Abedi, Naeim, Bhaskar, Ashish, Chung, Edward, & Miska, Marc
(2015)
Assessment of antenna characteristic effects on pedestrian and cyclists travel-time estimation based on Bluetooth and WiFi MAC addresses.
Transportation Research Part C: Emerging Technologies, *60*, pp. 124-141.

https://doi.org/10.1016/j.trc.2015.08.010

Assessment of Antenna Characteristic Effects on Pedestrian and Cyclists Travel-time Estimation based on Bluetooth and WiFi MAC Addresses

By

Naeim Abedi
Smart Transport Research Centre, School of Civil Engineering and Build Environment,
Science and Engineering Faculty, Queensland University of Technology (QUT),
Brisbane QLD 4001, Australia
e-mail: naeim.abedi@connect.qut.edu.au

Ashish Bhaskar*
Smart Transport Research Centre, School of Civil Engineering and Build Environment,
Science and Engineering Faculty, Queensland University of Technology (QUT),
Brisbane QLD 4001, Australia
e-mail: ashish.bhaskar@qut.edu.au

Edward Chung
Smart Transport Research Centre, School of Civil Engineering and Build Environment,
Science and Engineering Faculty, Queensland University of Technology (QUT),
Brisbane QLD 4001, Australia
e-mail: edward.chung@qut.edu.au

and
Marc Miska
Smart Transport Research Centre, School of Civil Engineering and Build Environment,
Science and Engineering Faculty, Queensland University of Technology (QUT),
Brisbane QLD 4001, Australia
e-mail: marc.miska@qut.edu.au

*corresponding author

ABSTRACT

Monitoring pedestrian and cyclists movement is an important area of research in transport, crowd safety, urban design and human behaviour assessment areas. Media Access Control (MAC) address data has been recently used as potential information for extracting features from people's movement. MAC addresses are unique identifiers of WiFi and Bluetooth wireless technologies in smart electronics devices such as mobile phones, laptops and tablets. The unique number of each WiFi and Bluetooth MAC address can be captured and stored by MAC address scanners. MAC addresses data in fact allows for unannounced, non-participatory, and tracking of people. The use of MAC data for tracking people has been focused recently for applying in mass events, shopping centres, airports, train stations etc. In terms of travel time estimation, setting up a scanner with a big value of antenna's gain is usually recommended for highways and main roads to track vehicle's movements, whereas big gains can have some drawbacks in case of pedestrian and cyclists. Pedestrian and cyclists mainly move in built distinctions and city pathways where there is significant noises from other fixed WiFi and Bluetooth. Big antenna's gains will cover wide areas that results in scanning more samples from pedestrians and cyclists' MAC device. However, anomalies (such fixed devices) may be captured that increase the complexity and processing time of data analysis. On the other hand, small gain antennas will have lesser anomalies in the data but at the cost of lower overall sample size of pedestrian and cyclist's data. This paper studies the effect of antenna characteristics on MAC address data in terms of travel-time estimation for pedestrians and cyclists. The results of the empirical case study compare the effects of small and big antenna gains in order to suggest optimal set up for increasing the accuracy of pedestrians and cyclists' travel-time estimation.

1. INTRODUCTION

Studying spatio-temporal movement of human has been recently focused especially in terms of crowd congestion control, safety, public transport and movement behaviour assessment. Various movement sensors have been developed by robust passive and active positioning technologies for capturing human's movement dynamics. The analysis of people movement's dynamic has received attention particularly in the field of visual analytics (Andrienko and Andrienko, 2007a). The study of big volumes of trajectory information of objects moving through geographical space has become a major subject of notice in research fields such as geographical information science (Ahlqvist et al., 2010, Shaw et al., 2008), computer science (Bogorny et al., 2009), urban evacuation (Nassir et al., 2013, Nassir et al., 2014), visual analytics (Andrienko and Andrienko, 2007b) and urbanism (Van Schaick and Van der Spek, 2008). However, the greater part of research has been applied to people mobility in different contexts and at various scales. For instance, the movement of athletes on a pitch (Laube et al., 2005), tourists on a regional (Ahas et al., 2008) and local scale (Kemperman et al., 2009, O'Connor et al., 2005, Shoval and Isaacson, 2007), and customers in a supermarket (Hui et al., 2009).

Interests in dynamic movement modeling of pedestrians and cyclists are increasing because of the pressure of urban growth on city infrastructure (Bierlaire and Robin, 2009, Duives et al., 2013, Kasemsuppakorn and Karimi, 2013, Kneidl et al., 2013, Weidmann et al., 2014). This has increased the demand for developing information and simulation tools in order to design new urban infrastructures as well as improvement of current urban foundations. Capturing movement data from pedestrians and cyclists plays a key role to model their travel behavior and habits especially for enhancement of urban transport systems. While various range of information acquisition methods have been introduced, each method is associated to noticeable issues such as precision, privacy, maintenance costs etc.

Surveys and video processing have been used as popular methods for recording information from people. Traditional survey has its limitations to the sample size, non-random sampling and excessive cost. Advanced video surveillance has a better capture rate but it's automatic data acquisition is highly sensitive to the weather conditions, viewing angles, illumination changes, density and brightness of crowd (Liebig and Wagoum, 2012). Video processing also requires considerable process time and complex algorithms in order to reconstruct individual movements across multiple camera angles. These drawbacks restricts video surveillance methods to capture the spatio-temporal paths of only limited objects in few spatial spaces (Dee and Velastin, 2008). Positioning the cell-phones based on Global System for Mobile (GSM) communication has also been explored to monitor people's movements. However, it has become less applicable especially due to the privacy objection concerns and large error range (for civilian use)(Giannotti and Pedreschi, 2008).

In response to the mentioned issues and given the ubiquity of WiFi and Bluetooth-enabled devices such as smart phones and tablets carried around by their owners, WiFi and Bluetooth technologies have increasingly attracted significant attention as a low-cost alternative for the reconstruction of spatial behaviour (Bullock et al., 2010, Wasson et al., 2008, Versichele et al., 2010, Mottram, 2007, Van LonderseLe et al., 2009) for various applications such as direct measurement of pedestrian and cyclist's travel time (Malinovskiy et al., 2012), space utilisation behaviour (Abedi et al., 2014) and location's popularity evaluation (Vu et al., 2010). Also, tracking individual in this method remains anonymous avoiding potential privacy infringements because each fixed Media Access Control (MAC) address cannot be associated to any personal information such as names or mobile numbers. Bluetooth and WiFi MAC address data are also increasingly being used for road traffic monitoring and management (Tsubota et al., 2011, Nantes et al., 2014, Bhaskar et al., 2014a, Bhaskar et al., 2014b, Bhaskar et al., 2015, Abbott-Jard et al., 2013). However, the major weakness of MAC address data is that its sample size may not represent the real sample number because there is the possibility of carrying more than one active WiFi and Bluetooth devices by a traveller and not all travellers will be carrying active WiFi and Bluetooth devices.

Antenna characteristic is a physical element that significantly impacts on the data range and accuracy of MAC address based movement tracking. Basically, higher gains of antenna provide wider scanning ranges. For travel-time estimation applications, setting up MAC address scanners equipped to an antenna with big gain is usually suggested for highways and main roads to track vehicle's movements. However, limited research has been done in order to offer optimal gain of antenna for travel-time estimation of pedestrians and cyclists. Pedestrian and cyclists mainly move in built-up districts and city pathways where plenty of fixed WiFi and Bluetooth devices may operate. Unlike vehicle transportation, people may travel in smaller scales with various speeds as a walker, runner or cyclist. Hence, the size of scanning area can significantly impact on the data range and capturing accuracy. This paper aims to investigate the effects of antenna's gain on the accuracy of collecting movement data from pedestrians and cyclists.

This study evaluates the results of different gains of antenna to the real-data in order to suggest an optimal set up for enhancing the performance and accuracy of MAC address data in terms of pedestrians and cyclists travel-time estimation. A case study based on experimental tests and scenarios has been carried out over a bridge allocated only to pedestrian and cyclists. The result of this study evaluates the advantages and drawbacks of antenna's gains in terms of capturing more relevant and less anomaly samples. The results of this research can be applied in the application of pedestrians and cyclists travel-time estimation for optimal and efficient data collection, decreasing processing time and enhancing tracking accuracy.

The rest of the paper is structured as follows. *Section 2* first presents the recent studies done on the analysis of human's movement behaviour and thereafter outlines the MAC address dataset as a technology for tracking people's movement. *Section 3* describes the details of the experiment and pre-processing on the data. *Section 4* presents the results of the analysis performed on the case study. Finally, the paper is concludes with the discussion on the importance of antenna characteristics on the accuracy of MAC address data set.

2. MONITORING PEDESTRIAN AND CYCLISTS

2.1. Human Movement Behaviour

Monitoring, simulating and predicting human's dynamic patterns of movement through space is becoming an increasingly important target of urban and transport planners interested in designing effective urban spaces for pedestrians (Batty, 2003). It is also an interesting area for studying and understanding human behaviour in terms of moving through pedestrian pathway environments such as corridors, urban and bridge pathways. However, such research and pattern extraction are complex due to a large number of variables related to pedestrian, situations and environments.

This section presents some insight relating to various elements of human movement behaviour in urban spaces. The most fundamentals include walking speed and various distances that people choose to maintain between themselves and other entities around such as obstacle, building, kerbs etc. The walking speed of pedestrians in urban spaces varies between 1 and 1.5 *m/s* (Polus et al., 1983, Virkler, 1998). Various factors may explain this walking speed variation. Personal factors such as gender and age significantly effect on walking speed (Boles, 1981, Knoblauch et al., 1996, Fugger et al., 2000). For instance, males walk faster than females and increasing age declines the speed (Bowman and Vecellio, 1994, Coffin and Morrall, 1995). Density of pedestrians also significantly effects on walking speed as demonstrated in fundamental speed-flow relationship (Fruin, 1992, Henderson, 1971, Abedi, 2014). Other situational factors such as level of mobility and group size play a role (Boles, 1981, Knoblauch et al., 1996) but such factors have not received much attention in the literature.

Environmental factor can also influence spontaneous walking speeds. Temperature affects people moving speed (Rotton et al., 1990). People moves more quickly when crossing roads (Lam et al., 1995). Overall function of pedestrian area such as shopping leisure, transport interchange, school route and business districts presumably varies pedestrian walking speed due to the differing priorities and targets of the people who populate them.

Studying the space preferences of pedestrians in urban and indoor spaces have essentially focused on establishing various levels of service criteria involving to pedestrian traffic in crowded or potentially crowded areas (Fruin, 1992, Pushkarev, 1975). Research suggested that people prefer to keep a buffer zone of approximately 0.45 *m* between themselves and buildings' edges (Ciolek, 1978, Fruin, 1992), and a larger distance of around 0.85 *m* between themselves and other pedestrians (Dabbs Jr and Stokes, 1975). Individuals also prefer to maintain the distance of around 0.1 *m* from stationary items of street equipment (Habicht and Braaksma, 1984). One research also reported that people like to stay around

0.75 *m* far as their companion(s) when walking (Burgess, 1983). However, most of these finding have remained actually uncorroborated (Kwon et al., 1998) as well as the influence of personal and environmental factors on these spacing behaviour. Nevertheless, these preliminary finding can be useful for designing of high-volume pedestrian facilities.

