BUS AND CAR TRAVEL TIME ON URBAN NETWORKS: INTEGRATING BLUETOOTH AND BUS VEHICLE IDENTIFICATION DATA

Le Minh Kieu, Smart Transport Research Centre, Faculty of Built Environment and Engineering, Queensland University of Technology, Brisbane, Australia

Ashish Bhaskar, Smart Transport Research Centre, Faculty of Built Environment and Engineering, Queensland University of Technology, Brisbane, Australia

Edward Chung, Smart Transport Research Centre, Faculty of Built Environment and Engineering, Queensland University of Technology, Brisbane, Australia

ABSTRACT

Travel time in an important transport performance indicator. Different modes of transport (buses and cars) have different mechanical and operational characteristics, resulting in significantly different travel behaviours and complexities in multimodal travel time estimation on urban networks. This paper explores the relationship between bus and car travel time on urban networks by utilising the empirical Bluetooth and Bus Vehicle Identification data from Brisbane. The technologies and issues behind the two datasets are studied. After cleaning the data to remove outliers, the relationship between not-in-service bus and car travel time and the relationship between in-service bus and car travel time are discussed. The travel time estimation models reveal that the not-in-service bus travel time are similar to the car travel time and the in-service bus travel time could be used to estimate car travel time during off-peak hours.

INTRODUCTION

Monitoring the traffic is a growing issue in any urban network. Travel time is one of the earliest and most important measures to evaluate the performance of an arterial or freeway system. Bus and car travel time estimations on urban areas have been extensively studied in the literature. However, there is limited knowledge on the relationship between multimodal (bus and car) travel times on the urban networks. A comprehensive understanding of the multimodal travel time is an important performance measure for managing the traffic.

Urban network has interrupted traffic flow due to conflicting areas such as intersections. Vehicles on urban networks are not conserved due to mid-link sources and sinks such as parking or side streets, and most of the urban links are shared by different modes of traffic such as cars, buses, bikes, and pedestrians etc. All these characteristics add to the complexities in travel time estimation on urban networks.

Models have been proposed for car travel time estimation (Ashish Bhaskar, Chung, & Dumont, 2010; A. Bhaskar, Chung, & Dumont, 2012) and bus travel time prediction (Abdelfattah & Khan, 1998; Chen, Liu, Xia, & Chien, 2004; Chien, Ding, & Wei, 2002; Jeong & Rilett, 2005; Kalaputapu & Demetsky, 1995; Shalaby & Farhan, 2004). Delay observed at the signal is a significant part of car travel time, and this results in significant variation in travel time between different cars (Ashish Bhaskar, Chung, & Dumont, 2011). Buses have different operational and mechanical behaviours than those of cars. Buses have to stop at bus stops for boarding/alighting of passengers. Their dwell time at stops is believed to accumulate up to 26% of the total travel time of buses (Levinson, 1983) and it is fluctuated due to changes in passenger demand, bus stop spacing, fare payment process, vehicle types, in-vehicle circulation and platform crowding. Buses also tend to use the left most lanes (for left hand driving) on the roads and their mechanical characteristics are also not similar to cars. Finally, there could be some operating Public Transportation Priority Systems that give priority to only the buses and affect their travel time. Those differences between bus and car increase the complexities of multimodal travel time estimation and complicate the relationship between bus and car travel time. The primary objective of this paper is to explore the relationship between bus and car travel time which is fundamental to the understanding of the multimodal travel patterns on urban networks.

The development of technology has resulted in many advanced traffic data retrieval systems from traditional loop detectors to modern on-board electronic systems. Installing the roadway infrastructure could help monitoring traffic flow and occupancy at specific locations on the network. On the other hand, mobile sensors such as probe vehicles collect data of the entire journey. Another method for traffic monitoring is direct measurement of travel time. Some examples of those methods are license-plate matching, toll tag tracking mechanical equipment attached to odometers or global positioning systems (Wasson, Sturdevant, & Bullock, 2008) and Radio Frequency Identification (RFID) of vehicles. In last couple of years there has been tremendous interest in collecting Machine Access Control (MAC) address along the road side using Bluetooth scanners and utilising it as a complementary data source for traffic monitoring. In this paper, two datasets of Bluetooth and Bus Vehicle Identification (VID) from Brisbane network are integrated to achieve the aforementioned objectives.

The paper is structured as follows: Next two sections focus on the issues and benefits related to Bluetooth data VID data, respectively. Thereafter, the methodology for the integration of both Bluetooth and VID Data from a study in Brisbane is presented and finally, the relationship between car and bus travel time is defined.

BLUETOOTH AS A COMPLEMENTARY TRAFFIC DATA SOURCE

Bluetooth technology provides an easy and wireless way for devices to connect with each other in short range. A variety of Bluetooth installed products such as mobile phones, cars, laptops, headsets, etc makes Bluetooth one of the most popular technologies in the modern society. Data package could be transmitted between Bluetooth-enabled devices within the distance up to 100 m, depends on the power rating of the each members of the connection. Each Bluetooth enabled device has a unique 48-bit identifier called Machine Access Control (MAC) address. This unique address is transmitted between the devices during communication. The availability and uniqueness of MAC address makes it possible to track the movement of these devices on the transport network, which opens avenues for its application for travel time estimation, OD estimation, and route identification.

Vehicle with activated Bluetooth devices passing Bluetooth scanner locations could be observed and their timestamp could be recorded. Travel time between scanners could be calculated by matching MAC addresses at different detectors as the difference in the observed timestamps at these detector locations (Haghani, Hamedi, Sadabadi, Young, & Tarnoff, 2010; Tsubota, Bhaskar, Chung, & Billot, 2011). Since each MAC address is unique, traditional matching algorithm similar to the ones used in license plate, cellular and toll tag tracking could be used (Wasson, et al., 2008). Wasson et al. (2008) proposed a method for collecting travel time from Bluetooth devices' MAC address along freeway and arterial street in Indianapolis, U.S. The authors concluded that arterial data has significantly much variance than that of freeway. However, even with large variance, the traffic patterns in peak and off peak hours are easily identified and it is possible to use MAC address matching for travel time estimation. Ahmed et al. (2008) claimed that their method could track cars travelling at 0 to 70 km/h with the accuracy of ± 10% - 15%. Other authors also believed that using Bluetooth for traffic monitoring is a cost effective and relatively accurate method to estimate car travel time (Barceló, Montero, Marquès, & Carmona, 2010; Blogg, Semler, Hingorani, & Troutbeck, 2010; Haghani, et al., 2010; Quayle, Koonce, DePencier, & Bullock, 2010; Tarnoff et al., 2009). Haghani et al. (2010) compared the performance of Bluetooth travel time collection method with the traditional floating car method. Their results illustrated that the difference between these two methods were not significant. Malinovskiy et al. (2011) also revealed that most travel time estimated by Bluetooth data were within 10% of the ground truth.

Although the possibility of using Bluetooth for travel time estimation is proved as positive, there are some issues related to the Bluetooth technology that could negatively affect the accuracy of the method. For instance:

- Bluetooth data corresponds to the Bluetooth device being transported by the traveller. The traveller can be using any mode of transport (car, bus, bike or pedestrian) resulting in data from multimode. Moreover, there can be more than one device being transported by a traveller (For instance, Bluetooth enabled headphone and mobile being used by a traveller in car) or a group of travellers with an active Bluetooth device sharing the same mode (for instance bus). In most of the studies it is assumed that the Bluetooth data points represent a car. The aforementioned issues, especially in urban environment, will result in significant bias and errors in travel time estimation.
- Bluetooth data is from a Bluetooth device being scanner over a zone (scanner coverage area, typically 100 meters) and hence the exact location of the device is not known. This is not a major issue on freeways, but on urban signalised networks can result in significant error in travel time estimation.
- The last issue is the low sampling rate compared to the traffic volume. Sharifi et al. (2010) in a case study in Maryland and Delaware found that the sampling rate (percentage of Bluetooth counts in total traffic flow) of the Bluetooth sensor is between 2% and 5% of the traffic volume. Malinovskiy et al. (2012) again experienced this low sampling rate between 2 and 5% in their study on pedestrian walking behaviour in Montreal, Canada and Seattle, U.S. The low sample rate could lead to the lack of sample size in some uncongested roads.

