Trajectory clustering

ABSTRACT

This paper illustrates a ride matching method for commuting trips based on clustering trajectories, and a modeling and simulation framework with ride-sharing behaviors to illustrate its potential impact. It proposes data mining solutions to reduce traffic demand and encourage more environment-friendly behaviors. The main contribution is a new data-driven ride-matching method, which tracks personal preferences of road choices and travel patterns to identify potential ride-sharing routes for carpool commuters. Compared with prevalent carpooling algorithms, which allow users to enter departure and destination information for on-demand trips, the proposed method focuses more on regular commuting trips. The potential effectiveness of the approach is evaluated using a traffic simulation-assignment framework with ride-sharing participation using the routes suggested by our algorithm. Two types of ride-sharing participation scenarios, with and without carpooling information, are considered. A case study with the Chicago tested is conducted to demonstrate the proposed framework's ability to support better decision-making for carpool commuters. The results indicate that with ride-matching recommendations using shared vehicle trajectory data, carpool programs for commuters contribute to a less congested traffic state and environment-friendly travel patterns.

## 1. Introduction

The massive use of private cars is associated with systemic congestion and negative impacts on the environment. To mitigate traffic congestion, the Federal Highway Administration (FHWA) has invoked the Dynamic Mobility Applications (DMA) and the Active Transportation and Demand Management (ATDM) programs for the development and deployment of dynamic traffic management strategies and tools (Yelchuru et al., 2013; Vasudevan and Wunderlich, 2013). An underlying premise of these strategies is to influence travelers' behavior towards sustainable travel choices in order to manage congestion, reduce externalities and maximize overall system efficiency.

Carpooling and other forms of ride-sharing emerged in the 1970s as an environment-friendly and sustainable travel alternative. The rationale is that effective usage of empty car seats by ride-sharing may lead to an important opportunity to increase occupancy rates and substantially increase the efficiency of urban transportation systems. Various benefits of carpooling have been shown in the literature, including congestion relief, emission reduction, and individual monetary savings (Katzev, 2003; Fellows and Pitfield, 2000; Chan and Shaheen, 2012; Shaheen et al., 2006, 2012), where the incentives for ride-sharing behaviors and the probability of

* Corresponding author.
*E-mail addresses:* zihanhong2012@u.northwestern.edu (Z. Hong), y-chen@northwestern.edu (Y. Chen), masmah@northwestern.edu (H.S. Mahmassani), shuang-xu@uiowa.edu (S. Xu).

http://dx.doi.org/10.1016/j.trc.2017.10.020
Received 14 July 2016; Received in revised form 11 October 2017; Accepted 16 October 2017
0968-090X/ 

mode and shift towards carpooling have been explored and estimated, primarily for commuting trips (Eash et al., 1974; Ben-Akiva and Atherton, 1977; Train and McFadden, 1978; Dumas and Dobson, 1979; Daniels, 1981; Levin, 1982; Huang et al., 2000; Li et al., 2007; Habib et al., 2011).

From the suppliers' perspective, carpooling serves as a form of public transit for a group of travelers, affording greater spatial coverage but requiring less significant public investment (Buliung et al., 2009). Therefore, planning authorities, especially in congested metropolitan areas, have generally encouraged carpool programs, regarded as an environment-friendly and cost-efficient element of traffic demand management in urban areas. Similarly, employers have facilitated carpooling by employees as a means of reducing on-site parking requirements.

From the users' perspective, the impedance towards greater adoption of carpool commuting is rooted in concerns regarding its reliability, flexibility, efficiency, and security (Li et al., 2007; Chan and Shaheen, 2012). The social, environmental, transportation, and land management benefits from carpooling (Chan and Shaheen, 2012; Shaheen et al., 2006, 2012) encourage the incentives for ride-sharing behaviors but do not provide solutions to its implementation. A practical solution would be to provide the answers to "when and where to carpool" (concerning flexibility and efficiency) and "with whom" (related to security and reliability).