Understanding of human crowds during evacuations and panic conditions were researched since the 1930s (Kholshevnikov and Samoshin, 2008). However, there is limited understanding on the behaviour of panicking groups and its impacts on the safety under emergency situations (Helbing et al., 2000). The development of mathematical simulation models based on the collective movements of animals has been done since the 1970s (Okubo, 1986). In terms of studying human movement behaviour in the panic conditions such as emergency evacuations, some studies have been recently done to develop evacuation and crowd control models based on assessing animal dynamics. Shiwakoti et al. (2011) derived a mathematical model for crowd panic based on collective animal dynamics. They developed and validated their model by data experimented with panicking Argentine ants.

In terms of crowd congestion study, Hoogendoorn and Daamen (2005) studied the microscopic pedestrian walking behaviour through wide and narrow bottlenecks. Basically, pedestrian form layers or trails inside bottleneck. The distance between pedestrians is measured approximately 45 *cm* which is less than effective width of a single pedestrian. This is called the phenomenon of "*zipper*" which corresponds overlapping of layers. Their finding shows that the phenomenon of "*zipper*" effect causes the capacity of the bottleneck to increase in a stepwise fashion with the width of the bottleneck. They found that two layers are formed in the narrow bottleneck (with of one meter), whereas four or five layers are formed for the wide bottlenecks (width of two meters). Wang et al. (2014) also presented a microscopic model from pedestrians' movement behaviour in terms of interacting visual attractors. In case of route choice behaviour study of pedestrians, Asano et al. (2010) developed a microscopic pedestrian simulation model combined with a tactical model. Zeng et al. (2014) developed a simulation model for analysis of pedestrian behaviour at signalised crosswalk.

Emerging technologies has increased the ability of extracting more valuable information from human's movement behaviour. Next section presents the capability of MAC address data as an emerging technology for tracking individuals' movement through spaces.

2.2. MAC Address as Movement Data

In terms of accessing to networks and data services with higher flexibility and mobility, wireless telecommunication networks are a widespread and fast-growing technology (Hossain and Wee-Seng, 2007). The advantages of wireless technologies are reducing the cable restrictions, easy deployment, low cost and dynamic communication formation. Bluetooth, WiFi, ZigBee, and UWB are four short-range wireless standards known as *IEEE 802.15.1*, *802.11 a/b/g*, *802.15.4,* and *802.15.3*, respectively. *IEEE* defines the MAC address and Physical Layers for mentioned wireless methods for an operation range of 10 to 100 *m* (Porter et al., 2012). Nowadays, majority of smart-phones and digital mobile devices use Bluetooth and WiFi technologies for data exchange and Internet access.

MAC addresses are unique identifiers and are used for various type of telecommunication networks and most of *IEEE 802* wireless technologies. Hence, they can be traced and this feature has motivated researches for various applications and data collection. Several factors associated with the hardware and software implemented may impact on the quality of MAC address data acquisition process (Bhaskar and

Chung, 2013). MAC address discovery time and antenna characteristics are important factors in terms of collecting efficient data during a time period. Bluetooth discovery time is theoretically 10.21 *sec* (Han and Srinivasan, 2012), whereas WiFi discovery time is almost 1 *sec* (Chakraborty et al., 2010).

The use of Bluetooth Media-Access-Control Scanner (BMS) has received significant interest from researchers and practitioners (Bhaskar and Chung, 2013) as complementary transport data. Time-synchronized BMSs positioned on motorways and road networks has the potential to provide live monitoring of vehicles' travel-time, assuming Bluetooth enabled-devices are transported by vehicles. This approach is one of the most cost efficient methods of travel-time estimation on the main roads. In case of signalized urban arterials, where travel-time estimation has always been very challenging with limited research (Bhaskar, 2009), BMS devices provide a well estimation from overall vehicle travel-time. Travel time from traditional matching of Bluetooth as ground truth travel-time can be considered for validating other travel time estimation models and forecasting future travel time values (Barceló et al., 2010). Bhaskar et al., (2014a) have also developed algorithm to estimate trajectory of the Bluetooth equipped vehicles on motorways. These trajectories provide detailed statistics of travel time between any two points on the network between the BMS scanner locations. Other applications of BMS data in transportation include the assessment of work zone effects, traffic congestion analysis (Tsubota et al., 2014, Nantes et al., 2015), route choice analysis (Xia et al., 2011), and multimodal travel time analysis (Kieu et al., 2015).

The success of BMS for road network monitoring has further attracted attention of exploring WiFi Media-Access-Control Scanner (WMS) (Abbott-Jard et al., 2013) as a complementary or alternative data source.

2.3. Related Works

Analysis of massive distributed movement data has been recently presented by new technologies as the popularity of using mobile devices has been increased (Jankowski et al., 2010, Andrienko and Andrienko, 2007a). Tracking mobile-devices and intercoms has motivated researches and scientist to collect movement information from individuals (Liebig and Wagoum, 2012, Stange et al., 2011). Recent research has been focused on the analysis of individuals' travelling behaviour in various applications such as the tourism industry (Jankowski et al., 2010), public transport utilisation in Graz (Weinzerl and Hagemann, 2007), movement behaviour assessment in shared areas (Abedi, 2014) and shopping malls (Millonig and Gartner, 2008) and pedestrian's density distribution during seasons (Andrienko et al., 2009).

Discovering Bluetooth enabled-devices has recently become an effective tool for human's movement monitoring purposes (Stange et al., 2011). Some research has been done on recording flows movements using Bluetooth and WiFi in outdoors and indoors. Versichele et al. (2010) studied the potential and implication of Bluetooth proximity-based tracking in moving objects. They placed a mesh of six BMSs at selected locations with distance of 50 to 200 *m*. Their study extracted the number of individuals with their route choice at particular locations. Pels et al. (2005) implemented various BMSs at Dutch train stations in order to track transit travellers. Weinzerl and Hagemann (2007) collected information from transit travellers and also tracked public transport busses by locating sensors inside buses . Abedi et al. (2014) analysed human behaviour in terms of shared space utilisation based on MAC address data. They presented MAC address data as effective information to extract features from human's spatio-temporal movement such as time spending, frequency of utilisation and group gathering . Vu et al. (2010) presented a joint Bluetooth/WiFi scanning way for evaluating the popularity of different locations as well as estimating people time spending in each location. Versichele et al. (2012b) used Bluetooth data as a tracking technology for extract features from spatio-temporal movement of music festival visitors. Versichele et al. (2012a) also developed an intelligent event management based on BMS sensor network . Abedi et al. (2013) compared the efficiency of WiFi and Bluetooth in terms of human movement data collection. Their research suggested that WiFi data range is more efficient and has higher scanning rate compared to Bluetooth enabled-devices. However, this study was only focused on collecting crowd MAC address data and did not discuss the impact of physical elements such as antenna gain. Stange et al. (2011) also employed Bluetooth tracking method to monitor visitors based on extracting their route choice. Delafontaine et al. (2012) investigated spatio-temporal sequences in Bluetooth tracking data to study movement behaviour of visitors at a major trade fair in Belgium. Danalet et al.,(2014) proposed a methodology based on Bayesian approach to detect pedestrian destination-sequence by capturing WiFi devices in the different locations. They empirically tested their algorithm at the *Ecole Polytechnique Fédérale de Lausanne* (*EPFL*) campus in *Lausanne*, *Switzerland*.

As evident from the above review, the literature has been mostly focused on applying MAC address tracking method for extracting movement features from people's movement for various applications. Physical elements such as antenna characteristics, scanner's hardware and environmental complexity have significant effects on the efficiency and precision of MAC address dataset. The impact of physical elements on the efficiency and accuracy of this dataset has not been thoroughly studied especially in terms of pedestrian and cyclist travel-time estimation.

2.4. Antenna Characteristics Effects

One of the primary stages in MAC address based data collection is to understand scanning equipment, especially antenna's type and detection range. WiFi and Bluetooth antennas are basically two types, directional and omni-directional. Omni-directional antennas send and receive signals from any direction and directional antennas only cover one direction and limited angles.

Antenna characteristic is one of the factors effecting on scanning data range. Porter et al. (2012) categorised six different antennas for assessing their capability and suitability in the Bluetooth data collection process for road traffic monitoring (that has a different environment than pedestrian monitoring). Their study shows that vertically polarized antennas with gains from 9 to 12 *dBi* are suitable for a Bluetooth based road traffic data collection. They also mentioned that the circular polarized antennas do not significantly improve the data collection process (Porter et al., 2012).

Comparing to MAC address data collection from vehicles transport, the role of antenna characteristics is more significant in crowd data collection field. For example, it is important to know that the antenna used for scanning MAC IDs is able to cover all area containing different types of environmental interference such as trees, tables, partitions, etc. Antenna can be designed in different power gains that highly impact on the antenna directivity and electromagnetic efficiency. The antenna power gain's unit is expressed in decibels and is called decibels-isotropic (*dBi*).

3. EXPERIMENTAL DESIGN

Here, first the equipment used for the study is introduced (*Section 3.1*) followed by the analysis performed on the antenna detection range (*Section 3.2*). Thereafter, the details of the study area for the case study are presented (*Section 3.3*). Finally, the pre-processing of the data obtained from the case study is described in *Section 3.4*.

3.1. Equipment

The hardware equipment used for data collection in this experiment is shown in *Fig. 1*. An integrated BMS and WMS termed as *CrossCompass* (manufactured by *Acyclica Inc.)* with the capability of scanning Bluetooth and/or WiFi addresses. This device can be configured to scan either Bluetooth or WiFi or both simultaneously. The device clock is either synchronised with PC clock or using GPS clock. It's WiFi and Bluetooth discovery times are experimentally computed from over 10,000 records. Our experiment's results show that it discovers WiFi and Bluetooth addresses every 1.37 and 5.57 seconds respectively. As can be seen from *Fig. 1*, capturing data can be stored on a flash memory and showed real-time on PC through LAN connection. Separate antenna connectors for WiFi and Bluetooth are available for plugging different antenna gains. An external wall charger or a battery source can power the scanner.

Fig. 1

3.2. Antenna Coverage Range Estimation

Antennas can be built in different gains. Manufacturers define the antenna's gain and the estimated operating range and these details are documented on the product's guide. Antenna gains are typically presented in absolute number of *1*, *2*, *3dBi* etc. However, the operating range can vary for equal gains from different manufacturers due to difference in the precision of the gain defined during the manufacturing conditions. For instance, a *3dBi* antenna manufactured by company *A* may have more or less coverage area compared to the same antenna gain made by company *B*. This variation is typically around 10 to 20 *m*. For travel time estimation using MAC address this variation in capturing range does not have significant impacts if travel time is estimated for vehicles. However, for monitoring pedestrian, this variation can have a significant role because pedestrians move slower and travel time estimation is over a short distance (few 100 *m*) compared to that of vehicles (over 2 *km* on motorways). Therefore, recording MAC address samples from pedestrians requires an accurate estimation of antenna's coverage range.

