Australasian Transport Research Forum 2013 Proceedings
2 - 4 October 2013, Brisbane, Australia
Publication website: http://www.patrec.org/atrf.aspx

Bluetooth and Wi-Fi MAC Address Based Crowd Data Collection and Monitoring: Benefits, Challenges and Enhancement

Naeim Abedi, Ashish Bhaskar, Edward Chung

2 George St., Queensland University of Technology

Brisbane, Queensland 400, Australia

Email for correspondence: naeim.abedi@student.qut.edu.au Reference Number: 318

Abstract

This paper firstly presents the benefits and critical challenges on the use of Bluetooth and Wi-Fi for crowd data collection and monitoring. The major challenges include antenna characteristics, environment's complexity and scanning features. Wi-Fi and Bluetooth are compared in this paper in terms of architecture, discovery time, popularity of use and signal strength. Type of antennas used and the environment's complexity such as trees for outdoor and partitions for indoor spaces highly affect the scanning range. The aforementioned challenges are empirically evaluated by "real" experiments using Bluetooth and Wi-Fi Scanners. The issues related to the antenna characteristics are also highlighted by experimenting with different antenna types. Novel scanning approaches including Overlapped Zones and Single Point Multi-Range detection methods will be then presented and verified by real-world tests. These novel techniques will be applied for location identification of the MAC IDs captured that can extract more information about people movement dynamics.

1. Introduction

In the past decade, customer research at social events has become an interesting topic for organising companies. Customers' movement and motivation understanding has been always a goal of marketing research and allows to develop new business strategies. It delivers key information about customers of a company and is also useful for the minimisation of the investment risk. The robust passive and active positioning technologies have motivated the development of sensors which have the capability of people's movement monitoring. Human movement behaviour analysis has received lots of notice particularly in the field of visual analytics (Andrienko & Andrienko, 2007) (Andrienko et al., 2007). In terms of visualisation of human movement, static maps with directed linear symbols (Vasiliev, 1996), space time cubes (Hgerstraand, 1970) and animated maps (Andrienko et al., 2000) are the most common techniques. The links between locations can be analysed when movement data is added into the flow maps and original destination matrices (Doantam, 2005). Sensors and analytical frameworks have been developed to allow for fine-grained locating and tracking moving objects (Hallberg et al., 2003) (Hay & Harle, 2009).

Surveys and video surveillance are the most common methods for customers and people monitoring. However, high cost and also hardly representing of surveys because of a non-random sampling process are always a problem. Video processing is also depended on weather conditions, density and brightness of crowd (Liebig et al., 2012). Increasing the popularity of cell phones has motivated researchers to collect crowd data based on recording people's mobile phones. Positioning the cell-phones based on Global System for Mobile (GSM) communication was proposed as the first popular method but it has become less applicable especially due to the privacy objection (Giannotti & Pedreschi, 2008).

Nowadays, majority of smart phones, laptops, and portable electronics devices use wireless communication, especially Bluetooth and Wi-Fi. The presence of Bluetooth and Wi-Fi networks in offices, buildings and campuses (Bisdikian, 2001) (The Bluetooth Special Interest Group, 2001) (Harwood, 2009) (Bray & Sturman, 2001) have been increased because of their wide availability on a huge number of personal portable electronic devices. Collection data from capturing wireless technologies such as Bluetooth and Wi-Fi which communicate based on MAC address standards have been recently applied successfully.

Because MAC address data allows for unannounced, non-participatory, and simultaneous tracking of people. Also, tracking individual in this method remains unknown 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. The traceability feature of MAC addresses has been used for crowd travel time estimation and also can be applied for other purpose such as the evaluation of space utilisation. This type of data collection can be complemented by other crowd data collection technique, such as surveillance cameras to improve crowd monitoring systems.

This paper aims to present the significant benefits and challenges of Wi-Fi and Bluetooth data for analysis of spatiotemporal dynamic of people movement. The previous research has been presented in section 2. Section 3 explains MAC address technology as the tracking tool for human movement analysis. Bluetooth and Wi-Fi technologies are overviewed and their performance in people tracking is compared. One of disadvantages of MAC address tracking is that the location of MAC device in detection zone is unknown. Section 4 presents two scanning approaches for estimation of MAC device's location. In section 5, the experimental results of mentioned scanning approaches are provided and analysed.

2. Previous Research

Recent studies have been done on the analysis of people's travelling behaviour in the tourism industry (Jankowski et al., 2010) and pedestrian's density distribution during seasons (Anderienko et al., 2009) for example. With increasing the popularity of using mobile devices, new techniques have been presented for analysis of massive distributed movement data (Jankowski et al., 2010) (Anderienko & Anderienko, 2007). Tracking mobile phones and intercoms have been recently noticed as a crowd data collection and monitoring system (Liebig et al., 2012) (Stang et al., 2011).

Bluetooth technology has recently become to an emerging tool for monitoring purposes (Stang et al., 2011). Some studies have been done on recording flows of outdoor movements using Bluetooth. Weghe et al. (2012) studied the applicability of Bluetooth technology for collecting movement data for big events. Leitinger et al. (2010) developed a Bluetooth-based mobility sensor for event monitoring at Szinger festival in Budapest. They placed a mesh of six sensors at selected locations with distance from 50 to 200 meters. Their work extracted the number of people with their route choice at specific locations. Pels et al. in 2005 implemented various scanners at Dutch train stations for capturing transit travellers [21]. Hagemann and Weinzerl in 2008 analysed the transit travellers and also tracked public busses by locating sensors inside the buses. Hallberg et al. (2003) presented an accurate positioning system by Bluetooth devices. Their work however focused on only the development of a Bluetooth-based locating system and is not effective when people are using their personal Bluetooth devices. Stange et al. (2011) also used Bluetooth tracking system for monitoring visitors with extracting their pathway choice. Vu et al. (2010) presented a joint Bluetooth/WiFi scanning framework for assessment of the location popularity and people time spending in a university campus area.

For crowd and people MAC address data collection purposes, the data collection process is more significant comparing with the traffic MAC address data collection. The reason is because of smaller size of the space. For instance, a highway or city road may be proposed for Bluetooth data collection that Omni-directional detection zones with the radius of 100 meters is perfect. For Bluetooth data collection from people in a small shopping centre, not only deciding what rages of detection zone must be used but also defining the number and location of scanners are important for optimal data collection.

