Applied Soft Computing 47 (2016) 460–482

Contents lists available at ScienceDirect

Applied Soft Computing

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

An online learning approach to eliminate Bus Bunching in real-time

Luís Moreira-Matias [a,*], Oded Cats [c], João Gama [b,e], João Mendes-Moreira [b,d], Jorge Freire de Sousa [f,g]

[a] NEC Laboratories Europe, Kurfürsten-Anlage 36, 69115 Heidelberg, Germany
[b] LIAAD-INESC TEC, 4200-465 Porto, Portugal
[c] Dep. Transport and Planning, Delft University of Technology, 2600 GA Delft, The Netherlands
[d] Dep. de Eng. Informática, Fac. de Engenharia, U. Porto, 4200-465 Porto, Portugal
[e] Faculdade de Economia, U. Porto, 4200-465 Porto, Portugal
[f] UGEI-INESC TEC, 4200-465 Porto, Portugal
[g] DGEI, Fac. de Engenharia, U. Porto, 4200-465 Porto, Portugal

ARTICLE INFO

Article history:
Received 21 November 2014
Received in revised form 1 June 2016
Accepted 19 June 2016
Available online 24 June 2016

Keywords:
Online learning
Bus Bunching
Stochastic Gradient Descent
Travel Time Prediction

ABSTRACT

Recent advances in telecommunications created new opportunities for monitoring public transport operations in real-time. This paper presents an automatic control framework to mitigate the Bus Bunching phenomenon in real-time. The framework depicts a powerful combination of distinct Machine Learning principles and methods to extract valuable information from raw location-based data. State-of-the-art tools and methodologies such as Regression Analysis, Probabilistic Reasoning and Perceptron's learning with Stochastic Gradient Descent constitute building blocks of this predictive methodology. The prediction's output is then used to select and deploy a corrective action to automatically prevent Bus Bunching. The performance of the proposed method is evaluated using data collected from 18 bus routes in Porto, Portugal over a period of one year. Simulation results demonstrate that the proposed method can potentially reduce bunching by 68% and decrease average passenger waiting times by 4.5%, without prolonging in-vehicle times. The proposed system could be embedded in a decision support system to improve control room operations.

1. Introduction

Major urban areas worldwide are intensely equipped with sensors able to monitor daily human activities. Vehicles, Smartphones, Inductive Loop Counter and License Plate Recognition are some examples of equipments that can collect this type of information in real-time. The recent development of communicational technologies such as GSM-based (Global System for Mobile Communications) 3G, Global Positioning System (GPS) and WiFi enable to *access* this data in *near real-time*.

These location-based technologies provide an unprecedented opportunity to develop large scale monitoring frameworks. Many industries are increasingly taking advantage of such data to improve their services and monitor their operations. Public transportation (PT) companies offer a good showcase of this trend: they operate in a highly competitive environment where the analysis of real-time data could potentially lead to service improvements and ultimately increase their market share. Service *reliability* is key in maintaining their profitability.

In the last decades, Bus Dispatching Systems have been deployed in many PT companies worldwide. These systems were first used to monitor fleet operations. Automatic Vehicle Location (AVL) data collected by these systems was later used for offline evaluation of service performance and reliability. These systems became therefore indispensable for PT service providers. The inherent uncertainty associated with PT operations could be mitigated in real-time by the *Control Centre*. However, it is often labour-intensive as traffic dispatchers need to take numerous decisions in response to rapidly evolving conditions.

Many researchers highlighted the potential of the stored AVL data to provide insights on how to improve PT reliability, e.g. [67,27]. In particular, previous work tended to focus on offline ex-ante evaluation. However, the real time availability of AVL data opens new research directions for improving PT reliability, namely introducing real-time decision models to support *Operational Control*.

PT reliability could be defined either in terms of punctuality – the extent to which operations adhere to the planned schedule – or in

* Corresponding author. Tel.: +49 6221 4342261.
E-mail address: luis.moreira.matias@gmail.com (L. Moreira-Matias).

