Using Bluetooth Detectors to Monitor Urban Traffic Flow with Applications to Traffic Management

by

Mohsen Hajsalehi Sichani

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the
requirements for the degree of
Doctor of Philosophy
in Computer Science.

Victoria University of Wellington
2020

Abstract

A comprehensive traffic monitoring system can assist authorities in identifying parts of a road transportation network that exhibit poor performance. In addition to monitoring, it is essential to develop a localized and efficient analytical transportation model that reflects various network scenarios and conditions. A comprehensive transportation model must consider various components such as vehicles and their different mechanical characteristics, human and their diverse behaviours, urban layouts and structures, and communication and transportation infrastructure and their limitations. Development of such a system requires a bringing together of ideas, tools, and techniques from multiple overlapping disciplines such as traffic and computer engineers, statistics, urban planning, and behavioural modelling. In addition to modelling of the urban traffic for a typical day, development of a large-scale emergency evacuation modelling is a critical task for an urban area as this assists traffic operation teams and local authorities to identify the limitations of traffic infrastructure during an evacuation process through examining various parameters such as evacuation time. In an evacuation, there may be severe and unpredictable damage to the infrastructure of a city such as the loss of power, telecommunications and transportation links. Traffic modelling of a large-scale evacuation is more challenging than modelling the traffic for a typical day as historical data is usually available for typical days, whereas each disaster and evacuation are typically one-off or rare events. Damage due to a disaster, combined with a sudden increase of demand due to the evacuation of people will likely result in increased pressure on the remaining, potentially fragmented, infrastructure. The lessons learnt from evacuation modelling can assist traffic operation teams and local authorities to provide safer and more efficient planning. The development of pervasive personal digital devices such as phones, watches, and headphones which

can be interconnected with technologies such as Bluetooth, has led to a disruptive change in the ways in which local governments can monitor traffic flows within their cities. Moreover, modern vehicles and navigation systems can interconnect to the personal devices of drivers and passengers primarily via Bluetooth technologies. By continuously monitoring such devices when they are discoverable and in range, traffic patterns can be estimated based on, not only the volume of detection, but also other characteristics of the devices that can be used to give more refined estimates of the real underlying traffic flows. This thesis examines Bluetooth traffic data collected from Bluetooth Traffic Monitoring Systems (BTMS) for modelling and monitoring the urban traffic. BTMS can monitor and track individual detected vehicles through a city. Installation, processing, data transmission, and maintenance of BTMS are easier, quicker and cheaper than existing standard monitoring systems such as CCTV cameras and inductive loops. Inductive loops are typically point-wise traffic monitoring systems that are installed in the roads and can measure the traffic flow. However, the use of BTMS devices presents several challenges: not every vehicle has a detectable device, some have many, and there are devices carried by pedestrians and non-motor vehicles as well as stationary devices. This thesis enumerates and investigates these challenges through statistical modelling, various protocols for cleaning and data preparation, dynamic estimation of the detection rate, and simulation through the case study of the city of Wellington, New Zealand. The city of Wellington experienced damage from the 2016 Kaikoura earthquake (a magnitude 7.8 earthquake), which led to road closures and other infrastructure damage. As part of modelling, performance evaluation, and identifying impacted routes by the 2016 Kaikoura earthquake, this thesis analyses three weeks of BTMS data from the periods before and after the earthquake. Furthermore, this thesis proposes a multi-disciplinary dynamic traffic modelling (TFDA2M) framework and evaluates the performance of TFDA2M on various large-scale evacuation scenarios. These scenarios cover a wide range of real-world use cases which may occur during a disaster such as power failure, an abrupt increase in demand, and damage to the main transportation infrastructure. The findings of this thesis highlight an immediate need for preparations of a

large-scale evacuation planning for Wellington to mitigate the consequences of a large-scale evacuation due to a future disaster.

Moreover, TFDA2M can assist traffic operation managers and authorities in making smarter decisions (both quantitative and spatially) through the simulation process. Since TFDA2M has a flexible schema, it can be set to monitor, assess, and manage the traffic flow on a daily basis and disaster occasions.

Acknowledgments

PhD is sometimes called for "Permanent Head Damage", which may emphasis on a genuinely life-changing experience, and I was not an exception. Finishing it would not have been possible without the support and guidance that I received from many people.

I would like to express my sincere gratitude to my supervisors Prof. Richard Arnold and A/Prof. Kris Bubendorfer, who have always been there for me, passionate, motivated, kind, and supportive. Richard is a fantastic teacher and scientist, and Kris is a real motivator (also motivated me to lose 20kg!). Their guidance helped me to develop my research spirit and be a newer version of me.

I would like to thank the examination committee: FGR, ECS , Dr. Andrew Meads, A/Prof. Ian Welch, and Prof. Peter Komisarczuk for their insightful comments and encouragement. I appreciate Andrew and Ian for their detailed comments.

I would like to thank my parents (Mohammadali and Parvin) who made many sacrifices and supported me in different stages of my life, especially at the beginning of this journey.

I cannot thank enough my wife, Hanieh, for being supportive, patient, kind, and motivator. She was there with me during the most challenging moments of my life, whenever I needed her, as there were some occasions that I was lost (and will be!).

I would also like to thank amazing people at the Student Learning Centre at my university who helped me a lot at the beginning of this journey.

I would also like to thank the researchers who supported me during my re-

Contents

Glossary

BTMS Bluetooth Traffic Monitoring Systems. 6

ExtoVT The cleaning algorithm which extracts Vehicles' Trips with the purpose of identifying vehicles which follow the regular traffic flow. 41

IAT Inter-Arrival Time, the time between two consecutive arrival times. 13

Macro trips A macro route/trip is the actual origin and destination of a vehicle rather than the two consecutive detections (i.e. micro route). Also referred as macro routes. 142

Micro trips The trip between two consecutive Bluetooth sensors/counters. Also referred as micro routes. 42

MTDiBT An algorithm to detect instances where multiple devices are carried in a single vehicle (i.e. multi-tenancy). This algorithm is the initial part of ExtoVT. 41

Multi-tenancy When multiple detectable Bluetooth devices are in a single vehicle. 13

Occupancy Occupancy is the ratio of the time that a vehicle is on top of traffic sensors. 3

OD Origin-Destination, the origin and the destination of a mobile node.. 3

1

Regular traffic flow The movement pattern of the majority of Vehicles. 41

Spatial layer A spatial layer is a data layer that is referring to the boundaries of an object on a map such as properties, roads, and points. 10

SUMO A microscopic traffic simulation . 108

TFDA2M A Traffic Framework for Decision Assessment, and Modelling of a traffic. 187

Chapter 1

Introduction

Urban transportation infrastructures are national assets that are mostly owned by the government, and have a massive impact on the economical [211], environmental [247], psychological and stress level of commuters [194]. Since the growth rate of population and demand are often higher than the rate of development of new transportation infrastructures, performance degradation is a dominant issue in urban transportation and imposes billions of dollars on taxpayers [172], [82], [217]. Monitoring, strategic management, and transportation planning of urban traffic are also time-consuming, expensive, and challenging tasks, which require a strong collaboration between policymakers, traffic control, information technology (IT), and traffic experts. Any transportation plan needs to be evaluated carefully to enhance the level of service (e.g. speed and travel time) of transportation infrastructures. Increasing the performance of transportation infrastructures is impossible to accomplish without having proper data, appropriate methodology, and decent tools (i.e. tools with the capability to process and visualise large datasets). Some of the traffic data that can inform these are speed, volume, and Origin-Destination (OD) of vehicles, which directly and indirectly contribute to the distribution of demand, Occupancy and density of vehicles on the road networks. Occupancy is the ratio of the time that a vehicle is on top of traffic sensors. To make the data collection more efficient, computer engineering and telecommu-

nication technology have been integrated into traffic systems and evolved manual road surveys into Intelligent Transportation System (ITS) [27]. ITS provides better, safer, and faster travel experience to its users through collecting, processing, and distribution of traffic data from/to vehicles and transportation infrastructure [109]. Proposing a flexible and accurate transportation model for addressing these requirements can not only provide a foundation for measuring the effectiveness of existing policies and infrastructures for traffic monitoring and management, but can also contribute toward ITS, utilising current infrastructure, and optimising the development of new transportation assets. Traffic management can be categorized into two general approaches, static and dynamic:

Static traffic management approaches are easier to implement such as by assigning traffic lights with a fixed phase but suffer from serious drawbacks like the lack of prioritisation of emergency vehicles, real-time information about the available capacity of transportation infrastructures and real-time monitoring of the traffic flow.

Dynamic traffic management systems through real-time traffic data collection are more challenging to implement but provide real-time information about road networks and demand. These systems also enhance the traffic flow through prioritization, dynamic traffic assignments, and dynamic traffic signs. In addition, such dynamic management systems could identify incidents or congestion in real-time and redirect traffic to alternative routes to mitigate the impact of an incident on the performance of road networks. Furthermore, these systems can update the traffic light phases based on the current demand and the distribution of demand.

Although dynamic traffic management is more complicated, it reduces the time and resource wastage represented by congestion, and increases the lifespan and the level of service of infrastructure assets [149]. These contribute to enhancement in freight and supply chains, reduction in the levels of stress of commuters, better air quality, a greener environment, and higher commuters satisfaction. Therefore, a dynamic traffic management system is an important part of an urban area and it is impossible to implement an efficient urban traffic management system without having an efficient traffic monitoring system. A well-developed

urban traffic monitoring system can assist traffic operation teams and local author-
ities to monitor the current traffic flow and identify any traffic disruption events
through collecting data from various resources. These collected data after pro-
cessing can be used in a traffic management system. Traffic data can be collected
in various ways, as is described in the next subsection.

1.1 Data collection

Data collection for an urban traffic monitoring system is done in various ways
with different technologies. Some of which are briefly explained below:

- Pneumatic road tubes are rubber tubes that detect vehicles when a pres-
  sure pulse triggers them [179]. These types of monitoring devices are not
  durable, but easy and cheap to install, but not suitable for high speed roads.

- Inductive loops consist of magnetic fields, wire and counting device [181].
  A single loop can only count the vehicles, but if work in pairs can detect
  speed by measuring the time that is required to pass these pairs. These
  are installed in a road and require road closures and traffic management for
  installation and maintenance.

- Weigh-in-motion sensors are triggered by inductive loops placed before
  them and can collect some information about the weight and axles of pass-
  ing vehicles [144]. These are similar to scales but measure the weight when
  vehicles are moving. These are more suitable for designing bridges and
  pavement.

- Radar detectors measure the speed and presence of vehicles through re-
  flected and returned radio waves.

- Video cameras (e.g. CCTV) can detect passing objects (vehicles, bikes,...)
  through image processing. These devices usually require heavy processing
  engines and may not have a good performance during bad weather. These

may also need a high network bandwidth for sending captured images and videos [61].

- Bluetooth technologies have enabled a newer kind of vehicle counting mechanism and are available through using Bluetooth detectors (sensors) at fixed points along roadsides [120]. Bluetooth Traffic Monitoring Systems (BTMS) can record Bluetooth devices carried in cars (principally mobile phones, but also on-board Bluetooth devices, tablets and laptops). Since each device identifies itself uniquely by its Media Access Control (MAC) address, this kind of data also allows tracking of individual vehicles on their routes around a city and formation of time-stamped Origin-Destination (OD) matrices. Each network interface controller (e.g. Bluetooth card/adaptor, Ethernet card/adaptor) has at least one unique identifier which is called MAC address.

- Smartphones can share their locations with service providers through General Packet Radio Services (GPRS) and Long-Term Evolution (LTE) as smartphones are equipped with GPS devices [135]. GPRS and LTE are types of services which enable continuous connection to the Internet. Cellular towers can also identify the proximity of mobile phones.

Such collected data can be used in a dynamic traffic monitoring system, which can monitor the traffic flow in real-time. This is explained in the next section.

1.2   Dynamic traffic monitoring system

A dynamic traffic monitoring system is composed of three high-level components: real-time traffic data collection as input, data processing as the main engine, and sharing the findings to connected objects as the output of the system. Data can be collected through traffic monitoring systems sensors installed within the city (e.g. Bluetooth sensors, CCTV cameras) and GPS devices (i.e. navigation systems) which are connected to the Internet. Data processing must be done in powerful servers which have road networks layers, traffic models, location and phasing of

traffic lights. Outputs of the processes can be shared with GPS devices which are connected to the Internet, traffic lights, traffic management teams, other connected systems, or as feedback to the system. The Google APIs, Navman, TomTom, and Garmin are some of the well-known companies which provide real-time information about the condition of road networks to the drivers through monitoring the road-networks and users [5]. Having an accurate picture of demand overlaying the road infrastructures can contribute to an efficient traffic monitoring system for daily usage. In addition to monitoring the traffic for a typical day, an urban area also needs a disaster traffic model to support modelling of the traffic flow under pressurised conditions (e.g. large-scale evacuation), which is explained in the next section.

1.3 Disaster traffic modelling

When a disaster occurs, existing traffic behaviour (e.g. speed, demand) changes rapidly as some transportation infrastructures can be damaged or disrupted. In all disasters, rapid movement of emergency teams is important, and it may also be necessary to evacuate an entire urban area (due to different reasons such as a large fire, toxic pollutant release, a cyclone, or a tsunami), or specific group of people (e.g. volunteers, vulnerable people, work-forces who commute to city from rural areas) [101], [124], [97]. A large-scale evacuation has a wide impact on the people and increases the stress level of evacuees. A disaster may damage some parts of road networks, telecommunication infrastructures, communication backbones, and power lines, which significantly increases the load on the remaining transportation infrastructures. For example, during a disaster a power outage may be observed in some parts of a city, vehicles collide, buildings collapse on streets and consequently some roads may be inaccessible. Due to a sudden increase in demand, heavy congestion and poor performance may be observed in some parts of the road networks. These may motivate evacuees to use their GPS systems (i.e. navigation systems) for finding better alternative routes. In addition, they may make calls, video chat, post on social media (e.g. Twitter) to share their location,

status, or conditions [195], [174]. This increases the load on the communication infrastructure (e.g. cell towers) and processing units (e.g. routing algorithms, data storing and retrieval engines) significantly, which may lead to performance degradation of communication infrastructure, navigation systems, or even misdirection of the traffic flow due to the low performance of processing units. Furthermore, it is not clear how navigation systems work under such scenarios as most of the famous navigation systems are commercial and closed source. Hence, it is critical for city planners, emergency teams, and authorities to consider all the possible risks and plan ahead of time to prepare the city for such a rare scenario.

1.4   Case study, Wellington, New Zealand

Obtaining a clear picture of what may happen during a disaster and planning for a large-scale evacuation is critical in Wellington, New Zealand, as it is located close to three fault zones, isolated by water, has a dense population of $207,000$ [63] (although the Wellington region has the population of $418,000$ [10]), and plans to become a smart capital by 2030 [63]. Wellington started using Bluetooth technology for traffic monitoring. This set of characteristics has motivated this thesis to focus on Wellington City for data collection and performance evaluation of the proposed models. This thesis examines the data collected by BTMS from 2016-10-31 to 2016-11-21. The details of Wellington profile are fully discussed in Chapter 2).

1.5   Motivation

Although many researchers have contributed toward transportation modelling, evacuation planning, driving behaviours, and policies and practices, most of this research is undertaken without combining multiple required disciplines and realistic criteria into the research topic, and targeted limited locations or time due to the complexity of the process [261], [223], [249], [241]. The core motivations of my work presented in this thesis are:

1. **Dynamic urban traffic modelling**

   Development of an efficient transportation model requires a bringing together of ideas, tools, and techniques from multiple overlapping disciplines such as traffic and computer engineering, statistics, urban planning, and behavioural modelling. For example, the use of modern technologies such as Global Positioning Systems (GPS) for finding better routes (i.e. navigational assistance) during an evacuation are mostly ignored in transportation evacuation planning [261].

   It is crucial to develop a generalized and flexible urban transportation model, which can be customised to address the unique characteristics of each city. Furthermore, this system must be able to model natural disasters and emergency situations by considering real-world parameters. Generally speaking, the current transportation models for disaster management and emergency evacuations have not fully captured or completely ignored some important traffic criteria (e.g. use of navigation systems, i.e. GPS, during a large-scale evacuation) and uncertainties such as vehicle incidents and breakdowns, and stressed and aggressive driving behaviours [186], [236], [140].

2. **Economics**

   The slow growth of transportation infrastructure compared to the higher growth of urbanisation has lead to traffic infrastructure performance degradation. Congestion imposes negative impacts on commuter, environment, and economy. Due to these negative impacts, congestion reduction has great value to local governments. For example, the US lost about 120 million USD in 2011 and New Zealand loses about 1% of its Gross Domestic Product (GDP) annually [220] due to congestion. To reduce congestion on the road networks, officials apply different strategies to the transportation such as E-lane for electrical vehicles [264], intelligent parking management and traffic light controls [252], [225], ramp controls [176],[229], and toll collection systems [100]. Although governments are allocating a vast amount of resources to transportation enhancement, having an efficient dynamic traffic

monitoring system is a big technical challenge [215].

3. **Traffic modelling before and after an earthquake**

   Wellington is spatially isolated by water and mountains, located on fault
   zones, has a dense urban area and only two exit land routes. A Spatial
   layer layer is a data layer that is referring to the boundaries of an object
   on a map such as properties, roads, and points. Wellington experienced
   a magnitude 7.8 earthquake on 14 November 2016, 00:02:00 AM, NZDT.
   I use the data collected from Wellington to investigate the impact of such
   a severe earthquake on the traffic flow as little research has been done on
   modelling the traffic characteristics before and after an earthquake using
   data collected by Bluetooth devices.

In this research, I use some of these modern data sources such as Bluetooth
traffic monitoring systems, road topologies, and the Google APIs, and combine
these with traffic simulation technologies to investigate the utilisation of dynamic
traffic monitoring systems in everyday operation, and contrast this with that data
collected during a crisis.

1.6   Research questions

My research questions are:

1. **How to detect Bluetooth devices in a vehicle which follows the regular traffic flow and distinguishing this from multiple vehicles with one Bluetooth device in, in a dense urban area?**

   Bluetooth traffic monitoring systems record the time-stamped location of
   any detectable MAC addresses which can belong to, for example, a static
   device, headphone of a pedestrian, a biker, a passenger of a bus or a GPS
   (i.e. navigation system) of a vehicle. Acquiring a clear picture of the char-
   acteristic of traffic flow through these collected data is a challenging task

especially in urban areas due to the detection of multiple devices in a vehicle, variation of traffic characteristics (e.g. speed), structure of urban areas such as numerous intersections, reflection of radio frequencies due to tall buildings, and blockage of line of sight of devices by big vehicles (e.g. trucks, buses). Investigation of the traffic characteristics in highways are usually less challenging due to the dedicated road structure and restrictions of highways (e.g. pedestrian and cyclists are not allowed in highways and highways can have concrete walls). In this research, I investigate the use of Bluetooth systems for traffic flow monitoring in urban areas. In addition, by utilising the Bluetooth data, information about the condition of the road networks can be obtained. These time-stamped observations, if matched in pairs and after data cleaning, can provide a clear picture of the distribution (Origin-Destination) of the demand within the city. This information can assist traffic operations to manage the traffic more efficiently through various ways such as measuring the demand of the road networks overtimes.

2. **What are the effects of the 2016 Kaikoura earthquake on traffic flow in Wellington City? Are Bluetooth Traffic Monitoring Systems able to identify these effects?**
Wellington experienced a severe earthquake known as Kaikoura earthquake in 2016. This earthquake followed by a localised flood in some parts of Wellington disrupted and damaged the infrastructure of Wellington. My research investigates the collected data from Bluetooth devices and models the traffic characteristics based on this data. This model highlights the roads and parts of the road networks within the city which were profoundly impacted by the earthquake.

3. **How to model a large-scale evacuation framework? What are the involved parameters/settings? Evacuees will use their navigation systems in a large-scale evacuation process. Is there any unforeseen outcome? How to predict the network performance during a large-scale evacuation modelling? What are the critical routes?**

Current traffic simulation models miss or partially ignore some important elements such as driver psychology, power failure, Geographic Information Systems (GIS), or use of navigation systems, which may significantly impact the accuracy of the model. This research aims to integrate these components into a model to provide a multi-disciplinary traffic framework and evaluate the performance of road networks and transportation infrastructures under different scenarios (e.g. historical data obtained from 2016 Kaikoura earthquake) for Wellington. Some realistic scenarios such as collapse of the buildings on some urban roads in the CBD, damage to main transportation infrastructures, the impact of distribution of traffic flow on the evacuation process, closure of the roads which direct the traffic outward the city, incidents due to the errors of drivers, and outage of telecommunication networks are investigated in my research. My novel framework can also be used to assist local authorities to quickly identify parts of the road networks with poor performance. In addition, this novel framework can also help authorities to assess their solutions for the traffic enhancement (e.g. optimisation of phases of traffic lights and construction of a new street) without disrupting the actual traffic flow. Furthermore, since a road network is composed of many streets, my developed framework illustrates the performance of road networks spatially through a visualisation engine (in addition to numerical summaries), which makes understanding and interpreting the results easier and quicker.

1.7   Research contributions

1. **A quick and efficient approach for monitoring urban traffic by utilising Bluetooth data**
   Data cleansing and outlier detection of collected Bluetooth traffic data is usually a time-consuming task and requires historical traffic data for model calibration and extraction of travel time information based on origin-destination (OD) of vehicles (i.e. demand distribution). This research develops and im-

plements an in-depth analysis of collected data and also a quick and efficient method for outlier detection along with a data management approach to accelerate the travel time extraction process.

2. **An efficient method for the detection of multi-tenancy devices in vehicles**

    Another contribution of this research is an efficient and novel algorithm for distinguishing an individual vehicle with several devices (i.e. Multi-tenancy), from several vehicles with one device. Lack of such an algorithm in the processing of the collected data overestimate the traffic density and underestimate vehicle Inter-Arrival Times (IAT). Inter-Arrival Time (IAT) is the time difference between two consecutive arrivals. This algorithm identifies public transportation vehicles such as buses, taxis and carpooling, based on flexible schema which can be configured based on various criteria.

3. **Modelling of detection of Bluetooth devices by Bluetooth sensors**

    This research examines the behaviours of Bluetooth traffic monitoring systems and statically models the detection of Bluetooth devices in vehicles. In addition, I use inferential statistical analysis to examine and model the Inter-Arrival Times (IAT). IAT modelling can be consumed in different engineering areas such as the Wireless Mesh Networks (WMN) [157], Wireless Sensor Networks (WSN) [89], and transportation.

4. **Large-scale evacuation modelling**

    This research proposes a multi-disciplinary evacuation framework which brings various key components together to provide a more realistic approach for a large-scale evacuation. I evaluate the performance of this novel disaster framework on Wellington and measure the evacuation time, flow and speed of vehicles. The outputs of this framework will assist policymakers and authorities to plan for future disasters. In this framework, I examine the impact of various important criteria such as using GPS (i.e. navigation system) in a large-scale evacuation, and driving behaviours. In addition, this framework not only identifies new challenges associated with a bulk

evacuation for cities with a dense population, but also proposes multiple solutions for addressing these challenges (e.g. optimisation of routing algorithm, placement of emergency teams, identification of roads with poor performance, changing the direction of streets, formation of a hybrid mobile network during a disaster).

1.8   Organization

The literature review in this thesis is distributed in various chapters due to the multi-disciplinary nature of this thesis (e.g. chapters three, four, five and eight have their dedicated literature review). However, Chapter two illustrates the spatial hazard layers of Wellington City and why dynamic traffic modelling and evacuation planning are critical for this city. Chapter three explains the principle of Bluetooth Traffic Monitoring Systems (BTMS) and various types of devices that may be detected by BTMS. Chapter four offers an algorithm for extraction of trips of vehicles from Bluetooth collected data. Chapter five investigates how modelling of Inter-Arrival Times (IAT) can be affected by Bluetooth traffic monitoring systems. Also, this chapter identifies how some characteristics of all passing vehicles can be estimated based on a proportion of detected vehicles by BTMS, and the approach for estimating this proportion is offered in Chapter six. Moreover, Chapter six models a digital transportation twin for traffic monitoring purposes. Chapter seven studies the impacts of the 2016 Kaikoura earthquake on the road networks of Wellington and identifies parts of the roads which were affected by the earthquake through identifying and examination of various types of vehicular trips. Chapter eight provides a literature review of overlapping disciplines such as traffic and computer engineers, statistics, urban planning, and behavioural modelling which can enhance transportation modelling. Chapter nine offers a framework for modelling the traffic flow of Wellington and evaluates the impact of a wide range of real-world scenarios on the evacuation process of Wellington, and Chapter ten concludes this thesis with some summary remarks, observations, and suggestions for future research.

Chapter 2

Wellington, an overview

This chapter presents the main features of Wellington City, highlights its characteristics, and clarifies why an appropriate emergency plan is critical for Wellington through the analysis of different spatial layers. A spatial layer is a data layer that joins the concept of 'where' to the non-spatial data. The concept of 'where' is referring to the boundaries of an object on a map such as properties, roads, and points. Processing of the spatial data is the main topic of Geographic Information Systems (GIS). Spatial indexing is a common method for querying and storing spatial data. Different processes can be done on spatial datasets such as intersection, within, and defining buffer zones.

2.1 Wellington

Wellington is the capital of New Zealand and one of the busiest urban areas of New Zealand. The total population of the Wellington region is about $418,000$ which is about 11% percent of the New Zealand population [10]. One out of four people who are living in Wellington is born overseas, which gives this city a rich cultural diversity (e.g. driving culture, communication culture). The temperature is between 19°C to 24°C in summer and 0°C to 6°C in winter. The average wind speed in Wellington is 22 km/h and the average annual rainfall is 1270mm. About

$23,000$ [67] ($44,000$ based on [10]) people commute to Wellington for work every day and 30% [67] (38% based on [10]) of people use public transportation. Wellington is isolated by water and terrain as shown in Fig. 2.1 and connects to the northern parts of the North island through limited land routes. Ferry terminals, roads, trains stations, and railways are located next to the water. Wellington can be described as a city sandwiched between mountains, hills, and sea as shown in Fig. 2.1.

2.2 Wellington spatial layers

In this section, I briefly examine some of the spatial layers of Wellington City.

2.2.1 Fault zones

Wellington is located on the fault zones as shown in figures 2.2 and 2.3. Major faults in Wellington can move vertically and horizontally, which make this city very vulnerable. The Wairarapa fault, Fig. 2.2, in 1855 moved 6 meters vertically and 15 meters horizontally [25].

2.2.2 Tsunami layer

As Wellington is isolated by water and terrain, and located on the Wellington fault zone and close to other fault zones, this city is also susceptible to a tsunami. The tsunami hazard layer is shown in Fig. 2.4. As demonstrated in this figure, all the coastal lines are located in the high-risk zone and the section which connects the east to the rest of Wellington has also a medium/low risk.

2.2.3 Flood layer

Due to the nature of hills of Wellington (high hills, mountains, and sea), the central area of Wellington has a high-risk of a flood as shown in Fig. 2.5.

Figure 2.1: Wellington aerial Imagery with main (land and sea) transportation infrastructures, Google maps

**This content is unavailable. Please consult
the print version for access.**

Figure 2.2: Wellington fault zones [3]

2.2.4   Building layer

The city centre of Wellington also has the highest buildings as visualised this Fig.
2.6 and Fig. 2.7. Fig 2.6 shows the 3D model of the distribution of heights of
buildings and Fig. 2.7 shows its heat map.

2.2.5   Power lines

Furthermore, the main power lines of Wellington are also located on the fault
zones as shown in Fig. 2.8.

2.2.6   Road layer

The road layer of Wellington is shown in Fig. 2.9. As demonstrated motorway
(SH1) passes through the city and connects inner parts of Wellington to the outer
parts.

This content is unavailable. Please consult the print version for access.

Figure 2.3: Wellington fault zones, close up, city centre [3]

This content is unavailable. Please consult the print version for access.

Figure 2.4: Wellington tsunami hazard layer [63]

This content is unavailable. Please consult the print version for access.

Figure 2.5: Wellington flood layer [63]

This content is unavailable. Please consult the print version for access.

Figure 2.6: Distribution of buildings and their heights in Wellington, 3D model, F4map

This content is unavailable. Please consult the print version for access.

Figure 2.7: Distribution of buildings and their heights in Wellington, heat map [63]

This content is unavailable. Please consult the print version for access.

Figure 2.8: Wellington main power lines [3]

Figure 2.9: Wellington road networks, Google

2.3 Conclusion, a combined view

As demonstrated Wellington is located on a hazardous zone. Ferry terminals (see Figures 2.1 and 2.10), airport, highways, and railways are located on the Wellington fault zones and tsunami hazard layer. The locations of these transportation infrastructures make them unsafe during an earthquake or tsunami. The only two routes which connect Wellington to the outer parts are located near or on the fault zones (the third route, which is also located on the fault zone, is Ohariu Valley Road, and the capacity of this route is so low that it can be discounted in this thesis). One is Burma road which is an arterial road and the other one is motorway (combination of SH1 and SH2). These are shown in Fig. 2.10. Therefore, in case of a tsunami or earthquake the ferry terminals, road networks, and rail stations may be damaged and only a proportion of vehicles can use land transportation. Evacuation may also be done by air by assuming the airport runway is undamaged as it located in the hazard zones, has enough resources, and/or that there are safe landing areas. To obtain a better picture of hazard zones, different geospatial layers (shown on the left of Fig. 2.11) are overlaid and the combined hazard layer is shown in Fig. 2.11. As shown in this figure, most of the hazardous zones are located inside the Central Business District (CBD) or on the main transportation infrastructures (ferry, exit points, airport, and railways). Moreover, CBD also has a high-risk of a local flood. Furthermore, as the main power lines of Wellington are also located on the fault zones, Wellington is susceptible to a power outage during an earthquake. This power outage can impact transportation infrastructure such as traffic lights, electric trains, communication infrastructures, and disrupt the traffic flow. Therefore, it can be concluded that the CBD is a high hazard zone due to four main reasons: located on the fault zones, high buildings may collapse during an earthquake, flood catchments, and tsunami hazards. The flood, Tsunami, Power lines, fault zones, and main transportation infrastructure layers are shown in Fig. 2.12 in one figure to show how these are correlated (use previous figures for studying map legends).

It is obvious that Wellington city (included CBD) is located on multiple haz-

ardous zones and in case of a disaster like an earthquake, the critical infrastructures of this city such as water and gas pipes, power lines, highways, and communication may be severely damaged making the evacuation of the city necessary for its residents. Therefore, it is necessary to have appropriate evacuation assessment for this city as Wellington suffers from a lack of exit-route redundancy. This evacuation assessment is impossible to achieve without an in-depth analysis of the transportation infrastructure of Wellington, which I will cover in this thesis.

This content is unavailable. Please consult the print version for access.

Figure 2.10: Wellington land and transportation infrastructures and fault zones (red lines) [25], [2]

This content is unavailable. Please consult the print version for access.

Figure 2.11: Wellington combined hazard layers [63]

This content is unavailable. Please consult the print version for access.

This content is unavailable. Please consult the print version for access.

(a) Flood layer

(b) Tsunami layer

This content is unavailable. Please consult the print version for access.

This content is unavailable. Please consult the print version for access.

(c) Wellington power lines

(d) Fault zones and transportation

Figure 2.12: Flood, Tsunami, Power lines, fault zones and main transportation infrastructure layers in one figure. Legends are available in previous figures (i.e. 2.4, 2.5, 2.8, and 2.10 .

Chapter 3

Bluetooth data

Bluetooth is a wireless technology that was invented in 1994 [127], and now is integrated into the day to day lives of people for a variety of purposes. For example, some navigation systems, smartwatches, and fitness trackers can be connected to smartphones through Bluetooth technology. Bluetooth devices use short-wave radio technology and operate at 2.4 GHz, which is an unlicensed Industrial, Scientific, and Medical (ISM) radio band. Device inquiry is a protocol that Bluetooth devices use to find nearby Bluetooth devices in proximity of several meters to hundreds of meters. Bluetooth networks are known as Piconets and composed of a master and several slaves as addressed in the 802.15 standard [83].

Bluetooth technology comes in different versions, 1 to 5. Each version shows improvements in various aspects such as compatibility, range, speed, and energy consumption [72]. The latest version (5) aims to enhance the Internet of Things (IoT) through a reliable and robust connection.

Due to the popularity of this technology, many researchers both from academia and industry adopted this technology for different purposes such as the investigation of social behaviour patterns [190], [111], modelling the connectivity between multiple devices [141], and estimation of vehicle travel times [258]. Bluetooth Traffic Monitoring Systems (BTMS), also known as Bluetooth counters, count the number of observed Bluetooth devices passing fixed locations on the road net-

works, where Bluetooth counters have been installed, by recording and storing their hashed MAC addresses in a timestamped format. Hashing allows devices to be identified and re-identified as they move between different BTMS sites, but prevents disclosure of the identity of their owners through studying their MAC addresses. Identification and re-identification of Bluetooth devices are important as these allow a vehicle to be tracked over a city and traffic flow be formed from detected vehicles.

Each Bluetooth counter installed on the road (also known as Road Side Units or RSU) communicates with servers to synchronise the collected data, time, and configurations [214]. Since BTMS are capable of detecting up to 250 devices [120] in one inquiry, they can easily detect all the vehicles which are approaching them in a typical roadside traffic setting. These Bluetooth counters can measure the traffic volume and provide operation teams with the distribution of demand over time.

The travel time for each route (two following locations) is the difference between the detection times of two consecutive observations (i.e. source (predecessor) → destination (successor)) as shown in Fig. 3.1. Since the distance between installed Bluetooth counters are known, the average speed can be calculated.

Finding the right location for the installation of Bluetooth sensors is a challenging task. Bluetooth counters should not be installed too close to each other (i.e. overlapping detection areas) as a Bluetooth device may be detected at multiple locations (Bluetooth counters) simultaneously. Furthermore, Bluetooth counters, depending on the location of their installation, can provide different levels of information. If they are installed at intersections, then the movement can be tracked at intersection level as the most granular level as typically a Bluetooth counter cannot detect the direction of detection (or approaching). This means identifying the exact road, which the Bluetooth device was on, is not always possible at an intersection level. If they are installed on a street and with the detection range limited to this street, then monitoring of the traffic can be done at street level, and the street which the Bluetooth device was on, can be identified accurately.

This content is unavailable. Please consult the print version for access.

Figure 3.1: Bluetooth traffic monitoring system [60]

3.1 Sample data of Bluetooth Traffic Monitoring Devices

Each Bluetooth counter, s, scans the surrounding (a radius of approximately 100 meters) and records the observation time, T, MAC address of the observed device, m, and some other customised attributes (depending on the type and brand). Table 3.1 provides a typical example of the collected data from a Bluetooth counter.

Table 3.1: Sample of the collected data from Bluetooth Traffic Monitoring Systems

m	T	Identifying label for location of s
ABCD30	2016-10-14 00:02:30	1
ABCD30	2016-10-14 00:15:24	2
ABCD30	2016-10-14 00:25:24	3

3.2   Types of detected Bluetooth devices

Observed Bluetooth devices can be categorised into different groups as suggested by [175], [246]:

1. Static nodes

   These nodes are located nearby an installed Bluetooth counter and are not observed in any other location for a long period of time (e.g. a printer). An example is shown in Table 3.2.

Table 3.2: An example of a static node as there are several detections at various times at the same location for the same device.

m	T	Identifying label for location of s
537ECD77B2530000	2016-10-14 00:02:09	472
537ECD77B2530000	2016-10-14 00:14:38	472
537ECD77B2530000	2016-10-14 00:18:28	472
537ECD77B2530000	2016-10-14 00:26:39	472
537ECD77B2530000	2016-10-14 00:34:00	472
537ECD77B2530000	2016-10-14 00:38:10	472

2. Regular vehicles

   These types of vehicles do not stop for a recognisable time anywhere between two consecutive locations and follow the mainstream of traffic (other vehicles). Table 3.3 provides an example.

Table 3.3: Two sequential observations of a vehicle that most likely moved continuously (i.e. without stopping) between the two detection locations. The average travel time for this route (484− >479) for this hour is about 45 seconds, which is similar to the travel time of this device (42 seconds). Therefore, it is possible that this detected device belongs to a vehicle (or was in a vehicle) which follows the mainstream traffic flow.

m	T	Identifying label for location of s
432B00E174480000	2016-11-06 14:36:32	484
432B00E174480000	2016-11-06 14:37:14	479

3. Stopped/Missed/rerouted vehicles

   These types of vehicles stop somewhere between two consecutive locations for a reasonably long time and do not follow the mainstream traffic flow, are missed at a Bluetooth counter, or have taken an alternative route. Table 3.4 illustrates this with an example. In this table the average travel time for this route is about 300 seconds, while the travel time for this specific MAC address was several hours.

Table 3.4: Two sequential observations of a vehicle (possibly) that most likely stopped between the two detection locations. This MAC address (i.e. hashed MAC), 1BC1BBC04ABD0000, has been detected at location 477 at time 09:12:44 and detected at location 622 at time 15:15:46. The average travel time for this route for a vehicle is about 5 minutes while this is about 6 hours for the owner of this device. This suggests that the owner of these devices may have stopped somewhere between these two locations.

m	T	Identifying label for location of s
1BC1BBC04ABD0000	2016-10-17 09:12:44	477
1BC1BBC04ABD0000	2016-10-17 15:15:46	622

4. Pedestrians and Cyclists

   These types of people usually do not travel at the average speed of vehicles *unless* in a heavily congested area, but Bluetooth counters detect their Bluetooth devices.

5. Bikes and electric bikes

   Bikes, electric bikes and e-scooters are other categories of a transportation mode that have similar movement characteristics (e.g. speed) to vehicles especially in urban areas, if riders respect the traffic rules (these cannot be separated from vehicles and probably should not).

3.3   Limitation of Bluetooth devices

Although Bluetooth counters can provide real-time traffic data, there are some important drawbacks associated with them as follows.

3.3.1   Only detect a proportion of passing vehicles

Not every vehicle is carrying an active and discoverable Bluetooth device. Therefore, Bluetooth sensors only capture a proportion of actual traffic flow. Without validating against other traffic data, the proportion of vehicles that carry Bluetooth devices cannot be measured. This detection rate varies over time in each region and depends on the proportion of people who use Bluetooth technology, type of antenna, placement of Bluetooth counters, and location of passing devices. The detection rate varies from about 2% to about 30% in different years and countries.

- O'Neil et al. in [197] reported about 7.5% detection rate for Bath, UK in 2006.

- Nicolai et al. in [190] reported 2%, and 6% in 2007 for Bremen, Germany and San Francisco, US respectively.

- ITERIS research group from Minnesota Department of Transportation in 2011 reported 3% to 11% for North America in [74].

- Young in 2012 reported 5% in [259] for Maryland State, US.