Here, we present the results of the experiments done to estimate the actual range of different omni-directional antennas used in this study. *Fig. 2* shows the experiment's equipment and environment. The environment was an open space sport field with very low environmental complexity. *Table 1* presents the assessment results of different antenna gains. As can be seen from *Table 1*, bigger antenna gains provide larger detection range. Also, for equal gains, WiFi has the bigger detection range compared to Bluetooth. This is because of the difference in Bluetooth and WiFi telecommunication architecture while they operate in same frequency (2.4 *GHz*). The difference in detection ranges are higher for smaller gains compared to bigger gains. Hence, higher antenna gains can capture more samples from human movements because they cover bigger areas. However, they may not be useful for smaller areas in terms of travel-time estimation applications as they could cover whole study area and decrease travel-time estimation accuracy.

Fig. 2

Table 1

3.3. Description of Study Area for Travel Time Estimation

Goodwill Bridge located in *Brisbane*, *Australia*, was selected as the study area. This bridge is a pedestrian-cyclist bridge and has been built over the *Brisbane River*. It connects *Southbank* suburb (including residential area, parks, museum, *Griffith University*, restaurants and cafes) to the *Gardens Point* campus of *Queensland University of Technology (QUT)* and Brisbane Central Business District (CBD). This bridge is 450 *m* long, 6.5 *m* wide and has distinct lanes for pedestrians and cyclists with speed limit of 10 *km/h*. Many walkers, runners, cyclists and skaters travel through this bridge daily. This bridge is also a popular pathway for runners and cyclists as an access to *Botanic Gardens*, *Southbank* or other city bikeways or pedestrian pathways (Musgrave, 2002). This bridge was selected as the study area for the following main reasons:

a) It is an important pedestrians and cyclists bridge for *Brisbane City Council*

This bridge is a major access route between *Southbank* and *Brisbane* CBD for pedestrians and cyclists. It plays a significant role in Brisbane transportation, connecting south and north of the city that is separated by the *Brisbane River*.

b) A noticeable number of people use this bridge daily for travel to their destinations

Over 40,000 pedestrians and cyclists use the bridge weekly and it can be counted as one of the important routes to study route-choice and travel-destination behaviour of pedestrians and cyclists. For example, *Mater Hill* and *Queen Stree*t (one of the CBD’s bus stops) bus stations are located in zone 2 and 1, respectively. People travelling from southern suburbs to CBD may tend to get off the bus in *Mater Hill* station and walk to CBD through *Goodwill Bridge* in order to pay less transportation fares as the fare increases if people travel into more zones.

c) The site are is surrounded by significant level of MAC address noises

The study site is not a simple pedestrians and cyclists bridge. As can be seen in *Fig. 3* captured from *Google Earth*, the southern gate of the bridge (where scanning point 1 is located) is surrounded by cafes and is near to *Southbank* campus of *Griffith University*. Northern gate of the bridge (where scanning point 2 is located) also is near to *QUT*’s *Gardens Point* campus and *Pacific Motorway* passes from above it. *City Ferries* also pass frequently.

d) There is a good level of environmental complexity

The spatial map of the study area, shown in *Fig. 3*, demonstrates environmental complexity level of the area. There are *QUT* and *Griffith University*’s building around the bridge as well as cafes and restaurants. In the northern side of the bridge is covered by trees and *Pacific Motorway* is built over the area.

Mentioned features of the study area are the key factors for transferability of the developed system in other complex environments. *S1* and *S2* in *Fig. 3* indicate the scanning stations at the northern and southern gates of the bridge. Four scanners were used at each station:

a) Two scanners with *2dBi* and *16dBi* antenna each for capturing WiFi addresses (WMS)
b) Two scanners with *2dBi* and *16dBi* antenna each for recording Bluetooth IDs (BMS).

In *Fig. 3*, the blue circle (inner circle) and orange circle (outer circle) illustrates the detection range of *2dBi* and *16dBi* antennas, respectively.

The data was collected during the morning peak period between 8:30 *AM* and 11:30 *AM*. During this period many students and staff cross over the bridge from *Southbank* towards *QUT*. Manual surveys were also conducted at the scanner locations to count the number of walkers, runners and cyclists using the *Goodwill Bridge* during the study period. In fact, the number of walkers, runners and cyclists passing the scanning station in each direction were recorded by volunteers. During 3 *hrs* of observation we have observed 2439 and 636 walkers, runners and cyclists moving from *S1* to *S2* and *S2* to *S1*, respectively (refer to *Fig. 7*).

Moreover, for validation of the travel time estimated using MAC data, additional surveys were performed where students were hired to act as probes (further details in *Section 4*).

Fig. 3

3.4. Pre-processing

The raw data includes *MAC address* and corresponding *detection timestamp* (see *Fig. 4*) individually for BMS and WMS. *Fig. 5a* and *Fig. 5b* shows a comparison between the percentage of WiFi and Bluetooth unique raw records in each scanning point by *2dBi* and *16dBi* antennas, respectively at both *S1* and *S2*. The results shows that the number of WiFi unique records is significantly more than Bluetooth ones captured by all four scanners. This suggests that WiFi is the more efficient MAC address dataset for tracking pedestrians and cyclists' spatio-temporal movements compared to Bluetooth. As expected, the scanners with *16dBi* antenna collected more unique records than scanners connected to *2dBi* antenna. Interestingly, more unique WiFi and Bluetooth MAC addresses were detected in point *S2* (*S2* in *Fig. 5a* and *Fig. 5b*). This is because *S2* is located near to *Pacific Highway* and QUT's Gardens Point campus where significant amount of student are present.

Each time a MAC address is detected it is stored. A Bluetooth (or WiFi) Device which is present in the scanning area for a large time will be detected multiple times. Similarly, a device detected at *S1* might not be detected at *S2* because of multiple reasons such as the device has not travelled the *Goodwill Bridge*. Therefore, a pre-processing stage was applied in to remove the aforementioned records which do not correspond to the travel time (anomalies). IE, in the pre-processing, the WiFi and Bluetooth IDs which were observed in only one scanning point were removed. Also, all IDs which were recorded during entire data collection period were presumed as fix devices and were removed from the dataset.

Fig. 4

Fig. 5

The number of unique WiFi and Bluetooth devices before (raw data) and after pre-processing are presented in *Fig. 6*. Here, *Fig. 6a* is for WiFi and Fig *6b* is for Bluetooth. Different colour represents data from *2dBi* (light colour) and *16dBi* (dark colour) antennas:

a) The number of raw WiFi addresses recorded by *16dBi* gain is almost twice bigger than *2dBi* gain dataset (refer to *before* in *Fig. 6a and Fig. 6b)*
b) Refer to *Fig. 6a*: Comparing the impact of pre-processing on the dataset of *2dBi* and *16dBi* for WiFi database indicates that the number of unique WiFi addresses in *16dBi* dataset is compressed by almost 74%, whereas this value is around 53% for the dataset collected by *2dBi* antennas. The results indicate that *16dBi* antenna scanned more anomaly WiFi MAC addresses than *2dBi* antenna. This is due to covering larger areas by *16dBi* antenna. *2dBi* antenna is then more efficient compared to *16dBi* antenna in terms of scanning less unique anomalies.
c) Refer to *Fig 6b*: The results of pre-processing on the Bluetooth database also show that *16dBi* antenna captured more anomalies compared to *2dBi* antenna.

Hence, while *16dBi* antenna collected more unique devices, it however recorded more anomalies MAC IDs compared to *2dBi* antenna. Therefore, higher gain antennas capture more unique devices but their dataset required more running time for filtering anomalies. On the other hand, lower gain antennas record fewer anomalies but they may capture less samples or miss some valuable unique records (especially those IDs that move faster).

Fig. 6

Fig. 7 presents the number of unique records travelling inbound (from *S1* to *S2)* and outbound (from *S2* to *S1)*. Here, real data (actual number of people observed from the manual survey), and pre-processed records from WiFi and Bluetooth are presented as separate bars. As can be seen from the bar chart, the number of people travelling in inbound is almost four times bigger than the number travelling outbound. This is as expected because during morning peak period the demand is high for inbound than that of outbound direction. Around 12% (=284/2439) of inbound travellers were scanned by WiFi scanner and only 0.6% (=14/2439) of them were captured by Bluetooth scanner. These proportions for outbound travellers were almost 9% (=58/636) for WiFi and 1.1% (=7/636) for Bluetooth, respectively. This indicates that BMS dataset does not have enough samples for data analysis and hence only WMS records are focused in the next section in terms of calculating travel-time of walkers, runners and cyclists in the real scenario.

Fig. 7

4. TRAVEL-TIME ESTIMATION RESULTS

Here, for a MAC-ID we define the travel time between two scanner locations as time gap between the last observation of the MAC-ID at the upstream scanner to the first observation of the MAC-ID at the downstream scanner. As we mentioned in *Section 3.3*, the system is tested in a real scenario. People passing over the bridge were counted manually and nearby Bluetooth and WiFi MAC addresses were scanned. We did not consider Bluetooth data due to a significant lower observation rate compared to WiFi and manual records. Hence, following dataset were used for further analysis:

a) Actual numbers of people travelled over the bridge inbound and outbound that is manually counted by volunteers – called as *Real-Data*
b) WiFi sample numbers based on *2dBi* Scanners' data
c) WiFi sample numbers based on *16dBi* Scanners' data

Table 2 and *Table 3* present the number of unique records travelling inbound and outbound, respectively. The results of these tables show the proportion of walkers, runners and cyclist who passed through the *Goodwill Bridge* non-stop. The actual number of walkers, runners and cyclists were counted by other volunteers and presented in *Table 2* and *Table 3* as *Real-Data*. Also, the records which passed through the bridge with a travel-time more than a typical walker are counted as devices which stopped during the journey. This group is named as *non-active travellers*. This group of records can represent people who spent some amounts of time on the bridge for taking a picture or having a coffee for example. The results of *Table 2* and *3* indicate that for *Real-Data* around 84% of people was observed as walkers; 12.5% are runners and 3.5% are cyclist, whereas for *16dBi* antenna around 73% were recorded as active walkers, around 8.5% as cyclist and around 6% as runners. In addition, near 12.5% of the records were counted as non-active travellers that can be walkers, runners or cyclists who stopped during their travel.

The line-graph in *Fig. 8* represents the number of real records (primary *Y-axis*, solid lines) and unique WiFi samples (secondary *Y-axis*, dotted lines) for inbound and outbound. As can be seen from *Fig. 8*, the number of inbound walkers (based on *Real-Data*) increased from 9:00 *AM* to 9:30 *AM* from almost 100 to 250 samples and then dropped to around 120 samples at 10:00 *AM*. WiFi samples for inbound travellers also clearly represent the peak period between 9:00 *AM* and 10:00 *AM*. However, the outbound did not experience any peak in sample size for both real and WiFi samples. Bluetooth dataset was not presented in *Fig. 8* due to the lack of data.

The percentage of real records to unique WiFi records are presented in *Fig. 9*, where *X-axis* is time (same as *Fig. 8*) with different bars for inbound (*S1* to *S2: dark colour*) and outbound (*S2* to *S1; light colour*). It can be seen that during the peak period (from 8:30 *AM* to 10:30 *AM*) around 8% to 12 % real-samples are represented by WiFi database.

