Liik enne vira sto

Finnish Transport Agency

HENRI SINTONEN

Bluetooth Based Travel Time Estimation

LITERATURE REVIEW

EasyWay

Co-financed by the European Union
Trans-European Transport Network (TEN-T)

Henri Sintonen

Bluetooth Based Travel Time Estimation

Literature review

Photo on the cover: Risto Kulmala

Abstract

This literature review explores the best approaches to Bluetooth based travel time estimation, since they have gathered much interest as of late. Focus is on the hardware and software design of the system and the technical and economic performance of the technology. Key components and software techniques are identified, such as antenna configuration and data filtering. Most of these topics along with the installation have multiple options, most of which have not been studied comprehensively. The approach performs well in comparison to other alternatives in the accuracy of travel time estimation under medium or heavy traffic, but tends to fail under low traffic. It costs less than the technologies it is compared against and offers a very high level of privacy protection.

Tiivistelmä

Tämä kirjallisuusselvitys kartoittaa parhaita käytäntöjä Bluetoothiin perustuvalle matka-aikojen arvioinnille, joka on herättänyt runsaasti kiinnostusta viime aikoina. Selvitys painottuu järjestelmän laitteisto- ja ohjelmistosuunnitteluun sekä teknolo-giaratkaisujen tekniseen toimivuuteen ja kannattavuuteen. Selvitys tunnistaa järjes-telmän tärkeimmät osat ja ohjelmistoratkaisut kuten antennijärjestelyt ja tiedon suo-datuksen. Näille samoin kuin laitteiston asentamiskäytännöille on useita vaihtoehtoi-sia ratkaisuja, joista moniakaan ei aineiston mukaan ole tutkittu kattavasti. Bluetoot-hiin perustuva matka-aikojen laskenta suoriutuu vaihtoehtoisiin laskentatapoihin nähden hyvin matka-aikojen arvioinnin tarkkuudessa melko ja hyvin vilkkaassa liiken-teessä mutta huonosti hiljaisen liikenteen aikana. Menetelmän kustannukset ovat vaihtoehtoisia teknologiaratkaisuja alhaisemmat. Menetelmä on erityisen toimiva yk-sityisyyden suojan suhteen.

Sammanfattning

Den här litteraturutredningen kartlägger de bästa sätten att mäta restid med Bluetooth, vilket har väckt mycket intresse under de senaste tiderna. Utredningen koncentrerar sig på planering av systemets hård- och mjukvara samt på hur bra de olika teknologilösningarna fungerar och hur lönsamma de är. Utredningen presenterar de viktigaste hårdvarorna samt lösningarna på mjukvara, som till exempel antennanordningar och filtrering av information. För dessa, som också för de olika installeringsalternativen, finns flera alternativa lösningar, varav många enligt denna utredning inte har undersökts uttömmande. I relativt livlig eller mycket livlig trafik ger beräkningen av restid med hjälp av Bluetooth relativt exakta restider jämfört med alternativa beräkningssätt, däremot under lugn trafik ger Bluetooth beräkningen mindre exakta restider än de alternativa metoderna. Kostnaderna för Bluetooth-metoden är lägre än för alternativa teknologilösningar. Därtill är metoden speciellt fungerande gällande integritetsskydd.

Foreword

One of the basic building blocks of traffic management is real time transport network status monitoring. In order to be effective, the quality of traffic management needs to be sufficiently high, and this relies on monitoring of sufficiently high quality. The road operator is in its role as the network manager constantly looking for more cost-effective monitoring information for its services. New technology solutions provide novel monitoring approaches, which need to be evaluated for their feasibility. One of the promising technologies is Bluetooth, which is claimed to provide a cost-efficient tool for monitoring of travel times in road transport. This study was initiated to compile today's knowledge of the technology solution, technical performance, costs and cost-efficiency of Bluetooth in travel time monitoring.

The report was produced at the Finnish Transport Agency by Student of Technology Henri Sintonen under the guidance of Dr Risto Kulmala.

Helsinki, December 2012

Finnish Transport Agency
Traffic Management

Table of Contents

1 Introduction

Travel time is an important performance measure in road operation and can help road users make informed trip decisions in the face of sudden variability in traffic and thus optimize the road network utilisation. Traffic operators can take advantage of the real-time data when making timely decisions and the information can be used to assess the bottlenecks of the road network in the hopes of providing high quality information for infrastructure planning.

Since 2008, a Bluetooth based travel time estimation system has received attention and research due to its non-invasiveness, cost-effectiveness and ease of installation. It derives the travel time information from the Bluetooth-enabled devices passengers carry with them in vehicles. The system is based on the concept of re-identifying vehicles at distinct sites and calculating the time it took the vehicle to travel the distance between them, but it lacks the privacy issues that haunt similar technologies such as automatic license plate recognition systems. Since the number of Bluetooth-enabled devices on vehicles is expected to grow, the concept is seen as a promising alternative to more traditional travel time measurements technologies.

This review seeks to incorporate the findings of the studies on the experiences and practices regarding the system design and installation. It will also try to clarify the general performance of the technology, factors affecting the performance and the cost of such a system by comparing them to other alternatives.

The review layout is as follows. Chapter 2 touches on the key concepts of the system. Chapter 3 presents the hardware design of the roadside unit. Chapter 4 looks at the software techniques used in the system. Chapter 5 takes a specific look at the antenna configurations. And finally Chapters 6 and 7 look at the technical and economic performance of the system, respectively.

2 Technology and concepts

2.1 Bluetooth

Bluetooth is a wireless telecommunications standard used for short-range data exchange in many consumer products such as laptops, mobile phones, personal navigation devices, and headsets. It uses the 2.4GHz short-range radio frequency spectrum, which it divides into 79 channels of 1 MHz bandwidth each to transmit data. Devices hop 1600 times a second from one channel to another to combat interference in the unlicensed spectrum shared by various other devices. In order for one Bluetooth device to find another, it needs to send inquiry packets to 32 of the 79 channels. Devices that are in a discoverable state will scan these channels and respond to the inquiry. (Bluetooth.com, n.d.; Bluetooth Special Interest Group, 2011).

Bluetooth radios are classified into three groups based on their communication range. Class 3 has a range of up to 1 metre, Class 2 up to 10 metres, and Class 1 up to 100 metres. However, only a minimum range is mandated by the Bluetooth specification and manufacturers can tune their implementations for needed ranges. (Bluetooth.com, n.d.).

2.2 MAC Address

Each discoverable device responds to the inquiry with their Media Access Control (MAC) address and clock information. Clock information is used to keep the devices' internal clocks in sync. The MAC address is a unique 48-bit identifier for the device, for example, in a human-friendly form, "01:23:45:67:89:ab". This address is given to the device by the manufacturer. It is divided into six octets (units of eight bits), where the first three ("01:23:45") are the Organizationally Unique Identifier (OUI) that represent the manufacturer. The last three octets ("67:89:ab") are assigned by the manufacturer in any way they see fit to make the devices unique.

If one device is identified by the MAC address at two sites and the exact time those identifications took place is known, the time it took the device to travel the distance between the two sites and the speed at which it did so can be calculated. Aggregating results from multiple devices would give an average travel time or speed between those sites.

2.3 Roadside Unit

A roadside unit (or just "unit") is a piece of equipment that detects passing Bluetooth devices and sends the information along or stores it. The papers, which have discussed the design of the units (e.g. Puckett and Vickich, 2010; Porter, Kim and Magaña, 2011; Beca Infrastructure Ltd, 2011; Iteris, 2011), have had the following elements in common and they can therefore be thought as the core components of the unit:

1) A Bluetooth radio that sends and receives information to and from other Bluetooth devices.