- Araghi et. al. [79] reported estimation of 27% to 29% in the same year, 2012. In 2015, the detection rate in Japan was reported from 2.7% to 6% [240],[192].

The variation of the detection rate indicates that this can be a localisation parameter. In addition, due to security reasons, some Bluetooth devices change to non-discoverable mode if they are not in use [99] for a period of time. Not all Bluetooth devices reply to inquiry packets when they are paired and connected to another Bluetooth device [4]. Currently, neither Navman [46] and Garmin [11]

navigation systems support multiple Bluetooth connections, which affects the detection rate of vehicles which carry these devices.

3.3.2 Uncertainties in the detection of the type of Bluetooth devices

Correct categorisation of the detected passing objects can be a challenging task as addressed in [175], [246], [242], [115], [182], [180]. For example, it is hard to distinguish a Bluetooth device, which belongs to a running person, from a slowly moving vehicle in a heavily congested area. This is even more difficult when historical data is not available as the traffic trends can not be identified through examining the historical data.

3.3.3 Route uncertainty

The other challenge with the collected BTMS data is that if multiple routes exist between two consecutive Bluetooth counters, the actual taken route by a vehicle can not be distinguished easily if Bluetooth sensors are not installed at all the inter between intersections. This concept is illustrated in Fig. 3.2. A driver may choose an alternative route which is less busy, but longer, and due to lack of congestion has a similar travel time. This is an important issue if the collected data is used for identifying any traffic disruption event on the route.

3.3.4 Multi-tenancy detection

Multi-tenancy refers to identifying more than one Bluetooth device in a vehicle. Uncertainty in distinguishing a vehicle type (e.g. buses), which has several Bluetooth devices, from several vehicles, which are travelling together and have one Bluetooth device each, is another challenging issue which has not received enough attention in the previous research. Table 3.5 shows the collected data with a high possibility of being a public vehicle.
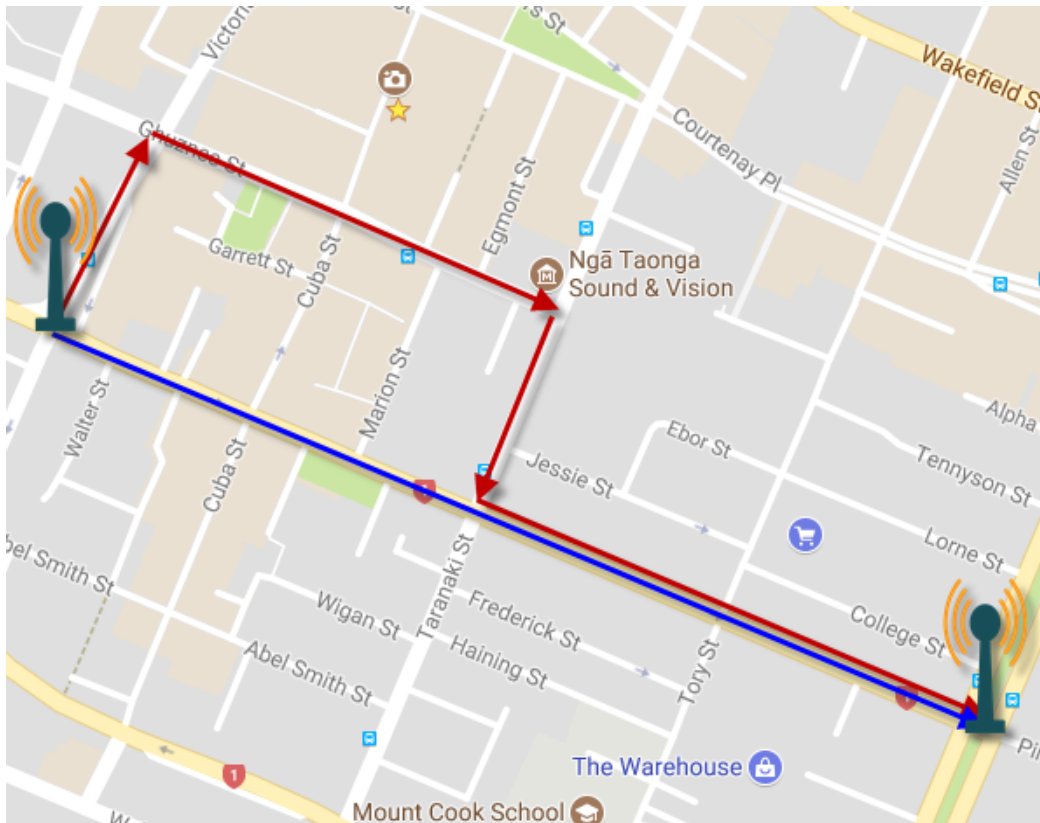
Figure 3.2: Uncertainty in route detection, Google Maps

Table 3.5: Possible public vehicle. As shown by the second and third columns, these three devices (three different hashed MAC addresses), are detected as the same locations (473 and 521) and times (19:33:50 and 19:42:32). Therefore, it is possible that these are travelling together in the same vehicle.

m	T	Identifying label for location of s
537ECD77BF7B0000	2016-11-18 19:33:50	473
537ECD77BFEA0000	2016-11-18 19:33:50	473
537ECD77BFCB0000	2016-11-18 19:33:50	473
537ECD77BF7B0000	2016-11-18 19:42:32	521
537ECD77BFEA0000	2016-11-18 19:42:32	521
537ECD77BFCB0000	2016-11-18 19:42:32	521

3.3.5 Privacy issues

As it is possible to track the movement of a node in a city where BTMS are installed, so personal behaviours of the owners of these mobile nodes might be detected especially when detected MAC addresses can be associated with a person. In other to be compliance with current privacy protocols (e.g. New Zealand ACT 1993), BTMS typically depersonalise the data. This can be done through a hashing function which enables re-identification of devices in a city but depersonalises the data, so they can not be tracked at the personal level.

3.3.6 Other issues

Other resources such as [173], [213], [214] reported issues with providing Bluetooth counters with continuous external power supply, clock synchronisation, and being unable to determine the exact location of observed Bluetooth devices.

3.4 Conclusion

This chapter investigated various types of nodes that can be detected by BTMS data, which are as follows: static nodes, vehicles which follow the regular traffic flow, Stopped/Missed/rerouted vehicles, Pedestrians and Cyclists, Bikes, and

electric bikes. Although some drawbacks are associated with BTMS data such as route uncertainty and clock synchronisation, but BTMS is popular technology and can provide an efficient traffic monitoring system for estimation of travel times.

Chapter 4

ExtoVT and MTDiBT

4.1 Introduction

Bluetooth traffic monitoring devices record any detected Bluetooth devices, which can belong to a vehicle, printer, smartwatch, etc. Identifying and removing unrelated traffic data, vehicles which stop or park between two consecutive detections for a recognisable time, and obtaining an accurate travel-time (TT) are still challenges associated with the utilisation of BTMS for real-time traffic monitoring as addressed by [110], [175], [246], [242], [115], [182], [180]. In this chapter, I introduce an algorithm for Multi-Tenancy Detection in the collected Bluetooth data (MTDiBT). MTDiBT is an algorithm to detect instances where multiple devices are carried in a single vehicle (i.e. multi-tenancy as this single vehicle carries multiple Bluetooth devices). It is of course important to remove such instances, since otherwise any system will overestimate the traffic density and underestimate vehicle inter-arrival times.

In addition, I develop a novel approach for Extraction of Vehicles' Trips (ExtoVT). These extracted vehicles are types of vehicles which follow the Regular traffic flow. I define the regular traffic flow (or stream or pattern) as vehicles which do not stop for a long and recognisable period between two consecutive detections in a typical condition. Equivalently, these are the vehicles which form the

41

typical movement patterns of the traffic flow. In this regard, emergency vehicles which move very quickly, or heavy trucks which may move very slowly, compared to other vehicles, are also considered as outliers as they are significantly different in terms of movement with the majority of vehicles.

MTDiBT is discussed before the ExtoVT algorithm, as ExtoVT utilises this algorithm for Multi-tenancy detection. In addition, MTDiBT can be used independently of ExtoVT to identify public vehicles (mass transit vehicles such as buses and taxis) in a transportation system.

The rest of this chapter is organised as follows. The next section is dedicated to a literature review for this chapter. MTDiBT is discussed in section 4.3. Section 4.4 discusses the details of ExtoVT. Section 4.5 validates this algorithm on a small part of a road, and its results are shown in Section 4.6. Section 4.7 generalises this approach for the whole urban area of Wellington City as a case study and its results are provided in the last section.

4.2 Literature review of the cleaning process

BTMS can provide an origin-destination (OD) matrix as they can track individual vehicles carrying detectable Bluetooth devices through a city. A simplified OD matrix is shown in Fig. 4.1 [40]. An OD matrix demonstrates the cost of the route from an origin (i.e. source) to a destination. For example, the cost of moving from D to E is 20. The cost can be measured in various ways such as travel time (e.g. seconds), distance (e.g. km), or the number of stops on a route. In this chapter, the default cost is travel time, and it is measured in seconds. This origin-destination might be different from the actual origin and destination of drivers. I define origin and destination in this context as the detection of a Bluetooth device at two consecutive locations (or time if it is a static node). The first detection is considered as the origin (source), and the second is considered as the destination. Therefore, the actual origin and destination of a driver may be composed of multiple smaller origin-destination pairs (or Micro trips) in BTMS. Every two consecutive detections can be used for travel-time calculation as the distance between these two

Bluetooth counters is known. Fig 4.2 shows how travel-time can be calculated.

This content is unavailable. Please consult the print version for access.

Figure 4.1: An example of OD matrices [40] which shows the cost of moving from a location (i.e. source) to another location (destination). The cost can be measured in various ways such as travel time or distance. The default is travel time in this thesis.

Identifying and removing outliers in the traffic data collected by BTMS are challenging tasks and previously investigated by other researchers. The authors in [78] proposed an approach to obtain the speed of a passing object at two consecutive locations and compare this with the expected speed as an option to only filter vehicles which follow the regular traffic flow. This approach has a lower accuracy in congested urban areas as, for example, cyclists may have similar speed to moving motor vehicles [110]. Southwest Research Institute in [142] and [106] proposed a similar algorithm called TransGuide, which is based on the travel time. TransGuide uses historical travel times, forms a range for each origin and destination, and excludes the travel times which are not in this range. Then it estimates the travel time by applying a rolling window on the remaining of the travel times. The authors in [110] suggested using classes of Bluetooth devices, identifiable by MAC address, to separately identify hands-free Bluetooth devices from others. This approach may provide a simple and efficient method for removing non-vehicle devices, but a drawback is removing Bluetooth devices of a vehicle

**This content is unavailable. Please consult
the print version for access.**

Figure 4.2: An example of BTMS [110]

which does not have a hands-free Bluetooth device and therefore preventing the
convenient usage of the multitudes of portable devices (such as phones) in the
monitoring of traffic flow. Araghi et al. in [79] investigated the minimum, max-
imum, average, and median of collected travel time in 15 minute periods on a 5
km motorway with a tunnel and reported the minimum and median travel times
(TT) approaches are reliable and trustworthy. The minimum travel time shows
the fastest speed and is not affected by slower vehicles or vehicles which do not
follow the regular traffic flow. The median is less affected by outliers, which are
unusually fast or slow vehicles. The approach in this study cannot be applied to
urban areas because in an urban area the frequency of different types of observed
Bluetooth devices is higher than a tunnel as usually in long tunnels (e.g. a 5 km
tunnel), pedestrians and cyclist are not allowed. Therefore, there is less ambiguity
in the detection of devices of vehicles which follow regular traffic compared to
other types of devices. Jang in [145] reported an algorithm based on the modified
median and standard deviation for the Dedicated Short-Range Communication
(DSRC), like Bluetooth frequencies, on rural highways. It should be noted since
highways are usually have dedicated routes and pedestrians and cyclist are not al-
lowed, the collected data have fewer outliers compared to dense urban areas like

Wellington City. Dion et al. in [107] assumed Lognormal distribution for travel time and calculated smoothed means and standard deviations of travel times from Exponential smoothing within a moving window. This moving window does not have a fixed temporal size, but rather a fixed number of required observations defines its size. This approach requires accurate estimates of the traffic flow parameters and obtaining these values are not always possible as reported by Jang in [145] and ITSKorea in [143].

4.3 Multi-tenancy detection problem

Since Bluetooth counters are capable of detecting multiple Bluetooth devices simultaneously (up to about 250 devices) [120], an algorithm to detect instances where multiple devices are carried in a single vehicle - a situation I refer to as multi-tenancy, is required. It is of course important to remove such instances, since otherwise the Bluetooth collected data overestimate the traffic volume and underestimate vehicle inter-arrival times after processing. The first subsection explains a very high level and simplified design of Multi-Tenancy Detection in the collected Bluetooth data (MTDiBT) algorithm as this novel algorithm uses advanced programming techniques (e.g. classes), and the next subsection explains the core idea of MTDiBT. The performance of MTDiBT is covered at the end of this section.

4.3.1 High-level design of MTDiBT

For multi-tenancy detection, the first step is looking at the data on an hourly basis. Many people leave in the morning and return to their homes in the afternoon. Public transport services offer most of their services during day-time [57]. Furthermore, traffic flows are highest from about 07:00 to about 20:00 [146], [53] (also is shown in Fig 7.19). Therefore, most Bluetooth devices are expected to be seen during these hours, and a daily basis (i.e. daily cycle) approach provides a granular level for the investigation of how often Bluetooth devices are carried

together.

MTDiBT identifies when Bluetooth devices are seen together (i.e. in a single vehicle) when the difference in their observation time is equal or less than a specified threshold time interval τ. This threshold, τ, is set to 1 second in my study as the granularity of data collection is 1 second. To have a confirmed detection, which suggests Bluetooth devices are moving together, at least two consecutive detections are required and in both cases a detector must register a putative pair of devices within the threshold interval of τ. This utilises the fact that these detected devices were together at least on one micro trip (source-destination). In this chapter, a micro trip is defined as a movement between the location of two consecutive Bluetooth counters with the detections being consecutive (i.e. the devices are not separately observed by any other Bluetooth counter during the micro-trip). The source is the location of the first Bluetooth counter, and the destination is the location of the second Bluetooth counter. Two consecutive detections (source-destination) are not necessarily sufficient to be certain as this means these Bluetooth devices were together only in one micro trip. This is because streets have several lanes and vehicles in a city may move slowly and close to one another. Therefore, two is not a good choice for this threshold. The number of required consecutive detections, α, is set to 3, so that at least two micro-trips (i.e. source-destination/source-destination) must be seen before a pair of devices are classified as being in the same vehicle. Consider a case when two devices, $m1$ and $m2$, were seen together at the first detection, later they were seen individually where the absolute difference of their times of observation is greater than τ seconds, and finally, they were seen together. Since these two MAC addresses were not seen together at two consecutive observations, they cannot be together in a micro trip. Hence, they are not in a single vehicle. A real use case is carpooling when people share a vehicle toward and outward from the city centre, but not within the city. However, these parameters, α and τ can have different settings for different cities with various conditions such as the expected frequency of Bluetooth devices, the number of vehicles equipped with Bluetooth technologies, population, and usage of Bluetooth technology to optimise the multi-tenancy detection algorithm. These

parameters can be identified through conducting field tests (explained in more depth in Chapter 6) or sensitivity analysis (similar to 4.3.5).

4.3.2 MTDiBT algorithm

The first goal is to identify all possible combinations of devices that might be travelling together. This is done by forming clusters of devices, and then seeing which clusters are seen consecutively at multiple locations. There are many possible ways to group the detections, particularly because there is a need to enable a rolling window in time of length τ, so a device may be associated with devices at a temporal distance of τ earlier or later than its own detection. At the start of the process each device detection may be assigned to more than one cluster, however as MTDiBT searches for persistent associations, it successively eliminates most of these putative associations and only a small number of clusters which might genuinely be sets of devices travelling in a single vehicle are left.

In this context, an upper-case letter, a lower case letter, and *DT* are references to the attribute, values for each attribute, and data table, respectively. This data structure is shown in Table 4.1 as this table has all the data. D, S, T, M, I are attributes for date, detection location (i.e. location of a sensor), time, MAC address, and unique detection identifier. A unique identifier is a value uniquely identify each row of data. MTDiBT defines a list as ℓ, which holds the values individual values of an attribute. For example, ℓ_D contains different value of dates (e.g. d_0).

Table 4.1: DT_{all}, sample collected data by BTMS

Date(D)	Location(S)	Time(T)	MAC (M)	Identifier(I)
d_1	s_0	10	m1	1
d_1	s_0	10	m2	2
d_1	s_0	11	m3	3
d_1	s_0	12	m4	4
d_1	s_1	20	m2	5
d_1	s_1	21	m3	6
d_1	s_3	31	m3	7
d_1	s_3	31	m4	8
d_1	s_4	45	m4	9
d_1	s_4	45	m5	10
d_1	s_5	70	m4	11
d_1	s_5	71	m5	12

The relation between a Bluetooth device and a MAC address (m) is typically a one-to-one relation, as most of Bluetooth devices have only one Bluetooth card, and each Bluetooth card has only one MAC address.

The high-level flow chart of this algorithm is illustrated in Fig. 4.3, and the Pseudo-code of the MTDiBT in algorithm 1. This Pseudo-code is explained below and some examples are provided after this. For a non-programmer reader starting with the example is suggested before reading the explanation.

Explanation

MTDiBT initially selects all unique dates in the collected data (DT_{all}). Then, for each unique date, d, it selects all unique locations, s. Then in the first Step, this algorithm selects all unique times (t)s from the data for each date and location and saves these (t)s in ℓ_T.

In Step 2, for each $t \in \ell_T$ this algorithm defines a rolling window with the size of $[t, t + \tau]$ and selects all observations for this date and location where they were seen within this time frame, $[t, t + \tau]$.
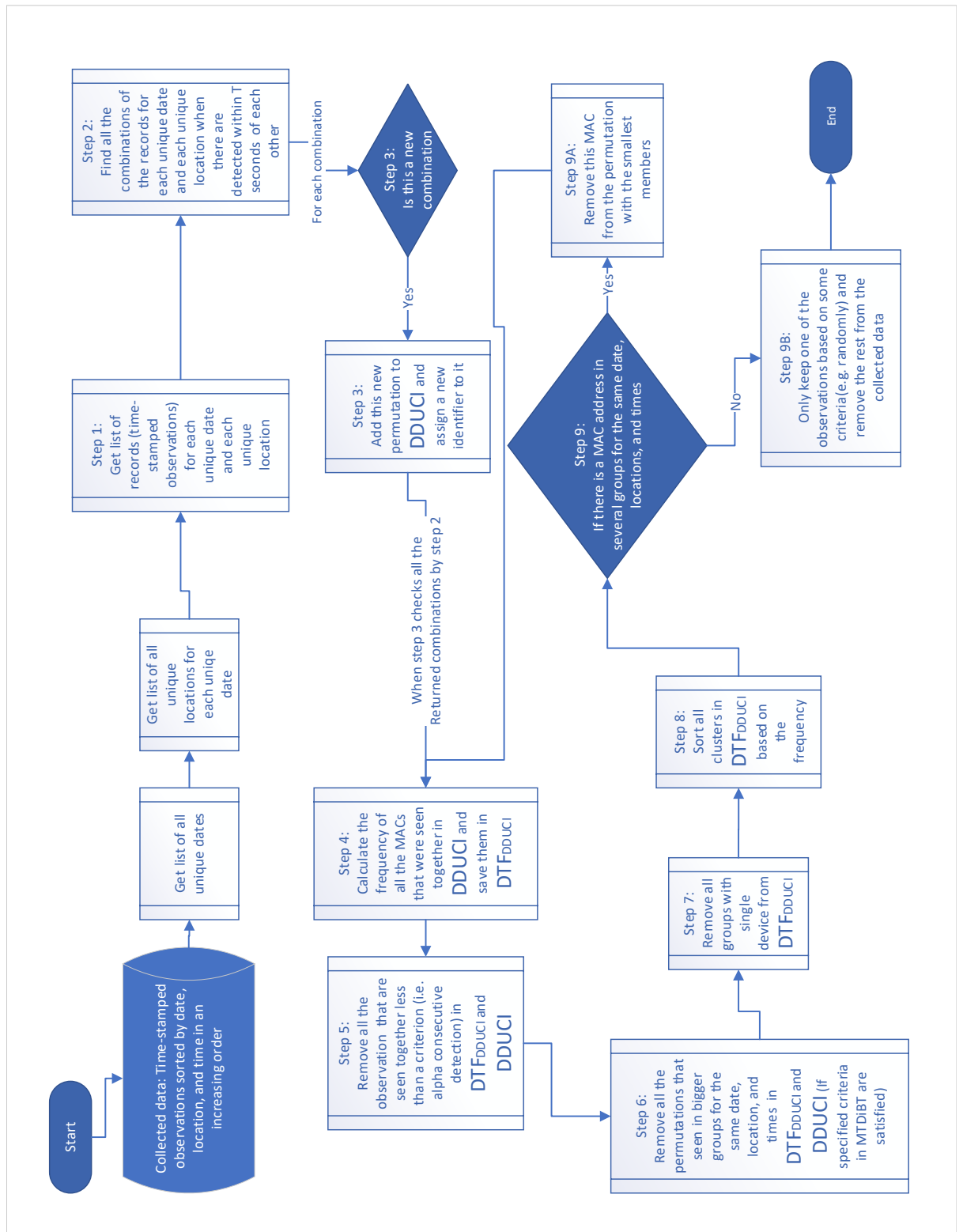
Figure 4.3: High-level flowchart of MTDiBT algorithm

Algorithm 1: MTDiBT algorithm

Data and definitions:
α: Number of consecutive detections
τ: Time difference in seconds
DDUCI: Data dictionary
DT_{all}: Data table of all the collected observations as
$(D_{d=0,\ldots}, S_{s=0,\ldots}, T_{t=0,\ldots}, M_{m=0,\ldots}, I_{i=0,\ldots})$
, where D, S, T, M are days, locations, observation times, and observed devices, respectively, which are sorted increasingly based on the D, S, T, M. Each sorted row gets a unique identifier as I.

Procedure:
foreach *Unique* $d \in DT_{all}$ **do**
    foreach *Unique s from* DT_{all} *where* $D_{DT_{all}} = d$ **do**
        step 1: ℓ_T=select unique t from DT_{all} where $D_{DT_{all}} = d$ and $S_{DT_{all}} = s$
        foreach $t \in \ell_T$ **do**
            Step 2: $DT_t^{d,s}$=select all records from DT_{all} where $t \leq t_{DT_{all}} \leq t + \tau$, $D_{DT_{all}} = d$ and $S_{DT_{all}} = s$
            Step 3: Find all the combinations of the observations in $DT_t^{d,s}$ and pass these combinations to AUCI($\{\ell_D, \ell_S, \ell_T, \ell_M, \ell_I\}$)
        end
    end
    Step 4: From *DDUCI*, calculate the frequency of each group of MACs (m)s, with a list of their (UCI)s and form data table DTF_{DDUCI}
    Step 5: From DTF_{DDUCI} and *DDUCI*, remove all clusters that seen in less than α consecutive detection
    Step 6: From DTF_{DDUCI} and *DDUCI*, remove all clusters which have seen in a bigger cluster with the same d, s, t, if (frequency of the bigger cluster $\geq$ frequency of the smaller cluster and all (m)s of the smaller cluster exists in the bigger cluster)
    Step 7: Remove all cluster with a single device.
    Step 8: Sort all clusters in DTF_{DDUCI} and *DDUCI* based on the frequency
    9A:**if** *several clusters share an observation (or several) for the same d,s,t, m, and frequency (achieved by joining* DTF_{DDUCI} *and DDUCI through UCI)* **then**
        Remove this shared detection from the clusters with smaller sizes and only keep it in the biggest cluster. If all have the same size, remove this detection from all of them except one randomly. Apply steps 4 to 7 if required.
    end
    else
        9B:for each remaining clusters remove all of the detection from DT_{all} except for one of them randomly, and consider this remained observation as the representative of this group (i.e. vehicle). Apply steps 4 to 7 if required.
    end
end
Function AUCI ($\{\ell_D, \ell_S, \ell_T, \ell_M, \ell_I\}$) **:**
    if *this* $\{\ell_D, \ell_S, \ell_T, \ell_M, \ell_I\}$ *exists in DDUCI* **then**
        ;
    end
    else
        generate a new identifier (UCI) as $UCI = Max(UCI \in DDUCI_{\{\{\ell_D, \ell_S, \ell_T, \ell_M, \ell_I\}, UCI\}}) + 1$, add this new UCI and $\{\ell_D, \ell_S, \ell_T, \ell_M, \ell_I\}$, to $DDUCI_{\{\{\ell_D, \ell_S, \ell_T, \ell_M, \ell_I\}, UCI\}}$
    end

In Step 3, MTDiBT finds all combinations of these observations, and assigns a unique identifier to each combination, and save them in a data dictionary $DDUCI_{\{\{\ell_D,\ell_S,\ell_T,\ell_M,\ell_I\},UCI\}}$, if does not already exist in *DDUCI*.

Step 4 extracts unique ℓ_Ms from *DDUCI*, which each ℓ_M is a list of devices that were seen together within $[t,t+\tau]$ time period for the same date and location, finds the frequency of each ℓ_M, and make a list of their (*UCI*)s and form a data table, DTF_{DDUCI}. Therefore, DTF_{DDUCI} contains three columns: the first column is the cluster of possible co-detected devices, the second column is the frequency of ℓ_M, and the third column is a reference to *DDUCI* for each of these clusters (ℓ_M). Because DTF_{DDUCI} keeps *UCI* in the second column, then this table can be joined with *DDUCI* by using *UCI*. In addition, because *DDUCI* keeps the identifier, *I*, of each detection, *DDUCI* can also be joined with DT_{all} by using *I*. Therefore, these two tables, DTF_{DDUCI} and DT_{all}, can be joined together.

Step 5 removes the MAC addresses from DTF_{DDUCI} and *DDUCI* that they were seen in less than α consecutive detections. This means all the entities where their frequency is lower than α are also removed.

In Step 6, MTDiBT removes smaller clusters that were seen in bigger clusters from DTF_{DDUCI} and *DDUCI*, if the bigger cluster contains all the devices of the smaller cluster, with the same date, location, time, and frequency.

In Step 7, this algorithm removes all remaining clusters with one device from DTF_{DDUCI} and *DDUCI* as they were not accompanied by any other device.

In Step 8, MTDiBT sorts the remaining clusters based on their frequencies to increase the performance for the rest of the steps.

Later, for DTF_{DDUCI} and *DDUCI*, in Step 9*A*, if any of the remaining clusters share a (or several) detection (a MAC address with its associated observed metadata (date, location, time)) with the same frequency, MTDiBT applies one of the two following actions based on the size of these clusters:

1. If the sizes of the clusters are not the same, then the cluster with the biggest size keeps this shared detection and it is removed from smaller clusters.

2. If the sizes of the clusters are the same, then this shared device (detection)

is removed from one of them randomly.

At the end, this algorithm applies Steps 4 to 7. In Step 9*B* If clusters do not share an observation, then this algorithm removes all the observations in a cluster except one of them randomly and applies Steps 4 to 7. This is done as these were in a single vehicle and only one of them must be considered. In this case, does not matter which one is a public vehicle and which one is a passenger if the purpose is only removing multi-tenancy detections. Otherwise, the shared element in different clusters is a public transport vehicle and the other ones are passengers.

Examples

MTDiBT is explained through some examples as follows.

Let's assume $\alpha = 2$ and $\tau = 1$ second, and the observations in Table 4.1 were collected from BTMS and sorted based on D, S, T, M.

A unique set of dates from Table 4.1 is selected as a list, ℓ_{D_U}. In this example there is only one date, and $\ell_{D_U} = \{d_1\}$. Then, for each unique date, which is d_1 in this example, a list of unique locations from observed data on that date, Table 4.1, are identified in the list as $\ell_{S_U(D_U)}$ in a **For loop**. In this case for $D_{DT_{all}} = \{d_1\}$, unique locations are $\{s_0, s_1, s_3, s_4, s_5\}$.

In Step 1, MTDiBT algorithm selects all unique times, (t)s, from Table 4.1 where $D_{DT_{all}} = d$ and $S_{DT_{all}} = s$ and they are saved in list ℓ_T in another **For loop**. This gives ℓ_T as $\{10, 11, 12\}$ for d_1 and s_0 for the first run of the **For loop**, ℓ_T as $\{20, 21\}$ for d_1 and s_1 for the second run of the *For loop*, ℓ_T as $= \{31\}$ for d_1 and s_3 for the third run of the *For loop*, ℓ_T as $\{45\}$ for d_1 and s_4 for the fourth run of the *For loop*, and $\{70, 71\}$ for d_1 and s_5 for the fifth run of the *For loop* in this example.

In Step 2, (in the second **For loop**) for each $t \in \ell_T$ this algorithm defines a rolling window with the size of $\{t, t + \tau\}$ and selects all observations from Table 4.1, where $D_{DT_{all}} = d$, $S_{DT_{all}} = s$ and $t \leq t_{DT_{all}} \leq t + \tau$, and saves them in $DT_t^{d,s}$. This finds all the possible MAC addresses which are seen together when their difference in their observation time is less than τ for the same date and location.

In Step 3, (still in the second **For loop**) MTDiBT finds all of the combinations (or clusters) of the observations and assigns a unique identifier to each, and saves them in $DDUCI_{\{\{\ell_D,\ell_S,\ell_T,\ell_M,\ell_I\},UCI\}}$. *DDUCI* is a data dictionary for holding the attributes of each cluster and a Unique cluster's identifier. A data dictionary, $<Key,Value>$, is a data structure which contains multiple key and value pairs. Each key must be unique in the whole dictionary, but duplicate values are allowed. An example is shown below for data dictionary d_{sample}.

$d_{sample}=\{\{1,Apple\},\{2,Orange\},\{3,Orange\}\}$,

where d_{sample} has three elements as $\{1,Apple\},\{2,Orange\}$, and $\{3,Orange\}$. For the first element the key is 1 and value is Apple, for the second element, the key is 2 and the value is Orange, and for the third element the key is 3, and the value is Orange. In this dictionary, all the keys are unique, but similar values exist (i.e. Orange). The key and value could also be a complex data structure or class of objects. For example, an array of arrays can be used as a key or value. The first element in *DDUCI*, $\{\ell_D,\ell_S,\ell_T,\ell_M,\ell_I\}$, is the key. This key is a class, which is composed of five lists: Date (ℓ_D), Location (ℓ_S), Times (ℓ_T), MACs (ℓ_M), and Identifiers (ℓ_I)s based on Table 4.1. This combination is always unique and that is the reason MTDiBT saves it as the Key in *DDUCI*. Although for this algorithm all dates and locations are the same and no list is required, but due to consistency this algorithm saves these two values in lists rather than single variables. Each collected row (cluster) has a Unique Cluster Identifier (*UCI*). This Value, *UCI*, is a 64-bit incremental integer (similar to a primary key or a unique identifier). Assigning *UCI* to a cluster is done by calling the function Assign Unique Cluster Identifiers (*AUCI*). This guarantees that the Value in *DDUCI* is also unique. Therefore, in *DDUCI* both Keys and Values are unique.

Based on the above steps, (see the first three steps of algorithm 1), MTDiBT operates as:

For $DT_{10}^{d_1,s_0}$ the rolling window is $\{10,11\}$. Then,

$DDUCI_{\{\{\ell_D,\ell_S,\ell_T,\ell_M,\ell_I\},UCI\}}=$
$\{$

$\{\{\{d_1\}, \{s_0\}, \{10\}, \{m1\}, \{1\}\}, \{1\}\},$
$\{\{\{d_1\}, \{s_0\}, \{10\}, \{m2\}, \{2\}\}, \{2\}\},$
$\{\{\{d_1\}, \{s_0\}, \{11\}, \{m3\}, \{3\}\}, \{3\}\},$
$\{\{\{d_1, d_1\}, \{s_0, s_0\}, \{10, 10\}, \{m1, m2\}, \{1, 2\}\}, \{4\}\},$
$\{\{\{d_1, d_1\}, \{s_0, s_0\}, \{10, 11\}, \{m1, m3\}, \{1, 3\}\}, \{5\}\},$
$\{\{\{d_1, d_1\}, \{s_0, s_0\}, \{10, 11\}, \{m2, m3\}, \{2, 3\}\}, \{6\}\},$
$\{\{\{d_1, d_1, d_1\}, \{s_0, s_0, s_0\}, \{10, 10, 11\}, \{m1, m2, m3\}, \{1, 2, 3\}\}, \{7\}\},$
$\}$

When $t = 11$, then the rolling window is $\{11, 12\}$, then AUCI does not add this
$\{d_1\}, \{s_0\}, \{11\}, \{m3\}, \{3\},$
to the *DDUCI* as it exists in *DDUCI* with the Value, *UCI*, of 3, but the followings are added to the *DDUCI*

$\{\{\{d_1\}, \{s_0\}, \{12\}, \{m4\}, \{4\}\}, \{8\}\}$
$\{\{\{d_1, d_1\}, \{s_0, s_0\}, \{11, 12\}, \{m3, m4\}, \{3, 4\}\}, \{9\}\}$

When $t = 12$, then the rolling window is $\{12, 13\}$. Since this key,
$\{\{d_1\}, \{s_0\}, \{12\}, \{m4\}, \{4\}\},$
exists in $DDUCI_{\{\{\ell_D, \ell_S, \ell_T, \ell_M, \ell_I\}, UCI\}}$ with the Value of 8, this Key is not added.

For $DT_{20}^{d_1, s_1}$, then rolling window is $\{20, 21\}$, following items are to the $DDUCI_{\{\{\ell_D, \ell_S, \ell_T, \ell_M, \ell_I\}, UCI\}}$

$\{\{\{d_1\}, \{s_1\}, \{20\}, \{m2\}, \{5\}\}, \{10\}\},$
$\{\{\{d_1\}, \{s_1\}, \{21\}, \{m3\}, \{6\}\}, \{11\}\},$
$\{\{\{d_1, d_1\}, \{s_1, s_1\}, \{20, 21\}, \{m2, m3\}, \{5, 6\}\}, \{12\}\},$

When $t = 21$, then the rolling window is $\{21, 22\}$ nothing new is added to $DDUCI_{\{\{D, S, T, I\}, UCI\}}$.

For $DT_{31}^{d_1, s_3}$, the rolling window is $\{31, 32\}$ and the followings are added to $DDUCI_{\{\{D, S, T, I\}, UCI\}}$.

$\{\{\{d_1\}, \{s_3\}, \{31\}, \{m3\}, \{7\}\}, \{13\}\},$
$\{\{\{d_1\}, \{s_3\}, \{31\}, \{m4\}, \{8\}\}, \{14\}\},$
$\{\{\{d_1, d_1\}, \{s_3, s_3\}, \{31, 31\}, \{m3, m4\}, \{7, 8\}\}, \{15\},$

For $DT_{45}^{d_1, s_4}$, the rolling window is $\{45, 46\}$ and the followings are added to $DDUCI_{\{\{D,S,T,I\}, UCI\}}$.

$\{\{\{d_1\}, \{s_4\}, \{45\}, \{m4\}, \{9\}\}, \{16\}\},$
$\{\{\{d_1\}, \{s_4\}, \{45\}, \{m5\}, \{10\}\}, \{17\}\},$
$\{\{\{d_1, d_1\}, \{s_4, s_4\}, \{45, 45\}, \{m4, m5\}, \{9, 10\}\}, \{18\}\},$

For $DT_{70}^{d_1, s_5}$, the rolling window is $\{70, 71\}$ and the followings are added to $DDUCI_{\{\{D,S,T,I\}, UCI\}}$.

$\{\{\{d_1\}, \{s_5\}, \{70\}, \{m4\}, \{11\}\}, \{19\}\},$
$\{\{\{d_1\}, \{s_5\}, \{71\}, \{m5\}, \{12\}\}, \{20\}\},$
$\{\{\{d_1, d_1\}, \{s_5, s_5\}, \{70, 71\}, \{m4, m5\}, \{11, 12\}\}, \{21\}\},$

Therefore, at the end of step three, $DDUCI_{\{\{\ell_D, \ell_S, \ell_T, \ell_M, \ell_I\}, UCI\}}$ is:

$DDUCI_{\{\{\ell_D, \ell_S, \ell_T, \ell_M, \ell_I\}, UCI\}}=$
$\{$
$\{\{\{d_1\}, \{s_0\}, \{10\}, \{m1\}, \{1\}\}, \{1\}\},$
$\{\{\{d_1\}, \{s_0\}, \{10\}, \{m2\}, \{2\}\}, \{2\}\},$
$\{\{\{d_1\}, \{s_0\}, \{11\}, \{m3\}, \{3\}\}, \{3\}\},$
$\{\{\{d_1, d_1\}, \{s_0, s_0\}, \{10, 10\}, \{m1, m2\}, \{1, 2\}\}, \{4\}\},$
$\{\{\{d_1, d_1\}, \{s_0, s_0\}, \{10, 11\}, \{m1, m3\}, \{1, 3\}\}, \{5\}\},$
$\{\{\{d_1, d_1\}, \{s_0, s_0\}, \{10, 11\}, \{m2, m3\}, \{2, 3\}\}, \{6\}\},$
$\{\{\{d_1, d_1, d_1\}, \{s_0, s_0, s_0\}, \{10, 10, 11\}, \{m1, m2, m3\}, \{1, 2, 3\}\}, \{7\}\},$

$\{\{\{d_1\},\{s_0\},\{12\},\{m4\},\{4\}\},\{8\}$
$\{\{\{d_1,d_1\},\{s_0,s_0\},\{11,12\},\{m3,m4\},\{3,4\}\},\{9\}\}$

$\{\{\{d_1\},\{s_1\},\{20\},\{m2\},\{5\}\},\{10\}\},$
$\{\{\{d_1\},\{s_1\},\{21\},\{m3\},\{6\}\},\{11\}\},$
$\{\{\{d_1,d_1\},\{s_1,s_1\},\{20,21\},\{m2,m3\},\{5,6\}\},\{12\}\},$

$\{\{\{d_1\},\{s_3\},\{31\},\{m3\},\{7\}\},\{13\}\},$
$\{\{\{d_1\},\{s_3\},\{31\},\{m4\},\{8\}\},\{14\}\},$
$\{\{\{d_1,d_1\},\{s_3,s_3\},\{31,31\},\{m3,m4\},\{7,8\}\},\{15\}\},$

$\{\{\{d_1\},\{s_4\},\{31\},\{m4\},\{9\}\},\{16\}\},$
$\{\{\{d_1\},\{s_4\},\{31\},\{m5\},\{10\}\},\{17\}\},$
$\{\{\{d_1,d_1\},\{s_4,s_4\},\{45,45\},\{m4,m5\},\{9,10\}\},\{18\}\},$

$\{\{\{d_1\},\{s_5\},\{70\},\{m4\},\{11\}\},\{19\}\},$
$\{\{\{d_1\},\{s_5\},\{71\},\{m5\},\{12\}\},\{20\}\},$
$\{\{\{d_1,d_1\},\{s_5,s_5\},\{70,71\},\{m4,m5\},\{11,12\}\},\{21\}\},$
$\}$

Thus, $DDUCI_{\{\{\ell_D,\ell_S,\ell_T,\ell_M,\ell_I\},UCI\}}$ contains all the possible permutations of de-vices that can be considered as co-detections, or simply all the possible combina-tions of clusters.

