2015

UniShuttle - A Small-Scale Intelligent Transport System in the Connected Mobility Digital Ecosystem

Edward Yong Dou
University of Wollongong

School of Information Systems and Technology
Faculty of Engineering and Information Sciences

UniShuttle - A Small-Scale Intelligent Transport System in the Connected Mobility Digital Ecosystem

Edward Yong Dou

This thesis is presented as part of the requirement for the
Award of the Degree of
Doctor of Philosophy
University of Wollongong

January 2015

CERTIFICATE OF ORIGINALITY

I, Edward Yong Dou declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, at the School of Information Systems and Technology, Faculty of Engineering and Information Sciences, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. Where I have co-authored with others, I have sought and received their permission to re-use materials for the purpose of this PhD thesis. The document has not been submitted for qualifications at any other academic institution.

Edward Yong Dou
January 2015

(Singed) .

ABSTRACT

Public transport contributes greatly to people's mobility. Intelligent Transportation Systems (ITS) are used to provide safer and better transport services to ensure people's mobility. Digital ecosystems, analogous to natural ecosystems, are self-evolving and self-sustaining.

A small scale ITS system that bears the characteristics of a digital ecosystem was designed and developed in this PhD study.

Based on the Internet Engineering Task Force presence model, the system was designed with three discrete but inter-connected components: an on-board device installed on a distributed network of shuttle buses; a central server; and a mobile application that allows users to access the system on the go. Using a combination of different communication protocols (e.g. HTTP, WSDL, SOAP, SSL), the system allows effective communication among the three components and ensures secured communication when needed. Using the Service Oriented Architecture (SOA) principle, the system software functions are designed to provide a stratum of services to the system, with the central layer of services designed to serve the other Connected Mobility Digital Ecosystem applications.

Passengers' travel patterns, shuttle bus travelling time on road, dwelling time at bus stops, and schedule adherence data was gathered and analysed. A historical data based prediction model was developed using an algorithm based on the historical travel data combined with the real-time bus travelling information. It incorporates variables of real-time bus travel speed, route distance, historical travelling time, and historical dwelling

time at each stop before reaching a particular bus stop. The prediction accuracy of travel time by using historical data based model and linear regression model is evaluated using the Root Mean Square Error (RMSE) method in comparison with the prescribed bus timetable. The evaluation results show clearly that the model based predictions of travel time from both linear regression and historical data based models show a significant advantage over using timetable. Furthermore, the prediction accuracy from the historical data based model consistently outperformed that using linear regression model in all the conditions, because the historical data based model has taken into account the dwelling time to improve the prediction accuracy while the linear regression model does not include the effect of dwelling time in its model.

The system collects data from its distributed sensors (i.e. network of shuttle buses equipped with the on-board devices, and the passengers who use the app through smartphones), performs data analysis and predicts bus arrival time. While passengers utilise the system to better plan their travel, the system identifies the repeatable passenger usage patterns, which allows the dynamic scheduling of buses that optimises throughput. This process continues and forms an on-going digital ecosystem cycle, which allows better system performance and management. This system behaviour embodies the CMDE's self-evolving, self-sustaining and mutual benefit design philosophy.

An extensive evaluation study was conducted to assess the system's usability and usefulness. Results showed that passengers highly value the system, with high ratings on the information sufficiency, accuracy, usefulness, relevance, and ease of use. The app's main function, prediction of bus arrival time, attracted the most use. Most

passengers believed the app provided them with the benefits of being able to plan ahead for a trip and elimination of the likelihood to miss a bus.

ACKNOWLEDGEMENTS

Completing a Ph.D. is definitely a vital event in my life, and I would not have been able to complete this long journey without the generous aids and supports from many people over the past four years.

First of all, I would like to express my huge appreciation to my principle supervisor, Professor Peter Eklund, for his faith in me, his inspiration, expertise, continuing support and the opportunities he has given to me. My huge thanks are also given to my co-supervisor, Professor Chris Cook, for his strong support, encouragement and guidance. To them, I am indebted.

I would also like to take this opportunity to express my sincere gratitude to A. Professor Ulrike Gretzel and Mrs Michelle Scarpino for their expertise and great collegial support on the evaluation study for the UniShuttle app I developed in this research work.

My appreciation also goes to my hardworking and intelligent colleagues, Tim Wray, Vu Tran and Jeffery Thom, for their great companionship and for their on-going collegial support during my study. I should also thank Mr Nikola Milosevic from the faculty IT support team for his help in setting up the central server in the UOW IT infrastructure for my study. I also would like to thank Dr Khay Wai See, for sharing his knowledge and experience in using Kalman filter.

Lastly, I would like to express my deep appreciation and love to my family- my parents (Shi Xue Dou and Hua Kun Liu), wife (Holly Dongqing Zhu), and my children (Victor Dou and Kaiyan Dou), for their immeasurable support, patience and encouragement. My love to them forever.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY OF ACRONYMS

3G	Third Generation of mobile telecommunications technology
ABS	Australian Bureau of Statistics
ANN	Artificial neural network
APC	Automatic passenger counter
APTS	Advanced Public Transportation Systems
ARP	Address Resolution Protocol
AVL	Automatic vehicle location
BT	Bluetooth
BAT	Bus arrival time
CBD	Central Business District
CMDE	Connected Mobility Digital Ecosystem
CPU	Central Processing Unit
DSRC	Dedicated Short-Range Communication
EKF	Extended Kalman filter
GIS	Geographic information system
GPRS	General Packet Radio Service
GPS	Global positioning systems
HTTP	Hypertext Transfer Protocol
IETF	Internet Engineering Task Force
iOS	iPhone Operation System
ICT	Information and Communication Technology
IT	Information Technology
ITS	Intelligent Transportation Systems
KML	Keyhole Markup Language
KNN	K Nearest Neighbours
LBS	Location Based Service
LIS	Location Information Server
MAC	Media Access Control
ML	Machine learning
SQL	Structured Query Language
PC	Personal Computer

PHP Hyptertext Preprocessor

RMSE Root Mean Square Error

SME Small and Medium Enterprises

SMS Short Messages service

SOA Service Oriented Architecture

SOAP Simple Object Access Protocol

SSL Secure Socket Layer

SSNN State Space Neural Networks

TCP/IP Transmission Control Protocol/Internet Protocol

UOW University of Wollongong

VANET Vehicular ad-hoc network

V2V Vehicle to vehicle

V2I Vehicle-to-infrastructure communication

WLAN Wireless Local Area Network

WSDL Web Service Description Language

WWAN Wireless Wide Area Network

PUBLICATIONS TO DATE ARISING FROM THIS RESEARCH

Dou, E., Eklund, P., Wray, T., & Cook, C. (2013). Connected Mobility Digital Ecosystem: A Case Study on Intelligent Transport Analytics.

Eklund, P., Thom, J., Wray, T., & Dou, E. (2011). Location based context-aware services in a digital ecosystem with location privacy. *Journal of Cases on Information Technology (JCIT)*, 13(2), 49-68.

CHAPTER 1. OVERVIEW OF THE STUDY

1.1 Background

Thomson Reuters (2014) has predicted a digital world by 2025 where everything, everyone and everywhere are digitally connected all the time. This connectedness will dominate the world in all economic activities and people's lives. Public transport contributes greatly to people's mobility, and public transport systems are a critical component of a digitally connected world or, as some say, the ultimate digital ecosystem.

Digital ecosystems, analogous to natural ecosystems, are self-evolving and self-sustaining (Badr & Caplat, 2010; Briscoe, 2009; W. Li, Badr, & Biennier, 2012). The digital ecosystem concepts and principles have inspired many researchers to design systems that are reminiscent of natural ones. Intelligent Transportation Systems (ITS), transportation systems that have intelligence built into them are used to provide safer and better transport services to ensure people's mobility (Giannoutakis & Li, 2012; Miles & Chen, 2004). Among many research areas within ITS, one focus is to design a system that can increase the use of public transport, improve passengers' satisfaction and increase the efficiency level of public transport management, while reducing traffic congestion and environmental pollution at the same time.

It was against this background, that the research work in this thesis has had the aspiration to build an ITS within the framework of the digital ecosystem that can serve the local community with better public transport information, and can self-evolve and self-sustain.

1.2 Objectives and Scope

The goal of this research was to build an ITS within the Connected Mobility Digital Ecosystem (CMDE) context at the University of Wollongong. Its key objectives are to design and develop an self-evolving and self-sustaining ITS to help both passengers and transport authority; to provide meaningful system functions to users, such as predicting bus arrival time at any particular bus stop; to develop a mobile application which provides users access the system with ease and on the go; and to evaluate the system on its usability and usefulness to further help the system to improve. With the many potential aspects of an ITS project, the scope of this research work is confined in the area of providing passengers with well-informed travel information.

1.3 Relevance and Significance

Wollongong is a typical Australian small city with a population of over 203,000 people (ABS, 2014). It is the home of two main campuses of UOW which has over 22,000 students and staff who study and work on these two campuses (UOW, 2013). New South Wales Transport and Infrastructure serves this community by providing fixed-route bus services (37 routes), a free shuttle around the city (3 routes), and the greater Sydney railway service. As a regional city, the frequency of the fixed-route bus service is lower than in larger capital cities.

To allow more passengers to take advantage of the public transport system, it is vitally important to provide a means to passengers to access more readily available information of bus’s arrival time. Doing so will allow passengers to save precious time on waiting for buses and to plan trips better and earlier.

On the other hand, due to the university’s teaching schedules, buses could be very full at times, while at other times, have few or no passengers on-board. The transport authority

had little ability to make informed service arrangements besides running to a fixed timetable. However, fixed timetables are often unreliable, which can cause students to miss classes, and resulted in numerous anxieties among passengers.

On a bigger picture, public transportation systems play a vitally important role in modern life. Providing travellers with real travel time information is an important aspect of ITS development. So far, Australia's ITS developments have been haphazard (PwC Digital Pulse 2014). Due to the market complexity, varied stakeholder interests and austerity in the government spending on public transport, there has been a lack of centralized approach towards ITS implementation in a regional city like Wollongong. Community and industry want ready technologies and can be deterred by the length of time between conducting the research and seeing tangible results.

This research strives to develop an ITS application that is research based but is also practical to the Wollongong community. The designs and implementations presented in this PhD research may be an alternative for the deployment of more efficient and cost-effective technological solutions. The research outcomes can be applied to the public transport systems in the other cities, but may also have implications to freight, logistics, and other aspects of transportation systems.

1.4 Contributions

This PhD study uses a digital ecosystem philosophy to underpin an ITS design. As one of the earliest projects delivered in CMDE, it fulfilled CMDE's vision of developing systems that exhibit the hallmark of a connected digital ecosystem using distributed sensors (network of shuttle buses and passengers with mobile devices).

The overall contributions of this PhD thesis include the system architectural design, component design and implementation, bus travelling data analysis, algorithm development for predicting bus arrival time, front-end user app design, system integration, and system evaluation. Its key contribution lies on its system design and implementation.

The research work delivered a low-cost and user highly valued ITS application. To this day, the system has been running for more than three years in the Wollongong community.

1.5 Chapter Organization

The chapter organisation of this thesis is briefly outlined as follows:

Chapter 1 describes the general background, objectives and motivation of the thesis, and the PhD candidate's contributions.

Chapter 2 is dedicated to a literature review in areas including digital ecosystem concept, principles and architecture; the CMDE at UOW; the ITS concept and developments; the major technologies in ITS; and the existing mobile applications for public transport.

Chapter 3 concentrates on the UniShuttle system design and development, including architecture design, software design, communication design, and the implementation of the three UniShuttle components (on-board device, central server and UniShuttle app).

Chapter 4 covers the acquisition and processing aspects of the system data analysis, development of a historical data based prediction method for bus arrival time prediction, and the prediction model's evaluation.

Chapter 5 explains the system evaluation study. The system evaluation involves a two-phased survey, including a respondent evaluation on the UniShuttle app and users' privacy needs relevant to a location-based mobile application.

Finally, Chapter 6 summarizes the overall PhD work outcomes and suggests future research work.

CHAPTER 2. LITERATURE REVIEW

2.1 Digital Ecosystems

2.1.1 The Emergence of Digital Ecosystems

Maley (1995) briefly described a digital ecosystem as a direct connection between biological properties and digital technology. This is possibly the earliest mention in literature of the digital ecosystem concept. Until 2002, the digital ecosystem concept rarely attracted academic attention.

In 2002, the European Commission's Framework Programme for Research and Technological Development funded digital ecosystem research to fulfil its ambition for higher growth, knowledge based economy and greater social inclusion (Nachira 2002). The project's major purpose was to empower small and medium sized enterprises (SME) to adopt and develop Information and Communication Technology (ICT) into their local innovation system ("Digital Ecosystems for eBusiness - Reference pages," n.d.). By that time, the digital ecosystem concept had gained prominence. Its context to ICT distinguishes it from any other business ecosystems set in a socio-economic context (Nachira 2002).

2.1.2 Defining Digital Ecosystems

Life and its physical environment interact with and are inter-reliant on each other in a natural ecosystem. A digital ecosystem transposes the concept of the natural ecology to the digital world, using the metaphors and the mechanism of natural ecosystems ("Digital Ecosystems for eBusiness - Reference pages," n.d.). It can be seen as the counterpart of a natural ecosystem only in the digital space (Briscoe, 2009; E. Chang,

Quaddus, & Ramaseshan, 2006; E. Chang & West, 2006; H. Dong, Hussain, & Chang, 2009).

A digital ecosystem has been described as an "open, loosely coupled, domain clustered, demand-driven, self-organizing and agent-based environment, in which each species is proactive and responsive for its own benefit and profit" (E. Chang et al., 2006). In describing the species inside a digital environment, Fu (2006) suggests the digital species or digital components include software components, applications, services, knowledge, business processes and contractual frameworks. Further, Francesco Nachira, Nicolai, Dini, Rivera Lèon, and Le Louarn, (2007) suggest the software components and agents inside a digital ecosystem can behave independently and can self-select and self-adapt. Other studies suggest that digital ecosystems have the characteristics of being self-organising, scalable, sustainable, (self-) adaptive and flexible (Badr & Caplat, 2010; Briscoe, 2009; W. Li et al., 2012).

Similarly, responding to lack of a holistic point of view in previous research, W. Li et al. (2012, p 119) describe a digital ecosystem as "a self-organising, scalable and sustainable system, composed of heterogeneous digital entities and their interrelations focusing on interactions among entities to increase system utility, gain benefits, and promote information sharing, inner and inter cooperation and system innovation." This all-inclusive definition, although lengthy, perhaps accounts for all of the digital ecosystem's characteristics including the fact that the interaction between digital entities inside a digital ecosystem has benefit for the entities themselves as well as the system. This critical characteristic has also been pinpointed by Eklund, Thom, Wray, and Dou, (2011), who used it as a key criterion for analysing a digital system.

2.1.3 Applications of Digital Ecosystems

Research has rarely been conducted to gain a big picture of the body of research work except for an effort by W. Li et al. (2012) who categorised digital ecosystems according to their applications, as business, knowledge management, service, and education (W. Li et al., 2012).

The majority of the digital ecosystem applications are in the category of business. The digital business ecosystem was the earliest application targeted by the European Commissions in order to support a diffused network of SMEs (Nachira 2002; Nachira et al., 2007). Digital business ecosystems are referred to as “digital environments” with an infrastructure that supports structural services and economic relationships. It denotes a self-organising and self-sustainable model including technologies, e-business models, services, and business knowledge sharing (Badr & Caplat, 2010).

A simplistic architecture describing the concept of a digital business ecosystem is included in Figure 2-1. As can be seen, a common digital ecosystem infrastructure is the foundation, and built on it are the structural services, such as accounting, billing, data or security. On the very top are the aggregated and personalised ICT-enabled digital ecosystem services. This logic of a digital business ecosystem’s architecture is not only suitable for guiding business application but also for applications in the four remaining categories.

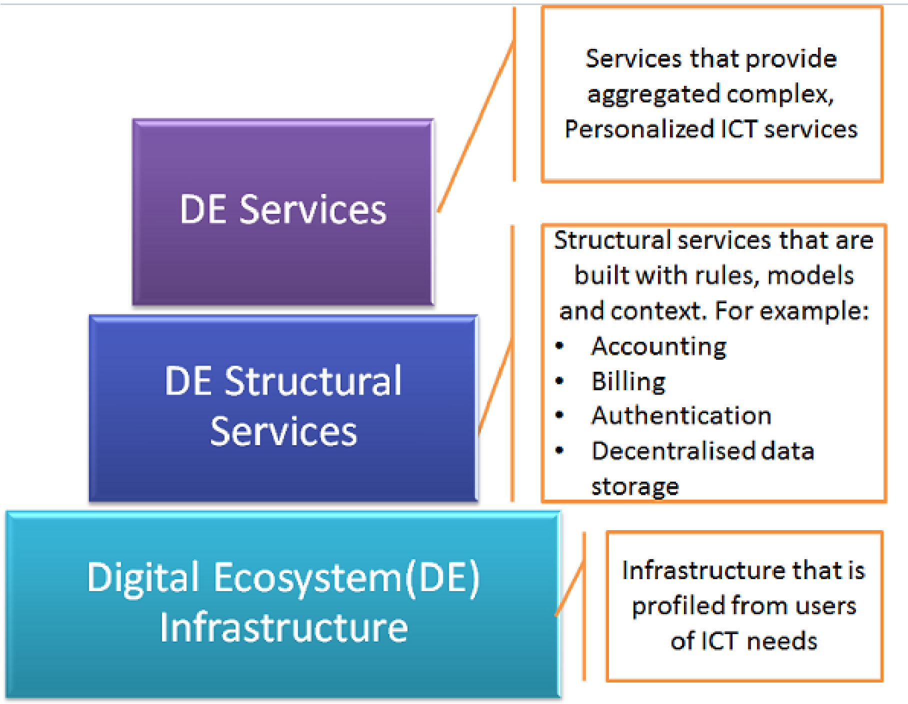


Figure 2-1: The Digital Ecosystem Concept Proposed by European Commission (Adapted from Nachira et al., 2007)

Numerous experimental digital ecosystems have been developed according to literature, such as Gary et al. (2013)'s ONE.MOTORING portal project which is a digital ecosystem developed and governed on a systematic architecture. Biuk-Aghai, Tang, Fong, and Si, (2009) argue that the Wiki is a good example of a digital ecosystem, as it is a system that has contributions by a community of volunteers, with all the characteristics of a digital ecosystem.

Similarly the concept of a digital ecosystem is applicable to knowledge management and other fields. Examples include a system that makes agriculture information easily accessible and further strengthens the links between knowledge, skills, technology and market to improve agricultural competitiveness (Chatterjee, Prabhakar, & Sarkar, 2007; Miah, 2012); a medical records system that enables efficient operation and diagnosis in hospitals (Hadzic, Dillon, & Chang, 2007); and Peter Eklund, Thom, Wray, and

Thomson, (2010)'s applied digital ecosystem to a university community for better services to students, staff and citizens.

2.1.4 Development of Digital Ecosystems

Many efforts have been observed in digital ecosystem development in the technological space. Briscoe and De Wilde, (2008) studied the computing technologies, such as Multi-Agent Systems, Service-Oriented Architectures (SOAs) and distributed evolutionary computing, that can contribute to allowing a digital system to possess features such as a self-organising and self-evolving nature. Using the Multi-Agent technology, Lurgi and Estanyol, (2010) proposed the first digital ecosystem of this kind.

From an architectural and methodological perspective, Briscoe, Sadedin, and De Wilde, (2011) expanded the existing digital ecosystem development by defining architectural principles using not only the technologies in Briscoe and De Wilde, (2008)'s study, but also using mobile agent systems collectively. Among other efforts (such as Hadzic et al. 2007; Hadzic & Dillon, 2008, who designed a health domain ecosystem), Stanley, and Briscoe, (2010) conceptualised a digital system's architecture into three layers: a coordination layer, which deals with distribution and coordination; a resource layer, which provides resources; and a service layer, which provides the end-user-accessible services. Notably, the authors also pointed out that whereas the architecture and functions seem to be conceivable in a distributed and heterogeneous ecosystem, it is more challenging to implement. It has been identified that most of the development in architecture in the literature only provides simple descriptions without the substantiation of the implementation of the concept into practical application (Li et al., 2012).

Further research also tackles methodologies and strategies that can optimize digital ecosystems (Briscoe & De Wilde, 2006; Briscoe et al., 2007; Briscoe & De Wilde,

2008; Briscoe, 2009; Briscoe, 2010; Briscoe et al., 2011; Briscoe & De Wilde, 2012, D'Angelo, Ferretti, Ghini, & Panzieri, 2011); design holistic framework that includes social, economic and environmental development (Huang, Hsueh, & Reynolds, 2013).

2.1.5 A Digital Ecosystem Based Intelligent Transportation System Concept

By combining the Intelligent Transportation System (Section 2.3) concept and the Digital Ecosystem concept (i.e. to design and implement an ITS system within the digital ecosystem principles), a digital ecosystem service in the transportation domain is conceived. In such a digital ecosystem, the "digital organisms" are passengers, bus drivers, community, operators, agencies and developers. These "digital organisms" interact in their environment comprising the city, vehicle routes, internet and mobile devices. All parties in this ecosystem benefit through their interactions and are self-evolving and self-sustaining.

2.2 Connected Mobility Digital Ecosystem

Eklund et al. (2011) proposed the CMDE concept, which is defined as the combination of the mobile software applications developed in the context of the University of Wollongong campus community and their shared infrastructure. The authors use the ecosystem metaphor to conceptualise the relationships between the players in the campus community: "The concept of a digital ecosystem resonates throughout the theme … and the term is used to describe the complex social-technical systems that result from multiple, spontaneous interactions with a particular ICT-enabled environment." (Eklund et al., 2011, p52). Four location-sensitive, context-ware applications (apps) were proposed (Eklund et al., 2010, Eklund et al., 2011). They are:

- UniShuttle – the front end application that is the subject of this thesis.

- Car-pooling application – a service that offers a demand driven marketplace for car-pooling and ridesharing.
- Freshman application - a ‘helper’ application for first-year university students.
- Art Collection Tour Guide - an application that allows custom tours of the artworks on a university campus based on a user’s location.

The authors further identified two similar characteristics they observed between a real ecosystem and a CMDE in order to shed more light on the CMDE concept. Firstly, as in a real ecosystem where interactions between species are simple, spontaneous, and can be self-motivated by the individuals’ willingness to survive, in the environment of the CMDE, the interactions in the University are self-governing, self-directed and often are for the users’ own benefits. Secondly, the interactions and their environment, both in the real ecosystem and in the CMDE, contribute to a sustainable whole, and both environments are capable of self-evolution.

Three key motivations for CMDE development can be summarised as below:

- Design of a robust architecture for location determination and location privacy;
- Provision of optimal services through aggregating and analysing individual inputs and actions;
- Provision of a platform for useful services in the university community.

2.3 Intelligent Transport System

2.3.1 Overview of ITS

The Intelligent Transport Systems (ITSs) concept emerged in 1970s (Qi, 2008), much earlier than the digital ecosystem. ITSs are defined as technologies, systems and services that provide safer and better transport services with the benefits of improving productivity, safety and service satisfaction (Casey et al., 2000; Miles & Chen, 2004; Qi, 2008).

The ITS systems have opened up many new opportunities (Giannoutakis & Li, 2012, PwC Digital Pulse, 2014). Literature has shown related works range from transportation safety, e.g. lane departure warning (Braitman, McCartt, Zuby, & Singer, 2010; Cheng, Jeng, Tseng, & Fan, 2006), to payment financial systems, e.g. electronic toll collection (C.-D. Chen, Fan, & Farn, 2007; W. H. Lee, Tseng, & Wang, 2008), to automatic highway systems, e.g. car platooning (Clement & Taylor, 2006), to travel time prediction (Chien & Kuchipudi, 2003; Ishak & Al-Deek, 2002; H. L. H. Liu, Lint, Zuylen, & Zhang, 2006; H. Liu, van Zuylen, van Lint, & Salomons, 2006; Vanajakshi, Subramanian, & Sivanandan, 2009).

In the public transportation domain, the connotation of ITS is synonymous to the Advanced Public Transportation System (APTS), which is defined as "a collection of technologies that increase the efficiency and safety of public transportation systems and offer users greater access to information on system operations" (Casey et al., 2000, p xii).

2.3.2 ITS Technologies Categorisation

Although diverged and seemingly complex, according to the technologies' relevance to transportation application, the U.S. Department of Transportation (Casey et al., 2000) has categorised ITS used in the public domain into five broader groups, they are:

I. Fleet management systems;
II. Travel information systems;
III. Electronic payment systems;
IV. Transportation demand management; and
V. Transit intelligent vehicle initiatives.

The ITS systems vary according to the technologies they use (Qi, 2008). For example, the technologies used in the Fleet Management Systems include Automatic Vehicle Location (AVL) systems, transit operations software, communications systems, Geographic Information System (GIS) and automatic passenger counters; whereas in the Transit intelligent vehicle initiatives, technologies such as lane change warning and collision avoidance are used. Technologies used in traveller information systems also include AVL, communications technology, GIS and some include pre-trip transit traveller information systems, in-terminal and in-vehicle transit information systems. The specific technologies used in ITS related to this thesis project are to be reviewed in sections 2.4, 2.5, and 2.6 of this chapter.

2.3.3 Traveller Information Systems in ITS

As a key application category in ITS, Traveller Information Systems (also called "Passenger Information Systems"), use various technologies to track bus locations in real time along with other data that are useful in predicting bus arrival timings at each stop along the route (Kale, 2014). The information provided to travellers allows passengers to plan their travel time efficiently and reach the bus stop just before the bus arrives rather than waiting for a delayed bus passively. The systems will also allow public transportation systems to attract many passengers by encouraging them to use public transport system rather than private vehicles (Lyons, 2001), which results in less traffic and reduces pollution (Lyons & Harman, 2002).

The advent of the internet has helped public transport agencies put public transport information online (Gildea & Sheikh, 1996). Since then, public transport agencies around the world have developed services such as interactive route maps, and trip planning online and email alerts by subscription (Transport For London, 2014; Trains

Transport Sydney, 2014). With the prevalence of smartphone and wireless broadband technologies, mobile optimized webpages or apps are also being offered.

Real-time traffic data acquisition and the sharing of traffic information with passengers are the essential parts of an advance traveller information system. Real-time vehicle tracking can lead to more efficient operations and also enhanced sense of safety for transit operators (Miller et al., 2008, Moreira, 2008). However, issues such as satellite coverage, cost and real-time requirements, make an efficient traveller information system solution challenging and hence has motivated academic research in this area. One exemplary work is a collaborative traffic information framework proposed by W. Lee, Tseng, and Shieh, (2010). The authors proposed architecture for traffic information generation and distribution. The framework (shown in Figure 2-2) has three parts: location-aware mobile devices (left on the figure), such as smart phones with GPS functions, which are used to gather location information; traffic information centre (central on the figure) for data acquisition, central services for data storage, fusion and prediction computation; and external traffic information data sources (right on the figure), which are plugged as extra information source to be computed by the information centre. All elements are connected through mobile and radio network.

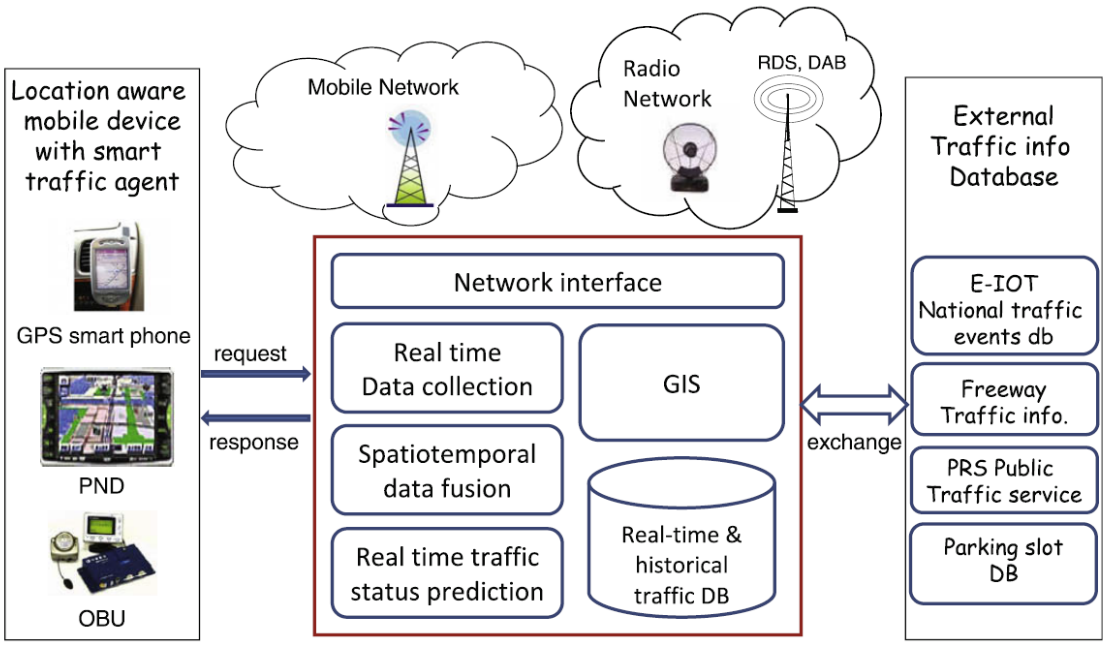


Figure 2-2: The collective traffic information generation and distribution system (W. Lee et al. 2010)

Recent traveller information system improvements have focused on real-time arrival information provision or the prediction of bus arrival data. An extensive review on the prediction techniques is included in Section 2.5.

2.4 Automatic Vehicle Location (AVL) Technologies

2.4.1 Overview of AVL

The identification of vehicle location can be a great challenge in day-to-day transport operations, as a vehicles location can be affected by many uncertain variables such as traffic congestion, unpredictable incidents, driver behaviours, and passenger flow. With many efforts made to develop technologies to locate vehicles in public transport system, Automatic Vehicle Location (AVL) technologies started to be in use the 1970s (Riter & McCoy, 1977) and have become the most commonly used technology in ITS. AVL fulfilled what Loukakos (2001) outlined as the primary objective of transport agencies by increasing the transport companies' capability to monitor information on vehicle

position and operational status, hence increasing vehicle utilization and reducing cost of fuel, labour and capital. AVL can also help improve vehicle safety by allowing quick location and response to incidents and emergencies.

The key modern AVL technologies include dead-reckoning, signpost and odometer, global positioning systems (GPS), and differential GPS (Abbott & Powell, 1999; Borenovic, Neskovic, & Neskovic, 2013; French, 1986). A single location technology is usually insufficient for determining vehicle position in situations where tall buildings block signals, which can result in multi-path errors. Hence it is common practice that the primary location technology is supplemented by another location technology.

The key AVL technologies are examined in the following sections.

2.4.2 Dead-reckoning Method

In vehicle location applications, the dead-reckoning method is used to predict a vehicle's current position based on a position previously determined and advancing that position using estimated speeds over elapsed time and course (Lezniak, Lewis, & Mcmillen, 1997). Although dead-reckoning is relatively inexpensive, this method has a number of drawbacks. For example, uneven surfaces and hills can cause errors in predicting vehicle positioning and travel distance (Loukakos 2001); it also requires the reset of equipment in certain situations. Thus, in practice, dead reckoning is commonly combined with other location techniques, for example Xu et al. (2009) developed a method to integrate Wi-Fi networks with the dead-reckoning method and Global Positioning System (GPS) to track a vehicle in an urban area to achieve a more precise vehicle location.

2.4.3 Signpost and Odometer

Signpost and odometer use a series of radio transmitters placed along bus routes (Dessouky, Hall, Nowroozi, & Mourikas, 1999; Leong, 1989). When a low powered signal is generated and sent by transmitters on poles along the route, a receiver on a vehicle detects the signal when this vehicle passes the poles. According to the odometer, the travelled distance is taken, and the vehicle's position is hence reported. Conversely, receivers can be installed along the bus routes while transmitters can be configured on the vehicles. The advantages of this technology include low in-vehicle cost and proven reliability. The disadvantages are that when the bus route changes, additional installation of transmitters or receivers will be required and the vehicle cannot be tracked when a bus is off-route (Loukakos, 2001).

2.4.4 Global Positioning System (GPS)

Among the AVL techniques, GPS is perhaps the most popular technique (Nguyen and Barth 2006). GPS is the worldwide satellite-based radio navigation system, consisting of a constellation of 24 satellites that are equally spaced in six orbital planes and orbit around Earth at a height of 20,200 kilometres (Djuknic & Richton 2001). The satellites send signals, which are picked up by the GPS receivers in vehicles or by mobile handsets. The receiver can calculate its own position based on signals received from at least 3 satellites through the technique known as triangulation (Bajaj, Ranaweera, & Agrawal, 2002). The vehicle location information can then be reported to a central server and compared with the local duplication of geo data to calculate its actual location on Earth (Porcino 2001).

Although GPS is widely used, efforts have been made to increase the vehicle location calculation accuracy. For example, Chadil, Russameesawang, and Keeratiwintakorn,

(2008) proposed improved two-dimensional and three-dimensional location identification means for vehicles. The advances in AVL systems based on GPS have provided the transit industry with tools to monitor and control the operation of their vehicles and manage their fleets in an efficient and cost effective way (Papadoglou & Stipidis, 2001).

2.4.5 Other Technologies

The database technologies such as GIS and communication technology integrated together have made vehicle tracking systems possible. Many related systems have been proposed in literature, such as centralized automatic vehicle location communications software as proposed by Al-Bayari, and Sadoun, (2005) which uses a GIS environment and GPS. Aloquili, Elbanna, and Al-Azizi, (2009) proposed a software tracking system which can pinpoint vehicle position, ground speed and fuel level. This system aimed at control fleet of vehicles by implementing geo-fencing and optimal path detection techniques.

Two types of wireless communication used in vehicle tracking systems are vehicle-to-vehicle communication (V2V) and vehicle-to-infrastructure communication (V2I). Thangavelul, Bhuvaneswari, Kumar, Senthilkumarl, and Sivan, (2007) developed a vehicle tracking systems called VANET(Vehicular ad-hoc network) by using both V2V and V2I, in which the vehicles are considered as mobile nodes in the network, and these vehicles and roadside equipment form an ad-hoc network which has no pre-arranged infrastructure (Doan, Berradia, & Mouzna, 2009). Further, Fukumoto et al. (2007) are developing Contents Oriented Communications by employing V2V and V2I wireless communication, which provides more functions for roadway security and management.

While large transit operators have set up their own communication network, many smaller agencies, especially regional/rural agencies with large service areas but fewer vehicles, have found it more economical to simply utilise their local cellular carrier.

2.5 Travel Time Prediction Technologies

Travel time information is a vital component of many ITS applications, particularly in public transportation systems. Static schedules and timetables are an important information source for public transport users. However, public transportation vehicles do not always run on time.

Literature has suggested different techniques for travel time prediction, including historical data based models, real-time algorithms, regression models and time-series techniques, Kalman filtering techniques, and machine learning models.

2.5.1 Historical Data Based Model

Vehicle travel and arrival time can be predicted from the historical travel data obtained from previous journeys in the same time period (Chien & Kuchipudi, 2003). This can be a reliable method for arrival time prediction if the traffic pattern remains stable over a long period of time. There are several ways to establish and use historical travel data. One of these is to use average travel time on each segment of the particular route and predict the future bus travel time in the same segment. In this model, the prediction accuracy largely depends on a function of the correspondence between the real-time and historical traffic patterns (Chien, Ding, & Wei, 2002). Variations in historical data or changes in the correspondence between historical data and real-time data could cause significant deviation in the prediction results (Smith & Demetsky, 1994). The problem in these methods is usually attributed to the change in dwell time at stops if the

prediction model is in real time (DeLurgio, 1998; Stephanedes, Kwon, & Michalopoulos, 1990). M. P. D'Angelo, Al-Deek, and Wang, (1999) used a non-linear time series model to predict corridor travel times on a highway. It was found that the single variable model using speed was actually better than the multivariable prediction model.

W.-H. Lin, and Zeng, (1999) developed a mathematical algorithm to provide real-time bus arrival information, in which they consider schedule information, bus location, difference between scheduled and actual arrival time, and waiting time at stops in their algorithm, although they could not consider traffic congestion and dwell time at stops. Williams, and Hoel, (2003) extended W.-H. Lin and Zeng's method to include traffic conditions that are assumed to have consistent daily routines and weekly patterns. In this case, the historical averages of the travel situations at a particular time of day and day of the week provide a reasonable prediction of future conditions. Williams and Hoel's models is reliable if the traffic pattern in the area of interest is relatively stable or reasonably periodic, which can often be the case in small regional cities and towns.

F. Li, Yu, Lin, and Min, (2011) proposed a statistical approach to predict the arrival time at each stop for public buses. Based on the assessment of all factors including departure time, driver characteristics, dwell time, intersections, traffic conditions etc., the authors proposed a linear model by analysing the historical data of the bus arrival time. Their approach has shown some improved travel time prediction accuracy of the prototype.

Another way to forecast travel time is to use average speed on each segment of the route from historical data. The travel data can be collected using GPS technology as the distance taken over the bus routes can be calculated using the vehicle position

information. F. Li et al. (2002) presented a method to estimate bus arrival times at bus stops using GPS information, which was implemented in the Information System for Urban Bus Transportation, in Brazil. The authors utilised the historical bus travel speed along the route segment and current speed of the bus derived from the GPS data to make an empirical calibration. As a result the modified model was found to be satisfactory in the implementation and experiment. The same kind of model was later developed by Sun et al. (2007) with a slight modification. The authors suggested that the current speed of a bus is usually a more important factor in influencing how fast the bus will travel over the coming stops. Their algorithm includes a real-time bus tracking model and a bus arrival time prediction model. Their results indicated that the proposed system achieved better prediction on bus arrival times.

Historical data based models require an extensive set of historical data, which may not be available in practice, especially when the traffic pattern varies significantly from time to time. These models are suitable for small and medium cities where both travel time and dwell time are relatively stable. Its accuracy largely relies on the similarity between the real time and historic traffic patterns. In the present study these models were used, as Unishuttle in Wollongong city does not have substantial variation from time to time.

2.5.2 Regression Model

Regression models predict and explain a dependent variable with a mathematical function formed by a set of independent variables (Chien et al., 2002). Unlike historical data based prediction models, these approaches can work better than using historical data based models under unstable traffic conditions. Regression models usually measure the simultaneous effects of various factors, which are independent of each other,

affecting the dependent variable. Patnaik, Chien, and Bladikas, (2004) proposed a set of multi-linear regression models to estimate bus arrival times using the data collected by automatic passenger counters (APC).

Jeong (2004) also developed multilinear regression models using different sets of inputs. The advantage of multilinear regression models is their capability to distinguish the different contribution of various variables to the travel time prediction.

2.5.3 Kalman Filtering Model

Kalman (1960) describes a recursive solution to the discrete-data linear filtering problem. With significant advances made in digital computing, the Kalman filter technique has become a major topic in both research and applications. The word "filter" refers to data processing language which is equivalent to a black box, or in practical applications, the filter is just a computer program in the central processor. The filter supports estimations of past, present, and even future states through continuing optimisation processes.

Welch and Bishop, (2006) describe the Kalman filter's working principle as follows. The Kalman filter estimates a process by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of measurements. There are two groups: time update equations (predictor) and measurement update equations (corrector), both are used for the Kalman filter calculation. The predictor equations are responsible for projecting forward the current state and error covariance estimates to obtain the priori estimates for the next time step, which is to incorporate a new measurement into a priori estimate to obtain an improved estimate using the corrector equations for feedback. A final estimation algorithm

resembles that of a predictor-corrector algorithm for solving numerical problems as shown below in Figure 2-3

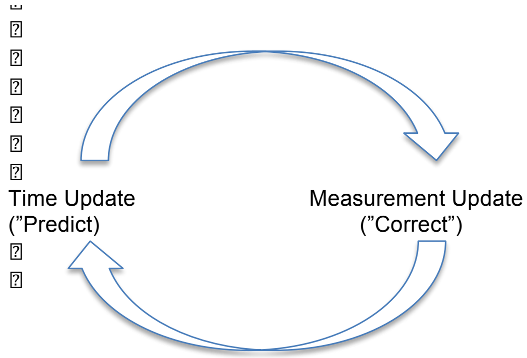


Figure 2-3: The ongoing discrete Kalman filter cycle.
The time update projects the current state estimate ahead in time. The measurement update adjusts the projected estimate by an actual measurement at that time (Welch & Bishop, 2006)

Because of its versatility, the Kalman filter has been widely applied in fields including navigation, space science, military and defence and more. Kalman Filter has also been used in transportation fields, in particular, for predicting bus travel time and arrival time (M. Chen, Liu, Xia, & Chien, 2004; Chien et al., 2002; Vanajakshi et al., 2009; Wall & Dailey, 1999). The basic function of the model is to provide estimates of the current state of the system. But it also serves as the basis for predicting future values or for improving estimates of variables at earlier times, i.e., it has the capacity to filter noise (Lezniak et al., 1997; Thomas, Weijermars, & Van Berkum, 2010).

Vanajakshi et al. (2009) used GPS data collected from public transportation buses on urban roadways in the city of Chennai, India, to predict travel times under heterogeneous traffic conditions using an algorithm based on the Kalman filter technique. The performance of the proposed algorithm is found to be promising and expected to be valuable in the development of advanced public transportation system (APTS) in India.

2.5.4 Artificial Neural Network Models

Artificial neural network (ANN) is one of the typical techniques used in Machine Learning (ML). ML is about the construction and study of systems that can learn from measurement data with two stages: choosing a candidate model, and predicting model parameters by learning the existing data (Jin & Sendhof, 2008). ML methods have certain benefits compared to the other prediction methods, for example it can handle complex variable relationships with large amounts of data present (Altinkara & Zontul, 2013); it can process non-linear relationships between various predictors, and it can treat complicated and noisy data (Recknagel, 2001).