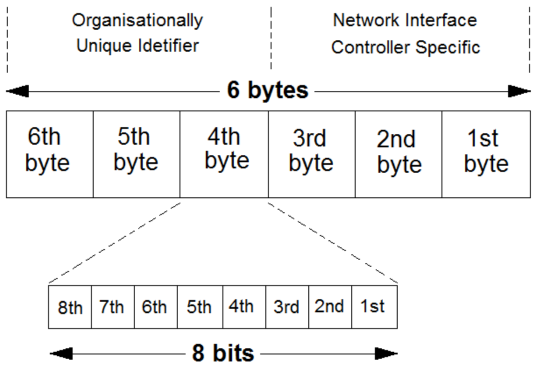
3. MAC Addresses as a Tracking Technology

3.1 Overview

In order to access networks and services with higher flexibility and mobility, wireless networks are a popular and fast-growing technology (Hussain & Wee-Seng, 2007). The benefits of wireless are reducing the cable restrictions, low cost, dynamic communication formation, and easy deployment. Bluetooth, Wi-Fi, ZigBee, and UWB are four short range wireless standards that respectively correspond IEEE 802.15.1, 802.11 a/b/g, 802.15.4, and 802.15.3. In fact, IEEE defines the MAC address and Physical Layers for mentioned wireless protocols for an operate range of 10 to 100 meters. Bluetooth and ZigBee are most efficient in terms of power consumption and UWB and Wi-Fi consume less normalized energy. Furthermore, ZigBee and Bluetooth have bigger transmission time and data coding efficiency associated to the data payload size (Porter et al., 2011).

MAC addresses are unique identifies and are used for various type of communication networks and most of IEEE 802 network technologies. Figure 1 shows the structure of a MAC address which consists six bytes. Bluetooth and Wi-Fi based networks are identifies by MAC addresses. Nowadays, majority of smart-phones and digital devices use Bluetooth and Wi-Fi technologies for communication. Because a Bluetooth or Wi-Fi device is defined by a unique code, it can be tracked and this feature has been a motivation for various applications and data collection purposes. Capturing Bluetooth devices in motorways and city roads has been recently used for vehicle's travel time estimation purposes.

Figure 1. A MAC address Architecture



Several factors may effect on the quality of MAC address data collection process that may be associated with the hardware and software implemented. Antenna type is one of those factors. For optimising the performance of MAC address scanning section, the antenna characteristics such as gain and polarization must be matched to specific applications. Porter et al. categorised six different antennas in (Bruno & Delmastro, 2003) for assessing their capability and suitability in the Bluetooth data collection process. They evaluated the antennas' performance for Bluetooth traffic data collection. Their study shows that vertically polarized antennas with gains from 9 to 12 dBi are suitable for a Bluetooth based traffic data collection. They also mentioned that the circular polarized antennas do not significantly improve the data collection process.

In the case of Bluetooth and Wi-Fi communications, the first step is scanning available Bluetooth or Wi-Fi devices in the detection zone. Various types of devices are designed that are only used for scanning Bluetooth, Wi-Fi or both. These MAC address scanners record the captured MAC IDs based the synchronized clock. However, there can be some limitations associated to MAC address scanners such as scanning frequency and maximum number of ID capturing in a same time frame.

3.2 Bluetooth Architecture and Scanning Protocol

Bluetooth (The Bluetooth Special Interest Group, 2001) (Bray & Sturman, 2001) known as IEEE 802.15.1 is actually designed for short wireless range and mostly cheap devices in order to remove cables for peripheral computer data transmission such as mouse, printer, keyboard, and printer. It is also very popular for short range transmission of data between electronic devices. Piconet and Scatternet are two connectivity topologies of Bluetooth technology. The Piconet is a Wireless Personal Area Network (WPAN) formed by the Bluetooth device and each Piconet is defined by a Frequency Hopping Channel based on the master's address (Porter et al., 2011). It is the Bluetooth network's building block that it is a small cluster of devices. Piconets indeed share a common physical channel and are synchronised to the same timeframe. They actually adopt the same Inquiry Hopping Sequence. Based on the Piconet's structure, one device plays the master role and all other devices assume the role of slave. The slave devices then drive the sequence of channel-hopping according to the function of the master's clock and address. The Time Division Duplex (TDD) scheme is the transmission scheme adopted in the piconet-based (Hussain & Wee-Seng, 2007). A set of operational Bluetooth piconets overlapping in a period of time is called Scatternet. Two Piconets can actually form a Scatternet (Porter et al., 2011).

Two main phases are required for the establishment of a Bluetooth-based connection. Inquiry is the first phase and it in fact allows the inquirer to discover the possible slaves' identity. The second and final phase is named Page. This phase indeed corresponds to initial connection setup. In this phase the pager informs the paged unit regarding its identification status and defines its clock as the main clock. The Bluetooth devices are synchronised based on the inquirer/pager's (master) clock. The connection is then created and the devices begin the process of exchanging data (Bray & Sturman, 2001) (Hussain & Wee-Seng, 2007) (Hallberg et al., 2003) (Bisdikian, 2001).

For Bluetooth MAC address based data collection, only the discovery section is needed and Bluetooth MAC address scanners never made a full connection with available Bluetooth device available. Inquiry and Inquiry Scan are the main parts of the Bluetooth device discovery protocol. The Inquiry part is run by a discovery device or Master and the Inquiry Scan part is run by a device willing to be discovered or slave (The Bluetooth Special Interest Group, 2001). The Inquiry part includes Standby/Connection and Inquiry states. A device can be initially in Standby or Connection mode which is basically the Link Manager Protocol (LMP) layer. LMP layer in fact decides when the baseband layer my initiate the Inquiry part. When the device is in Inquiry state, two types of state transition occur that are Tx (transmission) and Rx (reception) slots. Tx and Rx frequencies of a discovering device are defined based on the Inquiry Hopping Sequence which contains the set of 32 distinct hope frequencies and is generated by discovering device's native clock and General Inquiry Access Code (GIAC). The set of 32 frequencies also is partitioned into two subsets of 16 frequencies. It takes 8 Tx slots for transmission of a whole subset of 16 frequencies and each time slot is 625 us. Because of the interleaved feature of Tx and Rx slots, the total duration for covering each subset of 16 frequencies is 16 slots or 10 ms (=16 x 625 us). Each subset is repeated 256 times before the other subset is used and 4 subsets must be used for collection of all responses in an errorfree manner. The Bluetooth discovery stage therefore lasts for 10.24 s (=256 x 4 x 10 ms) (Chakraborty et al., 2010). The Bluetooth discovery time is investigated empirically and presented in the Bluetooth and Wi-Fi comparison section.