In Step 4, MTDiBT calculates the frequency of lists of MAC addresses, ob-tained from $DDUCI_{\{\{\ell_D,\ell_S,\ell_T,\ell_M,\ell_I\},UCI\}}$, and lists their (UCI)s to form data table DTF_{DDUCI} (Table 4.2). This is done to prepare the table for the next step. The output is shown in Table 4.2.

Table 4.2: DTF_{DDUCI}

$\{M\}$	Frequency	(UCI)s
$\{m1\}$	1	$\{1\}$
$\{m2\}$	2	$\{2, 10\}$
$\{m3\}$	3	$\{3, 11, 13\}$
$\{m1, m2\}$	1	$\{4\}$
$\{m1, m3\}$	1	$\{5\}$
$\{m2, m3\}$	2	$\{6, 12\}$
$\{m1, m2, m3\}$	1	$\{7\}$
$\{m4\}$	4	$\{8, 14, 16, 19\}$
$\{m3, m4\}$	2	$\{9, 15\}$
$\{m4, m5\}$	2	$\{18, 21\}$
$\{m5\}$	2	$\{17, 20\}$

To be able to address each row individually in this chapter, a column is added to following Tables 4.3, 4.4, 4.5, and 4.7.

Step 5 removes all rows from Table 4.2, where *Frequency* $< (\alpha = 2)$. For example, $\{m1, m2, m3\}$ are only seen in one location, so they cannot be any candidate for co-detections. This keeps the following rows (Table 4.3). This table (Table 4.3), contains all the clusters that are seen in at least $\alpha = 2$ detections as sub-condition of $\alpha = 2$ consecutive detections.

Table 4.3: DTF_{DDUCI}, where *Frequency* ≥ 2

$\{M\}$	Frequency	(UCI)s	Row
$\{m2\}$	2	$\{2, 10\}$	1
$\{m3\}$	2	$\{3, 11, 13\}$	2
$\{m2, m3\}$	2	$\{6, 12\}$	3
$\{m4\}$	4	$\{8, 14, 16, 19\}$	4
$\{m3, m4\}$	2	$\{9, 15\}$	5
$\{m5\}$	2	$\{17, 20\}$	6
$\{m4, m5\}$	2	$\{18, 21\}$	7

Furthermore, Step 5 also removes all the clusters which are observed in less than $\alpha = 2$ consecutive detections, by joining DTF_{DDUCI} and $DDUCI$ through *UCI*s and *I*s. This is accomplished by a reverse search using *I*. *I* is a unique

identifier which various data structures can be joined together if they keep it as a Foreign key (i.e. a pointer, which is the case here). DTF_{DDUCI} keeps the UCIs, and these UCIs are pointing to the associated records in *DDUCI*. *DDUCI* has the *I*s that point to the records of DT_{all}, therefore it is possible to track an observation from DTF_{DDUCI} to the actual collected data (i.e. Table DT_{all}).

Based on Step 5, cluster $\{m3, m4\}$ has to be removed as they are seen at two consecutive observations. $m4$ is not seen with $m3$ at s_1 with a time difference of fewer than τ seconds. Precisely, no observation is recorded for $m4$ at s_1 within time interval $[21-1, 21+1]$. This operation remains the following clusters (Table 4.4):

Table 4.4: DTF_{DDUCI}, $Frequency \geq 2$ and for 2 consecutive detections

$\{M\}$	Frequency	(UCI)s	Row
$\{m2\}$	2	$\{2, 10\}$	1
$\{m3\}$	2	$\{3, 11, 13\}$	2
$\{m2, m3\}$	2	$\{6, 12\}$	3
$\{m4\}$	4	$\{8, 14, 16, 19\}$	4
$\{m5\}$	2	$\{17, 20\}$	6
$\{m4, m5\}$	2	$\{18, 21\}$	7

Based on Step 6, by joining DTF_{DDUCI} and *DDUCI* through UCIs, MTDiBT finds $m2$ is seen at $UCI = \{2, 10\}$, and $m3$ is seen in $UCI = \{3, 11, 13\}$ in *DDUCI*. When $m2$ is seen at $UCI = 2$, $m3$ is seen at $UCI = 3$ and when $UCI = 6$, both $m2$ and $m3$ are seen together in a bigger cluster for the same d, s, t. Therefore, $UCI = 2$ and $UCI = 3$ should be removed. $UCI = 12$ is also a bigger cluster for $UCI = 10$ and $UCI = 11$. Hence, $UCI = 10$ and $UCI = 11$ should be removed. The same story applies to rows 4 and 6 in Table. 4.4. This was Step 6, and after applying Step 7 and removing single member clusters, the output is shown in Table 4.5:

Table 4.5: DTF_{DDUCI} at the end of Step 7

$\{M\}$	Frequency	(UCI)s	Row
$\{m2, m3\}$	2	$\{6, 12\}$	3
$\{m4, m5\}$	2	$\{18, 21\}$	7

So two clusters with the same frequency (Step 8) remain as follows: $\{m2,m3\}$, $\{m4,m5\}$

Since these two clusters do not share a MAC address m, then step 9B should be applied. For each cluster one of the m(s), must be kept randomly (say $m3$ and $m4$). Hence, $m2$ and $m5$ have to be eliminated from DT_{all} for the occasions they are clustered, which are $I \in \{2,5\}$ for $m2$, and $I \in \{10,12\}$ for $m5$. This is possible as this algorithm preserves (I)s as a list in $DDUCI$. Is are stored as values in data dictionary $DDUCI$ and they are pointing to the collected data (i.e. DT_{all}).

Step 9*A* was not considered in this example for simplicity (sharing a detection, i.e. a MAC address exist in different clusters). This situation is illustrated in the next example. Imagine the data shown in Table 4.6 was observed, and assume the same thresholds: $\alpha = 2$ and $\tau = 1$ second.

Table 4.6: DT_{all}, example 2

Date(D)	Location(S)	Time(T)	MAC (M)	Identifier(I)
d_1	s_0	10	m1	1
d_1	s_0	10	m2	2
d_1	s_0	11	m3	3
d_1	s_0	12	m4	4
d_1	s_3	20	m1	5
d_1	s_3	20	m2	6
d_1	s_3	21	m3	7
d_1	s_3	22	m4	8

The output of this algorithm at Step 8 is:

$DDUCI_{\{\{\ell_D,\ell_S,\ell_T,\ell_M,\ell_I\},UCI\}}=$

{

$\{\{\{d_1,d_1,d_1\},\{s_0,s_0,s_0\},\{10,10,11\},\{m1,m2,m3\},\{1,2,3\}\},\{UCI_i\}\}$,

$\{\{\{d_1,d_1\},\{s_0,s_0\},\{11,12\},\{m3,m4\},\{3,4\}\},\{UCI_{i+1}\}\}$,

$\{\{\{d_1,d_1,d_1\},\{s_3,s_3,s_3\},\{20,20,21\},\{m1,m2,m3\},\{5,6,7\}\},\{UCI_{i+2}\}\}$,

$\{\{\{d_1,d_1\},\{s_3,s_3\},\{21,22\},\{m3,m4\},\{7,8\}\},\{UCI_{i+3}\}\}$,

}

and the frequency is (Table 4.7):

Table 4.7: DTF_{DDUCI}, at the end of step 8

$\{M\}$	Frequency	(UCI)s	Row
$\{m1, m2, m3\}$	2	$\{UCI_i, UCI_{i+2}\}$	R_i
$\{m3, m4\}$	2	$\{UCI_{i+1}, UCI_{i+3}\}$	R_{i+1}

Since both clusters have the same frequency, and shared detections for a single device ($m3$ is observed in both clusters equally, $D = d_1, S = s_0, T = 11$, and $D = d_1, S = s_3, T = 21$), condition 9A is satisfied. Hence, the biggest cluster, which is $\{m1, m2, m3\}$, is kept and $m3$ is removed from the other cluster. This leaves the second cluster ($\{m3, m4\}$) as $\{m4\}$. $\{m4\}$ needs to be removed based on Step 7. Therefore, only one cluster is acceptable here as $\{m1, m2, m3\}$. However, another approach is to consider the frequency of the observation of $m3$ with both clusters ($m1, m2$ and $m4$) for other days, and cluster $m3$ with the devices that were seen more often, which requires higher processing power and time.

4.3.3   Possible limitation

This algorithm assumes all devices in a vehicle are detectable at all the times at each BTMS, and if a vehicle has multiple detectable devices, all of these devices should be detected in τ time interval of each other. Therefore, if the line of sight of some devices is blocked and some of these devices cannot be detected, then these devices will not be identified as co-detections. Therefore, the solution is releasing α from the condition of **consecutive**. This means if devices are detected together at any α locations during a day with the observation difference of τ seconds of each other, then they are considered as co-detections. Both of these criteria, consecutive and non-consecutive, are considered and evaluated during the execution of this algorithm.

4.3.4 Example from collected data

This section provides a couple of examples from the collected BTMS data from Wellington City.

Table 4.8 shows two devices which joined together at location 598 and travelled together to location 592 (i.e. $5 \leq I \leq 10$). Prior and after these time they were travelling individually.

Table 4.8: DT_{all}, example 3. As coloured by the yellow and green, these two devices (536B70FE00FA, CF52C6BEA1BEB) are detected by MTIBT as co-detections and possibly travelling together from location 598 to 592.

Location label	Date and time	Device	Identifier
638	2016-11-09 15:56:28	536B70FE00FA	1
472	2016-11-09 16:08:38	536B70FE00FA	2
645	2016-11-09 16:10:38	536B70FE00FA	3
594	2016-11-09 16:17:49	CF52C6BEA1BEB	4
598	2016-11-09 16:22:22	536B70FE00FA	5
598	2016-11-09 16:22:23	CF52C6BEA1BEB	6
597	2016-11-09 16:24:45	536B70FE00FA	7
597	2016-11-09 16:24:45	CF52C6BEA1BEB	8
592	2016-11-09 16:25:37	CF52C6BEA1BEB	9
592	2016-11-09 16:25:38	536B70FE00FA	10
595	2016-11-09 16:37:10	CF52C6BEA1BEB	11
596	2016-11-09 16:38:47	CF52C6BEA1BEB	12
632	2016-11-09 16:44:56	CF52C6BEA1BEB	13
595	2016-11-09 17:37:48	536B70FE00FA	14
592	2016-11-09 17:43:47	536B70FE00FA	15
595	2016-11-09 17:46:55	536B70FE00FA	16

In the second example, Table 4.9, it can be observed that the first time these two devices are seen close together is at location 479 and time '2016 − 11 − 09 12 : 15 : 39' and '2016 − 11 − 09 12 : 15 : 43', $ID \in \{3, 4\}$, which one of them arrived a few seconds earlier at the meeting point, and after getting in/on the same vehicle, they were seen together for one micro trip ($ID \in \{5, 6, 7, 8\}$), and later they were separated.

Table 4.9: DT_{all}, example 4. As coloured by the yellow and green, these two devices (C36BDEA26C5CC, 0FB2BD520A7BB) are detected by MTDiBT as co-detections and possibly travelling together from location 467 to 630. As the difference of detection time in location 479 is about 4 seconds, these are not considered as code-detections at this location. This is the location where these devices are assumed to be boarded by the same vehicle.

Location label	Date and time	Device	Identifier
588	2016-11-09 12:11:14	0FB2BD520A7BB	1
477	2016-11-09 12:13:11	0FB2BD520A7BB	2
479	2016-11-09 12:15:39	C36BDEA26C5CC	3
479	2016-11-09 12:15:43	0FB2BD520A7BB	4
467	2016-11-09 12:18:26	0FB2BD520A7BB	5
467	2016-11-09 12:18:26	C36BDEA26C5CC	6
630	2016-11-09 12:21:17	C36BDEA26C5CC	7
630	2016-11-09 12:21:18	0FB2BD520A7BB	8
473	2016-11-09 12:27:25	0FB2BD520A7BB	9
473	2016-11-09 12:47:55	C36BDEA26C5CC	10
629	2016-11-09 12:53:04	C36BDEA26C5CC	11
480	2016-11-09 12:55:47	C36BDEA26C5CC	12
625	2016-11-09 12:57:57	C36BDEA26C5CC	13
645	2016-11-09 13:00:06	C36BDEA26C5CC	14
632	2016-11-09 13:56:54	C36BDEA26C5CC	15
596	2016-11-09 13:59:12	C36BDEA26C5CC	16
630	2016-11-09 14:51:12	0FB2BD520A7BB	17
477	2016-11-09 15:00:27	0FB2BD520A7BB	18

4.3.5   Outputs of MTDiBT

The total number of mobile nodes in the data collected by BTMS from 2016-10-31 to 2016-11-21 in my dataset is $2,764,516$.

By setting α to 2 and τ to 0 seconds, 385 clusters of co-detected devices were detected for the three weeks of the collected data with the total number of 1529 observations, which is less than 0.001% of my dataset.

With setting α to 2 and τ to 1 second, MTDiBT found 3311 clusters of co-detected devices, with the total number of 13148 observations, which is less than

0.01% of my dataset.

By releasing the consecutive condition from α, MTDiBT found 5126 clusters of co-detected devices.

These co-detected devices may contain bikes, pedestrians, and vehicles. Therefore, the actual number of vehicles is much lower in my collected data, and this small proportion of co-detections do not affect the estimated traffic volume and traffic characteristics. However as explained in Chapter 3, Section 3.3 due to security reasons, some Bluetooth devices change to non-discoverable mode if they are not in use [99] and not all Bluetooth devices reply to inquiry packets when they are paired and connected to another Bluetooth device [4]. Therefore, the chance of multi-tenancy detection can be very low in the collected data as this is the case in the collected data.

The next sections present ExtoVT, applies ExtoVT on a specific route as an instance, and later applies ExtoVT on a whole city.

4.4 ExtoVT

In this section I develop and implement my algorithm for Extraction of Vehicles' Trips (ExtoVT), which is composed of these below steps:

1. **Identify and remove static nodes**
   Any kind of Bluetooth device that is static between two consecutive observation times is identified as a static node and must be excluded from data. An example is shown below (Table 4.10).

Table 4.10: An example of a static node.

m	T	Identifying label for location of s
537ECD77B25300	2016-10-14 00:02:09	472
537ECD77B25300	2016-10-14 00:14:38	472
537ECD77B25300	2016-10-14 00:18:28	472
537ECD77B25300	2016-10-14 00:26:39	472

Since this hashed MAC address (537ECD77B25300) was detected in one location (i.e. 472) between two consecutive detections (rows 1 and 2, rows 2 and 3, rows 3 and 4), therefore, it must be considered as a static node between two consecutive detection times (00:02:09 and 00:14:38,00:14:38 and 00:18:28,00:18:28 and 00:26:39). So these rows for this hashed MAC address must be flagged as static nodes. Alternatively, this data can be converted to the following format as shown in Table 4.11 and if a hashed MAC address is at the same location between two consecutive detection times, then this node is considered as a static node and must be removed.

Table 4.11: An alternative way to sort the data based on the detection time (ascending) and form a micro trip matrix.

m	source location	Detection time at source	destination location	detection time at destination
537ECD77B25300	472	2016-10-14 00:02:09	472	2016-10-14 00:14:38
537ECD77B25300	472	2016-10-14 00:14:38	472	2016-10-14 00:18:28
537ECD77B25300	472	2016-10-14 00:18:28	472	2016-10-1400:26:39

2. **Find multi-tenancy objects and group them together as one**
   This algorithm was discussed in MTDiBT. As shown the ratio is too low to impact the traffic flow, although I removed the co-detections.

3. **Identify traffic disruption events**
   Any traffic disruption event such as incidents and breakdowns should be identified and excluded from the data as these may affect the traffic characteristics, the arrival times and consequently inter-arrival times. If unknown, alternatively the pessimistic and optimistic values obtained from the valid range of travel times can be used to form a baseline for identification of the disruption event.

4. **Identify a valid range for OD matrices**
   To obtain the origin-destination (OD) matrices of travel times for micro trips, and recognise vehicles that follow the regular traffic stream, this research utilises OD matrices of travel times, which are provided by Google

and are accessible by an Application Programming Interface (API) [15]. ExtoVT utilises Google API (Google travel times) along with considering speed, travel time, distance, and Tukey's logic to only identify vehicles which follows the regular traffic flow in a dense urban area especially with historical traffic data is not available. Google travel times are considered as a reliable source of information as are used by many people for arrival time estimation, and also by academic and industrial researchers. Wang and Xu in 2011 used the OD matrices and routing information provided by the Google API to update the Geographical Information Systems (GIS) software applications with real-time traffic data [248]. The authors in [198] employed Google API for geocoding and travel time in GIS software. Seymour et al. in [222] used the Google API for Dallas transportation management centre (DalTrans) to provide users with traffic conditions. Chen et al. in [91] used the Google API to provide quick and reliable transportation service in 2014. They also integrated Fuzzy logic to find available parking spaces and estimate the availability based on the traffic data obtained from the Google API. The authors in [96] also applied traffic data provided by the Google API to enhance freight delivery. As highlighted, the Google API is widely used for OD matrices of travel time estimation. To obtain an upper band for the OD matrices of travel times, ExtoVT uses the pessimistic values offered by the Google API. The Pessimistic value is defined as the maximum duration in traffic in severe traffic congestion [52]. To find a lower bound for the OD matrices of travel times, ExtoVT considers the fastest travel time as the optimistic travel time through analysing of distance and speed. The optimistic travel time is the time that a car requires to reach a destination if it moves as fast as the speed limit of the route, when all traffic lights are green, and without any interruption (e.g. deceleration due to pedestrian crossing) in its movement patterns, which are unlikely during day time.

5. **Apply Tukey's logic and remove outliers**
Along with the construction of the OD matrix, ExtoVT uses Tukey's box-plot logic to identify possible remaining outliers. Boxplot logic, also known

as Tukey's boxplot logic is commonly used for outliers detection for univariate and multivariate anomalies. Inter Quartile Range (*IQR*) is defined as the difference between the upper quartile(*Q*3) and the lower quartile (*Q*1), $Q3 - Q1$, and data instances beyond $Q3 + (1.5IQR)$ and below $Q1 - (1.5IQR)$ are considered as anomalies [103]. Laurikkala et al. [167], Horn et al. [139], Solberg et al. [226], and Fialho et al. [116] used this logic for outlier detection in medical science. The authors in [159] used this logic for traffic outlier removal to detect incidents in real-time. The authors in [151] presented and investigated the accuracy of several techniques for modelling of railway dwell times and used Tukey's boxplot logic for outlier removal. Toshniwal et. al also in [237] used this approach for outlier detection in data streaming. Proto et al. also in [212] used this method for removing outliers of an intrusion detection system.

The pseudo-code of ExtoVT is shown below.

Algorithm 2: ExtoVT algorithm

Data:
Length=$L(m)$; Speed limit=$m(/s)$; Route=rr; Date=d; Hour=h ;Travel-Time of a vehicle=TT;

Procedure:
step 1: Find and remove static nodes
step 2: Find disruptive event and exclude related traffic data
step 3: Apply MTDiBT on the collected data

foreach *h* **of** *d* **on** *rr* **do**

    1. Find the pessimistic travel time obtained from Google API as TT_p (seconds)

    2. Calculate the fastest travel time (seconds), TT_f, as follows: $TT_f = \dfrac{L_{rr}}{m_{rr}}$

    3. Remove $TT > TT_p$ and $TT < TT_f$

    4. For the remaining TT calculate $IQR = Q3 - Q1$

    5. Remove any $TT < Q1 - 1.5IQR$ and $TT > Q3 + 1.5IQR$

end

In the next section I study a small segment of a road as a use case for evaluating

the performance of ExtoVT.

4.5 Pre-processing and exploratory analysis, a use case study

As an example, a route on Adelaide Road with a length of 750 meters and the speed limit of 50 km/h from the Basin Reserve to John Street (South direction) is studied as shown in Fig. 4.4. The departure time from the Basin reserve is from 2016-11-07 00:00:00 to 2016-11-07 23:59:59. No car incidents, breakdowns, or road closures were reported on this date for this route based on the local report. Except for removing static nodes, no further processing is done on this dataset.

The frequency of travel time of this data is shown in Fig 4.5 without differentiating hours. Fig. 4.6 shows the same concept for various hours.

As shown by Fig. 4.5 and 4.6, the collected data for this segment of the road contains inaccurate travel times. These values must be removed for any further processing by applying ExtoVT as addressed in the next section.

4.6 Results, the use case study

In order to show the effectiveness of the above approach, this study applies ExtoVT on the same section of the Adelaide Road as shown in Fig. 4.4 and considers traffic for the same date from 07:00:00 to 20:59:59 as hours of interest. The optimistic travel time for this route is 54 seconds. This means 54 seconds is the fastest time that a car can reach John Street if it travels 50 km/h when all traffic lights are green and there is no interruption and declaration on its route. The pessimistic value, which is the longest time needed to reach to the destination (i.e John Street), is obtained via the Google API. After removing these outliers by applying Tukey's boxplot logic (applying ExtoVT fully), the frequency plot is drawn and shown in Fig. 4.7. Fig. 4.8 also shows the same concept on an hourly basis.

As shown in these figures, the travel times increase in the afternoon compared

to the morning as this is the time that traffic flow is heavier outward the city on the South side due to high demand and limited capacity of the roads. Due to this limited capacity and the high volume of the cars, the travel time increases and congestion appears on the road. Furthermore, there are three schools (St Marks, Wellington East Girls College, and Wellington College) at this area (Basin Reserve), which cause an increase in the congestion in the afternoon as students finish for the day.

As demonstrated by these results, ExtoVT removed the outliers and constructed a valid travel time distribution for this route by identifying vehicles which follow the regular traffic streams. To evaluate the accuracy of this approach, the mean and standard deviation (as error bar) of travel times on an hourly basis are calculated. Fig. 4.9 shows these values before applying ExtoVT algorithm (only static nodes were removed), and the results after applying ExtoVT algorithm are shown in Fig. 4.10.

4.7 Pre-processing and exploratory analysis, a bigger picture

This section examines the collected data across all of the micro-trips from Wellington City. This data is obtained from 2016-11-07 00:00:00 to 2016-11-21 23:59:59. The only processing step was removing the static nodes. Fig. 4.11 shows the frequency of travel times without differentiating hours. The same concept on an hourly basis is shown in Fig. 4.12. As shown in these figures, the travel times vary from 1 to 3161492 second, which is an unrealistic range.

4.8 Results, the bigger picture

This section examines the collected data of BTMS after applying ExtoVT. The uncleaned data contains $2,434,005$ detections consisting of $1,837,416$ mobile nodes. The cleaning process left $205,830$ vehicles which follow regular traffic

patterns. Fig. 4.13 shows the frequency of travel times without differentiating the hours and Fig. 4.14 demonstrates the same concept on an hourly basis. As demonstrated in Fig. 4.14, a thicker tail (i.e. higher frequencies) with longer travel times exist in most of the hours in the afternoon compared to the morning hours on these micro routes. This suggests that most people commute to work early in the morning but finish at different hours in the afternoon. In addition, the frequency of detected vehicles decreases during evening hours (e.g. hours 19 and 20).

Finally, Fig. 4.15 and 4.16 show the mean and standard deviation of travel times before and after applying ExtoVT on the collected data from Wellington City, respectively.
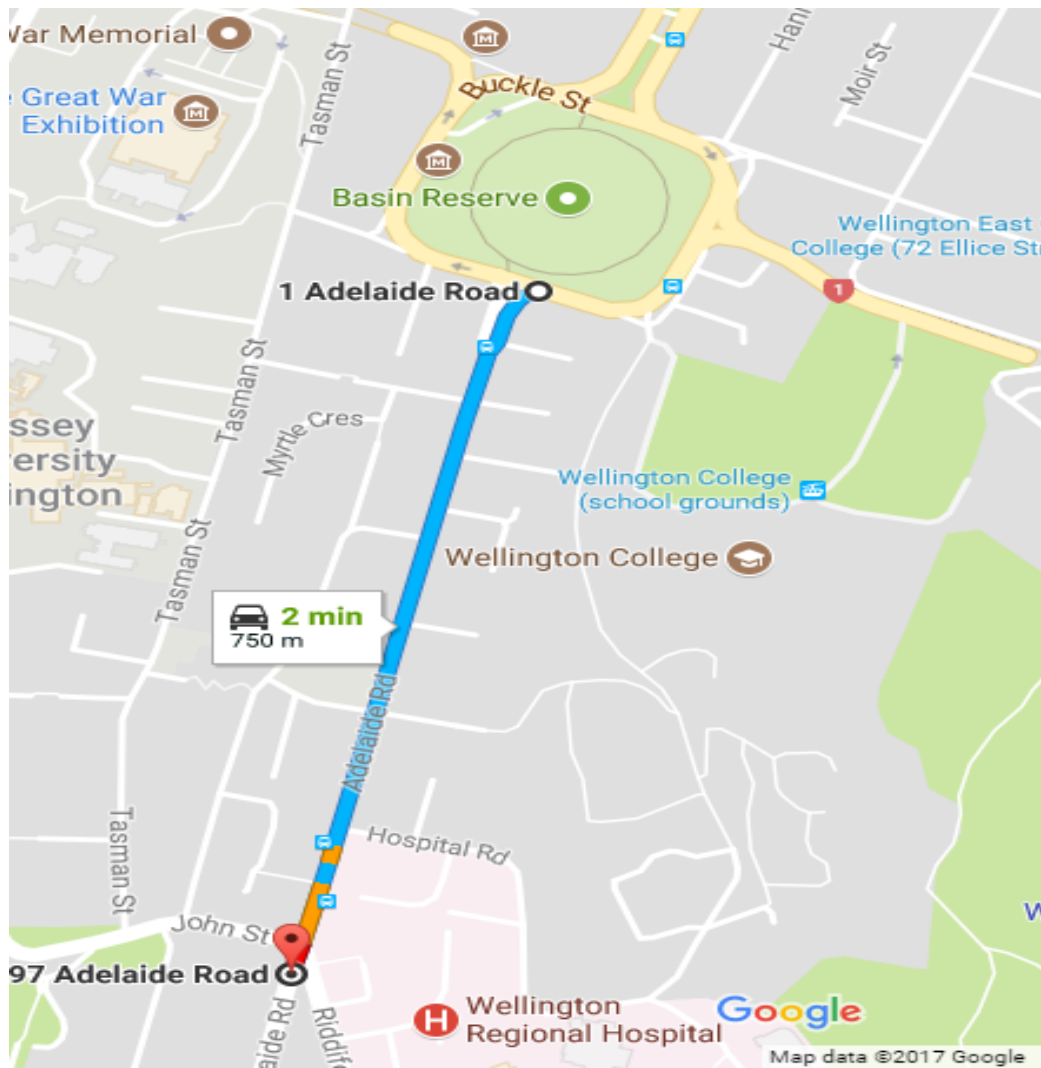
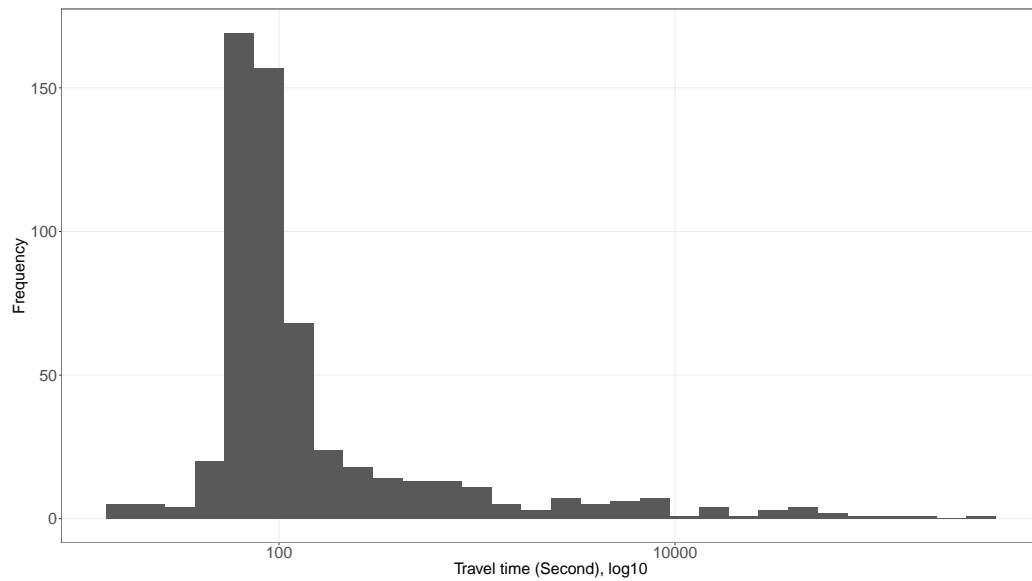Figure 4.4: A case study, Adelaide Road, South direction.

Figure 4.5: Frequency of travel time for Adelaide Road (x axis is scaled as log10), as observed, after removing static devices. This frequency plot shows a narrow tail with unrealistic travel times.
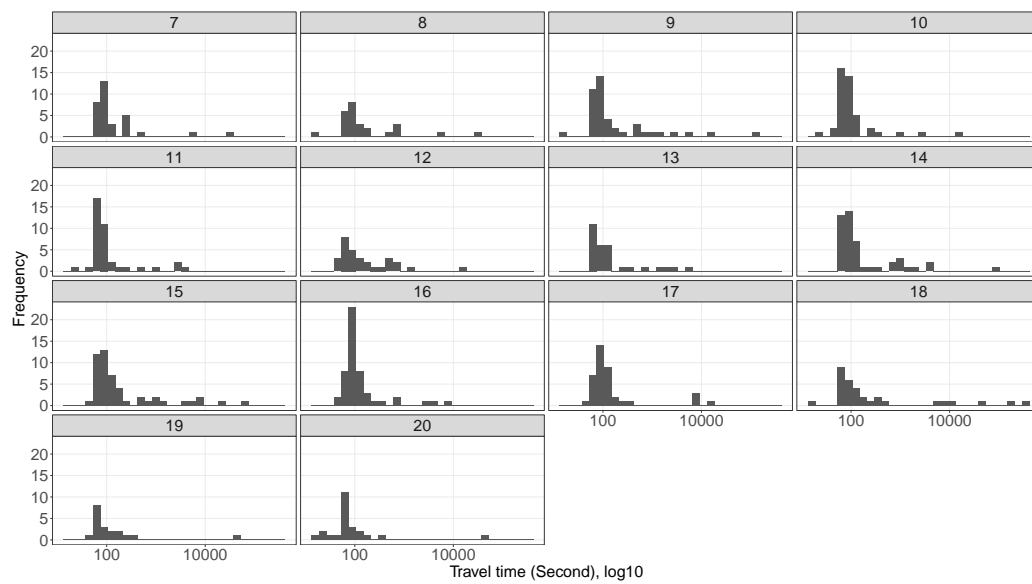

Figure 4.6: Frequency of travel times on an hourly basis (x axis is scaled as log10), Adelaide Road, as observed after removing static devices. Hours are shown on top of each histogram (7 to 20).
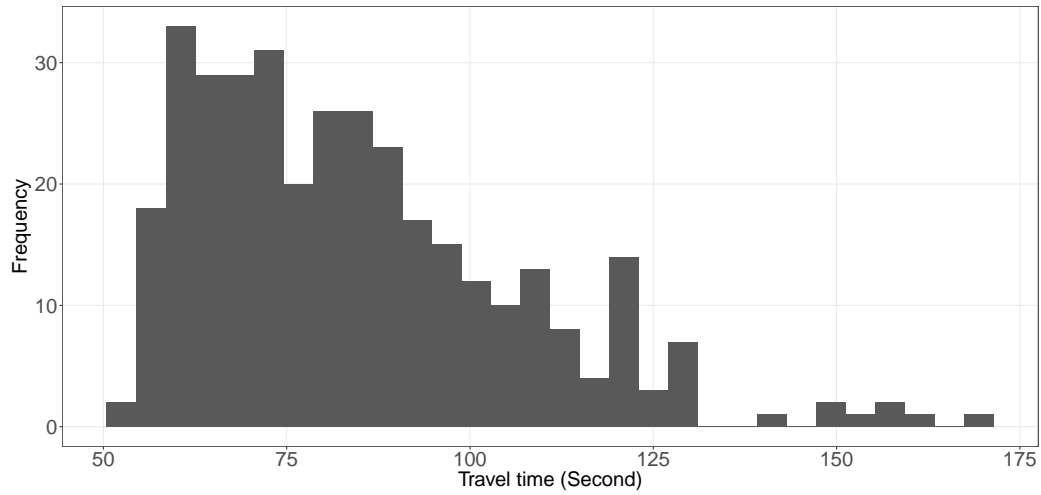
Figure 4.7: Frequency of travel times, Adelaide Road, after applying ExtoVT.
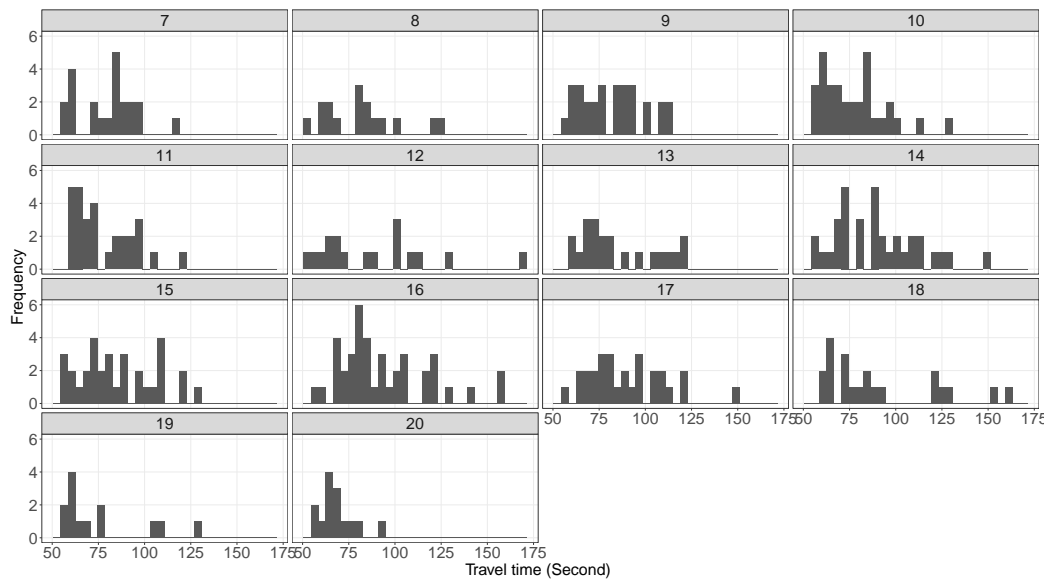

Figure 4.8: Frequency of travel timed on an hourly basis, Adelaide Road, after applying ExtoVT.
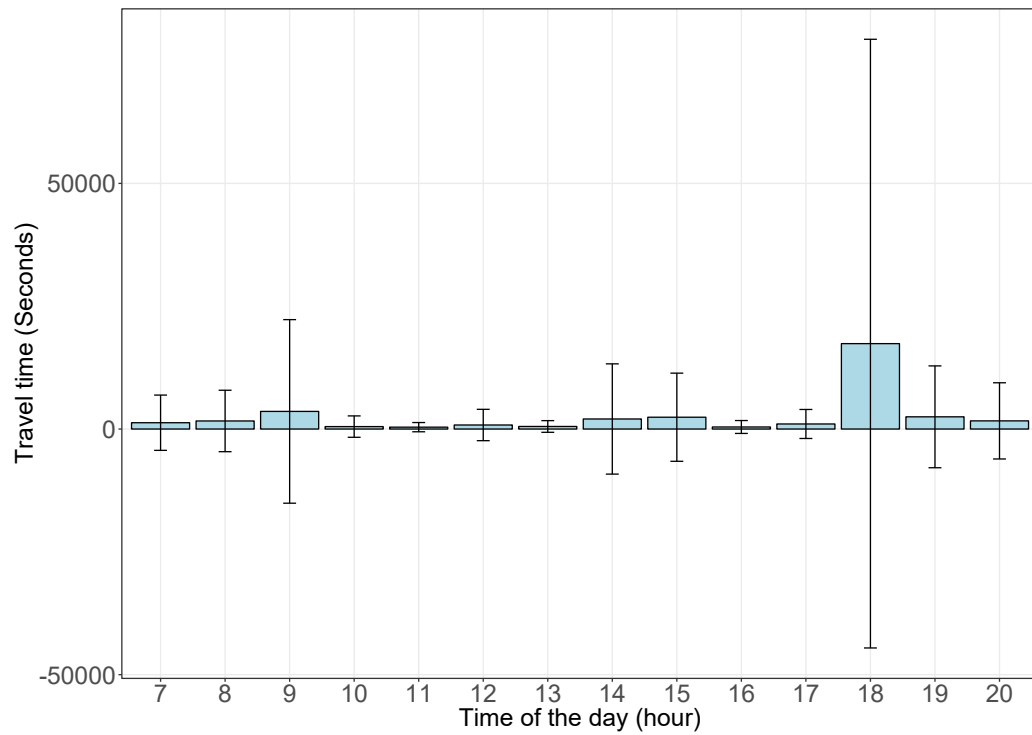
Figure 4.9: Adelaide Road, the mean and standard deviation of travel times before applying ExtoVT (only static nodes were removed).
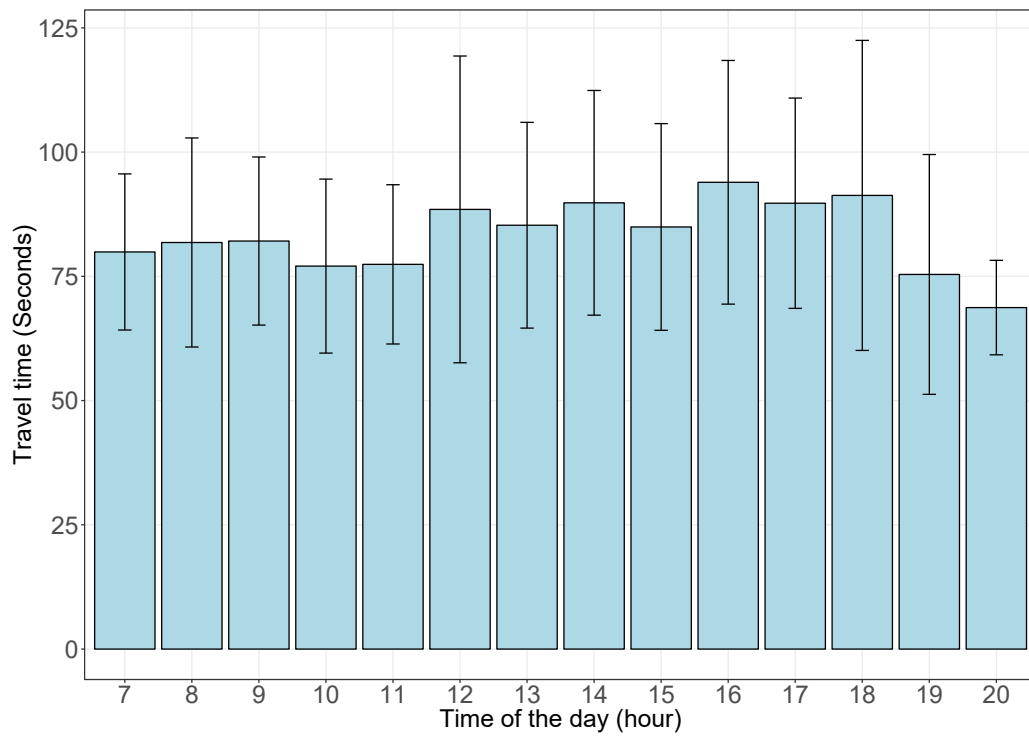
Figure 4.10: Adelaide Road, the mean and standard deviation of travel times after applying ExtoVT.
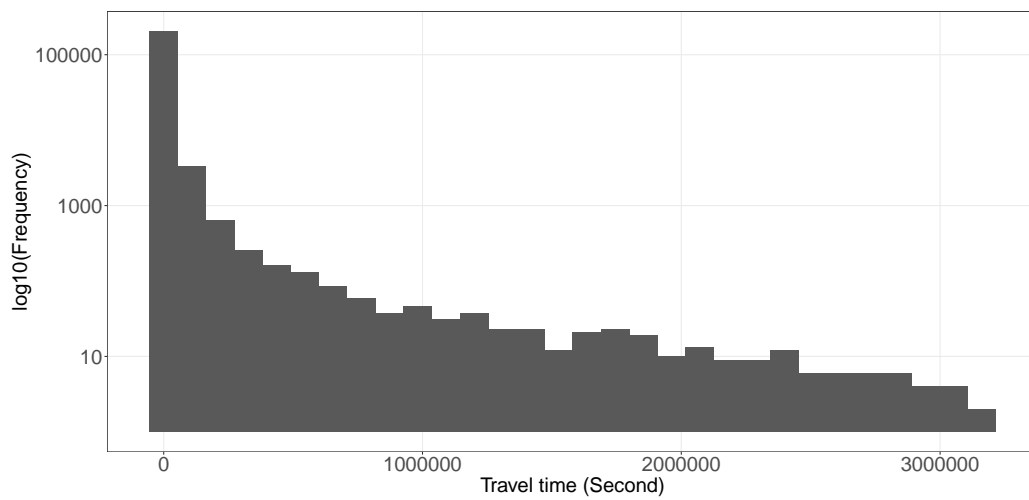