ANN is motivated by simulating the human brains' neural networks, and it is presented often as a mathematical or computational model. It is an adaptive system that changes according to the external and internal information during its learning processes through modelling complex relationships between inputs and outputs and identifying patterns in data. The ANN models are often constructed with multiple layers of processing units ("artificial neurons"). The neurons contain activation functions and are highly interconnected with one another. Information can be processed in a forward or feedback direction through fully or partially connected topologies (Hagan, Demuth, & Beale, 1996).

ANNs have been applied in forecasting public bus arrival time because of their ability to solve complex non-linear relationships (M. Chen et al., 2004; Chien et al., 2002; Y. Lin, Yang, Zou, & Jia, 2013; Park & Lee, 2004). Y. Lin et al. (2013) proposed ANN models to predict the real-time bus arrivals, using historical GPS data and automatic fare collection system data, and the authors have argued that ANN can outperform Kalman filter models if applying more sophisticated implementation. In a similar fashion,

Mazloumi, Rose, Currie and Moridpour (2011) proposed an approach and applied an ANN methodology to predict bus travel time over four bus route sections in Melbourne, Australia with a results showing robust prediction. Jeong and Rilett (2005) compared the performance of historical data based model, regression model and ANN model for bus arrival time prediction, and also claimed the ANN model performing better than the other two models in these circumstances.

The aforementioned cases show that the ANN models have advantages over other approaches for bus travel time predictions, especially when there are serious traffic congestions or high uncertain traffic conditions. However, the ANN approaches often require lengthy training and learning processes with large amounts of data with sufficient characteristic examples, which is often unpractical in real world situations (Jeong, 2004). Other disadvantages are also being identified. For example, ANN methods are more difficult to use in the field; ANN methods require greater computational resources; lastly, ANN models are empirical and have many methodological issues that haven't been resolved (Tu 1996, Yilmaz 2010).

2.5.5 Hybrid Model

The models described above all have their advantages and disadvantages. Some hybrid approaches have been proposed in research by combining two or more of these models to take advantages of the individual model strengths. For example, Wall and Dailey, (1999) propose an algorithm which uses two components: tracking using Kalman filter and prediction using historical statistical estimation. Chu, Oh, and Recker, (2005) develop an adaptive Kalman filtering-based travel time prediction method with consideration of the noises in the model system. H. L. H. Liu, Lint, Zuylen, and Zhang, (2006b) develop a hybrid model based on State Space Neural Networks (SSNN) and the

Extended Kalman Filter (EKF) approaches. While the SSNN model requires a large data set for offline training, they use an EKF model to train the SSNN. J. Dong, Zou, and Zhang, (2013) propose a mixed model to predict bus arrival times by using historical data in combination with real-time data to predict the future vehicle travel times with improved accuracy. The authors proposed short distance bus arrival time (BAT) prediction based on real-time traffic conditions and long distance BAT prediction based on K Nearest Neighbours (KNN) respectively. Here, the KNN is a non-parametric method used for classification and regression and a type of instance-based learning. In the empirical studies with real data from buses, their model outperformed the ANN approaches in accuracy and efficiency.

Zaki, Ashour, Zorkany, and Hesham, (2013) have compared the different techniques and concluded that overall, the historical approach, regression approach and the real time approach can perform well when the traffic patterns are stable over a long period of time and have periodical patterns, which means these techniques are applicable to small cities like Wollongong. The ANN techniques can deal with the more complicated scenarios where traffic patterns are varied and fluctuate significantly from time to time, however these techniques require large databases for accurate prediction. The hybrid approaches can take advantage of the individual models to have better prediction but, understandably, they are more complex and more difficult to implement.

2.6 Mobile Applications for Transport Information

2.6.1 Location Sensing Technologies

Location sensing technologies are also being referred as location tracking, location-aware, or positioning technologies in literature. The key location sensing technologies

used in mobile phones include the GPS technologies (covered in the previous section); Assisted GPS; Wi-Fi positioning; and cellular network positioning.

Assisted GPS (A-GPS) technology incorporates GPS with the cellular network in order to aid GPS receivers with the cellular network when required (Feng & Law, 2002). For example, if the GPS signals received are weak, the cellular network can provide necessary data, such as satellite health, atmospheric and satellite clock error coefficients, timing information, etc. The A-GPS technology has increased signal sensitivity and allows mobile devices to calculate location even without clear view of the sky (LaMance, DeSalas, & Jarvinen, 2002).

In contrast to the GPS and A-GPS that are capable of identifying the mobile phone in open areas, Wi-Fi positioning can be applied in a limited coverage range, such as inside a building or in a confined spatial area. Wi-Fi positioning uses the existing Wi-Fi infrastructure in an environment. An example is RADAR, an indoor tracking system developed and based on the IEEE802.11 wireless network technology (Hightower & Borriello, 2001), which can compute the indoor 2D positions by measuring mobile phone signal strength and signal-to-noise ratio. Research has reported that Wi-Fi positioning can attain meter-level positioning accuracy (Y.-C. Chen, Chiang, Chu, Huang, & Tsui, 2005). Wi-Fi is also cost effective, practical and accurate (N. Chang, Rashidzadeh, & Ahmadi, 2010; Y.-C. Chen et al., 2005; Lassabe, Canalda, Chatonnay, & Spies, 2009; B. Li, Salter, Dempster, & Rizos, 2006), although requiring device calibration (Y.-C. Chen et al., 2005; Vaupel, Seitz, Kiefer, Haimerl, & Thielecke, 2010).

Unlike GPS, Cellular network positioning is unaffected by the prerequisite of clear view of sky with cellular tower coverage, and it can also operate indoors (D’Roza & Bilchev, 2003). The popular techniques in this category include Cell Identification (Cell-ID),

Enhanced Cell-ID, Time of Arrival (TOA), Observed Time Difference (OTD), Enhanced-OTD and Time Difference of Arrival (Zeimpekis, Giaglis, & Lekakos, 2002). For example, Cell-ID identifies the approximate mobile phone position by identifying the cell site the mobile phone is using at a given time. The accuracy of cellular network positioning is impacted by the cell size. Densely covered areas such as in urban areas usually have higher accuracy (Giaglis & Pateli, 2002).

Smartphones have been increasingly integrating different location sensing technologies (Y. Chen, Lymberopoulos, Liu, & Priyantha, 2012; Ferster & Coops, 2013; Glasgow et al., 2014; Kim, Mankoff, & Paulos, 2013). The 3G iPhone was the first consumer device to have integrated three positioning technologies: A-GPS, Wi-Fi positioning and cellular network positioning (Zandbergen, 2009). Today, most smartphones are also embedded with low-cost navigation sensors, accelerometer, digital compass, and gyroscope, and have signal opportunity such as Radio Frequency (RF), Bluetooth and Radio Access Network (Y. Chen et al., 2012). Research has shown that locations identified using the 3G iPhone, although less accurate than those from regular autonomous GPS units, are sufficient for most Location Based Services (LBS) (Zandbergen 2009), and has the relative accuracy and reliability for the needs of community-based mapping approaches (Garnett & Stewart 2014).

2.6.2 Location-based Mobile Apps

Location based services (LBS) refer to services that are based on user's current geographic location (D'Roza & Bilchev, 2003; Korhonen, 2003; Mohapatra & Suma, 2005). The inclusion of location sensing technology (Section 2.2.1) within mobile phones has allowed the development of location based mobile apps to provide the LBS.

Locations-based mobile apps are widely used both by end consumers (e.g. apps for assisted navigation, mobile e-shopping) and businesses (e.g. apps for goods tracking, mobile customer support) (Giaglis & Pateli, 2002). Although there are variations in definition and categorisation, overall, when apps offer location information to not only the users who use the apps but also others, the apps are categorised as push or location tracking applications; conversely, if apps only offer personal location information to the requested app users, they are categorised as pull or location-aware apps (Annur & Gretzel, 2011). Location-based apps can be considered context-aware in that the user's geographic position and surroundings are the main context; the location information can trigger related information that is adjusted based on this context.

Not to be confined to location tracking or location aware functions on the other hand, mobile phone users who access mobile apps will leave digital footprints. Exploring these digital footprints can reveal great understanding of human behaviour and movement (Girardin & Blat, 2008). This has been another research interest in the further development of location based apps.

2.6.3 Real-time Public Transport Information Display

The provision of the next trains and bus arrival or departure time in a real time manner is a critical component to improve traveller information and service quality (Dziekan & Kottenhoff, 2007) and displaying such information has been proven to be appreciated by passengers (BMBF, 2002; Coogan, 2003; Dziekan & Kottenhoff, 2007; Infopolis, 1998; Lehtonen & Kulmala, 2002).

The display of the real time public transport travel information has been found to be used more often than other fixed timetables (Dziekan, K and Sedin, 2005; Science Applications International Corperation, 2003). Between 70% to 100% of people who

look at the real time display at a stop (K. Dziekan & Vermeulen, 2004; Infopolis, 1998; Schweiger, 2003; TriMet, 2002). Among other factors, Dziekan and Kottenhoff, (2007) explored the benefits of providing real time public transport information. Its proposed benefits can be seen in the mind map below (Figure 2-4). Overall it has been found that the provision of the real time travel information at stop can have a comprehensive impact on passengers. For example, it can reduce passenger's perceived wait time, created positive psychological effects, such as reduced uncertainty, increased easiness of use and increased feeling of security; allow passengers to adjust their travel behaviour, such as better use of their waiting time, arrange better travel plan, etc.

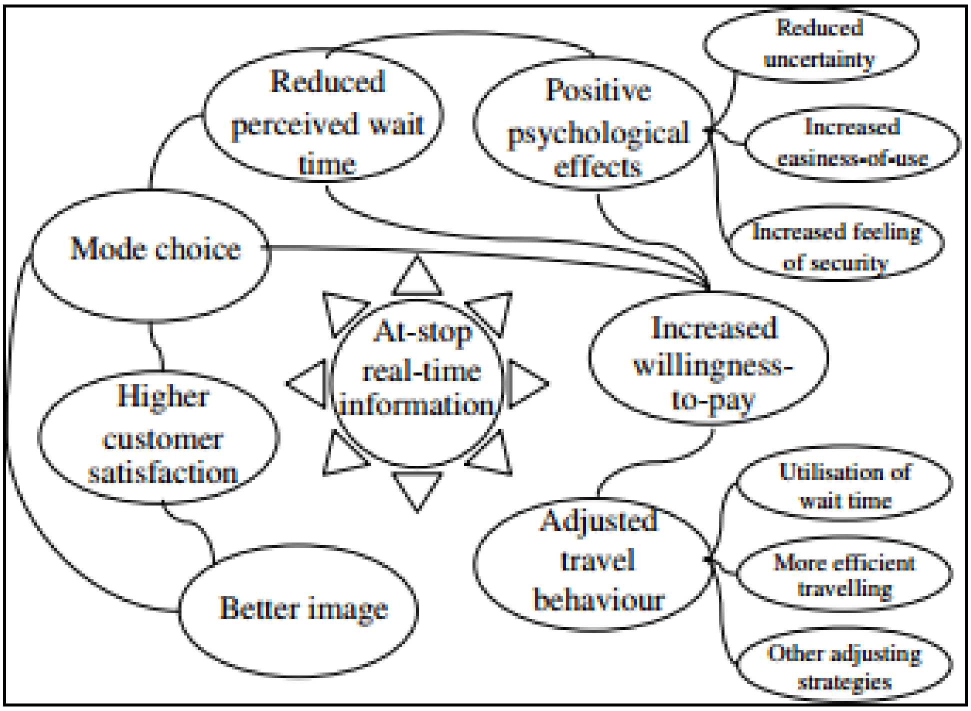


Figure 2-4: Mind map of possible effect of displaying at-stop real travel time information
(Copied from Dziekan and Kottenhoff, 2007)

Among the different means of presenting the real time travel information to passengers, the countdown displays, which shows route number, destination and time remaining until arrival or departure are state of the art (Dziekan and Kottenhoff, 2007, Infopolis2

1998). The evaluation of countdown function has been particularly evaluated by Schweiger (2003), he identified the following results in his survey study. Ninety percent of the riders look at the countdown signs at least once during their waiting time at the stop. Passengers experienced less stress while waiting for the bus when Countdown was present. Of passengers interviewed, 65% felt that they waited for a shorter period of time when countdown was present, with the perceived waiting time dropping from 11.9 to 8.6 minutes.

2.7 Privacy Concerns

The extensive use of location based technologies and apps may threaten users' location privacy and poses potential risk of privacy information abuse (Agre, 1995; Gruteser & Grunwald, 2003; Karger & Frankel, 1995; Want, Hopper, Falcão, & Gibbons, 1992). People do not always understand why privacy information is being collected (Beckwith, 2003). Studies found people prefer position-awareness systems (i.e., where a mobile phone determines location upon request and only shares with the user requesting it) to location-tracking systems (i.e. that expose users' information to others besides the users themselves) (Barkhuus & Dey, 2003). A study revealed that consumers oppose location tracking apps if they do not perceive the benefits as sufficiently compelling to overcome privacy concerns (Klasnja, Consolvo, Choudhury, Beckwith, & Hightower, 2009).

Research has also shown location information context and its use are major privacy concerns for information disclosure (Barkhuus et al., 2008). Consolvo et al. (2005) pointed out that a person's willingness to share location can be determined by four considerations: who to share with; why to share with; how the data is to be used, and ultimately users' willingness to share. While it has been found that younger people are more willing to share their personal information than older ones, it has been identified

that common privacy principles (such as user consent, purpose binding, and adequate date protection) when implemented properly can help reduce privacy breaches (Langheinrich, 2001).

2.8 Summary

Although the ITSs have been in development in the past four decades, the literature review has revealed that using the digital ecosystem principles to develop an ITS is a new approach. Identifying this gap, the PhD project was motivated to apply the digital ecosystem principles to ITS, and to develop a digital transport ecosystem that can provide safer and better transport services in a manner that is self-organising, scalable and sustainable.

Traveller Information System as a key application category in ITS, commonly utilises two key technologies – the AVL technologies and data prediction methods. Due to the advancement in the modern GPS technology, GPS has been chosen as the primary AVL in this research. Literature has shown that travel time information is a vital component of many ITS applications, especially in the public transportation systems. With many prediction methods recorded in literature, it has also shown that every prediction method has its advantages and disadvantages. In general the published research shows that the historical approach can perform well when the traffic patterns are stable over a long period of time and have periodic patterns. These insights have shown there is need for research to develop a historical data based method combined with the real time bus information for the arrival time prediction.

Regarding designing the public transport apps, the literature provided valuable direction by suggesting that there is considerable advantage in displaying bus arrival time in a real time manner, especially in a countdown manner, as a critical component for

improving the public transport service quality. Hence such functions have been incorporated in the app design in this research.

In choosing the smartphone platform, the literature showed the 3G iPhones have accurate location technologies and reliability (at the time of this research commencement). This finding led to the identification of 3G iPhones as a sufficient platform for most LBS and hence they been chosen as the end user device environment.

CHAPTER 3. UNISHUTTLE SYSTEM DESIGN AND IMPLEMENTATION

3.1 Background

The Gong Shuttle is a free public transport service for the City of Wollongong. Its route includes the UOW campuses, the city centre, city beach, and some other key attractions or facilities in the city. The Shuttle provides services to all the citizens, and is of particular use to UOW staff and students. With over 20,000 students enrolled for study in the campus and over 2000 staff, the parking situation at UOW campuses is under constant pressure. If driving to campus during the daytime, it is often difficult get a parking spot especially during the student study sessions. Hence, utilising the public transport travel to university can be the only possible option for most of students.

Gong Shuttles are partially scheduled services, ad hoc changes often occur during student exam periods and sessional breaks. In order not to miss buses or waste time at the bus stops, passengers need to know the real time shuttle bus timetable, and on the other hand, the university and the city transport authority also need to know when and how to arrange the shuttle bus in order to serve passengers better.

The Gong Shuttle buses are ticket-less, which means the traditional ticket counting based reporting is irrelevant and a new way to collect passenger statistics is required.

Three bus route services were chosen as the test bed in this research. These routes are highly patronized by the UOW students and staff, as well as the local residents. Figure 3-1 shows two of the three routes in blue and orange lines, the third route is the express version of the blue line, where it travels between the North Wollongong Train Station and UOW Northfields Ave.

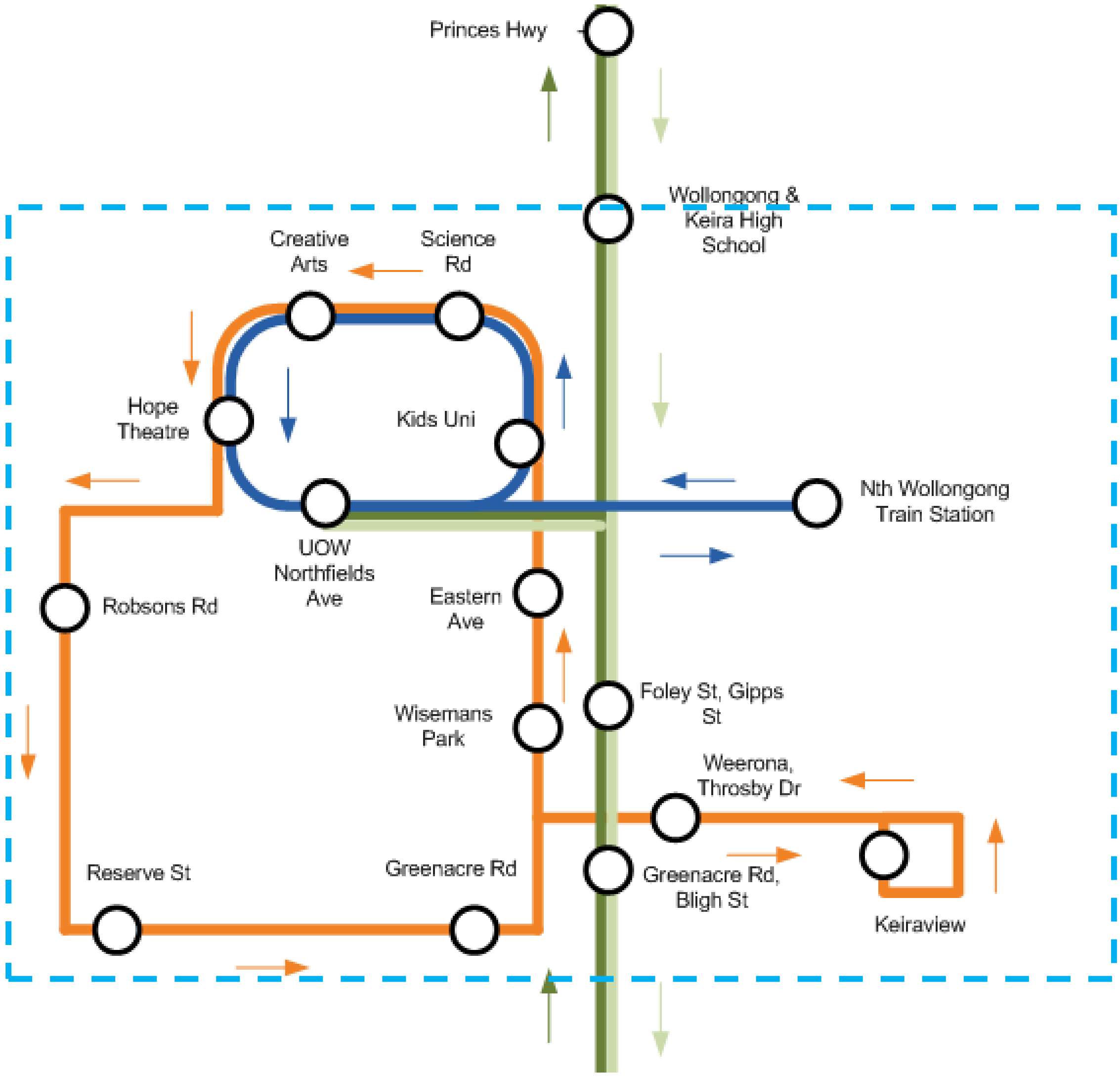


Figure 3-1: The free Shuttle Bus Network in the City of Wollongong

The UniShuttle system's interdependent digital components - shuttle buses equipped with on-board devices, central server which takes centralized management and control, passengers carrying mobile devices, and transport authority - interact with each other, where passengers' mobile devices enquire about shuttle bus route and travel time information in real time, and the central server computes information with the shuttle bus location information input from the shuttles and users' participation. The transport authority can better manage shuttle bus services with the information output from the central server. Together the system is self-evolving and self-sustaining. In the

Connected Mobility Digital Ecosystem context, the UniShuttle system is a transportation digital ecosystem application.

3.2 UniShuttle's System Architecture and Design

3.2.1 Architecture Overview

The UniShuttle system architecture is based on the IETF presence model (Day, Rosenburg, & Sugano, 2000). The architecture integrates the network of AVL sensors on shuttle buses to the central server and the central server to the end users with mobile devices. The presence model consists of three roles: *presentity* (short for presence entity), *presence service* and *watcher*. The *presentity* is the entity of interest to the *presence service* which is of interest to the potential *watchers* who seek status or other information about the *presentity* from the *presence service*. The simplified UniShuttle architecture is depicted in Figure 3-2.

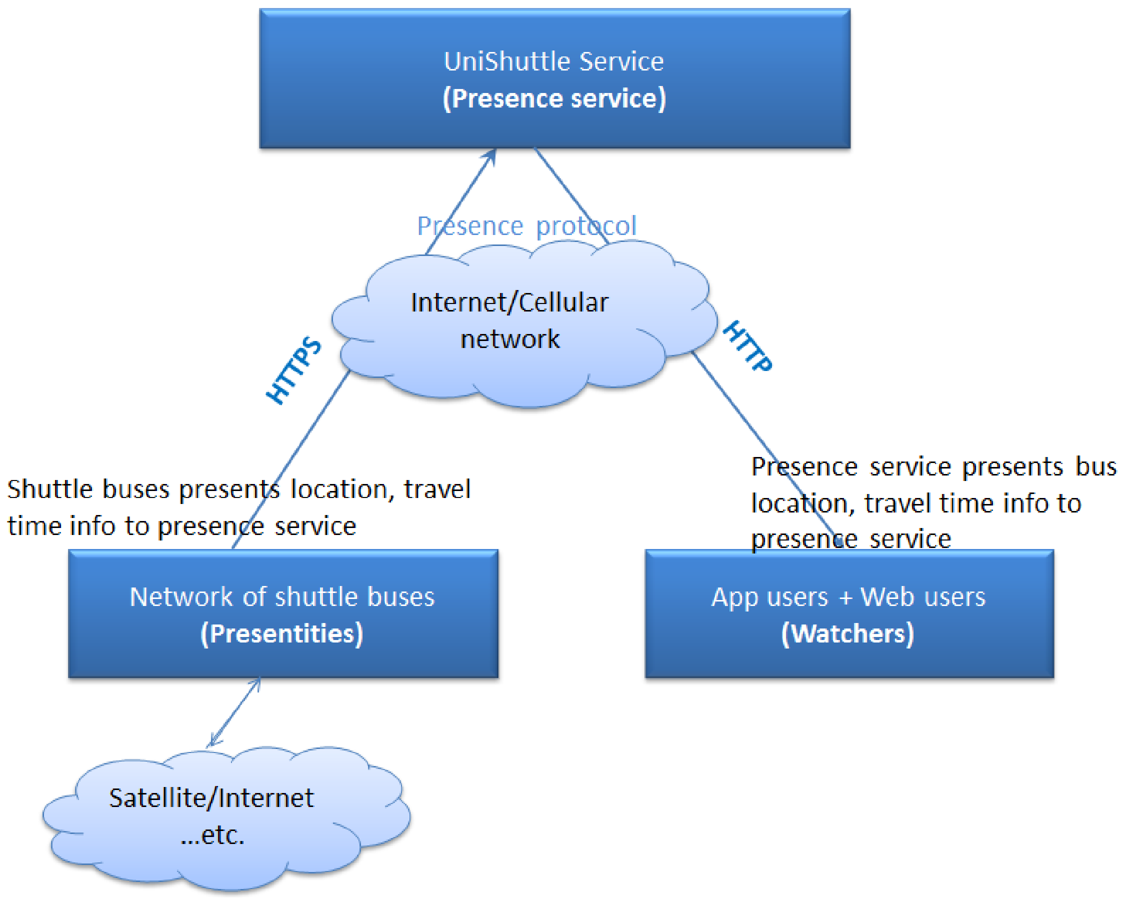


Figure 3-2: UniShuttle system's presence-based architecture overview

In the UniShuttle system, the shuttle buses gain location related information from their AVL sensors installed on-board. The Shuttle buses send the location related information to the central server. The Shuttle buses play the role of the *presentity*. The individual users use the UniShuttle app/website to seek the bus travel information, they play the role of *watchers*. The UniShuttle central service mediates the interaction between the *presentities* and the *watchers* and ensures that presence information is distributed to the users.

Under this architecture, the three components – shuttle buses, central server and the end users are integrated into one digital ecosystem. The communication protocols between the server and users are through HTTP, and between the server and the shuttle busses

are through HTTPS. The system uses a standard XML-based message for information representation and exchange.

Comparing to the collective traffic information architecture proposed by W. Lee et al. (2010) (Figure 2-2), this architecture allows a comprehensive central server which include functions such as data storage, historical data warehousing, bus travel time prediction, communication with the smartphone devices on the user end, and the overall control of the system; furthermore, this architecture not only allows the request-response communication, but also allows the integration of the passive system components (e.g. the GPS-equipped shuttle buses).

3.2.2 Architecture Design

The detailed explanation of the UniShuttle system architecture is depicted in Figure 3-3, where the three components and their interactions are outlined. The numbers on the figure indicate how information is accepted, distributed and used.

Inside the component of the shuttle bus, a GPS device (1) gathers and sends the location data to the Location Information Server (LIS). As a redundancy mechanism, in case the GPS signal fails and bus location data is unavailable, a Specific Dedicated Short-range Communications (DSRC) transmitter can be installed at some bus stops. The DSRC reports shuttle bus location to the DSRC reader installed on the shuttle bus. The location data identified is also reported to the LIS. The LIS ensures that the location data that correspond to bus movements are distributed to the system central server.

Inside the component of the shuttle buses, another key aspect is shuttle bus's mobile hotspot function. The hotspot wireless router function allows passengers' mobile device access to the Internet through Wi-Fi connection. The system maintains the list of

connected devices through the bus router logs. Through these logs, an indication of shuttle bus usage can be gained by examining the number of switched-on passengers' mobile devices in the logs (3).

Inside the component of the central service, the bus presence server receives the location data sent from the shuttle buses, and sends them to the web server, a Short Messages service (SMS) gateway and a database server. The database server (5) collects vehicle travel and passenger usage statistics, and these data can be mined to extract trends. Over time, bus scheduling can be optimised from repeatable patterns of travel and passenger data. This enables the shuttle bus network to be self-evolving. The web server hosts the UniShuttle web presentation, where users can refer to dynamic schedules of buses.

The component of the *watchers*, or the users, includes the UniShuttle users using a smartphone (6) or on the website or the usage of the available LCD screen at some bus stops. An SMS service (7) also allows users without smartphones to send SMS with the stop ID and receive the time of the next bus.

The communication between the shuttle and the central server use the 3G mobile networks and through 802.11e wireless networks when Wi-Fi is available (4).

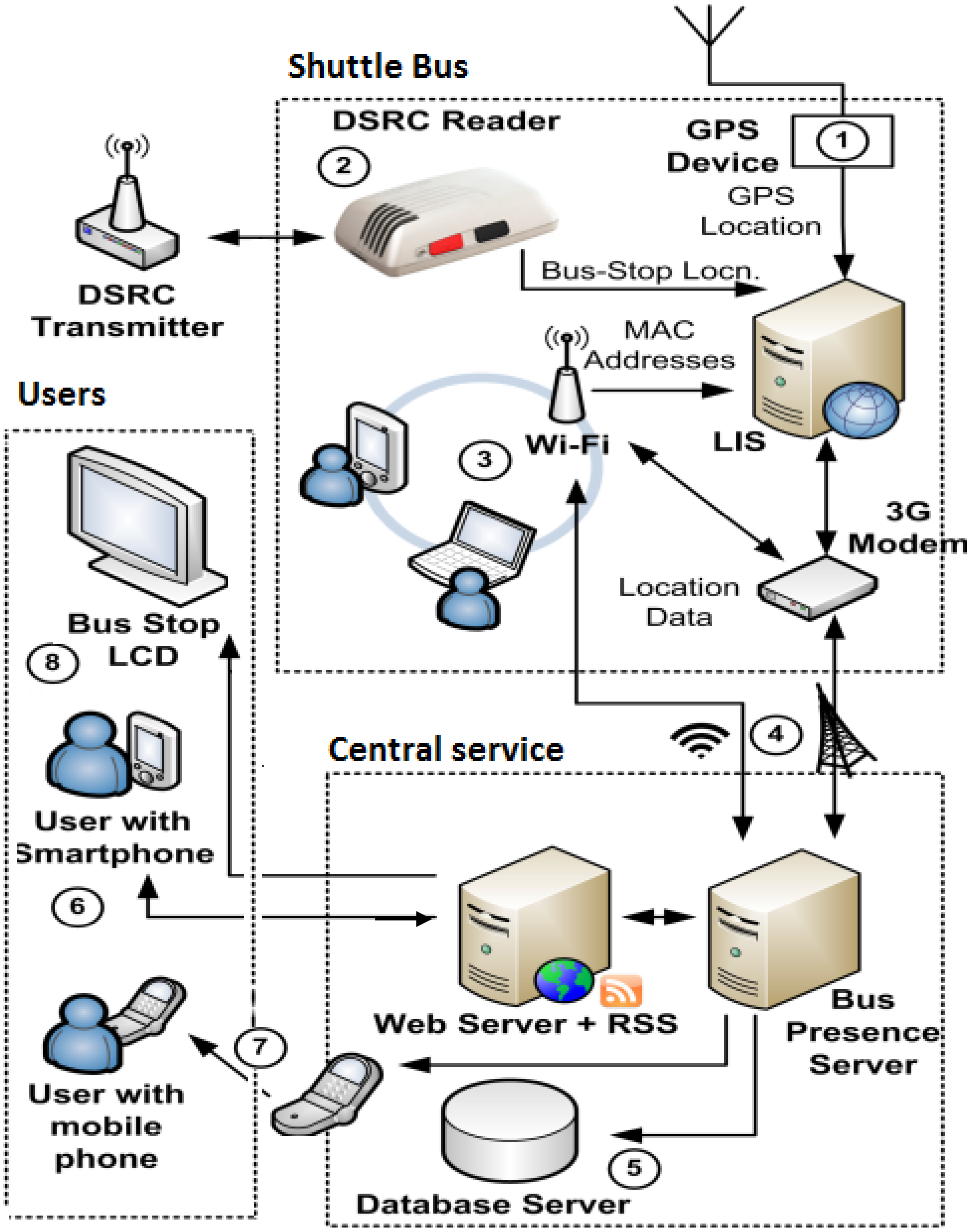


Figure 3-3: UniShuttle system's architecture design

3.2.3 Software Conceptual Design

The UniShuttle system is one of the earliest projects implemented for the CMDE concept. One vision of its system design is to lay a foundation that the future CMDE projects can build upon. In this light, its software design follows a Service Oriented Architecture (SOA) principle. Within the loosely-coupled SOA framework, business functions are linked building blocks of a system, and are defined as modules or services.

The SOA framework supports the services or modules' integration. (In this thesis, module and service are interchangeable.) The UniShuttle software's design framework is depicted in Figure 3-4 below.

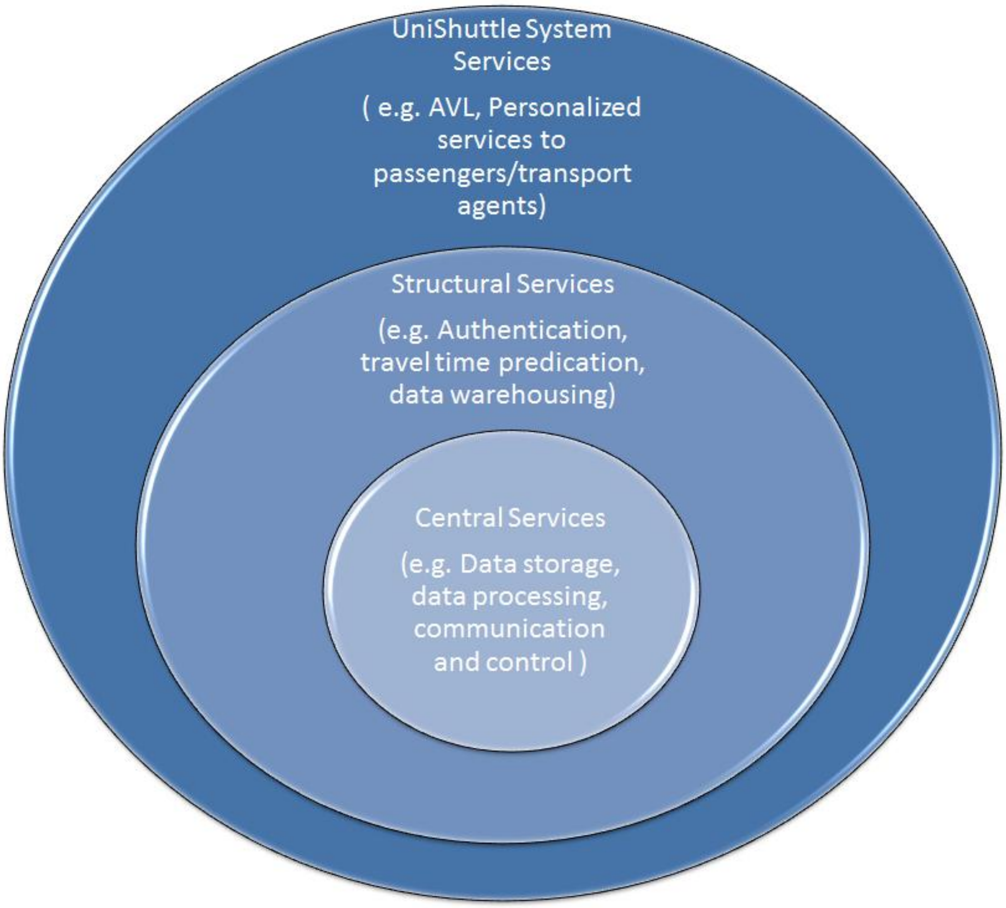


Figure 3-4: UniShuttle software conceptual design

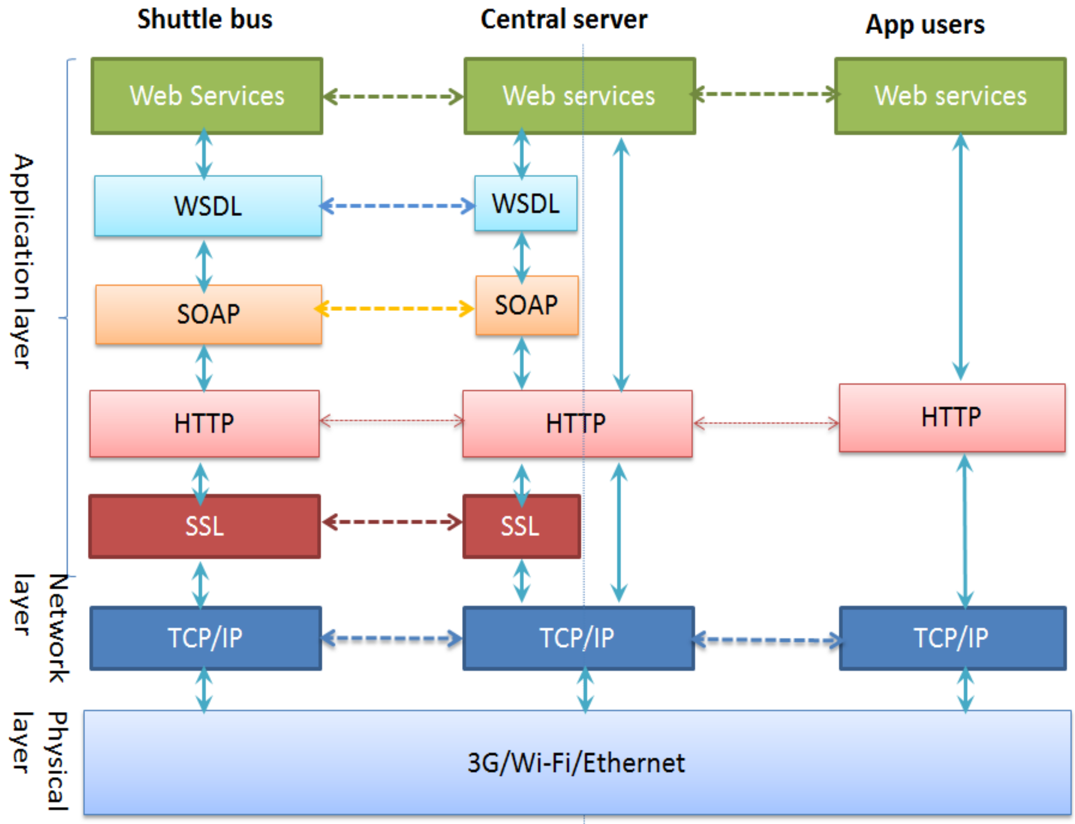
Based on the digital ecosystem concept proposed by the European Commission (Figure 2-1), the framework classifies the system software modules into three categories: central services, structural services, the UniShuttle specific services. At the centre of the framework are the central services that are responsible for the central system functions, such as data storage and processing, as well as communication and control of the system. It is envisioned that this group of services will act as the common infrastructure among the CMDE projects. The middle group of services consist of the structural

services, which include structural functions such as authentication, data security and data warehousing. The outer group refers to the specific UniShuttle services, which includes functions such as computing the shuttle bus travel time, displaying geographic map and presenting statistic information to transport authorities.

3.2.4 Communication and Messaging Design

The overview of the UniShuttle communication and messaging design is depicted in Figure 3-5. The shuttle bus packages its message into the Simple Object Access Protocol (SOAP) format based on the Web Service Description Language (WSDL) published by the central server. The message passes through the HTTP over Secure Socket Layer (SSL) and the TCP/IP layer and reaches the central server. The message reached on the central server goes through Ethernet, network layer (TCP/IP) and reaches SSL and HTTP handling on the server's application layer. The resulting SOAP message is processed through the WSDL, and the requested web service is called. The returned message from the central server passes through the same sequence in reverse as the message sent over from the shuttle bus.

When an app user requests a bus arrival time, the app sends a HTTP request through the network layer (TCP/IP) and the physical layer (wireless network) to reach the central server; the server returns the requested information back through the same path the request has travelled.



- A web service is a software module that provides a service and is accessible over the Internet.
- The interface to web services is defined in the XML-based Web Service Description Language (WSDL), which provides all the required information for an application to access the specified service. XML is a mark-up language that enables cross-platform data interchange via a standard format which encodes both the data structure and data definition.
- Simple Object Access Protocol (SOAP) is a network application protocol for the exchange of information between service instances in a distributed environment. It also enables cross-platform integration of any programming language and distributed object infrastructure.
- HTTP is a request-response protocol, which specifies that a client sends a request to server, and server return the response message to the client.
- Secure Sockets Layer (SSL) is a standard security technology for establishing an encrypted link between a server and a client.
- TCP/IP (Transmission Control Protocol/Internet Protocol) is the basic communication protocol of the Internet, which manages the assembling and packaging of messages and handles the message address so that they can get to the right destination.
- The physical network layer consists of the basic network hardware transmission technologies.

Figure 3-5: UniShuttle System communication and messaging network overview

3.3 UniShuttle's System Implementation

Figure 3-6 shows the UniShuttle system's schematic implementation in four parts: on-board device installed on shuttle buses, central server, frontend application and business intelligence generated from the data warehouse over time.

The on-board device installed on the shuttle buses consists of AVL technology, wireless router and LIS. It periodically transmits the shuttle buses' location information to the central server.

The central server is implemented with four sub-servers: a bus presence server, a database server, a data warehouse server, and a web service server. The bus presence server is responsible for receiving and processing data from the on-board devices, computing the shuttle bus travel time prediction, and saving the information into the database server. The database server stores the real time travel information. For business intelligence reporting purpose, a separated data warehouse server is responsible for data cleansing, loading and transformation. The web server provides web services to end user and displays geographic data in Google map format to passengers.

The frontend of the system is an iPhone application and a website. They are implemented to present personalized passenger travel information to users.

The aggregated data stored in the data warehouse can be mined to present the patterns of shuttle bus travel and the passenger usage of the shuttle services. This business intelligence function can be used to help the transport authority for better resource planning.

The detailed explanations on these four parts are covered in the sections below.