VEHICLE IDENTIFICATION (VID) DATA

Automatic identification procedures have been very popular in providing information about movement of people, animals and goods. One of the popular technologies in automatic identification is Radio Frequency Identification (RFID) technology. RFID is the technology that uses radio-frequency magnetic or electromagnetic for communication between a tag attached device and a reader (Finkenzeller & Muller, 2010). Unlike the bar code or smart cards technology, the tag does not need to be in direct contact or within line of sight of the reader and could be installed in the tracked object.

In traffic engineering, RFID technology could be used in toll application, parking management, vehicle identification, traffic control and theft protection of cars due to its ability in identifying and tracking of vehicles. Blythe (1999) provides a detailed review of the practices and issues of RFID applications in road tolling, road-use pricing and vehicle access control. Some other authors (Chon, Jun, Jung, & An, 2004; Lee, Oh, & Gerla, 2012) explored the possibility of using RFID for replacing or improving the accuracy of GPS in positioning.

Since each vehicle also has a unique ID tagged to its RFID device, this data could also be used for matching vehicle locations and estimating travel time between the locations similar to the Bluetooth technology mentioned in the previous section. Researchers have utilised the RFID technology to estimate travel time between toll gates (Swedberg, 2004) and OD matrices on motorway networks (Baek, Lim, Rhee, & Choi, 2010). Sriborrirux et al. (2008) also explored the possibility of using RFID technology in monitoring and scheduling of the bus fleet in Bangkok, Thailand. Their paper described the communication protocols, formats and functions in the design and development of the system. However, the performance of the system in terms of travel time estimation was not mentioned. Seo (2008) proposed a simulation study of a "futuristic study" where RFID tagging of cars is used for traffic data collection. Every car was assumed to have its own RFID tag and RFID readers were installed at intersections for collecting link velocity and travel time between the readers. The study was conducted on simulation without validation using observation data.

The RFID technology was invented a long time ago, from the World War II when it was used to differentiate between friendly and enemy aircrafts (Stockman, 1948). Even though the technology improvements have increased the detecting range, accuracy and reduced the equipment's' price, RFID technology has not been widely used in travel time estimation. The problems related to RFID technology could be classified into technical problems and security problems.

- On technological issues, the RFID technology has been implemented in different fields and by different ways. An official international standard is still missing, which could cause problems on conversions and upgrades. Secondly, the system could be jammed relatively easy both from the reader side (because of too many signals from tags) and tags side (because tags cannot response to simultaneous queries from readers). Finally, the costs of tags are still relatively high compared to other automatic vehicle identification systems
- The security and privacy concerns are the principal issues in the development of RFID technology in private vehicles. The RFID tags could be read from distance and without the knowledge or approval of the tag bearer. More importantly, unlike the Bluetooth technology, where each device has a unique MAC address but there is no MAC address database for matching and find information of the devices; for the RFID case because of the purpose of identification, there is always a RFID ID database. These two characteristics could lead to serious security and privacy problems for RFID tagged vehicles' owners.

INTEGRATING BLUETOOTH AND VID DATA FROM BRISBANE

Study site and data description

The Brisbane City Council has installed Bluetooth scanners at major corridors and RFID scanners at intersections of the major bus routes in Brisbane for monitoring traffic and transit operations. The Bluetooth data consists of the device ID (encoded from the unique MAC address of each Bluetooth device), the timestamp where each device is identified and the duration (time difference between the first and last identification of a Bluetooth device at a scanner location). The table 1 illustrates an example of Bluetooth data of a Device ID on 4th Sep 2011.

Table 1: Sample Bluetooth data

Sample ID	Device ID	Intersection ID	Timestamp	Duration
1	1	1	2011/08/04 09:23:26	105
2	2	1	2011/08/04 09:42:15	76
3	3	1	2011/08/04 11:32:07	65
4	4	1	2011/08/04 12:59:23	1

A sample from the VID data includes the Vehicle ID of an in-service or not-in-service bus, vehicle tag, intersection ID and the timestamp when the vehicle is identified. The table 2 shows an example of VID data on 4th Sep 2011.

Table 2: Sample VID data

Timestamp	Intersection Id	Vehicle ID	Vehicle tag
2011/08/04 09:15:24	1	1	BBT696,623700I103301
2011/08/04 09:42:58	1	2	BBT697,589000I103531
2011/08/04 10:54:11	1	3	BBT1878,622200I353011

The vehicle tag includes the bus identification number, service number, start time (according to the bus' schedule) and bus route number. The study site for the analysis is an urban corridor between two intersections in Brisbane, Queensland. In-service buses are the buses which are servicing at all the bus stops along the study area for dwelling. Only bus route 227 could be defined as in-service bus along the area because it is servicing all the stops. Other buses such as route 230, 235 and 236 are only serving part of the corridor and are therefore not included in the analysis. The not-in-service buses are the buses which do not stop at the stops along the corridor. They are from the bus routes which are not serving the corridor, but for some reasons they are travelling along our studied road. Examples of those reasons are travelling from bus terminal to depot stop in the morning, or travelling back to the bus terminal in the afternoon. The Figure 1 illustrates our study area between intersection A and B.

Figure 1: The study site between two intersections: Junction Rd/Wynnum Rd (Intersection A) and Hawthorn Rd/Wynnum Rd (Intersection B), Brisbane, Queensland.

The two intersections have Bluetooth scanners for collecting the MAC address and VID scanners for collecting the bus number, route, service number of each bus pass by the intersections. The total length of study site is 2.2 km and direction of study is inbound traffic (from A to B). Along the study corridor there is a short section with temporal bus lane, which only buses can use from 7-9 AM. The analysis has been carried out for 4.5 months data from 16th July 2011 to the end of November 2011. Because both MAC address and VID is unique of each vehicle, a simple algorithm has been used to match a vehicle upstream with the same vehicle downstream for calculating the travel time between A and B. The travel time is the time difference between the timestamp the vehicle leaves the upstream intersection to the moment it leaves the downstream intersection.

There are not many buses running along the studied corridor. Bus number 227 is servicing 2-4 bus an hour during peak hours and 1 bus an hour during off-peak periods. There are less than 10 not-in-service buses running between A and B per day during weekdays and almost no not-in-service bus is appeared during weekends. It is assumed that the RFID sensors could identify and keep records of all buses with RFID tag passing by. The travel time of buses between VID intersections are direct calculated by matching the same Bus Vehicle ID at upstream and downstream intersections and taking the difference of timestamps. This travel time value is then assumed as the representative of the bus travel time along the studied corridor.

As mentioned earlier, the Bluetooth technology has been used as a method for traffic monitoring and proved as adequate for travel time estimation. The issues related to segregating the Bluetooth data for different modes, bias in travel time estimation due to more than one device being transported by the traveller, etc are beyond the scope of this research. In this study we assumed that the travel time calculation from Bluetooth is the representative of the car travel time. The travel time is calculated by subtracting the exit time at downstream intersection (i.e. the entrance time plus duration) to the exit time at upstream intersection of the same matched device ID. The data is filtered for the outliers the details for which are presented in the next section.