2) A computer that processes the information. Processing can include for example privacy enhancement procedures and duplicate filtering.

3) Either means of communication or a local storage. If real-time data is required the unit should have, for example, a cellular modem for immediate data transmission to the host machine. Communications can also be used to maintain, troubleshoot, and update the software remotely.

4) A power source.

5) Antenna for the Bluetooth radio to increase the detection capabilities and range of operation.

6) Installation equipment and housing for the unit.

2.4 Host Machine

The host machine is a computer that receives the MAC addresses and time stamps from all the roadside units. It processes them, calculates the travel time, and refines its current average travel time information for the road segments. It should also offer easy exporting or viewing of the information for third parties.

Encryption and duplicate filtering can also be done by the host machine, but this is less safe due to the transmission of raw data, and it requires more bandwidth, because multiple instances of the same MAC address are going to be transferred.

2.5 On Privacy

Privacy issues should be taken seriously when extracting information from ordinary commuters. MAC address matching does not have overbearing issues in this department, since there is no database that would link any given MAC address to the user of the device (Haghani et al., 2010). In principle it could be possible to link the MAC address to the buyer of the device if the manufacturer revealed by the OUI had such a database. But even then the buyer may not be the user of the device, and it is possible to buy devices without leaving traces (e.g. cash). The concerned can turn their device's Bluetooth functionality off or set them to undiscoverable mode and so refuse to respond to the roadside unit's inquiry (Solon, Callaghan, Harkin, and McGinnity, 2006). Extra steps can be taken by anonymizing (e.g. removing octets) and encrypting (e.g. hashing) the MAC addresses before sending them to the host machine or storing them on the units. These steps are expanded in Chapter 4.5. Long-time storage of the addresses might enable tracking behaviour patterns that could risk people's privacy (Solon, Callaghan, Harkin, and McGinnity, 2006). Thus it is important that all information is deleted as soon as possible.

3 Roadside unit

The roadside unit is the main component of the whole system. It dictates the cost, effectiveness, longevity, and maintainability of the project, since it is the only component that needs to be built in quantities corresponding to the size of the monitored road network. In this chapter the various aspects of the design of the unit will be considered.

3.1 Overview

A trade-off exists between what should be done by the roadside unit and what by the host machine. A feature-rich unit can have lower data transmission needs, increased portability, better privacy protection, and offer more accurate results, but this comes with higher power requirements, more expensive units, and increased complexity.

There are multiple variables in the design and installation of the unit. These include all the internal and external components, power source, obstructions, and all the distances between the main unit, antenna, and all the lanes. Due to these varying environments it is not recommended to design a single model and installation method or equipment for all units and installation sites, but to first decide on a location and then customize the unit for it (Porter, Kim and Magaña, 2011). Failure to do so can lead to problematic situations, for example installing an exceedingly powerful computer into a tight and hot enclosure, which can cause system crashes (Puckett and Vickich, 2010), vandalism in cases where the unit is easily accessible (Steel and Kilburn, 2011), and in general a non-ideal coverage of the road near the unit's installation site.

3.2 Power

Powering the unit is usually done using the environment's available equipment. Powered road signs and traffic cabinets could be used as a source (Porter, Kim, and Magaña, 2011). It can be a combination of a street light and a battery, which is recharged during the night when the lights are on, but attention should be paid on the battery so that it lasts the whole day and can be sufficiently recharged from the power source during the night (Beca Infrastructure Ltd, 2011). Other possibilities include an AC socket, solar power, or just a long-lasting battery for a portable model, although a heavier battery can bring down the portability of a unit. A badly placed battery may cause interference with the short radio waves from the unit (Steel and Kilburn, 2011).

3.3 GPS

For accurate results the internal clocks of all the units need to be synchronized. To accomplish this many researchers have installed GPS modules in their units, because the GPS technology requires time synchronization in the nanosecond scale to calculate the position of the GPS receiver. However, if the module is connected via USB there might be some loss of accuracy due to USB latency, which is some milliseconds and may vary (Korver, 2003; Ramadoss and Hung, 2008). This would

mean that only an accuracy of some milliseconds should be expected from a USB GPS module for clock synchronization. This is similar to the several tens of milliseconds possible with regular clock synchronization over the Internet using Network Time Protocol (NTP) (Mills, 2012). More suitable ports for the module exist (e.g. RS232, PCI), but availability, compatibility, and cost may become an issue. USB GPS module can however be used to obtain the location of the unit, which is especially useful for a portable model.

3.4 Communications

3.4.1 Connection type

The simplest way to transfer data from the unit to the host machine is to install a removable memory card. Real-time traffic information is, however, impossible with this design, but it is sufficient for a portable unit that can be temporarily installed and used, for example, to study the effects of a finalized roadwork project or to prototype different antenna configurations (Porter, Kim, and Magaña, 2011).

A reliable, simple, and cheap method would be to use a wired Ethernet connection, but finding an available long-term connection port next to a road might be difficult.

The most common way to transfer real-time or periodical data is to use cellular networks through a USB modem. This should be done in close collaboration with the service provider, since problems experienced so far have included 3G data packet running out of credit, inaccurate data usage reports, and an uncooperative service provider (Beca Infrastructure Ltd, 2011).

3.4.2 Transferred Data

A standardized syntax for the information the roadside unit sends to the host machine would make integrating multiple systems easy. However, no such standard exists, so every manufacturer or researcher currently needs to design their own format. Porter, Kim, and Magaña (2011) proposed a simple method, where the information from each detected device is stored into a plain text file as a line with comma separated values:

"anonymized MAC address, date, time, unit's MAC"

This way a megabyte sized plain text file would store 22,795 detections and grow linearly. Puckett and Vickich (2010) had a similar format:

"date & time, location of unit, detected MAC address"

The location consisted of the primary street, where the unit was installed, and the nearest intersecting street separated by an underscore. However, the biggest difference between these two methods was in the transmission. Instead of periodically sending a file with data from multiple devices, the units of Puckett and Vickich (2010) sent individual detections immediately. This allows for a more real-time system.

Sending the date might be unnecessary in a real-time system, since for privacy reasons information should not be stored long enough to make it relevant. Even if it is

needed, for example to generate origin-destination data, it could be assigned by the host machine upon receiving since the transfers happen on the day of detection.

3.4.3 Protocols

TCP and UDP are two protocols that are used to transfer information from one computer program to another, when they are located on different computers, but share an Internet connection. Consideration should be used when deciding which protocol to use. When sending a text file with multiple detections, reliability should be emphasized to avoid big data losses. TCP has features that detect errors and lost information and the receiver sends an acknowledgement back to the sender on the status of the received packets (RFC 675; RFC 793). This way the sender can retransmit some information and in the end be sure that no information was lost or altered during the transmission. This however takes a toll on the latency and bandwidth of the transmission, since TCP requires much extra information to be transferred back and forth. Establishing a connection alone needs the back and forth transmission of three packets before any actual data can be sent (RFC 675).

UDP on the other hand lacks all of these features, except for a test on integrity (RFC 768). There is no way to know whether the data reaches its destination or not, which makes the transmission unreliable. The consequent lightweightness does, however, reduce latency. It is commonly used for real-time systems that do not suffer from packet loss or errors, for example streaming video. The Bluetooth based travel time system uses a sample of the traffic to calculate the travel time estimation. The sample is big enough for accurate estimation if the number of data points is over a certain threshold (Wieck, 2011; KMJ Consulting, 2010). If the data loss caused by UDP does not decrease the sample size below this threshold, it should not cause a significant difference in the estimation, but this was not studied in the material reviewed. Puckett and Vickich (2010) did mention using the UDP protocol to transmit data from the units to the host machine, but did not report experimenting with it or other protocols.