3.3 Wi-Fi Architecture and Scanning Protocol

Wireless Fidelity (Wi-Fi) also known IEEE 802.11 (Gast, 2005) (Arbaugh, 2002) (Crow et al., 1997) (Ferro & Potorti, 2005) is designed for wireless local area network connections. Its purpose is to provide wireless connectivity for the devices such as cell-phones and PDAs which need quick installation. It defines MAC address for access to the physical layer. Users can surf the internet by Wi-Fi at broadband speeds through connection to an Access Point (AP) or ad hoc mode. Wi-Fi architecture includes several components which interact to provide a wireless connection. Basic Service Set (BSS) is the basic cell of a Wi-Fi network. BSS is a collection of mobile or fixed stations and a station cannot directly communicate with other stations of BSS if it moves out of its BSS. Coordination function is a set of rules which control the access to the transmission medium. Wi-Fi defies two main functions; Distributed Coordination Function (DCF) and Point Coordination Function (PCF). Independent BSS (IBSS) is the simplest Wi-Fi network. DFC is the fundamental of Wi-Fi MAC protocol that a Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) channel access technique (Porter et al., 2011) (Harwood, 2009).

Wi-Fi has two modes, Active and Passive, and uses Scan, Authentication, and Association procedures. The scan procedure is applied for discovering MAC addresses and other Wi-Fi devices' parameters in the terminal's coverage area whether Wi-Fi is in the Active or Passive mode (Ferro & Potorti, 2005). Once a device has found an access point, it must operate Authentication with AP and then Association procedure. Since the Association phase has made with AP, the device can communicate with stations in other BSSes.

Same as Bluetooth MAC address scanners, Wi-Fi scanners also only operate the discovery section and never create a connection with around Wi-Fi devices. The discovery time for Wi-Fi is very less than Bluetooth discovery while both operate in a same radio frequency band. The reason is because of their difference in connection architecture. Han and Srinivasan in (2012) pointed that discovering time of Wi-Fi addresses is dependant to the environment type. They noted that the Wi-Fi discovery time in a park, home and office is respectively measured 0.52, 0.87 and 1.07 seconds. The Wi-Fi discovery time is investigated empirically and presented in the next section.

3.4 Bluetooth and Wi-Fi Comparison

3.4.1 Architecture

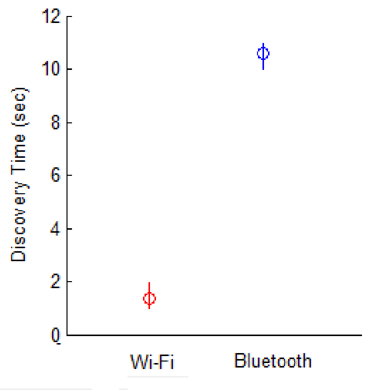
Comparing Wi-Fi and Bluetooth technologies, both handle the traffic by a central unit called AP in Wi-Fi and Master in Bluetooth. The AP and Master are responsible for routing packets between devices. Bluetooth has maximum 7 slave units and Wi-Fi has 2007. Also, the nominal range of Wi-Fi is 35 meters indoors and up to 100 meters outdoors and Bluetooth range is 10 meters for mobile devices. However, the range varies based on a function of transmission power and environment's complexity. A LAN Access Profile defines the connection with external networks for Bluetooth but a Wi-Fi AP is capable to perform as a bridge structurally. Both Wi-Fi IBSS and Bluetooth Scantternet need a global addressing procedure and a routing mechanism for ensuring stations global connectivity. A global addressing exists in Wi-Fi but Bluetooth does not provide any global addressing and it should be provided by the upper layer protocols such as the IP level. The major problems are in the node joining process and linking breaks caused by moving obstacles and terminals. These events do not make any changes in the structure of a Wi-Fi ad hoc network, whereas a modification is needed in the Scatternet structure for reorganising the underlying Piconets (Ferro & Potorti, 2004).

Due to switching the Bluetooth data signals based on the Frequency Hopping Spread Spectrum (FHSS) between RF bands, they have a strong resistance to the environmental factors and interference. In Direct Sequence Spread Spectrum (DSSS) technology, the system sends many redundant copies of data and a single copy is only needed for having full data transmission. DSSS then can minimise the interference and background noise effects. Comparing DSSS with FHSS, DSSS has better signal delivery and security whereas it is a sensitive technology to many environmental conditions (Harwood, 2009).

3.4.2 Discovery Time

MAC address discovery time is important in terms of collecting efficient data a short time period. As mentioned in the previous section, Bluetooth discovery time is 10.21 seconds whereas Wi-Fi discovery time is around 1 second. Wi-Fi and Bluetooth discovery times are also measured experimentally based of over 1000 records and are shown in Figure 2. Indeed, a specific Wi-Fi and Bluetooth ID were recorded separately for about two hours by a Wi-Fi and Bluetooth scanner. Based on the experimental results, Wi-Fi discovery time is 1.365 seconds in average and Bluetooth is almost 10.577 seconds. As the result, the data collection rate of Wi-Fi MAC address is almost 10 times theoretically and 8 times empirically bigger than Bluetooth.

Figure 2. Experimentally measurement of Bluetooth and Wi-Fi discovery time



3.4.3 Popularity of Use

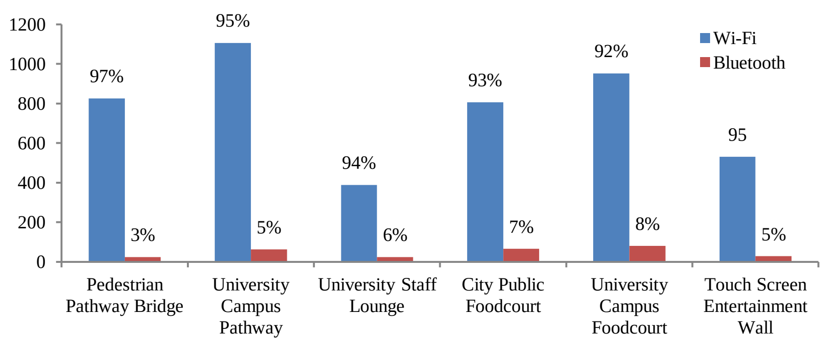
Another issue that is important in MAC address-based crowd data collection is the popularity of using MAC addresses. In terms of assessing the popularity of using Wi-Fi and Bluetooth for crowd data collection purposes, a Wi-Fi and Bluetooth scanner with 5 dBi antenna gain were placed in six different locations categorized into:

- **Pedestrian pathway**
    - ***Pedestrian pathway bridge (no Wi-Fi provided)*** – *Goodwill Bridge, Brisbane, Australia Between 10 AM and 12 PM on 7th of June 2013*
    - ***University campus pathway (Wi-Fi area)*** – *Queensland University of Technology (QUT) Gardens Point campus, Brisbane, Australia, Between 1 PM and 3 PM on 7th of June 2013*
- **Food court area**
    - ***City food court (no Wi-Fi provided)*** – *Myer Centre food court, Brisbane, Australia, Between 11:30 AM and 1:30 PM on 6th of June 2013*