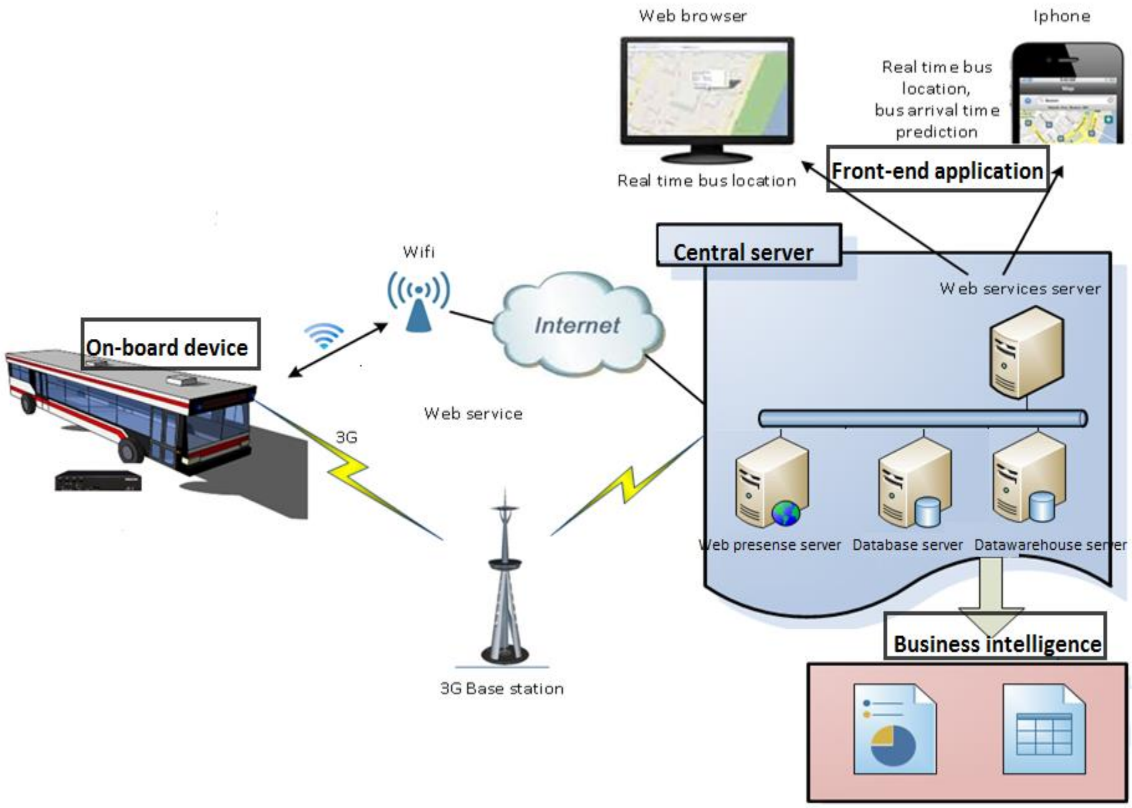


Figure 3-6: A Schematic view of UniShuttle system's implementation

3.4 The On-board Device

3.4.1 The Hardware Consideration

The primary function of the in-vehicle on-board device is to detect and transmit shuttle buses' location, speed and direction to the central server. As UniShuttle services are ticketless, the on-board device's second function is to collect the passengers' information, including counting the on-board passengers which is a very useful function for understanding the transport service usage. To implement this, the on-board device acts as a Wi-Fi router to provide Internet access to the passengers on-board. When passengers access the Internet through the on-board device, the system collects the passengers' digital footprint – their MAC (Media Access Control) addresses and time of boarding.

To track the shuttle bus location, the on-board device needs to have a robust and accurate AVL technology. Secondly, the shuttle bus needs to be able to send the location data to the central server periodically. This means the device should have communication ability that utilizes a reliable telecommunication network. Thirdly, the device should have Wi-Fi function to allow it to act as a wireless router to the passengers on-board. The device should also have the ability to store information locally for cross checking purposes. For maintenance purpose, the operation of the device should be stable enough with minimal intervention from the operator. For deployment purposes, the integration and management of the device should be relatively easy and costs kept relatively low. Other considerations also include operation temperature range, power ignition delay, low-power protection, and vibration protection that are directly related to transport applications.

Based on the considerations above, rather than developing a brand new piece of hardware for the on-board device, an in-vehicle PC as the device hardware is the most appropriate option. With similar CPU power and memory storage possessed by the potential choices, comparisons were focused on the key features (e.g. GPS, Wi-Fi). The table below (Table 3-1) shows some of comparisons made on four in-vehicle computers from different manufacturers.

Table 3-1 Comparison of Product spec from several manufacturers

In-Vehicle PC Key features to be considered		Possible in-vehicle PC platform			
		Newaye	**VTC2100 NEXCOM**	**eBOX310-830-FL Axiomtec**	**TREK 550 Advantech**
WiFi Function -enable wireless access inside the bus, and sending location information sent back to server		Not included	Included Internal wireless communication(3.5G, GSM/GPRS, WLAN, BT)	Included	Included Built-in wireless communications(WWAN, WLAN, BT)
3G/cellular –enable sending information back to server		Not included	Included	Not included	Included
GPS-Automatic Vehicle Location		Not included	Integrated GPS with dead reckoning	Not included	Included GPS with AGPS and dead reckoning technology
Operating Voltage		0-12V	8 to 60V (wider range) Selectable boot-up & shut-down voltage for low power protection; design ready for 8-level delay time on/off at self-configuration; Power on/off ignition, software detectable	Two modes: 9-16 or 18-32V	0-36V Power protection and car power management software (Ignition on/off, delay on/off, low battery monitor) prevent electrical noise and surges from impacting
Operation Roughness	**Operating temp (C)**	-30 to 60	-30 to 50 storage temperature -40-80	-10 to 50 storage temperature -20-80	-30 to 70
	Vibration	Not available	MIL-STD-810F, Method 514.5, Category 20, Ground Vehicle	Not available	MIL-STD-810G/ 202A, Method 516.5, EN60721-3 (5M3) compliant
Price (AU$)		$500	$1200	$1300	$2000

VTC2100 was chosen as the on-board device. Figure 3-7 shows the front and back view of the VTC2100 system. The main reasons for choosing VTC2100 include: it provides the required system features; its price is more affordable than others in the category, especially to the next level-TREK 550; it has local contact that can support repair, maintenance and technical support. GPS signals can be lost due to the canyon effect in busy city areas with high-rises. To overcome the loss of GPS signals, the dead-reckoning technique embedded in the on-board devices will be able to compensate the GPS signal loss.

Figure 3-7: View of VTC2100 (a) Front view and (b) Rear view

3.4.2 The Software Modules

The Linux operating system with the MySQL database was chosen as the software environment, and Python was chosen as the main programming language.

The software modules are developed for the on-board device. The main modules and processes are: to capture the bus location and travel information; to capture the on-board passenger information; to upload information to central server through wireless network; to monitor data capture and transmission.

Figure 3-8 shows the overview of the on-board device software modules and their required interaction with the other parts of the system.

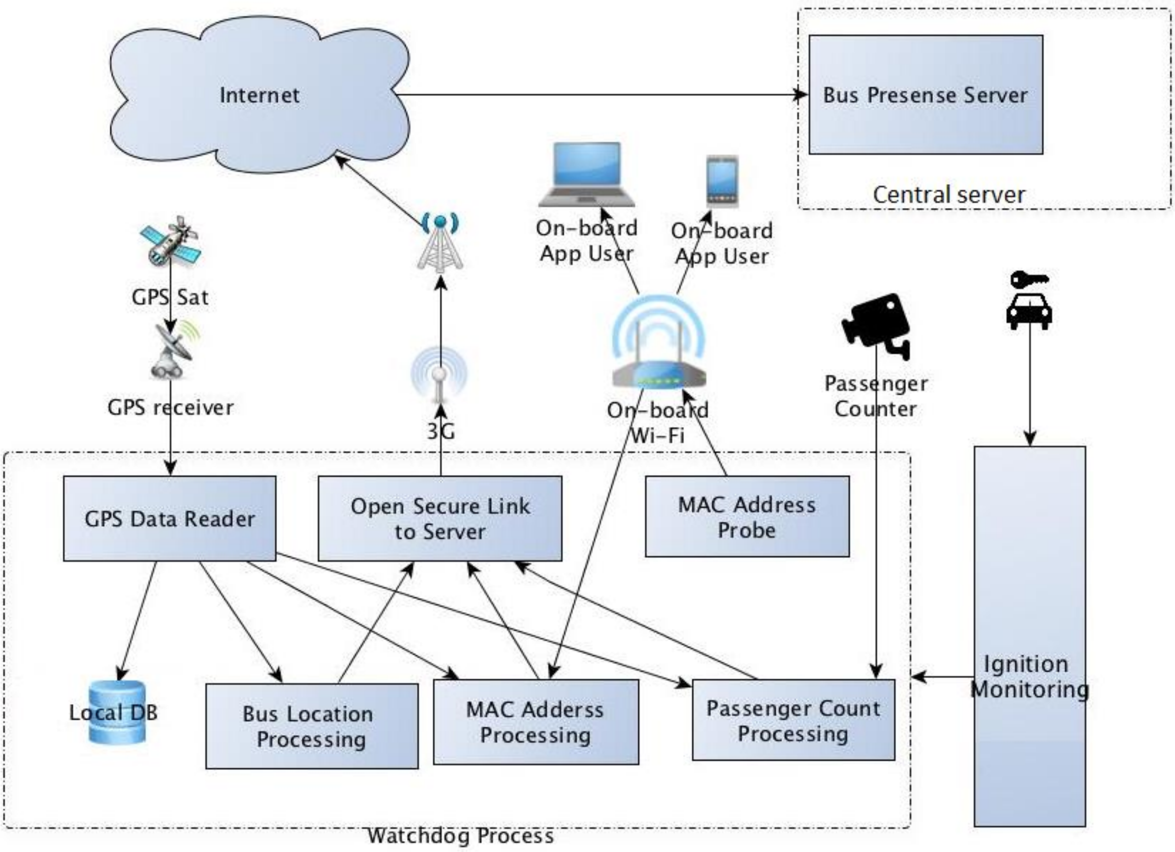


Figure 3-8: An Overview of the on-board device software modules and the interaction with the other parts of the UniShuttle system

From the on-board device's GPS receiver, the *GPS Data Reader* module extracts the latitude, longitude, speed, direction and timestamp of an in-service shuttle bus. The extracted geodetic coordinates are critical input for the further information process in the system. The geodetic information is input into the other modules (as illustrated in Figure 3-8). Simultaneously, it is also saved into the device's local DB, which can be used for backup and debugging purpose when issues arise. The *Bus Location Processing* module is responsible for packaging the geodetic coordinates into a message format that the *Bus Location Server* (located at the central server) can understand and

accept. The message format has to adhere to the XML schema in the WSDL published by the central server. The following example shows part of WSDL used for location packaging format in the system. Six elements - bus_id, longitude, latitude, time, speed, and direction - are the defined fields to communicate with the central server.

```
<xs:element name=”insertBusLocation”>
    <xs:complexType>
        <xs:sequence>
            <xs:element minOccurs="0" name="bus_id" nillable="true" type="xs:string"/>
            <xs:element minOccurs="0" name="longitude" nillable="true" type="xs:string"/
            >
            <xs:element minOccurs="0" name="latitude" nillable="true" type="xs:string"/>
            <xs:element minOccurs="0" name="time" nillable="true" type="xs:string"/>
            <xs:element minOccurs="0" name="speed" nillable="true" type="xs:string"/>
            <xs:element minOccurs="0" name="direction" nillable="true" type="xs:string"/>
        </xs:sequence>
    </xs:complexType>
</xs:element>
```

For the shuttle buses that have a passenger counter device, the *Passenger Count module* reads out the number of passengers on board. The *MAC Address Processing* module can also provide an approximate indication of the number of passengers on board.

The MAC Address Processing module is responsible for examining the mobile devices' MAC address through examining the ARP (Address Resolution Protocol) table. ARP is a known protocol for mapping IP addresses to MAC addresses. The ARP table has the records of the mobile devices' MAC addresses while these mobile devices are connected with the on-board gateway (Wi-Fi) function. When a passenger mobile device is newly connected, its MAC address appears in the ARP table. This MAC address, location and time of accessing are then packaged to be sent to the central server. Conversely, if a passenger mobile device is disconnected, the corresponding MAC address disappears from the ARP table. The entry that has the corresponding MAC, location and time of exiting will also be sent to the central server. By recording

the MAC addresses of the passenger mobile devices, the *MAC Address Processing* module also fulfils its second function, which is to keep a record of passengers' digital footprint (e.g. MAC address, time of connection, the travel and route information). When more digital footprints are recorded, the footprint can be mined to shed insights on passengers' behavioural patterns. The *MAC Address Probe* module ensures the disconnected MAC addresses are removed from the ARP table in the system.

There are also two process control modules for the on-board device. The *Watchdog Process* module monitors the operating status of the on-board device. If the modules mentioned above stop responding, the Watchdog restarts the unresponsive modules respectively and also sends the logging information to the central server. This logging information is for the system fault analysis. The *Ignition Monitoring* module monitors the shuttle bus ignition status. When the bus engine is off for 60s, the module will shut down the on-board device to conserve the bus's battery power.

The *Open Secure Link to Server* module is the communication gateway to the central server. It secures the messages between the three modules on the on-board device (*Bus Location Processing* module*, MAC Address Processing* module*, and the Passenger Count Processing* module) and the central server. Section 3.4.3 explains this module in detail.

3.4.3 Communication with the Central Server

The communication between the on-board devices and the central server is based on the secure Internet communication protocol, SSL, where through the verification of the provided digital certificate, two parties are assured of the other's identity. The security is first configured during installation. When the central server and the on-board devices

are installed, a public/private key pair and a self-signed certificate are generated respectively. The server's public key and certificate are imported to each on-board device, and the public key and certificate from the on-board device is also imported to the server. The secured link is then established when on-board devices communicate with the server where the *Open Secure Link* module is responsible for the secure communication. Figure 3-9 shows the overview of the communication process between the server and an on-board device (client) on shuttle bus. In its initial start-up, the client first sends a request to the server for secured resources. The server sends back its certificate. If the server's certificate is verified, the module sends the client's certificate to the server, and the server verifies its credential. Once the verification of the certificates is completed, the secured link is established between the server and the client, and all subsequent messages are encrypted to ensure message integrity. After the secured communication is established, the modules between the on-board device and the central server use SOAP for message communication.

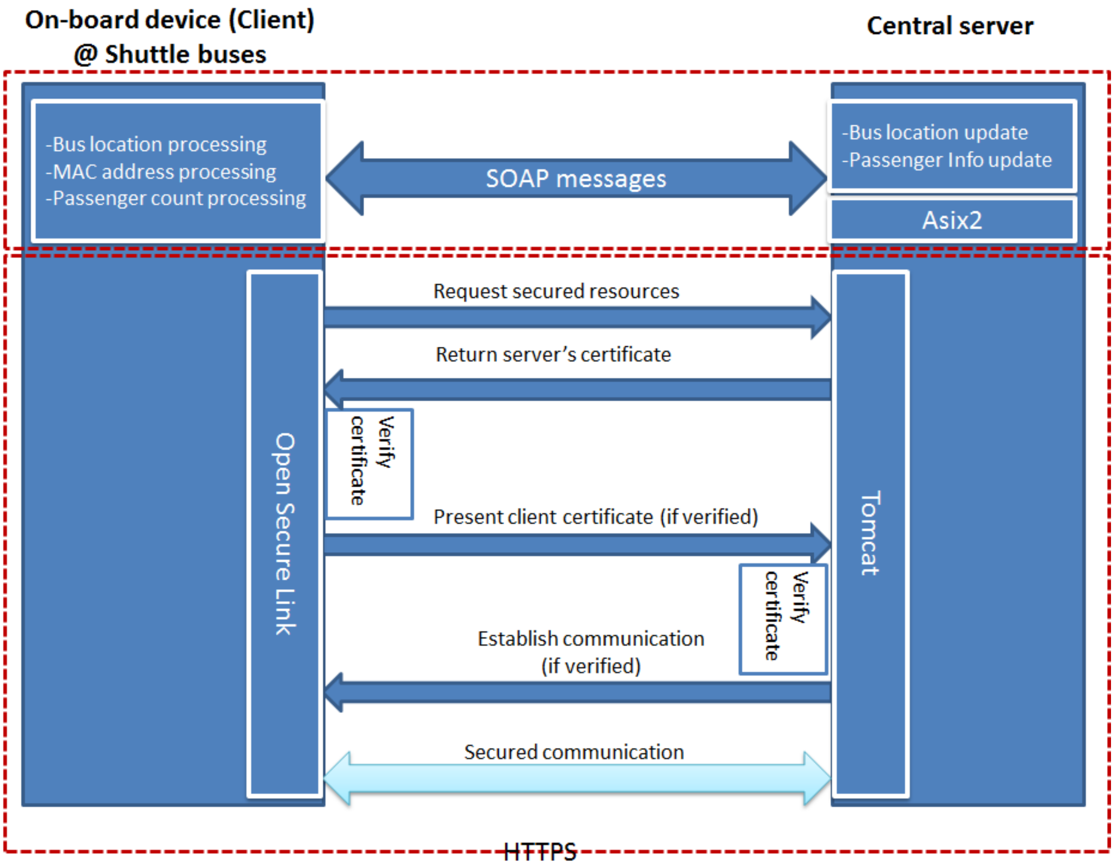


Figure 3-9: An Overview of communication between the on-board device and the central server

How frequently the shuttle buses report their locations to the central server is another important decision to make in the on-board device's communication design. If the communication frequency is too low, a bus stop may be missed to be reported. As the shuttle bus dwelling time at bus stops is a key factor in predicting the bus arrival time, such a situation would be detrimental for the later prediction work; however, if the frequency is too high, large amounts of cellular traffic data will be generated, which can be expensive for transport authority. In order to choose the proper balance, data traffic volume was monitored for the location update frequency of once per second, per five seconds, per 10 seconds, per 15 seconds, and per 30 seconds (Table 3-2). The consequent traffic volumes were then compared. The 10-second frequency gets the best

sampling on the bus route with a reasonable data usage approximately 240M bytes data per month per shuttle bus. The 10 sec reporting time was chosen.

Table 3-2 Data communication between on-board unit and server

Data collection period	Frequency of communication between on-board device and the server				
	1s	5s	10s	15s	30s
15 minutes	4.1MB	900KB	512KB	354KB	180KB
A day	98.4MB	21.6MB	12MB	8.5MB	4.3MB
A month	1.968GB	432MB	240MB	170MB	86MB

3.5 The Central Server

3.5.1 The System Environment

In comparison between Linux, Windows and UNIX operating system, the Linux operating system (server version) is chosen. Despite easy installation, the Windows (server version) needs frequent patch updates, requires reboot and is expensive. The enterprise graded UNIX (server version) requires proprietary hardware to install, and it is also more expensive. Conversely, Linux is open source software, can be installed on most of the hardware and is free. Although the free version lacks the technical support of Windows and UNIX, the system's stability and security is preferable.

The software environment and tools for the system development are described as follows. The widely used web server software – Apache HTTP server, Apache Tomcat (Java-based web application container that can run JSP and Java Servlets for web applications), and Axis2 (Apache extensible Interaction System, a web service/SOAP/WSDL engine) are chosen for running the system's web services. The relational database software - MySQL, is the database management tool using

structured query language (SQL). The programming language, PHP, is used for developing the web services (e.g. display bus route and remaining travel time) and Java is used when programming for the secured web-services (e.g. update bus location from the shuttle bus's message).

For secured communication between the central server and the shuttle buses, Apache Tomcat handles SSL and Axis2 handles SOAP messages. The communication between the central server and the passengers is through HTTP.

The central server is a virtual server within a larger server infrastructure hosted at UOW. It includes four sub-servers: bus location server, web server, database server and data warehouse server.

3.5.2 The Software Modules

The central server stores, controls and communicates with the distributed shuttle buses and the passengers' mobile devices. Its main functions include: recording and processing bus travel information; recording and processing passenger MAC address information; mapping bus route and trip information; predicting bus arrival time. Its key modules are depicted in Figure 3-10. These modules are explained in detail as below.

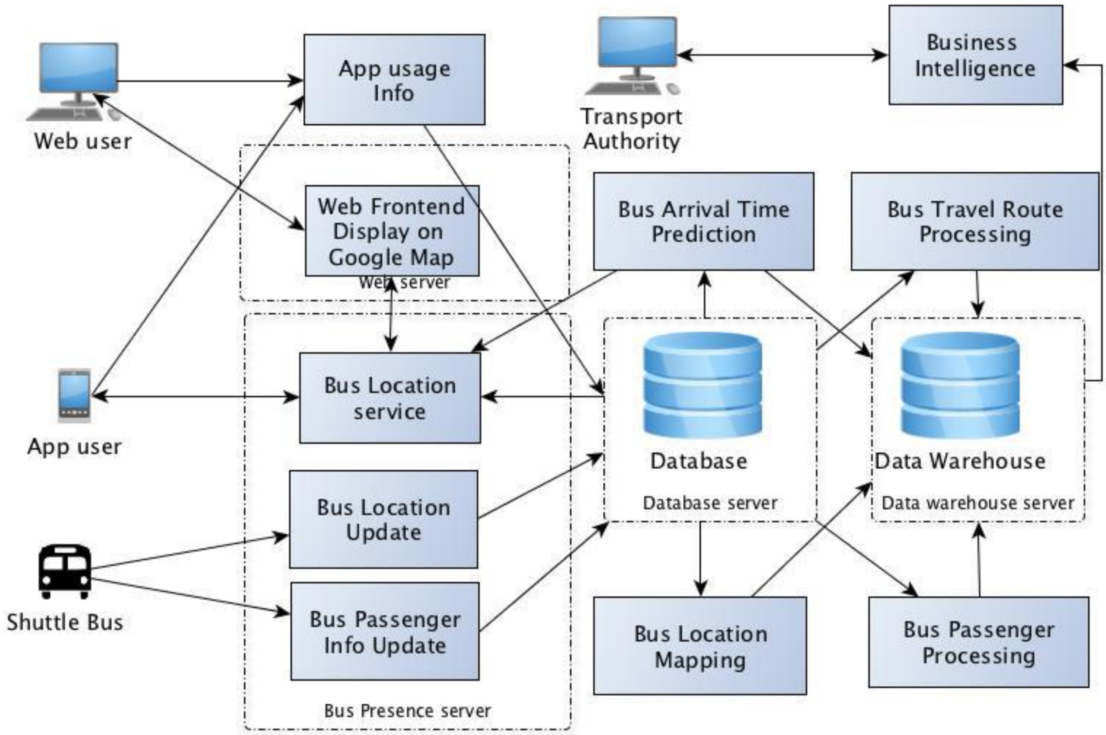


Figure 3-10: An overview of the central server software modules

The *Bus Location Update* and *Bus Passenger Info Update* modules are responsible for accepting the messages sent from the shuttle buses and updating the bus location and passenger information into the database server respectively.

The *Bus Location Service* module responds to passengers' request (requestor) with the bus location and arrival information through Keyhole Markup Language (KML) message format. Upon receiving the passengers' request, the module looks up the database server for the most up-to-date bus location and computes the arrival time. The information will then be packaged into KML format and sent back to the requestor.

The details of travel time prediction are explained in Chapter 4. To explain the trip specific modules - *Bus Location Mapping*, *Bus Passenger Processing*, *Bus Arrival Time Prediction*, *Bus Travel Route Processing*, it is worth mentioning here that bus routes

need to be split into smaller segments where each segment is the longest possible straight line. The segmented routes can be instrumental in estimating bus travel information. For example, adding up the travel time for each segment on a route can provide estimation for the travel time required on that route. The *Bus Location Mapping* module is used to identify the bus location on each specific segment and save the information into the database. To identify a particular route a bus travels on, with the reality that parts of routes often overlap with each other, the *Bus Travel Route Processing* module is responsible for the particular route identification. The *Bus Arrival Time Prediction* module uses the current bus location and speed, and based on the historical bus travel information. It predicts shuttle bus arrival time at a bus stop. The *Bus Passenger Processing* module processes passenger location information and generates an estimate of the time that passengers get on and off the buses at a particular stop.

The *App Usage Info* module is a simple service to gather passengers' information, such as location, time of access, etc., for the passengers' foot print gathering.

The *Web Frontend Display on Google Map* module is responsible for displaying the bus location from the geodetic coordinate format to a Google map format, which is more user friendly in bus travel information overview.

The *Business Intelligence* module is responsible for querying the data warehouse, reporting on aggregated trending information (e.g. shuttle bus usages at different time periods) for the transport authority to gain insights for more efficient service provision. Over time, with more passenger footprint data gathered, business insights can also be

gained on passenger behaviours, which can have wider social and management implications.

3.6 The UniShuttle App for Users

3.6.1 The App Development Consideration

The UniShuttle app is designed to be location aware by utilizing iPhone's own location sensing technologies to obtain users' geographic locations. At the time of the app development, iPhone has advanced location sensing technologies embedded into its hardware, which made it possible to choose the iPhone as the first choice for the app hardware platform as. To use the existing geo-positioning hardware built into the iPhones, the iOS operation system was chosen as the development platform.

The app software design follows the Model-View-Controller design pattern, where the app functions were implemented through the concept of Model, which encapsulates the required behaviours of the app functions and manages data, logic and rules of the app; the View manages the app screen presentation; and the Controller acts as a conduit between the Model and View by sending commands and accepting input.

Programming for the app was written with Object-C programming language to communicate with the bus presence server located at the central server and generate the user interface on the user device.

3.6.2 The App Functions

The UniShuttle app's key functions are designed around the provision of real-time bus travel information. Its four main functions are outlined in Figure 3-11. To illustrate how the app works, some screenshots are included in this section. Its landing page is shown

in Figure 3-12. The three routes shown on the landing pages correspond to the routes chosen in the system platform (routes highlighted in Figure 3-1). The app functions are accessible from this screen.

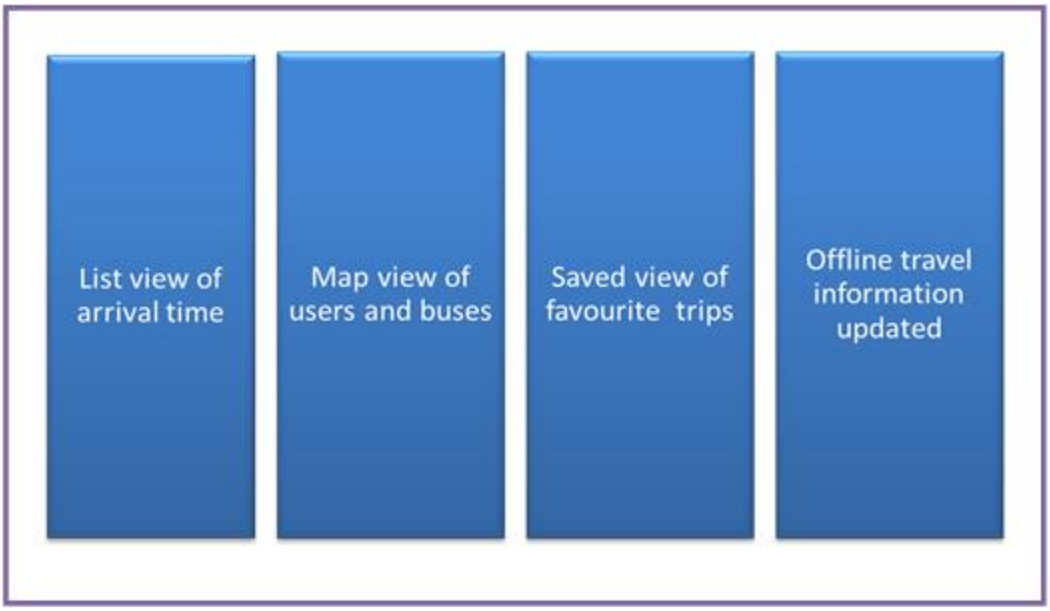


Figure 3-11: UniShuttle app four key functions for users

The app firstly allows users to see the arrival time of the next shuttle bus. The app does so by providing the arrival time at the shuttle bus station in a list view (Figure 3-13). Upon an app user's selection of a particular bus route, the app presents a list of bus stops along the route. A countdown arrival time for the next two shuttle buses is shown for each stop (see Figure 3-13). This function is developed based on the research results showed that a countdown function has a more positive psychological impact on passengers while waiting for buses (Section 2.6.3).

Figure 3-12: The landing page of the UniShuttle app

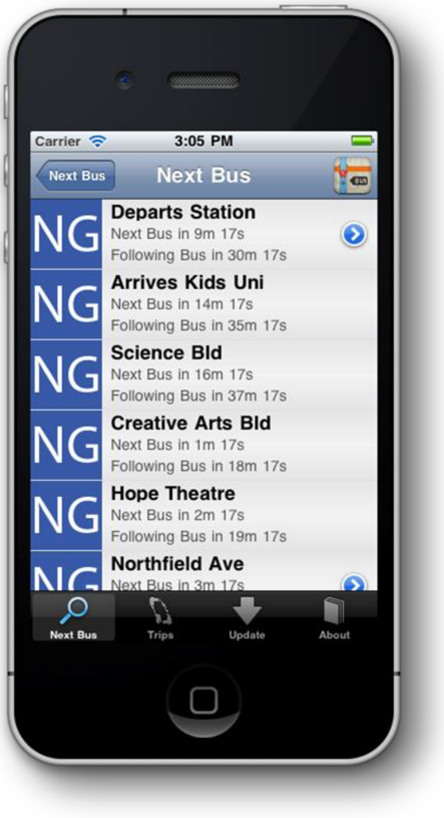


Figure 3-13: UniShuttle's view of next arrival time

Secondly, the app displays the geographic locations of users and shuttle buses through a Google map format, or a map view (Figure 3-14(a)). From the Next Bus Screen (Figure 3-13), by clicking the map icon on the top right hand corner of the screen, a Google map formatted information page appears with buses location, bus routes and bus stops highlighted on the map. The user's location is also displayed if it is within the vicinity of the bus routes. On this map screen, if a user clicks on a bus stop, the next two bus arrival times will be shown in a popped up information banner (Figure 3-14 (b)).

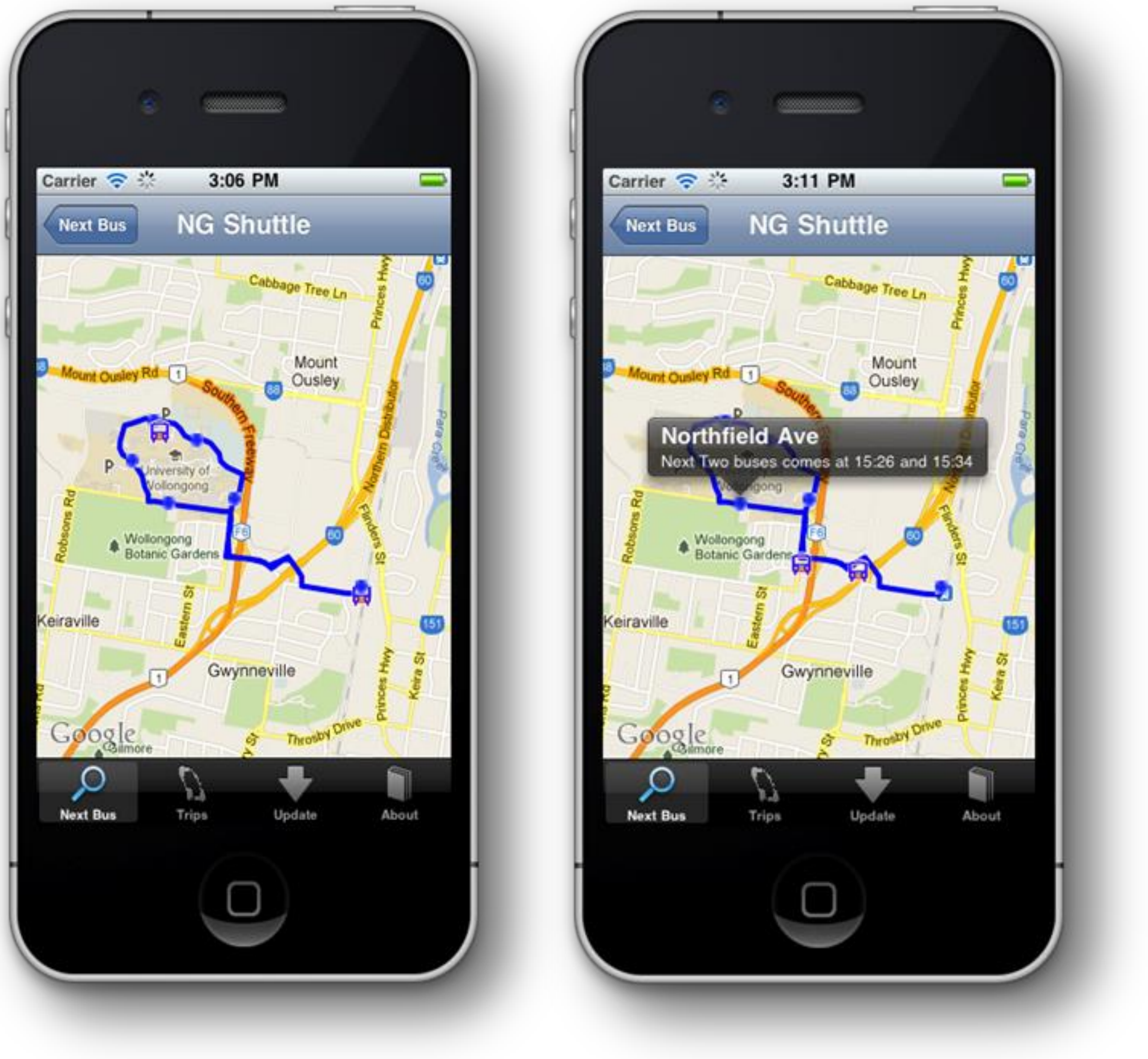


Figure 3-14: UniShuttle app displays current bus location on a Google map view. (a) The default current shuttle bus route display, (b) with pop up information of the next two shuttle bus arrival time

Its third function is to allow users to save their favourite trips with a designated starting bus stop and an ending bus stop into a shortcut access link, which users can access at any time they wish (Figure 3-15).

The fourth function is that the app notifies users of any shuttle bus timetable change (Figure 3-16). The users can then click one button to update their app timetable locally to ensure it keeps the latest timetable update. This function is particularly useful if users

have no access to internet, they at least can refer to more relevant timetable through the app.

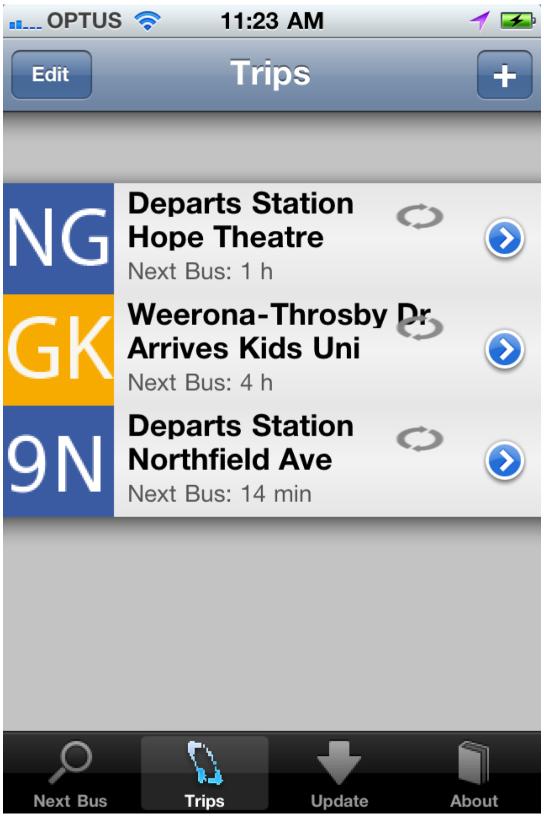


Figure 3-15: UniShuttle app's function to allow users to save favourite trips view

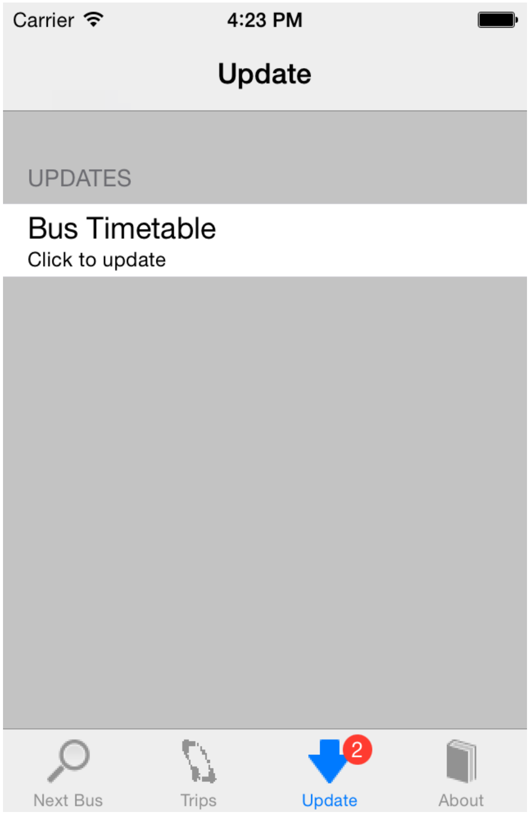


Figure 3-16: UniShuttle app's function to notify users with updated shuttle bus timetable

The app also performs the function to collect the app usage statistics such as session time, user location and the mac address of the iPhone, which are sent back to the system server for business intelligence reporting to the transport agent.

3.6.3 Showing the Map View Using KML

The UniShuttle app is developed based on client server architecture. As mentioned in section 3.5, the Bus Presence Server stores and computes the shuttle bus arrival time. The App acts as a client to retrieve this information and display it to the app users. One

unique characteristic of the app development is that app needs to transform the geographical data on the server side and display it on a map format. This challenge is overcome through using KML, which is a XML based language to display geographical information in 2D or 3D programs such as Google Earth or Google Maps. It has the ability to describe and express the geographical information. An example of using KML is in the implementation of the Current Bus Location View function. The KML messages returned from the server need to be parsed by the app to be displayed in a Google map format.

From the next Buses View (Figure 3-13), by clicking the map on the top right hand corner, the current bus location view will come into focus - the buses location, bus routes and shuttle stops are displayed on the map. The user’s location is also displayed if their location is within the vicinity of the bus routes. If the user clicks on a shuttle stop, the next two shuttle buses’ arrival time is shown (Figure 3-14 (b)). The App calls the Bus Location Service module on the Bus Presence Server (see Figure 3-10) through a HTTP web request, which returns a KML message that contains the current bus location, time and the arrival time information. The app then parses the KML message and displays the buses location on a map route on the user’s app. A brief request sequence can be viewed in Figure 3-17 below.

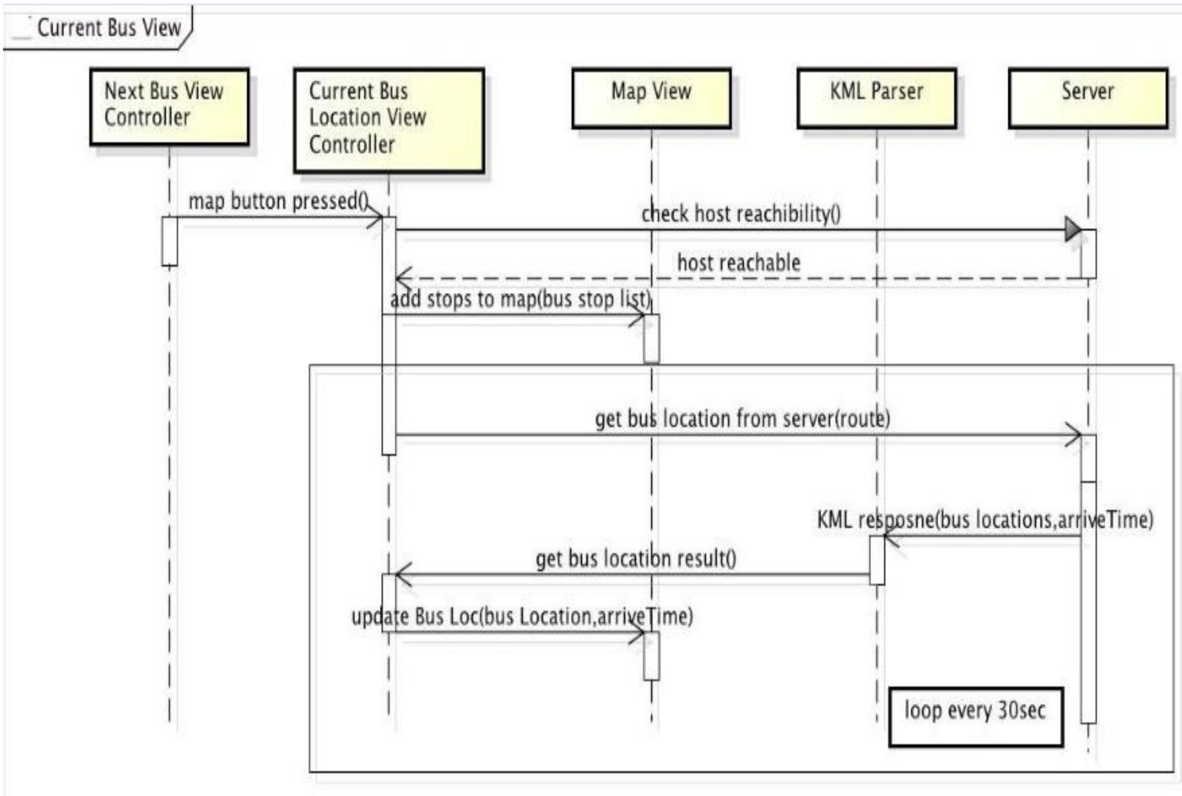


Figure 3-17: The Current Bus Location View Function's process sequence

The yellow square boxes represent the app controllers, views and modules except the last box "Server", which represents the boundary of the process and the entity of the server – the Bus Presence Server. The Next Bus View is a controller designed in the app which is triggered when the map button is pressed (Figure 3-13). It then calls the module of the Current Bus Location View controller, which is another controller responsible for commanding the shuttle buses location in a map format. It first checks whether it can reach the Bus Presence Server to establish communication. If so, it sends the user enquired shuttle stop to the next Module, Map View. The Map View module is responsible for displaying the geographical information in a Google map format. In this case, it generates the generic map information, which shows on the user's app screen as map loading. In the meantime, the Current Bus Location View Controller also sends the shuttle bus location information to the server – the Bus Presence Server, to request both

the route and the next two arrival time schedules. In response, the server responds with a KML message string, which is parsed by the KML Parser module - responsible for parsing the KML formatted messages sent from the server - so the geographical information can be displayed in a map format on the UniShuttle app. The following is an extract of the KML message received:


```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
	<Document>
		<Style id="unishuttle1">
			<IconStyle>
				<Icon>
					<href>/img/bus_icon.png</href>
					<scale>1.2</scale>
				</Icon>
			</IconStyle>
		</Style>
		 ...
		<Placemark>
			<name>NG unishuttle01</name>
			<description>Bus on Route NG, Arrives at Train Station in:
00m:15s</description>
			<styleUrl>#unishuttle01</styleUrl>
			<Point>
				<coordinates>150.888214108,-34.412499694,0 </coordinates>
			</Point>
		</Placemark>
		 ...
	</Document>
</kml>
```


This process is done through the Current Bus Location Module, which gets the deciphered message and sends it to the Map View Module to display the relevant information in a map format. The whole map display function is hence completed on the app. In order to display the route map and report the travel time prediction more accurately, the process will be refreshed every 30 seconds.

