Pervasive and Mobile Computing 25 (2016) 48–66

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

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

Predicting the passenger demand on bus services for mobile users[☆]

Chunjie Zhou[a,*], Pengfei Dai[b], Fusheng Wang[c], Zhenxing Zhang[a]

[a] School of Software, Ludong University, Shandong, China
[b] Institute of Network and Technology, ICT, Shandong, China
[c] Department of Biomedical Informatics, Stony Brook University, NY, USA

HIGHLIGHTS

- Build a passenger demand prediction system for mobile users.
- Propose a method for detecting parking places.
- Present the design of knowledge base.
- Build a location-based mobile application for mobile users.

ARTICLE INFO

Article history:
Received 25 November 2014
Received in revised form 15 September 2015
Accepted 2 October 2015
Available online 21 October 2015

Keywords:
Bus transport
Traffic congestion
Passenger demand prediction
Mobile users
Predictive models

ABSTRACT

Public transport, especially the bus transport, can reduce the private car usage and fuel consumption, and alleviate traffic congestion. However, when traveling in buses, the travelers not only care about the waiting time, but also care about the crowdedness in the bus itself. Excessively overcrowded bus may drive away the anxious travelers and make them reluctant to take buses. So accurate, real-time and reliable passenger demand prediction becomes necessary, which can help determine the bus headway and help reduce the waiting time of passengers. Based on a large database from a real bus system, this paper aims to present a passenger demand prediction system for mobile users. The system includes a server-side bus information data stream processing and mining program and a client-side mobile application for Android smartphones. The server program continuously monitors for each bus stop the number of passengers waiting at the bus stop, the number of passengers that will pass the bus stop, as well as the traffic conditions in the area around the stop. It delivers real time bus and traffic information to mobile users via restful web services. The client-side location-based mobile application consumes these services to help mobile users make informed transportation choices. For example the availability of buses might be a deterrent when they are too crowded. However, there are three major challenges for predicting the passenger demands on bus services: inhomogeneous, seasonal bursty periods and periodicities. To overcome the challenges, we propose three predictive models and further take a data stream ensemble framework to predict the number of passengers. We develop a prototype system with different types of Android based mobile phones and comprehensively experiment over a 22-week period.

[☆] This research was partially supported by the Project of Shandong Province Higher Educational Science and Technology Program (No. J12LN05); the grants from the National Natural Science Foundation of China (Nos. 61202111, 61273152, 61303017); the Project Development Plan of Science and Technology of Yantai City (No. 2013ZH092); the Doctoral Foundation of Ludong University (No. LY2012023); The Natural Science Foundation of Shandong Province and Special Projects (Grant No. ZR2013FL009); the Natural Science Foundation of Shandong Province China (No. ZR2011GQ001).

* Corresponding author. Fax: +86 535 6673241.
E-mail address: lucyzcj@gmail.com (C. Zhou).

http://dx.doi.org/10.1016/j.pmcj.2015.10.003

The evaluation results suggest that the proposed system achieves outstanding prediction accuracy among 86,411 passenger demands on bus services, more than 78% of them are accurately forecasted.

1. Introduction

With the rapid urbanization, travel demand becomes far beyond current road-network capacity, which directly results in serious traffic congestion. Although the government has been improving the urban transportation system continually, people still have to waste a lot of time on the road. On the other hand, a large number of exhaust emissions from the vehicles in traffic jams cause serious pollution to our environment. Public transport is an efficient approach to alleviate this problem. It can reduce the private car usage and fuel consumption, and alleviate traffic congestion. So public transport, especially the bus transport, has been well developed in many parts of the world. As one of the most comprehensive and affordable means of public transport, in 2011 the bus system serves over 3.3 million bus rides every day on average in Singapore with around 5 million residents. Nowadays, more and more people rely on public transportation in their daily work and life. Very often passengers need to make choices regarding their transportation mode based on constraints such as time to reach destination and so on. There are very few mechanisms for passengers to obtain the necessary information in a 'one-stop-shop'.

Recently, researches about bus arrival time prediction attracted the attention of both bus companies and researchers for the benefit of both transit agencies and passengers. However, when traveling in buses, travelers also care much about the crowdedness in the bus. Excessively overcrowded bus may drive away the anxious travelers and make them reluctant to take buses. One of the main obstacles to the passenger picking is the existing regulations limiting their activity, namely, the licensed areas and applicable fares. Thus, the equilibrium between passenger demands and bus companies is fundamental to maximize profit. An equilibrium fault may lead to one of two scenarios: (1) excess of vacant vehicles and lower passenger reliability or (2) larger passenger waiting time and excessively overcrowded buses. In order to explore passenger travel demand in different areas and in different time periods from transport management perspective, we predict the number of passengers to estimate passenger demands and congestion degree, a major focus of this paper.

Accurate, real-time and reliable passenger demand prediction can help determine the bus headway and help reduce the waiting time of passengers. However, due to a number of stochastic variables, we need to face the following three major challenges: (1) *inhomogeneous*. A periodicity in time on a daily basis that reflects the patterns of the underlying human activity, making the data appear non-homogeneous. (2) *seasonal bursty*. The demands in many stops can be often messed by seasonal bursty periods of expected events such as highly crowded holiday events, weather changes, and so on. (3) *other periodicities*. The passenger demands on bus services vary significantly at different times of a day, different days of a week, or even different seasons.

Aiming to address these challenges and to enable mobile users make informed transportation decisions, in this paper, we present a system that delivers real-time bus status and traffic information to passengers' mobile phones. Our system continuously monitors the numbers of passengers waiting at the bus stop and the number of passengers that will pass the bus stop. Furthermore, it also monitors the traffic conditions in the area around the stop. With real-time information of bus availability and traffic condition, passengers can make informed decisions. In order to achieve this, we predict the number of passengers (passenger demand) over space (bus stop) for a short-time period of P-minutes. Based on historical GPS location and service data of passenger pick-up and drop-off, time series histograms are built for each bus stop containing the number of services with an aggregation of P-minutes. To predict passenger demands on bus services, we develop a unique predictive model by adapting time series forecasting techniques to our problem. The time varying Poisson model is used to solve non-homogeneous data, as shown in Section 5.1. The time varying Poisson model can be faced as a time dependent average. In order to settle seasonal bursty periods of expected events, the weighted time varying Poisson model is proposed in Section 5.2. The two last models assume the existence of a regular (seasonal or not) periodicity in the bus service passenger demands. However, the demands can be different from stop to stop and the existence of other periodicities should be explored. So the autoregressive integrated moving average model is used to achieve a better accuracy prediction. Our goal is to predict at the instant `t` how many bus services will be demanded during the period `[t, t+P]` at each existing bus stop. The real service count in `[t, t+P]` extracted from the data can be reused to compute count of services for the next period `[t+P, t+2P]`, thus the approach can run continuously over streaming data. To the best of our knowledge, such approach has no parallel in the literature.

Such a system is feasible nowadays because most buses are now equipped with GPS devices and most passengers have mobile phones. For example, Singapore has more than 14,000 vehicles carrying GPS devices and the vehicles send location and status to a central server every certain seconds, and the mobile phone penetration in Singapore is 145% (many people carry more than one mobile phone) and 72% of mobile phones in Singapore are smartphones. Our system is built based on existing infrastructure.

In our work, we have conducted a real study using a dataset obtained by a large-sized bus network containing a total of 416 bus stops and 1326 vehicles running in the city of Yantai, China. Our test-bed is a computational stream simulation