4 Software

Software is an essential part of the Bluetooth system and can have a significant effect on the performance of the system. In this chapter we will describe various software parts of the system on both the roadside unit side and on the host machine side.

4.1 Operating System

The roadside unit's operating system is closely related to the hardware requirements and usability. Choosing a big and complex operating system increases the performance requirements of the roadside unit's computer just to run the operating system itself. Many of the powerful units are designed to support the Microsoft Windows operating system. Windows has been found to be too complex for the task, contribute to multiple processor and communications failures, lack some necessary processes, and be difficult to troubleshoot or maintain remotely (Puckett and Vickich, 2010; Porter, Kim and Magaña, 2011). For these and economic reasons almost all the researchers have either switched to or initially started with some Linux distribution. Linux distributions are a family of operating systems all built on top of the Linux kernel with varying software utilities and features that dictate their operational capabilities and environment of use. They range from fully fledged desktop computer operating systems to minimal ones in embedded systems used, for example, in the roadside units of Puckett and Vickich (2010).

4.2 Bluetooth Stack and API

The Bluetooth stack is a software layer that implements the Bluetooth protocol stack and is embedded in the operating system. For external software to access the stack it need to use the application programming interface (API). It provides the functionality for Bluetooth communication, such as the inquiry method used to capture the MAC addresses of devices. There is no official Bluetooth stack or API so multiple alternatives of varying features are available, but most operating systems provide very developed tools. For example, there is the Microsoft Windows Bluetooth stack and the BlueZ, which comes with the Linux kernel.

Puckett and Vickich (2010) ran into problems with both the Microsoft Windows Bluetooth stack and the BlueZ. They reported that the inquiry method provided was synchronous by nature meaning that it first scanned all the channels for about 10 seconds and then grouped and returned all the found devices. It also only found a maximum of 8 devices per scan. This led them to develop their own inquiry process that asynchronously reports all the devices immediately after they are found in mid-scan.

The about 10 second scan time is the results of a recommendation by the Bluetooth specification (Bluetooth Special Interest Group, 2011). It states that inquiring for devices involves scanning 256 times two 16-channel subsets of the 32 channels used for device detection. Subset should be switched three times, which means that both subsets are scanned twice. Scanning a subset once takes 10ms.

$$(2 \text{ subsets} * 256 \text{ times}) * 2 \text{ times} * 0.01 \text{ seconds} = 10.24 \text{ seconds}$$

This is the recommended duration and the default value for the inquiry, but the minimum is 1.28 seconds. The APIs have the option to set the duration parameter in 1.28 second intervals. It is unclear how Puckett and Vickich (2010) exactly changed the inquiry method. They used Python as their programming language, so it might be something similar to method detailed by Huang and Rudolph (2005, pp. 28-29). In any case, they reported that after implementing their tweak there was an immediate 50% increase in the number of non-duplicate Bluetooth device detections and the resulting match rate for a road segment went up by 51%. Most of the other studies reviewed did not report encountering this problem or using a similar asynchronous method, so it is left unclear whether their results are with or without a similar tweak. Notable exceptions are Stevanovic et al. (2011), who mentioned changing the duration parameter and the type of search to not report cached devices detected in previous searches, but noted that the discovered devices are processed after the whole search is over, and KMJ Consulting (2010), who reported that their units had a cap of 16 detections per minute per unit, which was only removed after the study.

4.3 Filtering and Travel Time Calculation

Duplicate filtering is removing detections of the same device at the same location during an appropriately short time. Simple ways to do this are, for example, preserving the first detection, the last detection, or taking the average of the two, the latter of which Porter, Kim, and Magaña (2011) found to provide the most accurate results of the three. Ideally, the travel time should be calculated exactly from the distance between the two units, but because both the units have a coverage area, where the detection is possible at some point and which is not usually perfectly symmetrical due to obstructions and installation decisions, the actual measurement may be done from a longer or shorter distance skewing the travel time. Even with the ideally spherical coverage area of an isotropic antenna, taking the first or last detections would not lead to perfectly accurate results, since the detections may not happen at the same point in the units' coverage areas. If two units are installed at separate intersections, taking the first detection in the first intersection or the last detection in the second intersection (first-last method) would count the time spent on both intersections, whereas taking the last detection of the first intersection and the first detection on the second intersection (last-first) would only include the time spent between the intersections (Wieck, 2011). First-first and last-last methods would count the waiting delay from one intersection.

Being able to measure the distance between the target device and the roadside unit at the time of detection could help. The means to do this are limited to begin with and are even more impaired due to the MAC address matching technology relying only on the device detection process and not on any actual data transmission. Capturing the received signal strength indication (RSSI) might help, since the signal strength is inversely proportional to square of the distance. However, the RSSI is very susceptible not only to the environment and its obstructions, but also to the transmission power, radio frequency, antenna characteristics, localization algorithm, and quality of the reference measurements (Awad, Frunzke, and Dressler, 2007). Thus it is unpredictable and unreliable as a direct measurement of distance, but might be of use when taking an average of multiple detections. Whether a clever RSSI processing algorithm that takes into account the parameters listed above is developed and what is requires from the units remains to be seen.

Duplicate filtering will cut down on the data transmission needs if done on by the units. When Puckett and Vickich (2010) implemented this by choosing only the first detection the MAC address, the transmissions were reduced by about 90% while preserving the same match rate. The actual travel times increased and it was presumed that this was because the data was now "cleaner" and processing more efficient, thus being more honest in representing real traffic. However, as discussed above, the distance between the first detections of the vehicle at both sites may not be equal to the distance between the units themselves. It might me longer, which would result in erroneously longer travel times with the same traffic.

Outliers in the data are travel times that are clearly too fast or too slow to portray the actual traffic situation. Thus filtering of outlier data points is important for ensuring a low error rate in the estimations (Wang, Malinovskiy, Wu, and Lee, 2011). Pedestrians, trains, bicyclers, and vehicles that stop for a moment or take a longer route will skew the estimation. There is no perfect one-size-fits-all way to filtering, so adaptive algorithms or tailor-made rules for specific road segments (Steel and Kilburn, 2011) are recommended. Simple rules, such as "if the travel time deviates more than 25% from the current average, discard it", may work well for freeways, where speeds are somewhat constant, but may not be flexible enough to react to quickly changing traffic conditions in arterial roads (Puckett and Vickich, 2010). Multistep processing methods have been demonstrated. For example, Haghani et al. (2010) devised a four step method, where each step filters data points with a different algorithm, but this was designed for offline data analysis.

Calculating the travel time involves updating the current value with new information. This can be done, for example, with moving mean analysis so that the average travel time for each minute is the mean of the travel times from the previous x minutes (e.g. 10 minutes) (Wang, Vrancken, and Seidel, 2011). The mean travel time can be unjustifiably harshly influenced by unfiltered outliers, so it could be substituted by the median, which on the other hand does suffer worse from very low traffic (Beca Infrastructure Ltd, 2011). The low traffic problems with the travel time calculation are worsened further since filtering of the outliers is difficult when there is no reliable baseline to which new data points could be compared to, but these problems could be alleviated by using a different calculation method designed for fewer samples during known periods of minor traffic (e.g. night) (Wang, Vrancken, and Seidel, 2011). Filters can spot undersampled time periods when a certain threshold of observations is not met (e.g. more than three cars every five minutes) (Haghani et al., 2010). This information could be used to report the travel time estimation only when it is thought as accurate if a certain level of veracity has been guaranteed or to inform the end user about possible inaccuracies in the estimation.