Table 2

Table 3

Fig. 8

Fig. 9

To validate the travel time estimation using the developed system, manual test scenarios have been performed. Seven students were hired as walkers, runners and cyclists. These students were equipped to Bluetooth and WiFi devices and their travel-times through the bridge were manually recorded. The MAC ID of the devices being transported by these students is known. They started their trip over 300 *m* far from the bridge's entrance and crossed the bridge non-stop in regular speed. Then, 35 trips were done in total. Each trip individually includes five different records as:

a) Actual Travel-Time recorded by volunteer's timer – called as *Ground-truth (or real sample)* travel-time
b) WiFi Travel-Time by *2dBi* antenna Scanners
c) WiFi Travel-Time by *16dBi* antenna Scanners
d) Bluetooth Travel-Time by *2dBi* antenna Scanners
e) Bluetooth Travel-Time by *16dBi* antenna Scanners

This survey is aimed to compare accuracy of WiFi and Bluetooth data with real records. *Fig. 10* illustrates a time-space trajectory plot for a MAC address device passing through a scanning zone. *Inquiry Train* is in fact the state that a MAC address scanner inquiries discoverable MAC IDs in the detection zone. The list of all symbols used in this section and their meaning are presented in *Table 4*.

Table 4

As can be seen from *Fig. 10*, there is a time delay between actual arrival time (t_A) and first scanning time (t'_A) as well as another time delay between last scanning time (t'_D) and actual departure time (t_D) of a device from the scanning zone. Actual time duration in scanning point can be calculated from

$$dt = t_D - t_A \tag{1}$$

and reported time duration by scanner is

$$dt' = t'_D - t'_A \tag{2}$$

Then, the temporal error of MAC address scanner in arrival time can be defined as

$$e_A = t'_A - t_A \tag{3}$$

Also, there is a temporal error of MAC address scanner in departure time that can written as follow

$$e_D = t_D - t'_D \tag{4}$$

The reported time period between the first and last observations can be also re-written as

$$dt' = dt - (e_A + e_D) \tag{5}$$

Fig. 10

Fig. 11 illustrates the systematic method of travel-time estimation based on MAC address dataset. *S1* and *S2* are two scanning points located Dx *m* far from each other. The actual exit-to-enter travel-time can be calculated from

$$TT_{ext-ent} = t_{A(S2)} - t_{D(S1)} \tag{6}$$

where $t_{D(S1)}$ is the actual departure time in *S1*'s scanning zone and $t_{A(S2)}$ is the actual arrival time in *S2*'s scanning zone. The estimated travel-time based on MAC address data is also

$$TT'_{ext-ent} = t'_{A(S2)} - t'_{D(S2)} \tag{7}$$

where $t'_{D(S1)}$ is the last scan time in *S1*'s scanning zone and $t'_{A(S2)}$ is the first scan in *S2*'s scanning zone.

Fig. 11

Table 5 provides the average travel-times between the scanner locations calculated for independently for walkers, runners and cyclists. Here, the first row represents the ground truth average ($\overline{Dt}_{Ground-truth}$), estimated from the manual survey. The second and third row represents travel time from WMS and BMS data, respectively using *2dBi* antenna. The fourth and fifth row represents travel time from WMS and BMS data, respectively using *16dBi* antennas. It is observed that the travel time estimates from *2dBi* antenna is higher than that from *16dBi* antenna. This is because of the smaller detection zone for *2dBi* antenna compared to that of *16dBi* antenna (See *Table 1*).

Table 5

Based on *Grand-truth* records, the average speed of test scenario samples can be calculated from

$$\overline{V}_{Ground-truth(S1-S2)} = \frac{Dx}{Dt} = \frac{x_{S2} - x_{S1}}{t_{S2} - t_{S1}} \quad (8)$$

where: Dx is distance between scanning points,

Dt is the travel-time between scanning points recorded by volunteers,

x_{S2} and x_{S1} are the position coordinates of the points where *S2* and *S1*, respectively, and

t_{S2} and t_{S1} are the time when the volunteer has crossed points *S2* and *S1*, respectively.

From MAC address data, the average speed of test scenarios can be calculated from

$$\overline{V}_{Scanner(S1-S2)} = \frac{dx}{TT_{ext-ent}} = \frac{Dx - 2\times R}{t_{A(S2)} - t_{D(S1)}} \quad (9)$$

However, the actual average speed is

$$\overline{V}'_{Scanner(S1-S2)} = \frac{dx'}{TT'_{ext-ent}} = \frac{Dx - 2\times r'}{t'_{A(S2)} - t'_{D(S1)}} \quad (10)$$

Here, for simplicity of explanation we have assumed that *r'* is same for both *S1* and *S2* for given antenna.

The radar graph in *Fig. 12a* shows the average speed of each test scenarios for *Ground-truth*, *2dBi* WiFi, *16dBi* WiFi, *2dBi* Bluetooth and *16dBi* Bluetooth dataset. As an example, a WiFi device in *Test 1* (see Fig. 12a) were moved by a volunteer 395 *m* from *S1* to *S2* within 275 *sec*, moving in average speed of approximately 1.44 *m/s*. Calculating the exit-enter travel-time ($TT'_{ext-ent}$) of the same WiFi device based on 2*dBi* and 16*dBi* dataset indicates 223 and 176 *sec* travel-time, respectively. Based on the antenna detection range estimation presented in *Table 1*, we expect the distance of 315 and 245 *m* ($dx = Dx - 2\times R$) between scanners detection zones for *2dBi* and *16dBi* antennas, respectively. Then, the reported average speeds ($\overline{V}'_{Scanner(S1-S2)}$) of the same WiFi device based on *2dBi* and *16dBi* scanners are respectively about 1.41 and 1.39 *m/s*. However, the actual radius of scanning point (r') is smaller than the estimated radius (R) as $t'_A > t_A$ and $t_D > t'_D$. Then, the actual radius of scanning zone (r') can be estimated based on

$$r' = \frac{Dx - \left(\overline{V}'_{Scanner(S1-S2)} \times TT'_{ext-ent}\right)}{2} \quad (11)$$

and

$$\Delta_R = R - r' \quad (12)$$

Hence, the measurement error can be defined as

$$e_{\overline{V}} = \overline{V}_{Ground-truth} - \overline{V}'_{Scanner(S1-S2)} \quad (13)$$

The results of all 5 tests for walkers indicate that *2dBi* antenna represent more accurate database comparing to *16dBi* in terms of travel time estimation. This is because the distance between detection zones (dx') of *2dBi* antenna is larger than *16dBi* antenna. We also observed that WiFi samples are more precise than Bluetooth data mainly because WiFi's scanning rate is about 1 *sec* while Bluetooth scanner captured MAC IDs almost every 5 *sec*. We were unable to do the same comparison on runners and cyclists as *2dBi* antenna failed to record some of runner and cyclists test samples.

In terms of travel-time, each dataset presents different value for travel-time of a test scenario. For instance, travel-time of *Test 1* is

- 275 *sec* based on *Ground-truth* where $Dx = 395m$,
- 223 *sec* based on *2dBi* WiFi scanners where $dx = 315m$,
- 176 *sec* based on *16dBi* WiFi scanners where $dx = 245m$,
- 253 *sec* based on *2dBi* Bluetooth scanners $dx = 340m$,
- 215 *sec* based on *16dBi* Bluetooth scanners $dx = 265m$.

All travel-times of a test scenario can be re-calculated for 100 *m* as

$$\overline{TT}_{Ground-truth/100m} = \frac{100}{\overline{V}_{Groun-truth}} \quad (13)$$

$$\overline{TT}'_{ext-ent/100m} = \frac{100}{\overline{V}_{Scanner(S1-S2)/100m}} \quad (14)$$

Hence, the travel-time error for each 100 *m* can be defined as

$$e_{\overline{TT}(100m)} = \overline{TT}_{Ground-truth/100m} - \overline{TT}'_{ext-ent/100m} \quad (15)$$

The bar chart in *Fig. 12b* compares the travel-time errors per 100 *m* between WiFi and Bluetooth dataset for each test scenarios. As can be seen from *Fig. 12b*, *2dBi* WiFi scanners had 3 *sec* error in travel-time estimation of *Test 1* every 100 *m*. In overall, lower gain antenna provided more accurate dataset for walker's travel time estimation.

Fig. 12

The sample size of WiFi addresses for active walkers were efficient to analyse the inbound and outbound travel-times. *Fig. 13* and *14* show the walkers' travel-time based on *2dBi* scanners' dataset for inbound and outbound, respectively. Because all active walkers were sampled by both of *2dBi* and *16dBi* antennas, there were not any significant changes in travel-time pattern and both present same shape but *2dBi* antenna presents more accurate values as mentioned in *Fig. 12*. The box plots actually present the walkers' travel-time (primary *y-axis*) and line-graph presents the walker's sample size (secondary *y-axis*) for 10 *min*. As can be seen from *Fig. 13*, the average travel-time of inbound (from *S1* to *S2*) walkers increased between *9* and *10 AM* when higher sample sizes were captured. This increase in walker's travel-time could be due to big number of pedestrians travelling inbound (from *S1* to *S2*) between *9* and *10 AM*. This increase in travel-time also indicates the impact of crowd congestion on pedestrian travel-time in narrow pathways. On the other hand, outbound does not have significant increase in travel-time value of walkers during the morning as not many people passed from *S2* to *S1*. As there was not any peak in travellers' sample size in outbound direction (from *S2* to *S1*), walker's travel-time fluctuated between 190 and 220 *sec* based on *Fig. 14*.

Fig. 13

Fig. 14

The results showed that the performance of both *2dBi* and *16dBi* antennas was same in case of detecting walkers. However, *2dBi* antenna provided less anomalies and lower percent of error in travel time estimation for shorter study corridors. The results presented in *Fig. 15* and *16* compare the performance of each antenna in order to detect walkers, runners and cyclists. As can be seen from *Fig. 15* and *16*, *2dBi* antenna has missed to capture noticeable numbers of WiFi devices carried by runners and cyclists due to covering smaller areas. Here, *2dBi* has captured only 2 cyclists and 5 runners, which is significantly lower than that captured by *16dBi* antenna. As the results, *16dBi* antenna could collect more unique samples from available WiFi devices carried by runners and cyclists compared to *2dBi* antenna. While *2dBi* antenna collected less efficient data from runners and cyclist, it provided more accurate estimation from walker's travel-time.

Fig. 15

Fig. 16

5. CONCLUSIONS and DISCUSSION

This paper empirically assessed the impact of small and big antenna gains on tracking movements of pedestrian and cyclists based on MAC address dataset, especially in the case of travel-time estimation application. The results have been verified by Ground-truth samples and test scenarios.

5.1. Antenna Characteristics Effects

It is observed that WMS has much higher capture rate than that of BMS. For the case study performed, WMS and BMS have captured around 12% and 1% of the target travellers (walkers, runners and cyclist), respectively. Based on this we can conclude that WMS should be used for monitoring pedestrian and cyclists. This is contrary to the monitoring of road traffic where BMS has better sample size. However, the higher rates of capturing WiFi MAC address than Bluetooth does not necessary correspond existing of higher number of enabled WiFi devices compared to Bluetooth ones. The main reasons of this phenomenon could be due to the differences in their operational architecture and utilisation popularity. WiFi has two operational modes (*OFF* and *ON*), whereas Bluetooth has three operational states (*OFF*, *ON-Visible*, *ON-Invisible*). Then, BMSs are able to only capture the Bluetooth devices that are in *ON-Visible* state. In terms of power consumption efficiency and security purposes, the default settings of most Bluetooth devices is *Visible* in case of no active connection and turning to *Invisible* immediately after a connection established. Because of this feature, there is the possibility of existing more active Bluetooth devices than WiFi in a scanning zone but the number of visible Bluetooth devices might be less than active WiFi ones.