Bluetooth Data cleansing

Various kinds of noise could be observed in the Bluetooth dataset. These noises could lead to overly long travel time along the studied corridor and should be removed from our analysis. Three filters are applied for removing outliers from the Bluetooth dataset:

- *Unrealistic travel time filter*: samples with travel time between intersection A and B larger than 1800 seconds or smaller than 120 seconds are considered as unrealistic and removed from our analysis. It is unlikely that it takes a vehicle more than 1800 seconds or less than 120 seconds (the speed limit along the corridor is 60 km/h)
- *Multiple Matching filter*: during matching of MAC address between the two intersections, we can experience multiple matchings of a device at upstream and downstream. This leads to one-to-many, many-to-one and many-to-may problem. The multiple matching issues are explained with the help of examples on Figure 2 and 3.The one-to-many problem is illustrated in the Figure 2

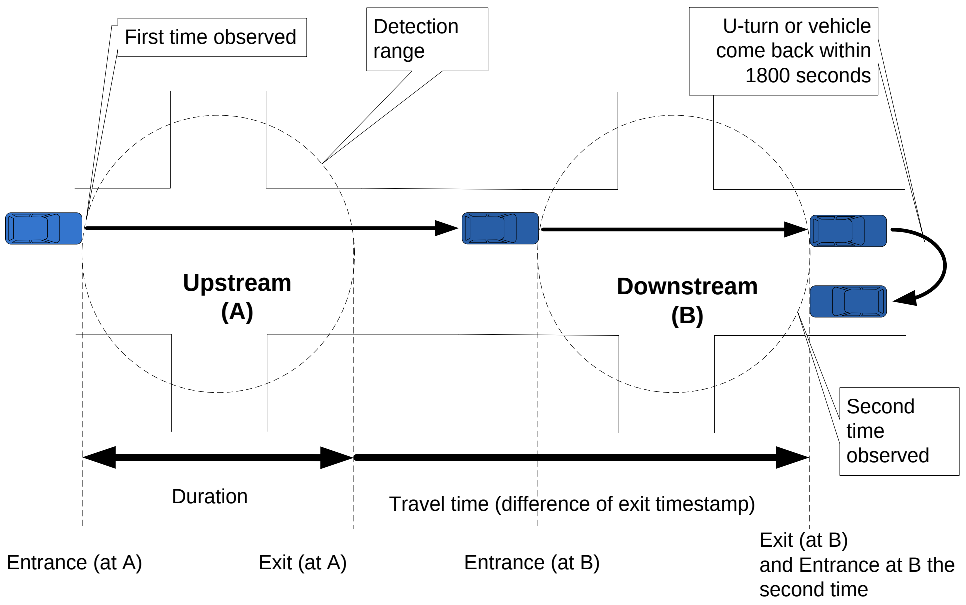

Figure 2: One-to-many problem of Bluetooth matching.

The one-to-many problem is when a vehicle observed once at A and then at B and then again at B due to U turn further downstream of B. There are two travel time samples are matched and the second one does not reflect the true travel time along the corridor. The Figure 5 illustrates the many-to-one problem

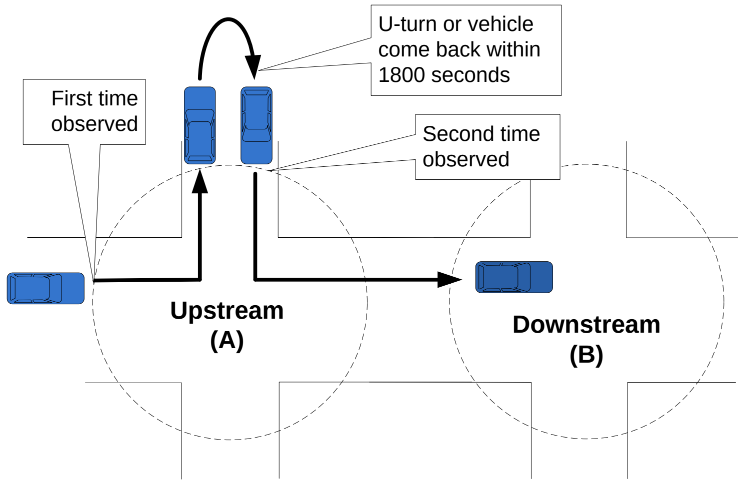

Figure 3: Many-to-one problem of Bluetooth matching.

Many-to-one: a vehicle is observed at A and then it makes a U turn between A and B and is observed again at A and it travels to B. Again the second travel time sample is a noise. Similarly, the many-to-many problem is also possible. The *Multiple Matching filter* removes all these noises by taking only the minimum travel time from all possible matches

- *Outlier filter*: This filter removes outliers by comparison with neighbour travel time observations within 10 mins interval. For each minute, a window of 5 mins before and 5 mins after is considered and this window is moved from the first to the last minute of the day. The outliers are identified if they are larger than a certain upper bound value or lower than a certain lower bound value of the current window.

For the outlier filter, two filtering techniques have been considered: the Box and Whisker (B&W) and the Median Absolute Deviation (MAD). The filter could be applied for each day, or for each travel pattern. The travel patterns are Working days (Monday to Friday), Holidays (Weekends and Public Holidays) and School Holidays (Work day but also School off day). Regarding the outlier filter, we have two candidates (B&W and MAD) with two filtering approach (apply filter for each day or for each travel pattern) to select only one for our analysis. In this section, each type of filtering for the outlier filter is discussed and the most effective filtering technique for our dataset is found.

DESCRIPTION OF BOX AND WHISKER AND MEDIAN ABSOLUTE DEVIATION TECHNIQUE

The Box and Whisker method is proposed by John Tukey in 1977 (Tukey, 1977).The outliers in the dataset are identified if they are larger than an upper bound value (UBV) or lower than a lower bound value (LBV). In the Box and Whisker technique, the UBV is calculated from an upper quartile (UQ) which is the median of the upper part of data or the 75th percentile. The LBV is calculated from a lower quartile (UQ) which is the median of the lower part of data or the 25th percentile. The gap between the UQ and LQ is Inter Quartile Range (IQR). It defines the noisiness of data. These values could be calculated by the following formulas:

$$UBV = UQ + 1.5 * IQR$$

$$LBV = LQ - 1.5 * IQR$$

$$IQR = UQ - LQ$$

The Median Absolute Deviation technique, also known as the Hampel Identifier, is the method which is described as very effective in practice (Pearson, 2002). A sample is considered as outlier if it is larger than a UBV or lower than a LBV. Basically, this method utilises the outlier-resistant median and median absolute deviation from the median (MAD) for identify the UBV and LBV.

$$UBV = median + \hat{\sigma} f$$

$$LBV = median - \hat{\sigma} f$$

Where $\hat{\sigma}$ is the standard deviation from the MAD, in which a normally distributed data can be approximated as $\hat{\sigma} = 1.4826 \times MAD$ and MAD is the Median of the absolute deviations from the data median.

$$MAD = median(|Xi - median(Xj)|)$$

Similar to the Box and Whisker method, the value of $\hat{\sigma} f$ defines the scatter of data, where *f* is a scale factor varies on a case by case basis. If *f* is small the gap between UBV and LBV to the median value is small and *vice versa*. The value of *f* has been suggested by some authors to be from 2 to 5 (Davies & Gather, 1993; Pearson, 2002). After testing the value of *f* from 1 to 5, the value *f=2* appears to be the optimal option for our dataset.

APPLY FILTERS ON EACH DAY

Firstly, the filters are applied on each day data. The Figure 4 shows the filtering results with B&W technique, where the black points are accepted travel time samples and green points are outliers. The Figure 5 shows the data cleansing results by using MAD technique with *f=2*. From the two figures, it could be noted that the MAD technique with *f =2* performs better in each day filtering. MAD could remove more noise than B&W method. However, a lot of noise could still be observed after applying the filters, especially in off-peak hours and in holidays. The days 5th (Work day), 18th (Holiday) and 19th (School holiday) have been used as examples. The reason for choosing these day are because 5th and 19th are both Mondays and day 18th is Sunday, so they are the best representatives of the Work days and Holidays.

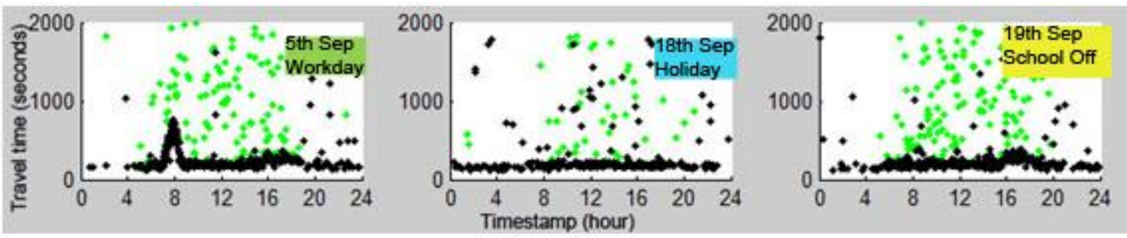

Figure 4: Filtering results by employing B&W technique on each day.