3.6.4 App Usage Statistics Collection

The UniShuttle app gathers the information about when, where and who uses the app every time the app is used. The information is collected inside the system database

server and data is mined in the system data warehouse server. Currently, the system reports on the app usage in terms of number of app users at different times. To facilitate future research the information collected by the app, such as location and user profile, will lead to more business intelligence. With increasing usage data gathered, the aggregated data will reveal the trends of users' behavioural patterns. For instance, if the usage pattern shows that there is a peak usage of the app around the university campus area around lunchtime on Mondays, this scenario might indicate a need to schedule more shuttle buses around that time in that area. The location data of the app usage might also indicate the areas people tend to congregate, such as a particular coffee shop in a certain area. Such information can be of use to transport authorities for public transport route planning. The app users profile can also be an insightful knowledge source for understanding public transport undertakings. For instance, if the trends revealed a certain age group of people are significantly less represented in the app usage, this might indicate that the transport agent may need to introduce certain actions such as conducting a targeted advertising campaign aimed at that age group. A key principle of the CMDE concept is to use user's interaction with the system to help the system self-organise and self-evolve just like an ecosystem does. This app's function is to provide the ability to gather usage data so that later these data can be analysed to help the system to improve, which can play a more and more important role for the UniShuttle's system's further development.

3.7 System Implementation Issues

The research project has been over four years in the making. From its conceptual design to its implementation, testing, releasing and evaluation, there were many challenges to

overcome. The reality is that there are always issues in a system implementation, which challenge the initial design and assumption.

3.7.1 Modifications Resulting from Implementation Issues

One issue which arose was not to rely on the Location Information Server (LIS) and GEOPRIV protocol to implement the shuttle bus location function in the system. GEOPRIV is an IETF standard for transmitting data over the internet, it specifies authorization, security and privacy requirement for LBS (Barnes et al., 2011) At the project initial stage, GPS equipped mobile devices were few and far between. A LIS was considered for use in the system to identify the location function of moving shuttle buses and App users. The research project's funding partner – Andrew Corporation (now CommScope), a telecommunication technology company, was to provide the LIS component in the project. The system's initial design also required the shuttle buses to use the campus Wi-Fi network while the network was available, which necessitated each Wi-Fi access point in the campus to be surveyed and recorded into the LIS. This needed LIS to be within the campus IT network, requiring the cooperation of the University's IT department. However with Andrews Corp's major restructuring and various other reasons we had to change our research implementation plan in response to changing requirement so that project can continue.

3.7.2 Mobility Program Requires Testing In Motion

There were many interesting findings during the software development of the system. One particular task is that the mobility program needs to be tested in a mobile environment as early as possible.

The initial software design of the on-board device used a single operating thread for the three modules - *GPS Data Reader, Bus Location Update and MAC Address Update*. The earlier version was to have these modules perform one after another, sleep for 10 seconds and restart the process again. The design worked in the lab. However, once the device was tested on board, the collected bus location data showed the bus making large hops on its route map, i.e. the bus appeared to be at a location for a 30-second to 1-minute period, then hopped to a location which was much further along on the route. The investigation on board showed that although the sleeping process was run, the GPS receiver continued receiving location data from satellites. The data was queued for the *GPS Data Reader* at a frequency around 5Hz, which was later largely read by the *GPS Data Reader* module. Further, the initial design relied on the timestamp on the on-board device, rather than on the GPS receiver. This caused the recorded location data to have the same location with a time gap corresponding to the length of the sleeping process (10 seconds). With the identification of the issue, the solution was quite easily achieved by using separate threads for the *GPS Data Reader* and the *Bus Location Processing* modules. The GPS data queue is read at every 0.2 second to ensure no buffer is built up and data lost.

Similar scenarios have reappeared in the app development. With the unexpected network connections encountered through the apps testing in various wireless networks, the asynchronous running threads were used rather than the previous synchronous ones, which freed up the app's resources and allowed a responsive user interface.

All of these experiences reinforced the lesson that a mobile program should be tested in its mobile environment as early as possible to ensure correctness of the design.

3.7.3 App Development Platform And Travel Pattern Identification

The market penetration rate of iOS devices was 36% in Australia at the time of the research project was developed (Dynamic Clarity 2011). Considering this statistics and the advanced location sensing technology of iOS devices (see Section 2.6.1), a decision was made to develop the UniShuttle app on iOS based platforms. Through the data experiments (as demonstrated in Chapter 4), the significant bus travel patterns (e.g., peak hour and off peak hour for both the normal services and the reduced services) have been identified. Hence developing the app on iOS proved to be sufficient to represent the traveller population for the purpose of this work - design and evaluate a digital ecosystem concept.

In the project's implementation, to identify the bus travelling patterns for predicating bus arrival time, MAC addresses of passenger devices were used to indicate individual passengers. And by anaylsing the collected MAC address information, the sessional travel patterns on all the routes were identified.

3.8 Summary

In this chapter, the UniShuttle system's architecture, design and implementation are described.

The system's architecture is based on the IETF presence model, which integrates the three major system components (i.e. on-board device, central server and iPhone App) into one integrated system. With the inspiration of building the Connected Mobility Digital Ecosystem, as one of the earliest projects at CMDE, the system software design follows the SOA principle. Its software functions, also called modules or service, are designed and implemented with the purpose of being utilised for the central services,

structural service and the UniShuttle system's specific services. The system's communication and messaging are designed to use a combination of communication protocols. When message security is of concern (e.g. the location sent by a shuttle bus needs to be saved at the central server), SSL is used for secure communication, with SOAP packaging format through the WSDL gateway to access the specific services at the destination. For security insensitive messages, HTTP protocol is used. The physical communication is based on the wireless network, 3G and Wi-Fi.

The on-board device, equipped with the AVL technology, is responsible for collecting vehicle location and passenger usage information. An in-vehicle PC (i.e. NEXCOM VTC2100) is chosen as the hardware platform for the shuttle buses based on the considerations of functionalities, maintenance and costs. The on-board device's software system uses a Linux operating system and Python language is used to develop its software modules. To avoid any risk of having the shuttle bus information being tampered with and the central server recording the incorrect information, the device uses a dedicated module (*Open Secure Link*) that utilises SSL and SOAP, to ensure secured communication with the central server web services.

The central server is a virtual server within a large IT infrastructure hosted at UOW. It includes four sub-server functions dedicated to different purposes. Its operating system is the server version Linux with Apache, Tomcat and Axis2 providing web services and MySQL for database management. The central server stores, controls and communicates with the distributed shuttle buses and the passengers' mobile devices. Its other key functions include predicting shuttle bus arrival times and analyzing historical data to provide business intelligence information. These functions (otherwise called modules or services) are designed to be linked and interacted together to perform the

central server functions following the SOA principle. They are developed using programming language Java and PHP.

The UniShuttle app is the third key component in the system. It is a location aware application utilizing iPhone's own location sensing technologies. Programming for the app was written using Object-C programming language. The app software design follows the Model-View-Controller design pattern. The UniShuttle app's key functions are designed around the provision of real-time bus arrival information. Its four key functions for users are: displaying a list view of arrival times; displaying a map view of users and buses; displaying a list view of saved favourite trips; and providing offline travel information updates. One unique characteristic of the app development is to transform the geographical data on the server side and display it in a map format on the app. The KML is used in the implementation of this feature. The app also gathers the information about who is using the app and when and where they were. The information can be analysed later to gain business insights on resource management, passenger behaviour understanding and so on. Together with the information collected from shuttle buses, the business intelligence gained can feedback to the system to allow self improvement, which is a key feature of a digital ecosystem.

Lastly, the chapter includes a brief mention of challenges and some lessons learnt during the research project, including overcoming obstacles in research environment and facilities, identifying and accommodating the unique characteristics of mobile programs and app development to ensure smooth development.

CHAPTER 4. SHUTTLE BUS TRAVEL DATA ACQUISITION AND ARRIVAL TIME PREDICTION

4.1 Description of Shuttle Bus Routes

As the university students are the key passenger group of the shuttle buses service, and many students rely on trains for their transportation, the shuttle bus timetables are scheduled with the university teaching sessions and the train timetable considered. The North Gong route (Route #1) and the express North Gong route (Route #2), both connecting the train station to the University, have 90 services per day during university sessions and 31 service per day during university holidays. The Gwynneville-Keiraville route (Route #3) only operates in the morning and afternoon during session time with 27 services per day.

Route #1, #2 and #3 are the test platform in this study. The details of these routes are described as follows. Route #1 has 7 bus stops: it starts at the North Wollongong Train Station, passes two freeway exits before entering into the University Ring Road, then returns to the Train Station. Route #2 is an express version of the first route with 3 stops: it does not pass the University Ring Road (see the blue lines in Figure 4-1). Route #3 has 13 stops: it goes through the University Ring Road and then passes through the surrounding suburbs including part of the Central Business District (CBD) (see the thin yellow lines in Figure 4-2). The details of the distances between the starting station and the other stations for the three routes are listed in Table 4-1.

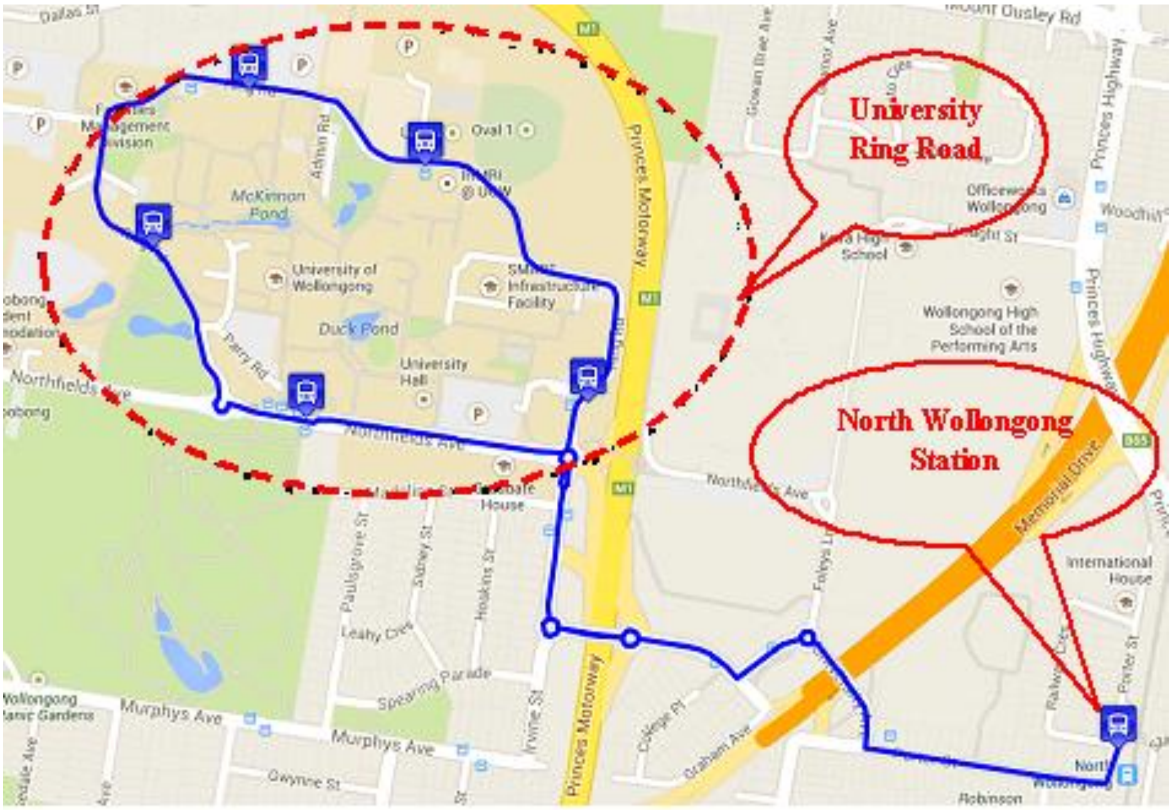


Figure 4-1: Shuttle bus Route #1 and Route #2 on Google map

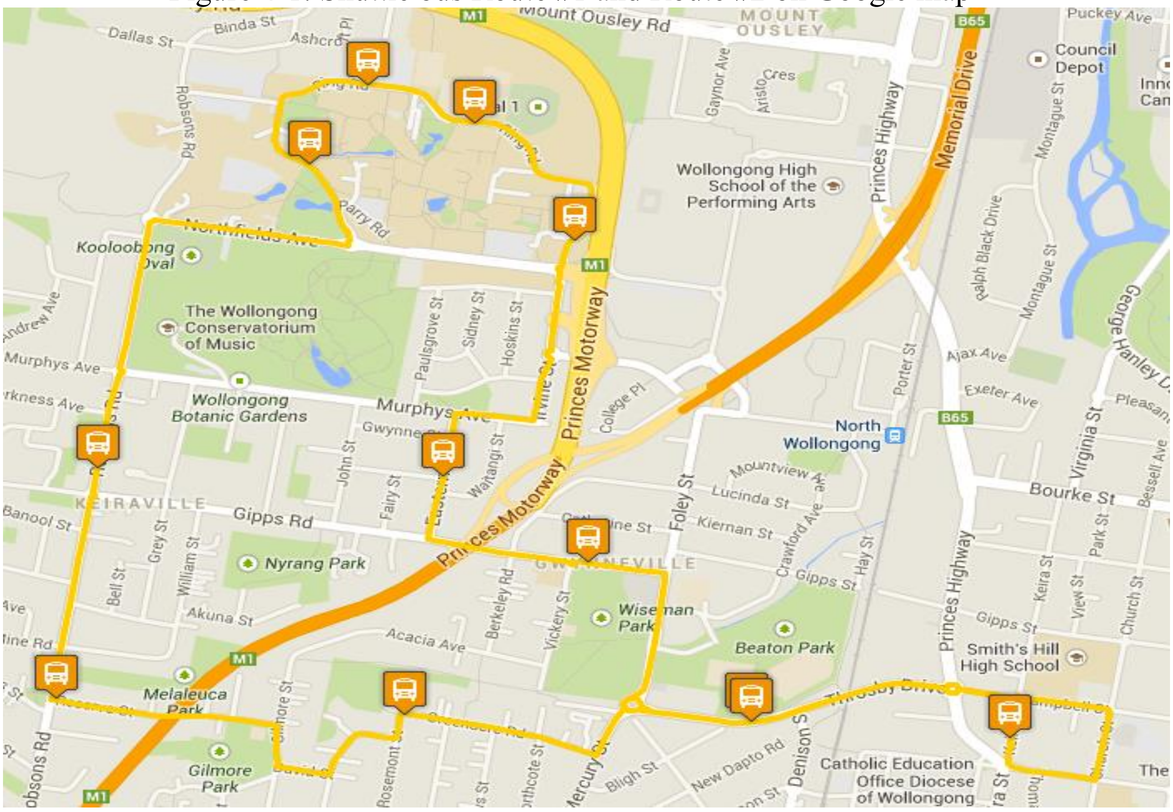


Figure 4-2: Shuttle bus Route #3 on Google map

Table 4-1: Distances of the bus stops from starting station on the three routes

Distance from the starting station (Km)								
Route #1			**Route #2**			**Route #3**		
Bus Stop	Bus Stop Name	Distance	Bus Stop	Bus Stop Name	Distance	Bus Stop	Bus Stop Name	Distance
Stop 1	Train Station	0	1	Train Station	0	1	Kids Uni	0
2	Kids Uni	1.31	2	Northfield Ave	1.81	2	Science Bld	0.55
3	Science Bld	1.85	3	Train Station	3.39	3	Creative Arts Bld	0.92
4	Creative Arts Bld	2.23				4	Hope Theatre	1.37
5	Hope Theatre	2.67				5	Robsons Rd	2.75
6	Northfield Ave	2.95				6	Reserve St	3.47
7	Train Station	4.54				7	Greenacre Rd	4.97
						8	Weerona-Throsby Dr	5.41
						9	Keiraview	6.21
						10	Throsby Dr	7.55
						11	Wiseman Park	8.43
						12	Eastern St	9.13
						13	Kids Uni	9.88

With different travel routes, Route #1, #2 and #3 serve passengers with different trips, and consequently they have different route characteristics. For example, they present different patterns for the length of dwelling time at a bus stop and the travel time on the road. To understand their travel patterns, the trip data is analysed against the factor of time, including the time of a day, school day, holiday, and the university session (since university students and staff are the main passenger groups).

4.2 The Passenger Usage Patterns

For shuttle bus travel time prediction, passengers' travel patterns also play a key role in terms of whether a shuttle bus is able to follow its planned schedule. Thus, it is important to establish the passengers' flow pattern.

The shuttle buses are a free and ticketless service, therefore there is no existing knowledge of where and how many passengers get on or off at a bus stop. Through the two approaches (described in Chapter 3) to gain the passenger information, we were able to establish an understanding of passenger usage patterns.

At the early deployment of the app, an average of 4,630 passengers per month were recorded in the system, which later increased to the range of 5,000 - 10,000 passengers during the university teaching sessions and the range of 2,000-3,000 during holiday periods. The recorded passenger data was processed to establish passengers' bus-travel daily pattern (or passenger usage pattern).

To describe a typical pattern identified in a month, statistics recorded for a typical shuttle service in the month of May 2012 with aggregation on time of day and day of week is shown in Figure 4-3.

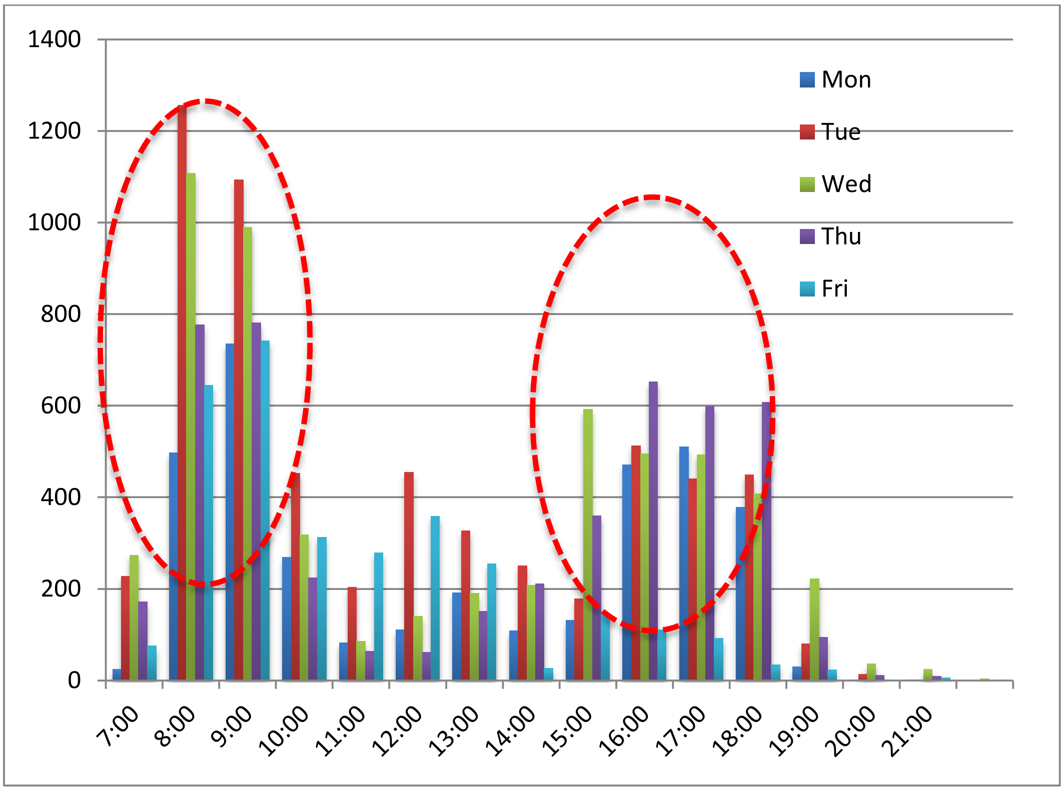


Figure 4-3: Typical passenger usage statistics recorded for May 2012
The vertical axis is the number of MAC addresses captured in the UniShuttle system, which represents the number of passengers on board.

As it is shown, there are more passengers who use shuttle buses during the morning peak hours (7:00-10:00 am), afternoon peak hours (4:00-6:30pm). It is also notable that the patterns on every week day (Mon-Fri) vary, as shown by the different coloured bars in the figure. For example, on Wednesday, there are more passengers in the afternoon while fewer passengers takes the bus in the afternoon on Friday.

To describe a typical pattern identified for a year, passenger usage statistics recorded for a typical shuttle service year, 1st Jan 2012 – 31st Dec 2012, is shown in Figure 4-4.

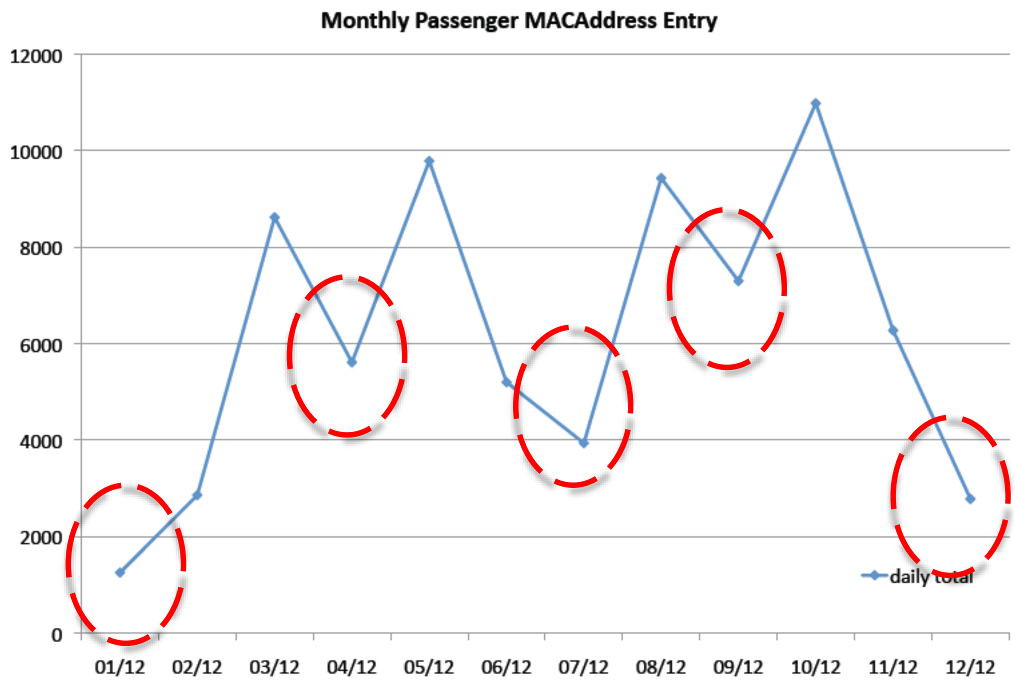


Figure 4-4: Typical passenger usage statistics recorded for 1/ 1/ 2012 – 31/ 12/ 2012
The vertical axis is the number of unique MAC addresses captured in the UniShuttle system, which represents the number of passengers on board.

It is evident from the figure that the number of passengers during the university summer holiday (from November to February) is the lowest; the second lowest is during the university winter holiday (in July). The university session breaks (in April and September) also have lower entries.

The distinct features of passengers' usage patterns are needed to be taken into account for shuttle bus scheduling and travel time predictions. Upon presenting these passenger statistics information, the UniShuttle operator was able to make informed shuttle bus scheduling adjustments, such as changing the number of shuttle buses and their running frequencies at different times of the day and month.

4.3 Acquisition and Processing of Shuttle Bus Travelling Data

4.3.1 The Shuttle Bus Travelling Data Acquisition

Recalling the GPS data acquisition covered in Chapter 3, the shuttle bus traveling data captured in the system includes bus location (latitude, longitude), direction and the velocity, as well as timing information. A sample of the dataset in the system is listed in Table 4-2 below.

Table 4-2: Bus travelling data obtained from on-board device

bus_id	latitude	longitude	time_stamp	speed	heading
unishuttle01	150.883873	-34.410943	2012/03/05 08:00:07	15.50115	278.2826
unishuttle05	150.88142	-34.415953	2012/03/05 08:00:08	10.6264	278.4948
unishuttle02	150.884468	-34.411062	2012/03/05 08:00:11	17.04035	290.4626
unishuttle01	150.883295	-34.410919	2012/03/05 08:00:18	10.40255	303.8089
unishuttle05	150.880561	-34.41584	2012/03/05 08:00:19	18.19845	281.8854
unishuttle02	150.883713	-34.410957	2012/03/05 08:00:21	8.58955	267.2321
unishuttle05	150.879511	-34.415621	2012/03/05 08:00:29	16.2134	283.5028
unishuttle02	150.883078	-34.410874	2012/03/05 08:00:31	15.26435	279.6703
unishuttle05	150.879034	-34.41539	2012/03/05 08:00:39	9.8975	17.1957
unishuttle01	150.882062	-34.410673	2012/03/05 08:00:40	10.4451	12.3273
unishuttle02	150.882373	-34.410837	2012/03/05 08:00:41	8.95585	260.3105
unishuttle05	150.879235	-34.41457	2012/03/05 08:00:49	22.7809	9.2856
unishuttle01	150.882228	-34.409827	2012/03/05 08:00:51	22.2962	7.7928
unishuttle02	150.882135	-34.410518	2012/03/05 08:00:52	14.5262	10.4583

The dataset has the shuttle bus identification (bus_id), which specifies which bus the data is reported from. The bus location (latitude, longitude), time, speed and direction of heading are imported from the GPS reading. The speed is the instantaneous speed the

bus is traveling at, the direction of heading is the angle between north and the bus travelling direction.

There were 2,298,249 raw data entries gathered for the entire year of 2012. These raw travel data need careful selection, such as they should only include the actual service routes, so the data should not include the travel information from a bus depot to the beginning of a route, and from the last stop back to a depot. A careful review process was conducted to fine tune the data, and the raw data was transformed into meaningful and useful information. This process is described in the section below.

4.3.2 The Shuttle Bus Travelling Data Processing

In order to understand and analyse the bus traveling patterns, the bus traveling data need to be processed. The routes are split into segments. A segment is a small section of a route which is in the form of a straight line. The first step of the data processing is to determine whether or not a bus is located on the defined bus routes or not. As route segments are the smallest units that split the bus routes with starting and ending position, as well as travelling direction (or heading), this step can be achieved by the determination of whether a bus location is on a route segment.

Theoretically, to determine whether a point is on a straight line, the mathematic formula is to sum up the distance from this point to the two ends of the line. When the difference between the sum and the length of the line is zero, this point is on the line. In the case of this system, in order to determine whether a vehicle location (p) belongs to a segment (S) of the bus route as shown in the Figure 4.5, the sum of the distance between p to starting point of x ($D_{p,sx}$) and the distance between p to the end point of y ($D_{p,sy}$) minus the length of the segment should ideally be zero or realistically be a minimum value the

system can compute. With the consideration of the GPS signal accuracy, a minimum value of 50 meters or less is chosen as the threshold for the comparison with the result of this value. The mathematic expression is shown in Eq. 4.1. From experiments, as vehicles can travel in opposite directions or take turns on road, we identified the imperative of the Eq. 4.1 is to require the vehicle (p) travel direction within 100 degrees of the segment (S).

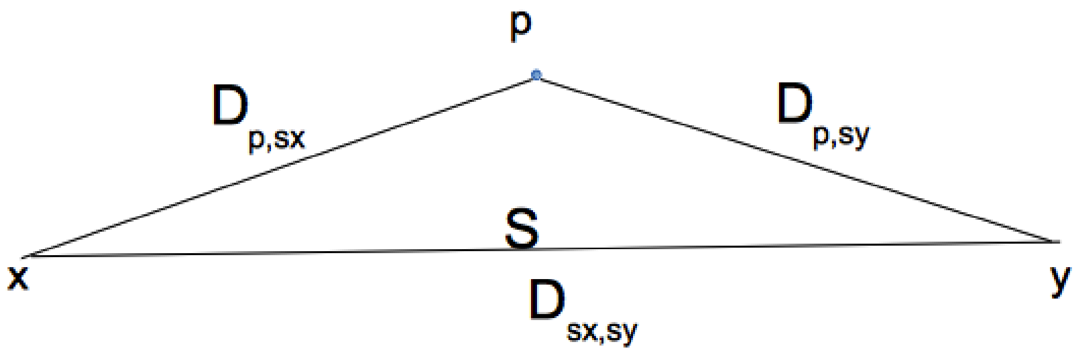


Figure 4-5: Schematic description of determination of vehicle location
p is the bus location; S is a segment on the route; x is the starting point of segment S; y is the ending point of segment S; sx is the starting location of the segment; and sy is the ending location of the segment.

$$MIN(\left|D_{p,sx} + D_{p,sy}\right| - D_{sx,xy}) \qquad \text{Eq. 4.1}$$

To calculate the distance between two locations, the Haversine formula (Inman 1849) is used. The Haversine formula calculates the shortest distance over the earth's surface. Its equation for calculating the distance between two stops on the selected shuttle bus route is expressed in Eq. 4.2

$$d = 2 \times R \times \sin^{-1}\left(\sin^2\left(\frac{\varphi_a - \varphi_b}{2}\right) + \cos(\varphi_a) \times \cos(\varphi_b) \times \sin^2\left(\frac{\lambda_a - \lambda_b}{2}\right)\right) \quad \text{Eq. 4.2}$$

φ_a : Latitude location a;

φ_b : Latitude location b;

λ_a : Longitude location a;

λ_b : Longitude location b;

R = 6371km (the Earth radius);

d : the distance between two stops.

After identifying a vehicle is on a defined route (by identifying that it is on a segment), the remaining steps of processing the location data are around the goal of calculating the travelled time of the particular shuttle on its trip. As in the historical data prediction module, the historical average travel time is a key factor in the prediction. The system implementation to achieve this goal includes marking the bus location which is within 30 meters to the nearest bus stops in order to identify whether a bus dwells at bus stop or travels on the road; converting the timestamp to an easily calculated time in order to use the travelled time to calculate travelled distance; calculating the travelled distance for a bus's moving location in order to predict the future travel time.

4.4 Results and Discussion of the Bus Travelling Data Analysis

Three key factors impact the shuttle bus arrival time in its service. These three factors are: bus travelling time on-road; dwelling time at bus stops; and the time adherence to the defined schedules. In this section, the shuttle bus travelling data is analysed against these three factors and is presented in a graph format. The calculated average travelling time, dwelling time and adherence to the timetable of the route and the standard deviation from the average value of these variables are summarized for each case.

4.4.1 Shuttle Bus Travelling Time at Different Routes and Time

The shuttle bus travelling data collected during the bus services on the three routes are presented in graphs. The horizontal axis shows the stop identities, the vertical axis

shows the travelling time used at the stops. Due to the unique characters of the different times of day, the data is analysed against the different hours of the day: morning peak time (7 am - 10 am), daytime off-peak (10 am - 2:30 pm), afternoon peak time (2:30 pm -6:30 pm) and evening off-peak time (6:30 pm - midnight). Except for the characters of the different times of the day, shuttle bus travelling data also shows distinct characteristics between school days and school holidays. Hence, the data is also analysed against these two categories of days: normal service time (i.e. school days) and reduced service time (i.e. school holidays).

Figure 4.6 to Figure 4.9 and Figure 4.10 to Figure 4.13 show the bus travelling time at different bus stops on Route #1 and #2 respectively, including eight different time periods in each route. Table 4.3 and 4.4 present the average travel time and standard deviation from average travel time on route #1 and #2 in each time period corresponding to the curves shown in Figure 4.6 to Figure 4.9 and Figure 4.10 to Figure 4.13. For Route 3 there is no reduced service, hence only its normal service data is presented in Figure 4.15 and Figure 4.16 for peak hours and off peak hours with normal service. The average travel time and standard deviation from average time are given in Table 4.5.

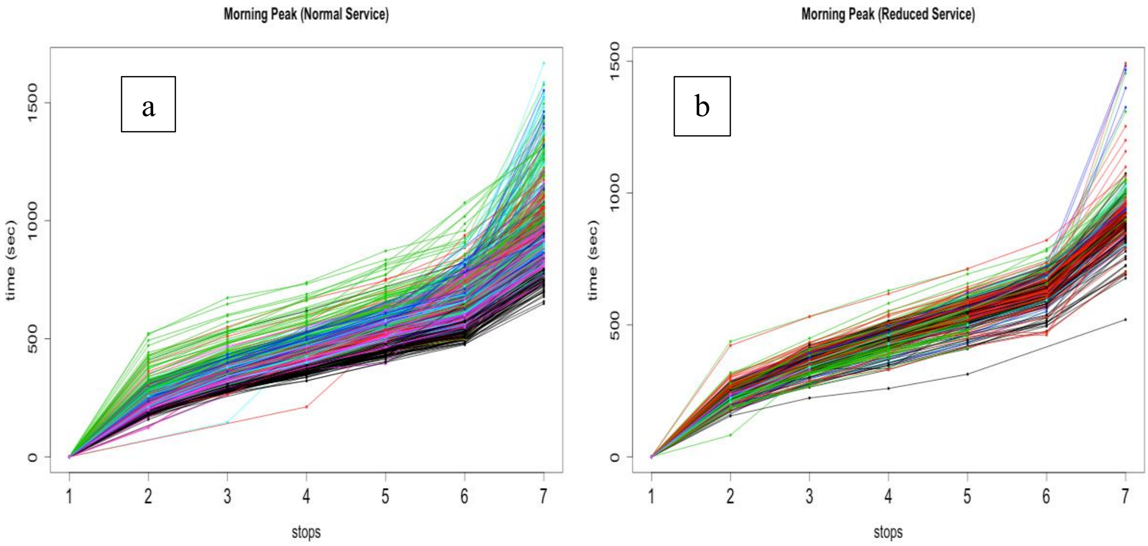


Figure 4-6 Shuttle bus travelling time at stops on Route #1 captured during the morning peak hour with normal (a) and reduced service (b)

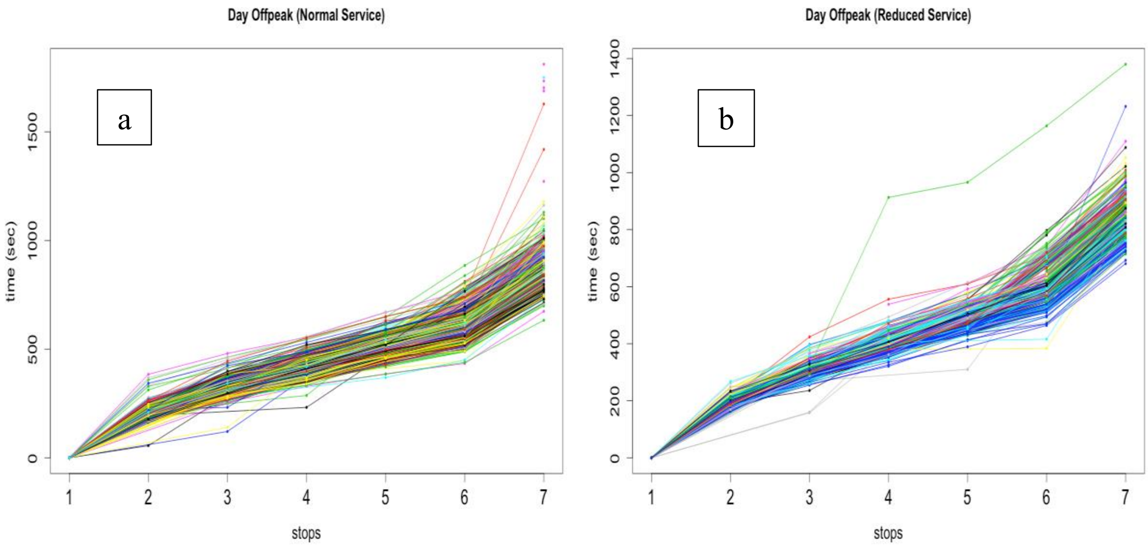


Figure 4-7: Shuttle bus travelling time at stops on Route #1 captured during day off-peak hours with normal (a) and reduced service (b)

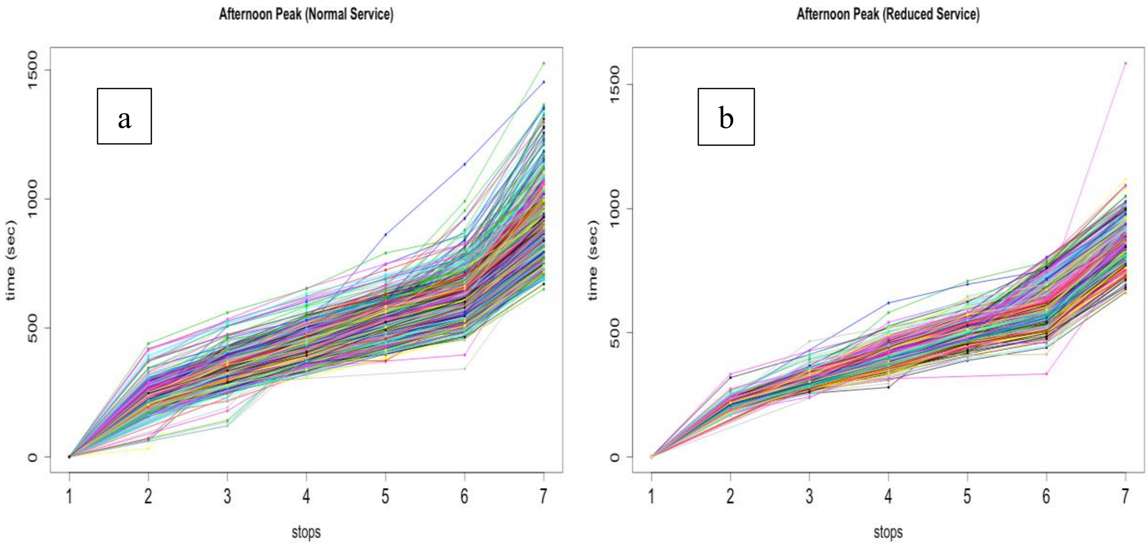


Figure 4-8 Shuttle bus traveling time at stops on Route #1 captured during the afternoon peak hour with normal (a) and reduced service (b)

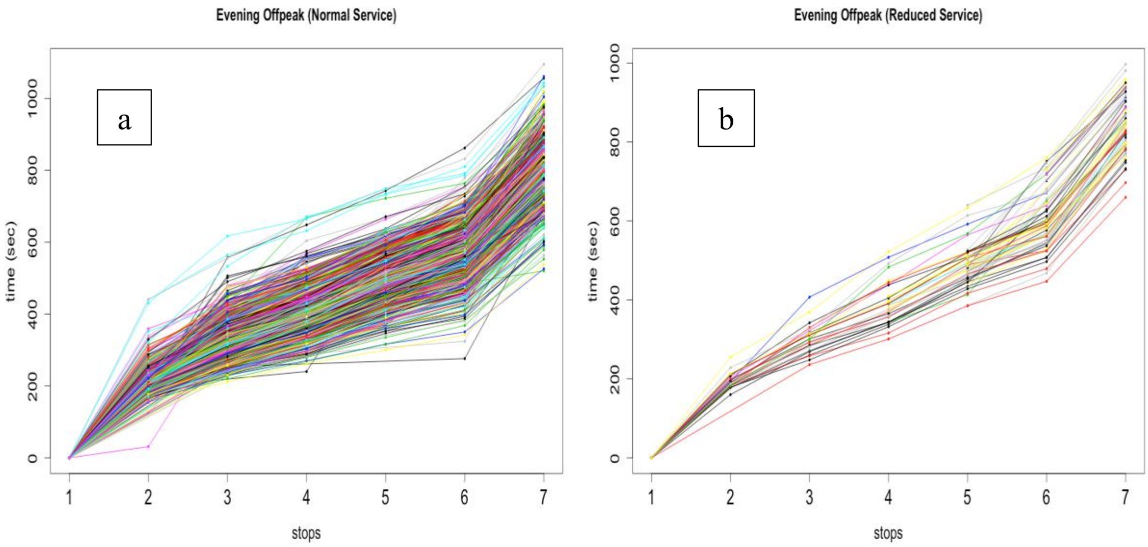


Figure 4-9: Shuttle bus travelling time at stops on Route #1 captured during the evening off-peak hour with normal (a) and reduced service (b)

Table 4-3: Average Shuttle bus travel time and standard deviation from average travel time for normal and reduced bus service on Route #1 during peak and off-peak hours

		Mean (sec)	Standard Deviation (sec)	Mean (sec)	Standard Deviation (sec)
		Normal Service		Reduced Service	
	Morning Peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Kids Uni	231.58	45.24	237.17	30.08
3	Science Bld	346.78	49.74	354.38	35.78
4	Creative Arts Bld	437.73	53.68	450.86	44.11
5	Hope Theatre	537.24	59.62	545.52	49.62
6	Northfield Ave	644.37	81.20	644.57	59.33
7	Arrive Train Station	933.87	158.73	925.85	116.26
	Daytime Off-peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Kids Uni	207.15	22.90	200.07	22.15
3	Science Bld	315.34	29.18	307.30	31.87
4	Creative Arts Bld	414.43	37.39	402.98	45.21
5	Hope Theatre	510.55	38.69	493.01	47.50
6	Northfield Ave	608.11	57.09	600.82	69.98
7	Arrive Train Station	864.24	99.62	851.83	76.57
	Afternoon Peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Kids Uni	224.46	33.52	205.95	18.63
3	Science Bld	337.56	44.63	307.33	27.12
4	Creative Arts Bld	432.64	46.66	406.81	40.20
5	Hope Theatre	524.12	51.65	503.04	43.18
6	Northfield Ave	616.63	66.28	604.33	70.09
7	Arrive Train Station	883.52	101.78	866.96	77.48
	Evening Off-peak				
1	Depart Train Station	0.00	0.00	0.00	0.00

2	Kids Uni	214.99	33.10	194.00	19.75
3	Science Bld	335.62	55.36	300.96	36.66
4	Creative Arts Bld	418.33	63.47	392.33	53.71
5	Hope Theatre	505.94	66.96	488.55	53.30
6	Northfield Ave	580.14	74.04	605.21	79.09
7	Arrive Train Station	815.54	84.57	843.49	73.03

