Transportation Research Part C 76 (2017) 51–70

Contents lists available at ScienceDirect

Transportation Research Part C

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

Understanding ridesplitting behavior of on-demand ride services: An ensemble learning approach

Xiqun (Michael) Chen *, Majid Zahiri, Shuaichao Zhang

College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China

ARTICLE INFO

Article history:
Received 30 June 2016
Received in revised form 12 December 2016
Accepted 30 December 2016

Keywords:
On-demand ride service
Rideplitting
Ensemble learning
Boosting
Decision tree

ABSTRACT

In this paper, we present an ensemble learning approach for better understanding ridesplitting behavior of passengers of ridesourcing companies who provide prearranged and on-demand transportation services. An ensemble learning model is a weighted combination of multiple classification models or week classifiers to form a strong classification model. The goal of ensemble learning is to combine decisions or predictions of several base classifiers to improve prediction, generalizability, and robustness over a single classifier. This paper employs the Boosting ensemble by growing individual decision trees sequentially and then assembling these trees to produce a powerful classification model. To improve the prediction accuracy of ridesplitting choices, we explored real-world individual level data extracted from the on-demand ride service platform of DiDi in Hangzhou, China. Over one million trips of the four service types, i.e., Taxi Hailing Service, Express, Private Car Service, and Hitch, are explored with descriptive statistics. A variety of features that may impact ridesplitting behavior are ranked and selected by using the ReliefF algorithm, such as trip travel time, trip costs, trip length, waiting time fee, travel time reliability of origins/destinations and so on. The Boosting ensemble trees with full features and selected features are trained and validated using two independent datasets. This paper also verifies that ensemble learning is particularly useful and powerful in the ridesplitting analysis and outperforms three other widely used classifiers. This paper is one of the first quantitative studies that empirically reveal the real-world demand and supply pattern by exploring the city-wide data of an on-demand ride service platform.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

One of the most important information required for demand modelers and transportation practitioners is to find out passenger's urban travel behavior. In fact, realizing the trip origin and destination, travel time, trip monetary costs and other individual travel information in an area leads to better understanding of travel behavior and demand patterns. This issue is of significance for policy makers to improve comprehensive transportation planning, and formulate urban traffic congestion mitigation strategies.

Most recently, Shaheen et al. (2016) comprehensively reported the various modes of shared mobility that enabled users to obtain short-term access to transportation as needed, rather than requiring ownership, such as carsharing, personal

* Corresponding author at: B828 Anzhong Building, College of Civil Engineering and Architecture, Zhejiang University, 866 Yuhangtang Rd, Hangzhou 310058, China.

E-mail address: chenxiqun@zju.edu.cn (X. (Michael) Chen).

http://dx.doi.org/10.1016/j.trc.2016.12.018
0968-090X/© 2017 Elsevier Ltd. All rights reserved.

vehicle sharing (i.e. P2P carsharing and fractional ownership), bikesharing, scooter sharing, ridesharing, and on-demand ride services. Ridesourcing or transportation network companies (TNCs) refer to an emerging urban mobility service that private car owners drive their own vehicles to provide for-hire rides. Zha et al. (2016) analyzed the ridesourcing market using an aggregate model and concluded that without any regulatory intervention a monopoly ridesourcing platform would maximize the joint profit with its drivers. Recent technologies enable on-demand ridesourcing services via smartphone applications, some of which enable passengers to choose to split a ride and fare in a ridesourcing vehicle, named ridesplitting (e.g., UberPOOL, Lyft Line, and DiDi Hitch). Shaheen et al. (2016) defined ridesplitting as "*a form of ridesourcing where riders with similar origins and destinations are matched to the same ridesourcing driver and vehicle in real time, and the ride and costs are split among users*".

This research analyzes the passenger ridesplitting behavior and travel patterns of an on-demand ride service platform using real-world data provided by DiDi in Hangzhou, China. In particular, understanding the emerging ridesplitting behavior needs a high quality and acquisitive dataset. Usually traditional manual methods of gaining knowledge and solving the problems are difficult and unreliable in case of complex engineering problems in real-world networks, thus advanced models and methods are of urgent necessity especially when there are abundant data. Nowadays, on the basis of exploring huge transportation data from roadways cameras, GPS, smartphones, traffic sensors and so on, traffic engineers can predict valuable traffic behavior and travel patterns in cities or solve the problems with innovative technologies, e.g., machine learning algorithms.