4.4 Maintenance

Since on-site troubleshooting visits are costly, time-consuming, and sometimes tricky, remote maintenance is important. Software problems can be fixed easily with a Secure Shell (SSH) connection, especially on a Linux system, or with similar remote connection. Issue with the communications could render remote means impossible. To prepare for such problems an automatic rebooting process could be implemented (Puckett and Vickich, 2010). The unit would monitor itself and reboot when experiencing a problem. Even this would not help if the whole unit would break down, so a downtime logging program on the host machine could write down the last time

each unit transmitted a MAC address to the host machine and monitor that the time passed since does not go over a set threshold (Puckett and Vickich, 2010). If it does it will notify the system operators. The unit should also send periodically information about the temperature and voltage for breakdown prevention.

4.5 Privacy Measures

Privacy measures are targeted at manipulating the MAC address to make is even more distant from the owner and user of the device. One way to do this is by removing some portion of it upon detection (Porter, Kim, and Magaña, 2011). Organizationally Unique Identifier (OUI) situated at the beginning of the MAC address is a unique three octet identifier for the manufacturer of the device. The rest of the address is chosen by the manufacturer is any way they like to make the address unique within the manufacturer's set of devices. Together they make the device unique across all Bluetooth devices. Removing some small section of the OUI, of the second half of the address or of both would make tracking the address through manufacturers even harder without sacrificing too much on the probability of two different devices now sharing the same modified address.

Another way, which can be used with or without the previous one, is cryptographic hashing (Puckett and Vickich, 2010). It modifies a string of characters to a usually fixed-length hash value in a specific manner while ideally preserving the uniqueness of it (i.e. there would not be two different strings with the same hash) and making it infeasible to convert the hash back to the original string. As an example, the MD5 hash of "01:23:45:67:89:ab" is "da959d984a9462f268db07d750cd8ab1". With a secure hash function and unit-side calculation, it would be infeasible to find out the original MAC address of the detected Bluetooth device while still allowing the travel time calculation to take place.

5 Antenna

A probabilistic evaluation of the Bluetooth technology by Bakula, Schneider IV, and Roth (2012) highlights why antenna decisions are important in ensuring a good sample size. It identified the bottleneck of the system to be the target radio's transmission power, but because that is hard to upgrade, efforts should be concentrated on increasing the sensitivity of the roadside unit. The antenna's characteristics and location play a major part in the size and shape of the coverage area, and the study found that the probability of detection is primarily influenced by the coverage area and secondly by the speed at which the vehicle is travelling. This chapter will look at the types of antennas used and the installation location of the antennas.

5.1 Characteristics of the Antenna

An ideal antenna in the MAC address capture context would allow for the detection of all devices in range, but only once and as close to the unit as possible. Interesting aspects of an antenna configuration can therefore be how much of the total traffic is detected and how many times a device is detected by the same unit in one drive-through (Porter, Kim, and Magaña, 2011). The number of duplicate detections can indicate how long vehicles are in the coverage area, which can be used to assess intersection performance (Wieck, 2011). In addition, the match rate should be measured to confirm that no local source of error is present and the accuracy of the samples should be calculated, since it is not unreasonable to prioritize accurate time travel estimation over a larger sample size.

Antennas can be defined by their directionality. An isotropic antenna is a hypothetical antenna that uniformly radiates in all directions with a spherical radiation pattern. Omnidirectional antenna radiates uniformly in horizontal directions, but not at all above or below the antenna. The radiation pattern can be described as "doughnut" shaped. A directional antenna radiates significantly more in one or more directions.

Polarization, linearly vertical or horizontal, is the orientation of the electric field radiating from the antenna. Most omnidirectional antennas have vertical polarization. There are dual polarization antennas and ones with circular polarization, which constantly change their polarization in a rotary manner. The reciprocity theorem states, that an antenna radiates and receives in the same way, i.e. a vertically polarized antenna radiates and receives vertically polarized fields, but cannot communicate with a horizontally polarized antenna. Polarization can be changed by simply turning the antenna so one could hypothesize that the orientation of the Bluetooth-enabled device in the car would matter and that dual or circular polarization antennas would produce better results.

The gain of an antenna can be thought as the transmission power in the direction of the peak radiation when compared to the isotropic antenna (dBi).

Gain seems to small positive effect on its own. Puckett and Vickich (2010) found that increasing the gain of an omnidirectional antenna from 1 dBi to 5 dBi offered only moderate improvement to the detection rate. Similarly, Wang, Malinovskiy, Wu, and Lee (2011) saw lower errors with higher gains using directional antennas. Porter, Kim,

and Magaña (2011) reported that the best antenna they tested had a 9 dBi gain, the second one 12 dBi, and the worst 11 dBi, but these were a mixture of different kinds of antennas. The first was omnidirectional, but the second directional, but both had a vertical polarization. Surprisingly, the worst was a dual polarized antenna and circular polarizations were in between.

Researchers from the University of Washington's STAR lab, when interviewed by Porter, Kim, and Magaña (2011), reported that based on their tests directional antennas provide less accurate results due to a lower sample size. The better accuracy of an omnidirectional antenna equipped system is supported by Stevanovic et al. (2011). No mention was made on either studies about the polarization of the antennas or how they were directed. In their own test, as mentioned above, Porter, Kim, and Magaña (2011) found that the second best antenna was directional, but with a vertical polarization, which most omnidirectional antennas also have. In their second test, they installed two units to measure the match rate. The best performer was clearly a 180 degree directional antenna, matching 9.56% of the total traffic. Admittedly, when installed next to a road, the portion of the coverage area that overlaps with the road is very similar between an omnidirectional and a 180 degree antenna. Again, the dual polarization performed worst.

Porter, Kim, and Magaña (2011) tested also the accuracy by comparing the travel times to those from GPS probe vehicles. Correlation was found between a higher match rate and a higher absolute error. This was suggested to emerge from the nature of the coverage area produced by different types of antennas. A large coverage area allows for a higher probability of detection, but also causes the distance between the detection and the unit to increase, even when taking the average of the first and the last detection. However, a completely different correlation was found by Wang, Malinovskiy, Wu, and Lee (2011). A higher match rate produced more accurate results, presumably because a bigger sample would even out some sources of error. Reason for the different conclusion might be related to using a better ground truth (automatic license plate recognition) and a different duplicate filtering method (last-last), but it could also result from unreported differences in local conditions.

It seems that the best option for most situations is an omnidirectional antenna with high gain. If directional antenna is needed, it should also have a high gain and a vertical polarization. Omnidirectionality and high gain increase the coverage area so that vehicles have a higher change of being detected, which might explain the better detection and matching achieved with them. According to Wang, Malinovskiy, Wu, and Lee (2011), these would translate to better accuracy.

An omnidirectional antenna cannot distinguish the target road's general traffic from that on parallel roads or High Occupancy Vehicle lanes. This can decrease the accuracy of the system, since filtering may get difficult. Directional antennas or omnidirectional antennas with shielding or concentrators may help, but this has only been suggested and not proven. (Puckett and Vickich, 2010).

5.2 Location of the Antenna

The location encompasses the characteristics of the installation environment and the vertical and horizontal placement of the antenna in relation to the traffic.

All three aspects became evident for Puckett and Vickich (2010) when they installed their unit with an internal antenna on the ground behind a concrete barrier. By removing the obstruction from the line of sight and raising the antenna higher (i.e. placing the unit on top of the concrete barrier) they noticed a significantly higher amount of detections from the lanes farther from the unit. They suggested that an optimal positioning of the antenna was at the windshield height of a typical passenger car. This might also be related to the finding that it is harder to detect devices that are in the pocket of the passenger compared to devices that sit on the vehicle's dashboard when the antenna is on top of a traffic cabinet (Stevanovic et al., 2011). This hints that there might be some obstruction in the car's body that aggravates the signal penetration.