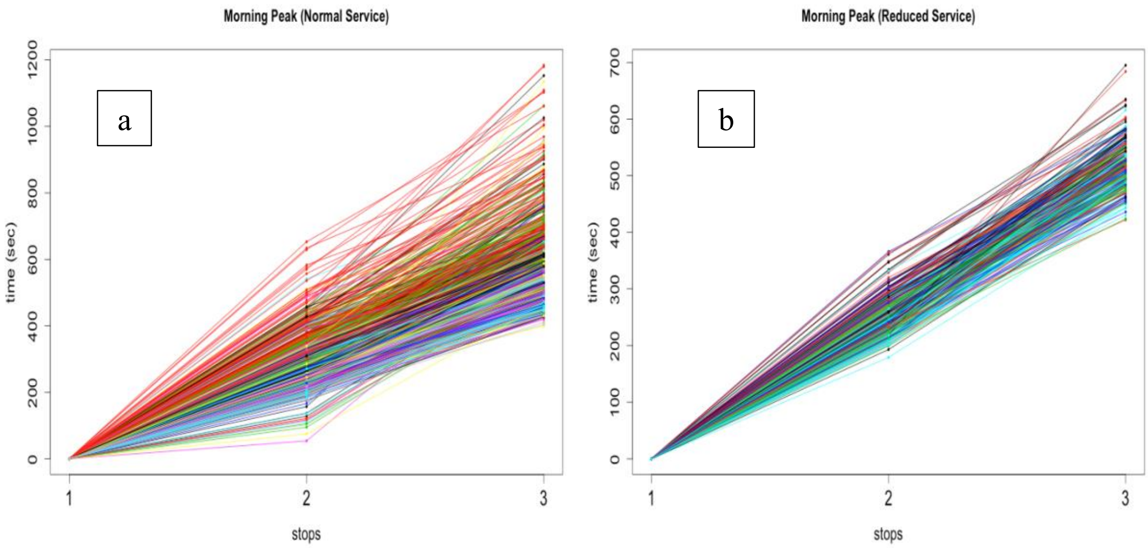


Figure 4-10**:** Shuttle bus travelling time at stops on Route #2 captured during morning peak hour with normal (a) and reduced service (b)

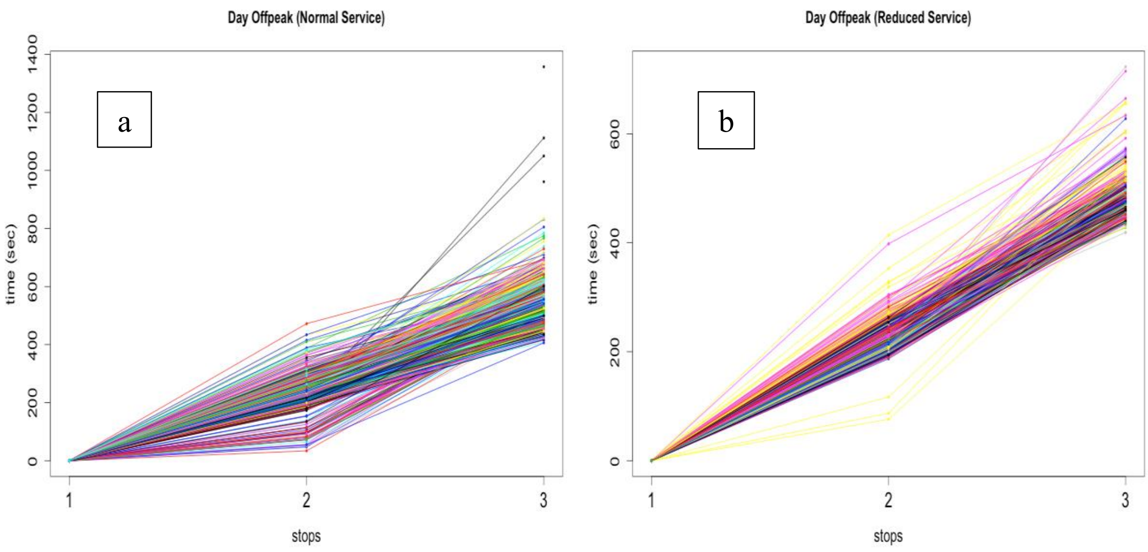


Figure 4-11**:** Shuttle bus travelling time at stops on Route #2 captured during day off-peak hour with normal service (a) and reduced service (b)

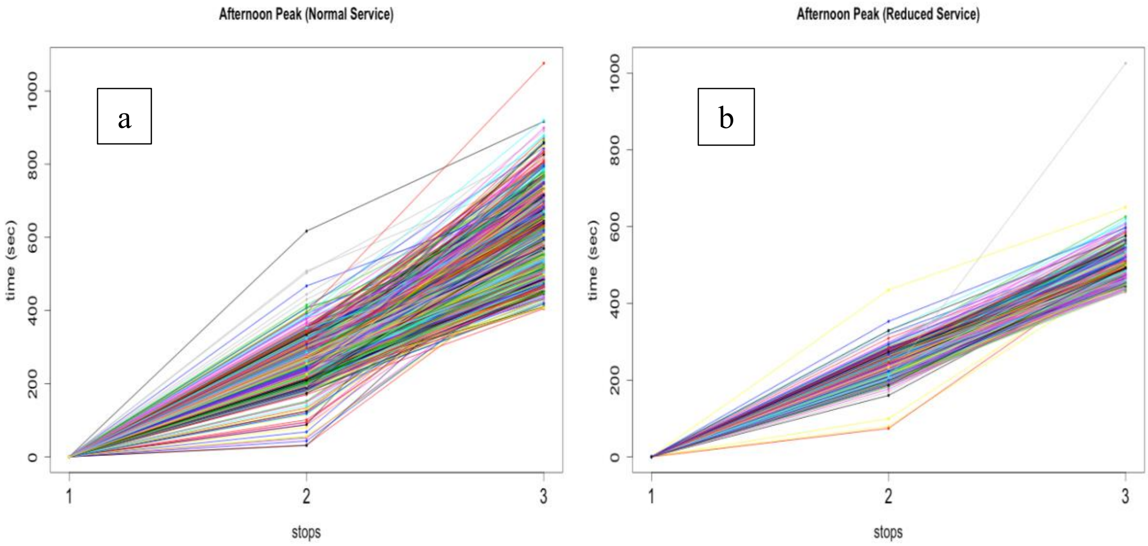


Figure 4-12: Shuttle bus travelling time at stops on Route #2 captured during the afternoon peak hour with normal service (a) and reduced service (b)

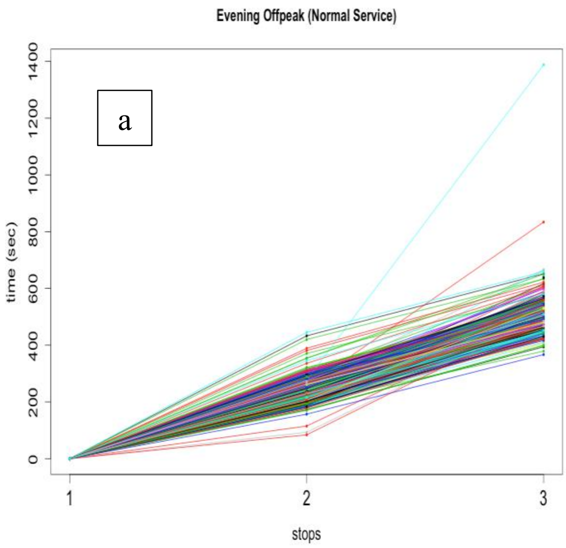


Figure 4-13: Shuttle bus travelling time at stops on Route# 2 captured during the evening off-peak hour

Table 4-4: Average Shuttle bus travelling time and standard deviation from average travel time for normal and reduced bus service on Route #2 during peak and off-peak hours

		Mean (sec)	Standard Deviation (sec)	Mean (sec)	Standard Deviation (sec)
		Normal Service		Reduced Service	
	Morning Peak				

1	Depart Train Station	0.00	0.00	0.00	0.00
2	Northfield Ave	271.78	55.39	258.33	34.22
3	Arrive Train Station	555.28	90.81	519.00	42.19
	Daytime Off-peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Northfield Ave	240.38	36.49	236.41	35.44
3	Arrive Train Station	531.40	54.00	497.48	57.95
	Afternoon Peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Northfield Ave	249.88	39.41	236.29	30.39
3	Arrive Train Station	564.78	73.26	496.56	39.46
	Evening Off-peak				
1	Depart Train Station	0.00	0.00		
2	Northfield Ave	238.22	37.03		
3	Arrive Train Station	503.43	52.99		

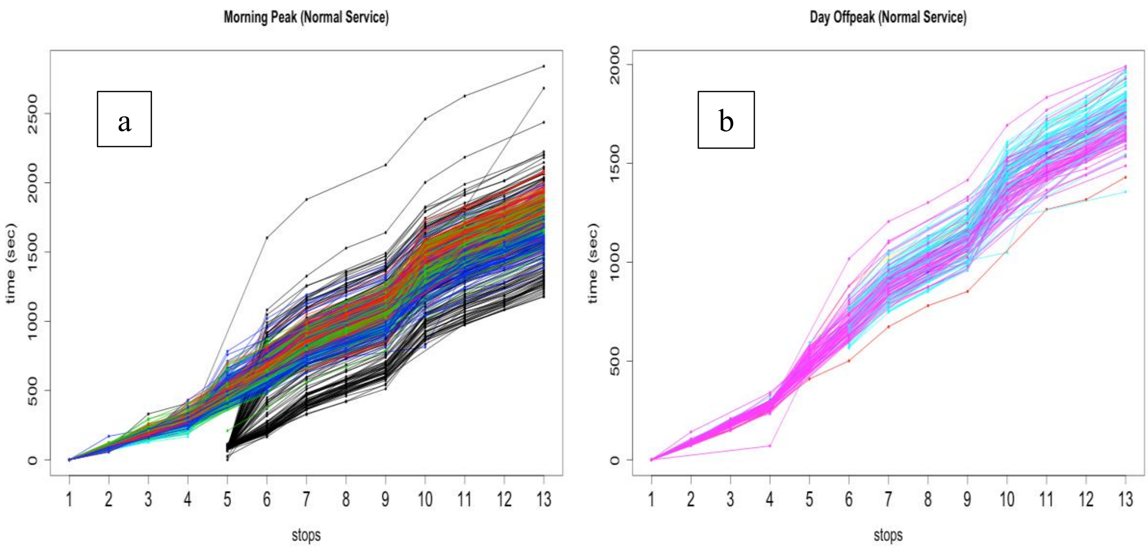


Figure 4-14: Shuttle bus travelling time at stops on Route #3 captured during the morning peak hour (a) and day off-peak hour (b)

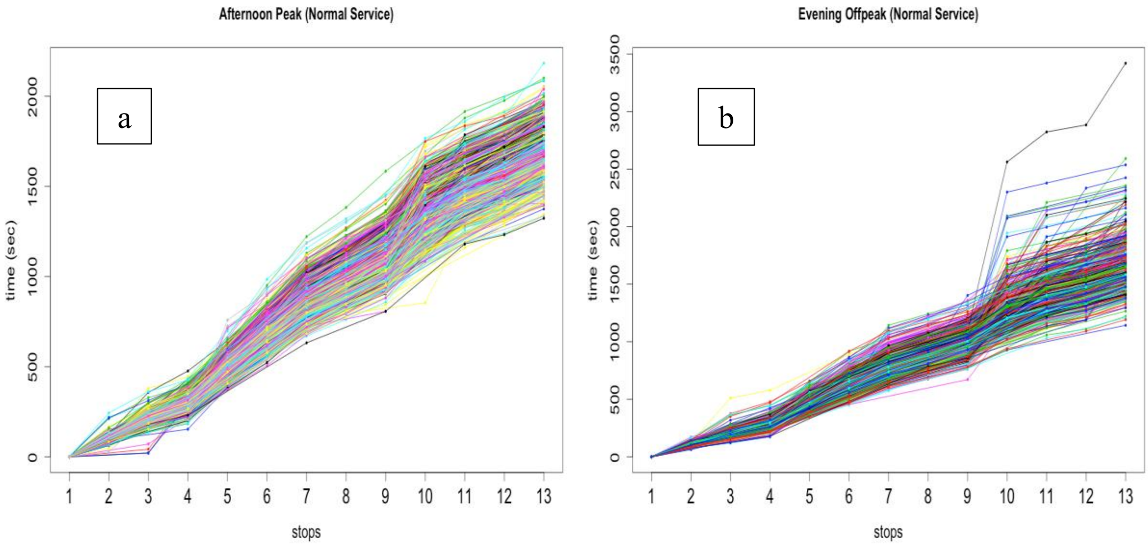


Figure 4-15: Shuttle bus travelling time at stops on Route #3 captured during the afternoon peak hour (a) and evening off-peak hour (b)

Table 4-5: Average Shuttle bus travelling time and standard deviation from the average travel time on Route #3 during peak and off-peak hours

		Mean (sec)	Standard Deviation (sec)
	Morning Peak		
1	Depart Kids Uni	0.00	0.00
2	Science Bld	91.14	11.73
3	Creative Arts Bld	184.73	22.65
4	Hope Theatre	279.51	36.87
5	Robsons Rd	374.46	198.47
6	Reserve St	589.66	181.02
7	Greenacre Rd	775.26	188.50
8	Weerona-Throsby Dr	898.81	193.98
9	Keiraview	1006.17	198.40
10	Throsby Dr	1332.26	212.39
11	Wiseman Park	1461.76	209.28
12	Eastern St	1553.87	209.96
13	Arrive Kids Uni	1682.49	217.57
	Daytime Off-peak		
1	Depart Kids Uni	0.00	0.00
2	Science Bld	88.20	10.33
3	Creative Arts Bld	180.30	13.22
4	Hope Theatre	273.32	31.62
5	Robsons Rd	515.16	43.19
6	Reserve St	699.86	85.65
7	Greenacre Rd	896.51	81.43
8	Weerona-Throsby Dr	1013.15	81.47
9	Keiraview	1121.43	89.95

10	Throsby Dr	1431.33	94.96
11	Wiseman Park	1554.75	102.34
12	Eastern St	1642.84	105.41
13	Arrive Kids Uni	1741.11	112.31
	Afternoon Peak		
1	Depart Kids Uni	0.00	0.00
2	Science Bld	104.25	16.69
3	Creative Arts Bld	205.01	32.37
4	Hope Theatre	298.09	39.92
5	Robsons Rd	506.78	57.00
6	Reserve St	689.79	70.79
7	Greenacre Rd	878.30	87.74
8	Weerona-Throsby Dr	1008.34	96.11
9	Keiraview	1122.46	105.65
10	Throsby Dr	1416.21	118.16
11	Wiseman Park	1524.34	123.73
12	Eastern St	1587.11	133.74
13	Arrive Kids Uni	1722.06	137.69
	Evening Off-peak		
1	Depart Kids Uni	0.00	0.00
2	Science Bld	107.75	17.31
3	Creative Arts Bld	197.40	37.16
4	Hope Theatre	279.77	46.10
5	Robsons Rd	476.08	62.10
6	Reserve St	641.37	80.51
7	Greenacre Rd	802.58	96.51
8	Weerona-Throsby Dr	915.69	104.67
9	Keiraview	1006.31	109.20
10	Throsby Dr	1354.90	198.54
11	Wiseman Park	1456.28	193.17
12	Eastern St	1518.68	201.59
13	Arrive Kids Uni	1657.34	213.09

The above graphs, showing the travelling time displayed against bus stops (the sequential bus stops are further distant from the starting bus stop), highlight some common characteristics for arrival time prediction.

Graphically, the vertical spread of the travelling time between the maximum and minimum time indicates the extent of the shuttle bus punctuality. The larger this spread is, the more diverse the travel pattern is, or in other words, the bus arrival time is more varied from the fixed scheduled time. It is notable that the vertical spread is smaller for

the reduced services compared to the normal services. This can be explained that, because during the normal services there tend to be more passengers resulting in more boarding/de-boarding time; further, the normal services are often on work days and non-holiday period, which means traffic conditions are often busier than otherwise. As a result, the normal services' graphs show a more scattered pattern than reduced service ones.

With this pattern observed, it is also notable that there is irregularity for the reduced services. This is shown by the number of scattered lines off the main trend in the relevant graphs. The limited number of reduced bus services operated is the likely causes. It is also noticeable that the graphs for peak hour are more scattered than the off-peak ones; this is likely due to the high flow of passengers and the busier traffic conditions during the peak hours.

Route #1 and #2 have a larger variation in the bus travelling time during the peak hour periods comparing to Route #3. This is because the two routes primarily serve the travellers to and from the university. For example, bus stop 6 (Northfields Ave, which is a key bus stop to the university) and stop 7 (train station) are very busy stops and often have delays as shown in all the graphs for Route #1, particularly during peak hours. Route #3 (Gwynneville-Keiraville route) also has similar variation during peak hours. Moreover, as it covers part of the busy CBD area, the large time variation also shows in the evening off-peak period.

The travelling time patterns are further quantified through calculating the average travel time and the standard deviation from the average travel time as shown in Table 4-3 for Route #1, Table 4-4 for Route #2 and Table 4-5 for Route #3. The results are in

agreement with the graphs shown in Figure 4-6 to Figure 4-15. It must be mentioned that some special cases such as the evening off-peak hour on Route 3 are supposed to be less scattered. But there is rapid increase in the deviation in the last five stops (Figure 4-15(b) and Table 4-5). This is likely to be caused by the fact that these stops are located in the University owned student accommodations near the CBD.

The important finding from the data analyses and presentation in this section is the not the identification that different routes, different times of day, and different days of a year can result in different bus travelling time patterns, but rather these patterns can be quantified so useful actions can be taken. The bus travelling time, along with the factors of route, time of day and day of year, should be used for the travel time prediction for the corresponding route and time period.

4.4.2 Shuttle Bus Dwelling Time at Different Bus Stops and Times

Dwelling time is the second factor impacting the prediction of the shuttle bus arrival time. It is the time a shuttle bus stops at a given bus stop, calculated by taking the difference between the arriving and departing time of a shuttle bus at a bus stop.

The shuttle bus dwelling data calculated during the bus services on the three routes is presented in graphs. The horizontal axis shows the stop identities, the vertical axis shows the dwelling time at the stops. Similar to the bus travelling time analysis, the dwelling time analyses are carried out for the distinct time groups: different time of a day (e.g. peak hours, off-peak hours) and different days of a year (e.g. holidays or non-holidays).

Figure 4-16 to Figure 4-23 shows the dwelling time at different bus stops and time groups for Route #1, #2 and #3. Table 4-6 to Table 4-8 shows the average dwelling time and the standard deviation from the average dwelling time at stops.

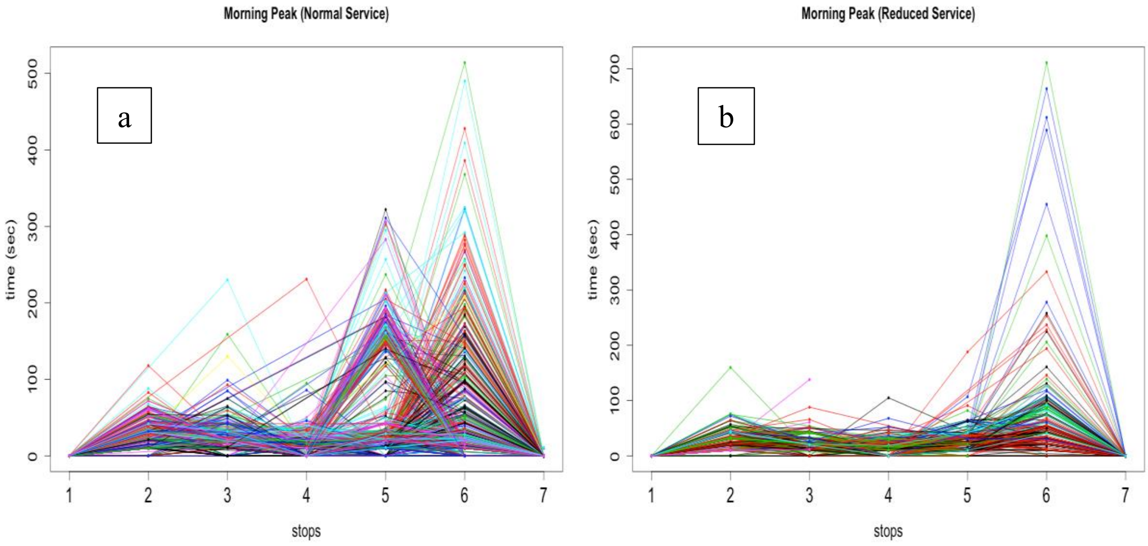


Figure 4-16: Shuttle bus dwelling time at stops on Route #1 during morning peak hour for normal service (a) and reduced service (b)

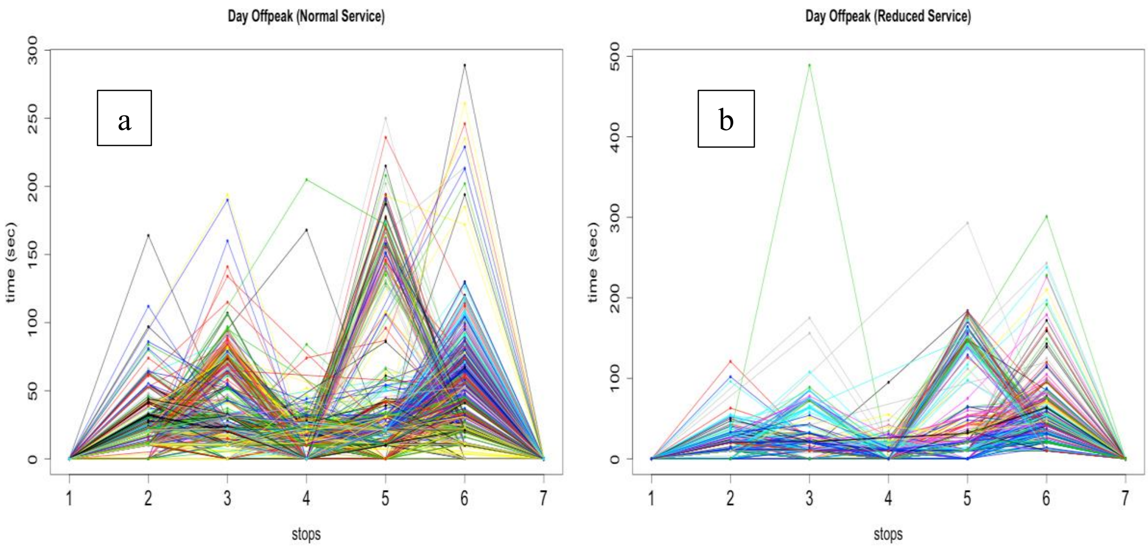


Figure 4-17**:** Shuttle bus dwelling time at stops on Route #1 during day off-peak hour for normal service (a) and reduced service (b)

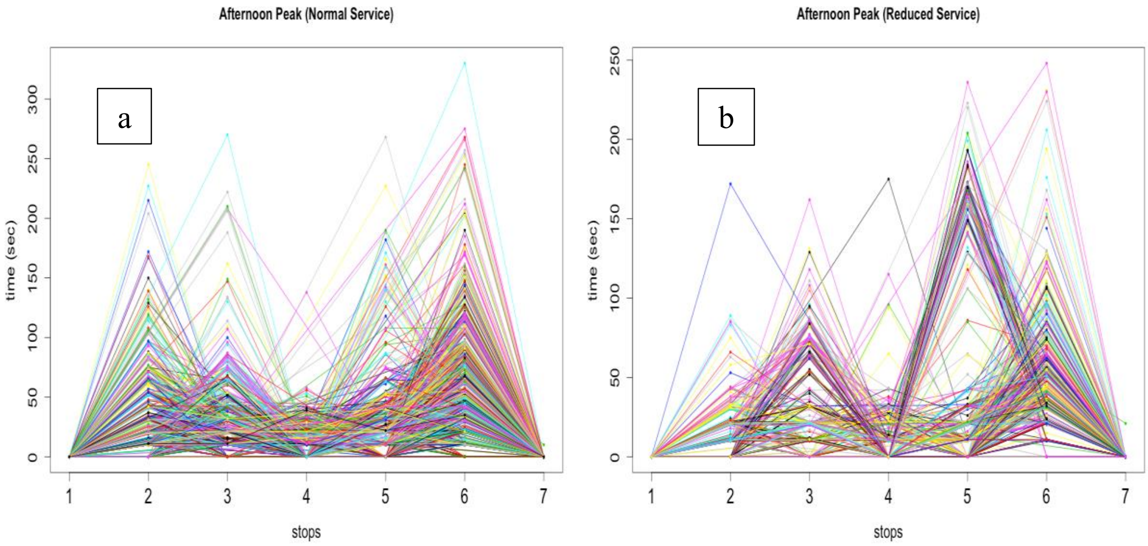


Figure 4-18: Shuttle bus dwelling time at stops on Route #1 during afternoon peak hour for normal service (a) and reduced service (b)

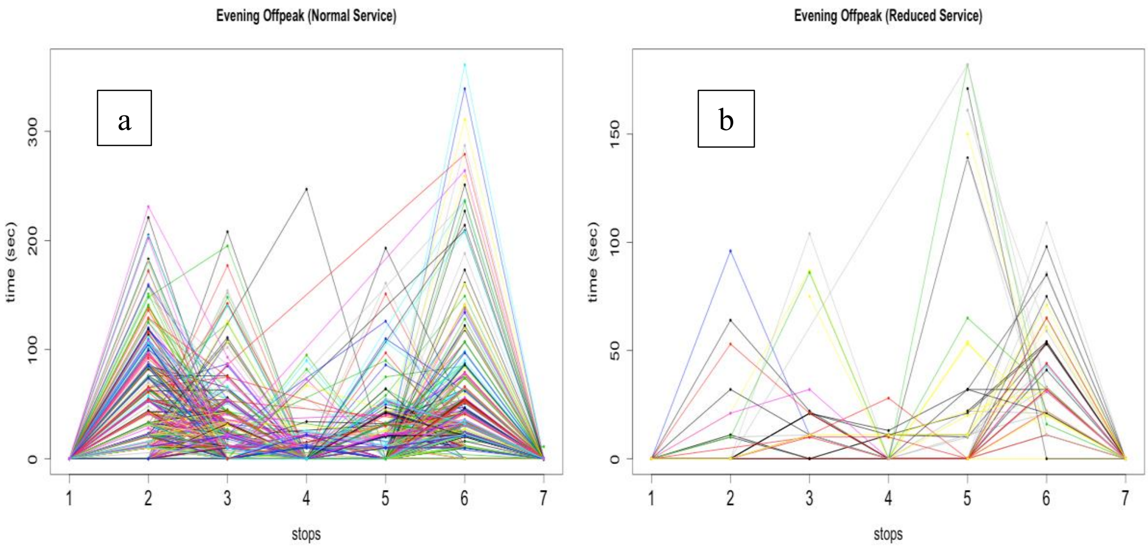


Figure 4-19 Shuttle bus dwelling time at stops on Route #1 during evening off-peak hour for normal service (a) and reduced service (b)

Table 4-6 Average Shuttle bus dwelling time and standard deviation from average dwelling time at stops for normal and reduced service on Route #1 during peak and off-peak hours

	Mean (sec)	Standard Deviation (sec)	Mean (sec)	Standard Deviation (sec)
	Normal Service		Reduced Service	

	Morning Peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Kids Uni	28.72	13.26	29.08	15.46
3	Science Bld	22.89	14.57	23.31	15.41
4	Creative Arts Bld	18.89	12.22	19.53	11.64
5	Hope Theatre	45.10	59.27	26.75	17.78
6	Northfield Ave	52.97	69.79	70.18	92.09
7	Arrive Train Station	0.00	0.00	0.00	0.00
	Daytime Off-peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Kids Uni	21.61	12.93	21.67	17.19
3	Science Bld	31.81	26.58	28.17	36.94
4	Creative Arts Bld	16.91	13.00	15.70	10.19
5	Hope Theatre	33.67	42.80	46.07	53.49
6	Northfield Ave	44.35	31.74	53.09	38.35
7	Arrive Train Station	0.00	0.00	0.00	0.00
	Afternoon Peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Kids Uni	27.52	25.61	19.34	16.54
3	Science Bld	26.55	23.44	33.54	28.21
4	Creative Arts Bld	16.84	9.57	21.90	24.95
5	Hope Theatre	26.49	24.47	54.45	63.45
6	Northfield Ave	52.70	38.76	52.24	33.89
7	Arrive Train Station	0.00	0.00	0.00	0.00
	Evening Off-peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Kids Uni	47.32	45.11	30.82	27.17
3	Science Bld	29.88	30.17	31.42	30.16
4	Creative Arts Bld	25.61	38.05	13.13	5.69
5	Hope Theatre	25.82	22.96	67.64	64.36
6	Northfield Ave	38.14	40.92	40.62	22.69
7	Arrive Train Station	0.00	0.00	0.00	0.00

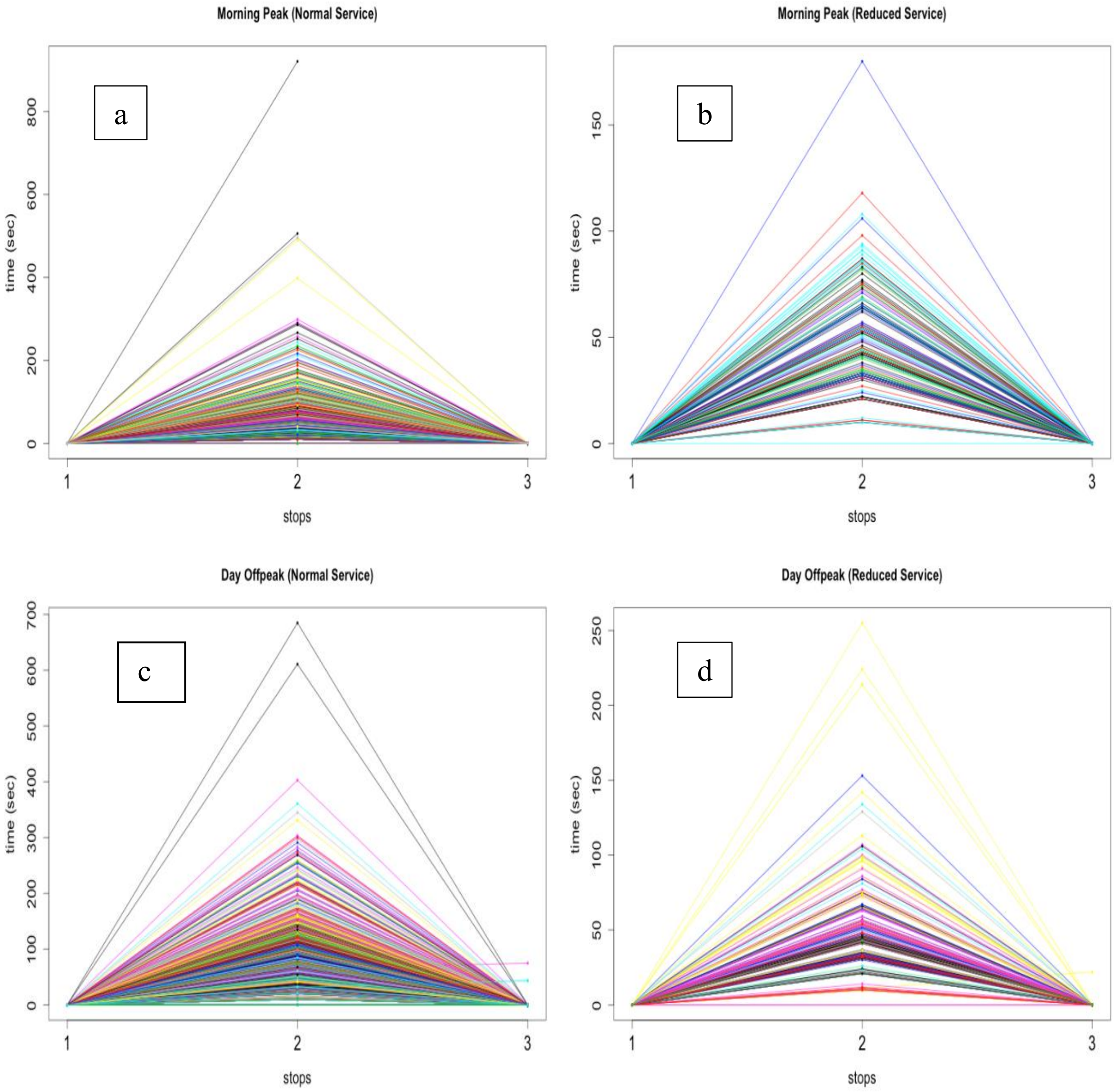


Figure 4-20: Shuttle bus dwelling time at stops on Route #2 during morning peak hour for normal service (a) and reduced service (b), and day time hour for normal service (c) and reduced service (d)

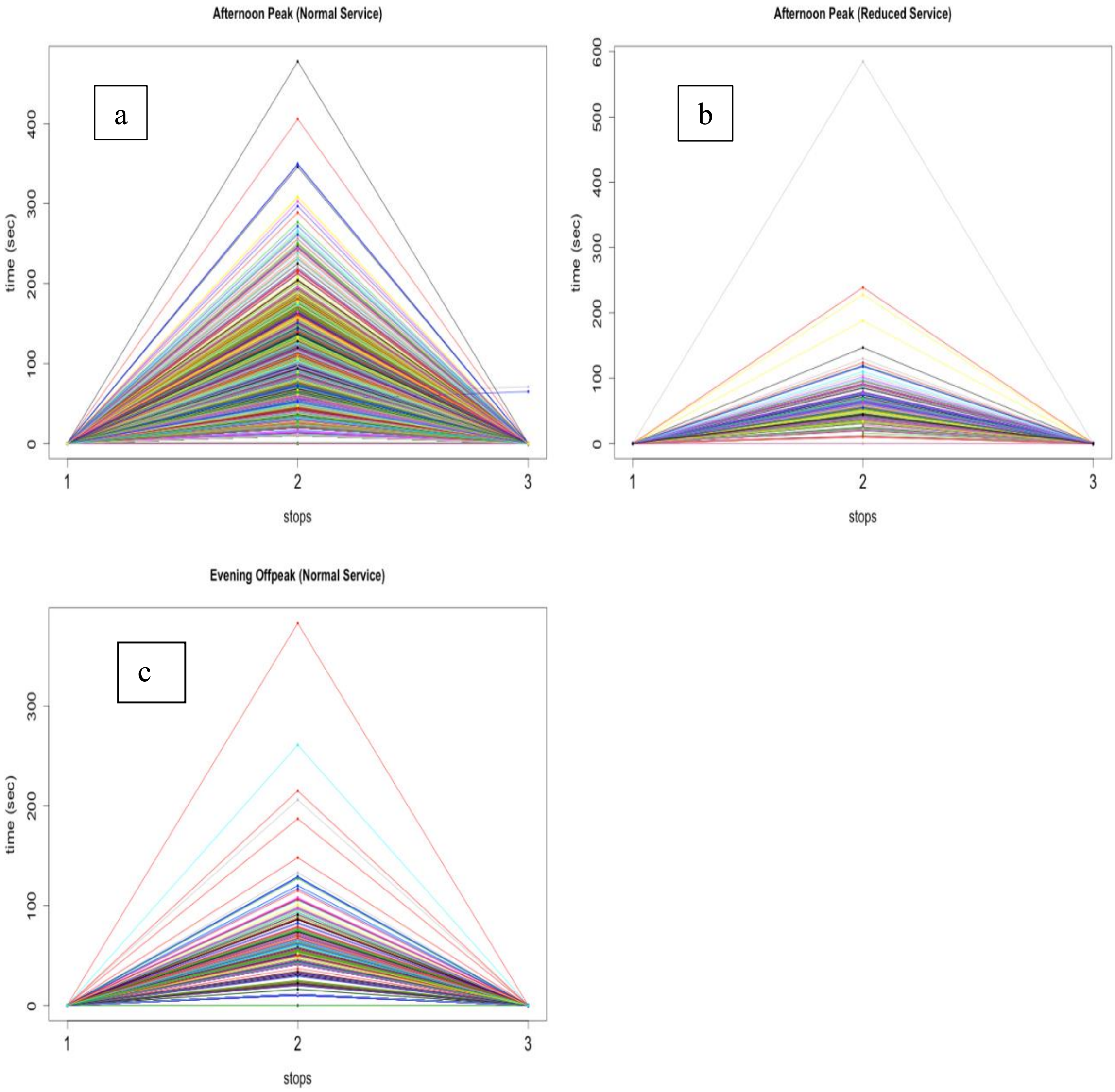


Figure 4-21: Shuttle bus dwelling time at stops on Route #2 during afternoon peak hour for normal service (a) and reduced service (b), and evening off-peak hour for normal service (c)

Table 4-7: Average shuttle bus dwelling time and standard deviation from average dwelling time for normal and reduced service on Route #2 during peak and off-peak hours

		Mean (sec)	Standard Deviation (sec)	Mean (sec)	Standard Deviation (sec)
		Normal Service		Reduced Service	
	Morning Peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Northfield Ave	55.28	45.09	46.69	21.68

3	Arrive Train Station	0.00	0.00	0.00	0.00
	Daytime Off-peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Northfield Ave	72.40	45.75	50.24	44.70
3	Arrive Train Station	0.00	0.00	0.00	0.00
	Afternoon Peak				
1	Depart Train Station	0.00	0.00	0.00	0.00
2	Northfield Ave	76.02	47.20	45.49	31.82
3	Arrive Train Station	0.04	1.75	0.00	0.00
	Evening Off-peak				
1	Depart Train Station	0.00	0.00		
2	Northfield Ave	43.66	30.24		
3	Arrive Train Station	0.00	0.00		

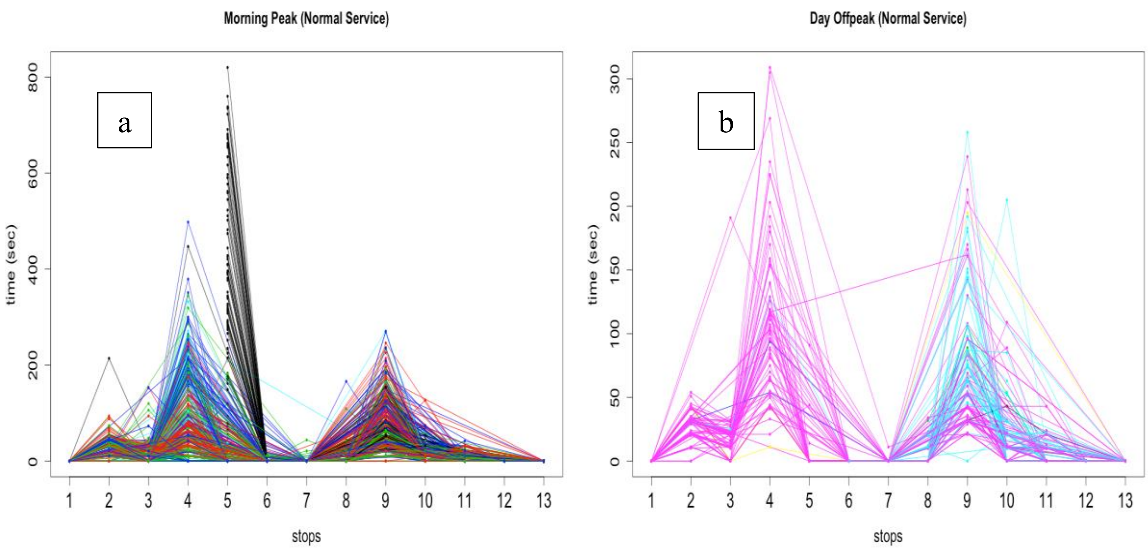


Figure 4-22: Shuttle bus dwelling time at stops on Route #3 for normal service during morning peak hour (a) and day-off peak hour (b)

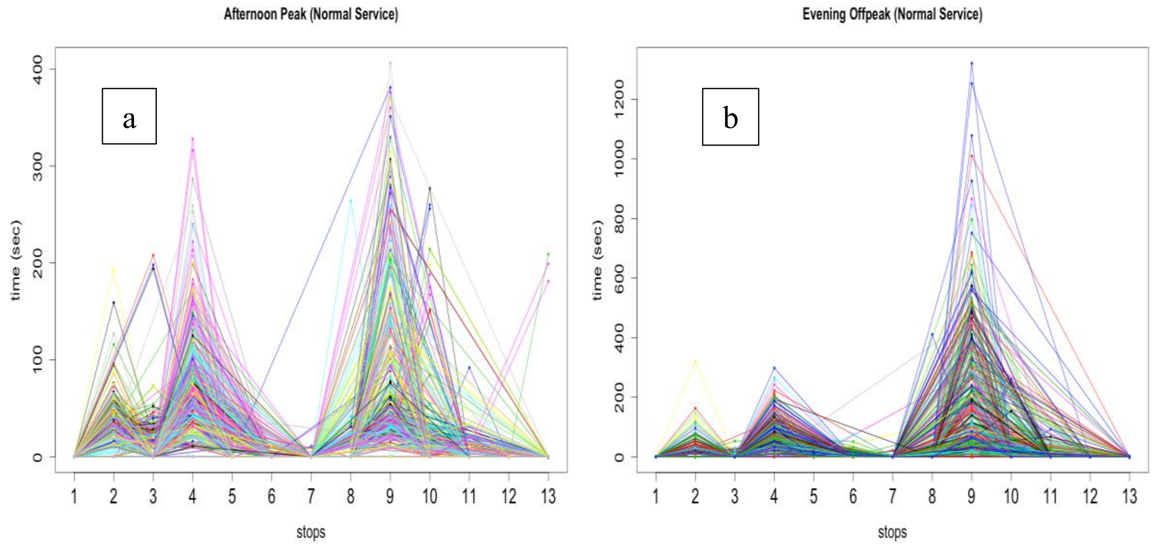


Figure 4-23: Shuttle bus dwelling time at stops on Route #3 for normal service during afternoon peak hour (a) and evening off-peak hour (b)

Table 4-8: Average Shuttle bus dwelling time and standard deviation from average dwelling time at stops for normal bus service on Route #3 during peak and off-peak hours

		Mean (sec)	Standard Deviation (sec)
	Morning Peak		
1	Depart Kids Uni	0.00	0.00
2	Science Bld	27.46	15.15
3	Creative Arts Bld	19.87	15.45
4	Hope Theatre	81.90	76.02
5	Robsons Rd	304.07	231.93
6	Reserve St	11.00	0.82
7	Greenacre Rd	32.50	11.50
8	Weerona-Throsby Dr	18.50	16.26
9	Keiraview	62.12	47.20
10	Throsby Dr	19.00	11.45
11	Wiseman Park	15.59	6.21
12	Eastern St	0.00	0.00
13	Arrive Kids Uni	0.00	0.00
	Daytime Off-peak		
1	Depart Kids Uni	0.00	0.00
2	Science Bld	29.74	9.42
3	Creative Arts Bld	23.45	24.15
4	Hope Theatre	110.39	64.09
5	Robsons Rd	49.67	31.69
6	Reserve St	0.00	0.00

7	Greenacre Rd	11.00	0.00
8	Weerona-Throsby Dr	15.33	6.83
9	Keiraview	77.35	50.12
10	Throsby Dr	25.71	29.99
11	Wiseman Park	15.18	7.07
12	Eastern St	0.00	0.00
13	Arrive Kids Uni	0.00	0.00
	Afternoon Peak		
1	Depart Kids Uni	0.00	0.00
2	Science Bld	29.25	16.51
3	Creative Arts Bld	19.89	19.70
4	Hope Theatre	56.03	39.92
5	Robsons Rd	14.23	6.85
6	Reserve St	11.60	1.36
7	Greenacre Rd	10.50	0.50
8	Weerona-Throsby Dr	15.04	14.28
9	Keiraview	54.06	62.14
10	Throsby Dr	41.05	65.68
11	Wiseman Park	16.27	10.30
12	Eastern St	0.00	0.00
13	Arrive Kids Uni	0.00	0.00
	Evening Off-peak		
1	Depart Kids Uni	0.00	0.00
2	Science Bld	26.43	24.82
3	Creative Arts Bld	14.61	6.91
4	Hope Theatre	56.20	46.06
5	Robsons Rd	15.24	9.07
6	Reserve St	19.53	16.75
7	Greenacre Rd	21.00	0.00
8	Weerona-Throsby Dr	18.27	41.80
9	Keiraview	151.80	190.77
10	Throsby Dr	42.78	71.72
11	Wiseman Park	20.00	16.00
12	Eastern St	0.00	0.00
13	Arrive Kids Uni	0.00	0.00

The above dwelling time analysis graphs show that dwelling time can significantly vary from stop to stop and at different times of a day and on different days of the year. It is discernible that dwelling time is largely impacted by the passengers' flow pattern. The dwelling times jump noticeably at stop 6 (Northfields Ave) and stop 5 (Hope Theatre stop) for Route #1 (Figure 4-16 -Figure 4-19), along with the big increase in dwelling time at stop 4 (Northfields Ave) and stop 9 (the Keiraview stop) for Route #3 (Figure

4-22 -Figure 4-23). These increases are apparently caused directly by the fact that a large number of university students often get on and off shuttle bus at these stops.