Fig. 1 shows the service model of TNCs, e.g., DiDi. The mobile connection and work flow between passengers and drivers are completed through the on-demand ride service platform. A TNC receives and analyzes requests of passengers and drivers according to program rules. For example, a passenger sends a request to the on-demand ride service platform for hailing a taxi or private car with the information such as the type of service, destination, departure time, number of passengers, car level and ridesplitting willingness. Then the TNC analyzes the request according to the real-time demand and supply nearby, car type, weather, holiday, responses of drivers and other passengers' requests. A matching procedure between the passenger and the most suitable driver nearby is implemented accordingly. The matched driver will receive the information about the passenger, price and suggested route. After finishing the ride, the passenger will pay a trip fee to the platform. The TNC will check out and charge commission at a certain rate. For example, DiDi charges zero commission for taxi drivers, while approximate 20% price paid by passengers of Express, Private Car Service and Hitch. The driver will then be paid by the platform at the remaining price. Different TNCs usually prepare a variety of promotion strategies for their customers according to their archived passenger information and real-time demand. For example, Leng et al. (2016) quantitatively analyzed the impact of promotion fees of two TNCs on the pattern of taxis services using 40-day trip data of over 9000 taxis in Beijing.

More research of ridesplitting is needed to better understand the behavioral characteristics on congestion, mobility, reliability and economy. In these years, solving transportation problems by machine learning has become more and more popular. Researchers can predict many travel activities from various perspectives. Many transportation problems can be

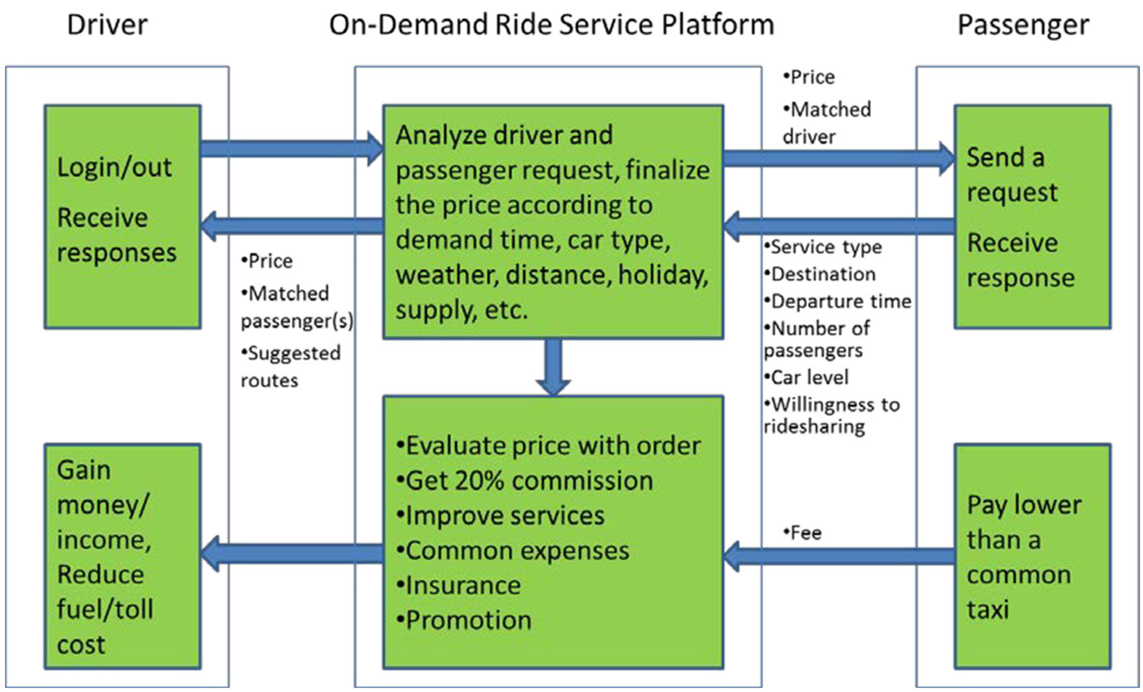

Fig. 1. Service model and work flowchart of an on-demand ride service platform.

Download English Version:

https://daneshyari.com/en/article/4968519

Download Persian Version:

https://daneshyari.com/article/4968519

Daneshyari.com