The vertical location was studied particularly by Brennan et al. (2010), who installed five units along a highway with masts of variable height with an omnidirectional antenna on top. The PVC tubing masts were between 0 and 3 metres tall at 76cm intervals. The 0 metre antenna didn't have mast as it was inside a weather-proof case with the rest of the components on ground-level. The results show that 2.29 and 3 metre masts produced over twice the number of detections as the 0 metre mast. Masts in between these extremes produced intermediate results, but followed the correlation that the higher the antenna is situated the better the detection rate. The difference between 2.29 and 3 metres was small, but not insignificant. The final recommendation was 2.5 metres. It is unclear whether this increase is due to clearing some obstruction in the car or just because of a bigger coverage area that results from an elevated installation.

Horizontal placement was also studied by Brennan et al. (2010) when the units were installed on a four-lane interstate. The closest lane was 5.5 metres and the farthest 34.7 metres from the unit. The unit acquired more MAC addresses from the lanes closer to the unit. With the ground-level unit 64% of the detections came from the two closer lanes. This effect diminished the longer the mast was, but did not disappear altogether. With the 3 metre mast 52% of detections came from the closer and 48% from the farther lanes. This bias was supported by RSSI measurements. Recommendation was that when installing units on a multilane road segment, either two units should be used or one situated in the middle of the lanes. The two unit setup, where the units communicated with each other via Wi-Fi, was tested by Wieck (2011) on a highway and was found unnecessary (Garbe, 2011). However, Wang, Malinovskiy, Wu, and Lee (2011) reported that multiple antennas greatly increased the detection (15.35% compared to the 9.37% of a single sensor system) and matching rate (7.92% versus 3.43%) and reduced errors in most cases, especially with combination of omnidirectional and directional antennas. The reason for Wieck (2011) and Garbe (2011) to find the use of two unit setup as not bringing any added value might be in the intersection installation, as it already increases the match rate due to lower speeds (Stevanovic et al., 2011).

6 Technical Performance

Now that the various parts of the system have been identified and studied, the technical performance of the system, i.e. the accuracy of the estimated travel times and the extent of the detection and matching rates, can be better assessed. This chapter will also look at the reliability of the system.

6.1 Veracity

Veracity represents how honest the measurements are in estimating the real average speed or travel time. Since no absolute truth about the average speed or travel time is available, for each comparison another technology must be taken as the ground truth.

Comparing the Bluetooth data to vehicle probe data may be difficult, since there are considerable differences in the amount of data points between the methods. Probe data can also suffer, especially on longer routes, from the driving behaviour and habits of the driver and from the flexibility that an individual vehicle has when manoeuvring in traffic (Haghani et al., 2010). However, there seems to be no statistically significant difference between the two methods when a highway segment is less than 1.6 kilometres long and the speed is under 97 km/h or when the length of the segment is over 1.6 kilometres and the speed is under 72 km/h (Haghani et al., 2010). Required amount of data was not available in the higher speed categories to confidently rule in one way or the other. In a study by Wieck (2011), conducted on a county highway segment of about 10 kilometres with 10,000 to 25,000 vehicles a day, the difference between the probe data and Bluetooth data ranged from 9 to 104 seconds with the average being 42 seconds (Garbe, 2011). Stevanovic et al. (2011) ran a probe test on arterials and found that there was no significant difference between the two methods.

Wang, Vrancken, and Seidel (2011) found that the Bluetooth technology produced travel times roughly similar to the ones measured by loop detectors on both 5 kilometre and over 10 kilometre long routes. On a 1.2 km highway section with a mean speed of 60 km/h and about 30,000 vehicles every day, the average difference compared to loop detectors was about 1 km/h with 2 km/h standard deviation and a 21 km/h maximum deviation, which occurred during low traffic when only a few data points were collected (Luber, Junghans, Bauer, and Schulz, 2011).

Comparison to automatic license plate recognition system (ALPR) showed that the travel times were an average of 4-7% above the ALPR travel time for the best tested configurations and across all configurations on average 8% above the ALPR time (Wang, Malinovskiy, Wu, and Lee, 2011). The reason for the bias for longer travel times was suggested to be the result of the probabilistic coverage area and its tendency to have a higher probability for slower vehicles, as echoed in Stevanovic et al. (2011).

KMJ Consulting (2010) examined the Bluetooth system in contrast with the EZPass toll tag reader system on a 4.67 km motorway section. The average travel time difference was less than 21 seconds in the westbound direction, and in the eastbound direction less than one minute, of which at least 16.5 seconds of the eastbound

difference is attributable to a 0.3 mile offset in device locations. Conclusion was that the travel times between the systems were comparable.

6.2 Availability

Availability expresses the time the system is operational. On the Bluetooth system this can be thought to mean that the system is not broken and that the data produced is not defective due to some perpetual, predictable, and repeatable reason.

Since the technology is relatively new (first introduced in the literature by Wasson, Sturdevant, and Bullock (2008)) there have been no long-term studies on the durability of the system that could give a confident estimate on the life-time of it or on the probability of a system failure. However, few of the short-term studies reviewed did encounter problems, which were battery related (Steel and Kilburn, 2011; Beca Infrastructure Ltd, 2011), vandalism (Steel and Kilburn, 2011), or communications problems (Iteris, 2011; Beca Infrastructure Ltd, 2011). However, these problems can be alleviated or removed by carefully selecting the components, elevated or secure (e.g. traffic cabinet) installation, and proper maintenance tools (see Chapter 4.4).

A harder case for the Bluetooth system is created by low traffic volumes. None of the reviewed papers that looked at nightly traffic reported accurate and reliable travel time estimations continuously during the night. Although the relative number of Bluetooth-enabled devices might be higher during the night than during the day, the absolute quantity will be significantly lower (Sharifi, Hamedi, Haghani, and Sadrsadat, 2011). This usually translates to an insufficient amount of data points, which cannot produce trustworthy estimates. This means, effectively, that the Bluetooth system is unavailable as a provider of reliable estimates during the night in these situations. The actual fleet volume varies locally, but if it is assumed that the low traffic lasts from midnight to 05:00 am, then it would account for ca. 21% of the daily uptime of the system. On the other hand, it would account for less than 2% of the daily traffic. Proposed ways to alleviate the issue are researching ways which permit using fewer samples (Wang, Vrancken, and Seidel, 2011), only reporting accurate results meaning that low traffic samples are filtered out (Haghani et al., 2010), and maybe aggregating multiple data sources.

6.3 Completeness

Completeness encompasses the detection and matching rates, i.e. what portion of the traffic volume is detected once or twice at different sites. The matching rate is bound by the detection rate, which in turn is bound by the penetration rate, or how many Bluetooth devices there are in the traffic, which varies by location and time. The matching rate is in practice always lower than the detection rate due to the probabilistic nature of the system as discussed in Chapters 4 and 5 and because vehicles can leave or enter the road section at many points, especially in longer sections (Wang, Vrancken, and Seidel, 2011.). However, with shorter distances the coverage area of the unit becomes a significant portion of the whole section, which increases errors (Haghani et al., 2010). Therefore delicate planning is needed for unit installation.

The rate percentages, however, cannot be used on their own to evaluate whether the units capture an adequate sample of the fleet. The real threshold is an absolute number, so the percentage should be accompanied by the average daily fleet volume since a low percentage may represent a significantly bigger sample on a busy road compared to a high percentage on a sedate road, and a good rule of thumb is for the number to be at least three matched pairs every five minutes, or nine matched pairs per 15 minutes, 36 matched pairs an hour, or 864 per day (based on research by the University of Maryland as cited in KMJ Consulting, 2010). Wieck (2011) noted that the sample size can be calculated by the usual statistical methods. For example, for a volume of 18,000 vehicles, 95% confidence, and 5% margin of error, the match count should be at least 375. For a 90% confidence, 267 matches are required.