It is also noted that no dwelling time occurred at some stops; this is because when no passengers are waiting at a stop, a bus passes by a stop without stopping.

The data analyses and presentation in this section allow the identification of the key factors (those are different routes, different times of day, and different days of a year) that can result in different bus dwelling time patterns. The bus dwelling time, along with these factors, should be used for the travel time prediction for the corresponding route and time period.

4.4.3 Shuttle Bus Schedule Adherence at Different Routes and Time

Schedule adherence is defined in this research as how adherent (or close) a shuttle bus is to its fixed timetable prescribed by the transport operator. Shuttle bus drivers try to adhere to the prescribed schedule; they may stay at a particular stop longer or skip a stop in order to keep its schedule close to the pre-defined one. Schedule adherence is measured by the difference between the prescribed arrival time and the actual arrival time. A positive value means the bus arrives at a stop behind schedule, while a negative value indicates the bus is ahead of schedule.

In order to decide whether the schedule adherence factor should be an important one in bus arrival time prediction analyses, similar to the previous analyses on bus travelling time and bus dwelling time, were conducted. Figure 4-24 to Figure 4-26 show the patterns generated. Further, the average and the standard deviation are also listed in Table 4-9 to Table 4-11.

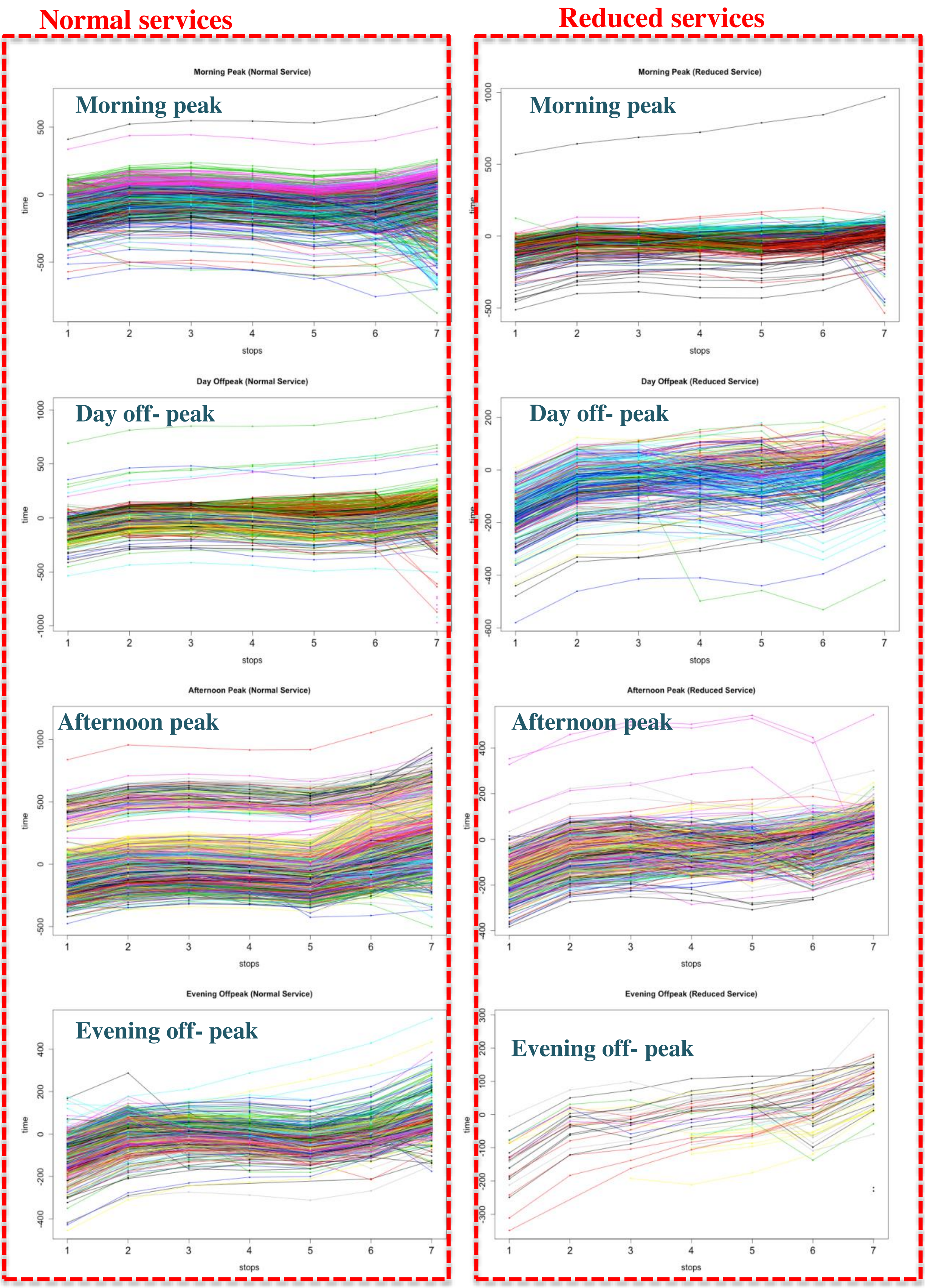


Figure 4-24: Schedule adherence during peak hour and off-peak hours on Route #1 for normal service (left-hand graphs) and reduced service (right-hand graphs)

Table 4-9: Average schedule adherence and standard deviation from average schedule adherence during peak and off-peak hours on Route #1 for normal and reduced service

		Mean (sec)	Standard Deviation (sec)	Mean (sec)	Standard Deviation (sec)
		Normal Service		Reduced Service	
	Morning Peak				
1	Depart Train Station	-82.08	95.40	-100.14	87.00
2	Kids Uni	-7.99	98.57	-28.27	81.37
3	Science Bld	-2.00	97.63	-25.48	76.60
4	Creative Arts Bld	-30.43	97.73	-27.85	82.62
5	Hope Theatre	-73.69	100.88	-35.44	93.01
6	Northfield Ave	-64.18	114.80	-9.43	93.42
7	Arrive Train Station	-39.61	178.87	19.49	115.89
	Daytime Off-peak				
1	Depart Train Station	-81.50	76.78	-153.32	86.05
2	Kids Uni	14.89	76.97	-47.66	83.71
3	Science Bld	27.08	76.73	-36.95	78.92
4	Creative Arts Bld	1.49	83.32	-24.80	79.37
5	Hope Theatre	-28.23	96.87	-25.25	87.96
6	Northfield Ave	-9.24	102.75	-18.30	84.20
7	Arrive Train Station	46.67	140.30	44.37	72.94
	Afternoon Peak				
1	Depart Train Station	-41.76	190.93	-154.68	81.62
2	Kids Uni	37.76	190.93	-55.39	78.79
3	Science Bld	42.70	189.95	-34.61	81.53
4	Creative Arts Bld	8.84	186.52	-33.00	71.79
5	Hope Theatre	-26.04	186.56	-32.72	74.03
6	Northfield Ave	88.23	199.53	-9.92	79.58
7	Arrive Train Station	195.06	215.46	50.51	70.30
	Evening Off-peak				
1	Depart Train Station	-83.25	74.45	-150.76	77.52
2	Kids Uni	12.81	62.98	-31.71	57.68
3	Science Bld	9.57	48.61	-33.00	67.07
4	Creative Arts Bld	-10.87	49.23	-22.27	68.61
5	Hope Theatre	-39.39	49.52	6.51	53.82
6	Northfield Ave	-0.87	52.95	13.26	65.29
7	Arrive Train Station	77.24	66.64	79.37	81.14

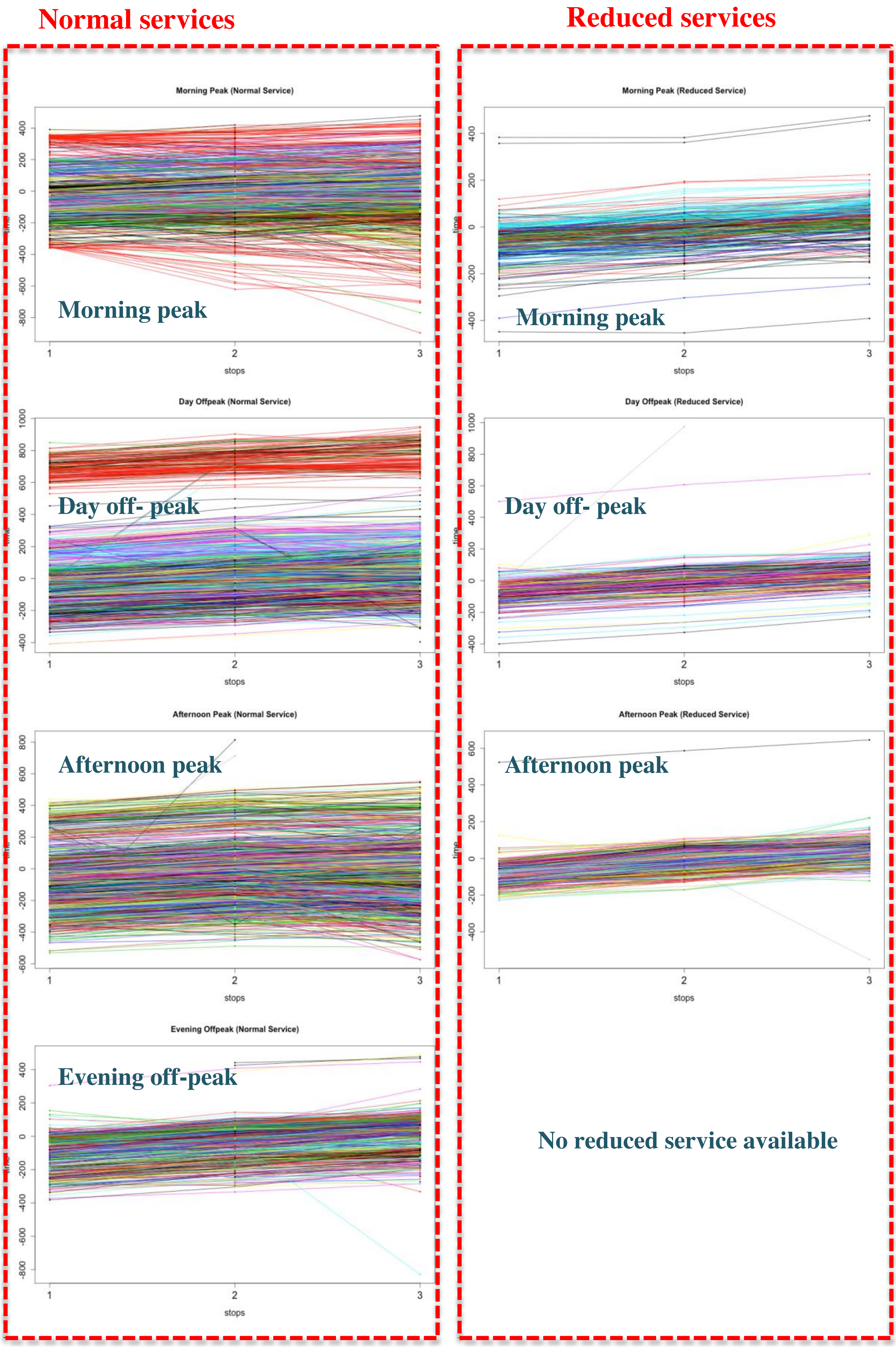


Figure 4-25: Schedule adherence during peak and off-peak hours on Route #2 for normal and reduced service

Table 4-10: Average schedule adherence and standard deviation from average schedule adherence during peak and off-peak hours on Route #2 for normal and reduced services

		Mean (sec)	Standard Deviation (sec)	Mean (sec)	Standard Deviation (sec)
		Normal Service		**Reduced Service**	
	Morning Peak				
1	Depart Train Station	-51.98	105.56	-58.95	83.09
2	Northfield Ave	-21.72	112.86	-13.07	85.11
3	Arrive Train Station	5.15	131.18	34.50	83.43
	Daytime Off-peak				
1	Depart Train Station	-8.73	181.04	-62.30	68.62
2	Northfield Ave	54.58	187.34	4.20	96.57
3	Arrive Train Station	76.46	188.78	59.41	74.82
	Afternoon Peak				
1	Depart Train Station	-60.34	126.85	-75.14	51.11
2	Northfield Ave	-8.27	128.28	-8.22	53.67
3	Arrive Train Station	-5.52	139.22	43.16	58.61
	Evening Off-peak				
1	Depart Train Station	-81.01	95.04		
2	Northfield Ave	-17.09	99.83		
3	Arrive Train Station	15.53	112.69	No reduced service available	

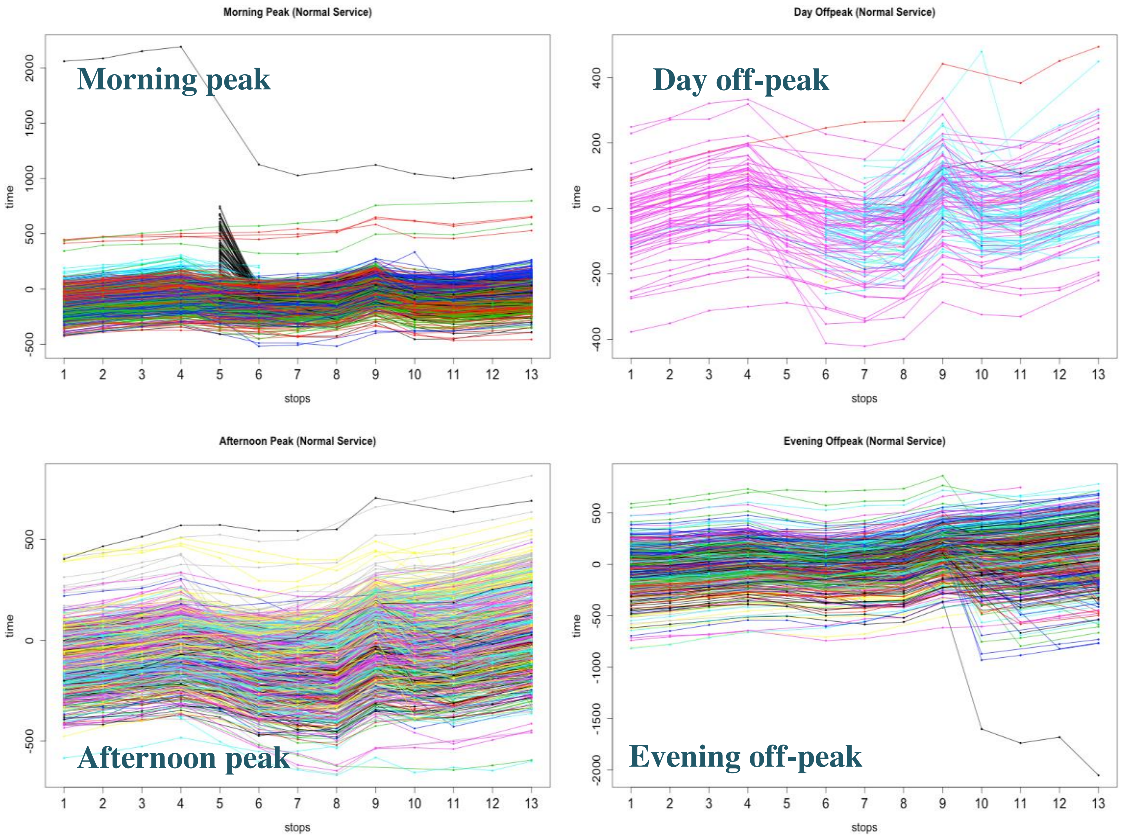


Figure 4-26: The schedule adherence during peak and off-peak hours on Route #3 for normal bus service

Table 4-11: Average schedule adherence and standard deviation from average schedule adherence during peak and off-peak hours on Route #3 for normal service

		Mean (sec)	Standard Deviation (sec)
	Morning Peak		
1	Depart Kids Uni	-105.13	151.19
2	Science Bld	-69.14	151.42
3	Creative Arts Bld	-41.46	155.68
4	Hope Theatre	-15.86	160.46
5	Robsons Rd	19.04	221.69
6	Reserve St	-90.30	122.09
7	Greenacre Rd	-105.33	119.64
8	Weerona-Throsby Dr	-114.70	116.56
9	Keiraview	19.72	124.17
10	Throsby Dr	-73.57	120.16
11	Wiseman Park	-76.62	122.88
12	Eastern St	-53.78	106.65
13	Arrive Kids Uni	-15.77	130.81
	Daytime Off-peak		

1	Depart Kids Uni	-45.39	109.14
2	Science Bld	-2.77	106.09
3	Creative Arts Bld	26.90	109.48
4	Hope Theatre	45.03	116.21
5	Robsons Rd	-47.42	101.54
6	Reserve St	-76.26	103.36
7	Greenacre Rd	-72.76	107.27
8	Weerona-Throsby Dr	-82.79	106.87
9	Keiraview	53.29	112.78
10	Throsby Dr	-22.88	108.53
11	Wiseman Park	-19.52	107.44
12	Eastern St	22.62	125.91
13	Arrive Kids Uni	75.39	115.62
	Afternoon Peak		
1	Depart Kids Uni	-86.61	130.29
2	Science Bld	-64.23	131.06
3	Creative Arts Bld	-42.63	133.30
4	Hope Theatre	-17.97	136.72
5	Robsons Rd	-48.36	126.45
6	Reserve St	-113.02	143.21
7	Greenacre Rd	-117.99	154.99
8	Weerona-Throsby Dr	-135.93	165.52
9	Keiraview	-3.54	167.03
10	Throsby Dr	-54.43	166.24
11	Wiseman Park	-41.78	173.22
12	Eastern St	9.34	181.31
13	Arrive Kids Uni	37.10	186.49
	Evening Off-peak		
1	Depart Kids Uni	-61.16	193.72
2	Science Bld	-39.45	186.58
3	Creative Arts Bld	-5.76	186.36
4	Hope Theatre	27.33	186.82
5	Robsons Rd	12.98	150.12
6	Reserve St	-30.32	176.59
7	Greenacre Rd	-16.50	184.80
8	Weerona-Throsby Dr	-16.96	190.00
9	Keiraview	135.05	192.21
10	Throsby Dr	23.48	269.70
11	Wiseman Park	40.03	263.24
12	Eastern St	95.55	259.59
13	Arrive Kids Uni	117.06	269.32

In previous sections three variables that affect the travel time prediction including bus arrival time, dwell time at stop and schedule adherences were analysed. However, these three variables are not independent of each other but are rather closely inter-related

because the schedule adherence as a variable is the consequence of the other two variables. In designing and formulating the historical model the two main variables: bus arrival time and dwell time are used for prediction of arrival time. In this approach, the variable of schedule adherence has been incorporated in these two main variables. Its contribution to the prediction formula is implicitly included through the other two variables.

4.5 The Prediction Model

As described in the literature review (in Chapter 2), there are several different techniques for travel time prediction. In this research, the historical data based model has been chosen. The key reason for this choice is because the Gong Shuttle service operates in the mid-sized city of Wollongong, where the city's traffic conditions have remained stable over the years. Further, the travelling time and dwelling time of shuttle buses are the two identified key impacting factors for arrival time prediction, and both have exhibited their relative stableness in their patterns. Thus, the historical data based prediction model is a suitable choice (Chien & Kuchipudi, 2003, Williams and Hoel, 2003) and sufficient for the purpose of the present work.

The historical data based model developed in this study is described as follows. The model calculates the predicted shuttle bus arrival time according to equation Eq.4.3 and Eq. 4.4. As illustrated in Figure 4-27, assume shuttle bus B's current position (p) is between bus stop 1 (s1) and stop 2 (s2). To predict the bus's arrival time at stop 3 (s3), the predicted arrival time should be the sum of the travelling time from position p to stop 2, $T_{p,s2}(B)$, the dwelling time at stop 2, $T_{s2,d}(B)$, and the B's travelling time to stop 3, $T_{s2,s3}(B)$. The equation is then expressed in Eq. 4.3:

$$T_{p,s3}(B) = T_{p,s2}(B) + T_{s2,d}(B) + T_{s2,s3}(B) \qquad \text{Eq. 4.3}$$

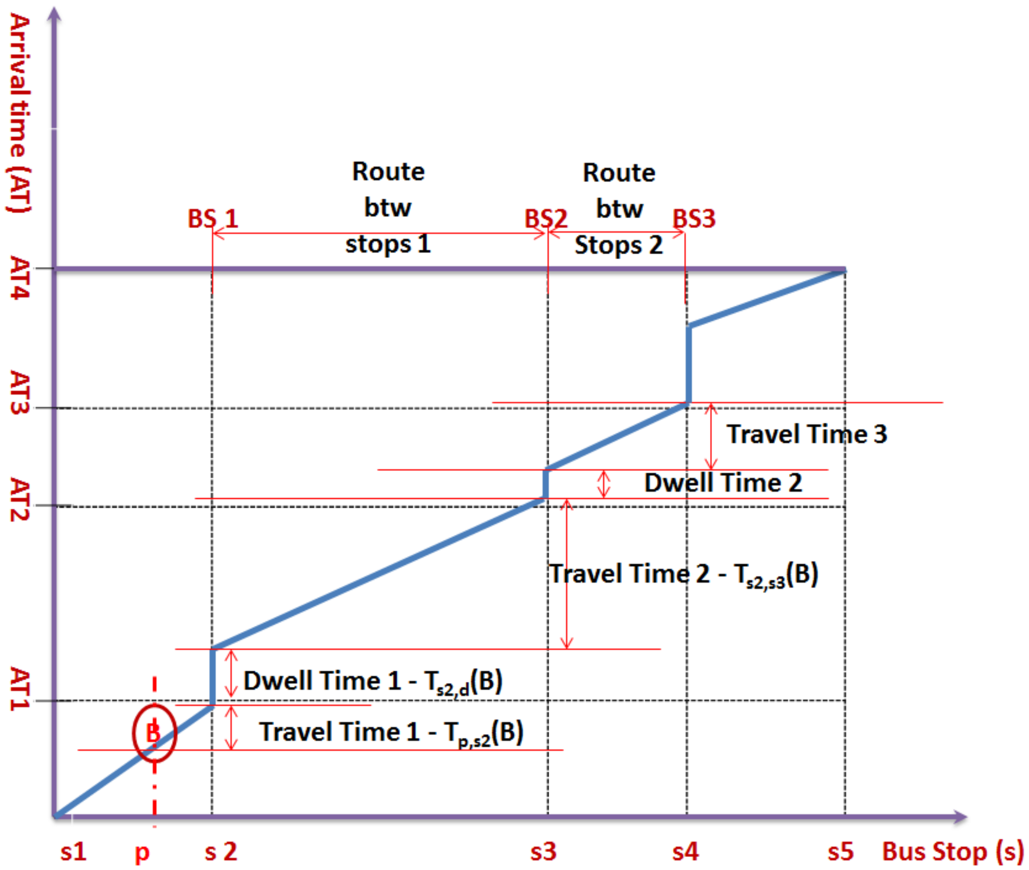


Figure 4-27: Bus arrival time, travelling time and dwelling time

To generalise shuttle bus (B)'s arrival time at a stop, for its current position, p, which is between stop i and stop i+1 to a stop j, the forecasted travel time at stop j $T_{p\text{-}j}(B)$ can be calculated with the following equation:

$$T_{p,j}(B) = T_{p,i+1}(B) + \sum_{i+1}^{j-1} T_{i+1,d}\,(B) + \sum_{i+1}^{j} T_{i+1,j}\,(B) \qquad \text{Eq. 4.4}$$

Where,

$T_{p,j}(B)$ is the predicted travelling time from the shuttle bus B's current position, p (which is between stop i and i+1), to stop j;

$T_{p,i+1}(B)$ is the travelling time from shuttle B's current position p to next stop i+1;

$\sum_{i+1}^{j-1} T_{i+1,d}\,(B)$ is the sum of the shuttle bus B's dwelling time from stop i+1 to stop j;

$\sum_{i+1}^{j} T_{i+1,j}\,(B)$ is the sum of the shuttle bus B's travelling time from stop i+1 to stop j;

For $T_{p,i+1}(B)$ and $T_{i+1,j}(B)$, as routes are divided into segments, travelling time is calculated by summing up the distance of each segment to travel divided by the bus's velocity on that segment.

The historical data sets are split into time of day and day of year to handle peak and off-peak times during the day and also session times and holiday periods. The travel time between the bus's current position and the next stop use real time information to give a more accurate prediction.

For comparison purposes, a simple linear regression model from R package (an open source software and provides a variety of statistical analysis techniques) is also used to process the data. The linear regression model uses distances travelled as the input variable and arrival time as output. The time a bus arrives at bus stop j is predicted by Eq. 4.5.

$$T_j = \beta_0 + \beta_1 \times D_j \qquad \text{Eq. 4.5}$$

Where,

β_1: the coefficient of the linear regression, it represents the inverse of the average speed between the starting stop to stop j;

β_0: the intercept of the linear regression, it represents the time at the starting stop which is ideally supposed to be zero otherwise is noise or error as result of regression process;

D_j: distance from the starting stop to stop j;

T_j: travel time to stop j.

The regression approach is similar to historical approach which can perform well under stable traffic patterns over a long period of time and have periodic patterns. This means both historical and regression techniques are applicable and reliable in small cities like Wollongong. While Kalman filter, ANN techniques and hybrid techniques may have advantages over the first two approaches for prediction under more complicated conditions such as serious traffic congestions, they often require lengthy training and learning processes, are more difficult to use, require greater computational resources and have other methodological issues to be resolved.

4.6 Prediction Model Evaluation

To measure the accuracy of the prediction model, the Root Mean Square Error (RMSE) measurement is used. RMSE is a standard statistical measurement for model performance evaluation. Its calculation is shown in Eq. 4.6.

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(y_{jm} - y_{jp}\right)^2} \qquad \text{Eq. 4.6}$$

Where,

y_{jm}: measured value at time j ;

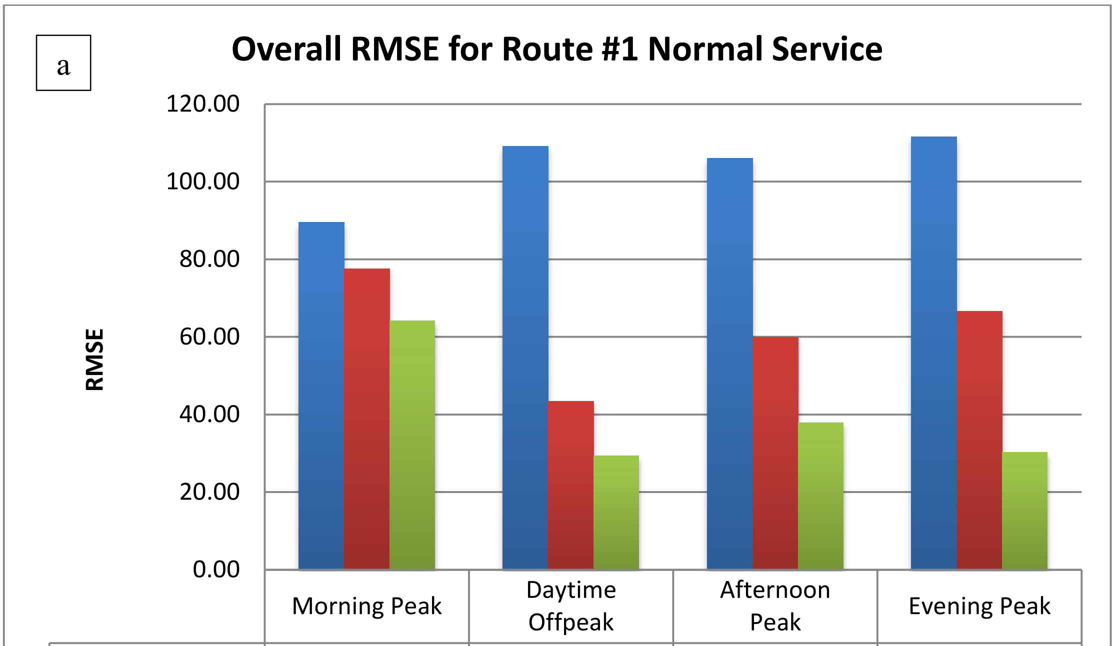
y_{jp}: predicted value at time j.

The evaluation has been conducted for three scenarios in prediction of shuttle bus travel times: bus timetable, simple linear regression model and historical data based model. The evaluation RMSE results calculated using Eq. 4.6 for the three cases are displayed in the Table 4-12, Table 4-13 and Table 4-14 for Route #1, #2 and #3. Figure 4-28, Figure 4-29 and Figure 4-30 show the graphical comparison for the three approaches. Included in these Figures are the tables for comparison on the average magnitude of

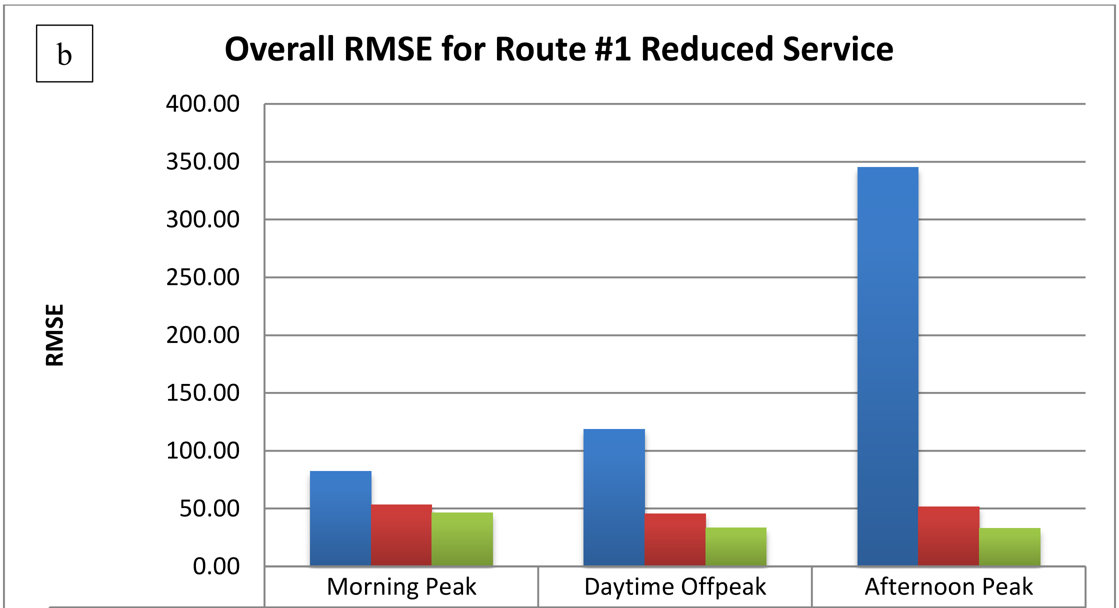
forecaste errors from regression model and historical data based model with the bus timetable.

Table 4-12: RMSE comparison for Route #1

	Timetable	Regression	Historical Model		Timetable	Regression	Historical Model
	Normal				**Reduced**		
Stops	**Morning Peak**						
1	106.47	0.00	0.00		113.59	0.00	0.00
2	75.46	40.26	40.02		70.04	35.89	34.52
3	72.68	42.97	42.26		61.77	39.94	37.88
4	74.12	46.64	45.79		74.66	46.32	44.40
5	96.23	52.98	52.06		100.32	49.40	46.54
6	96.43	73.27	71.27		82.79	52.25	45.03
7	98.81	149.41	106.78		57.71	83.22	64.98
	Daytime Off-peak						
1	133.02	0.00	0.00		180.32	0.00	0.00
2	102.89	23.72	16.00		110.58	18.29	21.50
3	103.60	30.96	20.99		98.79	19.58	23.61
4	85.37	38.34	25.35		104.03	32.40	26.17
5	110.89	39.91	26.68		127.01	36.80	25.64
6	105.16	49.60	34.33		102.34	63.21	42.99
7	117.59	65.44	43.83		79.10	73.63	49.01
	Afternoon Peak						
1	100.03	0.00	0.00		359.65	0.00	0.00
2	65.18	35.00	22.35		262.34	15.70	10.46
3	46.20	46.18	29.52		367.54	21.64	16.13
4	59.04	47.38	30.49		361.53	36.00	23.80
5	75.54	51.34	32.93		364.36	39.01	25.10
6	110.58	62.85	39.88		357.19	66.51	41.79
7	202.25	96.76	60.50		332.16	89.97	56.69
	Evening Off-peak						
1	100.03	0.00	0.00				
2	65.18	34.28	15.33				
3	46.20	54.58	24.90				
4	59.04	62.32	27.83				
5	75.54	68.37	31.10				
6	110.58	75.93	34.18				
7	202.25	90.17	41.44				



	Morning Peak	Daytime Offpeak	Afternoon Peak	Evening Peak
TimeTable	89.54	109.21	106.04	111.55
Regression	77.61	43.45	59.93	66.60
Historical Model	64.13	29.30	37.94	30.23

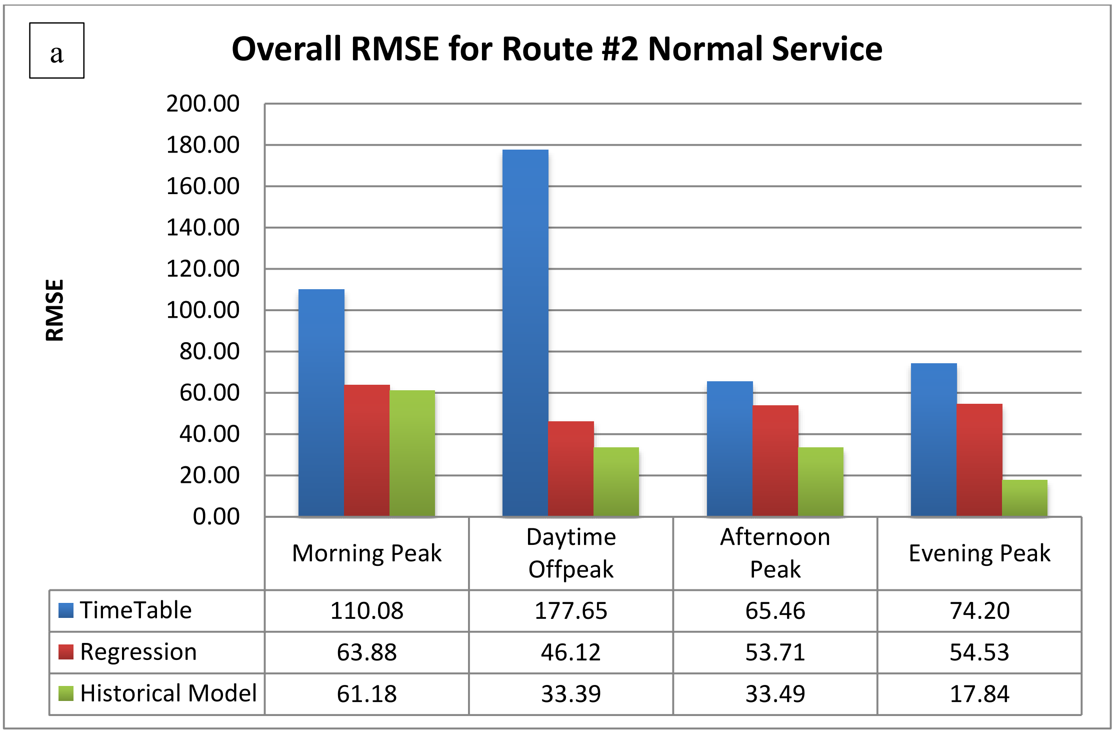


	Morning Peak	Daytime Offpeak	Afternoon Peak
TimeTable	82.33	118.43	345.30
Regression	53.42	45.71	51.72
Historical Model	46.57	33.20	32.98

Figure 4-28**:** Overall graphical RMSE comparison of Route #1 for the normal services (a) and the reduced services (b)

Table 4-13: RMSE comparison for Route #2

	Timetable	Regression	Historical Model		Timetable	Regression	Historical Model
	Normal				**Reduced**		
Stops	**Morning Peak**						
1	112.30	0.00	0.00		110.37	0.00	0.00
2	104.92	50.01	48.46		91.73	37.63	35.07
3	112.86	75.24	71.67		89.64	47.07	52.25
	Daytime Off-peak						
1	168.20	0.00	0.00		92.89	0.00	0.00
2	178.19	37.24	27.40		104.48	38.30	23.86
3	186.09	53.54	38.47		88.23	50.97	31.47
	Afternoon Peak						
1	57.58	0.00	0.00		95.45	0.00	0.00
2	63.97	38.14	23.75		58.04	31.71	21.61
3	73.80	65.69	40.97		74.75	43.04	29.62
	Evening Off-peak						
1	71.06	0.00	0.00				
2	57.03	42.20	13.55				
3	90.63	64.54	21.29				



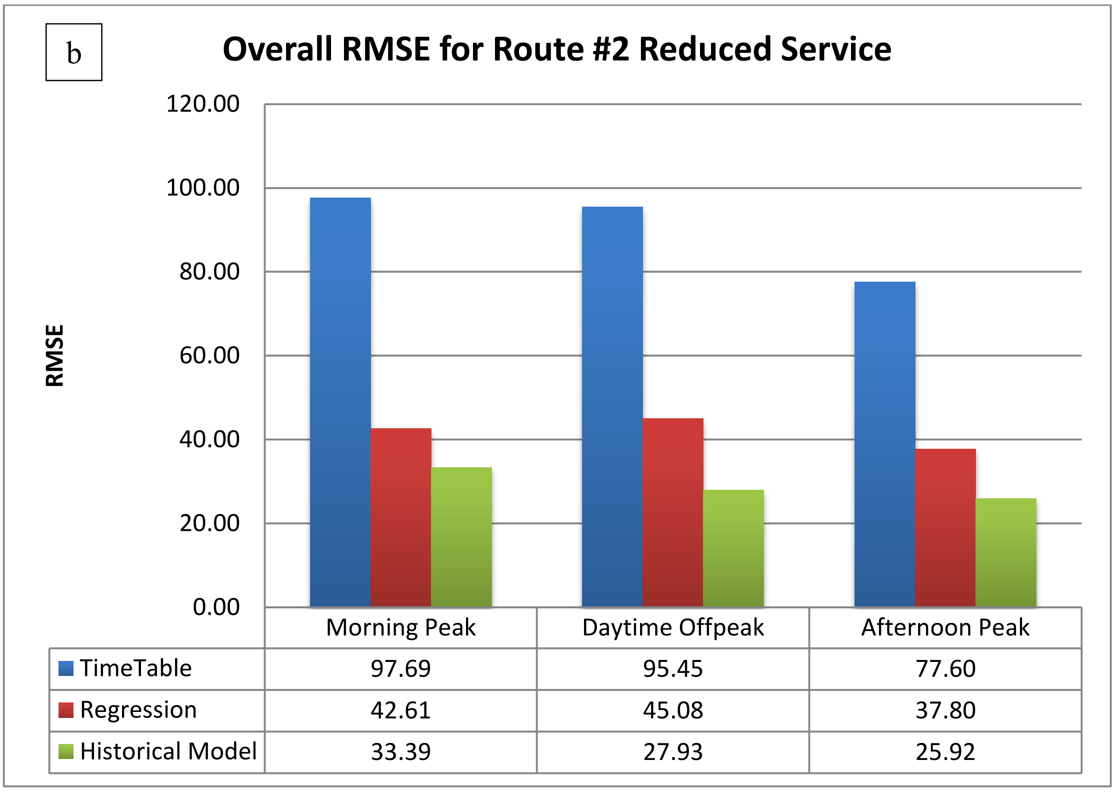


Figure 4-29: Overall graphical RMSE comparison of Route #2 for the normal services (a) and the reduced services (b)

Table 4-14: RMSE comparison for Route #3

	Timetable	Regression	Historical Model		Timetable	Regression	Historical Model
Stops	**Morning Peak**				**Daytime Off-peak**		
1	214.20	0.00	0.00		123.39	0.00	0.00
2	196.57	11.58	9.76		99.32	10.78	3.35
3	191.20	23.83	20.11		97.58	12.28	4.32
4	188.94	38.34	32.58		110.83	35.45	13.64
5	209.74	202.96	93.78		129.55	48.26	12.49
6	182.70	144.57	120.09		140.14	93.55	38.38
7	187.88	148.08	128.99		142.87	94.57	44.83
8	140.59	150.01	97.17		147.11	96.36	38.70
9	169.20	154.12	135.00		134.01	108.43	49.98
10	182.00	169.83	136.60		107.94	90.16	40.34
11	170.93	161.18	135.45		119.25	108.53	48.83
12	134.30	138.68	81.28		138.30	122.32	41.98
13	151.14	165.31	151.26		136.69	121.32	61.89
	Afternoon Peak				**Evening Off-peak**		
1	156.75	0.00	0.00		181.04	0.00	0.00
2	148.62	18.37	12.24		180.95	22.87	9.35
3	145.05	32.55	21.39		193.20	39.22	17.15
4	146.28	41.33	27.86		210.23	52.78	23.00
5	139.93	58.86	27.09		161.52	76.61	22.13

6	188.36	69.92	43.88		175.33	91.87	33.80
7	200.88	85.57	57.95		190.55	95.99	40.23
8	219.22	94.96	55.66		190.31	103.80	33.99
9	177.39	104.71	70.97		290.35	105.59	45.36
10	178.46	118.06	56.33		269.29	156.91	43.80
11	192.71	127.76	72.31		279.90	158.77	59.93
12	207.68	138.30	59.10		301.97	157.90	46.78
13	214.04	143.79	90.42		338.34	181.97	74.87

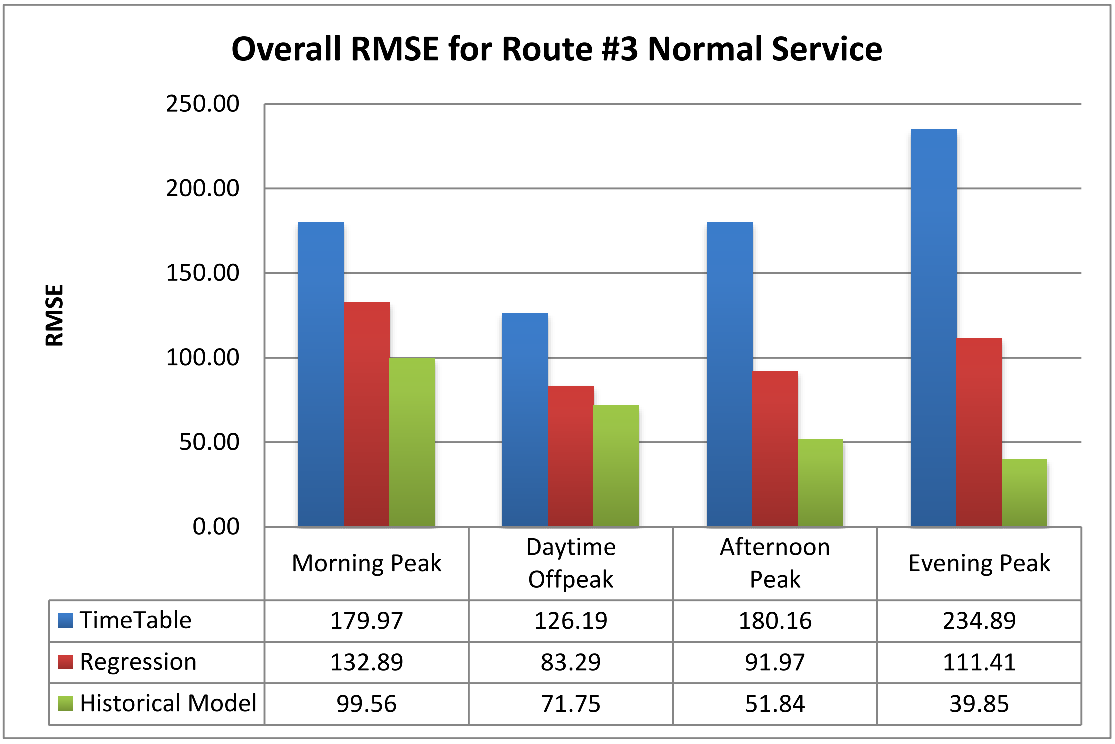


Figure 4-30: Overall graphical RMSE comparison of Route #3

From the above tables and figures, the comparison results clearly show that both the regression model and the historical data based model give much better predictions than using the bus timetable. From Figure 4-28 to Figure 4-30, the average errors from using the predictions based on bus timetable is about 3 times (range from1.4 to 5.9 times) larger than those predicted based on the historical data based model, and is about 2 times (range from 1.2 to 2.5 times) larger than those predicted using the regression model. In some special cases where the shuttle bus departure time was substantially

delayed, the errors incurred by using timetable were quite large as shown in Figure 4-28 for the afternoon peak hour in reduced services. The bus timetable is the only available means for travellers to rely on for their travel planning without any ITS input at work. As shown in Table 4-9 to Table 4-11 and Figure 4-24 to Figure 4-26 the buses adherence to pre-scheduled timetable is widely spread. It is not surprising that the average magnitude of the forecasted errors using timetable is much larger than model based predictions as the model approaches take advantage of the available historical data in various time categories.