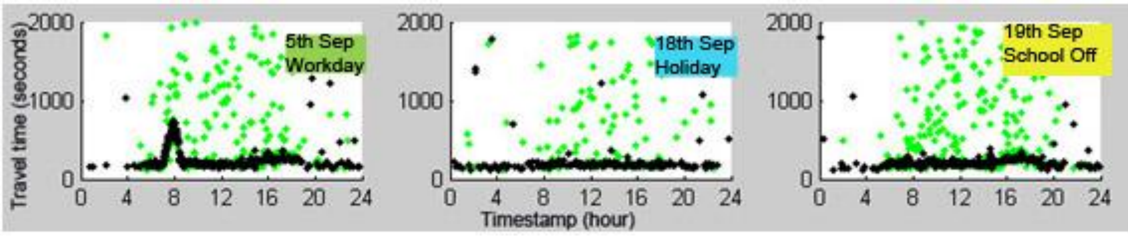

Figure 5: Filtering results by employing MAD technique on each day.

The small sample size when we apply the main filter could be the reason, because we applied the filter on each day data.

APPLY FILTERS ON EACH TRAVEL TIME PATTERN

The results of applying filters on each day are not promising. On the next step, the filters will be applied on each travel pattern (work day, holiday or school holiday). We assume that vehicles in the same travel time pattern would experience the same pattern of travel time. Hence, it is possible to merge them and apply the filter on the aggregated group of samples. By this approach, we apply the filters on a larger sample size. Our hypothesis is by utilising the historical travel time, we could have some advantages in identifying the outliers. If the travel time samples of the same travel patterns are similar to each other, it would bring us a better result than applying filters on each day. The Figure 6 shows the filtering results with B&W technique while the Figure 7 shows the same results by using MAD technique with *f=2*. The filters are applied on each travel pattern, but the results are plotted for each day for the purpose of comparison.

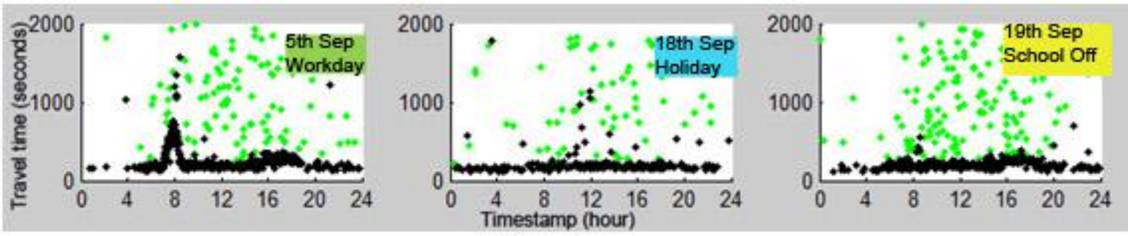

Figure 6: Filtering results by employing B&W technique on each travel pattern.

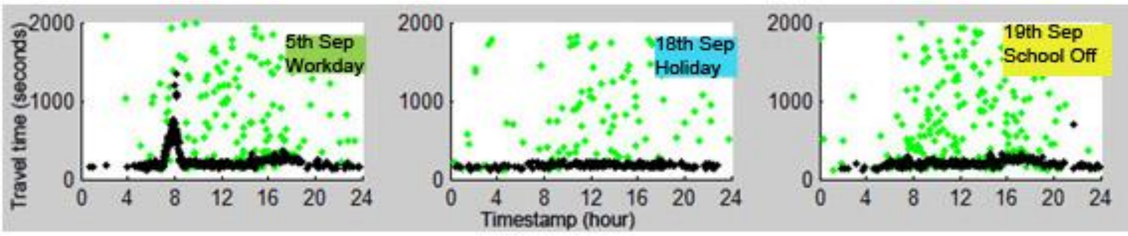

Figure 7: Filtering results by employing MAD technique on each travel pattern.

Compared to applying the filters on each day, this approach shows much less noise. The MAD technique still performs better than the B&W, especially in Holidays and School Holidays. The better filtering results show that our assumption of similar travel time pattern within the same travel pattern is correct. Hence, our final conclusion is selecting the MAD technique with *f=2* for cleansing Bluetooth data for each travel time pattern.

VID Data cleansing

As previously discussed, the number of VID samples is not large since VID sensors only identify the bus passing by. The sample size is not enough for us to apply B&W or MAD filter on it. Hence, after removing duplicated and uncompleted samples from the dataset, the *Unrealistic filter* and *Multiple matching filter* as mentioned earlier are applied. The other outliers are removed by inspection.

RELATIONSHIP BETWEEN BUS AND CAR TRAVEL TIME

This section aims to find the relationships between bus and car travel time. VID data includes both not-in-service and in-service bus data, hence here we have separately compared them with car travel time.

Not-in-service bus travel time and car travel time

As a not-in-service bus does not stop within the study area, it is predicted to operate similar to a car, except the differences in its operating capabilities as a heavy vehicle. For this section, our focus will be on Work days, as most of the buses are operating on these days. The analysis has been carried out for all 4.5 months data. The Figure 8 provides an example of car travel time (green dots) and not-in-service bus travel time (black stars) on all the Work days of September 2011.

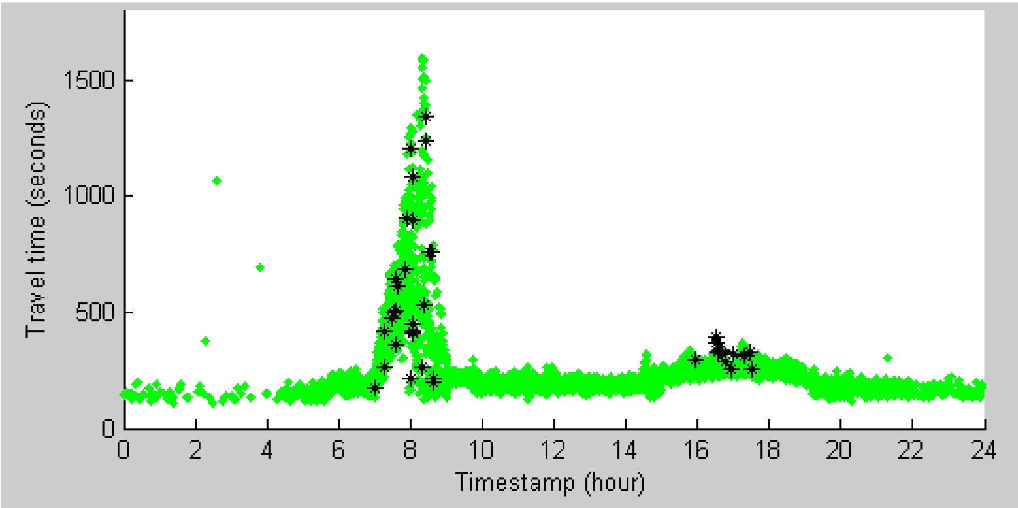

Figure 8: Car and not-in-service bus travel time of Work days September 2011.

Most of the not-in-service bus lies within the range of car travel time, which means the not-in-service bus travel time could be used to estimate the car travel time. For better understanding of the relationship between car and not-in-service bus travel time, the morning peak period and the afternoon peak period is studied separately in this section.

MORNING PEAK PERIOD (7 TO 9 AM)

The figure 9 shows the morning peak comparison of car and not-in-service bus travel time.