  - ***University campus food court (Wi-Fi area)*** *– Level 3 P Block, QUT GP campus, Brisbane, Australia, Between 11:30 AM and 1:30 PM on 4th of June 2013*
- **Office area**
  - ***University staff lounge (Wi-Fi provided) –*** *Level 7 S block, QUT GP campus, Brisbane, Australia, Between 11:30 AM and 1:30 PM on 4th of June 2013*
- **Entertainment facility area**
  - ***Touch screen entertainment wall (Wi-Fi provided) –*** *Level 4 P Block, QUT GP campus, Brisbane, Australia, Between 11 AM and 1PM on 3rd of June 2013*

For the two first mentioned categories, two places were selected for data collection, one place with no free Wi-Fi network and a free Wi-Fi network access was provided in another place. The purpose was to experimentally assess if a free Wi-Fi coverage can effect on an increase in the observation rate of unique Wi-Fi addresses or not. Figure 3 shows the observation rate of unique Wi-Fi and Bluetooth MAC addresses in the mentioned places during pick periods for about two hours. In fact, the duplicated MAC addresses were removed and only unique IDs were considered. The results indicate that more than 90% of all scanned unique MAC addresses in all places were Wi-Fi addresses and the popularity of using Wi-Fi devices is therefore significantly more than Bluetooth ones. In another word, the likelihood of collecting efficient Wi-Fi devices is much more that Bluetooth. Also, the findings show that a free Wi-Fi network can effect on the observation rate of unique Wi-Fi IDs, however the effect is not noticeable.

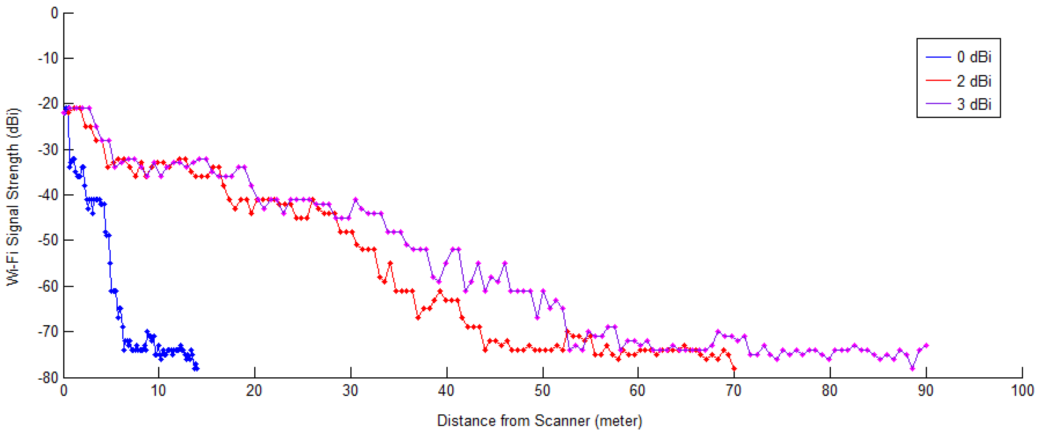
Figure 3. Experimentally assessment of Bluetooth and Wi-Fi popularity in different places



3.4.3 Signal Strength

The Wi-Fi and Bluetooth signal strength can be measured and its unit is dBm. Environment's type, distance and antenna's gain are the most important factors effecting on the signal strength value. If the impact of environment and antenna on signal strength were assessed, the distance between scanner and MAC devices can be approximately estimated based on the value of signal strength. Therefore, this feature can be used for localising the MAC addresses for crowd data collection aims. Dimitrova et al. (2012) experimental analysed Wi-Fi and Bluetooth signal strength and noted that the signal strength changes less between sensors of the same type than between mobile devices from different companies. The scanner used for experiments was only able to measure Wi-Fi signal strength. Figure 4 shows that the value of Wi-Fi signal strength is decreasing when the device is moving far from the scanner while the antenna type and environment's type is constant. The experiment was done for 3 different antenna gains in an open space environment in QUT Oval court situated in Kelvin Grove campus, Brisbane. The results show that the signal strength value indeed dropped significantly from zero distance to half the detection range and did not changed considerably until the end of the detection range. This indicates that the signal strength value can be used for MAC address location identification when their distance from scanner is half of detection area. In another word, if the radius of a scanner is R, the distance of device from scanner can be estimated by signal strength until R/2.

Figure 4. Experimentally Evaluation of Wi-Fi signal strength



3.4 Effects of Environmental Complexity on MAC ID Scanning

The interference of environment's obstacle on the wireless communication is a significant issue for setting a powerful and optimised wireless communication network. Some factors that cause considerable interference are (Harwood, 2009):

Table 1. Obstacle severity on wireless signals (Harwood, 2009)

Obstruction	Obstacle Severity	Sample Use
Wood	Low	Inside a wall or hollow door
Drywall	Low	Inside walls
Furniture	Low	Couches or office partitions
Clear glass	Low	Windows
Tinted glass	Medium	Windows
People	Medium	High-volume traffic areas that have considerable pedestrian traffic
Ceramic tile	Medium	Walls
Concrete blocks	Medium/high	Outer wall construction
Mirrors	High	Mirror or reflective glass
Metals	High	Metal office partitions, doors, metal office furniture
Water	High	Aquariums, rain, fountains

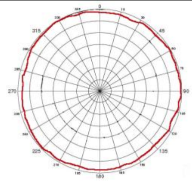
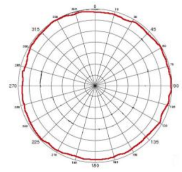
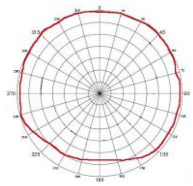
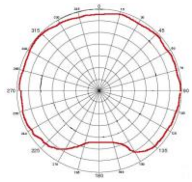
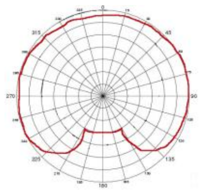
- **Physical objects** *(such as Trees, masonry, buildings, and other physical structures)*
- **Radio Frequency (RF) interference** *(such as microwave and cordless phones)*
- **Electronics Device interference** *(computers, refrigerators, fans and lighting fixtures)*
- **Environmental factors** *(such as weather conditions, fog, and lighting)*

While outdoor interference such as weather condition is not a serious problem, there are plenty of wireless obstacles in indoor spaces such as offices and homes. Table 1 presented the obstacle severity on wireless communication.

4. MAC Address Scanning Techniques

One of the important issues that can enhance the efficiency of MAC address based crowd data collection is to define the location of captured MAC IDs. A Wi-Fi or Bluetooth MAC address scanner only records the IDs based on the local time and date. This section present some approaches for approximately identifying the location of captured MAC IDs.