Figure 4.11: Frequency of travel time (y axis is scaled as log10), Wellington City, as observed, after removing static devices.
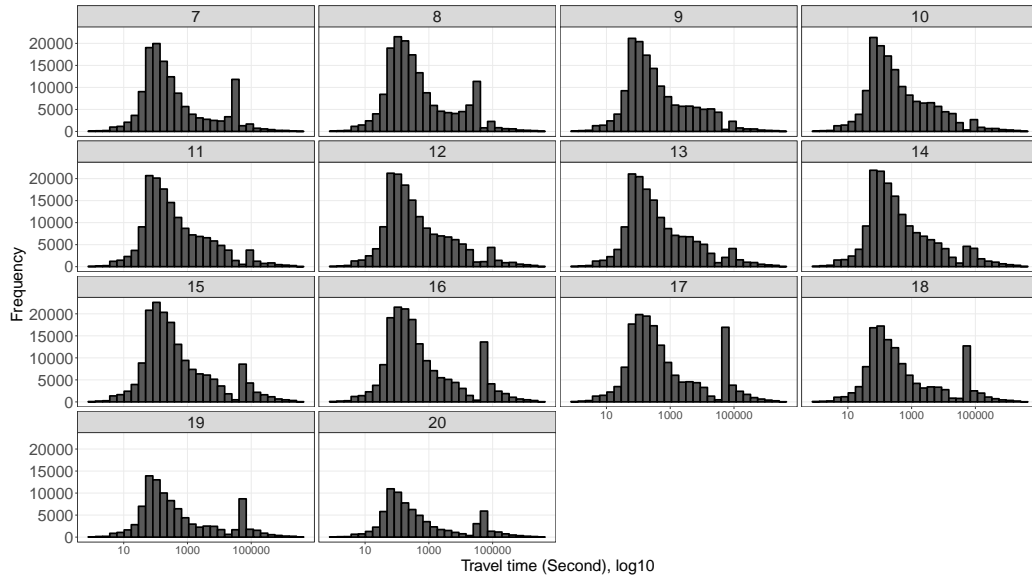
Figure 4.12: Frequency of travel times on an hourly basis (x axis is scaled as log10), Wellington City, as observed, after removing static devices.
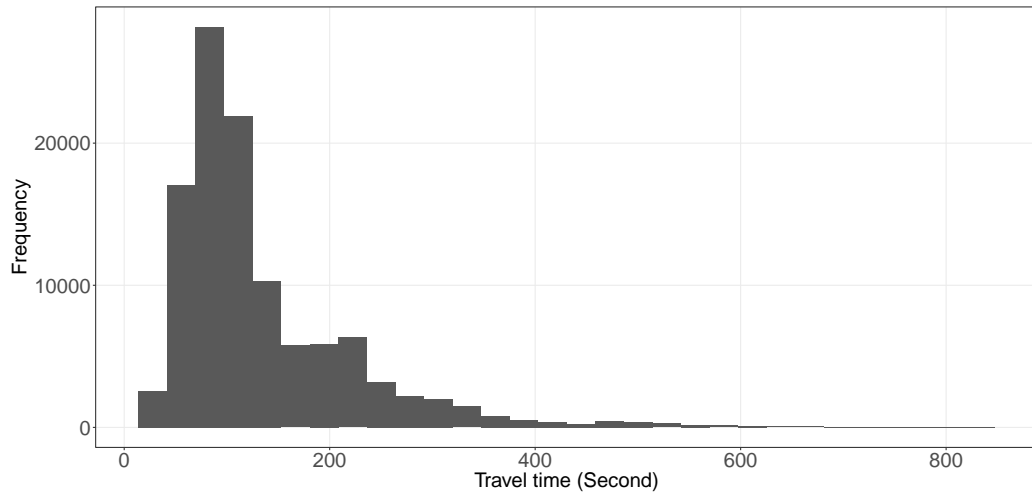
Figure 4.13: Frequency of travel times, Wellington City, after applying ExtoVT.
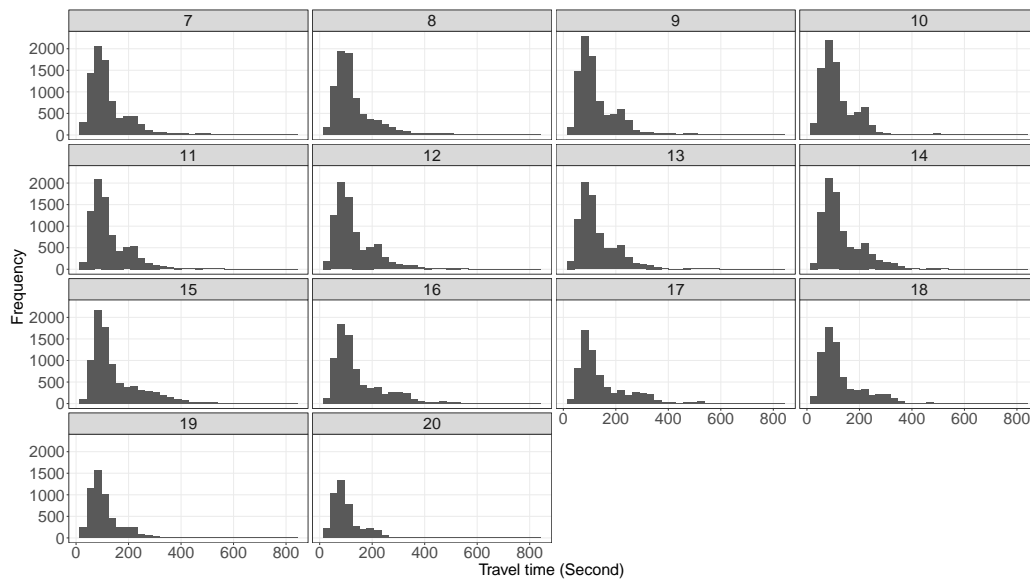
Figure 4.14: Frequency of travel times on an hourly basis, Wellington City, after applying ExtoVT. Trips are longer in the afternoon compared to the morning as thicker tails are observable in the most of afternoon hours (e.g. hours 15, 16).

Figure 4.15: Wellington City, the mean and standard deviation of travel times before applying ExtoVT (only static nodes were removed).

Figure 4.16: Wellington City, the mean and standard deviation of travel times after applying ExtoVT.

Chapter 5

Inter-Arrival Time Modelling

5.1   Introduction

Traffic simulations, on which scenario based urban planning is often based, require the correct characterisations of the traffic flow. The data that inf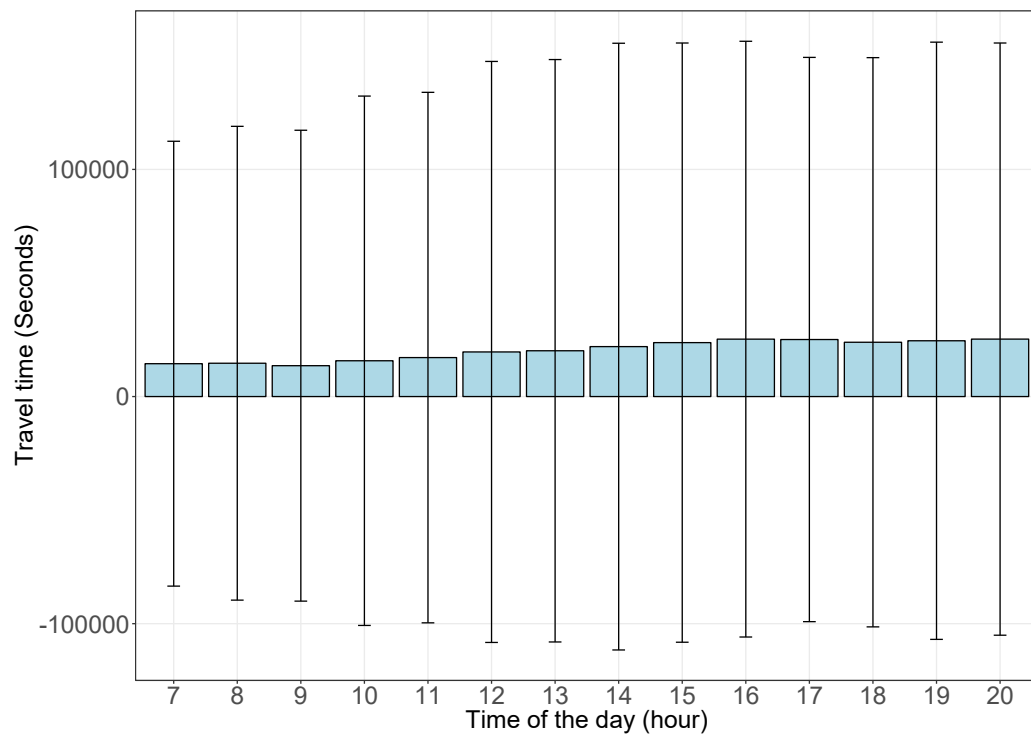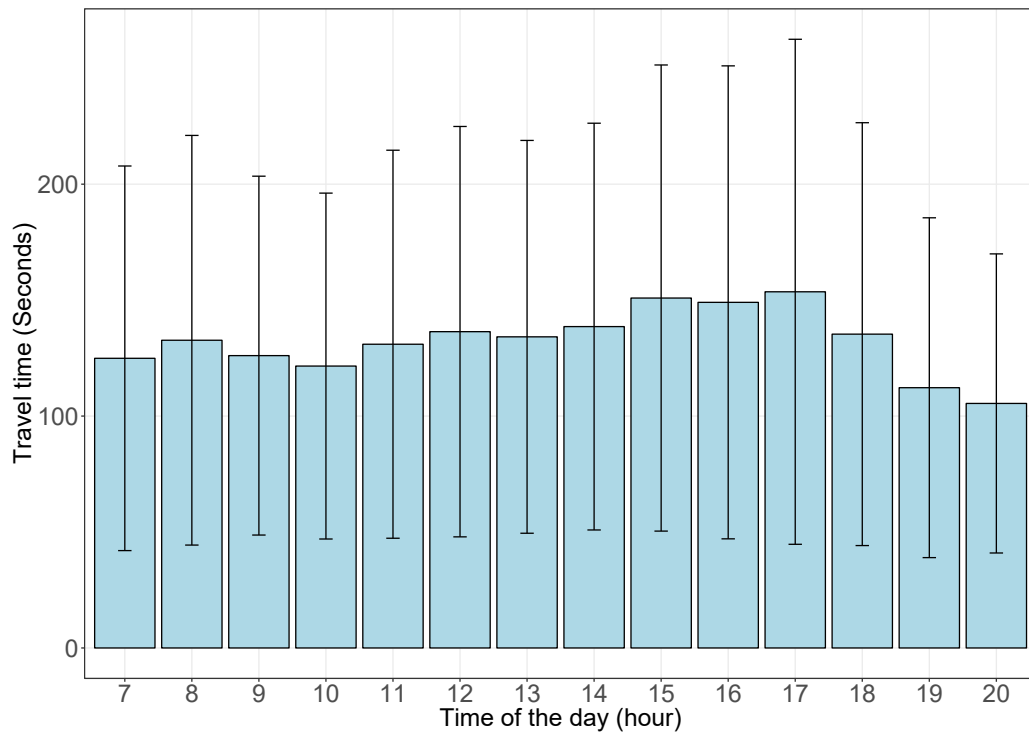orm these are counts of vehicles passing particular points on roads where cameras, counters, or sensors were installed. Inter-Arrival times (IATs) obtained from traffic data are used to model the traffic characteristics and also design scheduling policies. Inter-Arrival Time is the time difference between two consecutive arrivals. This chapter focuses on the examination, modelling, and extraction of Inter-Arrival Times (IAT) of all the passing vehicles based on the collected data by Bluetooth counters. Because, BTMS collects data at a very granular level (1 second) compared to some of the other traditional monitoring tools which used in Wellington and provide the data in bulk groups (e.g. pneumatic road tubes typically aggregate the data at 15 minutes intervals, based on the settings), this thesis examines this data for IAT modelling. There are four common distributions that many researchers either identified or used for IAT modelling and scheduling policies in transportation which are: Exponential, Gamma, Weibull, and Lognormal. Poisson process implies that the IAT between successive vehicles is exponentially distributed, arrivals are independent and identically distributed (iid), and the rate is constant.

The Exponential distribution is widely used for IAT modelling in different transportation areas such as modelling highway traffic accidents [105] and modelling the distribution of vehicles in vehicular networks [77]. For performance evaluation of vehicular communication algorithms, some researchers distribute vehicles randomly based on the Poisson process such as [136]. The authors in [128] investigated the performance of 802.11p, which is a broadcast schema, in vehicular communications and verified their model by simulation, through the distribution of vehicles by the Poisson process. Also, Vinel et. al. in [245], and Hafeez et. al. in [84] assumed a Poisson process for the distribution of vehicles on the road networks. Khabazian et. al. in [156] investigated the formation of clusters of vehicles and assumed that the arrival rates follow a Poisson process. The authors in [209] and [137], used a Poisson process for vehicle distributions.

The Lognormal distribution is another distribution for IAT modelling. Based on research published in [131], [81], [80], the discharge headway (i.e. time elapsed at a reference line for passing two consecutive vehicles in a queue) for most cars (except the first car)[147] follows a Lognormal distribution. The first car has different characteristics as the driver of the first car needs to pay attention to the traffic lights and other potential hazards (e.g. pedestrian crossing), while other subsequent vehicles are less/not affected by these factors.

The Gamma and Weibull distributions have also been used for IAT modelling in the transportation modelling. Jowon et. al. in [158] used the Gamma distribution for modelling the travel time in the transportation modelling. The authors in [256] used a Weibull distribution for modelling the headway distributions of vehicles.

This chapter explains the theory of IAT distributions, how their parameters are estimated, and how models are selected. In addition, as BTMS detect only a proportion of passing vehicles, this chapter examines the impact of this proportional detection on IAT modelling, parameter estimation, and estimation of the some of the characteristics (e.g. variance) of all the passing vehicles. This examination is done by investigation of IAT for two weeks of Bluetooth data collected in Wellington. To demonstrate IAT modelling, this chapter is organised as follows.

Section 5.2 briefly reviews the main characteristics of the four mentioned distributions. Section 5.3 examines and models the fact that BTMS can only detect a proportion of passing vehicles by introducing a thinning process. The fundamental of modelling and estimation of the parameters of distributions is covered in the fourth Section of this chapter, 5.4. Section 5.5 simulates the modelling of the IAT obtained from Bluetooth data, clarifies the methodology, and studies how the thinning process affects the IAT modelling through introducing the Monte Carlo simulation. Later, Section 5.6 finds the best fit for the IAT obtained from Bluetooth data and estimates the parameters for each hour. This chapter concludes in Section 5.7.

5.2 Overview of the four distributions

This section briefly reviews the probability density function (pdf), and the properties of the Exponential, Lognormal, Weibull, and Gamma distributions.

The probability density function (pdf) or simply density function, $f(x)$, for continuous variables, is the probability density of a random variable, X. The pdf has the following properties [188](p85):

$$(i). f(x) \geq 0 \quad , \text{for all } x$$
$$(ii). \int_{-\infty}^{\infty} f(x)dx = 1 \tag{5.1}$$

The probability that X takes a value in the interval $[a,b]$ is:

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx \tag{5.2}$$

Table. 5.1 provides the characteristics of the Exponential, Lognormal, Weibull, and Gamma distributions. The Weibull and Gamma distributions include the Exponential distribution when the shape parameter, (α), is equal to one. These four distributions are defined for a non-negative continuous random variable X. The Exponential distribution has only a single parameter, which restricts its flexibility,

while the other three distributions have more flexibility in the modelling of the variance and the mean.

Table 5.1: Probability density function (pdf), Expected value and variance of the popular distributions used for IAT modelling.

Distribution	Exponential	Lognormal	Weibull	Gamma
Symbol	$X \sim Exp(\lambda)$	$X \sim LN(\mu, \sigma^2)$	$X \sim Weibull(\alpha, \beta)$	$X \sim Gamma(\alpha, \beta)$
pdf	$\lambda e^{-x\lambda}$ (5.3)	$\frac{1}{\sqrt{2\pi}\sigma x} e^{\left(-\frac{[\ln(x)-\mu]^2}{2\sigma^2}\right)}$ (5.4)	$\frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-(\frac{x}{\beta})^\alpha}$ (5.5)	$\frac{1}{\Gamma(\alpha)\beta^\alpha} x^{(\alpha-1)} e^{-\frac{x}{\beta}}$ (5.6)
$E(X)$	$\frac{1}{\lambda}$ (5.7)	$e^{\mu+\frac{\sigma^2}{2}}$ (5.8)	$\beta\Gamma\left(1+\frac{1}{\alpha}\right)$ (5.9)	$\alpha\beta$ (5.10)
$Var(X)$	$\frac{1}{\lambda^2}$ (5.11)	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$ (5.12)	$1+\frac{2}{\alpha})- \Gamma^2(1+\frac{1}{\alpha}))$ (5.13)	$\alpha\beta^2$ (5.14)
Conditions	$\lambda > 0, x \geq 0$	$\sigma > 0, \mu \in (-\infty, +\infty), x \geq 0$	$\alpha, \beta > 0, x \geq 0$	$\alpha, \beta > 0, x \geq 0$

Each distribution has a distinctive shape. To demonstrate the possible shapes of each distribution, the pdfs of these four distributions for mean=1 and variance=1, and mean=1 and variance=0.1 are plotted in Fig. 5.1, and Fig. 5.2, respectively. As shown in Fig. 5.1, where $\alpha = 1$, the Gamma, Weibull, and Exponential are identical and they are overlaid on top of each other, whereas the Lognormal distribution has a sharp peak away from the rest. All of these four distributions are right-skewed. With the variance equal to 0.1, Fig. 5.2, the exponential distribution has the same shape as it has with variance equals to 1. This is because the exponential distribution is a single parameter distribution and is not affected by variance and only depends on the mean. By setting the variance to 0.1, the Lognormal distribution had minor changes, and the Gamma and Weibull change their shapes and are distinguishable from the Exponential distribution.
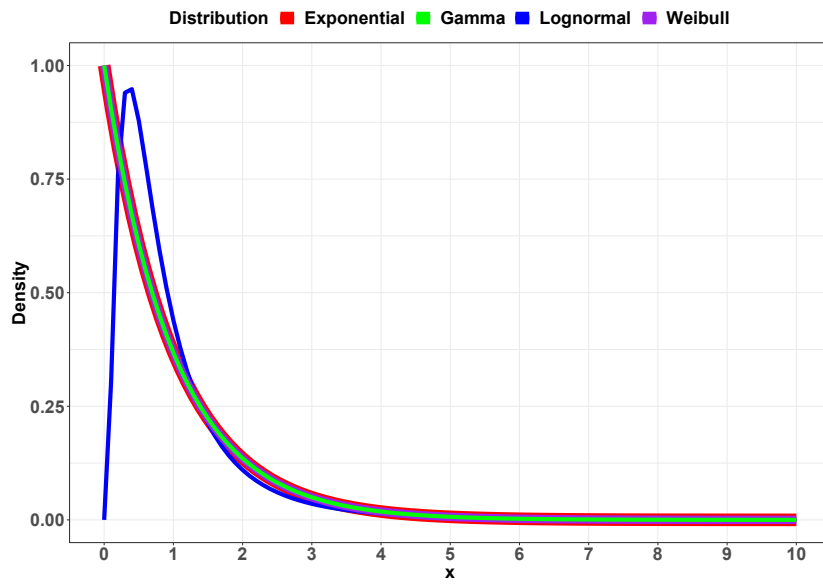
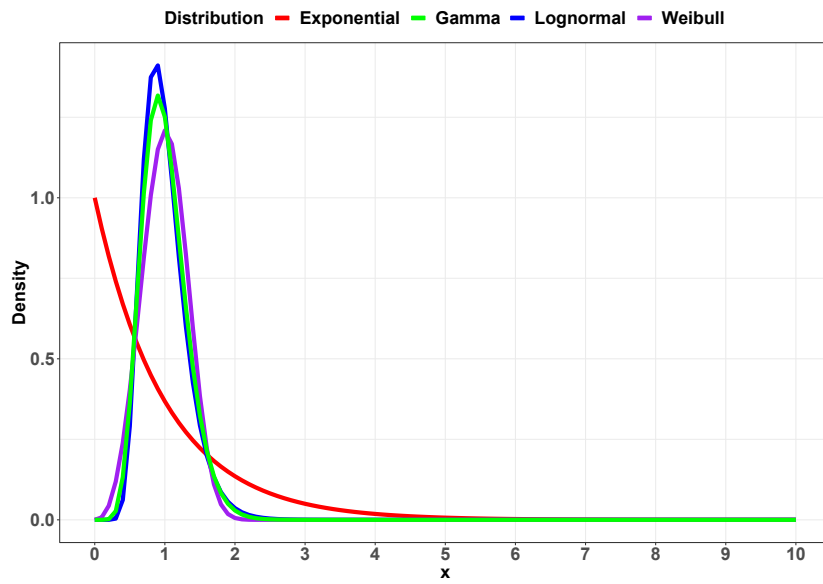Figure 5.1: Density plot, mean=1 and variance=1

Figure 5.2: Density plot, mean=1 and variance=0.1

5.3   Thinning process

As BTMS do not detect every passing vehicle and only detect a proportion of passing vehicles (of course with detectable Bluetooth devices), therefore rather than seeing the true (i.e. all the passing vehicles) vehicle arrival times(labelled with subscript t, i.e. $_t$), the data collected by BTMS only contains a thinned version of it (labelled with subscript o, i.e. $_o$): only a random subset of the vehicle arrivals, determined by the proportion of vehicles with detectable Bluetooth devices. IAT modelling using Bluetooth data needs to take proper account of this thinning. This detectability ratio is shown as θ ($0 < \theta \leq 1$). This thinning process describes how the IAT obtained from vehicles, which have detectable Bluetooth devices (observed dataset), can be generalised from all of the passing vehicles (true dataset). Let X be the inter-arrival times between successive vehicles, and Y be the time between successive detections. The procedure to parameterise the true expected value (true mean), true variance, observed expected value, and observed variance by introducing the mean, η_t, η_o, and squared coefficients of variation ζ_t and ζ_o for the two processes is:

$$E(X_i) = \eta_t \tag{5.15}$$

$$E(Y_i) = \eta_o \tag{5.16}$$

$$Var(X_i) = \zeta_t \eta_t^2 \tag{5.17}$$

$$Var(Y_i) = \zeta_o \eta_o^2 \tag{5.18}$$

Over long periods the parameters of the distribution will change with as the traffic volume varies through the day. However for periods of an hour or so, the true process is a sequence of inter-arrival times $X_1, X_2, \ldots$, which are independent and identically distributed. The jth observed inter-detection time is given by

$$Y_j = \sum_{i=1}^{N_j} X_{ji}, \tag{5.19}$$

where $\{X_{ji}, i = 1,\ldots,N_j\}$ are the true inter-arrival times between the $(j-1)^{\text{th}}$ and j^{th} Bluetooth device detection, relabelled X_{ji} to indicate the interval they belong to. However, N_j, the number of observations before seeing a car equipped with a detectable device, is a Geometric(θ) random variable with the following probability mass function (pmf), Expected value and variance:

$$f(N_j) = (1-\theta)^{N_j-1}\theta \qquad (5.20)$$

$$E[N_j] = \frac{1}{\theta} \qquad (5.21)$$

and

$$Var[N_j] = \frac{(1-\theta)}{\theta^2} \qquad (5.22)$$

If I assume: $E[X_i] = \eta_t$ and $Var[X_i] = \zeta_t \eta_t^2$,then

$$\eta_o = E[Y_j] = E[E[Y_j|N_j]] = E[N_j\eta_t] = \eta_t E[N_j] = \frac{\eta_t}{\theta}$$

and finally,

$$\eta_t = \eta_o\theta, \qquad (5.23)$$

and for the variance:

$$\begin{aligned}
\zeta_o\eta_o^2 = Var[Y_j] &= E[Var[Y_j|N_j]] + Var[E[Y_j|N_j]] \\
&= E[Var(X)N_j] + Var[E(X)N_j] \\
&= E[N_j\zeta_t\eta_t^2] + Var[N_j\eta_t] \\
&= \zeta_t\eta_t^2 E[N_j] + \eta_t^2 Var[N_j] \\
&= \frac{\zeta_t\eta_t^2}{\theta} + \frac{\eta_t^2(1-\theta)}{\theta^2} \\
&= \frac{\eta_t^2}{\theta^2}(\zeta_t\theta + 1 - \theta)
\end{aligned}$$

then,

$$\zeta_o = \zeta_t\theta + 1 - \theta$$

and finally,

$$\zeta_t = \frac{(\zeta_o - 1 + \theta)}{\theta} \tag{5.24}$$

Since ζ_t must be positive, then $\zeta_o > 1 - \theta$. Therefore, assuming I know which parametric distribution is generating the data, I can use the observed mean and variance, obtained from detected passing vehicles (observed), I can estimate parameters of the assumed true underlying distribution based on Table 5.2. Note that estimation of α for the Weibull distribution requires the solution of the below non-linear equation:

$$\zeta_t = \frac{\Gamma(1 + 2/\alpha)}{\Gamma(1 + 1/\alpha)^2} - 1 \tag{5.25}$$

Table 5.2: Parameters calculations for the candidates distributions based on the η_t and ζ_t

Distribution	Exponential	Lognormal	Weibull	Gamma
First parameter	$\lambda = \frac{1}{\eta_t}$ (5.26)	$\mu = \ln(\eta_t) - \frac{\sigma^2}{2}$ (5.27)	$\beta = \frac{\eta_t}{\Gamma(1 + \frac{1}{\alpha})}$ (5.28)	$\alpha = \frac{1}{\zeta_t}$ (5.29)
Second parameter		$\sigma^2 = \ln(\zeta_t + 1)$ (5.30)	$\zeta_t = \frac{\Gamma(1 + 2/\alpha)}{\Gamma(1 + 1/\alpha)^2} - 1$ (5.31)	$\beta = \frac{\eta_t}{\alpha} = \eta_t \zeta_t$ (5.32)
Conditions	$\lambda > 0, x \geq 0$	$\sigma > 0, \mu \in (-\infty, +\infty), x \geq 0$	$\alpha, \beta > 0, x \geq 0$	$\alpha, \beta > 0, x \geq 0$
$\zeta_t =$	1	$e^{\sigma^2} - 1$	solve for α : $\frac{\Gamma(1 + 2/\alpha)}{\Gamma(1 + 1/\alpha)^2} - 1$	$1/\alpha$

5.4   Model selection and estimation methodology

This section describes the approach for model selection and parameter estimation in this section. The model selection is accomplished by using Aikake's Information Criterion (AIC). Hirotugu Akaike developed AIC in 1973 [76], which balances the number of parameters and improvement of fit through the below

equation (5.33):

$$AIC = -2\ell(\hat{\phi}) + 2k, \tag{5.33}$$

where k is the number of parameters estimated, ϕ is the parameters of the model, $\ell(\phi)$ is the log-likelihood of the model given to the observed dataset, and $\hat{\phi}$ is the likelihood estimate of ϕ, given the data. The smallest AIC which maximises the balance between the number of parameters and fit must be selected. If the detectability ratio is one (i.e. traffic sensors detect all the passing vehicles), then the Maximum Likelihood Estimation (MLE) is used for parameters estimation as explained below.

For a sample of independent observations $x_1, x_2, \ldots, x_n$ from a random variable X with the probability function $f(x; \phi)$, the likelihood function is defined as:

$$L(\phi) = \prod_{i=1}^{n} L_i(\phi; x) = \prod_{i=1}^{n} f(x_i; \phi), \tag{5.34}$$

where ϕ is a vector of the unknown parameter(s) of the distribution, and n is the number of observations [205]. The maximum likelihood estimate (MLE) $\hat{\phi}$ of ϕ is the value of ϕ that maximises the value of the equation 5.34. Due to the monotonic transformation of logarithmic functions, the MLE ($\hat{\phi}_{ML}$) also maximises the logarithm of the likelihood, which usually has a simpler analytic form than the likelihood and defined as:

$$\ell(\phi) = \ln[L(\phi)] \tag{5.35}$$

The Hessian matrix is a square matrix of the second-order partial derivatives of a function which describes the local curvature of a function and can be calculated

as follows:

$$H = \begin{bmatrix} \dfrac{\partial^2 \ell(\phi)}{\partial \phi_1} & \dfrac{\partial^2 \ell(\phi)}{\partial \phi_1 \partial \phi_2} & \cdots & \dfrac{\partial^2 \ell(\phi)}{\partial \phi_1 \partial \phi_k} \\[2ex] \dfrac{\partial^2 \ell(\phi)}{\partial \phi_2 \partial \phi_1} & \dfrac{\partial^2 \ell(\phi)}{\partial \phi_2^2} & \cdots & \dfrac{\partial^2 \ell(\phi)}{\partial \phi_2 \partial \phi_k} \\[2ex] \vdots & \vdots & \ddots & \vdots \\[2ex] \dfrac{\partial^2 \ell(\phi)}{\partial \phi_k \partial \phi_1} & \dfrac{\partial^2 \ell(\phi)}{\partial \phi_k \partial \phi_2} & \cdots & \dfrac{\partial^2 \ell(\phi)}{\partial \phi_k^2} \end{bmatrix} \tag{5.36}$$

The Fisher information is the amount of the information that an observation of X, carries about the parameter(s) ϕ of the distribution which it came from, and defined as the negative expected value of the Hessian (5.36), as shown below:

$$\mathbf{I}(\phi) = -E \frac{\partial^2}{\partial \phi_i \partial \phi_j} \ell(\phi), \quad 1 \leq i,j \leq k \tag{5.37}$$

The relation between likelihood, Hessian, Fisher information and Variance can be defined as:

$$\mathbf{I}(\phi) = -E \frac{\partial^2}{\partial \phi_i \partial \phi_j} \ell(\phi), \quad 1 \leq i,j \leq k, \tag{5.38}$$

, where k is the number of parameters to be estimated.

$$\mathbf{H}(\phi) = \frac{\partial^2}{\partial \phi_i \partial \phi_j} \ell(\phi), \quad 1 \leq i,j \leq k \tag{5.39}$$

$$\mathrm{Var}(\hat{\phi}_{\mathrm{ML}}) = [\mathbf{I}(\hat{\phi}_{\mathrm{ML}})]^{-1} \tag{5.40}$$

$$\mathrm{SE}(\hat{\phi}_{\mathrm{ML}}) = \frac{1}{\sqrt{\mathbf{I}(\hat{\phi}_{\mathrm{ML}})}}, \tag{5.41}$$

where $(\hat{\phi}_{\mathrm{ML}})$ is the maximum likelihood of the distribution parameters.

5.5 Simulation study

Because the Exponential distribution is the only distribution among these four distributions that has a single parameter and this parameter is associated with the mean of IAT, where $Y = Exp(\frac{\theta}{\eta_t})$, therefore analytic results are only available for the thinned data of BTMS in the Exponential case. As Bluetooth devices only detect a proportion of passing vehicles, this section simulates the process of detecting Bluetooth devices by BTMS to investigate the effect of thinning on the properties of the data, and on the estimation methodology. The high level design of the simulation study is provided in the next subsection and Fig. 5.3 and more details are provided in the following subsection.

5.5.1 Description

BTMS only detect a proportion of passing vehicles with a rate of θ. These observed arrival times can be used to extract the observed IAT. These observed IAT have their mean, η_o, and variance, ζ_o, which can be different with the all of the passing vehicles (i.e. detectable and non-detectable, true dataset). Therefore, it is crucial to estimate the mean, η_t, and variance, ζ_t, of all the passing vehicles based on the observed dataset, which was the focus of the thinning process. In addition, because BTMS only detect a proportion of passing vehicles, the modelling of these observed IAT, may not be the same as the true dataset. Therefore, the way that BTMS works must be constructed through the Monte Carlo method. This is done to investigate whether the same distribution used for generating the true IAT (i.e. parent distribution) can be detected from the observed dataset (i.e. child distribution), under a thinning process, where BTMS only record a proportion of passing vehicles with a rate of θ. Moreover, the observed (thinned process) mean and variance are fixed, and for various detection rate values, the Monte Carlo method determines the parameter values of the underlying parent distribution that would lead to these thinned properties. This is done to ensure a comparison that is based on the same sample sizes (number of Bluetooth detections) rather than the same underlying vehicle flow rate. The θ can be unknown in a real-world sce-

nario, so the data is thinned by an unknown amount. Therefore, this simulation study regards ζ_o, η_o as constant, and for each θ derives the ζ_t, η_t, where these would generate the observed data properties. This Monte Carlo method simulated the behaviour of a traffic device counter for an hour of an artificial route. This Monte Carlo simulation is executed for a range of values of the detectability ratios, θ, observed mean and variance, with randomly generating vehicles based on the four mentioned distributions (i.e. parent distribution). The high-level idea is shown below.
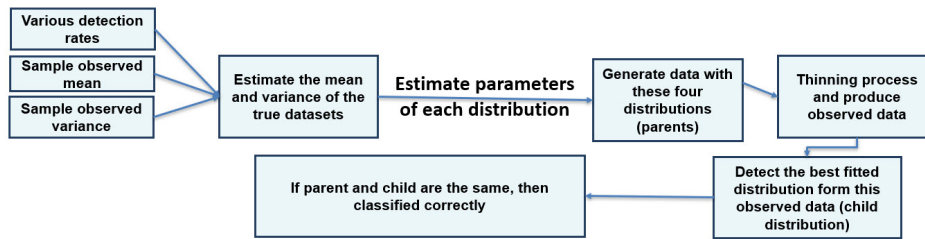

Figure 5.3: The high-level design of the Monte Carlo simulation

The Pseudo-code of the Monte Carlo simulation is shown in Algorithm 3. In the first five steps, this algorithm simulates the behaviour of BTMS and form the observed IAT set. This algorithm estimates the model parameters using Maximum Likelihood Estimation for each of the four parametric families, and select the best fitting model using the AIC criterion. This Monte Carlo simulation examines whether the fitted thinned distribution can correctly identify the distribution of the observed data. The motivation behind this is to examine whether in the absence of appropriate analytical results, the true parent distribution can be identified.

5.5.2   Monte Carlo results

The Monte Carlo algorithm, algorithm 3, is applied on common different sets of variances ($\{2,4\}$), means ($\{15,30\}$), and detectability ratios ($\{0.1,0.2,\ldots,1\}$). Each simulation is executed 100 times for each θ, and the results are shown in Fig. 5.4. In these figures, the X axis is the detectability ratio, and the Y axis represents the number of times that the identified child distribution is the same

---

Algorithm 3: Monte Carlo Algorithm

Data:

S_θ =Set of detectability ratios;S_{ζ_o} =Set of observed variances;

S_{η_o} =Set of observed means;

for *each θ in S_θ* **do**

    for *each ζ_o in S_{ζ_o}* **do**

        for *each η_o in S_{η_o}* **do**

            Calculate ζ_t and η_t based on Eq.5.24 and Eq.5.23 from θ, η_o and ζ_o

            for *Each candidate distribution* **do**

                1. calculate the parameters of this distribution based on η_t and ζ_t as shown in Table 5.2.

                2. Generate random inter-arrival times y_j, where $\sum y_j \leq 3600$ seconds by using the parameters of the distribution obtained in the previous step as the true IAT dataset.

                3. Add these inter-arrival together to form the actual ascending true arrival times.

                4. Make each vehicle observable with probability θ within y_j random arrivals by Bernoulli(θ) and eliminate the rest to form the observed arrival times.