Some on-demand personal mobility providers, such as Uber or Lyft, have emerged and are of growing significance in the urban transportation industry, redefining this market and addressing some of the above concerns. However, these providers primarily aim to provide services for occasional trips, which could occur during any time of the day without predictable nor recognizable patterns (Agatz et al., 2012). Commuting trips, viewed as recurring demand, are beyond their service scope. Unlike occasional trips, commuting trips usually require concentrated supply across spatial and temporal dimensions. The hidden patterns of shared origins, destinations, and compatible arrival times provide clues to a solution for carpooling commuters. The prevalent carpooling algorithms adopted by most providers allow users to enter departure time, origins and destinations for a specific trip; on the other hand, the proposed algorithm focuses more on commuting trips and provides a solution using the hidden patterns of the trips.

This paper aims to explore this solution from a technical point of view, rooted in the analysis of spatio-temporal patterns manifested through individual trajectories. With the widespread use of location sensing technology, spatial data is widely available and easily accessible through advanced information technology. Richer data availability and associated information processing technologies offer promising avenues to promote carpool programs. Particularly, the clustering analysis of drivers' daily trajectories (with locations published on social networks (Cheng et al., 2010) or vehicle trajectories (obtained through probe data) enables tracking patterns of human travel behaviors and activities, thus capturing potential carpool solutions (Han et al., 2012; Kharrat et al., 2008). The rationale behind the proposed solution is that travelers who share similar patterns of trips, including travel behaviors, route choices, and more personalized preferences (i.e. departure time, walking distance from the carpool location, and size of carpool group), could be grouped together. Following this rationale, this paper aims to solve three research questions: (1) what is the hidden pattern of commuter trips, and how to extract those; (2) how to integrate the trip patterns into ride-sharing route design, and (3) how to evaluate the effectiveness of proposed approach in a virtual simulation environment.

The paper first presents a new data-driven, trajectory-based ride-matching method, which tracks personal preferences and aggregated travel patterns, to identify potential ride-sharing routes for carpool commuters. The core part is a trajectory clustering algorithm, which takes advantage of the network topology to detect common sub-paths in a road network with archived vehicle trajectory data and provides potential routes for carpool commuters. To integrate the extracted patterns into ride-sharing route design, we assume that travelers sharing common sub-paths are likely to be matched into a carpooling group. If carpooling information is available in the back-end system (where the ride-matching algorithm is functioning) to extract additional information (e.g., departure time, origins, and destinations), more detailed solutions are achievable to answer implementation questions, such as "When and where do we start and end, and with whom do we carpool?" Finally, the potential effectiveness of the approach is evaluated using a traffic simulation-assignment framework with ride-sharing participation taking the routes suggested by our algorithm. The modeling and simulation framework thus "implements" the carpooling programs, and evaluates the resulting traffic performance with ride-sharing behaviors. Two types of ride-sharing participation tests, with or without carpooling information, are considered. A case study with the Chicago testbed is conducted to demonstrate the proposed framework's ability to support better decision-making for carpool commuters.

The next section presents the framework for the study, followed by a description of the specific algorithms that define the ride-matching and ride-sharing application. Section 4 focuses on the key trajectory matching algorithm, called TOPOSCAN. Section 5 introduces the Chicago testbed, describes the simulation modeling framework used to evaluate the effectiveness of the ride-matching process for commuter carpools, and discusses the results of the case study. Concluding remarks end the paper.

## 2. Framework

The process of this study is illustrated in Fig. 1 with five separate parts: data, method, results, application and evaluation. The input data include the predefined road network data and vehicle trajectory data. The road network consists of the geographic information of intersections and streets. The vehicle trajectory is a series of sequential locations and timestamps for each vehicle. It provides foundational information for tracking travel behaviors. The trajectory data could be obtained from probe dataset.

The TOPOSCAN algorithm is a network topology-based algorithm, which generates the candidates of shared trips and routes. It consists of four parts: network topology modeling, trajectory mapping, path clustering, and similarity evaluation. The first two parts belong to data preprocessing for path clustering. The path clustering is developed according to the density-based clustering (DBSCAN) principle. The similarity evaluation is a post-processing procedure that measures the accuracy of the clustering results. The results would reveal potential carpool commuting routes.