http://dx.doi.org/10.1016/j.asoc.2016.06.031
1568-4946/

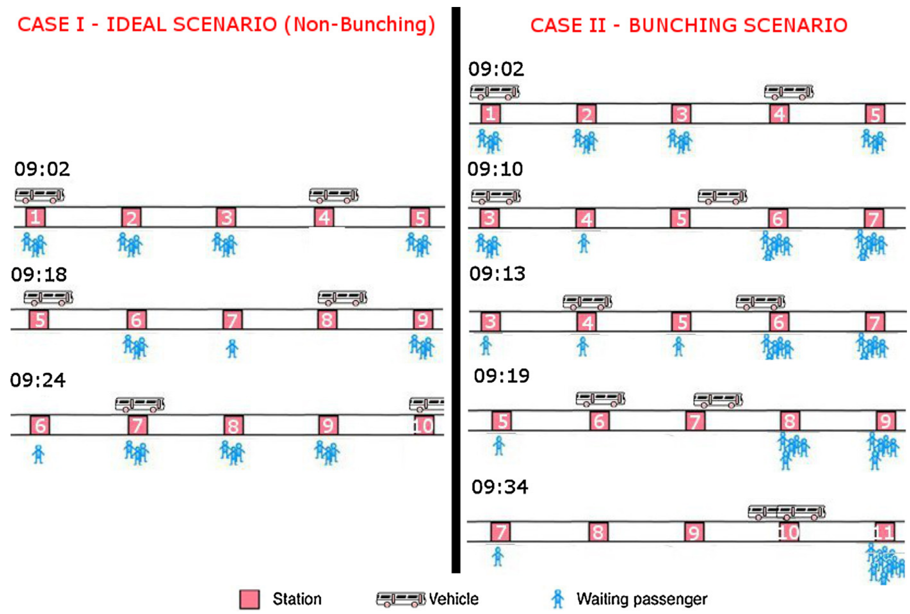

Fig. 1. Bus Bunching illustration.

terms of regularity – the extent to which vehicles are evenly spaced, implying even headways, the time interval between successive vehicles running on the same route [69]. In the case of high-frequency routes (headways of 10 min or less), regularity is the main indicator of service reliability since it is the main determinant of passenger waiting times [10]. Headways are inherently instable due to a positive feedback loop between the headway, the number of passengers waiting at the stop, dwell times and successive headways [20]. For example, a small delay provokes an increase in the number of passengers in the next stop. This in turn leads to an increase in the dwell time (bus service time at a stop) and consequently, it further increases the bus delay. At the same time, the next bus will have fewer passengers, shorter dwell times and will gradually catch up with the preceding bus. This snowball effect will result in pairs of buses forming a platoon as illustrated in Fig. 1. This phenomenon is denominated *Bus Bunching (BB)* [20,43,45].

The prevalence of BB is one of the most visible blueprints of an unreliable service. Two (or more) buses running together on the same route is an undeniable sign that something is going terribly wrong with the company's ser'vice. Operational Control can potentially address BB in real-time. This paper focuses on using both historical and real-time AVL data to deploy automatic control strategies, to mitigate BB while reducing the human workload required to make these decisions. It provides a *complete bottom-up methodology*, from fundamental theoretical aspects that capture the BB formation process to practical issues associated with the deployment of corrective actions, as well as with the evaluation of their impacts.

1.1. Literature review

Previous studies have deployed a range of analytical and simulation models to represent the dynamics of the bus service operations and evaluate the impacts of alternative control strategies. Early analytical studies that examined the BB phenomenon and the characteristics of its instability triggered by recurrent perturbations include Newell and Potts [50], Chapman and Michel [17] and Powell and Sheffi [55]. The latter devised a probabilistic model which builds a set of recursive relationships to calculate the p.d.f. to validate the hypothesis of generating a vehicle platoon formation at each stop. More recently, Daganzo [20] and Daganzo and Pilachowski [21] developed analytical models to assess the impacts of an adaptive control strategy which adjusts bus dwell times at stops and the running times between successive stops based on the respective headways.

Transit control strategies consist of a wide variety of operational methods aimed to improve transit performance and level of service. Holding strategies are among the most widely used transit control methods aimed to improve service regularity [2]. In order to implement corrective strategies and consequent actions, both the location – where the control decisions should be deployed [71,1,26,68,15] and the how – the criteria for intervening and its specification [32,39,13] – must be determined. In [24], a global control unit optimizes the holding times by solving a deterministic rolling horizon mathematical programming model which minimizes total passenger waiting times.

While previous studies might be effective in mitigating BB occurrences and therefore reducing service uncertainty, none of the abovementioned studies involved a systematic proactive approach to eliminate BB in real-time. This paper proposes online learning techniques to address headway instability by simultaneously considering historical and real-time data concerning service perturbations. Moreover, a comprehensive procedure for BB event detection and corrective action deployment is developed and applied.

1.2. Scope and objectives

Most of the abovementioned state-of-the-art research on this topic departs from the assumption that the probability of BB events is minimized by maximizing headway stability. This is achieved by either minimizing the difference between the actual headway and the scheduled one or by minimizing the discrepancies between successive headways. Notwithstanding its validity, this approach requires multiple control actions (i.e. speed modification, bus holding, etc.) which may impose high mental workload for both drivers and control centre staff, yielding results with sub-optimal decision making and low compliance rates.

Hereby, we propose a *proactive* rather than a *reactive* operational control framework. The fundamental idea is to estimate the likelihood of a BB event occurring further downstream and then deploy a corrective control strategy.