                5. Subtract each two consecutive observed arrival times to obtain the observed IAT dataset.

                6. Apply model selection and find which of the four distributions best fit the observed IAT. Use MLE to find the parameters of each distribution. If the detected distribution from observed IAT is the same as the used distribution for generating the true IAT, then this distribution is detected (classified) correctly, otherwise this is classified incorrectly.

            end

        end

    end

end

---

as the true parent distribution (correctly classified), and the legend demonstrates the four aforementioned distributions. In these figures, as θ increases the percentage of correctly classified distribution increases. This means that hypothetically, if BTMS devices were able to detect all the passing vehicles, then parent and child distributions can be the same (classified correctly) with a high probability. Furthermore, lower variance and mean lead to more samples which increase the accuracy of classification. As shown in these four figures the ratio of detecting vehicles (θ) is not a significant factor for correct classification of the Exponential distribution, while the mean, variance and the detection ratio affect the correct classification of the other three distributions. These set of results signify that if the true arrival dataset has the Exponential distribution, then Exponential distribution can be identified with a high probability of about 75% in the observed dataset independent of the detectability ratio, while the correct identification of other three distributions requires a higher θ. Moreover, in the absence of appropriate analytical results I can identify the true parent distribution if the detection ratio (θ) is very high, which may not be the case in a real-life situation.

Fig. 5.5 visualises the density and frequency plots of the true IAT and observed IAT for observed variance equal to 2, observed mean equal to 15 seconds, and detectability ratio equal to 0.1 ($\zeta_o = 2$, $\eta_o = 15$, $\theta = 0.1$), where the parent and child are Exponential distribution. These two data sets (true IAT and observed IAT) in this figure come from the same family, which means the parent distribution and the identified distribution (child) are the same. This figure also plots the exponential curve with the same estimated mean on these two density plots. The size of the true dataset is about 10 times more than the size of the observed dataset and the mean of the IAT of the true dataset is about 0.1 of the observed dataset.

(a) Observed variance=2,
observed mean=15, $\zeta_o = \frac{2}{15^2} = \zeta_a$

(b) Observed variance=4,
observed mean=15, $\zeta_b = 2\zeta_a$

(c) Observed variance=2,
observed mean=30, $\zeta_c = 1/4\zeta_a$

(d) Observed variance=4,
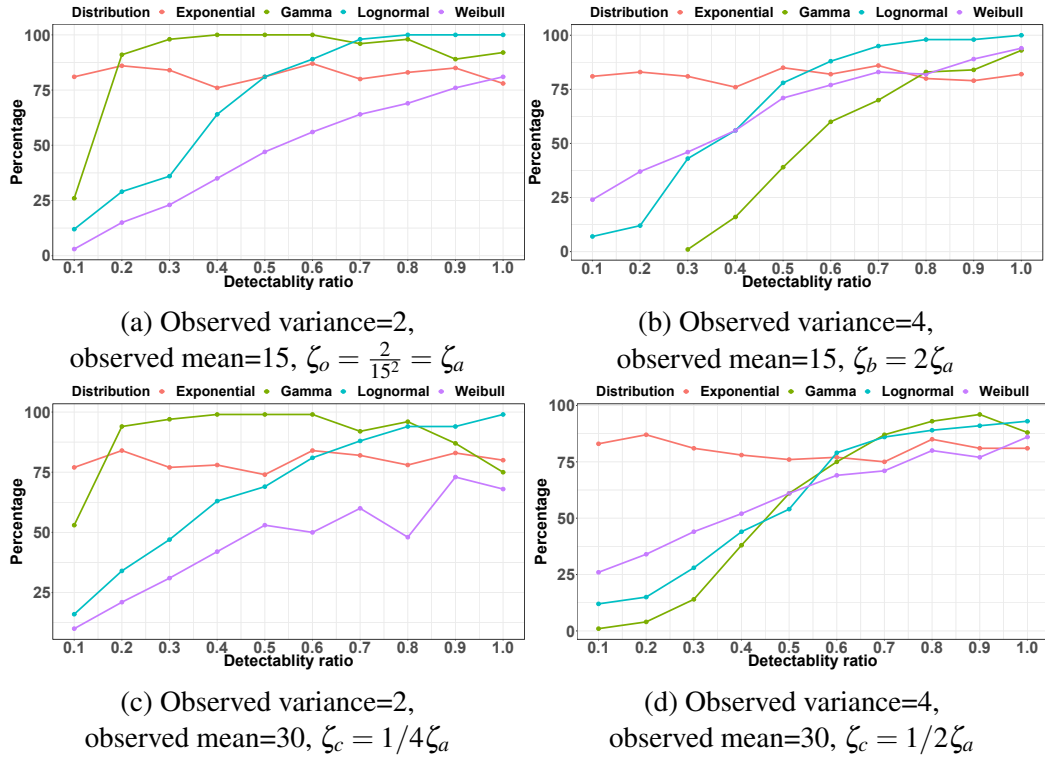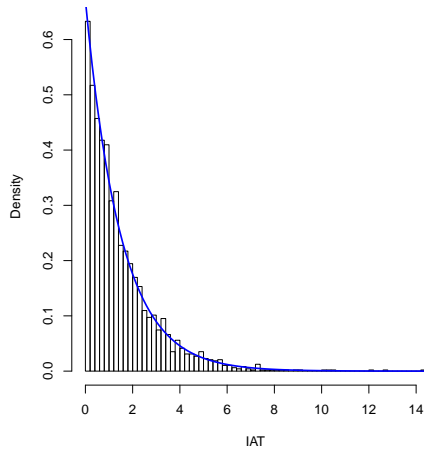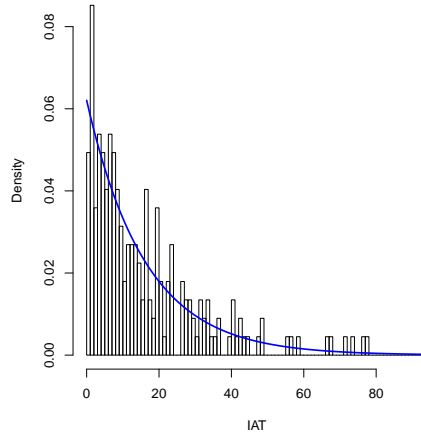observed mean=30, $\zeta_c = 1/2\zeta_a$

Figure 5.4: Results of Monte Carlo simulation for various detectability ratio, observed variance and mean. For each figure, ζ_o is computed based on ζ_a. As detectability ratio (θ) increases the probability of correct classification of the family of a distribution increases. The Exponential distribution is the only distribution that if the parent has this distribution, then the child can be classified correctly with a low detection ratio (θ).

(a) True IAT, density plot.

(b) Thinned IAT, density plot.
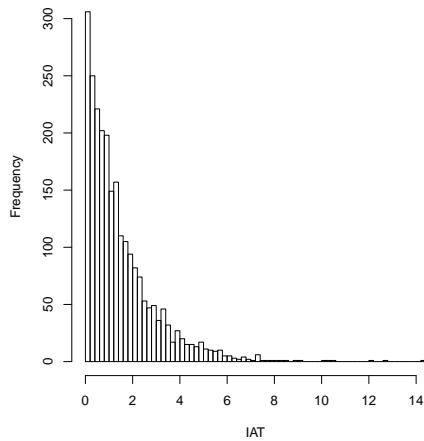
(c) True IAT, frequency plot.
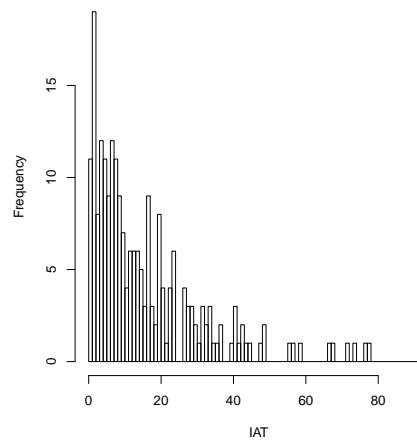
(d) Thinned IAT, frequency plot.

Figure 5.5: Two sample datasets from the Monte Carlo simulation for θ=0.1, observed variance=2, and observed mean=15. Density plots (a and b) with the Exponential distribution fit, and frequency plots (c and d). Since θ=0.1, the true mean is about 10 times less than the observed mean, and consequently the size of the true dataset (c) is about 10 times more than the size of the thinned (observed) dataset (d).

5.6 Inter-arrival modelling of Wellington City as a case study

This section scrutinises the collected Bluetooth traffic data of Wellington city for Inter-Arrival Time modelling.

5.6.1 Data collection and data processing

The proposed ExtoVT is applied for data cleansing as explained in Chapter 4 to only extract vehicles which follow the regular traffic flow. For IAT modelling, this study covers two weeks, 31 Oct (Monday) to 13 Nov (Sunday) 2016, from 07:00:00 to 20:59:59. First week is from 31 Oct (Monday) to 6 Nov (Sunday), and the second week is from 7 Nov (Monday) to 13 Nov (Sunday). This chapter mainly focuses on these hours as these are the hours which represent a typical road condition [47] when most people commute. This dataset has 30 routes, and $211,196$ detected vehicles out of $1,824,551$ mobile detections (i.e. static nodes are removed), which follow regular traffic patterns.

5.6.2 Inter-arrival modelling

This section examines the best detected distribution of the thinned IAT collected from BTMS on an hourly basis for each individual route and for each day. 14 days of collected data, 30 routes, and 14 hours per day provide 5880 sample IAT datasets for the fitting process.

Fig. 5.6 shows the percentage of each distribution which best fitted the observed IAT. As shown in this figure, the Exponential distribution is the best fit for about 70% of the cases (i.e. 70% of 5880 IAT datasets), and the Lognormal, Weibull, and Gamma, have a similar proportion, although these distributions may be affected by the thinning process during data collection by BTMS. Assume the underlying data follows exponential distribution as this is the distribution which is widely used for IAT modelling [128], [245], [84].
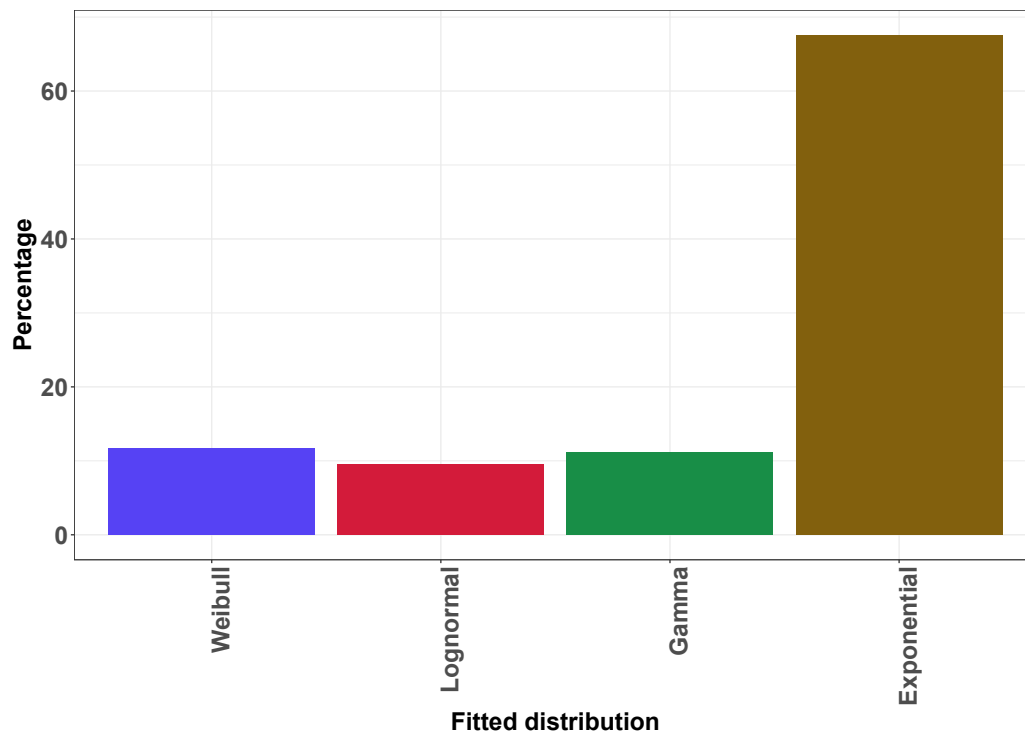
Figure 5.6: Relative proportions of the fitted distributions for the observed IAT.

For the Exponential distribution, this section examines the frequency of the detected Exponential distribution in a more granular level, per hour and per day, as the thinned exponential is the only distribution in the same family as its parent distribution (i.e. the exponential). For frequency, the hourly basis shows the total number of micro OD pairs (routes), which are modelled with each distribution per hour, and for the daily, the frequency is the total number of these hours per day.

The results of the fitting process are shown in Fig. 5.7 and 5.8. Fig 5.7 shows the total number of classifications of the distribution as an Exponential distribution. This section also studies the frequency in a more granular approach (hourly) as shown in Fig. 5.8).

The average of the observed rate (i.e. λ, Lambda) parameter of the Exponential distribution per day for all of the micro OD pairs (routes) is shown in Fig. 5.9. This figure shows that the mean between the arrival of vehicles in the weekend was higher (i.e. lower Lambda) compared to working days, which highlights that fewer vehicles were detected in the weekend due to lower trips due to working constraints of weekdays. Fig 5.10 shows the average of this parameter for all of the micro OD pairs (routes) per hour. As shown in Fig 5.10 the variations of IAT over normal weekdays is in the morning peak with tailing at the end of the day, while this variation is in the middle of the days of weekends with tailing in the morning and afternoon, due to working constraints.

Fig. 5.11 to 5.12 show the frequency plot and arrival times of vehicles at the south direction of the Adelaide Road (see Fig. 4.4) for the two Mondays of the two weeks of study periods from 17:00 to 17:59. The detected thinned distribution for the IAT of these two observed IAT dataset is the Exponential distribution. As demonstrated the frequency of the detected vehicles within this hour (17:00 to 17:59) on Monday of the first week is higher than the Monday of the second week.
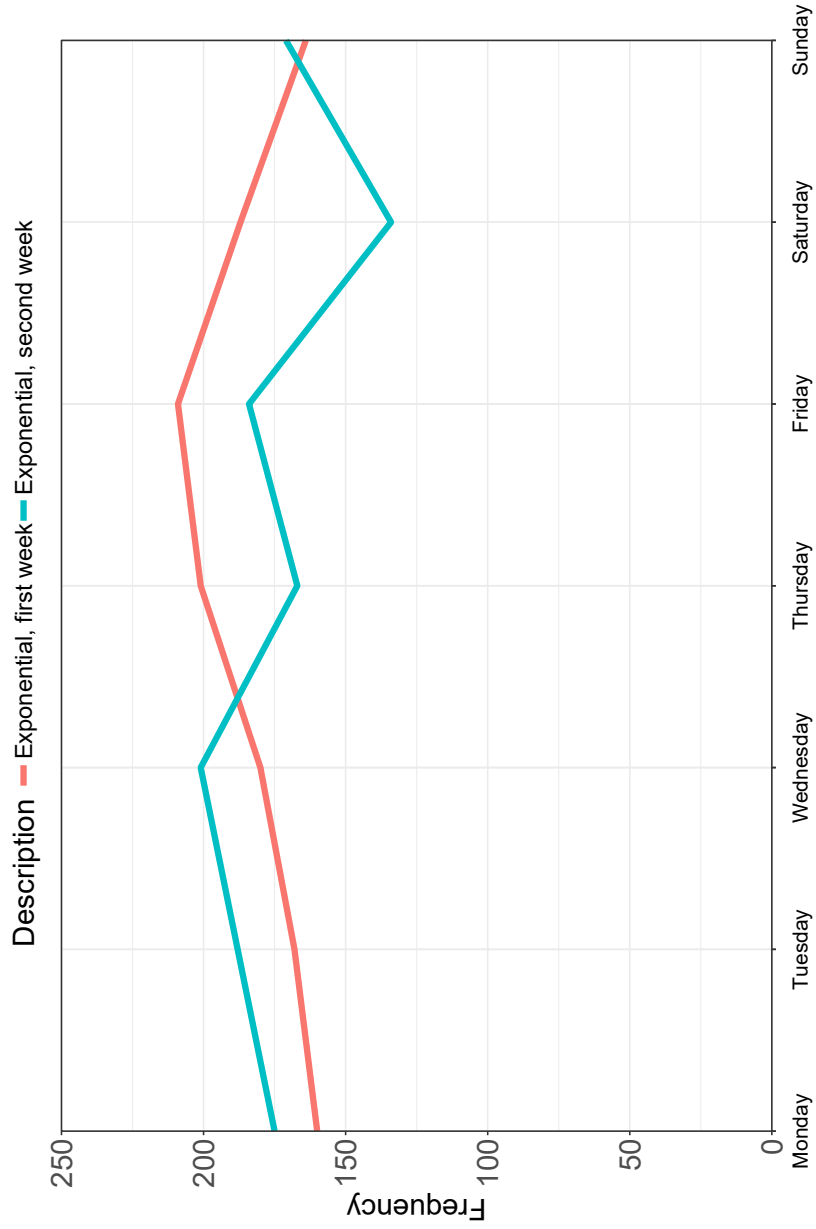
Figure 5.7: Frequency of the fitted thinned Exponential distribution, observed dataset, daily basis.
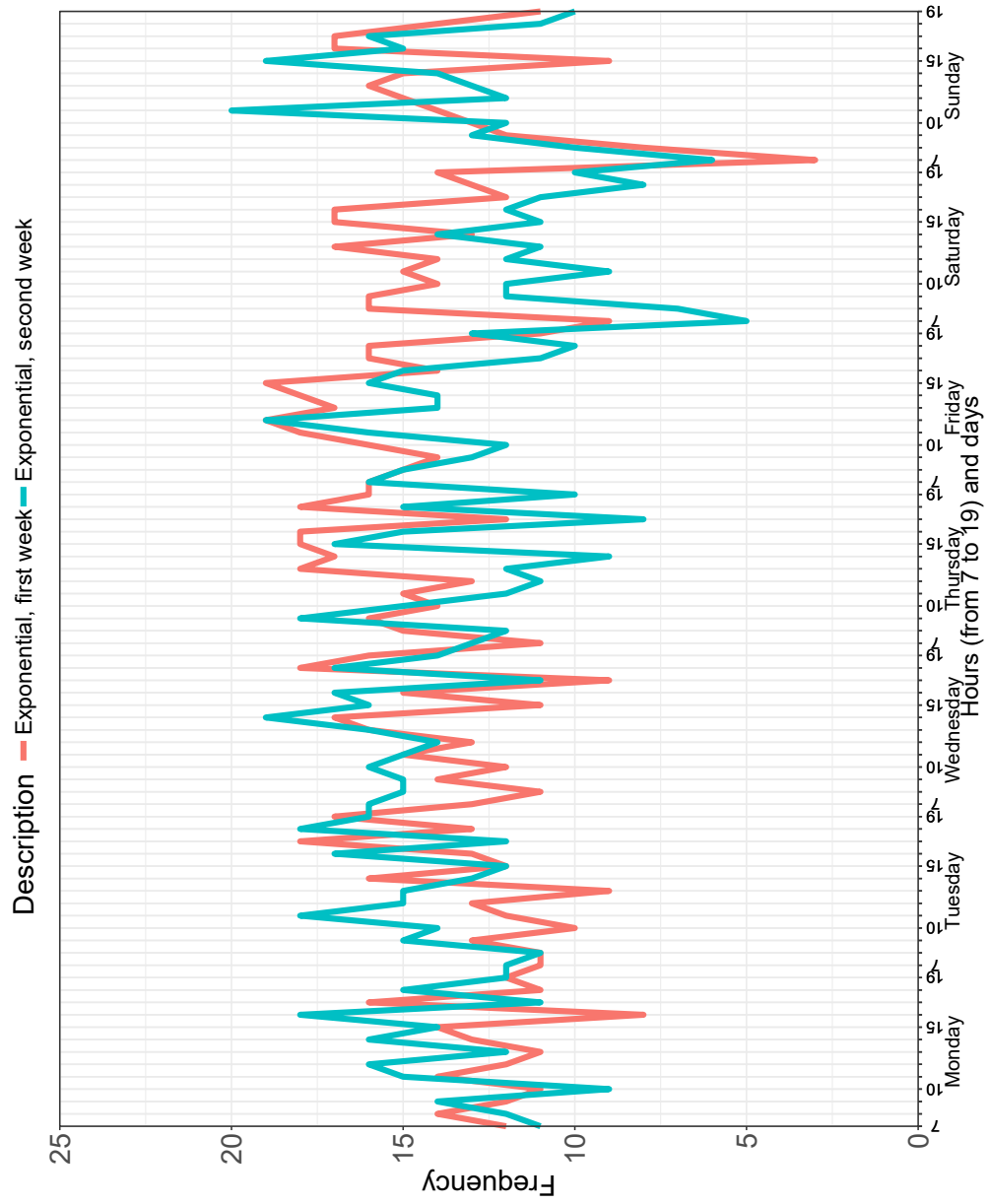
Figure 5.8: Frequency of the fitted thinned Exponential distribution, observed dataset, per hour per day.
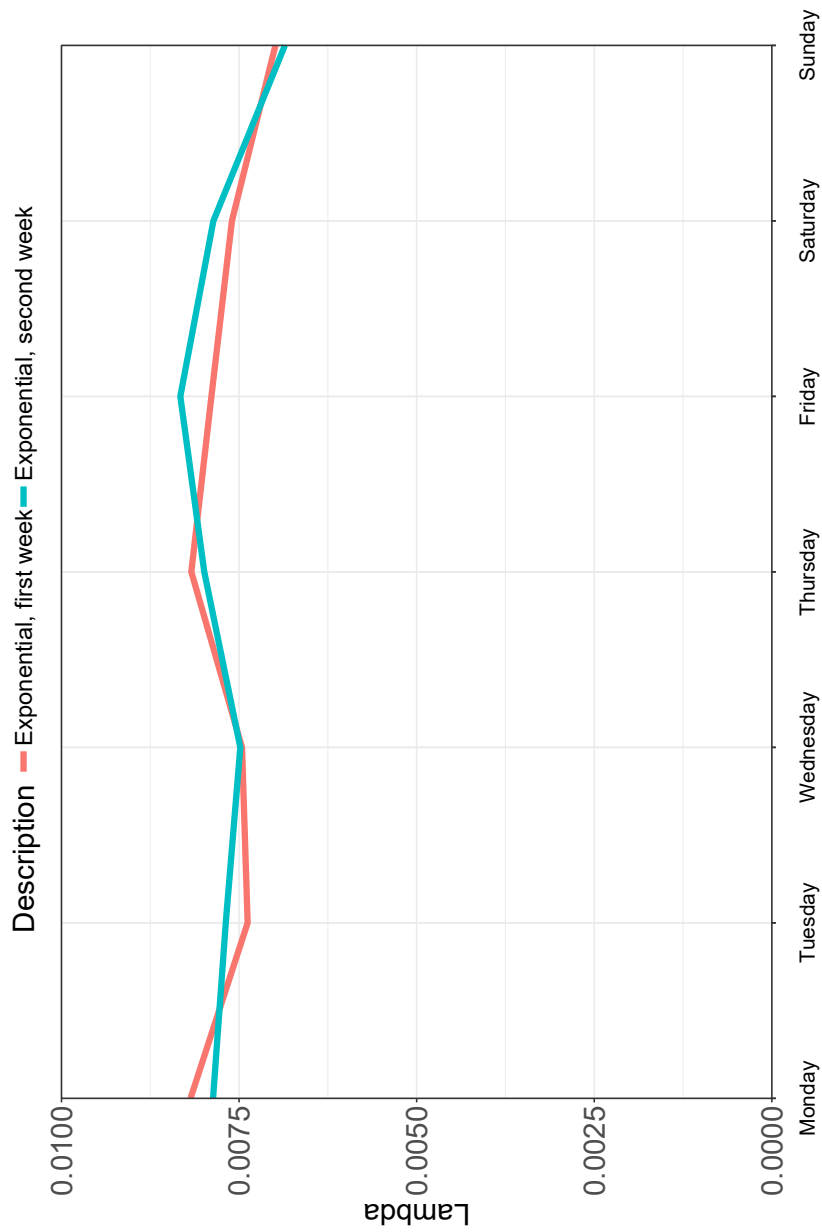
Figure 5.9: Average of the observed arrival rate parameter per day for all the routes, Lambda (λ_o), Exponential distribution, observed dataset.
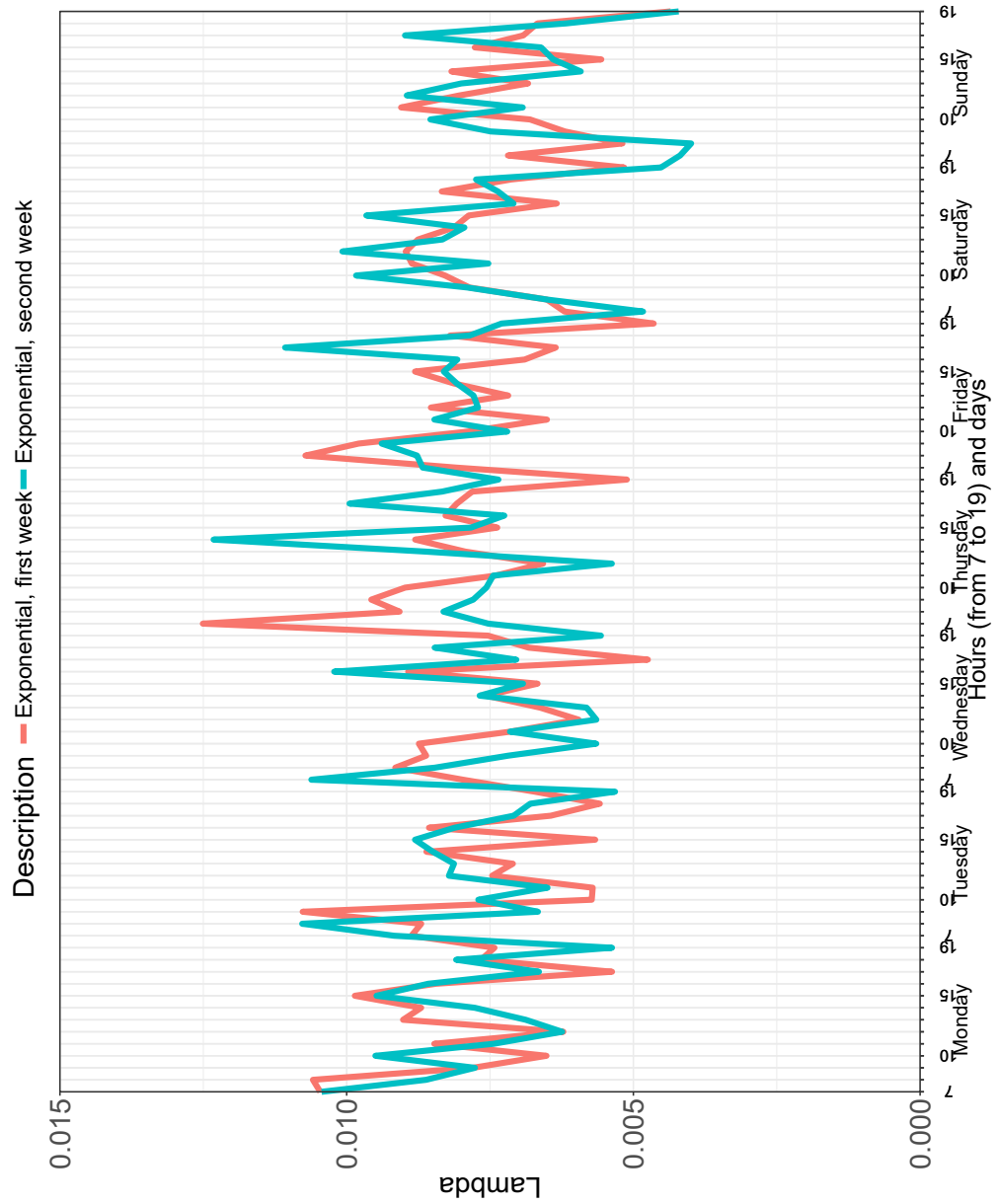
Figure 5.10: Average of the observed arrival rate parameter per hour per day for all the routes, Lambda (λ_o), Exponential distribution, observed dataset.

(a) Frequency

(b) Arrival times
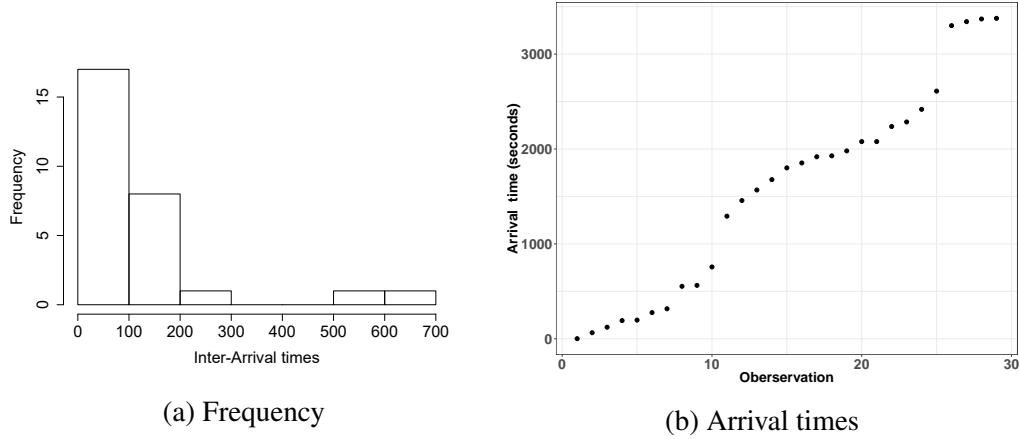
Figure 5.11: Adelaide Road, south direction, time 17:00 to 17:59, date=Monday, 2016-10-31, first week. 29 observation and the average observed IAT is about 130 seconds.

(a) Frequency

(b) Arrival times
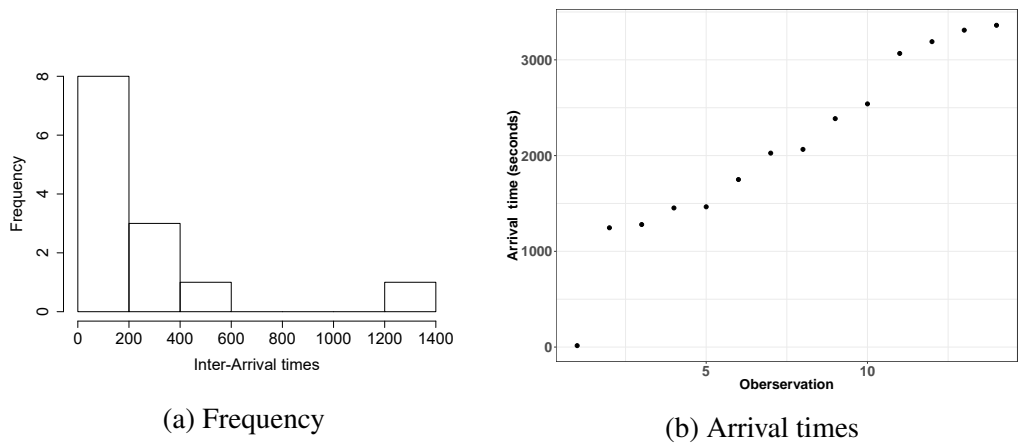
Figure 5.12: Adelaide Road, south direction, time 17:00 to 17:59, date=Monday, 2016-11-07, second week. 14 observation and the average observed IAT is about 260 seconds.

5.7 Conclusion

In this chapter, I statistically modelled the process of detection of a proportion of passing vehicles by BTMS. My results show that Exponential distribution is the only distribution among the Gamma, Lognormal, and Weibull distributions, which is correctly classified in the thinning process. This means that if the true dataset has the Exponential distribution, then the thinned dataset has also the same distribution. However, if the thinned distribution has the Exponential distribution, it is possible that the true dataset came from another distribution family. Also, for $\theta \leq 1$ thinned distributions are likely to be classified correctly from their true parent type. However for $\theta = 0.2$ or less (usually as is found in the real world, the average based on the literature review in Chapter 3, Section 3.3)) classification is unreliable. Moreover, these findings show that real thinned data do resemble Exponential data, suggesting that the IAT may be exponential (as has been found by other research).

Chapter 6

Transportation digital twin

A digital twin is a replica of the physical world in the digital world [119], [112], [6]. A digital twin models the behaviour of a physical object by its virtual twin [234]. The data for modelling of the behaviours of an object (i.e. physical asset) in the digital twin are collected by various sensors installed in the physical world. A digital twin can last for the entire life-cycle of an object and can be used to model the behaviour of an object in real-time by receiving and processing a live-stream of data, or replicate what has happened to the object in the past by receiving and processing the historical data. A digital twin can assist authorities to make smarter decisions based on the processing historical data and/or real-time data. The digital twin for monitoring traffic can be customised for management purposes. This concept is demonstrated in Fig. 6.1.

In this model, the physical twin submits the data to the digital twin for monitoring. In this digital transportation twin, received transportation data are cleaned and manipulated to be prepared for further processing and storage. The output of this digital twin can be demonstrated in forms of graphs and numerical summaries to inform traffic operation teams and local authorities about the characteristics of traffic flow. This digital monitoring twin can be connected to a digital management twin which allows managers and authorities to apply various scenarios on the almost live data and assess the outputs before applying those to the real-world.
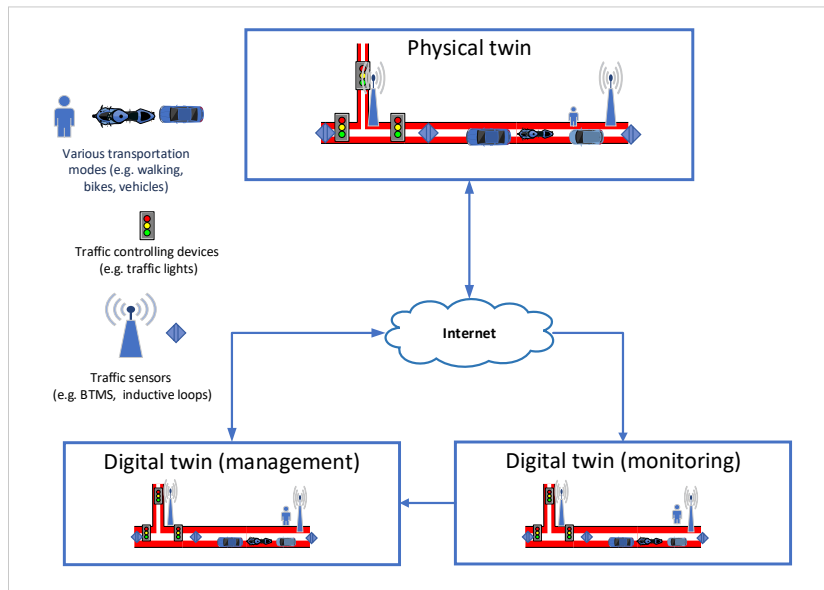
Figure 6.1: Physical and digital twins. These twins are connected by the Internet. Digital twin shall be developed in an elastic computing and powerful environment (e.g. cloud).

The development of a digital management twin is not carried out in this thesis, but is future work. Instead I only focus on modelling of the digital monitoring twin. Development of a digital twin requires strong backbone networking infrastructure and powerful processing units due to a large amount of the data [234], [183]. This approach cannot be fully real-time as there is the latency of data collection, data submission, and data processing. The delay for some traffic monitoring devices is higher than others as some are point-wise and some are route-wise. For example, inductive loops have a lower delay compared to BTMS. Inductive loops are typically point-wise traffic monitoring systems, are installed on the road, as shown in Fig. 6.2, and count the number of passing vehicles. These loops can also measure the time interval that each vehicle was on top of these loops (i.e. occupancy).

A Bluetooth counter in BTMS requires that a detected device reaches the successor Bluetooth counter (i.e. route-wise) before being able to classify the mode of transportation as BTMS detect detectable devices within the antenna range, which even may not be on the road. The successor detection in BTMS is required

This content is unavailable. Please consult the print version for access.

Figure 6.2: Inductive loops installed on a road [70].

because the travel time must be calculated for identifying the correct form of the transportation mode (e.g. walking, driving) and also for identifying the direction of movement. This concept is fully discussed earlier (Chapter 4). However, it may be possible to identify the direction of movement from the received signal by the antenna [160]. In contrast, inductive loops can determine the flow as soon as a vehicle passes on them (i.e. point-wise) as they are installed in the road and can only be triggered by vehicles.

This chapter focuses on building a digital transportation twin with the purpose of monitoring the traffic flow by utilising the data which are obtained by traffic sensors installed in the physical world. In this regard, the construction of this digital twin requires two sources of traffic data: BTMS and inductive loops.

In a traffic monitoring system, each type of traffic monitoring tool has some advantages and disadvantages. Inductive loops are dug into the road surface and need road closures for installation, and the maintenance process is not quick and easy as these are located under the asphalt, but can provide a complete detection of the passing traffic flow at the single point of installation (i.e. point-wise). Bluetooth Traffic Monitoring Systems (BTMS) are easier to install as these are not required to be installed on the road. Hence, no road closure is required for instal-

lation and maintenance. Consequently, the maintenance and installation of BTMS are easier, quicker, and cheaper than the inductive loops. However, these only detect a proportion of passing vehicles. Therefore, inductive loops can be utilised to assist BTMS to obtain an estimate of the passing traffic flow (i.e. true traffic volume) from the detected traffic flow (i.e. observed traffic volume of BTMS) on a micro route. In this case, in addition to the matrix of travel times for each Origin-Destination (OD) pair, which can be obtained by processing BTMS, the matrix of estimated flow (i.e. volume) for each OD will be available through the processing of the BTMS and inductive loops data. Therefore, both of these two technologies must be utilised in the physical and digital twins for monitoring the traffic flow and estimating the OD flow for each micro route. Achieving these requires finding the detection rate of BTMS with examining the inductive loops data. Specifically, in modelling of the transportation digital monitoring twin, this approach uses the data collected from inductive loops for calibration of BTMS data and estimating the detection rate (θ) and true flow (i.e. volume).

Since the processing of the digital twin is heavy and every single street does not have high importance in traffic modelling [161], this structure can be modelled for a small part of Wellington city as shown in Fig. 6.3 to imitate the real-world in a small-scale. In a real-world, there is no need to model traffic for every single street that is connected main streets, and as long as the information about the inflow/outflows over time are available, modelling can be done efficiently [257], [233].

For traffic modelling of the digital monitoring twin, this chapter uses SUMO (Simulation of Urban MObility). SUMO is an open-source software tool and is used in many academic and industrial projects [162]. This simulation tool can be integrated with other network simulators such as NS2, NS3, and OMNET++. It is widely used in computer science and transportation for modelling transportation and mobile nodes [243], [162], [108], [49]. In SUMO, roads are modelled through connecting various edges together, and a single vehicle moves in a set of specified roads (i.e. flow). SUMO is a microscopic simulation as each vehicle (or a mobile object such as a pedestrian) is modelled individually (e.g. brakes, acceleration,

Figure 6.3: Network topology of the physical twin and digital twin.

speed). The model in this chapter uses the default settings of SUMO for modelling the behaviour of a vehicle and traffic lights [12]. The next chapter explains the simulation details in more depth.

