Transportation Research Part C 37 (2013) 42–72

Contents lists available at ScienceDirect

Transportation Research Part C

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

Fundamental understanding on the use of Bluetooth scanner as a complementary transport data

Ashish Bhaskar [a,*], Edward Chung [b,1]

[a] Smart Transport Research Centre, School of Civil Engineering and Built Environment, Faculty of Built Environment and Engineering, Queensland University of Technology, Gardens Point Campus-S723, 2 George St., GPO Box 2434, Brisbane, QLD 4001, Australia

[b] Smart Transport Research Centre, School of Civil Engineering and Built Environment, Faculty of Built Environment and Engineering, Queensland University of Technology, Gardens Point Campus-P850, 2 George St., GPO Box 2434, Brisbane, QLD 4001, Australia

ARTICLE INFO

Article history:
Received 31 January 2013
Received in revised form 25 September 2013
Accepted 25 September 2013

Keywords:
Bluetooth
Traffic simulation
Communication simulation
Travel time
Signalised network
Bluetooth scanners
Data

ABSTRACT

Literature is limited in its knowledge of the Bluetooth protocol based data acquisition process and in the accuracy and reliability of the analysis performed using the data. This paper extends the body of knowledge surrounding the use of data from the Bluetooth Media Access Control Scanner (BMS) as a complementary traffic data source. A multi-layer simulation model named Traffic and Communication Simulation (TCS) is developed. TCS is utilised to model the theoretical properties of the BMS data and analyse the accuracy and reliability of travel time estimation using the BMS data.

1. Introduction and literature review

Transport agencies collect data from multiple sources for various applications related to monitoring, management and control, planning and finance. Advancement of technology has resulted in the production of numerous data retrieval systems, ranging from traditional loop detectors to Vehicle Information and Communication Systems (VICS). These systems are broadly classified as: (a) Fixed sensors (such as loops, Automatic Number Plate Recognition System (ANPR)) that provide traffic information at the location where the sensors are installed and (b) Mobile sensors (such as GPS equipped vehicles, Automatic Vehicle Location (AVL)) that provide data for the entire journey of the vehicle equipped with such sensors.

In early 2000, researchers explored the use of Bluetooth (BT) technology for the automotive industry. Nusser and Pelz (2000) presented the architecture of the Bluetooth network as an integral part of in-car communication and information systems. Researchers (Sawant et al., 2004; Murphy et al., 2002; Pasolini and Verdone, 2002) have tested the proof-of-concept for the use of BT for Intelligent Transport System services, and have verified that the BT equipped devices in moving vehicles could be discovered.

Recently, there has been significant interest from transport agencies in exploiting the Bluetooth Media Access Control Scanner (BMS) as a complementary transport data source. The concept behind BMS is rather simple. A BMS scanner has a communication range (say around 100 m in radius) that we term as *zone*. The *zone* is scanned to read the Media Access Control addresses (MAC-ID) of the *discoverable* BT devices transiting within the *zone*. The MAC-ID is a unique, alpha-numeric

* Corresponding author. Tel.: +61 7 3138 9985.
E-mail addresses: ashish.bhaskar@qut.edu.au (A. Bhaskar), edward.chung@qut.edu.au (E. Chung).
[1] Tel.: +61 7 3138 1143.

0968-090X/$ - see front matter
http://dx.doi.org/10.1016/j.trc.2013.09.013

string, that is communicated by the *discoverable* BT device. Most of the portable electronic devices such as mobile phones, car navigation systems and headphones are equipped with BT and its usage is increasing. Indeed, the strategic analytics of Special Interest Group (SIG) – an organisation devoted to maintain the BT technology – forecasts that around 70% of all the new cars will have a BT connectively by 2016; similarly, BT enabled mobile handsets will exceed 1.6 billion units by the end of 2015 (SIG, 2010).

Travel time (1) of a MAC-ID observed at time-synchronised BMS *zones* (u and d), can be directly obtained by matching the MAC-IDs and corresponding time stamp from the two *zones*.

$$TT(m,u,d) = T(m,d) - T(m,u) \tag{1}$$

where $TT(m,u,d)$ is the travel time from the *zone* u to the *zone* d for the device with MAC-ID = m; $T(m,k)$ is the time when the MAC-ID = m is observed at the kth BMS zone.