It is also evident that the prediction accuracy from the historical data based model has consistently outperformed those using the simple regression model in all conditions including peak hour, off-peak hour under both normal and reduced services. This is because, under relatively stable traffic conditions, the historical data based model has taken into account bus travelling time and dwelling time to improve the prediction accuracy, while the simple regression model has not incorporated the dwelling time.

4.7 Summary

Three shuttle bus routes were selected as the test platform for this study. Passengers' travel patterns were established during 2012. The patterns quantify the degree to which the passenger flow fluctuates from early morning to late evening during a day, which lead to the identification of peak and off-peak hours. The monthly passengers' travel patterns were used to define the normal service and reduced service for different days of a year.

The shuttle travel time, dwelling time and schedule adherence were obtained and calculated from GPS data, with the input from the passenger flow patterns analyses.

These data were grouped into 8 clusters: morning peak hour, day off-peak hour, afternoon peak hour and evening off-peak hours for normal service (school days) and reduced service (school holidays) for Route #1 and #2, and normal service for Route #3. These data show some clear common features and some irregularity. With large amounts of data entries gathered during the research, it is evident that the acquired datasets have high reproducibility and reliability in the transport network.

The historical datasets lay the foundation for choosing a statistical model for shuttle bus arrival time prediction. Based on the abundant data availability and patterns revealed, a historical data based model is developed by incorporating variables of travelling time in each segment and dwelling time at each bus stop along the three routes. The prediction accuracy from the linear regression model and historical data based model is evaluated using the RMSE method in comparison with the pre-scheduled bus timetable. The evaluation results show clearly that the model based predictions of travel time from both linear regression and historical data based models show a significant advantage over using the timetable. Furthermore, the prediction accuracy from the historical data based model consistently outperforms those using the regression model in all conditions. This is because, with a relatively stable traffic condition, the historical data based model has taken into account the travelling and dwelling time factors to improve the prediction accuracy while the simple linear regression model has not.

CHAPTER 5. UNISHUTTLE SYSTEM EVALUATION STUDY

The UniShuttle system evaluation had two main purposes. The first purpose is to assess the system's usability and usefulness through users' feedback on the UniShuttle app. The second purpose is to gain an understanding on user's privacy concerns in using the app, including mobile apps that are location-based. The latter purpose came more from the CMDE's perspective. Because the UniShuttle system was one of the earliest projects in CMDE, and location-based projects were considered to be the main development focus in CMDE, understanding users' location privacy perception on location based mobile applications will not only help this project but also help guide future project development in CMDE.

5.1 Research Evaluation Design

The evaluation study involved a pre-test – post-test experimental design. The study was conducted between November 2011 and August 2012 after extensive pilot testing of the survey questionnaire. Research assistants approached potential respondents at shuttle bus stops. Participants had to have an iPhone and had to be willing to download the app and complete the surveys. Once respondents gave their consent to participate, the research assistants opened the online survey questionnaire on an iPad and guided the respondents through the questionnaires. They encouraged the participants to download the app at the end of the survey but several participants preferred to download the app later when they had access to high speed Internet and/or did not have to rush.

Study participants were given a four-week time window to interact with the app. They were then sent a link to a second online questionnaire to complete the post-test. At least two reminder emails were sent to the participants.

Several challenges were encountered during the study. First, iPhone penetration proved to be low and it was difficult to recruit sufficient numbers of test subjects. Many of the transport users on the routes in question were students who use non-Apple smart-phones. Therefore, the iOS app proved rather limiting despite expectations to the contrary. Second, the weather was very bad throughout most of the study phase, making it very difficult to interact with participants at the bus stops. To compensate, an email was sent to students and staff to recruit further test subjects. This had the desired effect.

The test participants email address was used as the unique identifier to link survey 1 and survey 2 as well as track the download of the app and its actual usage.

5.2 Sample Description

A total of 53 participants provided valid responses for survey 1. All actually downloaded the app. 27 respondents who had downloaded the app interacted with it at least once. Only the responses of those actual users were considered for the evaluation of the app. A total of 18 respondents who had actually used the app completed the entire study process.

5.2.1 Respondent Profile Overall

There are slightly more female respondents (53.8%) in the sample than males. Their age ranges from 18 to 54 years, with about half being born before 1990 and half after. About half are full-time students, almost 10% are full-time employees and the rest are

part-time or casual employees. The test group are highly educated, with 71.2% having completed at least some university education. All but one respondent had previously downloaded a mobile app.

5.2.2 Respondent Profile of Actual Users Completing the Entire Experiment

Those respondents who completed Phase 1 and Phase 2 of the research and who actually used the app were slightly more female (55.6%) and a lot younger (max. age = 32) than the overall sample. They all scored highly on agreeableness and conscientiousness as a measure of their personality.

5.3 Results of Phase 1

5.3.1 Transport Use Behaviour

The respondents exhibited a range of shuttle bus use behaviours ranging from taking the shuttles more than 5 days a week to less often than one day a week, with 2-4 days a week being the most frequent response across all three shuttle routes.

Only 9.6% know exactly at what times the shuttle busses leave, 67.3% have a general idea and 23.1% have no idea. 80.8% think the shuttle schedules are convenient and 80.8% think the shuttle schedules are reliable.

The majority of respondents (60.5%) are willing to wait 10 minutes or less for a shuttle bus to arrive. Only 2 respondents indicated they are willing to wait for 30 minutes.

A clear majority (88.5%) indicated that they would use a Wi-Fi hotspot if it was available on a bus. Since the shuttle buses are equipped by the experiment with a WiFi hotspot this seems a popular feature of our transport solution.

5.3.2 Expectations of the App Functions

The UniShuttle app's function of telling the time of arrival at a particular stop attracted 83% users, followed by the function of displaying the map that shows the actual position of buses, and the function that shows a map to simply pinpoint the bus stop locations as the least useful. Figure 5-1 shows the expected usefulness of the UniShuttle app features.

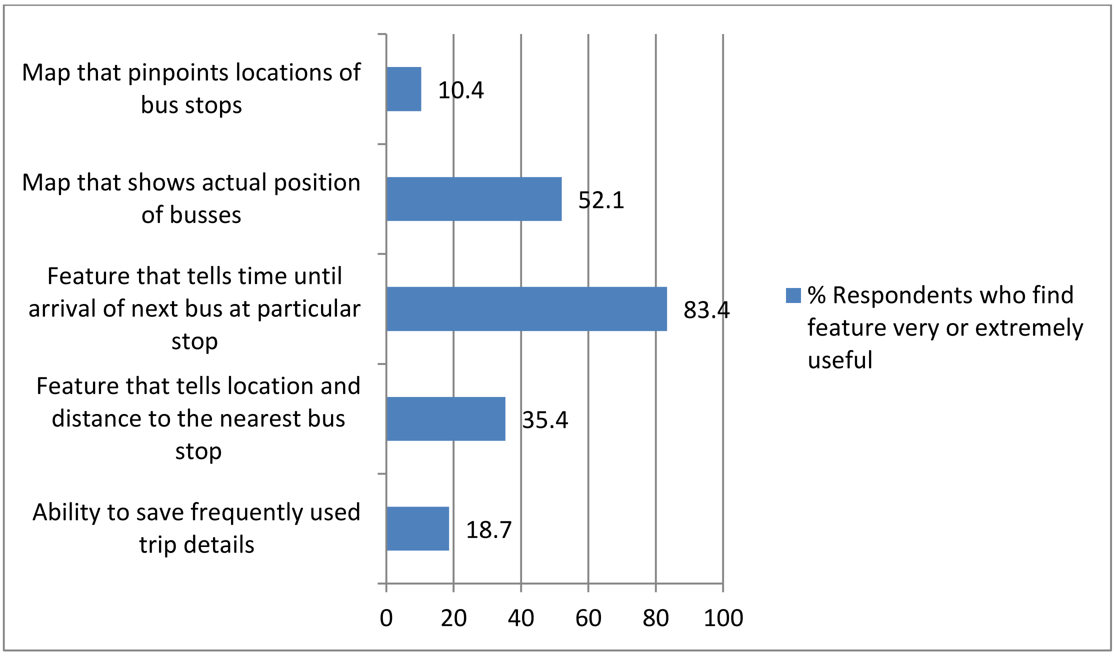


Figure 5-1: Evaluation of usefulness of the UniShuttle app functions

5.3.3 Privacy Needs and Concerns

One of the UniShuttle system's functions is to track users' location and keep record of their footprint data. In order to understand how users perceive the footprint tracking function, participants were asked what level of trust they would have in a mobile app in terms of their location information being safeguarded. About 19% indicated strong trust.

The majority (61.5%) indicated some trust and 9.6% proclaimed they are neutral. Only 7.7% reported some distrust and 1.9% strong distrust.

The survey also tested for institutional trust (Figure 5-2). The results clearly show that respondents trust the national government most, followed by universities/educational institutions and regional/state governments when it comes to respecting their privacy and not sharing their personal information without their permission.

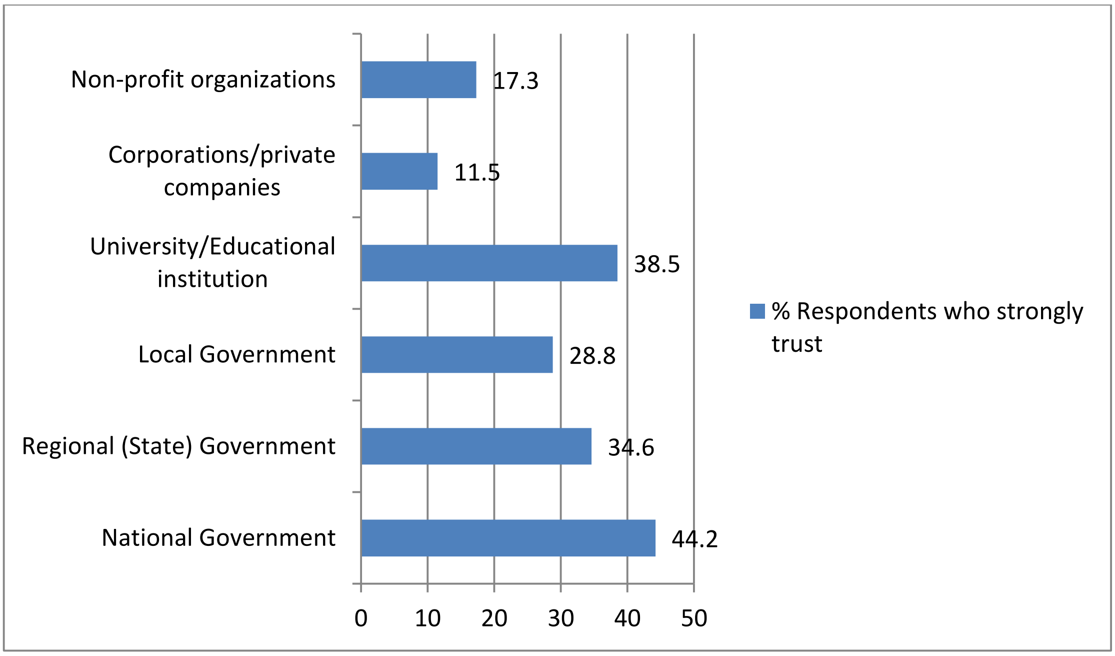


Figure 5-2: Institutional Trust

Regarding the overall participant's sentiments on privacy, the results show that the respondents value privacy highly but this does not necessarily mean that they would not provide personal information. However, it seems that giving permission first is important to them. Figure 5-3 shows the privacy needs. All the questions with privacy issues have concerns with high percentage.

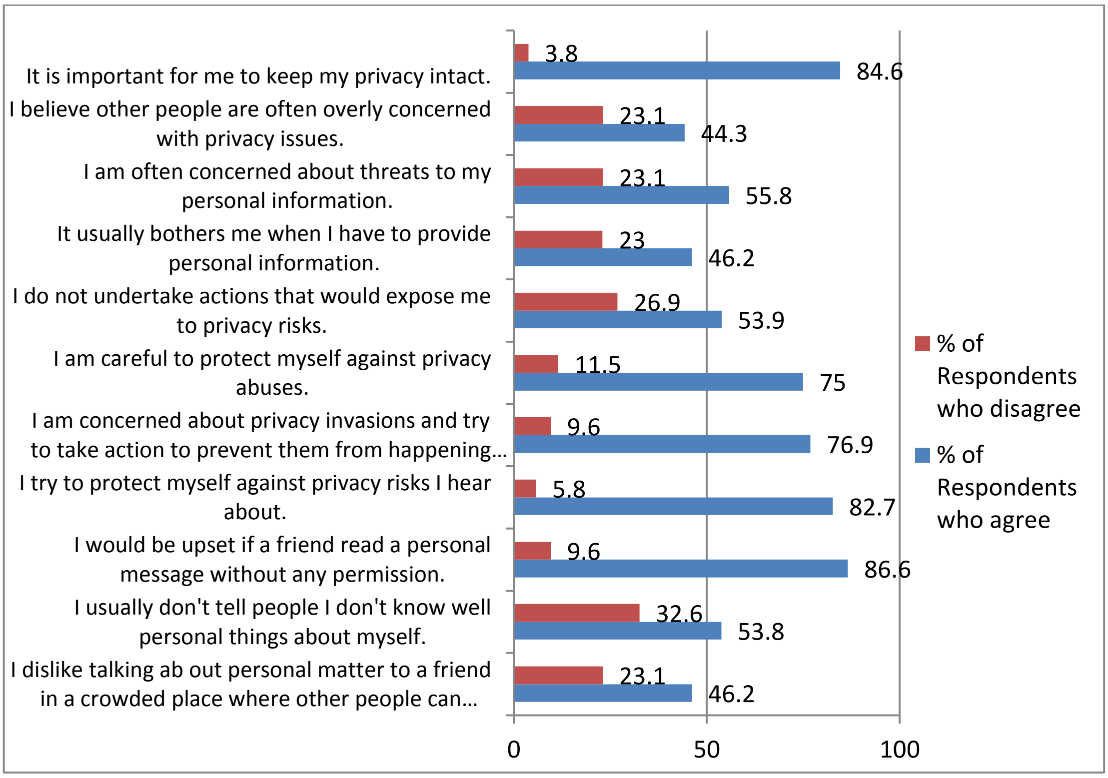


Figure 5-3: Overall participants' privacy sentiments

Regarding privacy concerns related to mobile applications, respondents are most concerned with having to provide too much information when downloading or registering a new app, followed by concerns about information sharing with third parties. Figure 5-4 shows the Mobile app related privacy concerns.

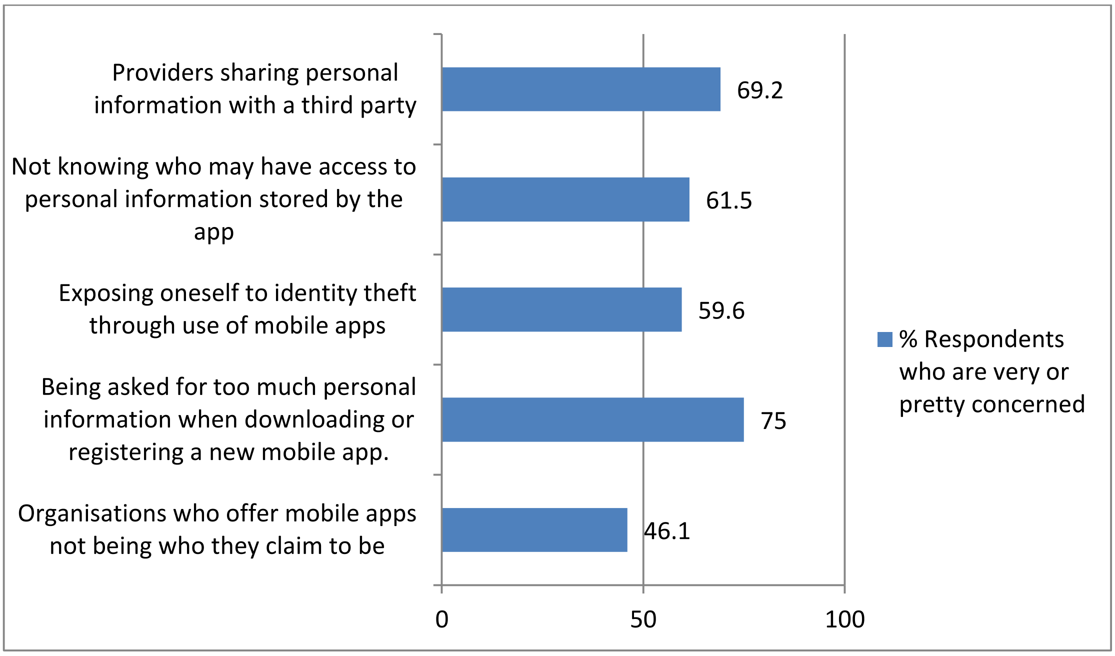


Figure 5-4: Overall privacy concerns on mobile apps

The survey further asked about specific behaviours the respondents might have engaged in to protect their privacy when engaging with mobile apps. A majority of respondents have refused to give information. Over half actually abandoned use or purchase processes due to mobile privacy concerns and a surprising 43% admitted to having provided false or fictitious information to protect their privacy. Figure 5-5 gives the actions to protect mobile app privacy.

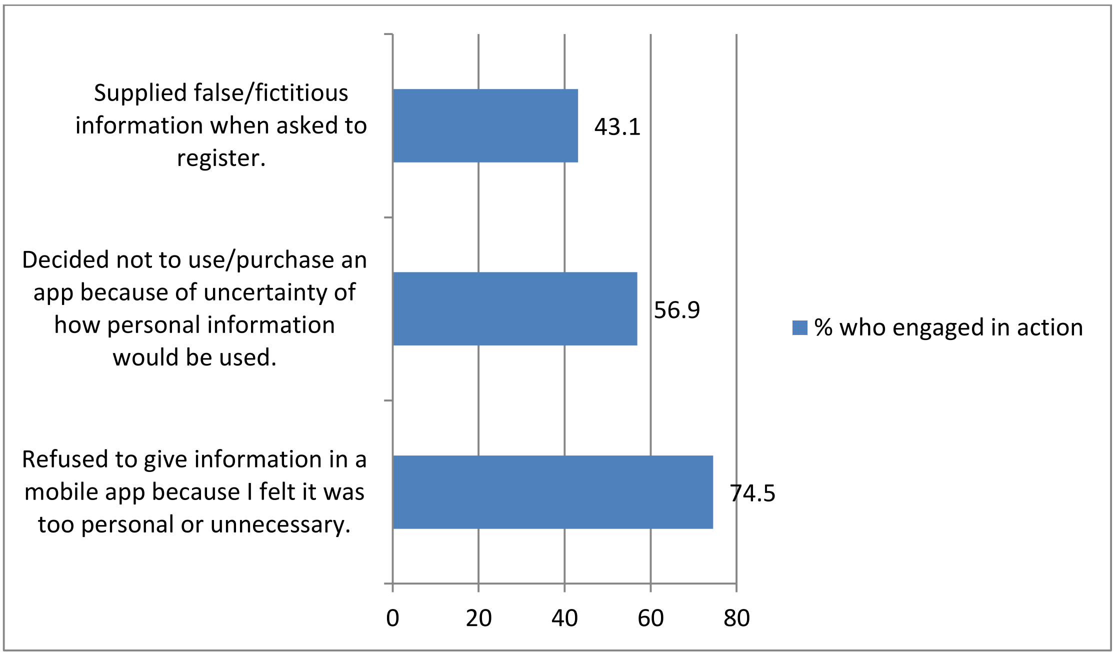


Figure 5-5: Actions to protect mobile app privacy

As far as location-based privacy is concerned, respondents were mostly worried about information being used for other services. Only about 42% were concerned about lack of their own awareness of tracking and 38% thought lack of accuracy could lead to problems. Figure 5-6 shows the location-based privacy concerns.

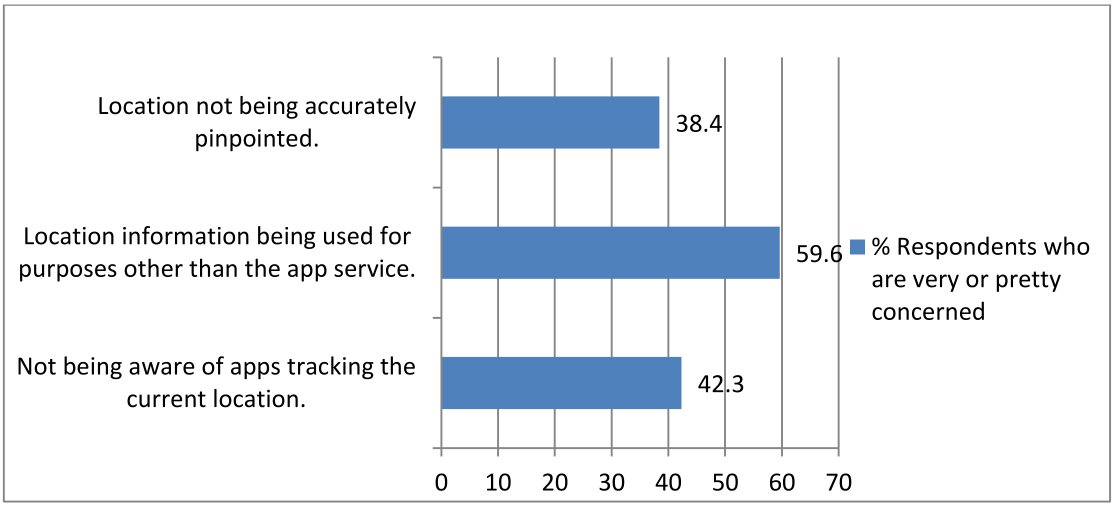


Figure 5-6: Location-based privacy concerns

A large majority (86.5%) of the respondents had used location-based services before participating in the study.

As far as willingness to have their location tracked is concerned, the results show that there is a clear drop in willingness to share this data with increasing granularity (or accuracy) of the location identification. Figure 5-7 shows the willingness to use location-based services at different levels of location granularity or accuracy.

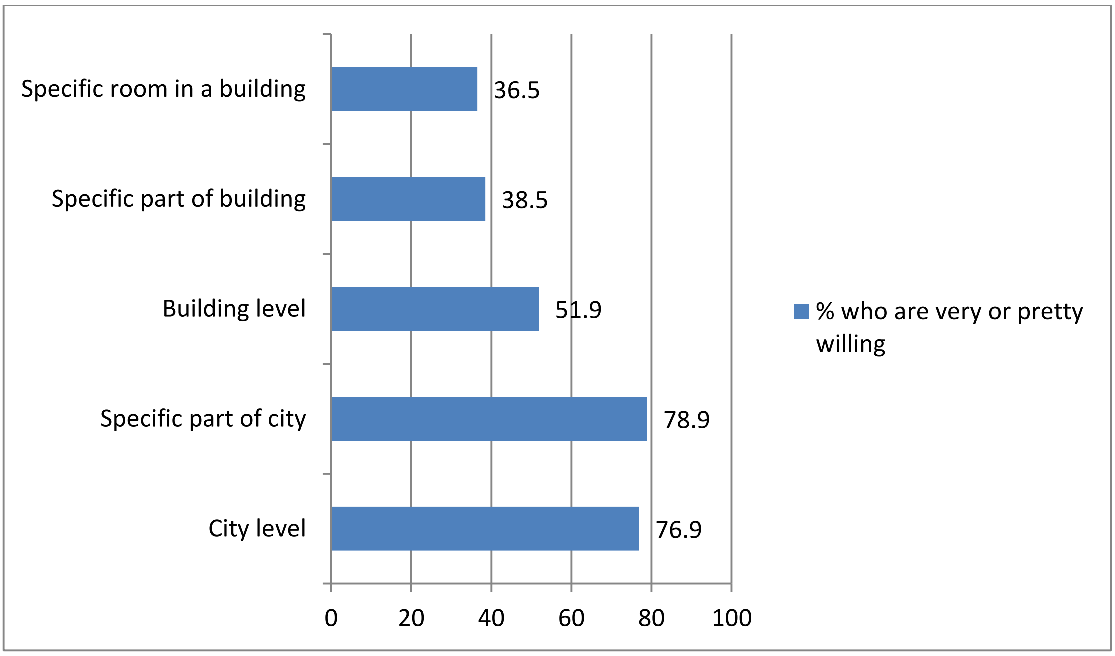


Figure 5-7: Willingness to use location-based services at different levels of location granularity or accuracy

5.4 Results of Phase 2

5.4.1 App Use

Of those respondents who used the app and completed Survey 2, 44.4% used the app regularly, 44.4% used it a couple of times and 11.1% used it only once. These numbers

largely correspond with the actual use behaviour tracked as part of the study. On average, these respondents used the app for 1691.48 seconds across 26.50 sessions.

Most of the respondents indicated that they used the feature (Figure 5-8) that tells the time until the arrival of the next bus at a particular stop (88.9%), followed by the map with the actual position of the buses and the feature that shows one's own position and distance to nearest bus stop.

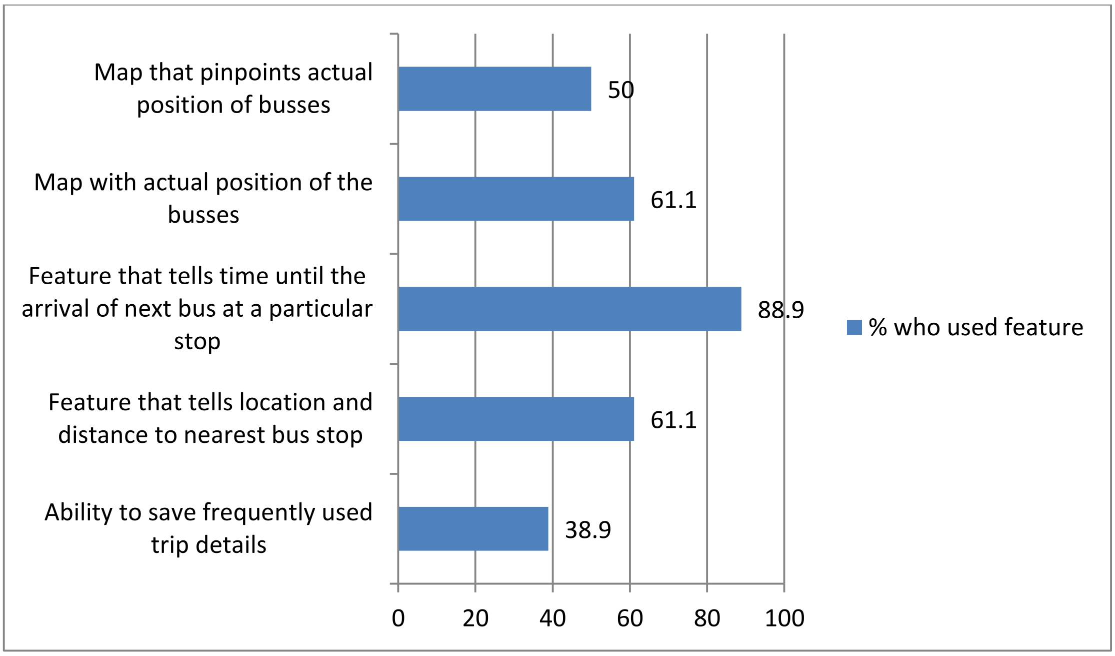


Figure 5-8: App feature usage

5.4.2 Perceived Usefulness

Overall, the respondents indicated that they found the app to be very or pretty useful (83.3%). Only 16.7% found it somewhat useful and nobody indicated that it was not useful.

Regarding the usefulness of specific features, the results are consistent with the use frequency questions and show that the feature that tells the time until the arrival of the

next bus was by far the most useful feature in the opinion of the majority of study participants but that other features were also assessed as very or pretty useful by a large portion of users. Figure 5-9 shows the perceived usefulness of individual features.

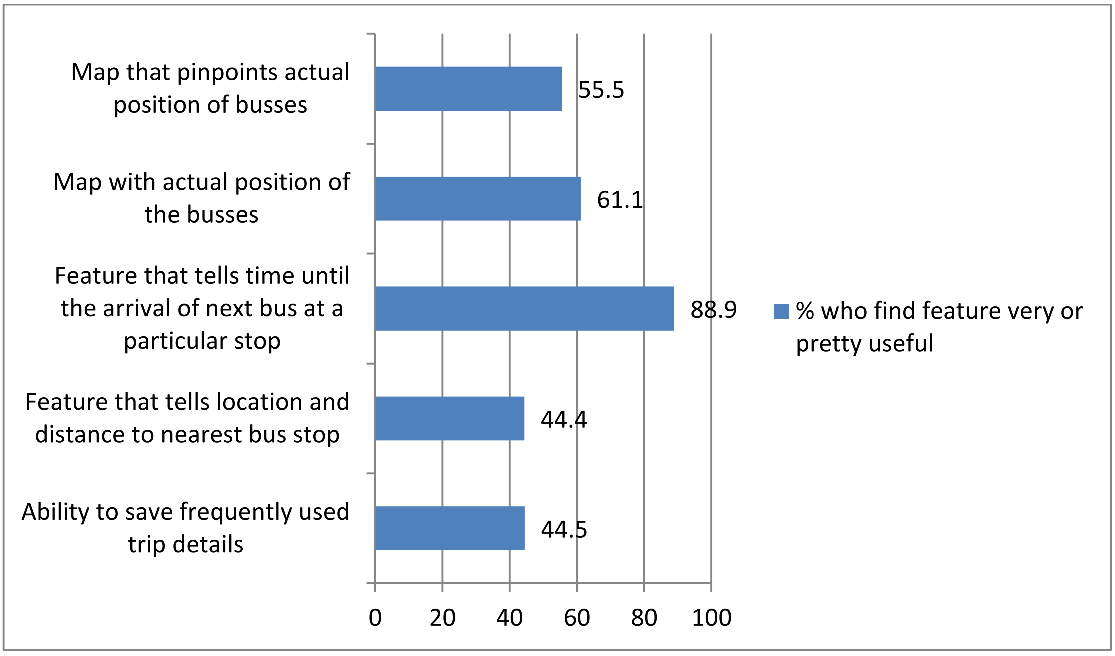


Figure 5-9: Perceived usefulness of the app features

5.4.3 User Experience

The users evaluated the user experience very positively (Figure 5-10). Only the fun element was not evaluated as highly as the other aspects of the App.

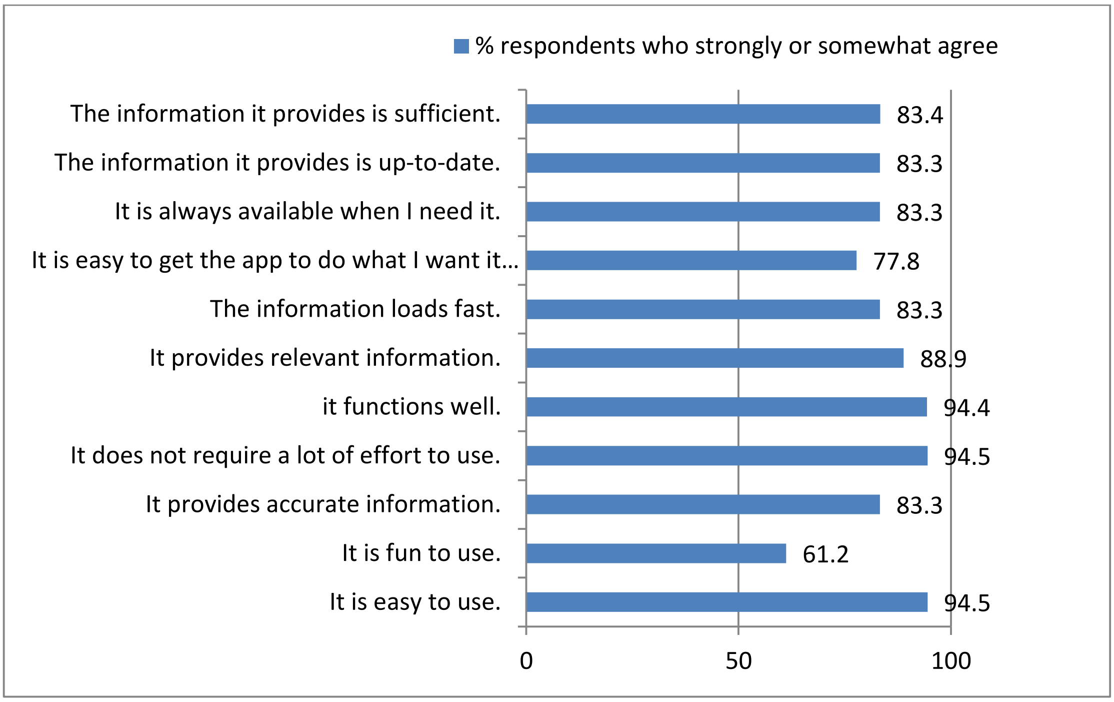


Figure 5-10: Use experience

5.4.4 Benefits Derived From App Use

Almost all respondents indicated that the app reduced their likelihood of missing a bus. They also found it improved their ability to plan ahead. Surprisingly, it reduced the waiting time only for 77.7%. Feelings of safety and entertainment were only experienced and reported by some users. Figure 5-11 gives the benefits derived (Feeling safe, easy of planning, reduce waiting time and etc.).

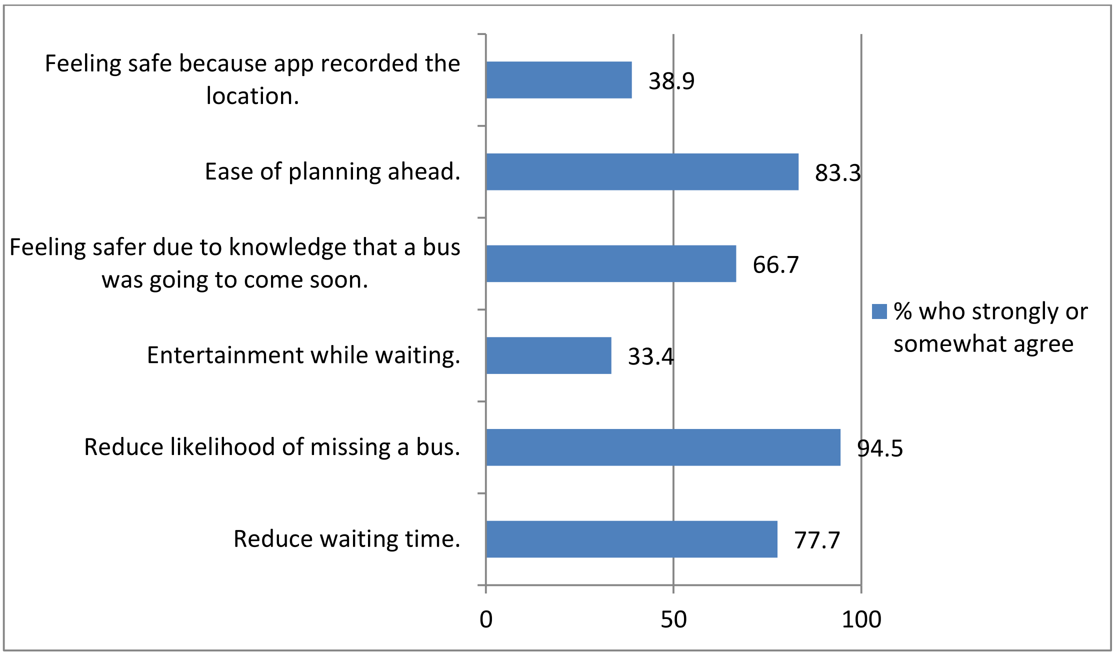


Figure 5-11: Benefits derived from the app use

5.4.5 Trust in the App

Only 16.7% indicated that they trusted the App in terms of their information being safeguarded by it and 66.7% indicated some trust. Nobody indicated distrust, which is very encouraging for this PhD study.

5.4.6 Satisfaction and Intentions to Continue Use

Almost 39% indicated they were very satisfied with the app and 50% indicated they were somewhat satisfied. Nobody indicated dissatisfaction.

The majority of users indicated that they would continue using the app (61.1% very likely, 22.2% somewhat likely). Importantly, 83.3% are very or somewhat likely to recommend the app to others.