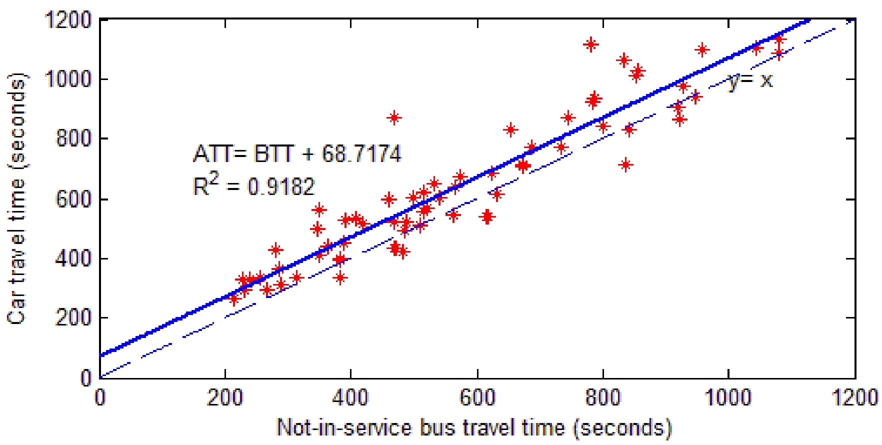

Figure 9: Comparison between car and not-in-service bus travel time during morning peak period.

As could be seen on Figure 9, the not-in-service bus travel time is very similar to the car travel time. Most of the comparison points are very close to the *y=x* lines. Formula (1) has been proposed for estimating the car travel time from not-in-service bus travel time. The formula (1) denotes that the not-in-service bus travel time is the car travel time added a parameter value.

$$ATT = \alpha + BTT \tag{1}$$

Where:

ATT = estimated car travel time

BTT = Observed not-in-service bus travel time

α = Parameters for calibrating

Linear regression analysis has been carried out on all the data to find the parameter α in formula (1). The result of linear regression is shown in Figure 9, in which the bold line is the model (1) and the dashed line is the line *y=x*. The model has a very good fit with R squared value is 0.9182. For further quantification of the estimation accuracy, the Mean Absolute Percent Error (MAPE) was calculated for the average car travel time and the estimated car travel time (from not-in-service bus travel time). The MAPE is calculated by the following formula:

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{ATT - (BTT + \alpha)}{ATT} \right|$$

Where:

ATT, BTT = travel time of car and not-in-service bus

n = total number of not-in-service bus samples.

The value of MAPE for the model (1) is 10.36 % which is an acceptable accuracy level. These results show that the not-in-service travel time and car travel time are similar and one could be used to see the travel time of the other during the morning peak period.

It could be noted from the model (1) that the α value is larger than zero, which means the in-service bus travel time is actually less than car travel time. The comparison of car average speed along the corridor and not-in-service bus average speed is shown in Figure 10.

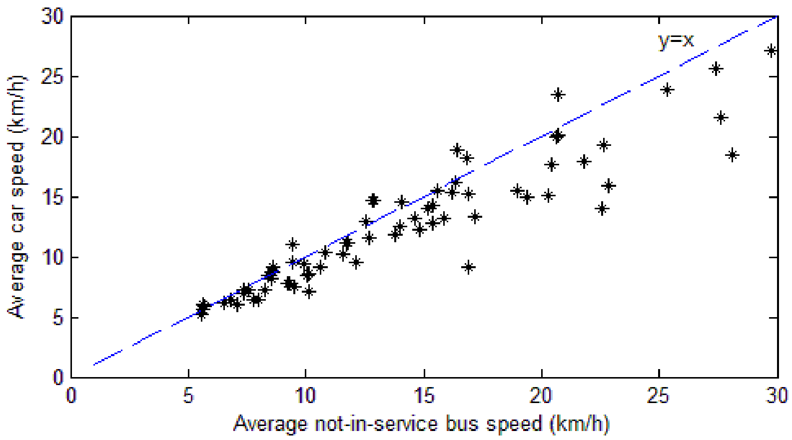

Figure 10: Comparison between car and not-in-service bus average speed during morning peak period.

During congested conditions the average not-in-service bus speed is very close to average car speed. The difference between the two speeds is significant during non congested conditions. Overall, the not-in-service bus has higher speed and less travel time along the corridor than car during morning peak period. The reason for this pattern is the temporal bus lane between 7 and 9 in the morning. The bus speed is increased due to the utility of the bus lane during morning peak hours.

AFTERNOON PEAK PERIOD (15 TO 19 PM)

The Figure 11 shows the comparison of car travel time with not-in-service bus travel time in the afternoon peak hours.

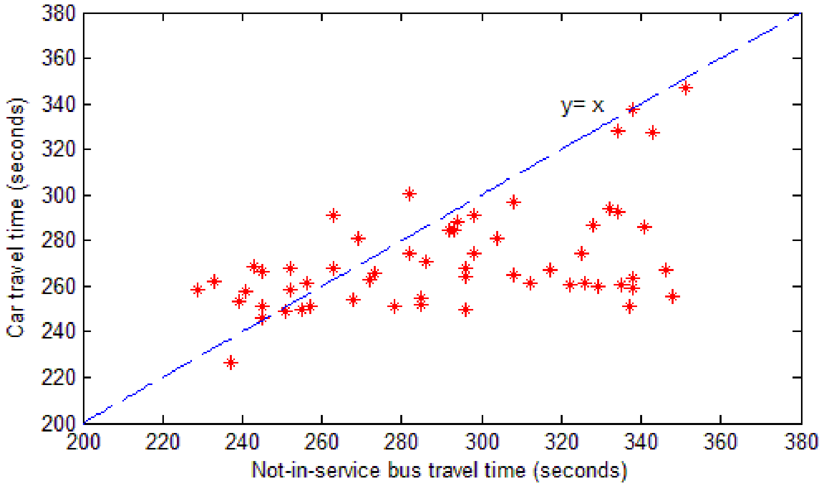

Figure 11: Comparison between car and not-in-service bus travel time during afternoon peak period.

During the afternoon peak period, because the bus lane is non-operational, not-in-service buses experience longer travel time than cars. The reasons for that are the difference in mechanical characteristic, with buses as heavy vehicles; and driver behaviour, where buses tend to stay strictly with the speed limit. However, the difference between them is relatively small and getting smaller as the travel time increased. We could conclude that during afternoon peak period the travel time of not-in-service buses are slightly higher than those of cars, even though not-in-service buses do not have to stop at bus stops.

The Figure 12 shows the full day comparison of car and not-in-service bus travel time. The morning and afternoon peak travel time pattern could clearly be seen on the graph.

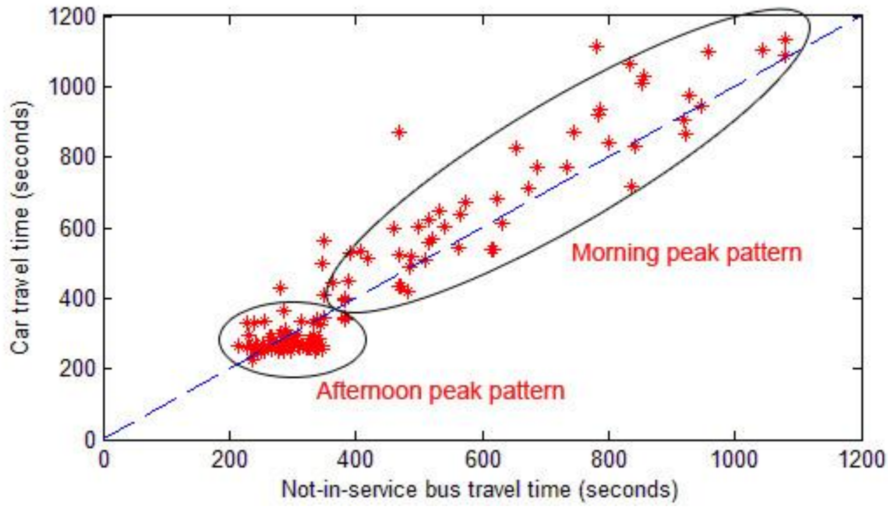

Figure 12: Comparison between car and not-in-service bus travel time

In-service bus travel time and car travel time

Along the study corridor there is only one servicing bus route. These in-service buses have to stop at bus stops for boarding/alighting of passengers. Because of this stopping time, they are predicted to experience longer travel time than the not-in-service buses and cars. There are 7 stops along the corridor where the bus might have to stop for dwelling. Figure 13 shows the car travel time (green dots) and in-service bus travel time (black dots) for an example of Work days of August 2011.