Popularity of utilising WiFi could also play a substantial role in presence of more enable WiFi than Bluetooth in a crowd zone. As WiFi's bandwidth allows for higher data exchange rates, it is normally aimed for Internet access and people may tend to keep their devices' WiFi turned on most of the time due to significant daily needs of the Internet. While most of smart phones benefit cellular network (3G and 4G) technologies for Internet access, WiFi is reasonably in priority because it is cheaper and faster than 3G and 4G networks. It can be assumed that smart device users usually tend to keep their device's WiFi turned on to increase the chance of connectivity to any nearby WiFi networks. Bluetooth, on the other hand, is used when it is required. For example, mobile users use Bluetooth technology if they want to stream music to Bluetooth headsets or stereos. Then, users tend to keep their Bluetooth turned off when there is no demand mainly because of saving in their device's battery life. Hence, the possibility of scanning more WiFi MAC addresses than Bluetooth could be assumed as the difference in their nature of utilisation.

Comparison of *2dBi* and *16dBi* antennas in terms of collecting accurate data indicates that bigger antenna gains collect more unique samples as they cover wider areas. However, they may not be useful for small scales of monitoring environment due to overlapping possibilities and scanning more anomaly samples such as fix WiFi or Bluetooth devices. *2dBi* antenna gain collected less samples compared to *16dBi* but its dataset is more optimised and accurate. However, *2dBi* antenna was not suitable for collecting MAC address samples from runners and cyclists as they normally move faster than walkers and spend less time in scanning zones. On the other hand, while *16dBi* antenna collected more anomalies, it had a better performance in capturing cyclists and runner's WiFi and Bluetooth enabled devices.

MAC address is a useful dataset for tracking object's spatio-temporal movement. Research in vehicle transportation suggested big antenna gains as suitable equipment. This study showed that antenna gain significantly effects on the dataset's accuracy in terms of pedestrians and cyclists' movement tracking. This research showed than both small and big antenna gains have benefits and some drawbacks. Small antenna gains have been suggested in this study as optimal equipment for monitoring walkers and slow runners. Also, this research recommends bigger values of antenna gains for recording cyclists and fast runners' WiFi and Bluetooth enabled devices. The finding of this research can effectively apply for collecting efficient and effective database in order to monitoring pedestrians and cyclists.

5.2. Added values to Transportation Research

The findings of this study can be useful to improve the accuracy of people's movement monitoring by considering people's movement speed and environment's scale (small or large corridors). This can effectively help transportation research for detailed study of pedestrians and cyclists' movement behaviour that can be applied to various applications (e.g., pedestrian and cyclists destination modelling and route-choice analysis, crowd safety, crowd congestion control, evacuation strategies and urban pathway design etc.). Precision in movement monitoring of pedestrians and cyclists in urban areas can be effectively usable to optimise and enhance urban transport system by considering travel destination of active travellers.

ACKNOWLEDGEMENTS

The authors would like to acknowledge School of Civil Engineering and Build Environment, Science and Engineering Faculty, Queensland University of Technology (QUT) for supporting this research.

REFERENCES

ABBOTT-JARD, M., SHAH, H. & BHASKAR, A. Empirical evaluation of Bluetooth and Wifi scanning for road transport 36th Australasian Transport Research Forum (ATRF), 2 - 4 October 2013 Brisbane, Australia.

ABEDI, N. 2014. *Monitoring spatiotemporal dynamics of human movement based on MAC address data.* Masters of Engineering, Queensland University of Technology.

ABEDI, N., BHASKAR, A. & CHUNG, E. Bluetooth and Wi-Fi MAC Address Based Crowd Data Collection and Monitoring: Benefits, Challenges and Enhancement. Australasian Transport Research Forum (ATRF), 36th, 2013, Brisbane, Queensland, Australia, 2013.

ABEDI, N., BHASKAR, A. & CHUNG, E. 2014. Tracking spatio-temporal movement of human in terms of space utilization using Media-Access-Control address data. *Applied Geography,* 51**,** 72-81.

AHAS, R., AASA, A., ROOSE, A., MARK, Ü. & SILM, S. 2008. Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. *Tourism Management,* 29**,** 469-486.

AHLQVIST, O., BAN, H., CRESSIE, N. & SHAW, N. Z. 2010. Statistical counterpoint: Knowledge discovery of choreographic information using spatio-temporal analysis and visualization. *Applied Geography,* 30**,** 548-560.

ANDRIENKO, G. & ANDRIENKO, N. Extracting patterns of individual movement behaviour from a massive collection of tracked positions. Workshop on Behaviour Modelling and Interpretation (BMI), 2007a Bremen. 1-16.

ANDRIENKO, G., ANDRIENKO, N., BAK, P., KISILEVICH, S. & KEIM, D. Analysis of community-contributed space- and time-referenced data (example of flickr and panoramio photos). Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on, 12-13 Oct. 2009 2009. 213-214.

ANDRIENKO, N. & ANDRIENKO, G. 2007b. Designing Visual Analytics Methods for Massive Collections of Movement Data. *Cartographica: The International Journal for Geographic Information and Geovisualization,* 42**,** 117-138.

ASANO, M., IRYO, T. & KUWAHARA, M. 2010. Microscopic pedestrian simulation model combined with a tactical model for route choice behaviour. *Transportation Research Part C: Emerging Technologies,* 18**,** 842-855.

BARCELÓ, J., MONTERO, L., MARQUÉS, 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**,** 19-27.

BATTY, M. 2003. Agent-based pedestrian modelling. *Advanced spatial analysis: The CASA book of GIS***,** 81-106.

BHASKAR, A. 2009. *A methodology (CUPRITE) for urban network travel time estimation by integrating multisource data.* Doctor of Sciences (Ph.D.), Ph.D. Thesis ,The Ecole Polytechnique Fédérale de Lausanne (EPFL)

BHASKAR, A. & CHUNG, E. 2013. Fundamental understanding on the use of Bluetooth scanner as a complementary transport data. *Transportation Research Part C: Emerging Technologies,* 37**,** 42-72.

BHASKAR, A., KIEU, L. M., QU, M., NANTES, A., MISKA, M. & CHUNG, E. 2015. Is bus overrepresented in Bluetooth MAC Scanner data? Is MAC-ID really unique? *International Journal of Intelligent Transportation Systems Research (http://dx.doi.org/10.1007/s13177-014-0089-9),* 13**,** 119-130.

BHASKAR, A., QU, M. & CHUNG, E. 2014a. Bluetooth Vehicle Trajectories by Fusing Bluetooth and Loops: Motorway Travel Time Statistics *IEEE Transactions on Intelligent Transportation Systems (in press, DOI: 10.1109/TITS.2014.2328373)*.

BHASKAR, A., TSUBOTA, T., KIEU, L. M. & CHUNG, E. 2014b. Urban traffic state estimation: Fusing point and zone based data. *Transportation Research Part C: Emerging Technologies,* 48**,** 120-142.

BIERLAIRE, M. & ROBIN, T. 2009. Pedestrians choices. *Pedestrian Behavior***,** 1-26.

BOGORNY, V., KUIJPERS, B. & ALVARES, L. O. 2009. ST-DMQL: A Semantic Trajectory Data Mining Query Language. *International Journal of Geographical Information Science,* 23**,** 1245-1276.

BOLES, W. 1981. The effect of density, sex, and group size upon pedestrian walking velocity. *Man-environment systems*.

BOWMAN, B. L. & VECELLIO, R. L. 1994. Pedestrian walking speeds and conflicts at urban median locations. Transportation Research Board.

BULLOCK, D., HASEMAN, R., WASSON, J. & SPITLER, R. 2010. Automated Measurement of Wait Times at Airport Security. *Transportation Research Record: Journal of the Transportation Research Board,* 2177**,** 60-68.

BURGESS, J. 1983. Interpersonal spacing behavior between surrounding nearest neighbors reflects both familiarity and environmental density. *Ethology and sociobiology,* 4**,** 11-17.

CHAKRABORTY, G., NAIK, K., CHAKRABORTY, D., SHIRATORI, N. & WEI, D. 2010. Analysis of the Bluetooth device discovery protocol. *Wireless Networks,* 16**,** 421-436.

CIOLEK, M. T. 1978. Spatial arrangements in social encounters: An attempt at a taxonomy. *Man-Environment Systems*.

COFFIN, A. & MORRALL, J. 1995. Walking speeds of elderly pedestrians at crosswalks. *Transportation Research Record***,** 63-67.

DABBS JR, J. M. & STOKES, N. A. 1975. Beauty is power: The use of space on the sidewalk. *Sociometry***,** 551-557.

DANALET, A., FAROOQ, B. & BIERLAIRE, M. 2014. A Bayesian approach to detect pedestrian destination-sequences from WiFi signatures. *Transportation Research Part C: Emerging Technologies,* 44**,** 146-170.

DEE, H. & VELASTIN, S. 2008. How close are we to solving the problem of automated visual surveillance? *Machine Vision and Applications,* 19**,** 329-343.

DELAFONTAINE, M., VERSICHELE, M., NEUTENS, T. & VAN DE WEGHE, N. 2012. Analysing spatiotemporal sequences in Bluetooth tracking data. *Applied Geography,* 34**,** 659-668.

DUIVES, D. C., DAAMEN, W. & HOOGENDOORN, S. P. 2013. State-of-the-art crowd motion simulation models. *Transportation research part C: emerging technologies,* 37**,** 193-209.

FRUIN, J. 1992. Designing for pedestrians. National Transportation Library: Research and Innovation Technology Administration (RITA) (http://ntl.bts.gov/DOCS/11877/Chapter_8.html).

FUGGER, T. F., RANDLES, B. C., STEIN, A. C., WHITING, W. C. & GALLAGHER, B. 2000. Analysis of pedestrian gait and perception-reaction at signal-controlled crosswalk intersections. *Transportation Research Record: Journal of the Transportation Research Board,* 1705**,** 20-25.

GIANNOTTI, F. & PEDRESCHI, D. 2008. Mobility, data mining and privacy: A vision of convergence.

HABICHT, A. T. & BRAAKSMA, J. P. 1984. Effective width of pedestrian corridors. *Journal of transportation engineering,* 110**,** 80-93.

HAN, B. & SRINIVASAN, A. eDiscovery: Energy efficient device discovery for mobile opportunistic communications. 2012.

HELBING, D., FARKAS, I. & VICSEK, T. 2000. Simulating dynamical features of escape panic. *Nature,* 407**,** 487-490.

HENDERSON, L. 1971. The statistics of crowd fluids. *Nature,* 229**,** 381-383.

HOOGENDOORN, S. P. & DAAMEN, W. 2005. Pedestrian behavior at bottlenecks. *Transportation Science,* 39**,** 147-159.

HOSSAIN, A. K. M. M. & WEE-SENG, S. A Comprehensive Study of Bluetooth Signal Parameters for Localization. Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on, 3-7 Sept. 2007 2007. 1-5.

HUI, S. K., FADER, P. S. & BRADLOW, E. T. 2009. Research Note—The Traveling Salesman Goes Shopping: The Systematic Deviations of Grocery Paths from TSP Optimality. *Marketing Science,* 28**,** 566-572.