As the coverage area the unit lays around it can be seen as a zone of detection probability, the ways to increase the probability and thus the detection and matching rates are usually antenna enhancements. These can include the type of the antenna and the vertical and horizontal location of it. In Chapter 5, where these issues are discussed at length, it was determined that the best antenna for a general installation, was an omnidirectional one with an elevated installation and centered in the middle of the lanes. Thus only the studies that come close to this optimal situation are reported here.

KMJ Consulting (2010) had a match rate of 3.5-4.1% on an interstate with volumes of about 51,000 to 58,000 in westbound and 83,000 to 91,000 in eastbound direction. The absolute number of matches was 2.3 to 3.8 times the daily minimum of 864. The EZPass toll tag reader system in comparison had a 10-37% match rate. Bluetooth results were with the 16 detections per minute per unit cap discussed in Chapter 4.2.

Luber, Junghans, Bauer, and Schulz (2011) used loop detectors as their ground truth on a highway with a mean speed of 60 km/h and about 30,000 vehicles every day. The detection rate was 6.3%.

Wang, Malinovskiy, Wu, and Lee (2011) compared the Bluetooth system against automatic license plate recognition (ALPR) on a 3-mile freeway segment with average speeds of 97 km/h. During an hour long test the ALPR detected 1957 vehicles for site 1 and 1368 for site 2. For Bluetooth the counts were 432 and 190 devices. At site 2 there was a concrete barrier shielding the unit, which partly explains the lower number. Bluetooth matched 116 of the possible 190 devices (61%) in the segment, while the ALPR matched 533 out of the 1368 (39%). ALPR had a larger sample, but only from one lane and one direction, while the Bluetooth system collects data from all lanes from all directions and was able to match more of its detections. They conducted another test on a segment with over 50,000 vehicles a day and found that multiple antennas greatly increased the detection (15.35% compared to the 9.37% of a single sensor system) and matching rate (7.92% versus 3.43%).

Wieck (2011) studied the technology on an arterial corridor with a daily volume of 10,000 to 25,000 vehicles and speeds between 60 and 90 km/h. The corridor had 6 intersections where the units were installed with 2.7 to 4.8 km spacing. The matching rate varied from 3% to 11.4%. (Garbe, 2011; Iteris, 2011).

Steel and Kilburn (2011) installed 30 units on a highway ring road for one week with en elevated installation on existing roadside infrastructure. They studied origin-destination patterns and the percentage of devices that were matched between units

and had one or more Automatic Traffic Recorders (ATR) in their path varied from 6.5% to 9.4% with an average of 7.4% when the total volume at the ATR sites for the whole week was 877,006 vehicles.

The detection and matching rates seem to vary from 3% to 11% in all tested road types and seem to grant a statistically good enough sample given adequate fleet size.

6.4 Treatment of Confounding Factors

Parallel roads, multi-passenger vehicles, and footpaths may produce error in the estimations. Ways to deal with these issues can be split into two categories: software and hardware means.

On the hardware side there is not much that can be done. The unit propagates a zone of probability, the coverage area, around itself that catches Bluetooth devices passing by and so it does not know anything about the travelling speed of the device as this calculation is done by the host machine nor can it now much about the location of the device inside the zone (see Chapters 4 and 5). The shape of the coverage area could be modified by directional antennas or by shielding omnidirectional antennas (Puckett and Vickich, 2010).

On the software side there is more flexibility. Filtering can effectively remove outliers from the data, such as people walking by, as described in Chapter 4.3. It gets harder the more close the speeds are between the wanted and unwanted traffic and the outcome is tied with the performance of the filtering algorithms, which are a constant area of research and can be updated on the host machine independently from the units. Similarly for multi-passenger vehicles, they can only be dealt by filtering, for example based on the notion that a big portion of the passengers should have very similar travel times and detection times. This was not discussed in any of the papers reviewed, but KMJ Consulting (2011) noted that the system studied claims to do this.

It should be noted that none of the papers evaluated the extent of the problem presented here in a controlled environment. It should be first determined how big of a problem the said issues are.

6.5 Other Performance Related Remarks

Increasing the sample size could be achieved by combining other real-time data sources. A similar MAC address matching technology seems to work for Wi-Fi and the vehicles detected by Bluetooth and Wi-Fi seem to be distinct from each other (Luber, Junghans, Bauer, and Schulz, 2011). The Wi-Fi technology had a mean matching rate of 1% a highway with a mean speed of 60 km/h and about 30,000 vehicles per day. Mean difference against loop detectors was 2 km/h with 3 km/h standard deviation.

Wang, Malinovskiy, Wu, and Lee (2011) noted that units should not be installed near bus stops. Even if buses and other High Occupancy Vehicles (HOV) themselves were not an issue, devices which have first been detected inside a bus and then walking away from a bus stop could be interpreted as causing extra delay which might go unnoticed by filtering algorithms.

7 Economic Performance

There are multiple ways to measure travel time and the cost of the method, both up front and during operation, is taken into account when deciding to adopt a system. This chapter looks at the various mentions on the economics of the Bluetooth system.

7.1 Investment Costs

The Bluetooth investment cost information of the reviewed literature are presented in Table 1 along with the comparisons made in the studies with alternative methods.

Study	Type	Cost per unit		Additional costs	Notes
		without installation	with installation		
Porter, Kim, and Magaña, 2011	BT	$280-$350	< $2,000		Research unit
Rajbhandari, 2008; Fink, 2011	BT	$500-$1,500	$2,000-$3,500		Product: AWAM
Iteris, 2011	BT		$5,000 or $6,000	$3,000 for misc. hardware for 8 units, $3,000 for host machine	Product: StreetWAVE
Fink, 2011	BT		$8,000		Solar powered
KMJ Consulting, 2010	BT		$9,700-$12,000		Solar powered
Wang, Malinovskiy, Wu, and Lee, 2011	ALPR	$10,000			One unit per lane
KMJ Consulting, 2010	EZ		$34,000-$36,000		Several units for multilane roads

Table 1   *BT stands for Bluetooth, ALPR for Automatic License Plate Recognition, and EZ for EZPass, the RFID toll rag reader system. The unit cost for Porter, Kim, and Magaña (2011) was calculated from the links provided by the authors or from component manufacturer's website.*

The low investment cost of the Bluetooth is not limited to the cost of a single unit, but to the absolute number of units needed. One four-lane location would require eight ALPR units (one for each lane and direction) bringing the total cost for one location to about $80,000 excluding installation (Wang, Malinovskiy, Wu, and Lee, 2011), which is close to the $90,000 per location for toll tag readers (City of Houston, 2011). The Bluetooth system requires only one unit per location.

7.2 Operation and Maintenance Costs

The operation costs are limited to power consumption and data transmission feeds. These vary depending on the unit design, the country of operation, and service provider and were not reported in any of the studies reviewed, with the exception of Iteris (2011), which mentioned the 3G data transmission cost in Minnesota, US from Sprint in 2010 to be an average of $340 per month for six units to transfer the data collected by eight units (two units reported their data via Wi-Fi to close-by units). However, a back-of-the-envelope calculation can be made to work out a ballpark estimation.