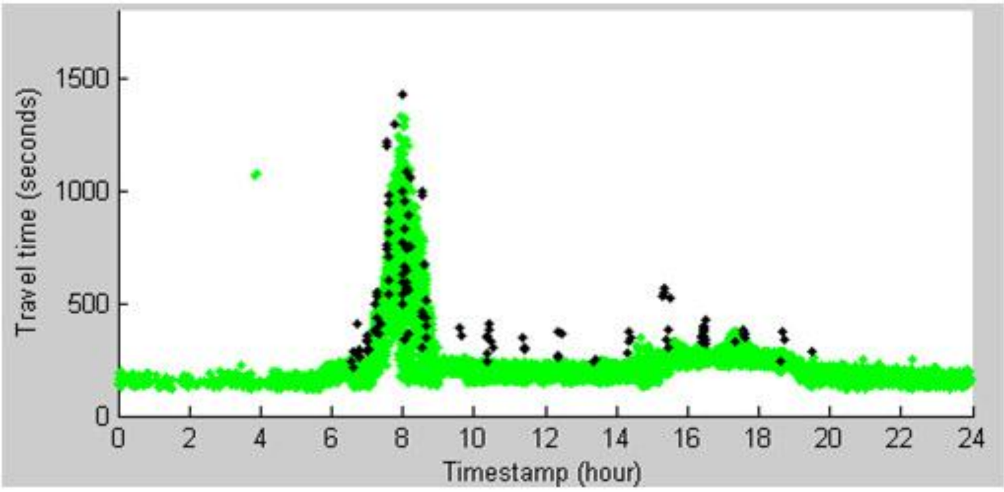

Figure 13: Car and in-service bus travel time of Work days August 2011.

In general, the in-service bus travel time is higher than the car travel time, especially after 9 AM. During the morning peak period, the in-service bus travel time does not show any significant difference to the car travel time. The reason for that is the temporal bus lane in the short section along the corridor, where only buses could use during 7 to 9 in the morning (see Figure 1).

In off-peak hours between 9 AM to 16 PM we could also find some interesting patterns. The Figure 14 illustrates the in-service bus travel time (dark blue boxes) and car travel time (green dots) in mid-day off-peak hours on our dataset of 4.5 months.

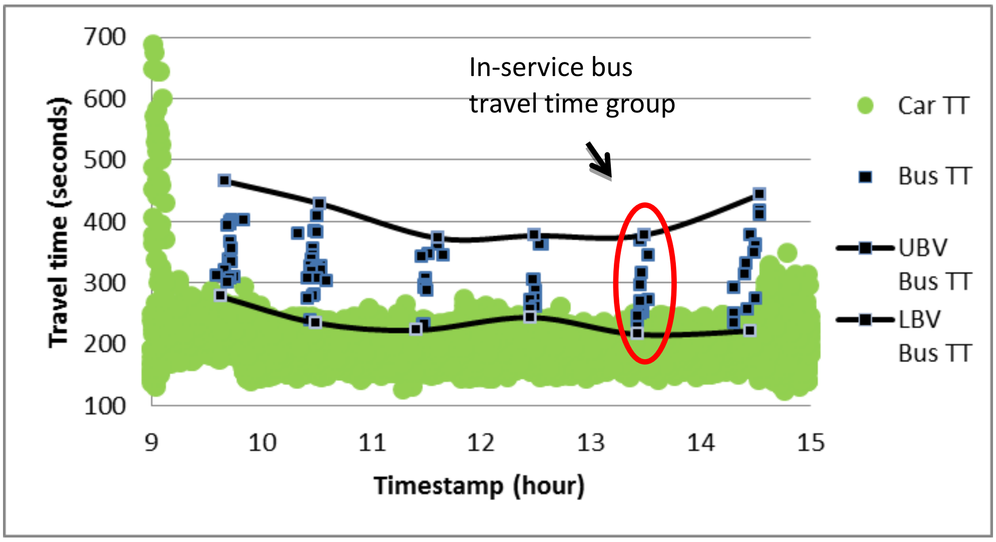

Figure 14: In-service bus travel time and car travel time during off-peak hours.

The in-service bus travel time samples (black boxes) forms 6 groups of travel time, in which the travel time samples line up in a nearly vertical line. The timestamp difference between each group is around 1 hour, similar to the headway of bus line 227 along the study corridor during off-peak hours. The reason for this pattern could be the fact that the in-service buses could stay with the schedule during off-peak periods. Because there is not much congestion at upstream intersection during off-peak periods, everyday a bus arrives at the same time at upstream location. As the bus travels from upstream to downstream intersection, the travel time could be varied, but the arrival time is almost the same. That is how these groups are formed.

The black lines draw the lower bound (LBV) and upper bound (UBV) of the in-service bus travel time. As could be seen on the Figure 14, the in-service bus travel time and car travel time form two bands of travel time, in which the LBV of in-service bus travel time are very close to the UBV of car travel time. These fastest in-service buses could be the buses that did not have to stop at bus stops along the corridor. The next two sections discuss the findings from the patterns of travel time on Figure 14.

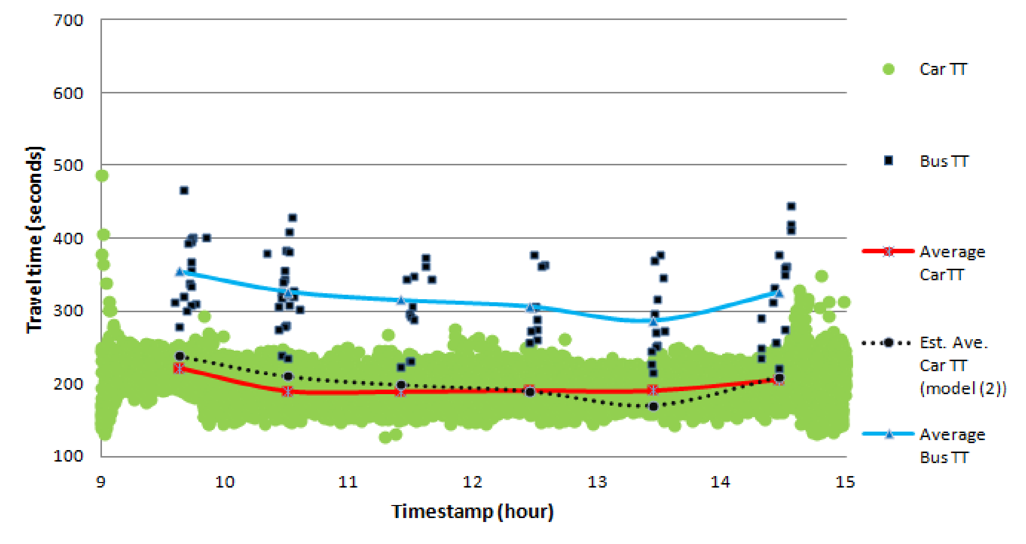

Figure 15: Using not-in-service bus travel for estimating car travel time.

The in-service bus travel time could be used to briefly estimate the car travel time during the mid-day off-peak hours. As could be seen on the Figure 15, from 9:30 to 14:30 bus travel time has similar pattern to car travel time in the sense that the travel time is higher when the timestamp is close to peak hours and lower at the middle of the mid-day off-peak period. Hence, if we take the average value of each in-service bus travel time group, we can draw a line similar to the average of car travel time. Moreover, this line could be used to estimate the average car travel time. This pattern could not be used for the peak period case, where heavily congestion varies the arrival time to the upstream intersection of bus, and no travel time group is formed.

In the Figure 15, the dotted black line is our estimated average of car travel time from in-service bus travel time, which is actually derived from the average values of the travel time groups by the estimated formula (2). The estimation was carried out through inspection.

$$ECTT = ABTT - 116 \text{ seconds} \tag{2}$$

Where:

ECTT = Estimated car travel time at each group's time period (seconds)

ABTT = Average values of in-service bus travel time groups (seconds).