The matched travel time data do contain noise due to reasons such as:

(a) *Unknown mode*: Obtained travel time is for the BT device transported by a traveller utilising any mode (car, bus, bicycle, pedestrian, etc.) of transport. Different modes have different travel time depending on their operational and behavioural characteristics. If one is interested in car travel time, then the presence of pedestrians or bicycles can result in unrealistic high travel time values and vice-versa.
(b) *No information outside the zone*: The estimate is only from the data available at *zones*, hence the actual travel pattern of the vehicle between the *zones* is unknown. A vehicle can rest along the route or can take a different route with significantly different travel time than that of the assumed route.
(c) *Multiple matches*: Especially on arterial networks, a device can be observed at a *zone* and then it might take a detour, return to the same zone, and thereafter travel to the next *zone*. In such situations, a device can be observed twice at the first *zone* and only once at the second *zone*, resulting in two travel time values. Similarly, other combinations of multiple matches can occur, resulting in the noise.
(d) *Missed observation*: A BT device has a probability of being discovered at a zone and not all devices passing the *zone* are discovered. For instance, say a device travels twice between zones u and d. During its first trips, the device was observed at u at time tu1, however it was missed at d. During its second trip, it was observed at d at time td2, but missed at u. Such observations will result in noisy travel time from u to d as (td2 − tu1). Similarly, other combinations of observations can result in inaccurate travel time values.

In addition to the above, there are other issues related to the BMS data acquisition, which will be discussed later. In literature, filtering techniques, such as Moving Median Filter (Wang et al., 2011), Median Absolute Deviation (MAD) (Kieu et al., 2012), Box-and-Whisker filter (Tsubota et al., 2011) and other techniques utilising *Greensheld's model* and *least median of squares* (Van Boxel et al., 2011) and *multiple matched filter* (Kieu et al., 2012) have been utilised to reduce noise from the directly matched travel time values.

Especially, on urban arterials, it is very challenging to estimate travel time from loops. Models have been proposed to estimate and predict travel time using traditional data sources on arterials (Bhaskar et al., 2011; Bhaskar et al., 2010; Bhaskar et al., 2009; Bhaskar et al., 2012; Kwong et al., 2009) and motorways (Zhang and Rice, 2003; van Lint et al., 2005; Sun et al., 1999; Soriguera and Robusté, 2011; Paterson and Rose, 2008; Li and Rose, 2011; Khosravi et al., 2011; Fei et al., 2011; Coifman and Krishnamurthy, 2007; Coifman and Kim, 2009). BMS data provides significant benefit to the road operators for estimating the travel time on road networks in a very cost effective manner. In literature, travel time from BMS data is compared with that from video cameras for motorways (Wang et al., 2011) and arterial (Mei et al., 2012) and promising results are reported.

For other studies (Haghani and Aliari, 2012), travel time obtained from traditional matching of BMS data is being considered as ground truth travel time. BT tracking is not only being explored for car travel times estimation, but also for other applications such as bicycle travel time (Mei et al., 2012), travel patterns of people movement in airports and shopping malls (Bullock et al., 2010; Malinovskiy and Wang, 2012; O'Neill et al., 2006; Abedi et al., 2013), work zone delays (Haseman et al., 2010), Origin–Destination estimation (Barceló et al., 2012; Blogg et al., 2010; Barceló et al., 2010), route choice analysis (Hainen et al., 2011; Carpenter et al., 2012), and freeway travel time variability (Martchouk et al., 2011).

The amount of data collected from a BMS location depends on numerous factors related to the penetration of BT in the vehicles, software and hardware related with the Bluetooth protocol, etc. Researchers have experimented with the BMS antennae types, position and number against the quantity of data collected. A difference in the quantity of data collected by different antenna type, has been reported. Porter et al. (2011) recommends vertically polarised antennae with gain between 9 and 12 dBi. Vo et al. (2012) and Click and Lloyd (2012) reports that the use of more than one BMS at the site can increase the quantity of data collection.

As can be inferred from the above, literature is primarily focussed on the applications of direct match of BT MAC-ID's at different BMSs; there is limited fundamental understanding on the Bluetooth protocol based data acquisition process, and its impact on the accuracy and reliability of the analysis performed using the data. The objective of this paper is to fill this gap by modelling the theoretical properties of the BMS data and analysing the accuracy and reliability of travel time estimation using the data. To fulfill this objective, firstly the Bluetooth communication and data acquisition process is reviewed. Then, a