If we assume that one megabyte of data can carry detections from 22,795 devices (Puckett and Vickich, 2010) and that one unit would detect the minimum 864 devices a day (KMJ Consulting, 2010), or about 25,920 a month, needed for suitable travel time measurements, the monthly data transmitted per unit would be about 1.14 megabytes. It should, however, be presumed that the units will detect much, much more than the cited minimum number of devices and that there will be some overhead from the transmission itself. In any case, the amount of data transmitted will be small. Similarly for the power, Puckett and Vickich (2010) measured their computer and cellular modem to consume 2.88 watts at 13.1 V, so the cost of the electricity should also be modest.

With the maintenance tools discussed in Chapter 4.4, most of the activities will be automatic without actual labour or can be done remotely. In the case of a breakdown or a communications problem, which cannot be fixed by an automatic reboot, is an on-site visit required. Because the units can be assembled from commercial off-the-shelf components, replacing one should not require considerable resources. In cases where replacing a component is not enough, the cost of the repair should be at most the cost of a new unit and its installation.

7.3 Cost-benefit and Cost-efficiency

Virtually all papers reviewed mentioned that the Bluetooth MAC address matching technology is a cost-effective way of gathering traffic information. This is the result of low initial and operational costs and good maintainability along with a large number of data samples, which can be collected from multiple lanes with traffic going to multiple directions with singular units at each installation sites (Sharifi et al., 2011; KMJ Consulting, 2010).

The ways the collected data samples can be used are versatile. Besides travel time information, they can be used to gather average speed, origin-destination data, sample volumes of traffic, traffic movement patterns (e.g. passing by, zigzagging in the city, taking a break in the city, local trips) (Steel and Kilburn, 2011), and border

crossing times (Rajbhandari, 2008). Steel and Kilburn (2011) compared the technology to manually observing traffic. While manual observations have the advantage of differentiating between vehicle types, it costs $88,000 for 192 hour data collection, whereas the Bluetooth system had 30 units each running for 168 hours for a total of 5,040 hours at a cost of $75,000. Conventional origin-destination studies would be much more expensive.

University of Maryland (2008) compared the Bluetooth approach to manually driven vehicle probes and depending on assumptions the cost-per-data-point ratio states that the Bluetooth is estimated to be 500 to 2,500 times cheaper.

8 Summary and discussion

The review started with the assumption that the Bluetooth based travel time estimation technology is on par with its alternatives, but more cost-effective and respective of users' privacy. For the most part this assertion was substantiated with the evidence gathered from the papers reviewed. The performance of the system under normal to heavy traffic seemed to be roughly in line with technologies such as automatic licence plate recognition (ALPR), loop detectors, probe vehicles and toll tag reader systems and is adequate to provide informative and true travel time estimations. The cost of the system was less than that of any of the alternatives. The relation between a data point and the user from whom the data was extracted can be obscured so well that revealing the relation it is practically impossible. However, the Bluetooth technology does share the downside that many techniques, which take a sample of the total traffic and calculate the travel time from that, also have. It is susceptible to producing uncertain information when the sample size is small, for example during low traffic. Currently the ways to combat this are very limited and need more research.

The Bluetooth system is comprised of multiple components, many of which have a direct relationship on the sample size and travel time the system can produce. This modularity, however, can be seen as a positive thing. It allows the roadside units to be customized to their environment, which raises the probability of accurate results in varying situations, and raises the adaptability to new research. If, for example, future research develops a new algorithm for filtering outliers or a new type of antenna that is perfect for the system, these can be incorporated painlessly and results seen immediately. The only thing that matters is that the units send detections in a specific format to the host machine regardless of how they are captured. Similarly, it only matters for the end-user that the host machine calculates the travel time in some way, as long as the results are presented in a certain format. Anything in between these can be constantly changed, updated and upgraded. Integrating multiple different Bluetooth systems would be easy if the syntax of the sent detections was standardized, but this is currently not the case.

The travel time system reviewed differs from the rest of the alternatives it was compared against in that it is dependent on another technology. It is strictly tied with the rise and fall of Bluetooth itself. The sample size is in the end bounded by the penetration rate of the Bluetooth devices with the <u>Bluetooth functionality enabled, set in discoverable mode</u>, and situated inside vehicles. If the adoption of Bluetooth drops and it becomes a passing fad, the travel time system based on it will not have a long life-time. As loop detectors and ALPR systems are not probably going to have similar problems in the future, it should be taken into account when calculating possible costs and savings of different systems.

In terms of performance the biggest decision on the hardware design is the choice and installation of the antenna. Due to the number of variables involved no clear rule can be made at this point on what would be the best choice, but it does seem that for the most cases an omnidirectional antenna, or something that resembles it in terms of the road coverage pattern, is a good choice. Directional antennas with a vertical polarization can be used for special environments, where a more focused coverage is needed. A higher gain of an antenna seems to have a modest positive effect. The antennas should be installed at least 2.5 metres above ground between the

carriageways (or on the side of the road if only one carriageway) with no obstructions between the antenna and the traffic. Multiple antenna installations should increase the detection rate, especially when combining multiple types of antennas.

The requirements for the rest of the components seem less important. They should only have modest power consumption, especially for units powered by solar panels, and work well with the software, which on the other hand does have a considerable effect on the performance. The three main critical software algorithms are duplicate filtering, outlier filtering and the travel time calculation itself. The most common way to calculate the travel time is by using a moving average that, for example, calculates for each minute the average travel time from the past ten minutes. Little research has been done comparing the different mathematical methods that can be used to calculate the travel time estimation from the Bluetooth data specifically.

Duplicate filtering plays a role especially when the units are installed on intersections, where the time between the first and the last detection of the same unit may be minutes. Taking the first, last or some combination of the two will determine how intersection delay time will be taken into account in the travel time. No clear guidelines were presented so it is left for the consideration of the planner of the system how they want waiting at intersections to affect the travel time.

Because the Bluetooth technology uses a sample based estimation technique, anything that affects the sample will impact the estimation. Outlier filtering removes some data points from the sample and thus skews the estimations. An ideal algorithm would cause the skewing to always happen towards the actual average travel time by removing data from vehicles that took a different route or stopped during the trip and travel times that are clearly too little or too long, all while being able to adapt to quickly fluctuating traffic. More research is needed to approach this ideal.

The Bluetooth system, especially the hardware, has been demonstrated to be able to provide a good enough sample size for travel time estimation. The future research should especially explore the software techniques that, in the end, make or break the system.

References

Awad, A., Frunzke, T., Dressler, F., 2007. Adaptive Distance Estimation and Localization in WSN using RSSI Measures. *Digital System Design Architectures, Methods and Tools. DSD 2007. 10th Euromicro Conference*

Bakula, C., Schneider IV, W.H., and Roth, J., 2012. Probabilistic Model Based on the Effective Range and Vehicle Speed to Determine Bluetooth MAC Address Matches from Roadside Traffic Monitoring. *Journal of Transportation Engineering*, 138(1)

Beca Infrastructure Ltd, 2011. *Bluetooth Deployment - Puhoi to Warkworth Pilot Study*. Prepared for New Zealand Transport Agency. [pdf] Available at: http://www.bliptrack.com/fileadmin/pdf/traffic/NZ1-4002843-Bluetooth%20Deployment%20-%20Puhoi%20to%20Warkworth.pdf [Accessed 7 June 2012]

Bluetooth.com, n.d. *Bluetooth Basics*. [online] Available at: http://www.bluetooth.com/Pages/Basics.aspx [Accessed 18 July 2012]

Bluetooth Special Interest Group, 2011. *Adopted Bluetooth Core Specifications. Core Version 4.0*. [online] Available at: https://www.bluetooth.org/Technical/Specifications/adopted.htm [Accessed 2 July 2012]