Table 2. Antenna detection range for Bluetooth and Wi-Fi

Antenna Picture	Gain	Wi-Fi (Radius)	Bluetooth (Radius)	Horizontal Phase Plane
	2 dBi	70 m	45 m	
	3 dBi	90 m	75 m	
	5 dBi	140 m	100 m	
	7 dBi	150 m	110 m	
	10 dBi	150 m	120 m	

4.1 Antenna Detection Range Assessment

One of the primary stages in MAC address based data collection is to understand scanning equipment, especially antenna's type and detection range. Wi-Fi 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.

Comparing to MAC address data collection for transport purposes, the role of antenna characteristics is more significant in crowd data collection field. The most important reason is the effect of environment's complexity and interference on the discovery range. 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 and so on. 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). Therefore, a detail assessment of antenna's detection range is required. The proposed scanning approaches in the following sections are dependent on how accurate the detection range of antennas is assessed. Table 2 shows the experimental results for five different antenna's gains defined by their producers. The experimental results indicate that the Bluetooth discovery range in all gains is less than Wi-Fi detection range. Also, there is not a noticeable difference for discovery range in higher gains for both Bluetooth and Wi-Fi.

4.2. Multi-Range Scanning Technique

In this technique, as shown in Figure 5, the MAC addresses will be recorded by different range of antennas from one point in a same time. Indeed, the whole detection zone is divided into several regions. This approach is effective to estimate the distance of captured MAC IDs from the scanning point. In Figure 5 for example, three antennas in different detection ranges are employed for scanning MAC IDs. Based on Table 3, MAC IDs captured by all scanners in a same time are in zone #1, those are only recorded by scanner #2 and #3 are in zone #2, and those just observed by zone #3 are located in region #3.

Figure 5. Multi-range scanning approach using three different ranges of detection

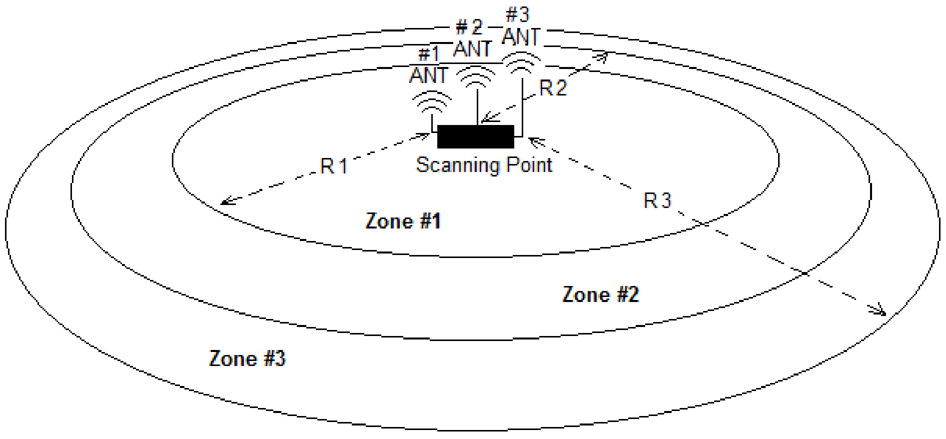


Table 3. Multi-range approach table

	Zone #1	Zone #2	Zone #3	. . .	Zone #n
Scanner #1	✓	×	×	. . .	×
Scanner #2	✓	✓	×	. . .	×
Scanner #3	✓	✓	✓	. . .	×
.	.	.	.		.
.	.	.	.		.
.	.	.	.		.
Scanner #n	✓	✓	✓	. . .	✓

4.3. Overlapping Detection Zones Technique

Overlapping the detection zones of scanners is another scanning technique to identify the location of MAC addresses. In this scanning strategy, the scanning location is divided into more scanning zones. As can be seen from Figure 6, the data collection location is broken down into seven zones while three MAC address scanners are employed. Table 4 shows how to identify the location of MAC addresses using three scanners. For instance, if a MAC ID is captured by all three scanners, it is in Zone #1.

Figure 6. Overlapping detection zones approach using three scanner

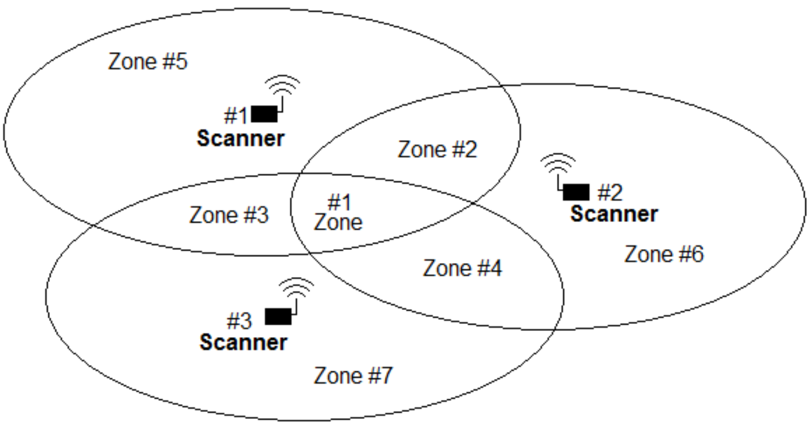


Table 4. Overlapping detection zones approach table for three scanners

	Zone #1	Zone #2	Zone #3	Zone #4	Zone #5	Zone #6	Zone #7
Scanner #1	✓	✓	✓	×	✓	×	×
Scanner #2	✓	✓	×	✓	×	✓	×
Scanner #3	✓	×	✓	✓	×	×	✓

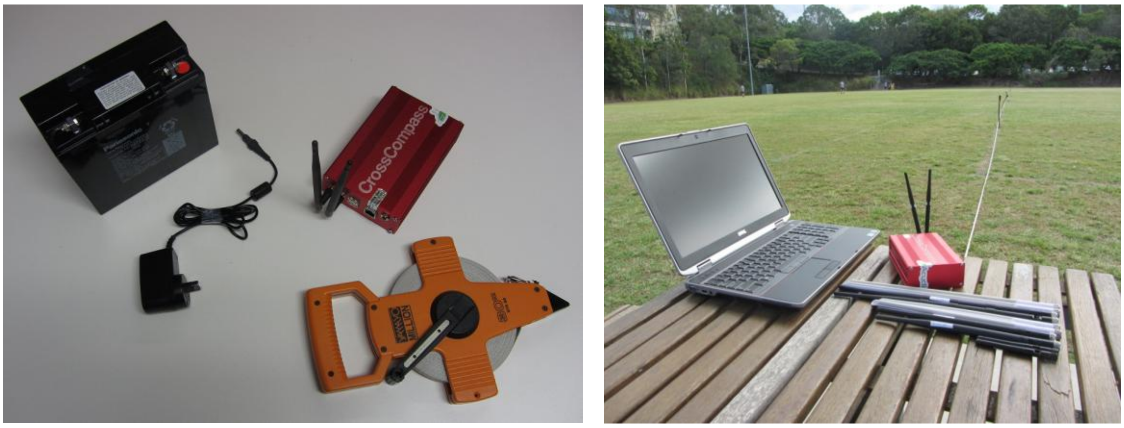
5. Experimental Results and Scanning Methods Validation

In this section, the empirical results of several real-world experiments are presented. The first experiments are done for assessment of antennas' detection range and scanner's capability in detecting Bluetooth and Wi-Fi addresses.