In this research I simplify the development of the digital monitoring twin using the following assumptions:

- One virtual Bluetooth network card per device
  Assume each device in the real-world can only have one virtual Bluetooth network card. A device is electronic equipment like cell phone, tablet, or on-board navigation system, and a Bluetooth network card is a network drive that enables Bluetooth communication. Virtual means that there is no need to model the full details of the process of wireless communications (e.g. signal processing, modulation, handshaking, TCP/IP). Therefore, the modelling of carrying Bluetooth devices is done by naming of an object as a Bluetooth carrier (through tagging the name, e.g. BTCar1). This name can demonstrate whether this object has a wireless card. To model the multi-tenancy detections (i.e. multiple detectable Bluetooth devices in a car), this research replicates a few trips and change their detection times by few seconds randomly to model the multi-tenancy detection, which is explained fully in the next section.

- All virtual Bluetooth network cards are detectable
  Since the possibility of having a virtual Bluetooth card is denoted through the naming convention (i.e. tagging by the name such as BTcar1) in this approach, all Bluetooth cards are discoverable. To model the detection, at each simulation step, the Euclidean distance of an object from the location of all the Bluetooth counters in the detector networks are calculated, and if the distance is less than or equal to 150 meters of a particular Bluetooth counter, then this device is detected by the Bluetooth counter.

6.1 Physical and Digital twin settings

Since the digital twin should be the replica of the physical twin, this section explains the traffic modelling of the physical twin below, which also demonstrates the assumed settings of the digital twin as shown in Fig. 6.4. In this Figure, the
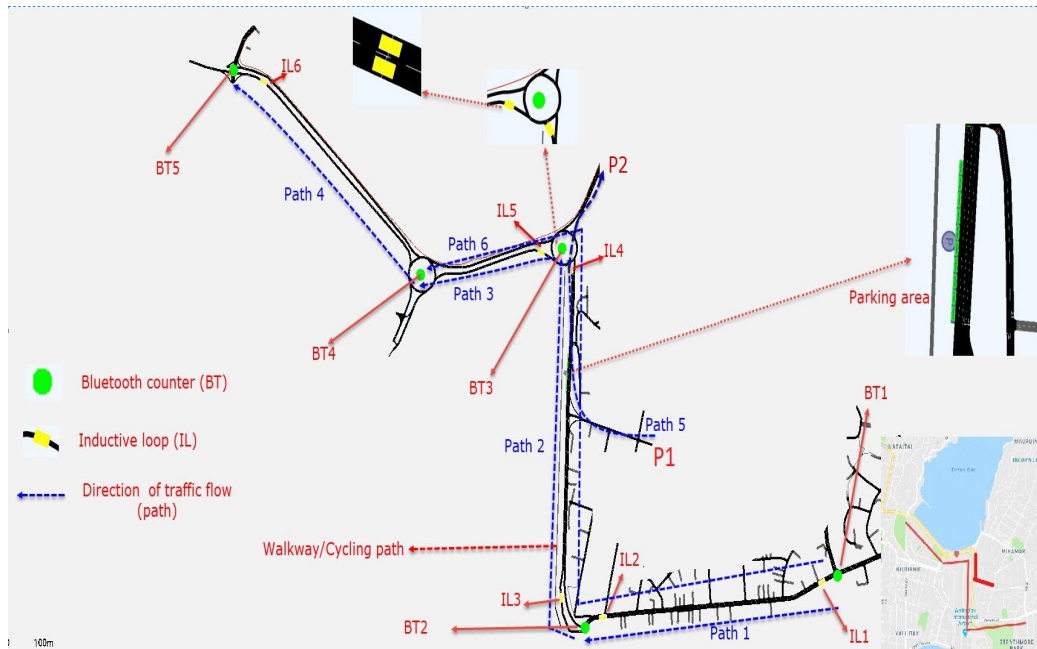
Figure 6.4: Settings of physical and digital twins. Traffic flows are shown by Path 1 to 6. The traffic flow of Path 6 starts at BT1 and ends at BT4. Path 3 and 4 model closed micro routes (i.e. no intersection which leads to no change in the passing traffic volume). Path 5 starts at P1 where no measuring device is available and ends at P2. Path 1 starts at BT1 and ends at BT2, and Path 2 starts at BT2 and ends at BT3. Inductive loops are shown by yellow rectangles and Bluetooth counters are shown by filled green circles (or points).

directions and paths of traffic flows (origin-destination or micro routes) are shown by Path 1 to Path 6. The beginning of a blue arrow (i.e. a path) is the source and the end of an arrow is the destination of the traffic flow. Traffic flow is the movement of vehicles along each path. For example, the movement of vehicles in Path 1 start at a location close to BT1 and end at a location close to BT2 and so forth. The flow of Path 5 starts at P1 and ends at P2 after passing IL4. BT1 to

BT5 show the location of Bluetooth counters installed on the physical twin. IL1 to IL6 show the location of inductive loops (as pairs, one per lane). A walking/cycling path is also shown as a parallel line to Path 2 along with a parking area on the left side of Path 2. Although this is a sample and small network due to the limitation of resources, but each component models specific behaviours of the real-world (i.e. physical twin). These components are explained in the list below, and shown in Fig. 6.5. These components of the Physical world (i.e. digital twin) are shown (throughout this paragraph) on the left (green boxes), the approach for modelling these components of the physical twin in the digital twin are plotted on the right (blue boxes), and the directed arrows show the relationship between these components. For example, there are some mobile nodes such as bikes and emergency vehicles (see Chapter 3) that can have different movement patterns with vehicles in a real-world and these are shown through tagging as Minority in the modelling. Another example is the challenge of estimating the detectability ratio (see Chapter 5) in a real-world which is modelled through utilising the flow of Path 3 and Path 4. However, the mapping details are fully discussed below. In addition to modelling of these components of the physical twin, the flow of this twin must be also modelled in the digital twin. The flow is modelled through the orange boxes which are located at the bottom of this figure and represent the foundation of the transportation system, which is the traffic flow. The purpose of the items listed below are to demonstrate how this model can utilise the inductive loops and BTMS data to obtain the matrix of traffic flow for each OD pair, and how these components must be modelled in a digital twin to be a mirror of the physical world.

Each of these components is fully discussed below:

1. As demonstrated in Fig. 6.4, vehicle flows exist in Path 1, 2, 3, 4, and 5. These vehicles do not stop anywhere in these paths, form regular traffic flow, and denoted as Non-stopping vehicles.

2. In addition to the above vehicles, there might be some cases that a vehicle stops somewhere on the route for various reasons (e.g. shopping, dropping
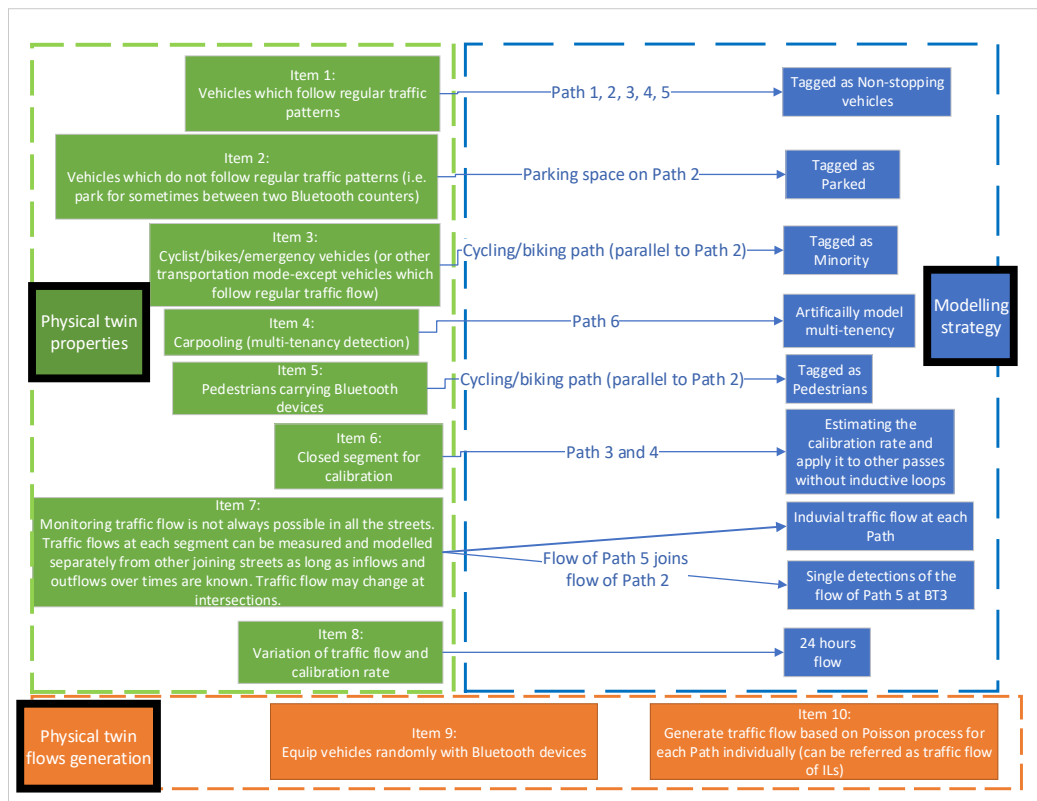
Figure 6.5: The components of the physical twin (left, green boxes), the methodology for modding these components in the digital twin (right, blue boxes), the interaction between them (directed blue arrows), and the foundation of a transportation system (flow, orange boxes).

off, picking up). To investigate the impact of these stops on the identification of the transportation mode (e.g. static/bikes/pedestrians, discussed in Chapter 3) through examining BTMS data, this approach simulates a parking area where a few vehicles stop for 20, 40, 60, 90, or 120 seconds (denoted by *parked* in the results). This parking area is marked on Path 2 in Fig. 6.4.

3. To model the misclassification of transportation mode, I insert a few mobile nodes on the walking/cycling/emergency vehicles path (Parallel to Path 2) with various speeds and tagged these as Minority. This miss-classification may happen on some roads, where the congestion is very high, and the movement of cyclists (or pedestrians) is similar to the slowed movement of vehicles. An alternative scenario is that a few cars, bikes, or emergency vehicles drive faster than most of the cars. This is modelled by moving objects (denoted by *Minority* in the results), which move with the speed of 15, 30, 70, or 100 km/h. The differentiation of the path of vehicles, bikes (or cyclists), and pedestrians is done only for the purpose of visualisation.

4. In order to model multiple-detection in a vehicle without using wireless communication, these 10 (can be more, but will impact the procedure and output) trips from a location close to BT1 with the destination of a location close to BT4 are generated. This flow passes BT2, BT3, IL1, IL2, IL3, and IL4 along its route and shown by Path 6 in Fig. 6.4.

5. In order to model the detection of pedestrians by BTMS, a flow of pedestrians with Bluetooth devices moving from BT2 to BT3 on the walking/cycling path with a Poisson process with the speed of about 4 km/h is generated.

6. Previous research mainly accomplished the calibration (i.e. estimating the Bluetooth detection rate, θ) of BTMS through examining the traffic flow on an hourly basis for a closed segment of a road (i.e. part of the road without any intersection like a bridge or a tunnel), where mostly motor vehicles

are allowed and flow cannot change at intersections [129], [115], [224]. Alternatively, the calibration can be done on routes where image processing can recognise the licence plates of vehicles to uniquely associate them with a detected MAC address collected by BTMS (e.g. toll locations) [259]. In the data collected by BTMS, it is not possible to identify the transportation mode (e.g. walking, cycling, driving) accurately for all passing objects only by examining a single detection. A typical method is when a device is detected a second time by another Bluetooth counter. Then, based on the traffic characteristics (e.g. distance, speed, travel time) between these two points of detection (Bluetooth counters), the mode of transportation may be deduced. This setting is shown by Path 3 and 4 in Fig. 6.4, where both are closed micro routes without any intersection (i.e. the traffic flow/volume along the route does not change).

7. The traffic flow of some streets of an urban area is not the point of interest, as they have low priority in traffic modelling, so there is no need to monitor the traffic flow for every single street of a city. For example, the streets which join a highway have a much lower priority compared to the highway. Therefore, the focus is mainly measuring the inflow and outflows of the streets which are connected to the highway (i.e. ramps) over time. As long as the flows which are entering and exiting the highway over time are known, there is enough information to model the traffic flow in the highway, and no need to model those streets. The same logic can be applied to an urban street. Since in an urban street, an intersection can be considered as the main point that the volume changes (vehicles enter and leave), therefore it is crucial to measure these changes for traffic modelling over time. Moreover, a long street may be composed of several intersections in which only a few of these segments (i.e. part of the street between two intersections) may have high priority in traffic modelling. An example is shown in Fig. 6.6. In this figure, the traffic flow is not consistent over the length of Adelaide Road. The flow at the South end is lower (shown in blue) than the North side (shown in orange). This is due to changes in traffic flow at the

intersection between Adelaide Road, Riddiford Street, and Wallace Street (shown by a black rectangle). Therefore, the North section of this street may have a higher priority in modelling as this is the location that the traffic volume changes and demand increases, whereas the South part may have lower priority in the modelling due to lower demand.

These two examples highlight that the focus of modelling should be segments of streets with high priority, the flow may change at intersections, and the flow for each segment of a street can be modelled separately as long as the inflow/outflows are known.

In modelling the twins, it is assumed that the traffic flow does not change at every single intersection, and only changes at the main intersections (i.e. where traffic monitoring devices are installed), and also where the flow of Path 5 joins the flow of Path 2.

Since it is almost impossible in the real-world to monitor the traffic of every single street, it is therefore possible that a vehicle may be detected at a Bluetooth counter and then proceed to follow a route which does not have a Bluetooth counter (due to a lower priority), although it may be detected at some other locations later. An example of this case is demonstrated by Path 5 in Fig. 6.4, where vehicles are detected at Bluetooth counter 5 (BT5), but they are not detected at any other location during my modelling. This is because the flow of Path 5 starts at location P1, where there is no traffic monitoring system on the first part of this path (i.e. before joining the flow on Path 2), and ends after passing the first roundabout (close to IL4) at Location P2, where Location P2 does not have any traffic monitoring device. Since this flow passes IL4, therefore, vehicles in this flow are detected at IL4 and partially at BT3 (when a car is tagged as a Bluetooth carrier). This flow does not enter Path 3 (Between Bt3 and BT4). Therefore, the difference in measured flow at IL3 and IL4 demonstrates the flow of Path 5.

8. The traffic flow for this network for a random day (24 hours, denoted by 0 to 23) is modelled as a prototype, although I am aware that the traffic flow
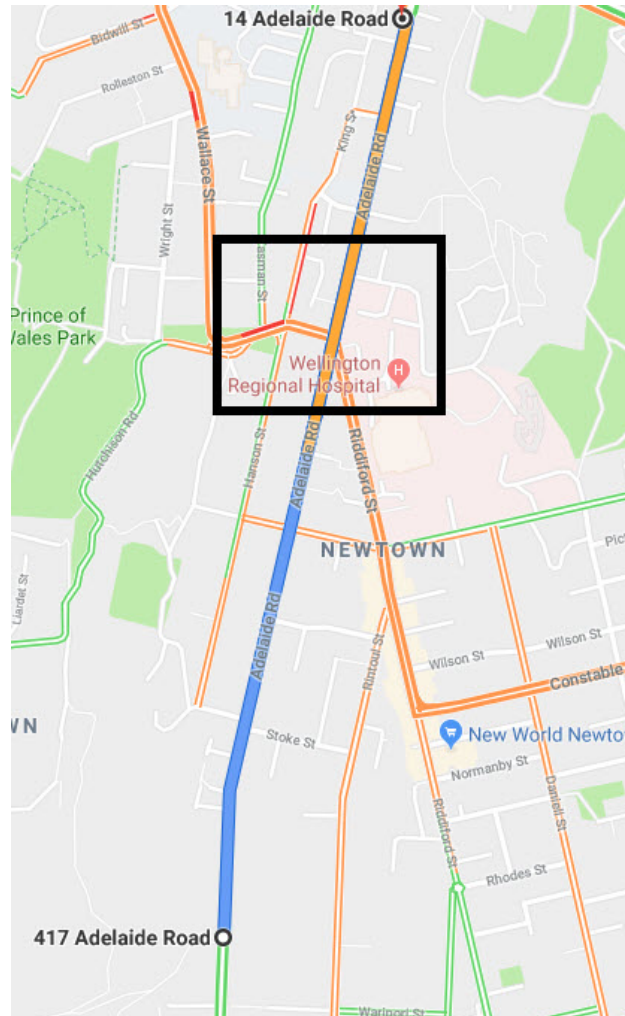
Figure 6.6: An example that demonstrates that the traffic flow may change significantly at an intersection, Google Maps. The North side of Adelaide Road has heavier congestion (shown in orange) compared to the South side (shown in blue). This is because the flow changes at the intersection between Adelaide Road, Riddiford Street, and Wallace Street (shown by a black rectangle).

for weekends, weekdays, school holidays and public holidays are different from each other (an example is shown in Chapter 7, Fig. 7.19).

9. Modelling carrying a Bluetooth device is done through randomly (with Bernoulli distribution) equipping mobile objects with a Bluetooth network card (i.e. tag the name) with the detectability rate between 0.5% and 40% (based on the literature review as discussed in Chapter 3, Section 3.3), for Paths 1, 3, 4, and 5. The flow of Path 6 was explained in Item 4. For Path 2, I also apply a logic that the morning and afternoon peak hours have a higher detection rate as the congestion is heavier during these peak hours and possibly more people use their navigation systems to find better routes. This is done for the purpose of simulation and visualisation and can be changed. In this modelling, the detectability rate is constant within an hour.

10. This model generates multiple traffic flows at various micro routes (Path 1 to 5) by a Poisson process. As inductive loops count the number of passing vehicles, the measured flow of each set of inductive loops (i.e. IL1 to IL6) demonstrates the inserted flow in the simulation environment (also in the physical twin as physical and digital twins are replicas of each other). These flows are demonstrated in the section which is allocated to the results.

A screenshot of the cyclist, pedestrians, two types of vehicles (with detectable Bluetooth (BT) device, without detectable Bluetooth, BT, device), and parked vehicles in the simulation environment is shown in Fig. 6.7.

6.2   Simulation results

Initially, this section examines Path 6 to discuss the simulation of multi-tenancy detection. Then, this section examines micro route 2 (from BT2 and BT3, shown as Path 2) in detail as this is a part of the network that multiple scenarios occur (i.e. change of traffic flow, pedestrian flow, vehicles which stop at the parking area). Later, this section discusses the details of the calibration approach and also the demonstration of the generated flow in the last subsection.
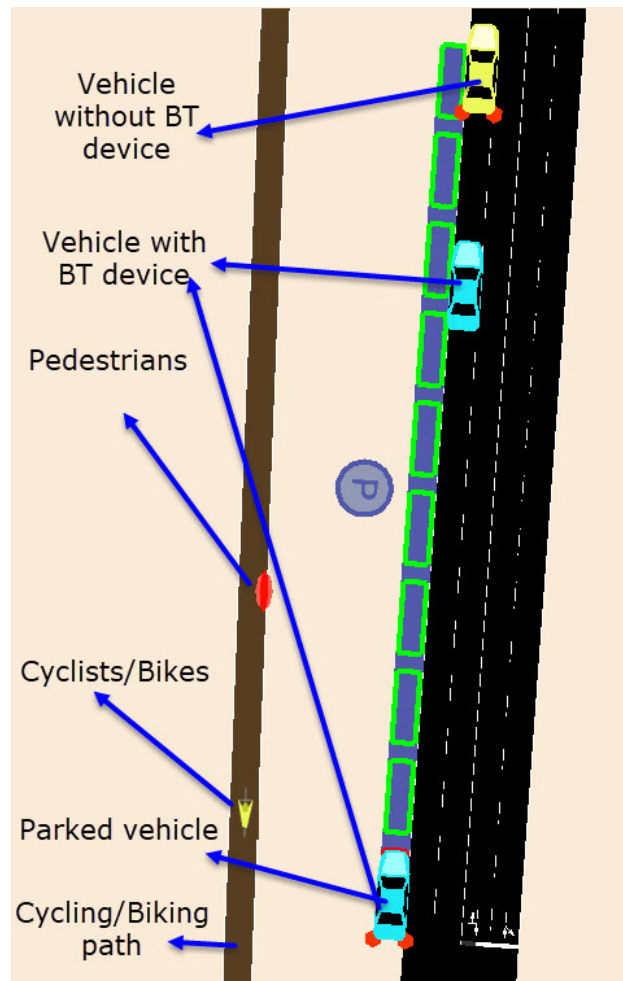
Figure 6.7: Cyclist/bikes, pedestrians, vehicles in SUMO in the digital twin.

6.2.1 Path 6

The ExtoVT algorithm (the cleaning process) contains an algorithm, Multi-Tenancy Detection in the collected Bluetooth data (MTDiBT), for multi-tenancy detection. The details are fully discussed in Chapter 4 Section 4.3. MTDiBT examines the collected data and clusters various detections and identifies when Bluetooth devices are seen together when the difference in their observation time is equal or less than a specified threshold time interval τ. To have a confirmed detection, which suggests Bluetooth devices are moving together, at least two consecutive detections are required ($\alpha \geq 2$) and in both cases a detector must register a putative pair of devices within the threshold interval of τ. To demonstrate these 10 vehicles which travel from BT1 to BT4 (Path 6) are generated as shown in Fig. 6.8. τ is set to 1 second and α to 3 detections for MTDiBT algorithm. X axis shows the index number (i.e. car label), just for the visualisation purpose, and Y axis shows the detection time. The colours and shapes demonstrate the location of the detection (BT1 to 4).

Since, modelling wireless communication (e.g. handshaking, wireless communication protocols is known as IEEE 802.11 standard [206]) is out of the scope of my research, and without wireless communication, it is not efficiently possible to equip a vehicle with multiple wireless network card, 4 of these cars are selected (index 4, 5, 6, 7) randomly and replicated a few times randomly (one time for indexes 4, 6, and 7 and three times for index 5). During the replication, I changed the detection time by a few seconds to model the multi-tenancy detection. After this replication, it is not clear which device corresponds to a car and which corresponds to a tenant, and this is the reason for referring to each as a mobile node instead of a car. These new mobile nodes along with their detection times are shown in Fig. 6.9 and provided to the MTDiBT algorithm as the input. Detection times are shown as a label at each location of detection in this figure and each new object is shown by an index. The difference in detection times between two symbols/colours for each index (i.e. car), on Y axis, demonstrates the travel time between two consecutive Bluetooth counter (i.e. travel time of a micro route).

After applying MTDiBT, this algorithm identified the following objects as co-
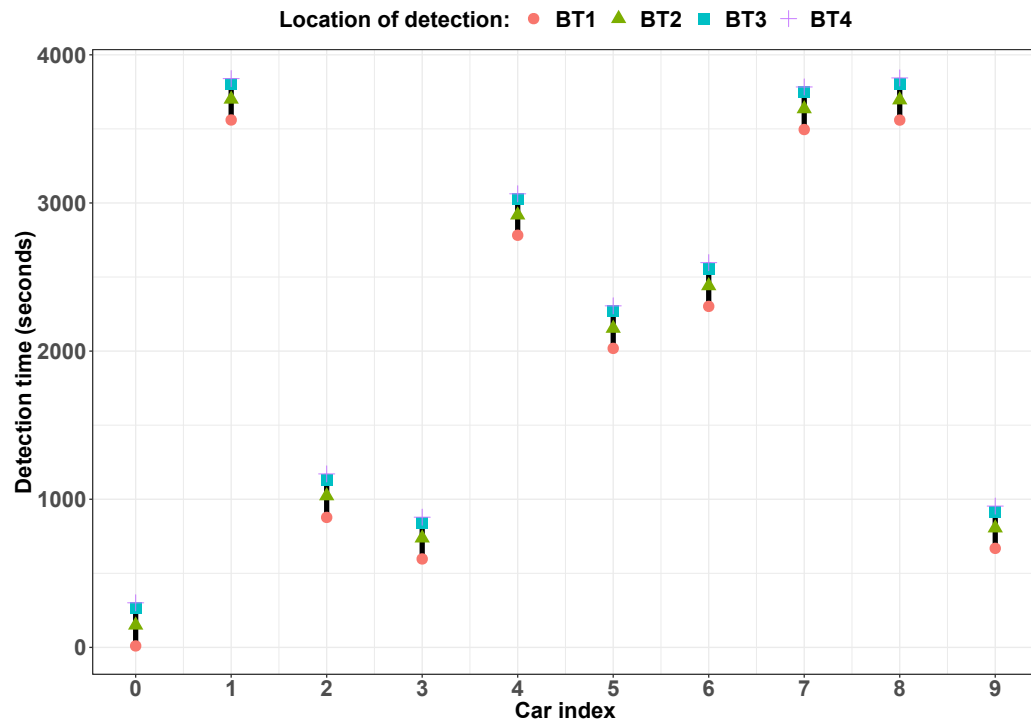
Figure 6.8: 10 trips for 10 vehicles (index 0 to 9) from BT1 to BT4 (Path 6). X axis shows the indexing number (car label), Y axis shows the arrival time (i.e. detection time) at each specific Bluetooth counter. Since the flow of Path 6 starts at BT1, ends at BT4 and passes BT2 and BT3, therefore, the color and symbol demonstrate the detection time (i.e. arrival time) at each Bluetooth counter. The difference of detection times between two symbols/colours for each index (i.e. car), on Y axis, demonstrates the travel time between two consecutive Bluetooth counter (i.e. travel time of a micro route).
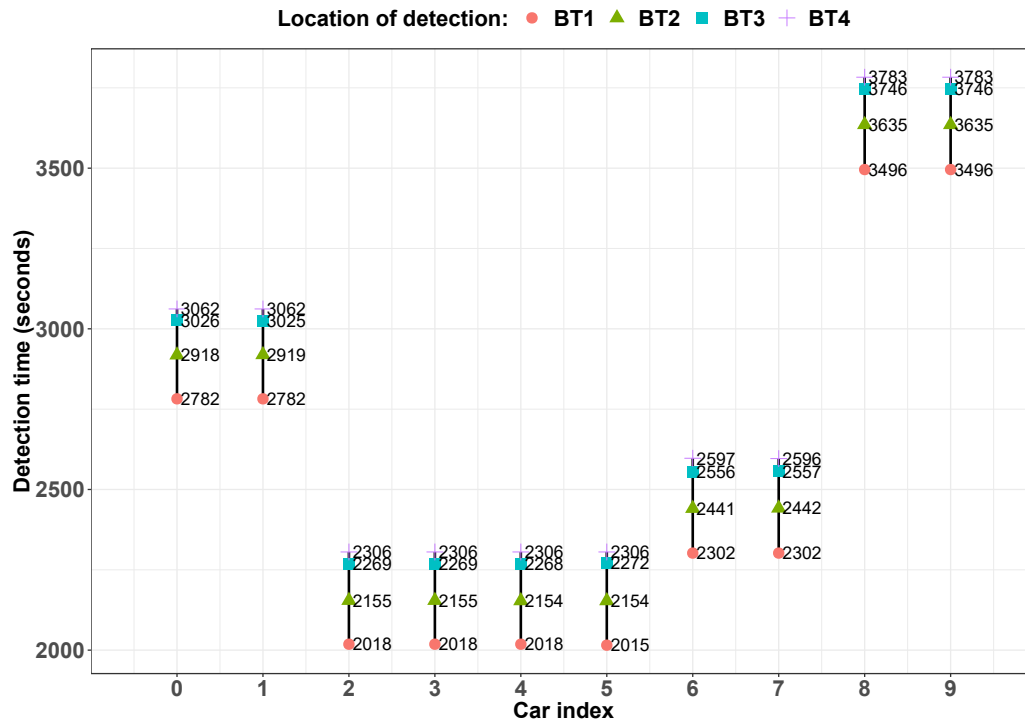
Figure 6.9: Simulating multi-tenancy vehicles. Data labels show the detection time at each location. The mobile node with Index 5 is not travelling with mobile nodes 2, 3, and 4 as it is detected 3 seconds before the rest of the group (indexes 2, 3, and 4) at location BT1 (2915 vs 2018 seconds), and 4 seconds later at location BT3 (2269 vs 2272 seconds). Therefore, mobile node 5 is a separate car and 2, 3, and 4 are within the same car (i.e. co-detections). Mobile nodes 1 and 2 are co-detections. 6 and 7 are also. 9 and 10 are also.

detections:

- index 0 and 1
  These two mobile nodes are seen together in more than three locations where $\tau \leq 1$ second in each location. Therefore, possibly this is a co-detection and these two mobiles nodes are tenants of a car.

- index 2, 3, 4
  These three mobile nodes are seen together in more than three locations where $\tau \leq 1$ in each location. The mobile node with index 5 is not categorised as a co-detection with mobile nodes with indexes 2,3, and 4 as the detection time for this mobile node (index 5) at BT3 is 2272, where the difference in the detection (τ) time is greater than the threshold (i.e. 1 second). Furthermore, the detection time at BT1 for this mobile node has a difference greater than 1 second with the rest of the cluster members (cars with indexes 2 to 4). Therefore, mobile nodes 2, 3, 4 are co-detections and possibly are travelling together within one car, and mobile node 5 is in another car.

- index 6,7
  Similar to the first item.

- index 9,10
  Similar to the first item.

As the next step after identifying the co-detection devices, this algorithm keeps one of the co-detections and eliminates the rest. In this case, the following mobile nodes are left in the dataset for the further processing: 0, 2, 5, 7, and 8, and the rest are eliminated. The next subsection focuses on Path 2.

6.2.2   Path 2

Fig. 6.10 shows the travel time and arrivals of the detected Bluetooth carriers (i.e. a mobile object with detectable Bluetooth devices) at the micro route between BT2 and BT3, as observed (i.e. without any cleaning process). Values on the Y

axis of this figure are travel times and arrival times are shown on X axis. The arrival time is the time that each device is detected at a second (or successor, see Chapter 3) Bluetooth counter and travel time is the time difference between two consecutive observations (i.e. arrival time − departure time). Since the flow in Path 5 is only detected at BT3 and there is no traffic monitoring device at P1, no second detection is reordered for the vehicles in Path 5. This means that these vehicles from Path 5 only have one record in the collected BTMS data. Since these vehicles only have one record, the travel time cannot be calculated for these vehicles as travel time requires the subtraction of two observations sorted by time. Therefore, these single detections must be removed from the collected data.

The pessimistic (350 seconds as reported by the Google API) and optimistic (60 seconds) thresholds for this path remove some of the detections (mostly pedestrians and single detections) as shown in Fig. 6.11, although some minority/parked mobile objects are still in the valid range. This step was the initial part of the Ex-toVT algorithm (cleaning algorithm) proposed in Chapter 4.

The next step applies Tukey's logic in accordance with the ExtoVT algorithm to identify the remaining outliers. The result is provided in Fig. 6.12. The window for the Boxplot is one-hour intervals (from 0 to 23). As displayed in Fig. 6.12, some of the minority/parked data (i.e. belonging to mobile nodes, shown by red points) are not detected as outliers (e.g. *path2_Minority_15kmh_18* at hour 10) as their travel time was similar to the patterns of the typical traffic flow. After removing outliers, shown by blue points, this dataset demonstrates the typical traffic flow collected by BTMS.

The next subsection discusses the calibration approach by using collected inductive loops data.
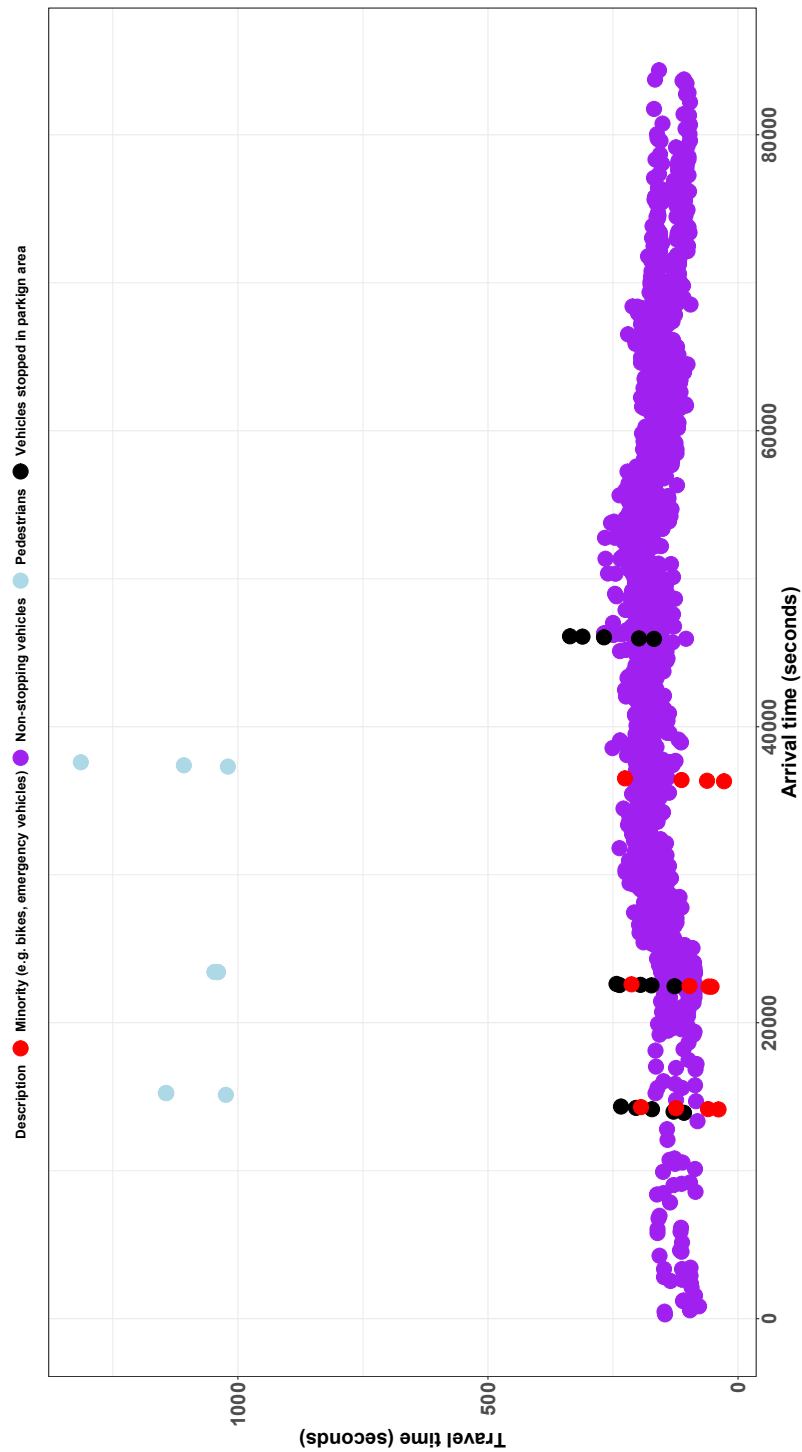
Figure 6.10: Observed travel times (without any cleaning process) at the micro route between BT2 and BT3. The light blue points on top show the travel time of pedestrians. The purple points are the travel times of vehicles which did not stop anywhere on the Path 2. The black points are cars which stopped in the car park, and red points are the Minority mobile nodes (e.g.cyclist, emergency cars).
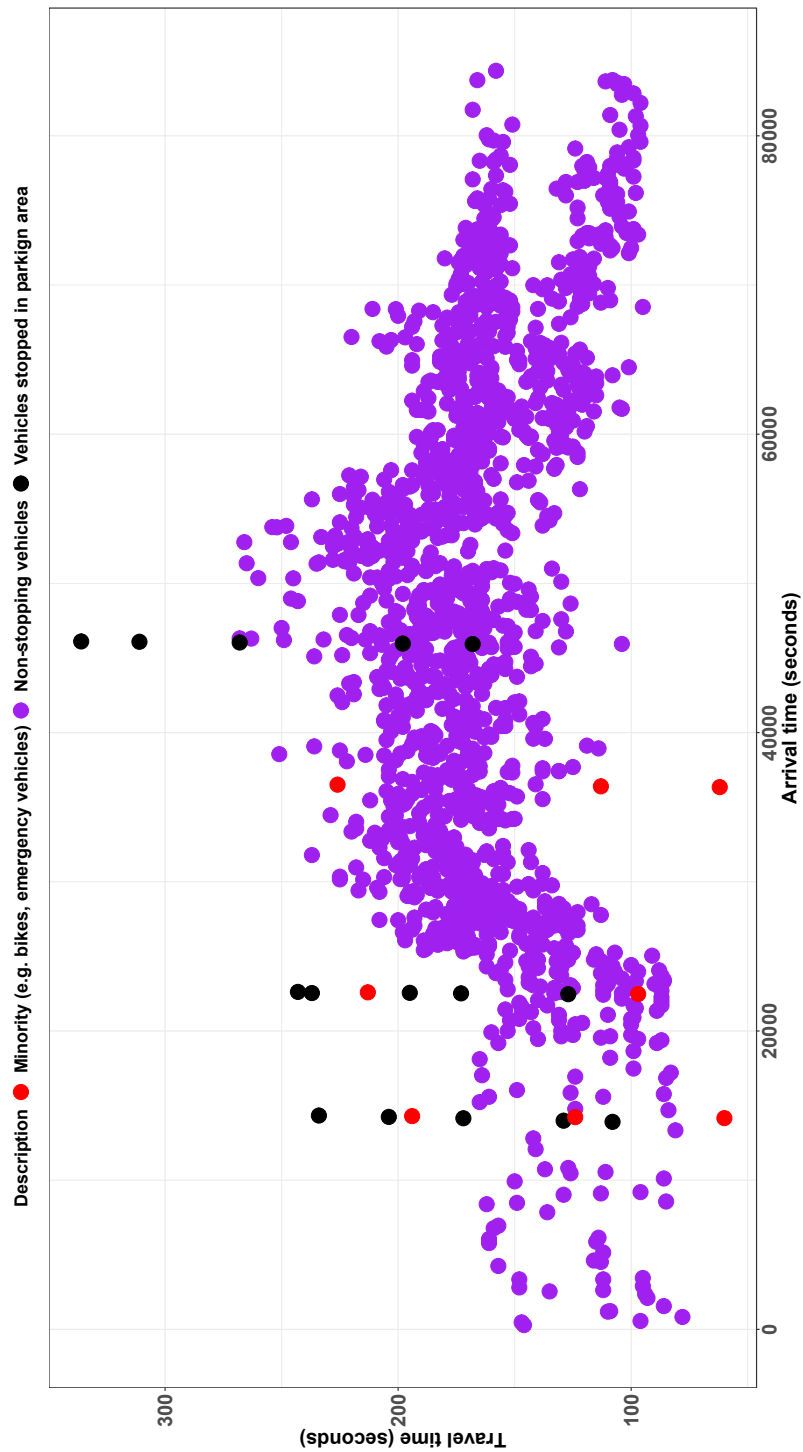
Figure 6.11: After removing values higher than pessimistic and lower than optimistic bound. Pedestrians and single detections are identified and removed but some Minority/parked mobile objects are still in the data.
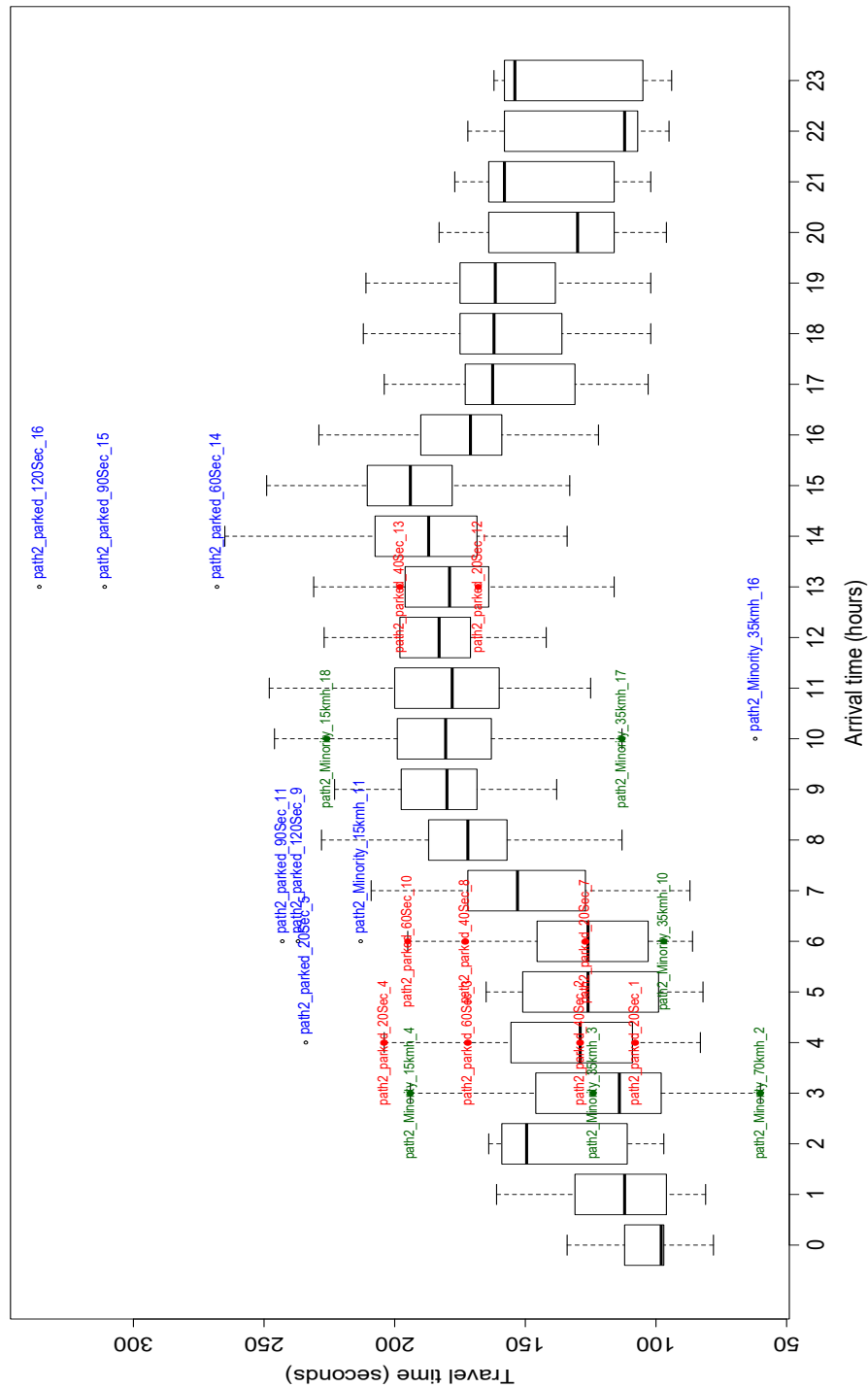
Figure 6.12: After applying Tukey's logic on the data. To make spotting the Minority/parked vehicles easier, red and green points are added to the boxplot to show the location of Minority and parked mobile nodes, respectively. Detected outliers by Tukey's logic is shown by blue points. Some of the Minority/parked data have similar travel time to the typical traffic flow and are not recognised as outliers.