Brennan Jr., T. M., Ernst, J. M., Day, C. M., Bullock, D. M., Krogmeier, J. V., and Martchouk, M., 2010. Influence of Vertical Sensor Placement on Data Collection Efficiency from Bluetooth MAC Address Collection Devices. *Journal of Transportation Engineering*, 136(12), pp.1104-1109

City of Houston, 2011. *City of Houston Launches New Arterial Travel Time Tool in West Houston*. [press release], 3 March 2011, Available at: http://www.houstontx.gov/mayor/press/20110303.html [Accessed 9 July 2012]

Fink, D., 2011. *Anonymous Wireless Address Matching (AWAM) for Travel Time Data Collection*. [pdf] Available at: http://tig.transportation.org/Documents/AdditionallySelectedTechnologies-AST/AWAM-summary.pdf [Accessed 9 July 2012]

Garbe, S., 2011. *Innovative Arterial Bluetooth Project Along CSAH 81*.[pdf] Available at: http://matc.unl.edu/itsheartland/documents/2011_Presentations/Garbe,%20Steve%20-%20Innovative%20Arterial%20Bluetooth%20Project%20Along%20CSAH%2081.pdf [Accessed 3 July 2012]

Haghani, A., Hamedi, M., Sadabadi, K.F., Young, S., and Tarnoff, P., 2010. Data Collection of Freeway Travel Time Ground Truth with Bluetooth Sensors. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2160, Transportation Research Board of the National Academies, Washington, D.C., pp. 60–68

Huang, A., Rudolph, L., 2005. *Bluetooth for Programmers.* [pdf] Available at: http://people.csail.mit.edu/rudolph/Teaching/Articles/BTBook-march.pdf [Accessed 26 June 2012]

Iteris, 2011. *ITS Innovative Idea Project. Arterial Travel Time Monitoring System Using Bluetooth Technology.* Final Report. Submitted to: Minnesota Department of Transportation. [pdf] Available at: http://www.dot.state.mn.us/guidestar/2006_2010/att/Final%20Report.pdf [Accessed 12 July 2012]

KMJ Consulting, Inc., 2010. Bluetooth Travel Time Technology Evaluation Using the BlueTOAD. [pdf] Available at: http://trafficcast.com/docs/PennDOT_BlueTOAD_final_report_incl_charts_4_Jan_20 10.pdf [Accessed 6 July 2012]

Korver, N., 2003. *Adequacy of the Universal Serial Bus for real-time systems.* Internal report, University of Twente. [pdf] Available at: http://doc.utwente.nl/56344/1/Korver03adequacy.pdf [Accessed 26 June 2012]

Luber, A., Junghans, M., Bauer, S., and Schulz, J., 2011. On Measuring Traffic With Wi-Fi and Bluetooth. In: *ITS America, the 18th World Congress on Intelligent Transport Systems*, Orlando, Florida, United States, October 16-18, 2011.

Mills, D., 2012. *Executive Summary: Computer Network Time Synchronization.* [online] Available at: http://www.eecis.udel.edu/~mills/exec.html [Accessed 29 June 2012]

Porter, D.J., Kim, D.S., Magaña, M.E., 2011. *Wireless Data Collection System for Real-Time Arterial Travel Time Estimates.* Oregon State University. Report for Oregon Department of Transportation. [pdf] Available at: http://www.oregon.gov/ODOT/TD/TP_RES/docs/reports/2011/wirelessdata.pdf [Accessed 18 June 2012]

Puckett, D.D., Vickich, M.J., 2010. *Bluetooth®-Based Travel Time/Speed Measuring Systems Development.* Final Report. University Transportation Center for Mobility. Texas Transportation Institute. [pdf] Available at: http://utcm.tamu.edu/publications/final_reports/Puckett_09-00-17.pdf [Accessed 19 June 2012]

Rajbhandari, R., 2008. *Measuring Crossing Times of Passenger Vehicles Using Bluetooth Technology at U.S. - Mexico Border.* In: Cross Border Mobility Meeting, City of El Paso – Ciudad Juarez, City Hall, September 22 2008. [pdf] Available at: http://ttihouston.tamu.edu/Bluetooth/docs/presentation_stakeholders_elpaso.pdf [Accessed 9 July 2012]

Ramadoss, L., Hung, J.Y., 2008. A study on universal serial bus latency in a real-time control system. Industrial Electronics. IECON 2008. 34th Annual Conference of IEEE

RFC 675. Cerf, V., Dalal, Y., and Sunshine, C., 1974. *Specification of Internet Transmission Control Program.* Available at: http://tools.ietf.org/rfc/rfc675.txt [Accessed 18 July 2012]

RFC 768. Postel, J. ed., 1980. *User Datagram Protocol.* Available at:
http://tools.ietf.org/rfc/rfc768.txt [Accessed 18 July 2012]

RFC 793. Postel, J. ed., 1981. *Transmission Control Protocol.* Available at:
http://tools.ietf.org/rfc/rfc793.txt [Accessed 18 July 2012]

Sharifi, E., Hamedi, M., Haghani, A., and Sadrsadat, H., 2011. Analysis of Vehicle
Detection Rate for Bluetooth Traffic Sensors: A Case Study in Maryland and Delaware.
In: *ITS America, the 18th World Congress on Intelligent Transport Systems*, Orlando,
Florida, United States, October 16-18, 2011.

Solon, A.J., Callaghan, M.J., Harkin, J., and McGinnity, T.M., 2006. Case Study on the
Bluetooth Vulnerabilities in Mobile Devices. *IJCSNS International Journal of Computer
Science and Network Security*, 6(4), April 2006.

Steel, P.H.A., Kilburn, P., 2011. Using Bluetooth Technology to Monitor Traffic
Patterns Around Urban Centres in Alberta. In: *ITS America, the 18th World Congress on
Intelligent Transport Systems*, Orlando, Florida, United States, October 16-18, 2011.

Stevanovic, A., Olarte, C. L., Galletebeitia, A., Galletebeitia, B., and Kaisar, E.I., 2011.
Testing Accuracy and Reliability of MAC Readers to Measure Arterial Travel Times. In:
ITS America, the 18th World Congress on Intelligent Transport Systems, Orlando,
Florida, United States, October 16-18, 2011.

Technologic Systems, n.d. *TS-7800.* Available at:
http://www.embeddedarm.com/products/board-detail.php?product=ts-7800
[Accessed 9 July 2012]

University of Maryland, 2008. *Bluetooth Traffic Monitoring Technology --- Concept of
Operation & Deployment Guidelines.* [pdf] Available at:
http://www.catt.umd.edu/documents/UMD-BT-Brochure_REV3.pdf [Accessed 9 July
2012]

Wang, Y., Malinovskiy, Y., Wu, Y., and Lee, U.K., 2011. *Error Modeling and Analysis for
Travel Time Data Obtained from Bluetooth MAC Address Matching.* [pdf] Available at:
http://www.transnow.org/files/final-reports/TNW2011-01.pdf [Accessed 9 July 2012]

Wang, Y., Vrancken, J., Seidel, P., 2011. Measure travel time by using Bluetooth
detectors on freeway. In: *ITS America, the 18th World Congress on Intelligent Transport
Systems*, Orlando, Florida, United States, October 16-18, 2011.

Wasson, J. S., Sturdevant, J.R., Bullock, D.M., 2008. Real-Time Travel Time Estimates
Using MAC Address Matching. *Institute of Transportation Engineers Journal*, ITE, Vol.
78, No. 6, pp. 20-23.

Wieck, M., 2011. Use of Bluetooth Based Travel Time Information for Traffic
Operations. In: *ITS America, the 18th World Congress on Intelligent Transport Systems*,
Orlando, Florida, United States, October 16-18, 2011.