5.1 Equipment Used

For capturing MAC addresses, a Wi-Fi/ Bluetooth scanner called CrossCompass manufactured by Acyclica Inc with the capability of scanning both Bluetooth and Wi-Fi addresses in the same time and also the record time synchronisation option with GPS or PC clock. In order to have the minimum range of detection, no antenna was plugged to the scanners. The scanners used for the experiments can scan Wi-Fi devices up to 15 meters and Bluetooth devices up to 10 meters without using any external antenna. This scanning range is actually because of the antenna connectors' gain. A stopwatch was employed to collect data for the validation of experiments' result. All empirical tests were done in an outdoor environment with minimum environmental interference. Figure 7 shows the picture of equipment used for doing and validating the real-world experiments.

Figure 7. The equipment used and environment for the real-world experiments



5.2. Validation Framework

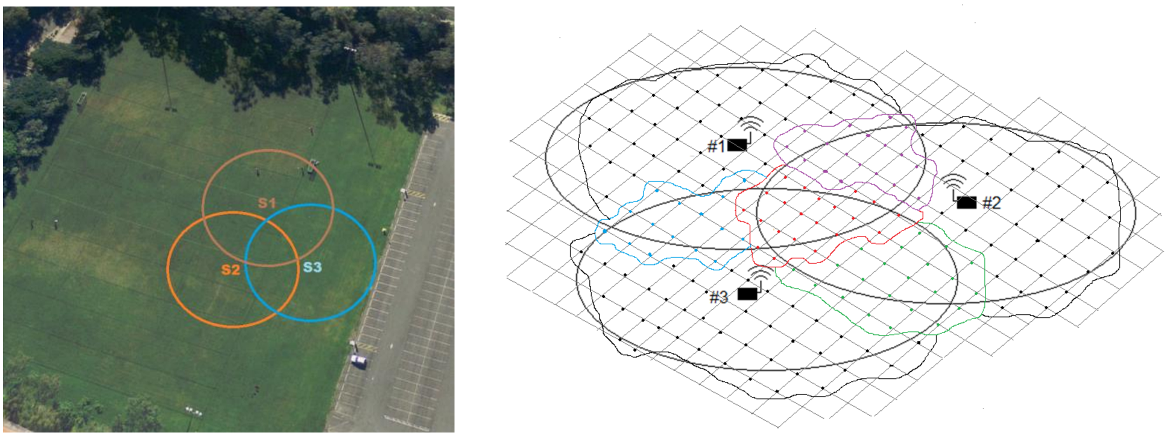
For validation of the real-world experiments, the locations of MAC device were recorded based on a synchronised time with MAC scanners' clock. For running the experiments, the scanners are fixed in determined locations and Wi-Fi and Bluetooth devices are moved to the different locations of experiment space. In order to have more accurate data for validation purposes, the experiment place was marked every couple of meters and a stopwatch recorded the time when MAC devices were been in each marked location.

5.3. Experimental Results and Analysis

In this section, the results of scanned MAC IDs based on the overlapping detection method are compared by the recorded data in three rounds. The multi-range scanning technique is validated by the antennas' detection range measured experimentally and provided in Table 1. Figure 9 shows the scanners set up and the locations where MAC device was moved in different locations. The scanners were positioned 20 meters far from each other because their range of detection is from 12 to 14 meters and the MAC device was stoped around 10 seconds in each marked locations.

The red points in the figure are the places where the Wi-Fi device was scanned by all three scanners. The purple points indicated the locations that both scanner 1 and 2 captured the Wi-Fi device. The locations were highlighted by green are the places where the Wi-Fi ID was scanned by scanners 2 and 3. The blue points are also the results of common record between scanners 1 and 3. Finally, the black points present the locations only observed by each scanner. The results show that the area is divided into 7 detection zones while 3 scanners were employed. The size and shape of the zones are not exactly same as the theoretical ones. However, the all 7 zones were created and tested successfully. Therefore, overlapping method can be applied for crowd data collection in order to approximately identify the location of if MAC devices.

Figure 9. Overlapping detection zones approach setup and results



This scanning approach is next applied in Goodwill Bridge on 7th of June 2013 between 10:30 AM and 12:30 PM in order to identify spatiotemporal location of pedestrians and cyclists passing through this pedestrian bridge. Antenna gain of 5 dBi is used for both Wi-Fi scanners. Based on the experimental results of antenna characteristics, each scanner had around 150 meters detection zone in radius. Figure 10 shows the area map of experiment location and the position of each scanner.

After 1 hour data collection by both scanners simultaneously, only unique Wi-Fi IDs that were scanned by both scanners were used for data analysis. Also, some mobile devices with Wi-Fi in on mode were moved through the bridge for validation. Figure 11 presents the scanning results for each scanner. The scanning area was in fact divided into 3 detectable zones.