The red line shows the 1-hour average of car travel time. As could be seen on the Figure 15, the average car travel time line (red line) and the estimated average car travel time line (dotted line) are very close to each other. Table 3 compares the observed average car travel time with the estimated average car travel time on each time points where bus travel time is available. The absolute deviation values are always less than 21 seconds and the percentage differences are less than 12 precents. The MAPE of the equation (2) is 6.11 % which means the estimated average car travel time is very close to the actual average car travel time.

Table 3: Comparison between average car travel time and its estimated value

Timestamp (hour)	Observed Average Car travel time (s)	Estimated Average Car travel time (s)	Absolute deviation	Percentage Difference (%)
9.623611	221.4423	239	17.5577	7.63
10.49389	190.5764	210.8095	20.23312	10.08
11.40861	189.7103	199.1538	9.443546	4.86
12.45278	191.0328	190.3	0.7328	0.38
13.43611	190.7282	170.5833	20.14487	11.15
14.46167	205.6169	210.1333	4.516433	2.17

USING IN-SERVICE BUS TRAVEL TIME FOR BRIEF ESTIMATION OF BUS DWELL TIME AND VARIANCE OF BUS ARRIVAL TIME

The *vertical dimension* of each in-service bus travel time group could give us an insight into the approximate total dwell time along the corridor. If we assume that the fastest in-service buses (lowest travel time) are the buses which did not have to stop, the highest bus travel time samples could be the ones which have to stop at most of the bus stops. Hence the vertical spread of each group is the total dwell time variance at that time period. These variances are illustrated in the Figure 16.

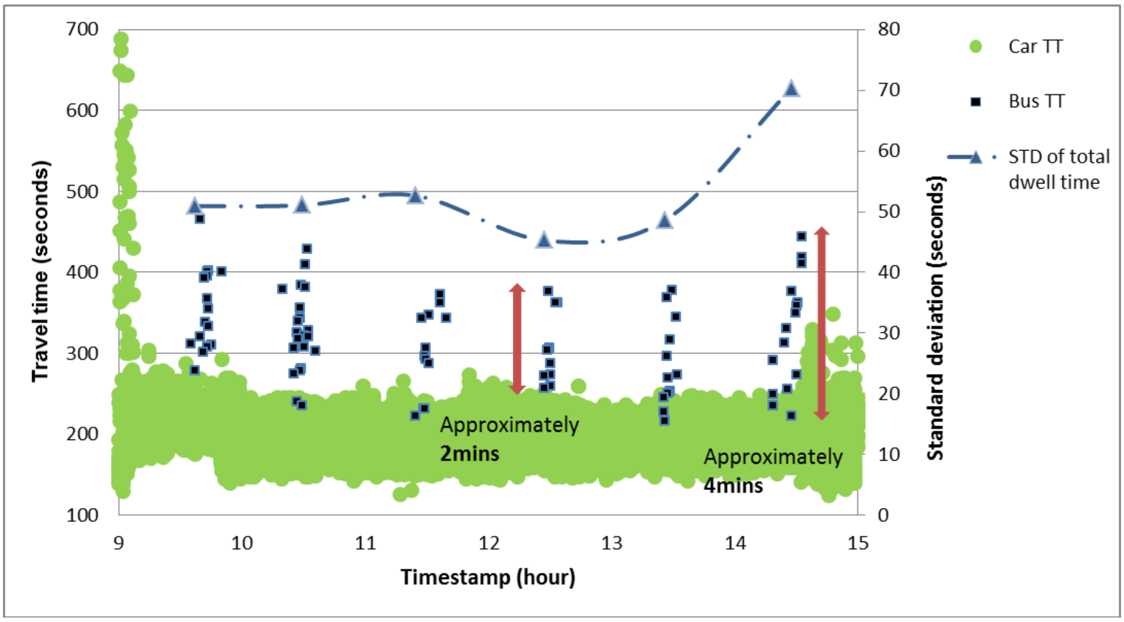

Figure 16: Variance of total dwell time (vertical spread of each travel time group).

As could be seen on the Figure 16, the maximum dwell time varies from around 2 minutes to 4 minutes during mid-day off-peak period. The maximum dwell time is smaller at the middle of the off-peak period, due to the small passenger and traffic demand and larger at timestamps which are close to the peak hours. Similarly, the standard deviation of the total dwell time is lowest during the noon period (12-14 PM) and highest when it is close to the afternoon peak period (14-15 PM).

As there is only one bus route is servicing along the study corridor, the schedule of all the in-service buses are the same. If all the buses stayed strictly to the schedule, we would see a straight vertical line at each travel time group. Because the arrival times to the entrance of the study corridor (upstream intersection) are varied due to congestion and dwell time variability, the travel time groups are not straight lines and they have some horizontal spreads. Thus, the *horizontal dimension* of each group shows us the variance of arrival times to the intersection. The Figure 17 illustrates the variances of in-service bus arrival times to the intersection. The maximum variance is around 15 minutes at timestamps which are close to the peak hours. At the middle of the off-peak period the arrival time variance is only around 5 minutes. Similar to the variance of total dwell time, the standard deviation of bus arrival time to the intersection is lowest during the noon period (12-14 PM) and highest when it is close to the afternoon peak hours (14-15 PM).

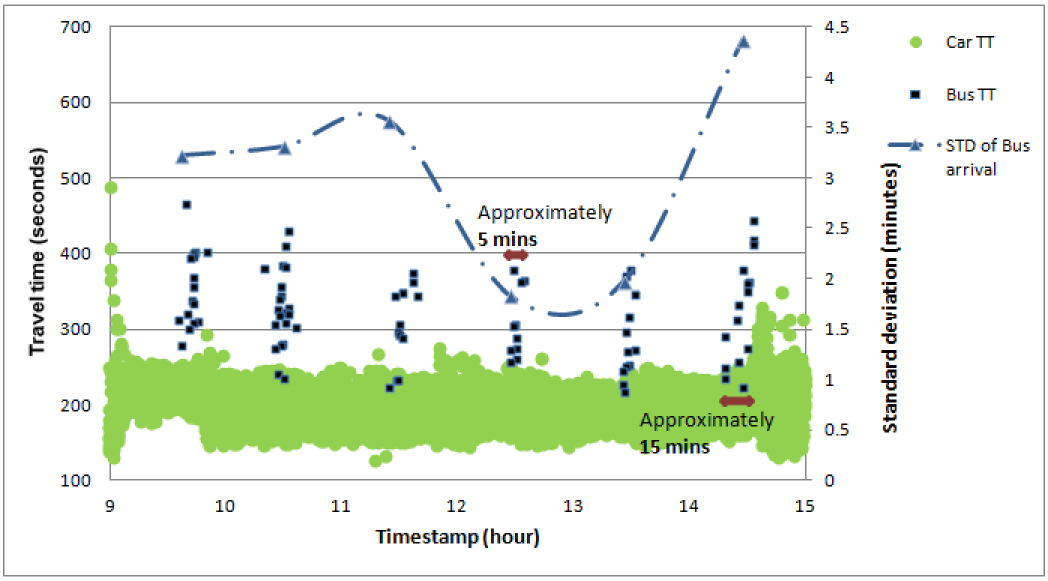

Figure 17: Variance of arrival times to the intersection (horizontal spread of each travel time group).

CONCLUSION

Buses and cars have different operating and mechanical characteristics. The Bluetooth and Bus Vehicle Identification data have been explored in this paper to develop a relationship which helps us to understand their operations and could be used to estimate the car travel time from bus travel time. The modelling results show that the not-in-service bus and car travel time are comparable. The car travel time is slightly higher than the not-in-service bus due to the utility of the bus lane during morning peak period. Whereas, during the afternoon peak period the car travel time is lower than the not-in-service bus travel time due to the difference in mechanical properties and driver behaviours of the two modes.