JANKOWSKI, P., ANDRIENKO, N., ANDRIENKO, G. & KISILEVICH, S. 2010. Discovering Landmark Preferences and Movement Patterns from Photo Postings. *Transactions in GIS,* 14**,** 833-852.

KASEMSUPPAKORN, P. & KARIMI, H. A. 2013. A pedestrian network construction algorithm based on multiple GPS traces. *Transportation research part C: emerging technologies,* 26**,** 285-300.

KEMPERMAN, A. D. A. M., BORGERS, A. W. J. & TIMMERMANS, H. J. P. 2009. Tourist shopping behavior in a historic downtown area. *Tourism Management,* 30**,** 208-218.

KHOLSHEVNIKOV, V. V. & SAMOSHIN, D. A. 2008. Laws of motion of pedestrian flow—basics for evacuation modeling and management. *Resilience of Cities to Terrorist and other Threats.* Springer.

KIEU, L. M., BHASKAR, A. & CHUNG, E. 2015. Empirical modelling of the relationship between bus and car speeds on signalised urban networks. *Transportation Planning and Technology,* 38**,** 465-482.

KNEIDL, A., HARTMANN, D. & BORRMANN, A. 2013. A hybrid multi-scale approach for simulation of pedestrian dynamics. *Transportation research part C: emerging technologies,* 37**,** 223-237.

KNOBLAUCH, R. L., PIETRUCHA, M. T. & NITZBURG, M. 1996. Field studies of pedestrian walking speed and start-up time. *Transportation Research Record: Journal of the Transportation Research Board,* 1538**,** 27-38.

KWON, Y.-I., MORICHI, S. & YAI, T. 1998. Analysis of pedestrian behavior and planning guidelines with mixed traffic for narrow urban streets. *Transportation Research Record: Journal of the Transportation Research Board,* 1636**,** 116-123.

LAM, W. H., MORRALL, J. F. & HO, H. 1995. Pedestrian flow characteristics in Hong Kong. *Transportation Research Record***,** 56-62.

LAUBE, P., IMFELD, S. & WEIBEL, R. 2005. Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science,* 19**,** 639-668.

LIEBIG, T. & WAGOUM, A. U. K. Modelling Microscopic Pedestrian Mobility using Bluetooth. International Conference on Agents and Artificial Intelligence (ICAART) (2), 2012. 270-275.

MALINOVSKIY, Y., SAUNIER, N. & WANG, Y. 2012. Analysis of pedestrian travel with static bluetooth sensors. *Transportation Research Record: Journal of the Transportation Research Board,* 2299**,** 137-149.

MILLONIG, A. & GARTNER, G. Shadowing-Tracking-Interviewing: How to Explore Human Spatio-Temporal Behaviour Patterns. BMI, 2008. Citeseer, 1-14.

MOTTRAM, C. Collective choreography of space: modelling digital co-Presence in a public arena. 19th International Conference on Systems Research, Informatics and Cybernetics, July 29-Aug 4 2007 Baden-Baden, Germany. UCL Discovery, 59-63.

MUSGRAVE, E. 2002. Goodwill overture. *Architecture Australia,* 91**,** 66-69.

NANTES, A., MISKA, M., BHASKAR, A. & CHUNG, E. 2014. Noisy Bluetooth traffic data? *Road & Transport Research,* 23**,** 33-43.

NANTES, A., NGODUY, D., BHASKAR, A., MISKA, M. & CHUNG, E. 2015. Real-time traffic state estimation in urban corridors from heterogeneous data. *Transportation Research Part C: Emerging Technologies doi:10.1016/j.trc.2015.07.005*.

NASSIR, N., HICKMAN, M., ZHENG, H. & CHIU, Y.-C. Network Flow Solution Method for Optimal Integrated Traffic Routing and Signal Timing to Evacuate Urban Networks with Varying Threat Levels. Transportation Research Board 93rd Annual Meeting, 2014.

NASSIR, N., ZHENG, H., HICKMAN, M. & CHIU, Y.-C. Optimal Traffic Routing for Large-Scale Evacuation in Urban Networks with Various Threat Levels. Transportation Research Board 92nd Annual Meeting, 2013.

O'CONNOR, A., ZERGER, A. & ITAMI, B. 2005. Geo-temporal tracking and analysis of tourist movement. *Mathematics and Computers in Simulation,* 69**,** 135-150.

OKUBO, A. 1986. Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. *Advances in biophysics,* 22**,** 1-94.

PELS, M., BARHORST, J., MICHELS, M., HOBO, R. & BARENDSE, J. 2005. Tracking people using Bluetooth: Implications of enabling Bluetooth discoverable mode. *Final report, University of Amsterdam*.

POLUS, A., SCHOFER, J. L. & USHPIZ, A. 1983. Pedestrian flow and level of service. *Journal of Transportation Engineering,* 109**,** 46-56.

PORTER, J. D., KIM, D. S., MAGAÑA, M. E., POOCHAROEN, P. & ARRIAGA, C. A. G. 2012. Antenna Characterization for Bluetooth-Based Travel Time Data Collection. *Journal of Intelligent Transportation Systems,* 17**,** 142-151.

PUSHKAREV, B. 1975. Urban space for pedestrians. *Cambridge, MA: MIT Pres***,** 212.

ROTTON, J., SHATS, M. & STANDERS, R. 1990. Temperature and Pedestrian Tempo Walking Without Awareness. *Environment and Behavior,* 22**,** 650-674.

SHAW, S.-L., YU, H. & BOMBOM, L. S. 2008. A Space-Time GIS Approach to Exploring Large Individual-based Spatiotemporal Datasets. *Transactions in GIS,* 12**,** 425-441.

SHIWAKOTI, N., SARVI, M., ROSE, G. & BURD, M. 2011. Animal dynamics based approach for modeling pedestrian crowd egress under panic conditions. *Transportation research part B: methodological,* 45**,** 1433-1449.

SHOVAL, N. & ISAACSON, M. 2007. Tracking tourists in the digital age. *Annals of Tourism Research,* 34**,** 141-159.

STANGE, H., LIEBIG, T., HECKER, D., ANDRIENKO, G. & ANDRIENKO, N. Analytical workflow of monitoring human mobility in big event settings using bluetooth. Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness, 2011. ACM, 51-58.

TSUBOTA, T., BHASKAR, A. & CHUNG, E. 2014. Macroscopic Fundamental Diagram for Brisbane, Australia: Empirical Findings on Network Partitioning and Incident Detection. *Transportation Research Record: Journal of the Transportation Research Board,* 2421**,** 12-21.

TSUBOTA, T., BHASKAR, A., CHUNG, E. & BILLOT, R. 2011. Arterial traffic congestion analysis using Bluetooth Duration data. *Australasian Transport Research Forum* Adelaid, SA.

VAN LONDERSELE, B., DELAFONTAINE, M. & VAN DE WEGHE, N. 2009. Bluetooth tracking. *GIM International,* 23**,** 23-25.

VAN SCHAICK, J. & VAN DER SPEK, S. C. 2008. *Urbanism on track [electronic resource]: application of tracking technologies in urbanism*, IOS Press.

VERSICHELE, M., DELAFONTAINE, M., NEUTENS, T. & VAN DE WEGHE, N. Potential and implications of bluetooth proximity-based tracking in moving object research. 2010. 111-116.

VERSICHELE, M., HUYBRECHTS, R., NEUTENS, T. & VAN DE WEGHE, N. Intelligent Event Management with Bluetooth Sensor Networks. Intelligent Environments (IE), 2012 8th International Conference on, 2012a. IEEE, 311-314.

VERSICHELE, M., NEUTENS, T., DELAFONTAINE, M. & VAN DE WEGHE, N. 2012b. The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities. *Applied Geography,* 32**,** 208-220.

VIRKLER, M. R. 1998. Prediction and measurement of travel time along pedestrian routes. *Transportation Research Record: Journal of the Transportation Research Board,* 1636**,** 37-42.

VU, L., NAHRSTEDT, K., RETIKA, S. & GUPTA, I. 2010. Joint bluetooth/wifi scanning framework for characterizing and leveraging people movement in university campus. *Proceedings of the 13th ACM international* *conference on Modeling, analysis, and simulation of wireless and mobile systems.* Bodrum, Turkey: ACM.

WANG, W. L., LO, S. M., LIU, S. B. & KUANG, H. 2014. Microscopic modeling of pedestrian movement behavior: Interacting with visual attractors in the environment. *Transportation Research Part C: Emerging Technologies,* 44**,** 21-33.

WASSON, J. S., STURDEVANT, J. R. & BULLOCK, D. M. 2008. Real-time travel time estimates using media access control address matching. *ITE Journal (Institute of Transportation Engineers),* 78**,** 20-23.

WEIDMANN, U., KIRSCH, U. & SCHRECKENBERG, M. 2014. *Pedestrian and Evacuation Dynamics 2012*, Springer Science & Business.

WEINZERL, J. & HAGEMANN, W. 2007. Automatische Erfassung von Umsteigern per Bluetooth-Technologie. *Nahverkehrspraxis,* 3**,** 18-19.

XIA, J., CHEN, M. & HUANG, W. 2011. A multistep corridor travel-time prediction method using presence-type vehicle detector data. *Journal of Intelligent Transportation Systems,* 15**,** 104-113.

ZENG, W., CHEN, P., NAKAMURA, H. & IRYO-ASANO, M. 2014. Application of social force model to pedestrian behavior analysis at signalized crosswalk. *Transportation Research Part C: Emerging Technologies,* 40**,** 143-159.

Figures

Fig. 1. Wi-Fi and Bluetooth MAC address scanning hardware used for data collection: computational unit (1), Wi-Fi (2) and Bluetooth (3) antenna connector, USB storage (4), omni-directional antenna (5), LAN cable (6) for data connection to PC, 240 *v AC* to 5 *v DC* power convertor (6) and rechargeable 14 *v* acid batter.

Fig. 2. Experiment equipment and place (Kelvin Grove (KG) Oval, QUT KG campus)

Fig. 3. Spatial map of study area. *S1* and *S2* show the location of scanning points. The detection range of *2dBi* and *16dBi* antennas are shown in blue and orange circles, respectively.

Time (RTC)	MAC Address
1382438075	38:e7:d8:02:9d:a7
1382438075	00:3b:ff:7c:4e:2c
1382438077	00:3b:ff:7c:4e:2c
1382438087	3c:5a:37:0a:20:ff
1382438097	3c:5a:37:0a:20:ff

Fig. 4. An example of MAC address scanner's raw data.