5.4.7 Willingness to Use in Other Contexts

In order to survey whether users will be willing to use similar apps developed from CMDE but used in different areas, participants were asked if they would like to use an app like UniShuttle in a different context. The survey results show that the respondents are split in terms of whether they would like to use an app like UniShuttle in a different context, although a majority thought that such an app could be used to track their vehicle use for tax and emission tracking purposes. Figure 5-12 indicates the willingness to use in other contexts.

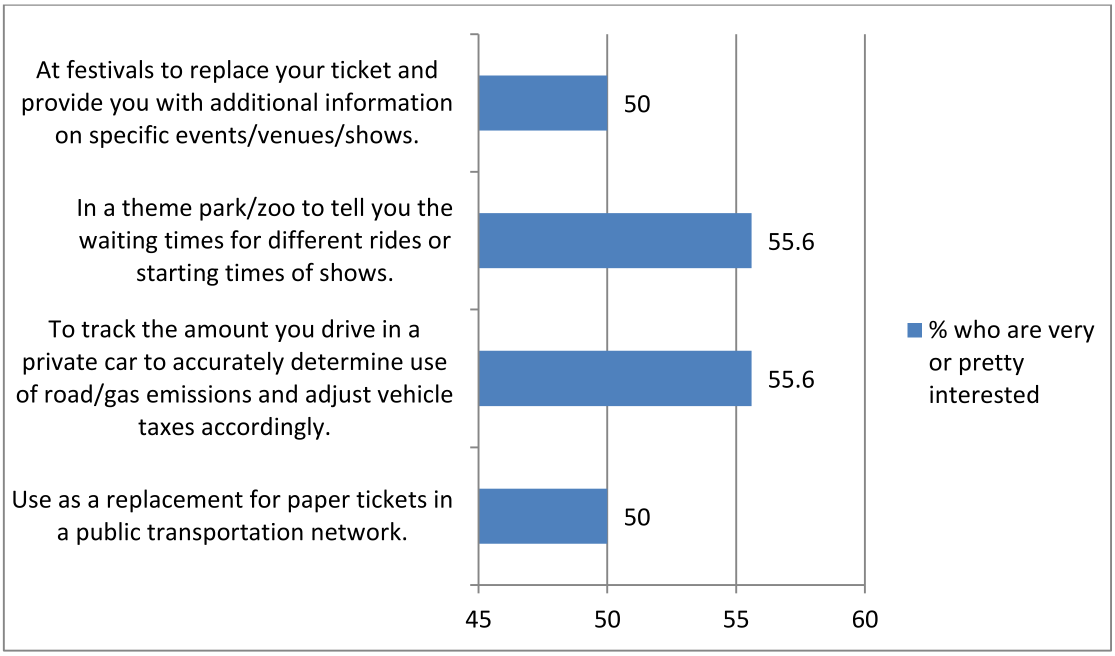


Figure 5-12: Willingness to use an app like UniShuttle in other contexts

5.4.8 Determinants of Use

A discriminant analysis was conducted to determine what distinguished those who used the app from those who downloaded it but did not use it. Only 66.7% could be correctly classified, which is better than by chance but not a stellar result.

The results further indicate that transport behaviour was the driving force, with the number of shuttle buses used, perceptions of the convenience of the bus schedule and knowledge of the bus schedule driving users' behaviour is more important than privacy concerns.

5.4.9 Privacy Concerns

Paired t-tests were conducted to examine whether after using the UniShuttle app, respondents' privacy concerns have been altered. This required constructing scales for the privacy measures collected both in Phase 1 and in Phase 2. Factor analyses and Cronbach Alpha were used to establish that the scales were uni-dimensional, internally consistent and reliable.

Need for privacy was identified with a single indicator: "It is important to me to keep my privacy intact".

General privacy concerns were constructed using four privacy concern questions ("I try to protect myself against privacy risks I hear about"; "I am concerned about privacy invasions and try to take actions to prevent them from happening to me"; "I am careful to protect myself against privacy abuses"; "I am often concerned about threats to my personal information"). Cronbach Alpha for this scale is 0.901.

Mobile privacy was derived from the five questions asked about concerns regarding mobile applications (see Figure 5.4). Cronbach Alpha for this scale is 0.893.

Location-based privacy was constructed using two of the questions used in the survey about concerns regarding location-based services (Location information being used for

purposes other than the app service; Not being aware of apps tracking the current location). The Cronbach Alpha score for this scale is 0.741.

The paired t-tests revealed no significant differences. A look specifically at location-based privacy shows that the score indeed stayed the same for most, while it changed to the positive for some and to the negative for others. Overall, however, it is very stable (Table 5-1). Correlations with extent of use (time and number of sessions) also revealed no significant patterns.

Table 5-1: Change in privacy concerns

privacychange					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-2.00	1	5.6	5.6	5.6
	-1.50	1	5.6	5.6	11.1
	-1.00	1	5.6	5.6	16.7
	-.50	3	16.7	16.7	33.3
	.00	7	38.9	38.9	72.2
	.50	2	11.1	11.1	83.3
	1.00	2	11.1	11.1	94.4
	1.50	1	5.6	5.6	100.0
	Total	18	100.0	100.0	

T-tests and regression analyses were conducted to see whether differences in privacy concerns could be explained by socio-demographic differences.

There were no significant differences between male and female respondents. Older respondents had greater general privacy concerns but did not differ from younger respondents regarding mobile app privacy and location-based services privacy concerns.

5.5 Summary

An extensive survey has been conducted in two phases on a broad range of issues including the acceptance and evaluation of the UniShuttle system app and users' perception of personal privacy towards location-based apps.

The survey results show that the respondents evaluated a very positive user experience of the app, valued highly on the information sufficiency, up-to-datedness, ease of use, accuracy, usefulness and relevance the app provided to them. Nearly 50% respondents have used the UniShuttle app on a regular basis. More than 83% respondents found the app to be very or pretty useful. The app's main feature, prediction of arrival time of shuttle bus at a particular stop, attracted nearly 90% users, followed by the feature that shows a map with actual position of a shuttle bus, which attracted over 60% users. Almost all respondents indicated that the app reduced their likelihood of missing a bus, with over 83% and 94% users reported the benefits of using the app for planning ahead and elimination of the likelihood to miss the desired bus. The majority of users indicated they would continue using the app, and over 83% respondents would like to recommend the app to others.

A majority of respondents, regardless of demographic profile, have strong concerns on personal privacy in disclosure of their personal information, but this doesn't necessarily mean they would not provide personal information. Regarding privacy concerns related to location based apps, respondents are most concerned with having to provide too much information followed by information sharing with third parties. A majority of respondents admitted that they have refused to give information in order to protect their privacy when using a mobile app. However, most respondents are willing to provide required information with their permission.

CHAPTER 6. SUMMARY AND CONCLUSIONS

The major objective of this PhD research is to design, implement and evaluate an intelligent transport system (ITS) within the UOW CMDE context. These objectives have been fulfilled. The system has now been fully implemented and has been deployed into the Wollongong community for over three years. Over this period, new improvements and evaluations have been carried out.

6.1 System Review

The UniShuttle intelligent transport digital ecosystem is designed and developed by integrating the digital ecosystem concepts and principles into the ITS design and technologies. With its characteristics of being self-sustaining and self-evolving, the system sets out to be an extraordinary ITS system that has the potential to have longevity and progress over time.

The thesis presents the holistic work of the ITS project, including the architectural design, hardware consideration, software design and implementation, communication design, app development, testing methods, system integration, data collection and analysis, and system evaluation through user survey. Highlights of this work are presented below.

From an architectural design perspective, this research presents a foundation based on the IETF presence model, which enables it to integrate distributed components together, and gather and send the right information to the right parties. Based on the digital ecosystem design concept proposed by the European Commission and the SOA

principles, the research exploits a software architecture that allows its software components to be reusable and expandable.

In choosing the on-board device installed on the shuttle buses, rather than to design specific hardware, a commercial in-vehicle PC is used as the hardware platform on shuttle buses. This provides the required key functions at an economic price and with straightforward maintenance.

The central server is designed with multiple sub-server functions to provide storage, control and communications with the large network of distributed devices. Combinations of varied communication protocols are employed to ensure smooth and secure communication.

In predicting shuttle bus arrival time, this research develops a prediction model based on historical data combined with the current bus information. Using the RMSE method, the prediction accuracy of the model is compared with simple linear regression and scheduled timetable. The developed prediction model produces a much better performance, and also proves to be satisfactory to users as revealed by a customer survey.

In addition to an online web interface, the research develops a dedicated app to provide users with the key functions of telling real-time bus arrival information and the whereabouts of the shuttle buses on iPhones. The app is evaluated by users for the system's usability, usefulness and users' reaction. The results show that most of the respondents value the system highly, with high ratings on the information sufficiency, information accuracy, useful and relevant information, and easiness of use.

The system has been in operation since late 2011 with only one hardware failure of an on-board device requiring replacement. With its front end app being downloaded over 10 thousand times, and hundreds of trips recorded each week, the system has demonstrated its viability and reliability.

6.2 System Benefits

This PhD work for the first time built an innovative intelligent transport system that has been fully implemented and extensively trialled on a shuttle bus system in the City of Wollongong. This UniShuttle system is readily available to city dwellers, students and commuters.

To city passengers, the system enables them to make better decisions to use shuttle buses, or in a wider sense, public transport network, by their ability to access real time bus scheduling information anytime, anywhere. The availability of the real time scheduling information has proved to be helpful to set up passengers' expectations, reduce anxiety, and eliminate the likelihood of missing a bus. The passengers' active use of the UniShuttle app also allows the system to capture the passenger use pattern and identify the repeatable passenger use patterns, which allows the dynamic scheduling of buses to optimise throughput. This process continues and forms an on-going digital ecosystem cycle, which consequently allows the system to self-evolve and self-sustain

For transport operators/authorities and university planners, this research provides a cost effective solution to make real time bus arrival information available. The provision of real arrival time can increase the customer satisfaction level. Also, the real time arrival information can equip them to proactively manage expectations and react to them if necessary. Secondly, and perhaps more importantly, the bus travelling and passenger

use data gathered over time in the system could also be analysed to generate patterns. The identified patterns can reveal meaningful information about the shuttle buses' network usage and resource management. The better management practice applied to the transport network, allows the system to evolve with more passenger participation.

It is also worth mentioning that, although the majority of the focus in this study was focused on the passengers, the bus drivers also gain a better understanding of how they are traveling against schedules and so they can act correspondingly to offer a positive impact into the system.

Further, by increasing the service level of the transport network, more passengers will be encouraged to take the shuttle buses, rather than resorting to private transportation. With more people taking public transport, there will be less green gas emission to the environment, and less energy consumption which will benefit our community at large.

6.3 Future work

The UniShuttle system was built on the CMDE context. A great amount of data, such as when and where passengers get on and get off buses, has been gathered through this project. However, this dataset alone can be limited in terms of gaining a more realistic modelling in understanding the population behaviours. It would be greatly beneficial if this dataset could be analysed across with the other datasets, such as the patron data in nearby shops around the bus stops, residents' health data, and city event activities. The cross analysis of this data could potentially reveal insights for a better behavioural understanding of the population, which can guide better practices for businesses, organisations and the community with both social and economic impact.

PwC (PwC pulse 2014) has highlighted that a key aspect of building a common digital platform for modern transportation system is to have more and more developers contribute to the development of the system. It is a wish of this research that the foundation built for this thesis can be used as the open infrastructure basis for the other developers to build new projects. In such a manner the current digital ecosystem can grow and advance.

REFERENCES

Abbott, E., & Powell, D. (1999). Land-vehicle navigation using GPS. *Proceedings of the IEEE*, *87*. doi:10.1109/5.736347

Agre, P. (1995). Looking down the road: Transport informatics and the new landscape of privacy issues. *CPSR Newsletter*, *13*, 15–20.

Al-Bayari, O., & Sadoun, B. (2005). New centralized automatic vehicle location communications software system under GIS environment. *International Journal of Communication Systems*, *19*, 833–846.

Aloquili, O., Elbanna, A., & Al-Azizi, A. (2009). Automatic vehicle location tracking system based on GIS environment. *IET Software*. doi:10.1049/iet-sen.2008.0048

Altinkaya, M., & Zontul, M. (2013). Urban Bus Arrival Time Prediction: A Review of Computational Models. *Int'l Journal of Recent Technology and Engineering*, *2*, 164–169.

Anuar, F., & Gretzel, U. (2011). Privacy Concerns in the Context of Location-Based Services for Tourism. In *ENTER 2011 Conference, Innsbruck, Austria*.

Badr, Y., & Caplat, G. (2010). Software versioning and evolution in digital ecosystems. In *4th IEEE International Conference on Digital Ecosystems and Technologies - Conference Proceedings of IEEE-DEST 2010, DEST 2010* (pp. 381–386). doi:10.1109/DEST.2010.5610616

Bajaj, R., Ranaweera, S. L., & Agrawal, D. P. (2002). GPS: location-tracking technology. *Computer*, *35*. doi:10.1109/MC.2002.993780

Barkhuus, L., Brown, B., Bell, M., Hall, M., Sherwood, S., & Chalmers, M. (2008). From awareness to repartee: Sharing Location within Social Groups. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08* (p. 497). doi:10.1145/1357054.1357134

Barkhuus, L., & Dey, A. (2003). Location-Based Services for Mobile Telephony : a Study of Users ' Privacy Concerns. In *Proceedings of the INTERACT 2003, 9TH IFIP TC13 International Conference on Human-Computer Interaction* (pp. 1–5). doi:10.1.1.10.527

Barnes, R., Lepinski, M., Cooper, A., Morris, J., Tschofenig, H., and H. Schulzrinne (2011) "An Architecture for Location and Location Privacy in Internet Applications", *BCP 160, RFC 6280*

Beckwith, R. (2003). Designing for ubiquity: The perception of privacy. *IEEE Pervasive Computing*, *2*, 40–46. doi:10.1109/MPRV.2003.1203752

Biuk-Aghai, R. P., Tang, L. V. S., Fong, S., & Si, Y. W. (2009). Wikis as digital ecosystems: An analysis based on authorship. In *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies, DEST '09* (pp. 581–586). doi:10.1109/DEST.2009.5276695

BMBF. (2002). *Mobility in Conurbations – First results*. *Bonn: Federal Ministry of Education and Research (BMBF).*

Borenovic, M., Neskovic, A., & Neskovic, N. (2013). Vehicle positioning using GSM and cascade-connected ANN structure. *Intelligent Transportation Systems, IEEE Transactions on*, *14*, 34–46.

Braitman, K. A., McCartt, A. T., Zuby, D. S., & Singer, J. (2010). Volvo and Infiniti drivers' experiences with select crash avoidance technologies. *Traffic Injury Prevention*, *11*, 270–278. doi:10.1080/15389581003735600

Briscoe, G. (2009). *Digital ecosystems*. Imperial College, London.

Briscoe, G. (2010). Complex adaptive digital EcoSystems. In *Proceedings of the International Conference on Management of Emergent Digital EcoSystems - MEDES '10* (p. 39). doi:10.1145/1936254.1936262

Briscoe, G., & De Wilde, P. (2006). Digital ecosystems: Evolving service-orientated architectures. In *2006 1st Bio-Inspired Models of Network, Information and Computing Systems, BIONETICS*. doi:10.1109/BIMNICS.2006.361817

Briscoe, G., & De Wilde, P. (2008). Digital ecosystems: Optimisation by a distributed intelligence. In *2008 2nd IEEE International Conference on Digital Ecosystems and Technologies, IEEE-DEST 2008* (pp. 192–197). doi:10.1109/DEST.2008.4635157

Briscoe, G., & De Wilde, P. (2012). Ecosystem-oriented distributed evolutionary computing. In *Computational Intelligence (UKCI), 2012 12th UK Workshop on* (pp. 1–8).

Briscoe, G., Sadedin, S., & De Wilde, P. (2011). Digital ecosystems: Ecosystem-oriented architectures. *Natural Computing*, *10*, 1143–1194. doi:10.1007/s11047-011-9254-0

Briscoe, G., Sadedin, S., & Paperin, G. (2007). Biology of applied digital ecosystems. In *Proceedings of the 2007 Inaugural IEEE-IES Digital EcoSystems and Technologies Conference, DEST 2007* (pp. 458–463). doi:10.1109/DEST.2007.372015

Casey, R. F., Labell, L. N., Moniz, L., Royal, J. W., Sheehan, M., Sheehan, T., … Parker, D. (2000). *Advanced Public Transportation Systems: The State of the Art Update 2000*.

Chadil, N., Russameesawang, A., & Keeratiwintakorn, P. (2008). Real-Time Tracking Management System Using GPS , GPRS and Google Earth. In *Proceedings of ECTI-CON 2008* (pp. 393–396). doi:10.1109/ECTICON.2008.4600454

Chang, E., Quaddus, M., & Ramaseshan, R. (2006). The Vision of DEBI Institute: Digital Ecosystems and Business Intelligence.

Chang, E., & West, M. (2006). Digital Ecosystems A Next Generation of the Collaborative Environment. *The Eight International Conference*, *214*, 3–23. Retrieved from http://research.mercubuana.ac.id/proceeding/iiWAS2006_21-41.pdf

Chang, N., Rashidzadeh, R., & Ahmadi, M. (2010). Robust indoor positioning using differential Wi-Fi access points. *IEEE Transactions on Consumer Electronics*, *56*, 1860–1867. doi:10.1109/TCE.2010.5606338

Chatterjee, J., Prabhakar, T. V., & Sarkar, R. (2007). Evolution of a Digital Ecosystem

for Knowledge Services to Indian Agriculture. In F. Nachira, P. Dini, A. Nicolai, M. Le Louarn, & L. Rivera Lèon (Eds.), *Digital Business Ecosystems* (pp. 176–182). Office for Official Publications of the European Communities.

Chen, C.-D., Fan, Y.-W., & Farn, C.-K. (2007). Predicting electronic toll collection service adoption: An integration of the technology acceptance model and the theory of planned behavior. *Transportation Research Part C: Emerging Technologies*, *15*, 300–311.

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

Chen, Y., Lymberopoulos, D., Liu, J., & Priyantha, B. (2012). FM-based indoor localization. In *MobiSys'12 - Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services* (pp. 169–181). doi:10.1145/2307636.2307653

Chen, Y.-C., Chiang, J.-R., Chu, H., Huang, P., & Tsui, A. W. (2005). Sensor-assisted Wi-Fi indoor location system for adapting to environmental dynamic. In *Proceedings of the 8th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems* (pp. 118–125).

Cheng, H.-Y., Jeng, B.-S., Tseng, P.-T., & Fan, K.-C. (2006). Lane Detection With Moving Vehicles in the Traffic Scenes. *IEEE Transactions on Intelligent Transportation Systems*. doi:10.1109/TITS.2006.883940

Chien, S. I.-J., Ding, Y. Q., & Wei, C. H. (2002). Dynamic bus arrival time prediction with artificial neural networks. *Journal of Transportation Engineering-Asce*, *128*, 429–438. Retrieved from <Go\nto\nISI>://000177512600005

Chien, S. I.-J., & Kuchipudi, C. M. (2003). Dynamic Travel Time Prediction with Real-Time and Historic Data. *Journal of Transportation Engineering*. doi:10.1061/(ASCE)0733-947X(2003)129:6(608)

Chu, L., Oh, J.-S., & Recker, W. (2005). Adaptive Kalman Filter Based Freeway Travel time Estimation. *Transportation Research Board 2005 Annual Meeting*, 1–21. Retrieved from http://www.www.clr-analytics.com/Files/Adaptive Kalman Filter Based Freeway Travel time .pdf

Clement, S., & Taylor, M. (2006). The simple platoon advancement model of ITS technologies applied to vehicle control at signalised intersections. *Journal of Advanced Transportation*, *40*, 1–21.

Consolvo, S., Smith, I. E., Matthews, T., LaMarca, A., Tabert, J., & Powledge, P. (2005). Location Disclosure to Social Relations: Why, When, & What People Want to Share. *CHI 2005 Conference on Human Factors in Computing Systems*, 81–90. doi:10.1145/1054972.1054985

Coogan, M. . (2003). Recent Developments in Integrated Passenger Information Systems in North America. In *3rd UITP Conference on Travel Information, Gothenburg/Sweden*.

D'Angelo, G., Ferretti, S., Ghini, V., & Panzieri, F. (2011). Mobile Computing in Digital Ecosystems: Design Issues and Challenges. In *2011 7TH INTERNATIONAL WIRELESS COMMUNICATIONS AND MOBILE*

COMPUTING CONFERENCE (IWCMC) (pp. 2127–2132).

D'Angelo, M. P., Al-Deek, H. M., & Wang, M. C. (1999). Travel-Time Prediction for Freeway Corridors. *Transportation Research Record: Journal of the Transportation Research Board*. doi:10.3141/1676-23

D'Roza, T., & Bilchev, G. (2003). An overview of location-based services. *BT Technology Journal*. doi:10.1023/A:1022491825047

Day, M., Rosenberg, J., & Sugano, H. (2000). A Model for Presence and Instant Messaging. *Request for Comments*. Retrieved from http://www.ietf.org/rfc/rfc2778.txt

DeLurgio, S. (1998). Forecasting principles and applications. *Statistics*.

Dessouky, M., Hall, R., Nowroozi, A., & Mourikas, K. (1999). Bus dispatching at timed transfer transit stations using bus tracking technology. *Transportation Research Part C: Emerging Technologies*. doi:10.1016/S0968-090X(99)00019-4

Digital Ecosystems for eBusiness - Reference pages. (2004). Retrieved September 24, 2014, from http://www.digital-ecosystems.org/

Djuknic, G. M., & Richton, R. E. (2001). Geolocation and assisted GPS. *Computer*, *34*. doi:10.1109/2.901174

Doan, Q. C., Berradia, T., & Mouzna, J. (2009). Wireless communication-based measurement of vehicle speed and traffic volume. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC* (pp. 703–708). doi:10.1109/ITSC.2009.5309867

Dong, H., Hussain, F. K., & Chang, E. (2009). A hybrid service metadata clustering methodology in the digital ecosystem environment. In *Proceedings - International Conference on Advanced Information Networking and Applications, AINA* (pp. 238–243). doi:10.1109/WAINA.2009.205

Dong, J., Zou, L., & Zhang, Y. (2013). Mixed Model For Prediction Of Bus Arrival Times. In *Evolutionary Computation (CEC), 2013 IEEE Congress on* (pp. 2918–2923). IEEE.

Dynamic Clarity (2011). Global & Australian Smartphone and Apple iOS statistics in June 2011. Retrieved June 11, 2015, from http://www.slideshare.net/dynamicclarity/global-australian-smartphone-and-apple-ios-statistics-in-june-2011

Dziekan, K., & Kottenhoff, K. (2007). Dynamic at-stop real-time information displays for public transport: effects on customers. *Transportation Research Part A: Policy and Practice*, *41*, 489–501. doi:10.1016/j.tra.2006.11.006

Dziekan, K., & Vermeulen, A. . (2004). The added value of real-time information and effects to customer behaviour-A before-after study of the implementation of a new real-time information system. *Project Report*. *Stockholm/The Hague: KTH/HTM*.

Dziekan, K and Sedin, S. (2005). Customer reactions to the implementation of a trunk bus network in Stockholm. In *56th UITP World Congress*.

Eklund, P., Thom, J., Wray, T., & Dou, E. (2011). Location Based Context- Aware Services in a digital Ecosystem with Location Privacy. *Ournal of Cases on*

Information Technology (JCIT), *13*, 49–68.

Eklund, P., Thom, J., Wray, T., & Thomson, M. (2010). Location privacy in a digital ecosystem for context-aware applications. In *Proceedings of the International Conference on Management of Emergent Digital EcoSystems* (pp. 137–144).

Feng, S. F. S., & Law, C. L. L. C. L. (2002). Assisted GPS and its impact on navigation in intelligent transportation systems. *Proceedings*. *The IEEE 5th International Conference on Intelligent Transportation Systems*. doi:10.1109/ITSC.2002.1041344

Ferster, C. J., & Coops, N. C. (2013). A review of earth observation using mobile personal communication devices. *Computers & Geosciences*, *51*, 339–349. doi:10.1016/j.cageo.2012.09.009

French, R. L. (1986). Historical overview of automobile navigation technology. *36th IEEE Vehicular Technology Conference*, *36*. doi:10.1109/VTC.1986.1623457

Fu, H. F. H. (2006). Formal Concept Analysis for Digital Ecosystem. *2006 5th International Conference on Machine Learning and Applications (ICMLA'06)*. doi:10.1109/ICMLA.2006.24

Fukumoto, J., Sirokane, N., Ishikawa, Y., Wada, T., Ohtsuki, K., & Okada, H. (2007). Analytic method for real-time traffic problems by using Contents Oriented Communications in VANET. In *ITST 2007 - 7th International Conference on Intelligent Transport Systems Telecommunications, Proceedings* (pp. 40–45). doi:10.1109/ITST.2007.4295830

Garnett, R., & Stewart, R. (2014). Comparison of GPS units and mobile Apple GPS capabilities in an urban landscape. *Cartography and Geographic Information Science*, *42*, 1–8.

Giaglis, G., & Pateli, A. (2002). On the potential use of mobile positioning technologies in indoor environments. *The Proceedings of …*, 413–429. Retrieved from https://domino.fov.uni-mb.si/proceedings.nsf/Proceedings/470EE786EC116F7DC1256E9F0034EBF0/$File/giaglis.pdf

Giannoutakis, K. N., & Li, F. (2012). Making a Business Case for Intelligent Transport Systems: A Holistic Business Model Framework. *Transport Reviews*, *32*, 781–804. doi:10.1080/01441647.2012.740096

Gildea, D., & Sheikh, M. (1996). Applications of technology in providing transit information. *Transportation Research Record: Journal of the Transportation Research Board*, *1521*, 71–76.

Girardin, F., & Blat, J. (2008). Assessing pervasive user-generated content to describe tourist dynamics. *International Workshop on Trends in Pervasive and …*. Retrieved from http://www.girardin.org/fabien/publications/girardin_trends_giscience_2008.pdf

Glasgow, M. L., Rudra, C. B., Yoo, E.-H., Demirbas, M., Merriman, J., Nayak, P., … Mu, L. (2014). Using smartphones to collect time–activity data for long-term personal-level air pollution exposure assessment. *Journal of Exposure Science and Environmental Epidemiology Jean Others*. doi:10.1038/jes.2014.78

Gruteser, M., & Grunwald, D. (2003). Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. *Proceedings of the 1st International Conference on Mobile Systems Applications and Services MobiSys 03*, *Proceeding*, 31–42. doi:10.1145/1066116.1189037

Hadzic, M., Dillon, T., & Chang, E. (2007). Use of digital ecosystem and ontology technology for standardization of medical records. In *Proceedings of the 2007 Inaugural IEEE-IES Digital EcoSystems and Technologies Conference, DEST 2007* (pp. 595–601). doi:10.1109/DEST.2007.372043

Hadzic, M., & Dillon, T. S. (2008). Application of digital ecosystems in health domain. In *2008 2nd IEEE International Conference on Digital Ecosystems and Technologies, IEEE-DEST 2008* (pp. 543–547). doi:10.1109/DEST.2008.4635222

Hagan, M. T., Demuth, H. B., & Beale, M. H. (1996). *Neural network design*. *Boston Massachusetts PWS* (Vol. 2, p. 734). Retrieved from http://ecee.colorado.edu/academics/schedules/ECEN5120.pdf

Hightower, J., & Borriello, G. (2001). Location systems for ubiquitous computing. *Computer*, *34*, 57–66. doi:10.1109/2.940014

Huang, J. S., Hsueh, K., & Reynolds, A. (2013). A framework for collaborative social, economic and environmental development: Building a digital ecosystem for societal empowerment. In *Digital Ecosystems and Technologies (DEST), 2013 7th IEEE International Conference on* (pp. 166–171).

Infopolis. (1998). *Review of Current Passenger Information Systems* (p. 39). Retrieved from ftp://ftp.cordis.europa.eu/pub/telematics/docs/tap_transport/infopolis2_d1.pdf

Ishak, S., & Al-Deek, H. (2002). Performance Evaluation of Short-Term Time-Series Traffic Prediction Model. *Journal of Transportation Engineering*. doi:10.1061/(ASCE)0733-947X(2002)128:6(490)

Jeong, R. H. (2004). *The prediction of bus arrival time using Automatic Vehicle Location Systems data*. Texas A & M University.

Jeong, R. H., & Rilett, L. R. (2005). Bus arrival time prediction model for real time applications. *Transportation Research Record: Journal of the Transportation Research Board*, *1927*, 195–204.

Jin, Y., & Sendhoff, B. (2008). Pareto-based multiobjective machine learning: An overview and case studies. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*. doi:10.1109/TSMCC.2008.919172

Kale, R. (2014). Real Time Passenger Information System. *International Journal of Advanced Research in Computer Science & Technology*, *2*(3), 153–155.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME-Journal of Basic Engineering*, *82*, 35–45. doi:10.1115/1.3662552

Karger, P. A., & Frankel, Y. (1995). Security and privacy threats to ITS. In *Steps Forward*. *Intelligent Transport Systems World Congress*.

Kim, S., Mankoff, J., & Paulos, E. (2013). Sensr: Evaluating A Flexible Framework for Authoring Mobile Data-Collection Tools for Citizen Science. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work - CSCW '13*, 1453–

1462. doi:10.1145/2441776.2441940

Klasnja, P., Consolvo, S., Choudhury, T., Beckwith, R., & Hightower, J. (2009). Exploring privacy concerns about personal sensing. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5538 LNCS, pp. 176–183). doi:10.1007/978-3-642-01516-8_13

Korhonen, J. (2003). *Introduction to 3G mobile communications*. Artech House, Inc.

LaMance, J., DeSalas, J., & Jarvinen, J. (2002). Assisted GPS: a low-infrastructure approach. *GPS World*, *13*.

Lassabe, F., Canalda, P., Chatonnay, P., & Spies, F. (2009). Indoor Wi-Fi positioning: Techniques and systems. *Annales Des Telecommunications/Annals of Telecommunications*, *64*, 651–664. doi:10.1007/s12243-009-0122-1

Lee, W. H., Tseng, S. S., & Wang, C. H. (2008). Design and implementation of electronic toll collection system based on vehicle positioning system techniques. *Computer Communications*, *31*, 2925–2933. doi:10.1016/j.comcom.2008.05.014

Lee, W., Tseng, S., & Shieh, W. (2010). Collaborative real-time traffic information generation and sharing framework for the intelligent transportation system. *Information Sciences*, *180*(1), 62–70. doi:10.1016/j.ins.2009.09.004

Lehtonen, M., & Kulmala, R. (2002). Benefits of Pilot Implementation of Public Transport Signal Priorities and Real-Time Passenger Information. *Transportation Research Record*. doi:10.3141/1799-03

Leong, R. (1989). An unconventional approach to automatic vehicle location and\ncontrol for urban transit. *Conference Record of Papers Presented at the First Vehicle Navigation and Information Systems Conference (VNIS '89)*. doi:10.1109/VNIS.1989.98766

Lezniak, T. W., Lewis, R. W., & Mcmillen, R. A. (1997). Dead Reckoning/Map Correlation System for Automatic Vehicle Tracking. *Vehicular Technology, IEEE Transactions on*, *26*(1), 47 – 60. doi:10.1109/T-VT.1977.23656

Li, B., Salter, J., Dempster, A., & Rizos, C. (2006). Indoor positioning techniques based on wireless LAN. *… Conference on Wireless …*, *CD-ROM pro*, 13–16. Retrieved from http://www.gmat.unsw.edu.au/snap/publications/lib_etal2006a.pdf

Li, F., Yu, Y., Lin, H., & Min, W. (2011). Public bus arrival time prediction based on traffic information management system. *Proceedings of 2011 IEEE International Conference on Service Operations, Logistics and Informatics*, 336–341. doi:10.1109/SOLI.2011.5986581

Li, W., Badr, Y., & Biennier, F. (2012). Digital ecosystems: challenges and prospects. In *MEDES '12 (International Conference on Management of Emergent Digital EcoSystems)* (p. 117). doi:10.1145/2457276.2457297

Lin, W.-H., & Zeng, J. (1999). Experimental Study of Real-Time Bus Arrival Time Prediction with GPS Data. *Transportation Research Record Journal of the Transportation Research Board*, *1666*, 101–109.

Lin, Y., Yang, X., Zou, N., & Jia, L. (2013). Real-Time Bus Arrival Time Prediction: Case Study for Jinan, China. *Journal of Transportation Engineering*, *139*, 1133–

1140.

Liu, H. L. H., Lint, H. van, Zuylen, H. van, & Zhang, K. Z. K. (2006a). Two distinct ways of using kalman filters to predict urban arterial travel time. *2006 IEEE Intelligent Transportation Systems Conference*. doi:10.1109/ITSC.2006.1706849

Liu, H. L. H., Lint, H. van, Zuylen, H. van, & Zhang, K. Z. K. (2006b). Two distinct ways of using kalman filters to predict urban arterial travel time. *2006 IEEE Intelligent Transportation Systems Conference*, 99–108. doi:10.1109/ITSC.2006.1706849

Liu, H., van Zuylen, H., van Lint, H., & Salomons, M. (2006). Predicting Urban Arterial Travel Time with State-Space Neural Networks and Kalman Filters. *Transportation Research Record*. doi:10.3141/1968-12

Loukakos, D. (2001). AUTOMATIC VEHICLE LOCATION-ITS DECISION REPORT. *ITS Decision Report*.

Lurgi, M., & Estanyol, F. (2010). MADBE: A Multi-Agent Digital Business Ecosystem. In *4th IEEE International Conference on Digital Ecosystems and Technologies - Conference Proceedings of IEEE-DEST 2010, DEST 2010* (pp. 262–267). doi:10.1109/DEST.2010.5610637

Lyons, G. D. (2001). Towards integrated traveller information. *Transport Reviews*, *21*, 217–235.

Lyons, G. D., & Harman, R. (2002). The UK public transport industry and provision of multi-modal traveller information. *International Journal of Transport Management*. doi:10.1016/S1471-4051(01)00002-7

Maley, C. C. (1995). Artificial Biology. *The Artificial Intelligence Lab*.

Mazloumi, E., Rose, G., Currie, G., & Moridpour, S. (2011). Prediction intervals to account for uncertainties in neural network predictions: Methodology and application in bus travel time prediction. *Engineering Applications of Artificial Intelligence*, *24*, 534–542. doi:10.1016/j.engappai.2010.11.004

Miah, S. J. (2012). The Role of End User in E-Government Application Development: A Conceptual Model in the Agricultural Context. *Journal of Organizational & End User Computing*, *24*, 69–85. doi:10.4018/joeuc.2012070104

Miles, J. C., & Chen, K. (2004). *The Intelligent Transport Systems Handbook*. (2nd ed., p. 365). Swanley : Route 2 Market, 2004.

Miller, D. L., Dot, K., Ekern, D. S., Dot, V., Walton, C. M., & Cockrell, E. H. (2008). *SYNTHESIS 73 AVL Systems for Bus Transit : COOPERATIVE A Synthesis of Transit Practice*. *Traffic Safety* (p. 114). Retrieved from http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_syn_73.pdf

Mohapatra, D., & Suma, S. B. (2005). Survey of location based wireless services. *2005 IEEE International Conference on Personal Wireless Communications, 2005. ICPWC 2005*. doi:10.1109/ICPWC.2005.1431366

Moreira, J. P. C. L. M. (2008). *Travel Time Prediction for the Planning of Mass Transit Companies: a Machine Learning Approach*. Universidade do Porto.

Nachira, F. (2002). *Towards a network of digital business ecosystems fostering the local*

development. *Ecosystems* (p. 23). Retrieved from http://www.digital-ecosystems.org/doc/discussionpaper.pdf

Nachira, F., Nicolai, A., Dini, P., Rivera Lèon, L., & Le Louarn, M. (2007). Digital Business Ecosystems. In F. Nachira, A. Nicolai, P. Dini, L. Rivera Lèon, & M. Le Louarn (Eds.), *Digital Business Ecosystems* (pp. xi–xvi). Office for Official Publications of the European Communities.

Nguyen, L. N. L., & Barth, M. (2006). Improving Automatic Vehicle Location Efficiency through Aperiodic Filtering. *2006 IEEE Intelligent Transportation Systems Conference*. doi:10.1109/ITSC.2006.1706755

Papadoglou, N., & Stipidis, E. (2001). Investigation for a Global AVL System. *IEEE Transactions on Intelligent Transportation Systems*, *2*, 121–126. doi:10.1109/6979.954545

Park, T., & Lee, S. (2004). A Bayesian Approach for Estimating Link Travel Time on Urban Arterial Road Network. In *Computational Science and Its Applications--ICCSA 2004* (pp. 1017–1025).

Patnaik, J., Chien, S., & Bladikas, A. (2004). Estimation of Bus Arrival Times Using APC Data. *Journal of Public Transportation*, *7*, 1–20.

Porcino, D. (2001). Location of third generation mobile devices: a comparison between terrestrial and satellite positioning systems. *IEEE VTS 53rd Vehicular Technology Conference, Spring 2001. Proceedings (Cat. No.01CH37202)*, *4*. doi:10.1109/VETECS.2001.944146

PwC Digital Pulse. (2014). Better transportation customer experiences through digital: Part four.

Qi, L. (2008). Research on intelligent transportation system technologies and applications. In *Proceedings - 2008 Workshop on Power Electronics and Intelligent Transportation System, PEITS 2008* (pp. 529–531). doi:10.1109/PEITS.2008.124

Recknagel, F. (2001). Applications of machine learning to ecological modelling. *Ecological Modelling*, *146*, 303–310. doi:10.1016/S0304-3800(01)00316-7

Reuters, T. (2014). THE WORLD IN 2025 10 PREDICTIONS OF INNOVATION. Retrieved October 12, 2014, from http://sciencewatch.com/sites/sw/files/m/pdf/World-2025.pdf

Riter, S., & McCoy, J. (1977). Automatic vehicle location—An overview. *IEEE Transactions on Vehicular Technology*, *26*. doi:10.1109/T-VT.1977.23649

Schweiger, C. (2003). Customer and media reactions to real-time bus arrival information systems. *Transportation Research Board, Report No*, *48*.

Science Applications International Corperation. (2003). *Oregon Regional Intelligent Transportation Systems (ITS) Integration Program Final Phase III Report: Transit Tracker Information Displays*. California-Oregon. Retrieved from http://www.fta.dot.gov/documents/final_COATS_Phase_III_reportmemo_(rev).doc

Smith, B. L., & Demetsky, M. J. (1994). Short-term traffic flow prediction: Neural

network approach. *Transportation Research Record*, 98–104.

Stanley, J., & Briscoe, G. (2010). The ABC of digital business ecosystems. *Communications Law*, *15*, 12–25.

Stephanedes, Y. J., Kwon, E., & Michalopoulos, P. (1990). ON-LINE DIVERSION PREDICTION FOR DYNAMIC CONTROL AND VEHICLE GUIDANCE IN FREEWAY CORRIDORS. *Transportation Research Record*, *1287*, 11–19.

Sun, D., Luo, H., Fu, L., Liu, W., Liao, X., & Zhao, M. (2007). Predicting Bus Arrival Time on the Basis of Global Positioning System Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2034*, 62–72.

Thangavelul, A., Bhuvaneswari, K., Kumar, K., Senthilkumarl, K., & Sivan, S. (2007). Location Identification and Vehicle Tracking using VANET. In *IEEE-ICSCN 2007,* (pp. 112–116).

Thomas, T., Weijermars, W., & Van Berkum, E. (2010). Predictions of Urban volumes in single time series. *IEEE Transactions on Intelligent Transportation Systems*, *11*, 71–80. doi:10.1109/TITS.2009.2028149

Trains Transport Sydney. (2014). Timetables - Sydney Trains. Retrieved October 06, 2014, from http://www.sydneytrains.info/timetables/#landingPoint

Transport For London. (2014). Transport For London. Retrieved October 06, 2014, from https://www.tfl.gov.uk/

TriMet. (2002). *Transit Tracker Evaluation*.

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of clinical epidemiology*, 49(11), 1225-1231.

Vanajakshi, L., Subramanian, S. C., & Sivanandan, R. (2009). Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses. *IET Intelligent Transport Systems*, *3*, 1. doi:10.1049/iet-its:20080013

Vaupel, T., Seitz, J., Kiefer, F., Haimerl, S., & Thielecke, J. (2010). Wi-Fi positioning: System considerations and device calibration. In *2010 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2010 - Conference Proceedings*. doi:10.1109/IPIN.2010.5646207

Wall, Z., & Dailey, D. J. (1999). An algorithm for predicting arrival time of mass transit vehicle using automatic vehicle location data. In *Transportation Research Board 78th Annual Meeting*.

Want, R., Hopper, A., Falcão, V., & Gibbons, J. (1992). The active badge location system. *ACM Transactions on Information Systems*. doi:10.1145/128756.128759

Welch, G., & Bishop, G. (2006). An Introduction to the Kalman Filter. *In Practice*, *7*, 1–16. doi:10.1.1.117.6808

Williams, B. M., & Hoel, L. A. (2003). Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results. *Journal of Transportation Engineering*. doi:10.1061/(ASCE)0733-947X(2003)129:6(664)

Xu, L., Zhang, S., Quan, J., & Lin, X. (2009). Vehicle Positioning Using Wi-Fi

Networks and GPS/DR System. In *Mobile Ad-hoc and Sensor Networks, 2009. MSN'09. 5th International Conference on* (pp. 287–293).

Yilmaz, I. (2010). Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. *Environmental Earth Sciences*, 61(4), 821-836.

Zaki, M., Ashour, I., Zorkany, M., & Hesham, B. (2013). Online Bus Arrival Time Prediction Using Hybrid Neural Network and Kalman filter Techniques. *International Journal of Modern Engineering Research*, *3*(4), 2035–2041.

Zandbergen, P. A. (2009). Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning. In *Transactions in GIS* (Vol. 13, pp. 5–25). doi:10.1111/j.1467-9671.2009.01152.x

Zeimpekis, V., Giaglis, G. M., & Lekakos, G. (2002). A taxonomy of indoor and outdoor positioning techniques for mobile location services. *ACM SIGecom Exchanges*. doi:10.1145/844351.844355