Figure 10. Area map of overlapping detection approach implementation

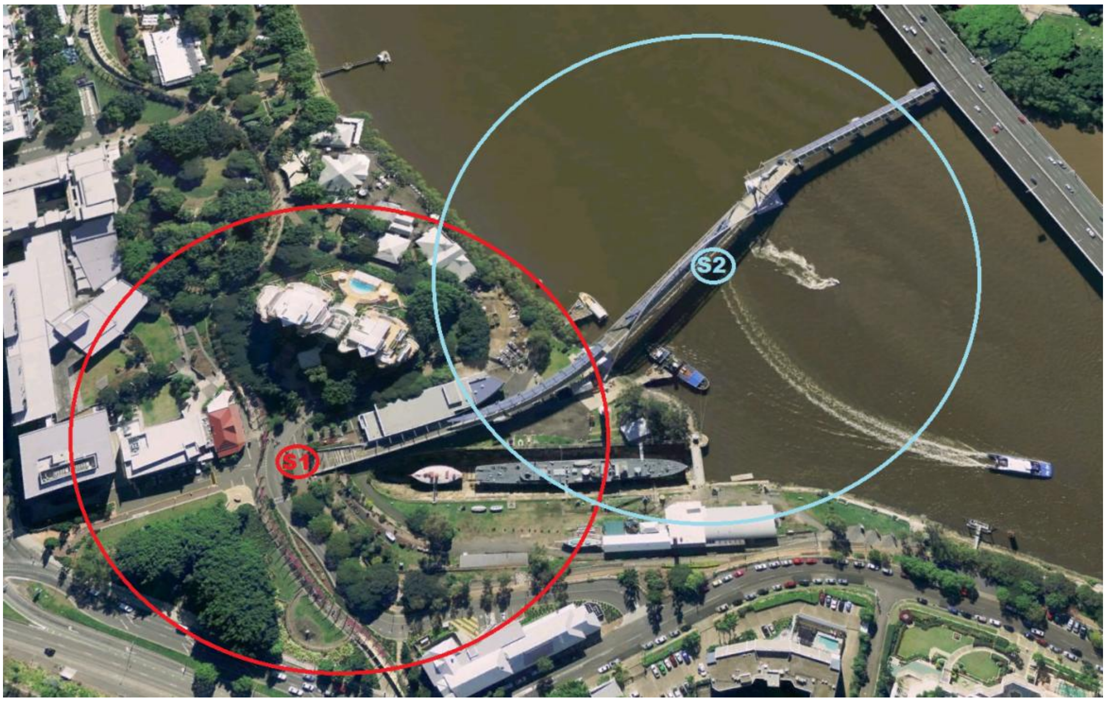
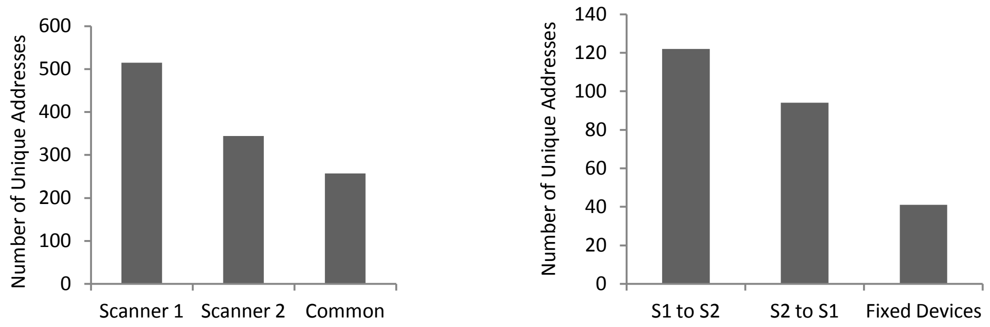


Figure 11. Scanning results



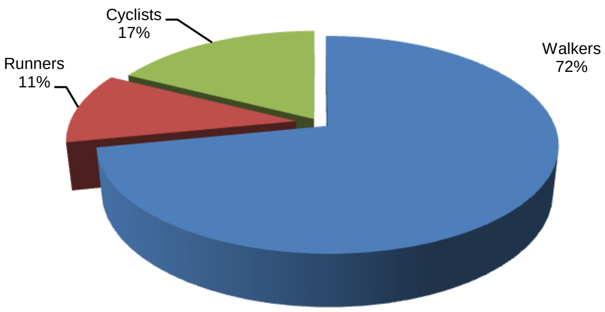
As can be seen from Figure 11, scanner 1 scanned unique Wi-Fi addresses more than scanner 2. 57 unique IDs were captured by both scanners. This means that majority of unique IDs observed by scanner 2 were also captured by scanner 3. Around 120 devices travelled from scanner 1 to scanner 2, 94 of them moved from scanner 2 to scanner 1, and 47 of them were scanned during entire experiment period.

We moved a mobile phone with active mode Wi-Fi during the experiment from scanner 1 to scanner 2 and opposite direction. We then matched the real-time records of our mobile Wi-Fi with scanner records. We moved the Wi-Fi device by average walking speed of 1.4 m/s. The scanner results showed that our MAC address was scanned by both scanners around 37 seconds. Therefore, the detection area in overlapped zone was almost 52 meters long (dx=V x dt). Furthermore, a cyclist with 5 m/s will spend around 10 seconds in the overlapped area. Figure 12 and Table 5 shows the proportion of walker, runner, and cyclists according to their time period in overlapped area.

Table 5. Categorising people movement by their time spent in overlapped zone

Category	Time Spent
Walker	30-40 sec
Runner	20-30 sec
Cyclist	10-20 sec

Figure 11. Proportion of walker, runner, and cyclists



According to the results of the overlapping technique, the spatiotemporal dynamics of people movement was tracked in three consecutive detection branches. The results indicate that the area was divided into more detectable zones by overlapping scanners coverage area. This approach can provide more information about people movement dynamics. We could be able to track people and measure their spent or travel time in each consecutive detection zones.

Conclusion

The significant benefits and challenges on the use of Bluetooth and Wi-Fi for crowd data collection and monitoring have been discussed in this paper. The traceability function of this type of crowd data collection has counted as an advantage compared to other methods. Collecting efficient crowd data by scanning MAC addresses can be matched with other crowd data collected by other methods in order to enhancement of crowd movement dynamic analysis and monitoring. Antenna characteristics, environmental interference and scanning features were explained as the major factors effecting on the process of MAC Address scanning for crowd data collection purposes. The impact of different antenna gains on the scanning process was empirically tested. The result recommends that the antenna type and gain must be selected based on the data collection space size and environmental complexity.

Wi-Fi and Bluetooth standards were compared in terms of architecture, discovery time, popularity of use and signal strength. The theoretical and experimental results of discovery time indicate that Wi-Fi has shorter discovery time (around 1 second) comparing to Bluetooth discovery time (almost 10 seconds). Also, the result of scanning MAC addresses in six different places showed that more than 90% of observed unique IDs were Wi-Fi addresses. It was expected at first that the scanning rate of Wi-Fi addresses in the places with a free Wi-Fi network is more than other places because a free Wi-Fi network may increase the number of people who have turned on Wi-Fi devices. However, the results showed that free network places have more number of turned on Wi-Fi devices but the difference is not noticeable. The signal strength of Wi-Fi and Bluetooth is measureable and the distance between scanner and MAC device can be estimated based on the signal strength value. The empirical results of different antenna gains for Wi-Fi signal indicated that signal strength value drops dramatically from zero distance to almost half of detection range and then does not change significantly until the end of detection range. In overall, Wi-Fi is accepted as the more suitable MAC standard compared to Bluetooth for crowd data collection.

Estimating the location of MAC addresses is discussed as an important issue that can enhance the process of crowd data collection and analysis. Two scanning approaches are presented in this paper, multi-range and overlapping detection techniques. In these methods, the scanning area is divided into different detectable zones. The method was applied in real experiments. The real-world results successfully proved that the methods can be applied and will extract more information about people movement dynamics. Presented scanning approaches are useful for estimating the location of MAC addresses. The implementation of mentioned scanning approaches in large scale can deliver significant information from spatiotemporal dynamics of people movements.