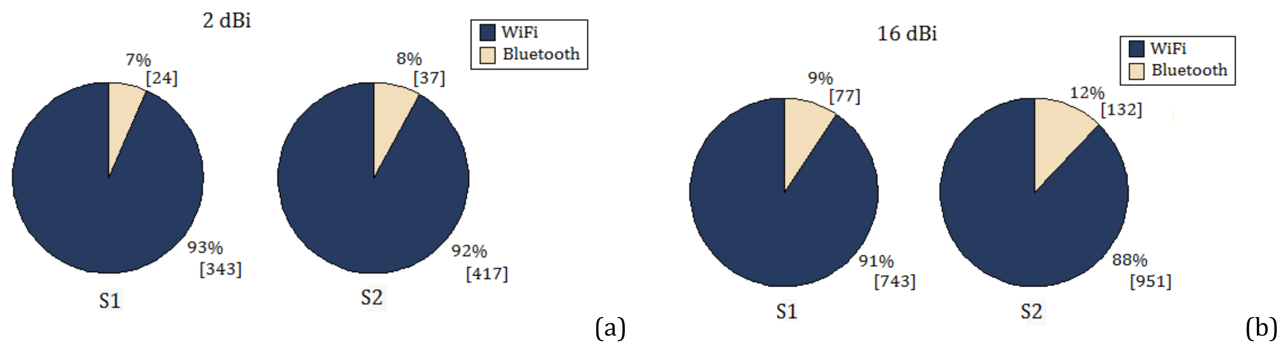

Fig. 5. The percentage of WiFi and Bluetooth unique MAC addresses in each scanning point for (a) *2dBi* and (b) *16dBi* antenna gain

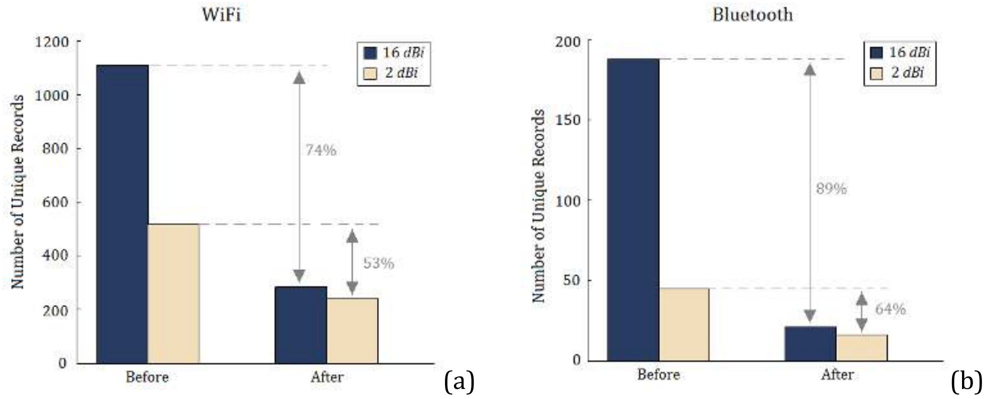

Fig. 6. The number of unique (a) WiFi and (b) Bluetooth devices after and before pre-processing stage.

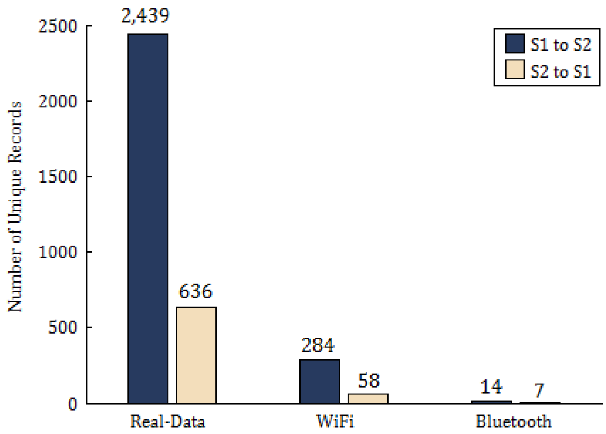

Fig. 7. The number of (a) real, (b) WiFi and (c) Bluetooth unique records

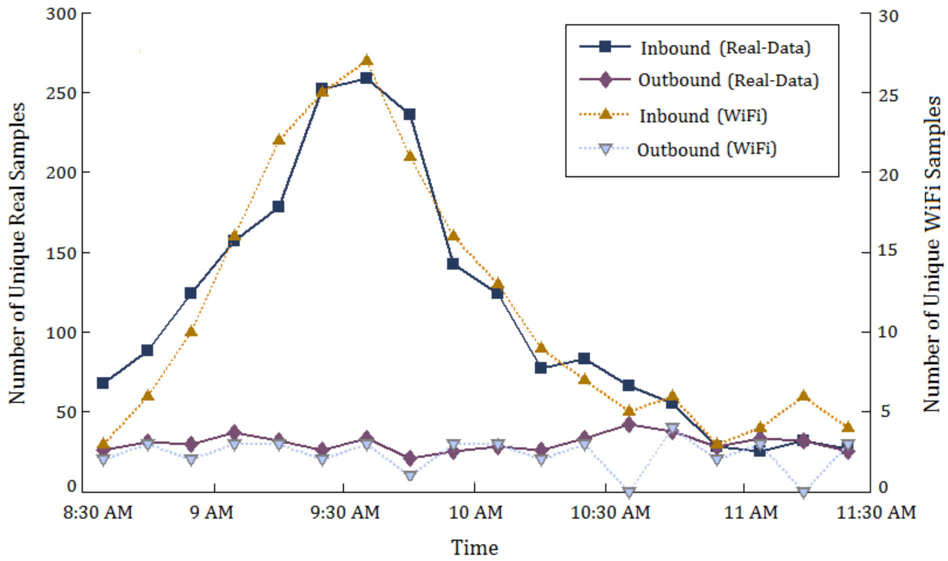

Fig. 8. The walker's sample size of real and WiFi unique records travelling inbound and outbound every 10 *min*

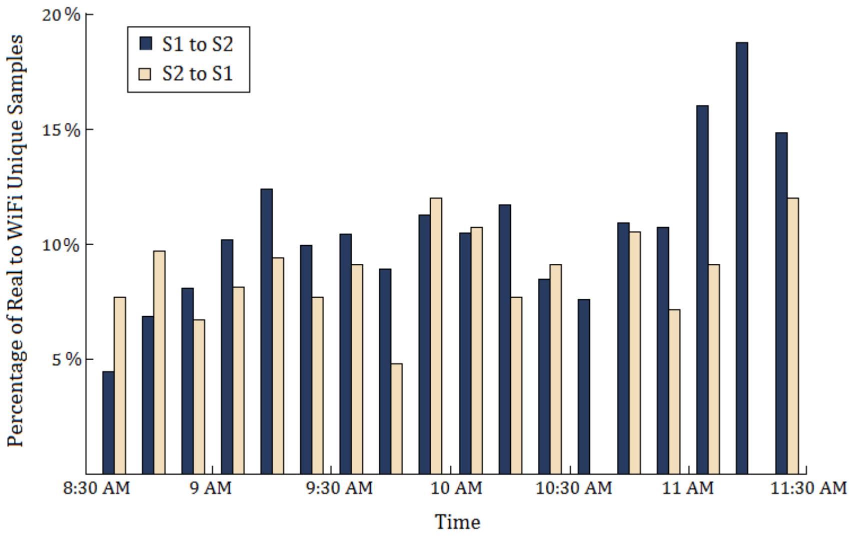

Fig. 9. The percentage of unique WiFi samples to the number of real samples for walkers travelling inbound and outbound every 10 *min*

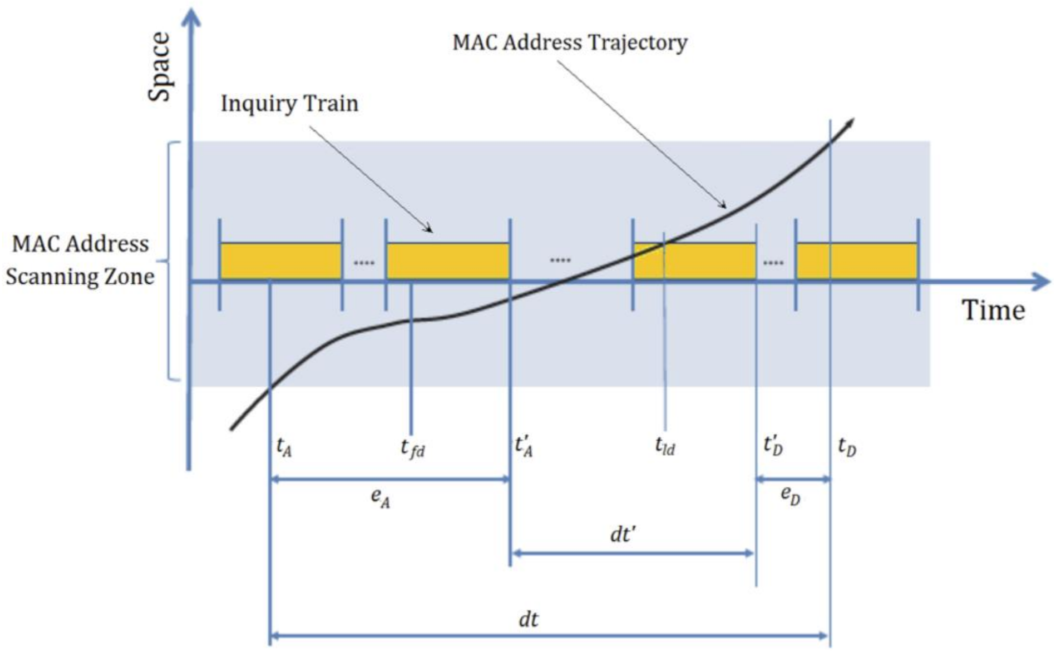

Fig. 10. Time-Space trajectory plot for a MAC ID device through a MAC Address scanning zone

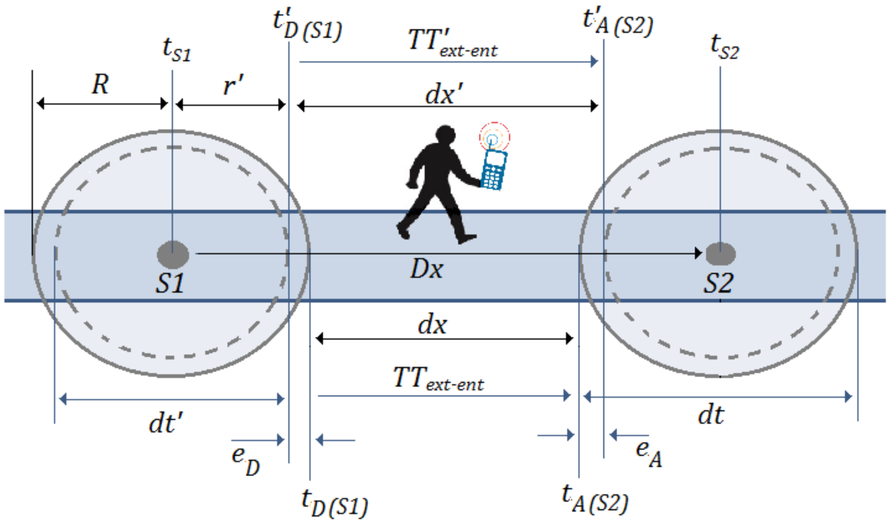

Fig. 11. Systematic illustration of estimating pedestrian and cyclist's Travel-Time