6.2.3   Detection rate and calibration

In this section, the detection rate is defined as follows:

$$\theta_t = \frac{F_{BT_t}}{F_{IL_t}}, \tag{6.1}$$

where t is the measurement period (e.g. hourly, daily, weekly, monthly), F_{BT_t} is the measured total Bluetooth volume for the measurement period t, and F_{IL_t} is the measured total number of induction loop detections for the measurement period t. The estimated true volume (f) based on the F_{BT_t} and θ for measurement period t is defined as:

$$f_t = \frac{F_{BT_t}}{\theta_t}, \tag{6.2}$$

and in this chapter t is set as 60 minutes (an hour).

In a real-world settings there might be a few rare occasions that there are enough resources and inductive loops can be installed next to each Bluetooth counter and in every location that the traffic flow changes to obtain an accurate estimate of traffic flow, although usually these inductive loops cannot track vehicles over the city, while BTMS can. In this way, every intersection (i.e. flow changes) is artificially converted to a closed micro route (i.e. artificial closed micro route) by installing two Bluetooth counters (one at the beginning and one at the end of the path between two intersections) and one set of inductive loops. Or more realistically, inductive loops are only installed in fewer locations, on closed micro routes (no intersection between two consecutive Bluetooth counters like the path between BT3 and BT4, and the path between BT4 and BT5). The second approach can be called a natural closed micro route and some examples are: a bridge, a segment of a motorway between two junctions (a route with no entry or exit points for traffic), or a tunnel. Therefore, this section models both of these approaches. Since in the simulation environment, inductive loops are installed at every part of a road where flow changes, I can therefore demonstrate both of these approaches in my modelling.

In the first approach when inductive loops are installed at every lane of each

intersection, and there is no processing and financial limitation, then the estimated flow from BTMS, f_t, overlays the measured flow of inductive loops, F_{IL_t}. This happens as for each micro-route (the path between two Bluetooth counters), there is at least one inductive loop. Therefore, the detection rate can be accurately calculated as the flow (i.e. volume) does not change in these closed micro routes (no vehicle joins or leaves), and all vehicles are detected by the inductive loops installed on these particular micro routes. This is unlikely in a real-world due to limited resources.

In the second approach which is more realistic, the average measured θ for multiple closed micro routes is used to estimate f_t of other micro routes without inductive loops. This does not provide an accurate estimation of traffic flow for all routes (compared to the first approach), but this is a more practical way of estimating f_t. Accuracy can be increased by installing more inductive loops where there are enough resources. Also, a smoothed mean or other statistical methods can be used to enhance the accuracy of the estimation. As an example, this section uses the rolling mean (4 previous points, with replacing the null value of the first three points with the fourth one) for the second approach to smooth the detection rate over multiple closed routes.

The result of the first approach is shown in Fig 6.13. In this approach the estimated true flow, f_t, overlays the measured flow of inductive loops, F_{IL_t} as each micro-route has at least one set of inductive loops (one per lane). Hence, only the BTMS flow and inductive loops flow are demonstrated. Since IL4 measured the flow for both Path 2 and Path 5, and IL3 measured the flow for Path 2, these two flows are shown in Fig. 6.14. This figure uses the flow of IL3 as the representation of the flow of Path 2 as the vehicles of Path 2 were on this route for the whole duration of their trips.

For the second approach, Paths 3 and 4 are closed micro route as there is no intersection within these routes (flow does not change). The average detection rate, θ, of these micro routes is obtained along with the smoothed mean with a rolling window (as shown in Fig. 6.15) and is used for the calibration of Path 1 and Path 2. Results of the calibration for both mean and smoothed mean of the

(a) Frequency of Path 1

(b) Detection rate of Path 1

(c) Frequency of Path 2

(d) Detection rate of Path 2

(e) Frequency of Path 3

(f) Detection rate of Path 3

(g) Frequency of Path 4

(h) Detection rate of Path 4
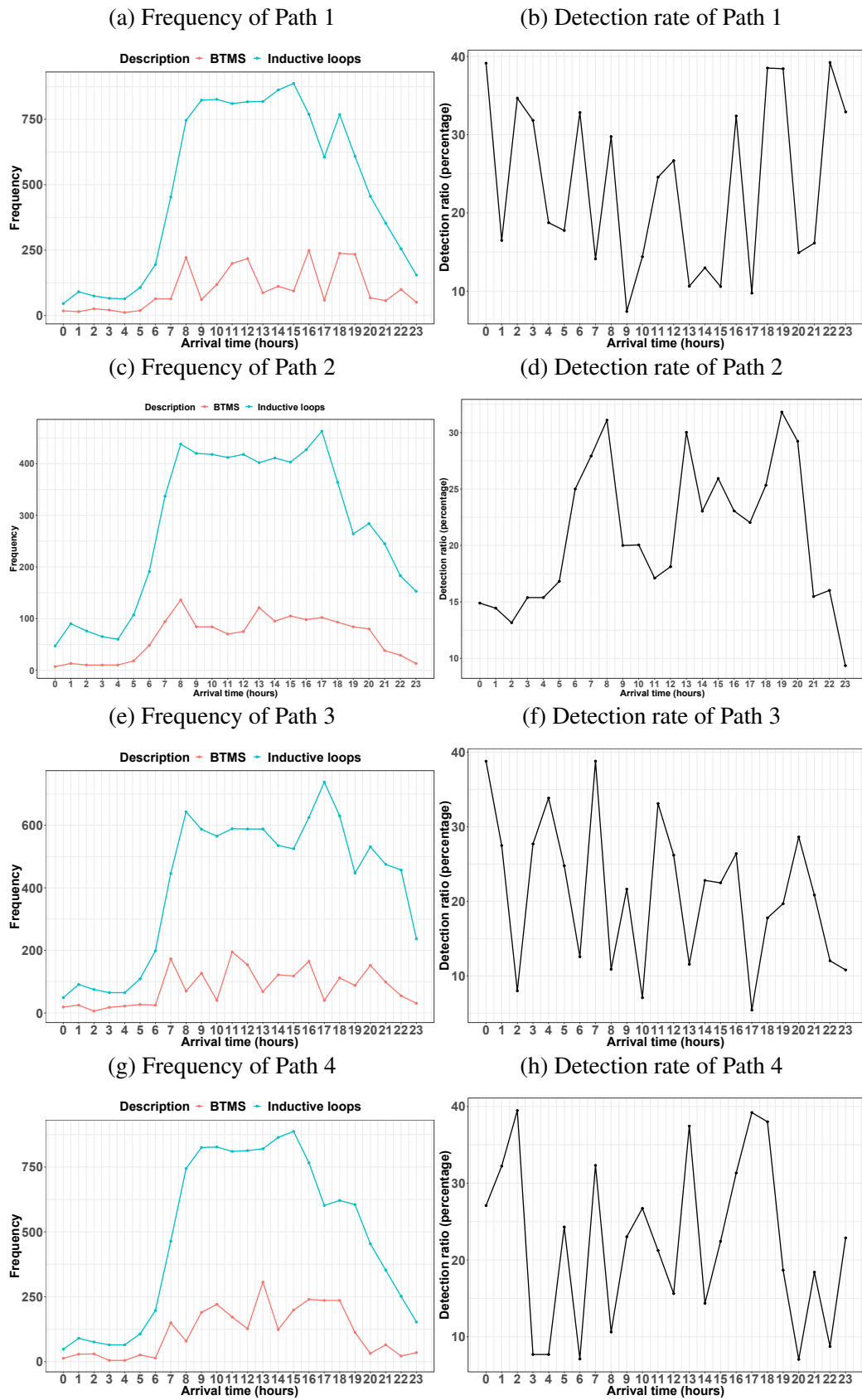
Figure 6.13: Total flow (i.e. Frequency of the detected vehicles) and the detection rate of Path 1, 2, 3, and 4. The estimated true flow overlays the inductive loop flow as this is the first approach (i.e. unlimited resources).
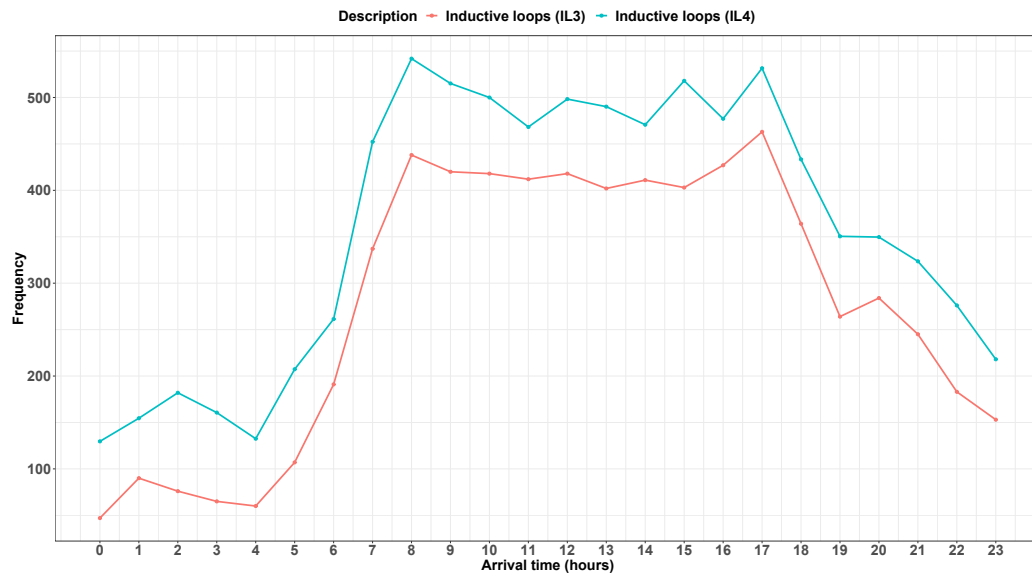
Figure 6.14: Detected flow by inductive loops at locations IL3 and IL4. The measured flow at IL3 shows the number of vehicles which used this path to reach to BT4 from BT3. Therefore, the flow of IL3 is used for the calibration.

average rate are shown in Fig. 6.16.

6.3 Conclusion

In this chapter, I proposed a monitoring transportation system as a digital twin and demonstrated how the monitoring and calibration of BTMS can be achieved. Moreover, I explained the key components of a physical twin and how these key elements can be modelled in a digital twin to represent a true replica of the physical world (i.e. digital twin). For constructing the OD matrix of estimated flow (i.e. calibration), I also proposed several approaches to estimate the flow for each micro route based on the collected data from BTMS and inductive loops. This model can be implemented in a cloud environment, where elastic computing and huge storage are not constraints, and can assist traffic operation teams and local authorities to monitor the traffic flow and make smarter decisions based on the provided outputs from this model through the whole life-cycle of transportation
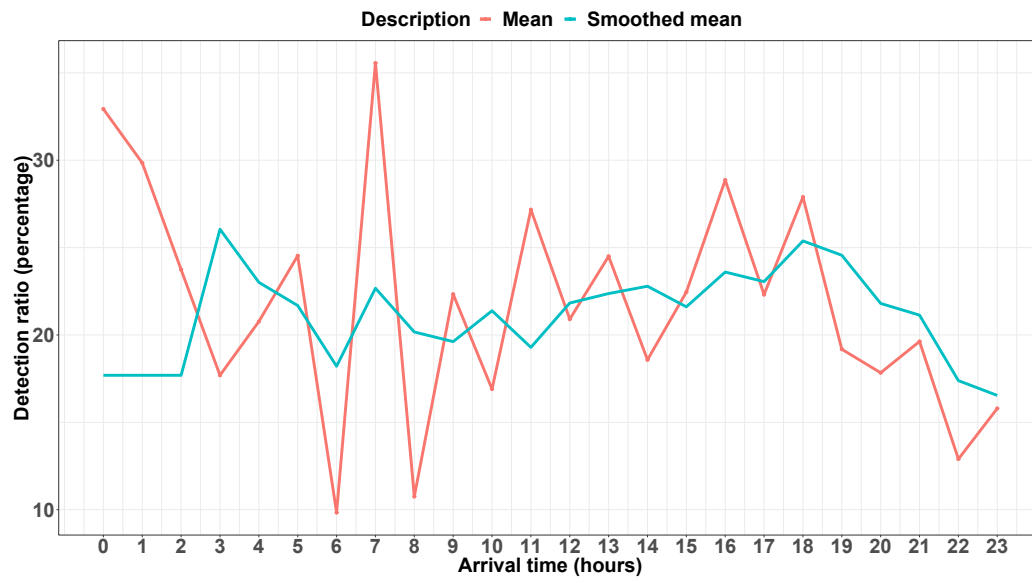
Figure 6.15: Hourly average θ of Paths 3 and 4, obtained from Fig. 6.13f and 6.13h.

assets.
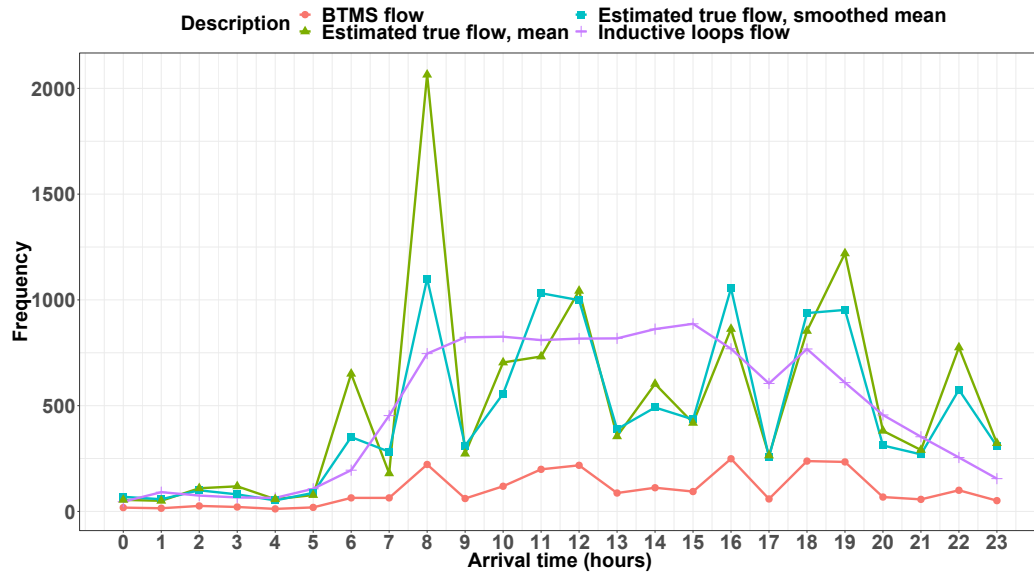
(a) Frequency of Path 1
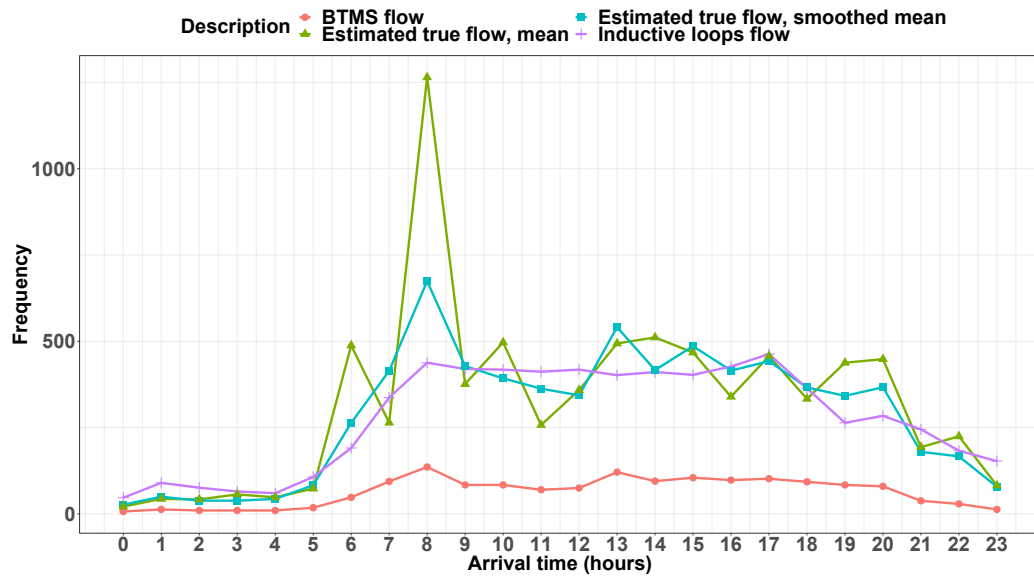
(b) Frequency of Path 2

Figure 6.16: Measured and estimated flow of Paths 1 and 2 by using the mean and smoothed mean of detection rate of Path 3 and 4 (displayed in Fig 6.15).

Chapter 7

Effects of the 2016 Kaikoura earthquake on traffic flow in Wellington City

7.1 Introduction

On 14th of November 2016 (Monday) at 00:02:56 NZDT a magnitude (Mw) 7.8 earthquake hit the South Island of New Zealand. This earthquake caused significant physical movement of land, and generated landslides, disrupted roads, and rail systems, damaged buildings, and infrastructure. The centre in the affected regions are predominantly small rural towns. The main tourist town, 60 km away from the centre of the earthquake, is Kaikoura. This earthquake was about two minutes long and caused severe damage to the main infrastructure (e.g. transportation). This research uses BTMS collected data to investigate the impact of this earthquake on traffic flow.

This earthquake caused faults to move both horizontally and vertically (12 meters) [28], [253]. Some of these faults were located at sea and movement of these faults caused tsunami waves in some areas as shown in Fig. 7.2. Geonet,

135

which is an online hazard monitoring system in New Zealand, reports: "15,840 Felt Rapid Reports were submitted online in the first hour and our website took a massive 250 million hits in 24 hours (peaking at 35,000 hits per second)" [28]. Fig 7.1 shows the location and distribution of these submitted reports along with the location of the earthquake and Wellington. Woods et al. reports that thousands of people were affected by this earthquake along with isolated communities and other critical infrastructure [253]. This earthquake caused damage in Wellington City, some of which has not been repaired two years later [253], [219]. Some of this damage is shown in Fig 7.3 to 7.10. In addition to the streets in these photos, Featherston Street in centre of Wellington was also declared as a No-Go area on the earthquake day [22], [7]. Damage to the buildings in Murphy Street was also reported[8]. Fig. 7.11 shows the location of some of these streets in Wellington in.

This content is unavailable. Please consult the print version for access.

Figure 7.1: Location of the 2016 Kaikoura earthquake is shown by a circle and the distribution of the shakes reports are shown by rectangles [28].

The only available official report on the damage is related to the damage to Molesworth Street. Based on a web announcement published by the Wellington City Council, on 16th of November (two days after the earthquakes) [68], and as shown in Fig. 7.8, there was a closure on this road due to a partial building collapse. The closed area is shown by a green rectangle in Fig. 7.7.

This content is unavailable. Please consult
the print version for access.

Figure 7.2: The measured height and recorded time of the tsunami waves after
the 2016 Kaikoura earthquake. The height of the largest wave in Wellington coast
was 70 centimetres and the first wave arrived 30 minutes after the earthquake [28].

This content is unavailable. Please consult
the print version for access.

Figure 7.3: Displacement of the road and house by about 10 meters due to the
2016 Kaikoura earthquake [28].

This content is unavailable. Please consult the print version for access.

Figure 7.4: Damage to the main trunk railway line and State Highway 1S, near Kaikoura [16].

This content is unavailable. Please consult the print version for access.

Figure 7.5: One of the street in the Wellington CBD (Tory Street) which was closed due to damage[62].

This content is unavailable. Please consult the print version for access.

Figure 7.6: Demolition of a damaged building (61 Molesworth Street) in the Wellington CBD [7].

This content is unavailable. Please consult the print version for access.

Figure 7.7: Molesworth Street on the map. Location of the damaged building is shown by the green rectangle.

This content is unavailable. Please consult the print version for access.

Figure 7.8: Road closure on Molesworth street based on the official report available at [68].

This content is unavailable. Please consult the print version for access.

Figure 7.9: Damaged buildings on Wakefield street [69] (Photo/Mark Mitchell).

This content is unavailable. Please consult the print version for access.

Figure 7.10: Damaged buildings on Cable street [69] (Photo/Mark Mitchell).

This content is unavailable. Please consult the print version for access.

Figure 7.11: Some of the damaged buildings are located on these streets (indicated by red rectangles), Bing map.

The aim of this chapter is identifying the impact of the 2016 Kaikoura earthquake on Wellington traffic flow using BTMS data. This chapter examines the collected Bluetooth data to see whether it is possible to identify directly from the data these disruption events (e.g. road closures in the streets like Molesworth and Featherston Streets). Road closure between two Bluetooth counters (i.e. in a micro route) may be detected by the investigating of the changes in traffic flow between sensors (Bluetooth counters) before and after the closure event. In addition, if a road closure occurs between two Bluetooth counters, vehicles may take an alternative route. If this new route has the same origin and destination and similar traffic characteristics (e.g. travel time), then this road closure cannot be distinguished as the pattern of traffic characteristics is unchanged between these two Bluetooth counters before and after the road closure, but if these routes have different traffic characteristics (e.g. travel time), then it shall be possible to identify this road closure (see Chapter 3). Furthermore, an increase in the number of observed vehicles after a natural disaster (e.g. an earthquake) does not necessarily show a demand increase. It is possible that a greater proportion of the drivers used their navigation systems, with detectable Bluetooth cards, to obtain real-time information about the conditions of the road networks during their trips to assist them to reach to their destinations quicker.

In order to identify affected areas, this chapter examines the traffic flow in three main steps in this chapter. The first step, section 7.2, investigates any possible changes in the rate of the detection of Bluetooth devices within the proximity of each Bluetooth counter before and after the earthquake. In the second step, Section 7.3, Macro trips (or macro routes) are constructed and the traffic flow (i.e. volume) within each macro route is investigated. A macro trip is the actual origin and destination of a vehicle rather than the two consecutive detections (i.e. micro route). The third step, Section 7.4, explores the collected traffic data at the micro route level. Steps two and three investigate the pre/post earthquake changes in the traffic trends for each route by utilising the processed data obtained from applying previous algorithms (ExtoVT). Since the analysis of the traffic flow at the macro route level is more complex than the analysis of the flow at the micro routes due

to the bigger size of the Geospatial dataset, this chapter briefly analyses the flow at the macro route level, and focuses on the micro routes.

7.2 Result and analysis at the Bluetooth counter level

The initial step is an investigation of any sudden drop to a level close to zero detections in the number of the detected devices (as collected, i.e. no cleaning) at each Bluetooth counter during the three weeks of study (i.e. two weeks before the earthquake and one week after the earthquake). This is done to identify any failures of Bluetooth counters due to various reasons (e.g. power outage), or any road closure within the proximity of the detection range of Bluetooth counters. Fig. 7.13 shows the location of these devices and the detection ratio of these devices pre/post earthquake are shown by a pie chart at the location of each Bluetooth counter, although this pie chart may not be the best type of chart for this purpose, but this is the only built-in tool for visualising the results on a map in my spatial visualisation tool (i.e. Power BI).

The blue area is the daily average detection number of the Bluetooth devices (i.e. Equation 7.1) for the two weeks before the earthquake and the grey area in the pie charts is the daily average detection of the Bluetooth devices after the earthquake (one week). As demonstrated, most Bluetooth counters had fewer detections after the earthquake compared to the time before the earthquake (because there were fewer Bluetooth devices on their detection range after the earthquake) except a few locations (mostly on the East side (right) of Wellington). A bigger bubbles size demonstrates more Bluetooth detections.

The number of detection per day for the three weeks of study is shown in Fig. 7.12. As shown in this figure, no sudden drop to a level close to zero detections is observable for these Bluetooth counters, although there is a drop in the number of detected devices after the earthquake (14th of November). In addition, the number of detected devices is lower after the earthquake due to less activity around those Bluetooth sensors (or counters) at this time.

$$\text{Average daily number of detections} = \frac{\text{Total number of detections}}{\text{Number of days}} \tag{7.1}$$
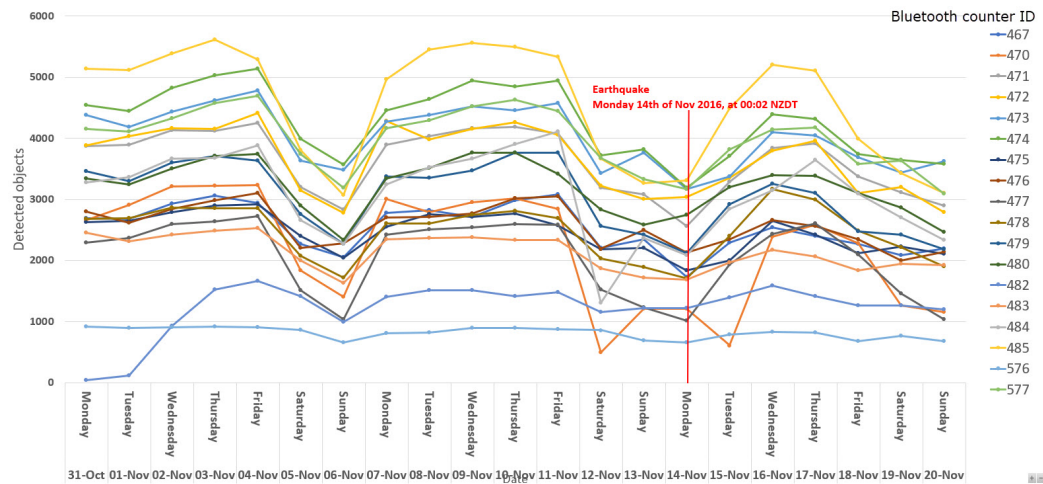
Figure 7.12: Total number of the detected Bluetooth devices (objects) by traffic Bluetooth sensors per day.

As demonstrated in Figures 7.12 and 7.13, there is no sudden drop to a level close to zero in the number of the detection of Bluetooth devices at the installed Bluetooth counters (sensors). This suggests there was no road closure in the proximity of these Bluetooth counters. In addition, since these Bluetooth counters recorded devices for the three weeks of the study period, no malfunction or power failure affected these counters. Even if a connection failure had occurred, BTMS store the collected data in their internal storage, and later when they are reconnected, then they submit this data to servers. The BTMS maintenance team also confirmed this matter and emphasised that none of the BTMS experienced any power failure during the study period. As the earthquake occurred on Monday, 14th of Nov, Monday is considered as the beginning of the week and Sunday is the last day of the week in this chapter, so that the third week of the study is entirely after the earthquake.
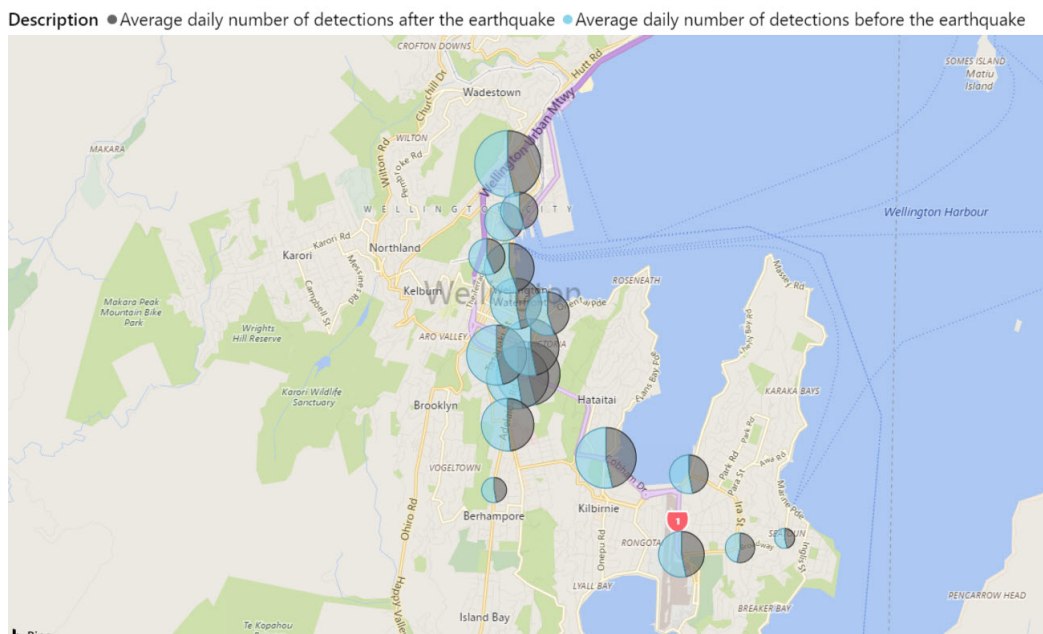
Figure 7.13: The average daily number of detected Bluetooth devices before (blue) and after (grey) the 2016 Kaikoura earthquake, Bing Maps. Most locations experienced a decrease, while a few (east of Wellington) experienced an increase. Size is a reflection of the quantity of the detection.

In order to investigate whether the collected travel times (i.e. the time difference between two consecutive detections at a predecessor and successor Bluetooth counters) from Bluetooth devices, without any processing for mobile nodes (i.e. Travel time >0), are significantly different within these three weeks (i.e. Alternative Hypothesis H1) or not (i.e. Null Hypothesis), this section utilises the Analysis of Variance (ANOVA). ANOVA is a statistical test which demonstrates whether the variation between the mean of travel times in each week is a true difference or because of detection variability [203]. ANOVA answers this question by calculating the F statistic, which is the ratio of travel time variation among different weeks over the travel time variation within each week. In this ANOVA test, the travel times of the two weeks before the earthquake are considered as one set and the travel times of the week with the earthquake are considered as another set. The results of this ANOVA test are shown in Table. 7.1. A P value $Pr(>F)$ less than 0.05 suggest significant deference [71]. These results demonstrate that these two periods are significantly different from each other as $Pr(>F)$ is lower than 0.05.

Table 7.1: ANOVA results for the weeks before and after the earthquake. Results show the week with the earthquake is significantly different from the study time before the earthquake.

Comparison between	F Value	$Pr(>F)$
Before and after the earthquake	43.17	5.02e-11

Since results show that there is a significant difference between these two periods (i.e. before and after the earthquake), the next sections focus on identifying the affected routes at a macro and micro level analysis.

7.3   Result and analysis at the macro route level

As BTMS can track individual vehicles, it is therefore possible to study the origin and destination of vehicles at a macro level. This means each individual vehicle which follows the regular traffic pattern can be tracked over the city during a day and their detection locations (i.e. latitude and longitude) can be connected in

ascending order of the detection times. This forms a spatial path (i.e. combination of latitudes and longitudes) that represents the movement of each individual driver during a day. Since it is possible to track each vehicle over various days, the behavioural analysis of drivers can also be done which is part of the future research.

An example of the movement of a vehicle is shown in Tables 7.2 and 7.3. Table 7.2 shows a sample of recorded observation for the vehicles which follows the regular traffic flow between two consecutive Bluetooth counters. Table 7.3 shows the formation of each macro trip (or macro route), which is obtained by connecting the micro trips (or micro route) based on the detection times in ascending order. Table 7.3 also shows the frequency of each macro trip per day which is obtained from Table 7.2. To clarify this concept, Fig. 7.14 spatially visualises a sample route taken by vehicle $m1$ by red lines. A red line shows a path between two consecutive Bluetooth counters, in which a vehicle follows the typical traffic trend (i.e. micro path, (473$-$>475 and 577$-$>578). In this figure, vehicle $m1$ was following the regular traffic patterns between locations 473 and 475, and also between locations 577 and 578, but not in between locations 475 and 578 (either because there is no information about the route, or it did not follow the regular traffic patterns). Therefore, the complete spatial path (i.e. macro route) for vehicle $m1$ is demonstrated as a green line in this figure.

Since this was a big Geo-spatial dataset, I only consider macro routes which have the demand of greater than 14 trips (at least 1 trip per day).

Table 7.2: A sample example of micro trips. The details of each micro trip for each individual vehicle which follows the regular traffic flow between two consecutive locations of Bluetooth counters.

Vehicle identifier	Departure time	Arrival time	Identifying label for source and destination (micro trip)
m1	2016-11-18 09:33:50	2016-11-18 09:43:50	473 → 475
m1	2016-11-18 12:42:32	2016-11-18 12:45:32	577 → 578
m2	2016-11-18 08:33:50	2016-11-18 08:43:50	473 → 475
m2	2016-11-18 11:42:32	2016-11-18 11:45:32	577 → 578

Table 7.3: Formation of macro trips (routes) per day from micro trips (routes) and their frequencies obtained from Table 7.2.

Date	Macro route (macro origin and destination)	Frequency of macro trips (route)
2016-11-18	473 → 475 → 577 → 578;	2
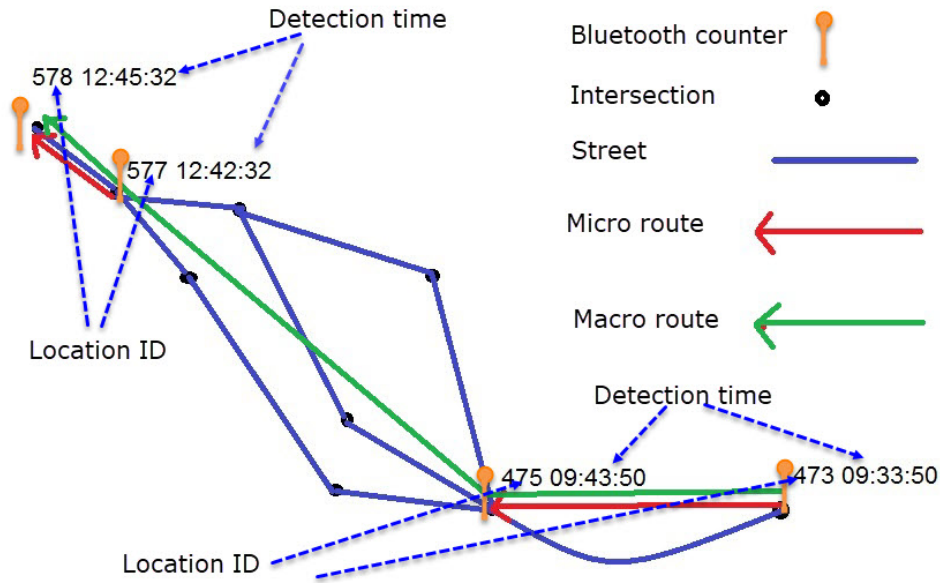
Figure 7.14: Visualised demonstration of the formation of a macro path (displayed by a green arrow) from micro paths (displayed by red arrows) for vehicle $m1$, based on the recorded observation in Table 3.5.