It is observed that the in-service bus travel time could be used to estimate the average of car travel time during off-peak hours. The vertical and horizontal spreads of in-service bus travel time during off-peak period provides an insight into the maximum total dwell time along the corridor and the variances of arrival times to the entrance of the study corridor (upstream intersection), respectively.

This paper explores the relationship between bus and car travel time and the results are very promising. We are currently analysing the whole Brisbane network to generalise our findings.

ACKNOWLEDGEMENTS

The Bluetooth and VID data used in this study was provided by the Brisbane City Council. This research contributes to the Integrated Traveller Information Research Domain at the QUT-Smart Transport Research Centre. The conclusions in this paper reflect the understandings of the authors, who responsible for the accuracy of the data.

REFERENCES

Abdelfattah, A. M., & Khan, A. M. (1998). Models for predicting bus delays *Transit: Bus, Paratransit, Rural, Intermodal, Rail, Commuter and Intercity Rail, Light Rail* (pp. 8-15).

Ahmed, H., EL-Darieby, M., Abdulhai, B., & Morgan, Y. (2008). *Bluetooth-and Wi-Fi-based mesh network platform for traffic monitoring.* Paper presented at the Transportation Research Board 87th Annual Meeting, Washington DC, U.S.

Baek, S., Lim, Y., Rhee, S., & Choi, K. (2010). Method for estimating population OD matrix based on probe vehicles. *KSCE Journal of Civil Engineering, 14*(2), 231-235. doi: 10.1007/s12205-010-0231-4

Barceló, J., Montero, L., Marquès, L., & Carmona, C. (2010). Travel time forecasting and dynamic origin-destination estimation for freeways based on bluetooth traffic monitoring.

Transportation Research Record: Journal of the Transportation Research Board, 2175(-1), 19-27.

Bhaskar, A., Chung, E., & Dumont, A.-G. (2010). Analysis for the Use of Cumulative Plots for Travel Time Estimation on Signalized Network. *International Journal of Intelligent Transportation Systems Research, 8*(3), 151-163.

Bhaskar, A., Chung, E., & Dumont, A.-G. (2011). Fusing Loop Detector and Probe Vehicle Data to Estimate Travel Time Statistics on Signalized Urban Networks. *Computer-Aided Civil and Infrastructure Engineering, 26*(6), 433-450. doi: 10.1111/j.1467-8667.2010.00697.x

Bhaskar, A., Chung, E., & Dumont, A. G. (2012, January 22-26, 2012). *Urban Route Average Travel Time Estimation Considering Exit Turning Movements.* Paper presented at the Transportation Research Board 91st Annual Meeting, Washington, D.C.

Blogg, M., Semler, C., Hingorani, M., & Troutbeck, R. (2010). *Travel Time and Origin-Destination Data Collection using Bluetooth MAC Address Readers.* Paper presented at the Australasian Transport Research Forum 2010, Canberra, Australia.

Blythe, P. (1999, 1999). *RFID for road tolling, road-use pricing and vehicle access control.* Paper presented at the IEE Colloquium on RFID Technology (Ref. No. 1999/123).

Chen, M., Liu, X., Xia, J., & Chien, S. (2004). A Dynamic Bus-Arrival Time Prediction Model Based on APC Data. *Computer-Aided Civil and Infrastructure Engineering, 19*(5), 364-376. doi: 10.1111/j.1467-8667.2004.00363.x

Chien, S. I. J., Ding, Y. Q., & Wei, C. H. (2002). Dynamic bus arrival time prediction with artificial neural networks. [Article]. *Journal of Transportation Engineering-Asce, 128*(5), 429-438. doi: 10.1061/(asce)0733-947x(2002)128:5(429)

Chon, H. D., Jun, S., Jung, H., & An, S. W. (2004). Using RFID for accurate positioning. *Journal of Global Positioning Systems, 3*(1-2), 32-39.

Davies, L., & Gather, U. (1993). The identification of multiple outliers. *Journal of the American Statistical Association*, 782-792.

Finkenzeller, D. K., & Muller, D. (2010). *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards, Radio Frequency Identification and Near-Field Communication, Third Edition*: John Wiley & Sons.

Haghani, A., Hamedi, M., Sadabadi, K. F., Young, S., & Tarnoff, P. (2010). Data collection of freeway travel time ground truth with bluetooth sensors. *Transportation Research Record: Journal of the Transportation Research Board, 2160*(-1), 60-68.

Jeong, R., & Rilett, L. (2005). Prediction model of bus arrival time for real-time applications *Transit: Planning, Management and Maintenance, Technology, Marketing and Fare Policy, and Capacity and Qualtiy of Sevice* (Vol. 1927, pp. 195-204). Washington: Transportation Research Board Natl Research Council.

Kalaputapu, R., & Demetsky, M. J. (1995). *Modeling schedule deviations of buses using automatic vehicle location data and artificial neural networks*: Transportation Research Board.

Lee, E.-K., Oh, S. Y., & Gerla, M. (2012). RFID assisted vehicle positioning in VANETs. *Pervasive and Mobile Computing, 8*(2), 167-179. doi: 10.1016/j.pmcj.2011.06.001

Levinson, H. (1983). Analyzing transit travel time performance. *Transportation Research Record*(915).

Malinovskiy, Y., Lee, U. K., Wu, Y. J., & Wang, Y. (2011). *Investigation of Bluetooth-Based Travel Time Estimation Error on a Short Corridor.* Paper presented at the 90th Annual Transportation Research Board Meeting, Washinton DC, U.S.

Malinovskiy, Y., Saunier, N., & Wang, Y. (2012). *Pedestrian Travel Analysis Using Static Bluetooth Sensors.* Paper presented at the 91st Annual Transportation Research Board Meeting, Washington D.C.,U.S.

Pearson, R. K. (2002). Outliers in process modeling and identification. *Control Systems Technology, IEEE Transactions on, 10*(1), 55-63. doi: 10.1109/87.974338

Quayle, S., Koonce, P., DePencier, D., & Bullock, D. (2010). Freeway Arterial Performance Measures Using MAC Readers: Portland Pilot Study. *Transportation Research Record: Journal of the Transportation Research Board*(2192), pp 185-193.

Seo, G., Yazici, A., Ozguner, U., & Cho, J. (2008, 17-20 Feb. 2008). *An approach for data collection and Traffic Signal Control in the futuristic city.* Paper presented at the Advanced Communication Technology, 2008. ICACT 2008. 10th International Conference on.

Shalaby, A., & Farhan, A. (2004). *Bus Travel Time Prediction Model for Dynamic Operations Control and Passenger Information Systems*. Paper presented at the TRB 2003 Annual Meeting CD-ROM, Washington D.C., 2003.

Sharifi, E., Hamedi, M., & Haghani, A. (2010). *Vehicle detection rate for bluetooth travel time sensors: a case study in Maryland and Delaware.* Paper presented at the 91st Annual Transportation Research Board Meeting, Washinton DC, U.S.

Sriborrirux, W., Danklang, P., & Indra-Payoong, N. (2008). *The design of RFID sensor network for bus fleet monitoring.* Paper presented at the 8th International Conference on ITS Telecommunications, 2008.

Stockman, H. (1948). Communication by Means of Reflected Power. *Proceedings of the IRE*, 1196-1204. doi: citeulike-article-id:169507

Swedberg, C. (2004). RFID drives highway traffic reports. *RFID Journal,[Online]. Available: http://www.rfidjournal.com/article/articleview/1243/1/1*.

Tarnoff, P. J., Bullock, D. M., Young, S. E., Wasson, J., Ganig, N., & Sturdevant, J. R. (2009). *Continuing evolution of travel time data information collection and processing*.

Tsubota, T., Bhaskar, A., Chung, E., & Billot, R. (2011). *Arterial traffic congestion analysis using Bluetooth duration data*. Paper presented at the Australasian Transport Research Forum 2011. http://eprints.qut.edu.au/46312/

Tukey, J. W. (1977). Exploratory data analysis. *Reading, MA*.

Wasson, J. S., Sturdevant, J. R., & Bullock, D. M. (2008). Real-time travel time estimates using media access control address matching. *ITE Journal, 78*(6).