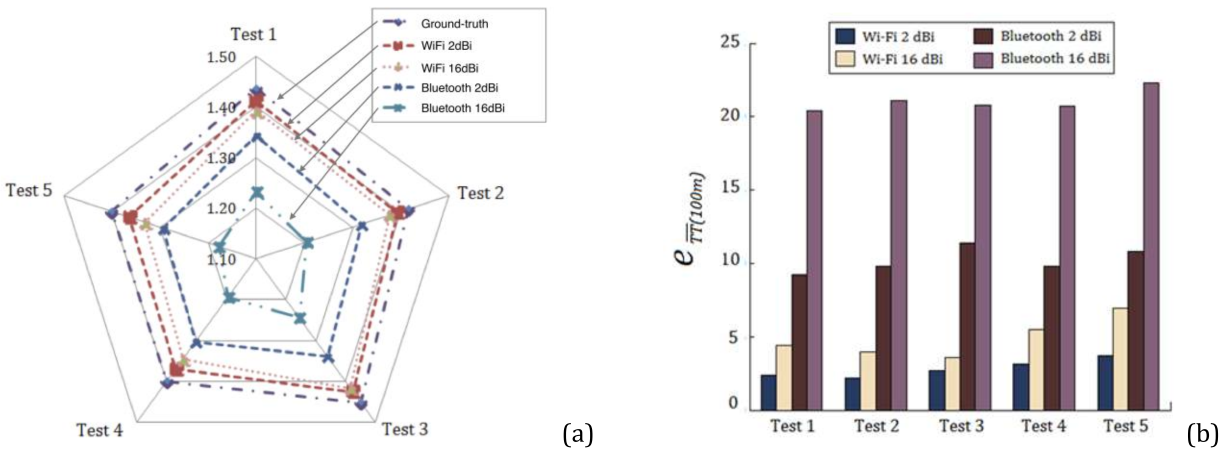

Fig. 12. (a) The test scenarios' average speed for real and scanners' samples and (b) the estimated difference of scanners' travel-time from real travel-time every 100 *m*

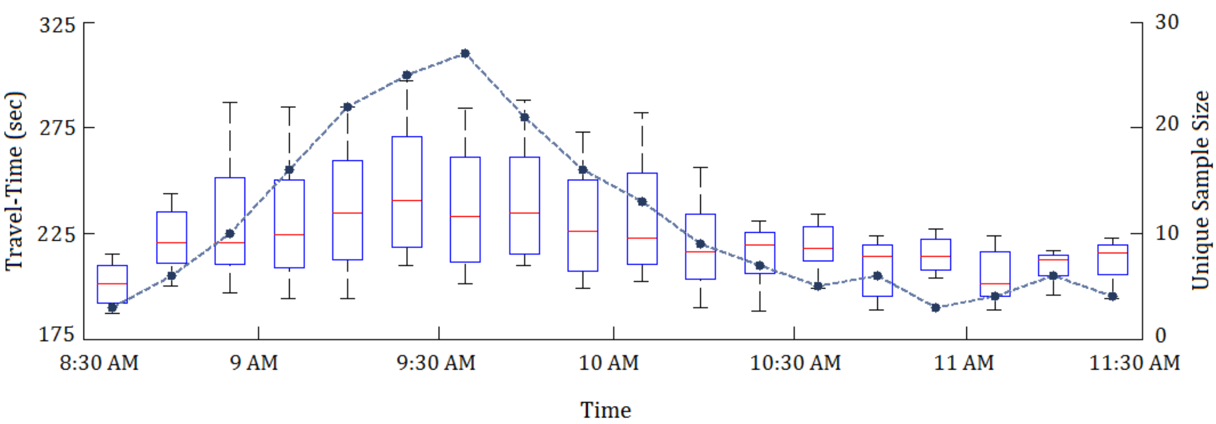

Fig. 13. The travel-times (box-plot) and sample size (line-graph) of walkers travelling from point *S1* to *S2* (inbound) based on *16dBi* scanners' dataset

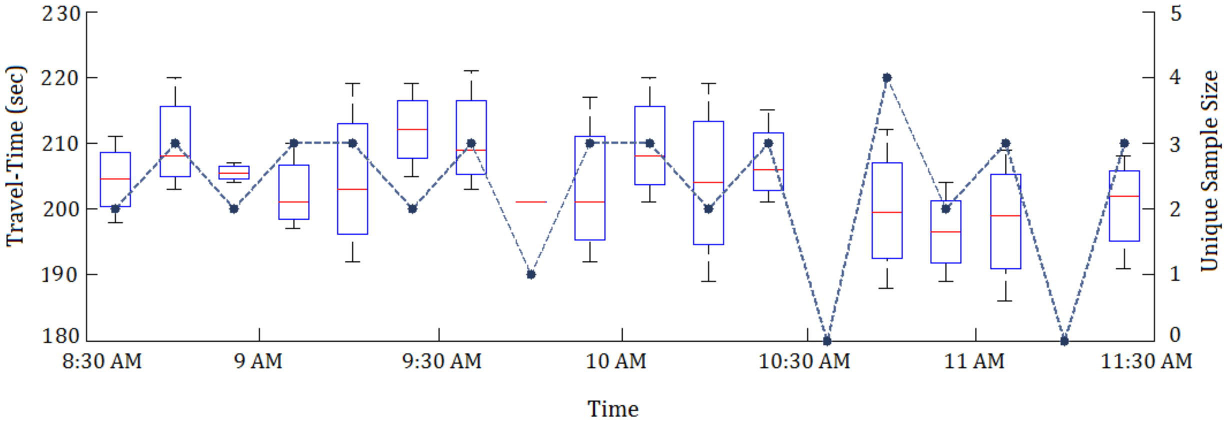

Fig. 14. The travel-times (box-plot) and sample size (line-graph) of walkers travelling from point *S2* to *S1* (outbound) based on *16dBi* scanners' dataset

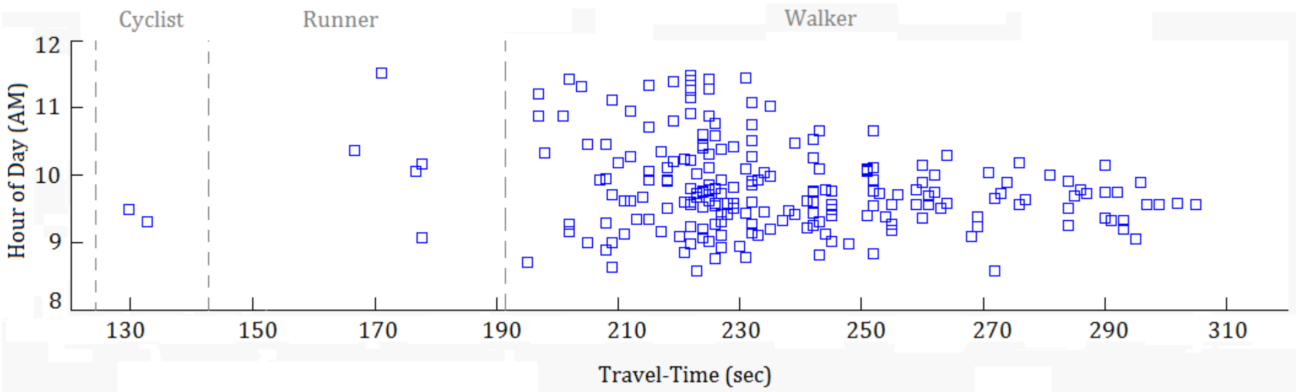

Fig. 15. Inbound (*S1* to *S2*) travel-time values of *2dBi* scanners

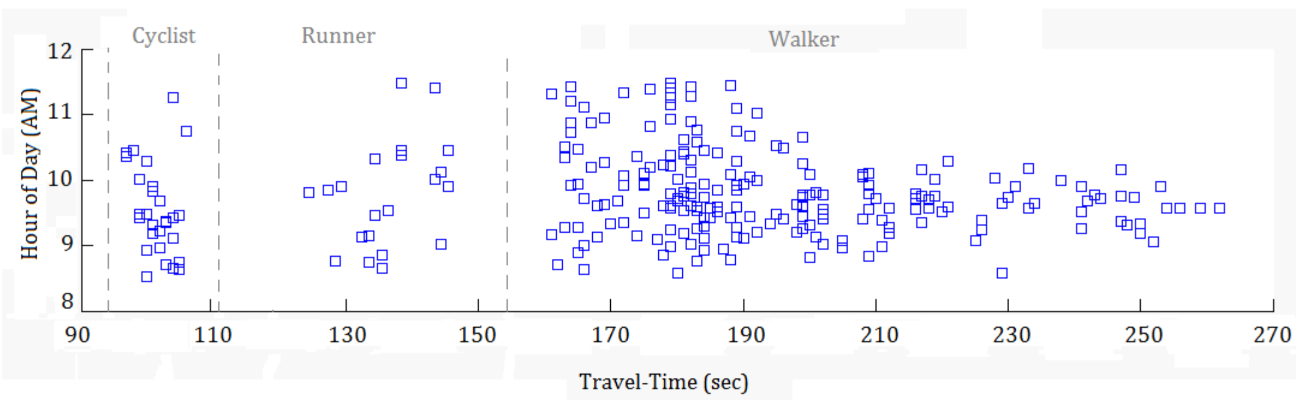

Fig. 16. Inbound (*S2* to *S1*) travel-time values of *16dBi* scanners

Table

Tables

Table 1. Antenna detection range for Bluetooth and Wi-Fi

Antenna Gain	Wi-Fi (Radius)	Bluetooth (Radius)
2 dBi	85 m	55 m
3 dBi	100 m	85 m
5 dBi	130 m	100 m
7 dBi	140 m	110 m
10 dBi	145 m	120 m
16 dBi	150 m	130 m

Table 2. The number of pedestrians, runners, and cyclists travelled from point S1 to S2 (inbound)

Dataset	Total	Walker	Runner	Cyclists	Non-active Traveller
Real Data	2,439	2,021	85	322	-
2 dBi Scanner	242	203	5	2	32
16 dBi Scanner	284	203	21	28	32

Table 3. The number of pedestrians, runners, and cyclists travelled from point S2 to S1 (outbound)

Dataset	Total	Walker	Runner	Cyclists	Non-active Traveller
Real Data	636	546	23	67	-
2 dBi Scanner	55	42	1	0	12
16 dBi Scanner	58	42	3	1	12

Table 4. List of symbols and their meanings

Symbol	**Meaning**
R	*Estimated radius of scanning zone*
r'	*Accrual radius of scanning zone*
t_A	*Actual arrival time*
t'_A	*Arrival time based on MAC address data*
t_D	*Actual departure time*
t'_D	*Departure time based on MAC address data*
$TT_{ext\ \ ent}$	*Actual exit-to-entre Travel-Time*
$TT'_{ext\ \ ent}$	*Reported exit-to-enter Travel-Time*
Dx	*Distance between scanning points*
dx	*Actual distance between scanning zones*
dx'	*Reported distance between scanning zones*
dt	*Actual duration in scanning point*
dt'	*Reported duration in scanning point*
S_1	*Scanning point 1*
S_2	*Scanning point 2*
e_A	*Temporal error of MAC address scanner in reporting the arrival time*
e_D	*Temporal error of MAC address scanner in reporting the departure time*
t_{fd}	*The time when MAC address is first time is discovered*
t_{ld}	*The time when MAC address is last time is discovered*

Table 5. WiFi and Bluetooth average travel-time based on 5 test scenarios

Gain	Travel-Time (sec)	Walker	Runner	Cyclists
-	$\overline{D}t_{Ground\ truth}$	278	215	186
2 dBi	$\overline{D}t_{WiFi}$	226	164	132
	$\overline{D}t_{Bluetooth}$	258	198	152
16 dBi	$\overline{D}t_{WiFi}$	178	127	95
	$\overline{D}t_{Bluetooth}$	219	175	122