References

Andrienko, G. & Andrienko, N. V. (2007). Extracting Patterns of Individual Movement Behaviour from a Massive Collection of Tracked Positions. Paper presented at the BMI, Gottfried, Björn.

Andrienko, N., & Andrienko, G. (2007). Designing Visual Analytics Methods for Massive Collections of Movement Data. Cartographica: The International Journal for Geographic Information and Geovisualization, 42(2), 117-138.

Andrienko, G., Andrienko, N., Bak, P., Kisilevich, S., & Keim, D. (2009). Analysis of community-contributed space-and time-referenced data (example of Panoramio photos).

Andrienko, N., Andrienko, G., & Gatalsky, P. (2000). Supporting visual exploration of object movement. Paper presented at the Proceedings of the working conference on Advanced visual interfaces, Palermo, Italy.

Andrienko, G. L., Andrienko, N., & Wrobel, S. (2007). Visual analytics tools for analysis of movement data. SIGKDD Explor. Newsl., 9(2), 38-46.

Arbaugh, W. A., Shankar, N., Wan, Y. C. J., & Kan, Z. (2002). Your 80211 wireless network has no clothes. Wireless Communications, IEEE, 9(6), 44-51.

Bisdikian, C. (2001, December). An overview of the Bluetooth wireless technology. IEEE Communications Magazine, 45, 86-94.

Bruno, R., & Delmastro, F. (2003). Design and analysis of a Bluetooth-based indoor localization system. Paper presented at the International Federation for Information Processing.

Bray, J., & Sturman, C. (2001). Bluetooth: connect without cables. Upper Saddle River: Prentice Hall PTR.

Chakraborty, G., Kshirasagar, N., Chakraborty, D., Shiratori, N. & Wei, D. (2010). Analysis of the Bluetooth device discovery protocol. Wireless Netw ,16:421-436

Crow, B. P., Widjaja, I., Kim, L. G., & Sakai, P. T. (1997). IEEE 802.11 Wireless Local Area Networks. Communications Magazine, IEEE, 35(9), 116-126.

Dimitrova, D. C., Alyafawi, I., and Braun, T. 2012. Experimental Comparison of Bluetooth and WiFi Signal Propagation for Indoor Localisation. WWIC 2012, LNCS 7277, 126–137, 2012.

Doantam, P., Ling, X., Yeh, R., & Hanrahan, P. (2005, 23-25 Oct. 2005). Flow map layout. Paper presented at the Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on.

Ferro, E. & Potorti, F. (2004, 30 Jun). Bluetooth and Wi-Fi wireless protocols: a survey and comparison. IEEE Wireless Communication Magazine.

Ferro, E., & Potorti, F. (2005). Bluetooth and Wi-Fi wireless protocols: a survey and a comparison. Wireless Communications, IEEE, 12(1), 12-26. doi: 10.1109/MWC.2005.

Gast, M. (2005). 802.11 Wireless Networks: The Definitive Guide (2nd ed.): O'Reilly Media.

Giannotti, F., & Pedreschi, D. (2008). Mobility, Data Mining and Privacy Springer.

Hagerstraand, T. (1970). What about people in regional science? Papers in Regional Science, 24(1), 7-24.

Hallberg, J., Nilsson, M., & Synnes, K. (2003, 23 Feb.-1 March 2003). Positioning with Bluetooth. Paper presented at the Telecommunications, 2003. ICT 2003. 10th International Conference on.

Han, B. & Srinivasan, A. (2012). eDiscovery: energy efficient device discovery for mobile opportunistic communication.

Harwood, M. (2009). CompTIA network+ N10-004 Cram (3rd ed.): Pearson IT Certification.

Hay, S., & Harle, R. (2009). Bluetooth Tracking without Discoverability. Paper presented at the Proceedings of the 4th International Symposium on Location and Context Awareness, Tokyo, Japan.

Hossain, A. K. M., & Wee-Seng, S. (2007, 3-7 Sept. 2007). A Comprehensive Study of Bluetooth Signal Parameters for Localization. Paper presented at the Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on

Jankowski, P., Andrienko, N. V., Andrienko, G. L., & Kisilevich, S. (2010). Discovering Landmark Preferences and Movement Patterns from Photo Postings. T. GIS, 14(6), 833-852.

Leitinger, S., Gr¨ochenig, S., Pavelka, S., and Wimmer, M. (2010). Erfassung von Personenstrmen mit der Bluetooth-Tracking-Technologie. In: Angewandte Geoinformatik 2010. Addison Wesley Longman Inc., New York, 15th edition.

Liebig, T., & Wagoum, A. U. K. (2012). Moddeling microscopic pedestrian mobility using Bluetooth. Paper presented at the Proceedings of the 4th International Conference on Agents and Artificial Intelligence.

Paper presented at the Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, Washington.

Pels, M., Barhorst, J., Michels, M., Hobo, R., & Barendse, J. (2005). Tracking people using Bluetooth: Implications of enabling Bluetooth discoverable mode.

Porter, J. D., Kim, D. S., Magana, M. E., Poocharoen, P., & Arriaga, C. A. G. (2011). Antenna characterization for Bluetooth-based travel time data collection. Paper presented at the TRB 90th Annual Meeting Compendium of Papers DVD.

Stange, H., Liebig, T., Hecker, D., Andrienko, G., & Andrienko, N. (2011). Analytical workflow of monitoring human mobility in big event settings using Bluetooth. Paper presented at the Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness, Chicago, Illinois.

The Bluetooth Special Interest Group. Baseband specification version 1.1. http://www.bluetooth.com, 2001.

Vasiliev, I. (1996). Mapping time: An analysis of the cartographic problem of representing spatiotemporal information. (Doctor of Philosophy (PhD)), Syracuse University.

Versichele, M., Neutens, T., Delafontaine, M., & Van de Weghe, N. (2012). The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities. Applied Geography, 32(2), 208-220.

Vu, L., Nahrsedt, K., Retika, S. and Gupta, I. 2010. Joint Bluetooth/EiFi scanning framework for characterizing and leveraging people movement in university campus. MSWiM’10, October 17-21, Bodrum, Turkey.

Weinzerl, J., & Hagemann, W. (2008). Automatische Erfassung von Umsteigern per Bluetooth-Technologie. Nahverkerspraxis.

Willig, A. (2003, 2-6 Nov. 2003). An architecture for wireless extension of PROFIBUS. Paper presented at the Industrial Electronics Society, 2003. IECON '03. The 29th Annual Conference of the IEEE.