To compare the average daily frequency of detected vehicles before and after the earthquake for each macro route (R), the average daily traffic flow for each macro route (R) is obtained from the following equation:

$$\overline{F}_R^e = \frac{\sum_{k=1}^{k^e} \sum_{d=1}^{7} \sum_{h=0}^{23} f_{kdhR}}{n^e},$$

(7.2)

where e is a period before or after the earthquake ($e \in \{a =$ after earthquake, $b =$ before the earthquake$\}$), k^e is the number of the week e in the period of study, d is the day of the week k^e of period e, and h is the hour of day d of week k of period e. f_{kdhR} is the total number of vehicles counted in week k, day d, and hour

h for macro route R ($R \in \{1, \ldots, n_R\}$). The value of $n^b = 7k^b$ is 14 days before the earthquake as there are two weeks before the earthquake in my dataset and the value of $n^a = 7k^a$ is 7 days after the earthquake as there is only one week of data after the earthquake in my dataset.

The ratio of $\overline{F}_R^a$ to $\overline{F}_R^b$ is computed to obtain the ratio of mean flows in the week after the earthquake over the baseline, and the result is shown in Fig. 7.16. A ratio below 1 demonstrates a decrease in a macro route and is shown by a warm color, and a value above 1 demonstrates an increase in the traffic flow after the earthquake and is shown by a cool color. A ratio of 1 denotes no changes in the traffic flow. The ratio can be a more informative measure compared to the traffic volume and difference, as the ratio signifies the proportion of changes in the traffic flow regardless of the capacity of the road or the total number of lanes. As demonstrated in this figure, there is an increase in traffic flow on the West side (left) of the Basin Reserve (denoted by number 1A, 1B, 1C). Location 1 is the entrance to the Wellington Urban Motorway (also called motorway in my thesis), which is also an exit from the city, with no further Bluetooth counter. Thus, the cars at Location 1 are not detected again. As there was damage to some of buildings and infrastructure in the CBD (e.g. Tory Street and Molesworth Street), therefore it may be concluded that drivers used the Wellington Urban Motorway as a preferred route compared to the routes which pass through CBD to move from South and East of the Basin Reserve to the North parts of Wellington to avoid the traffic disruption events (e.g. road closures).

Furthermore, there is an increase in traffic flow after the earthquake between Location 3 and Location 4 (see the light green line between Location 3 and Location 4). This increase between these two locations suggests that some of the drivers took alternative routes to commute between North parts of Wellington and location 3. They possibly used the coastal Oriental Bay route, or other routes which do not pass through the Basin Reserve and closed roads in the CBD (e.g. Tory Street and Wakefield Street). Two possible routes from Location 3 to 4 which are suggested by Google, are shown in Fig. 7.17. One of them passes through the coastal Oriental Bay route (shown by a blue line). The other one which is shown

by grey passes the Basin Reserve and also the roads which were closed after the earthquake for some time (e.g. Tory Street and Wakefield Street, also see Fig. 7.11). Therefore, drivers possibly used the blue route and avoid using the grey route after the earthquake to avoid traffic disruption events (e.g. road closures). Wellington Airport (domestic and international) is located in the East of the city (with high demand) and generally there are four main routes which connect East of Wellington to the rest and all pass through a natural barrier (a ridge of hills). These four routes are shown on the map in Fig. 7.18. Two of these three routes are also shown on the map in Fig. 7.17 and are Oriental Bay and Mount Victoria Tunnel. The third one passes through Constable Street and the fourth one, which is the longest, passes through Houghton Bay as shown at the bottom of Fig. 7.18.

Numerical values of these four paths, 1A, 1B, 1C, 3 to 4) $\overline{F}_R^a$ for each macro route from the average daily flow of the baseline $\overline{F}_R^b$ for each macro route to find the difference of the flow before and after the earthquake for each macro route. The result is shown on the map in Fig. 7.15.

A negative value (shown by warm colours (e.g. red)) of shows a decrease in the daily average flow after the earthquake for each macro route, while a positive value (shown by cool colours (e.g. green)) of demonstrates an increase in the daily average traffic flow after the earthquake for each macro route.

7.4   Result and analysis at the micro route level

Fig. 7.19 shows the total number of detected vehicles by BTMS for the three weeks of study in all of the micro routes. As demonstrated in this figure, the total number of detected vehicles is lower in the week with the earthquake compared to the other two weeks. As demonstrated here, the weekends have a different pattern compared to weekdays due to work constraints. The peak hours during weekends is about the lunchtime, which in the weekdays there is a morning and afternoon peak.

To identify the changes in the traffic flow after the earthquake by examining
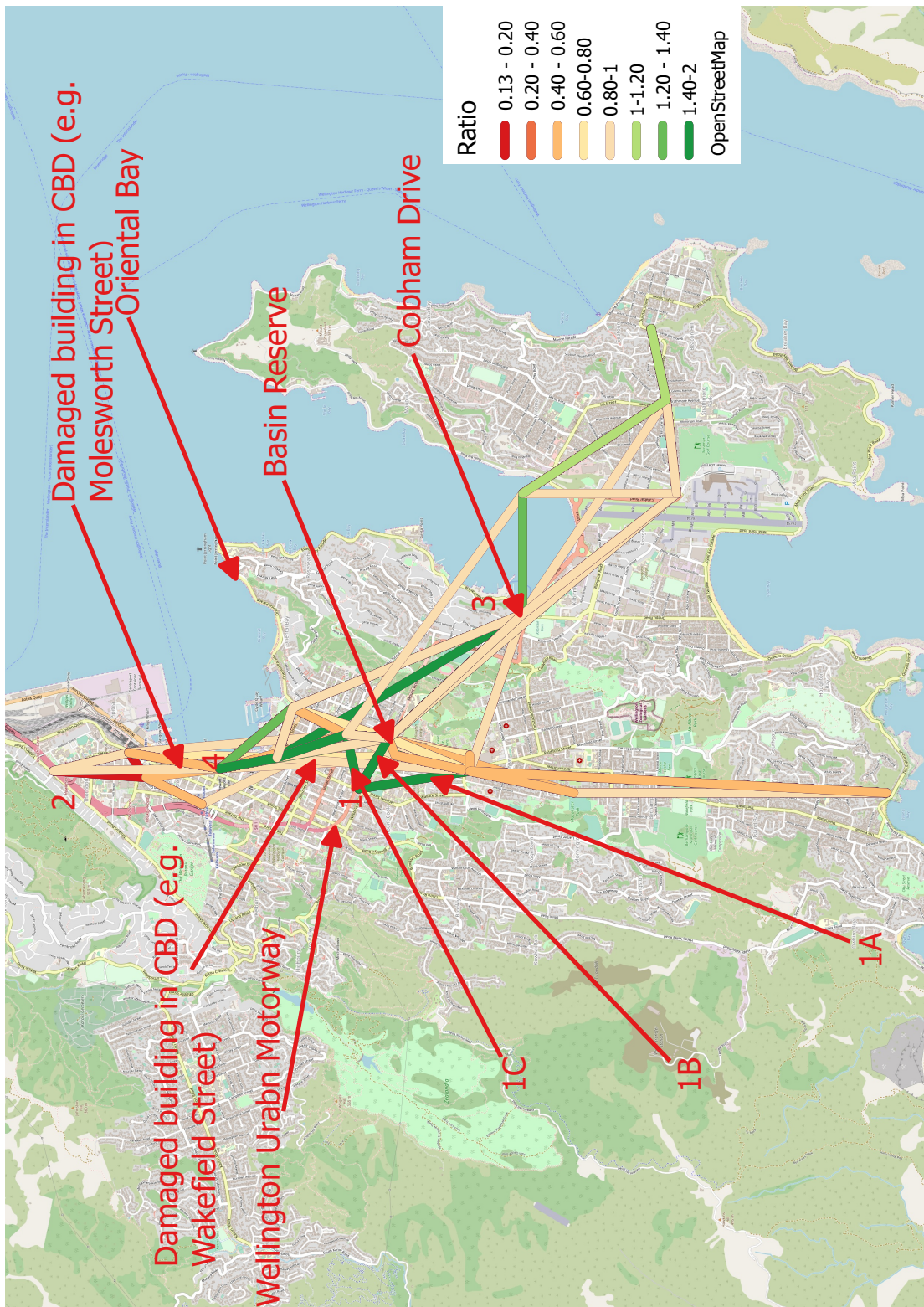
Figure 7.15: The ratio of average daily traffic flow (frequency) for each macro route after the earthquake over the baseline, Open Street Maps (OSM). Warm colour (e.g. red) shows a decrease and cool colors (e.g. green) shows an increase in the average daily traffic flow after the earthquake compared to the baseline (i.e. before the earthquake).
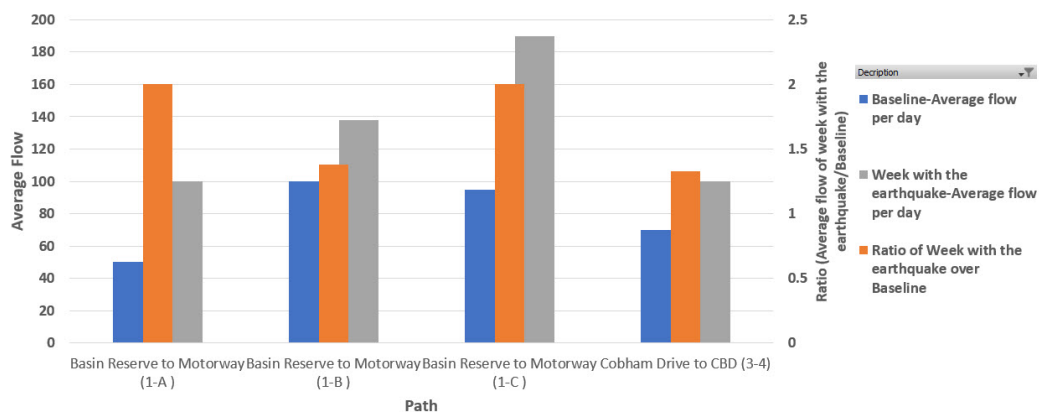
Figure 7.16: Difference of average daily traffic flow (frequency) for route 1*A*, 1*B*, and 1*C* before and after the earthquake in accordance with Fig. 7.15. Blue shows the time before the earthquake (baseline), and grey shows the week after the earthquake. Orange shows the ratio of the week with the earthquake over baseline as the second Y axis on the right. X axis shows the name of these locations and Y axis on the left shows the value of detected cars for the baseline and the week after the earthquake. This figure shows an increase in the average daily traffic flow for these locations after the earthquake compared to baseline (i.e. before the earthquake).
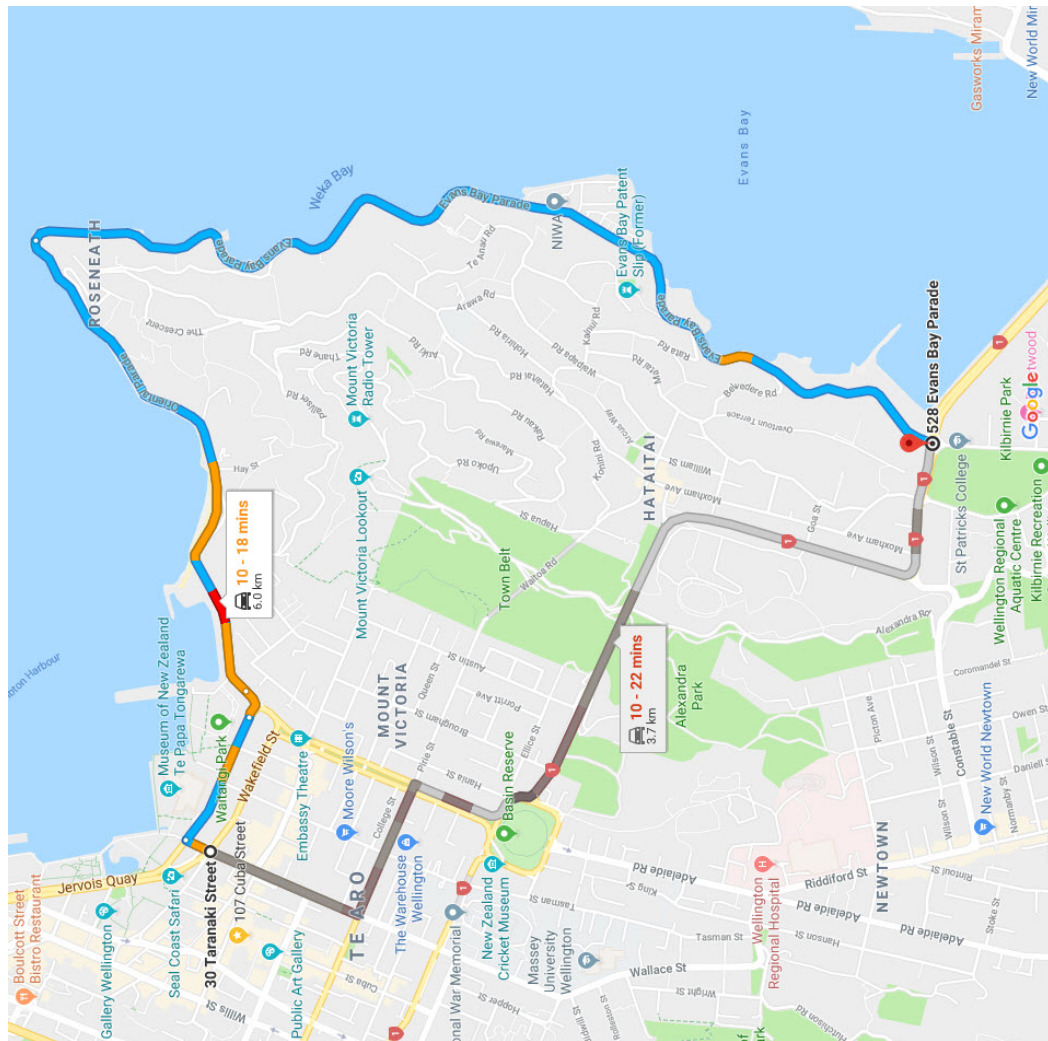
Figure 7.17: Best routes suggested by Google for moving between Locations 3 and 4 (shown in Fig. 7.15). As suggested by Google, Oriental Bay is a better option than the grey route.
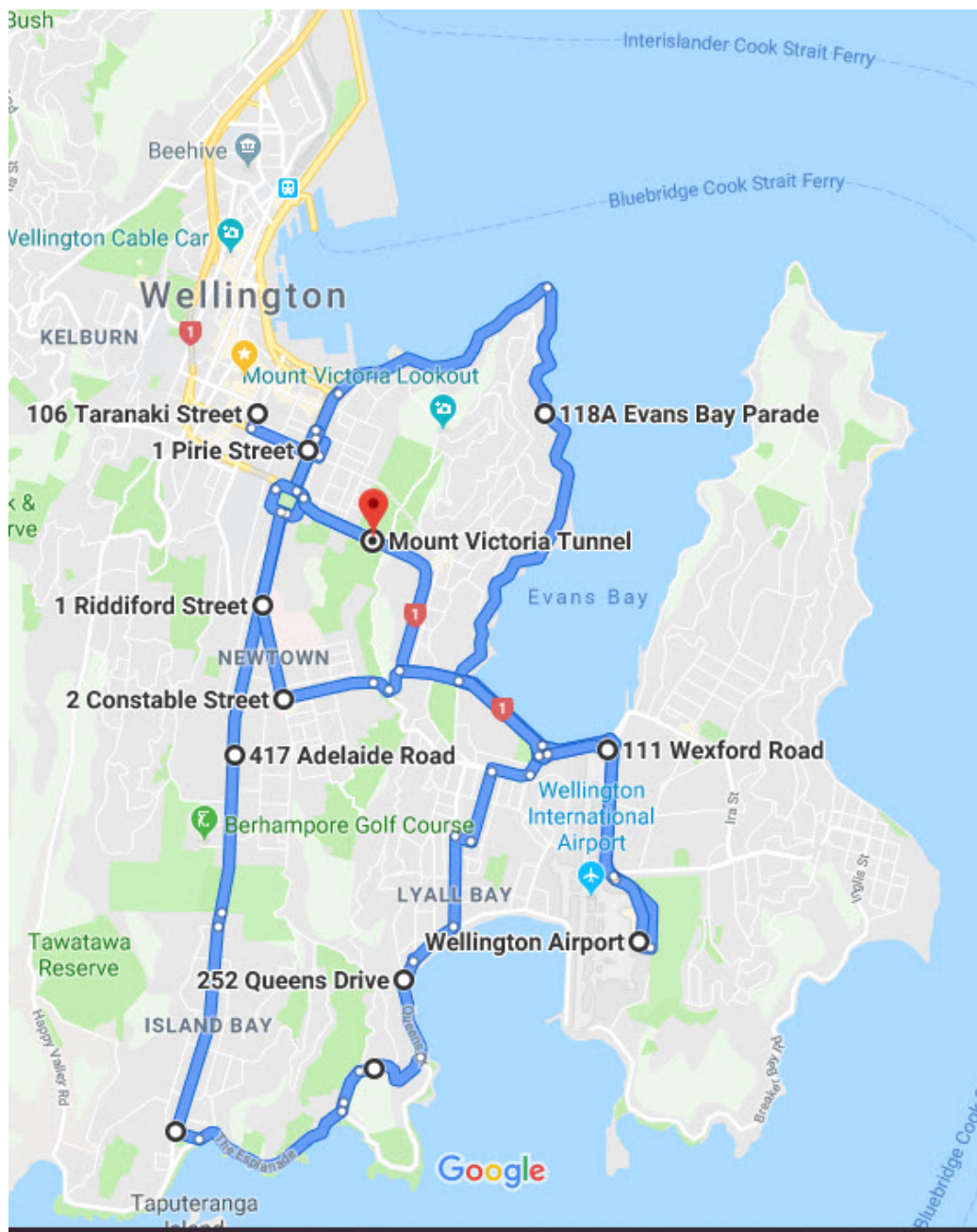
Figure 7.18: Four main routes which connect East to the rest of Wellington are Oriental Bay, Mount Victoria Tunnel, Constable Street, and the route which passes south through Island Bay and Adelaide Road. The fourth one is the longest path. Among these four paths, the route which passes the coastal Oriental Bay route (top blue line) does not go through Basin Reserve.
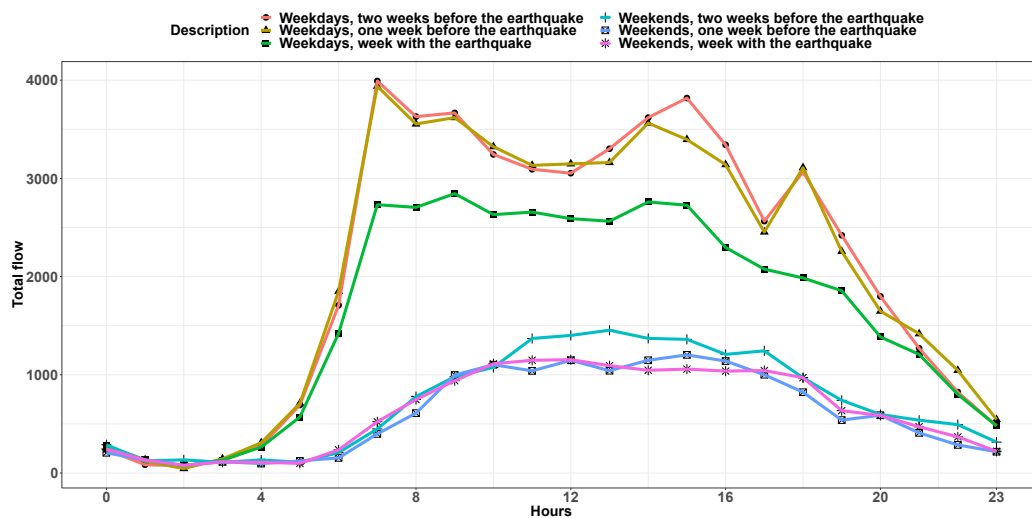
Figure 7.19: The total number of detected vehicles by BTMS per hour for each day of the study period. Fewer vehicles are detected in the week with the earthquake compared to the two other weeks before the study. Weekends and weekdays have different patterns due to working constraints. Weekdays have a sharp peak around lunchtime, but the peak points are in the morning and afternoon in the work days (weekdays) as these are the times that people commute to work. Also, the number of detected vehicles during day time is much higher than the number of vehicles during night time.

the trend at micro routes, this chapter also forms a baseline with the two weeks of data before the earthquake for each micro route.

Fig. 7.20 shows the average daily traffic flow before the earthquake (called the baseline) per weekday $\overline{F}_d^b$, next to the average daily flow per weekday for the week after the earthquake $\overline{F}_d^a$. Calculation of $\overline{F}_d^e$ is shown in the following equation:

$$\overline{F}_d^e = \frac{\sum_{k=1}^{k^e} \sum_{h=0}^{23} \sum_{r=1}^{r_n} f_{kdhr}}{k^e},\qquad(7.3)$$

where e is period before or after the earthquake ($e \in \{a = $after earthquake, $b = $before earthquake$\}$), k is the number of the week in the period of study (e), d is the day of week k of period e, and h is the hour of day d of week k of period e. f_{kdhr} shows the total number of counted vehicles (flow) in week k, day d, and hour h for micro route r, $r \in \{1, \ldots, n_r\}$. The value of k^b is two before the earthquake as there are two Mondays,..., Sundays in the two weeks of data before the earthquake. The value of k^a is one for the week after the earthquake as there is only one Monday,..., Sunday in the week after the earthquake. Since the period of the study time for the traffic flow before the earthquake contains two weeks and the study period after the earthquake contains one week, therefore the average of the daily flow per day for the two weeks before the earthquake must be compared with the total flow per day (or average daily flow as $k^a = 1$) for the one week after the earthquake. Therefore, this section uses the average of daily traffic flow before the earthquake with the total flow per day after the earthquake. The total flow per day for one week after the earthquake can be considered as the average flow per day as there is only one Monday,..., Sunday in the week after the earthquake.

As shown in this figure, Fig. 7.20, the week with the earthquake has a lower traffic flow compared to the baseline. This is possibly because fewer people commuted to work, because some employees were told to not work in their workplaces [21], and some public assets (e.g. schools) were closed and waited to be inspected by engineers and inspectors [21]. The weekend after the earthquake also has a lower traffic flow compared to the weekend of the baseline (i.e. before the earthquake) and this is because some people restricted their non-essential trips over the weekend after the earthquake.
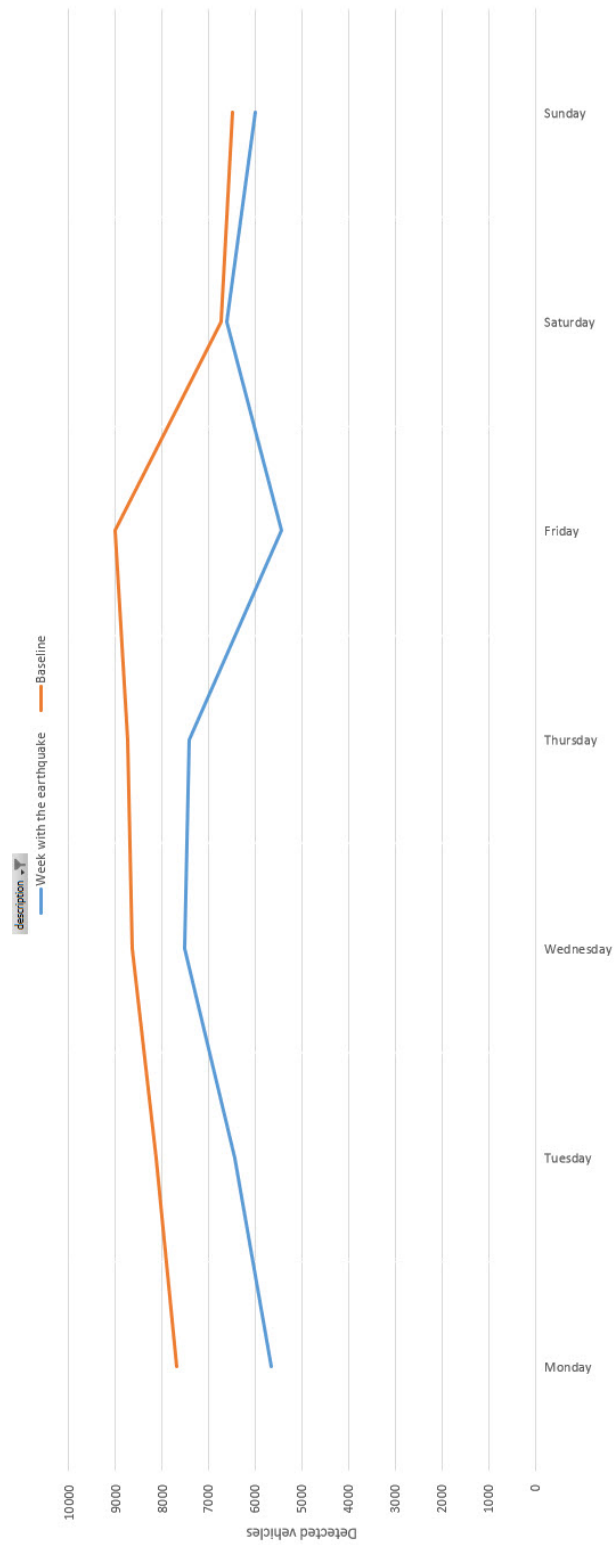
Figure 7.20: Average daily traffic flow per micro routes per day of the week pre/-post earthquake, which demonstrates a lower traffic flow after the earthquake.

Fig. 7.21 demonstrates the average weekly traffic flow for each micro route for the two weeks before the earthquake (i.e. baseline) and the average weekly flow after the earthquake for each micro route. The average weekly traffic flow for each micro route is calculated by the following equation:

$$\overline{F}_r^e = \frac{\sum_{k=1}^{k^e} \sum_{d=1}^{7} \sum_{h=0}^{23} f_{kdhr}}{k^e}, \tag{7.4}$$

where e is period before or after the earthquake ($e \in \{a = $after earthquake, $b = $before earthquake$\}$), k is the number of the week in the period of study (e), d is the day of week k of period e, and h is the hour of day d of week k of period e. f_{kdhr} shows the total flow (i.e. frequency) of counted vehicles in week k, day d, and hour h for micro route r ($r \in \{1, \ldots, n\}$). The value of k^b is 2 weeks before the earthquake as there are two weeks before the earthquake in my dataset and k^a is 1 after the earthquake as there is only one week of data after the earthquake in my dataset.

As illustrated in this figure, Fig 7.21, a few routes experienced an increase (e.g. West of Basin Reserve) in demand, but most experienced a decrease. The majority of these decreases happened in the streets with damaged buildings (e.g. Wakefield Street, Cable Street, Molesworth Street). Some of this damage is shown in Fig. 7.3 to 7.10. In addition, based on Fig. 2.6 and 2.7, most of the tall buildings at the greatest seismic risk are located in the CBD. However, as part of future work, this research will investigate the differentiation of non-earthquake tourism from earthquake tourism. Earthquake tourism refers to the trips that people make to see the impact of the earthquake on their city and non-earthquake tourism is referring to typical trips other than earthquake tourism.

Furthermore, this chapter also examines the difference between the week with the earthquake and the baseline along with the ratio of the traffic volume of the week with the earthquake over the traffic volume of the baseline, based on Equation 7.4 and Fig. 7.21. These two quantities (difference and relative traffic flow between the baseline and the week after the earthquake) are utilised based on Fig. 7.21 to illustrate the changes in the traffic flow in a more informative approach. A negative difference in the traffic flow shows a decrease in the detected vehicles after the earthquake. A value of the ratio lower than one of traffic flow after/be-
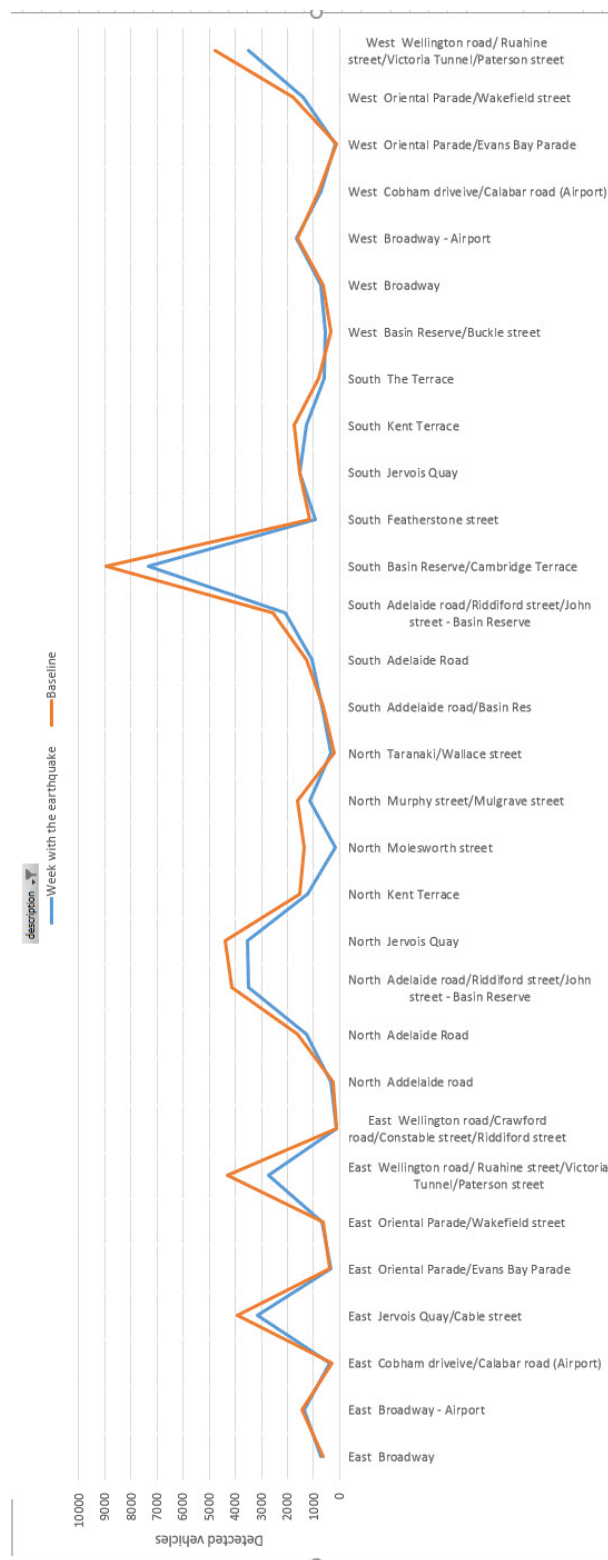
Figure 7.21: Average weekly traffic flow per micro routes pre/post earthquake. Most micro routes experienced demand reduction after the earthquake.

fore earthquake indicates a lower quantity of the identified vehicles in the week with the earthquake compared to the baseline. Fig. 7.22 and 7.23 show the sorted ratio and flow difference for each route based on the highest value, respectively. This plot shows the numerical values. In order to make the concept clearer Fig. 7.25 plots these values on a map. The variation of the colour in this heatmap demonstrates various ratio values. Warmer colour indicates a reduction of traffic flow after the earthquake and a cooler colour demonstrates an increased traffic flow after the earthquake. As shown in Figures 7.22 and 7.25, Molesworth Street has the lowest ratio, which highlights that this route experienced a severe reduction in traffic flow due to the earthquake (Fig 7.6, 7.7, and 7.8). Furthermore, the heatmap shows most of the affected areas and road closures are within the CBD where the highest buildings are (see Figures 2.6 and 2.7). In addition, as demonstrated in Fig. 7.25, the traffic to the west (left) of the Basin Reserve has increased, which suggests that drivers avoid passing through the CBD to reach the North part of Wellington and they possibly used Wellington Urban Motorway. In addition, there is a decrease on the East (right) side of the Basin Reserve while there was an increase in the alternative routes such as Oriental Parade (Oriental Bay) which suggests that drivers used the Oriental Bay to move between Cobham drive (i.e. Location 3 in Fig. 7.15) and CBD rather than using the Mount Victoria Tunnel. These agree with the findings of the analysis of routes at the macro level (i.e. drivers avoiding passing through CBD and using alternative routes). Based on the Fig 7.23, there was some traffic disruption events around the Basin Reserve and Mount Victoria Tunnel after the earthquake, which these streets are also colored by warm colors in Chapter 2, Fig. 2.7.

Fig. 7.24 visualises the difference and ratio together for each route for the week with the earthquake and the baseline to make the comparison easier.

7.5   Conclusion

As demonstrated in this chapter, my proposed traffic monitoring approach, which utilises the collected data from Bluetooth devices, was able to identify the changes
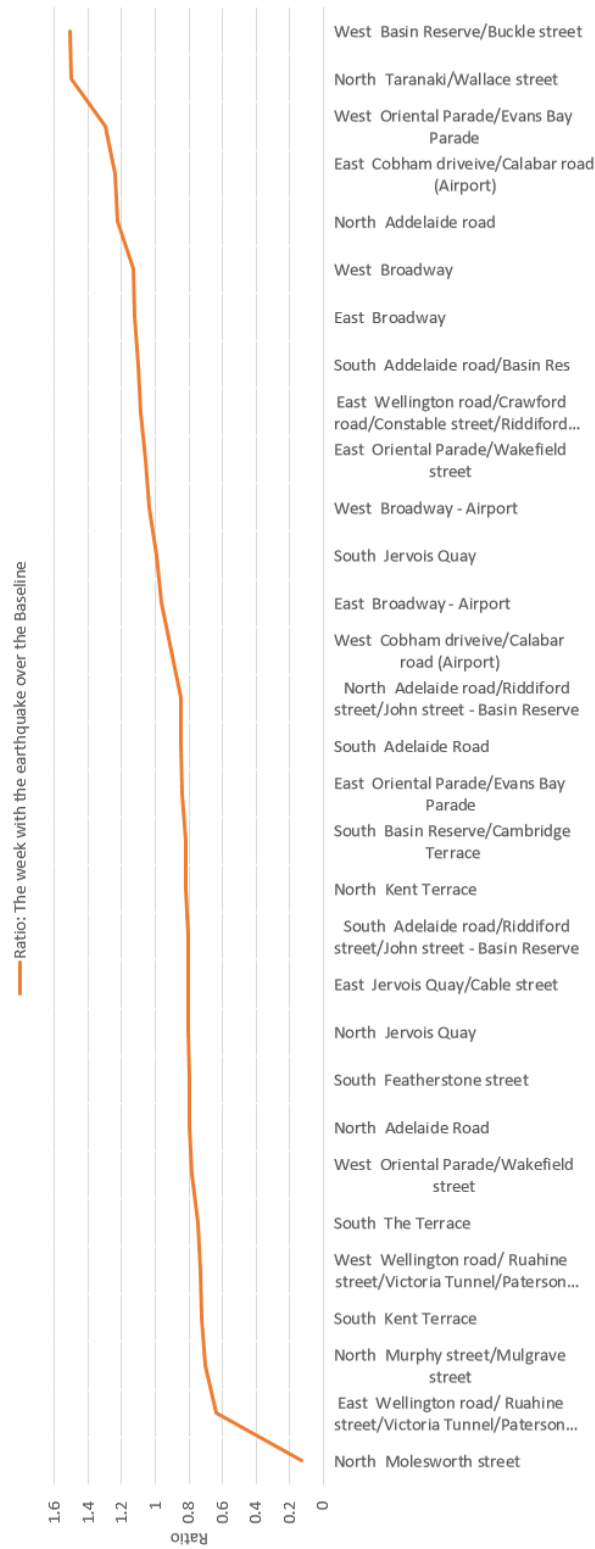
Figure 7.22: Sorted relative difference (ratio) per micro route based on Fig. 7.21.
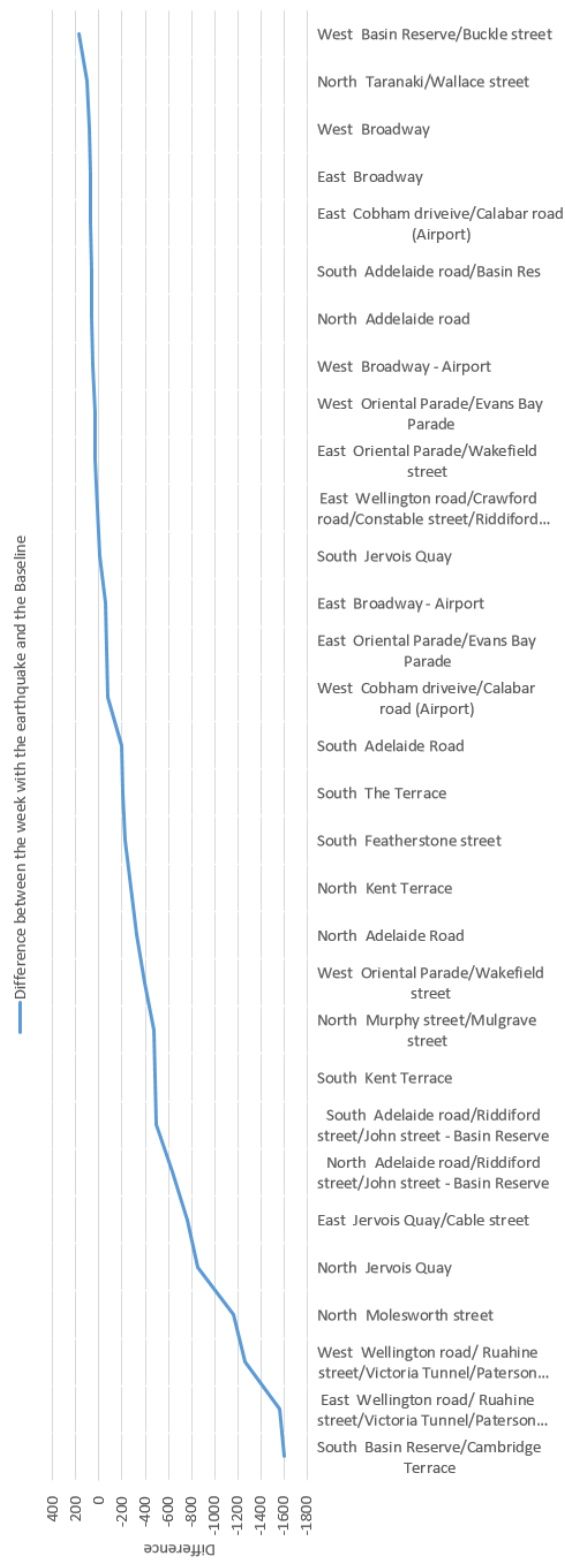
Figure 7.23: Sorted difference per micro route based on Fig. 7.21.
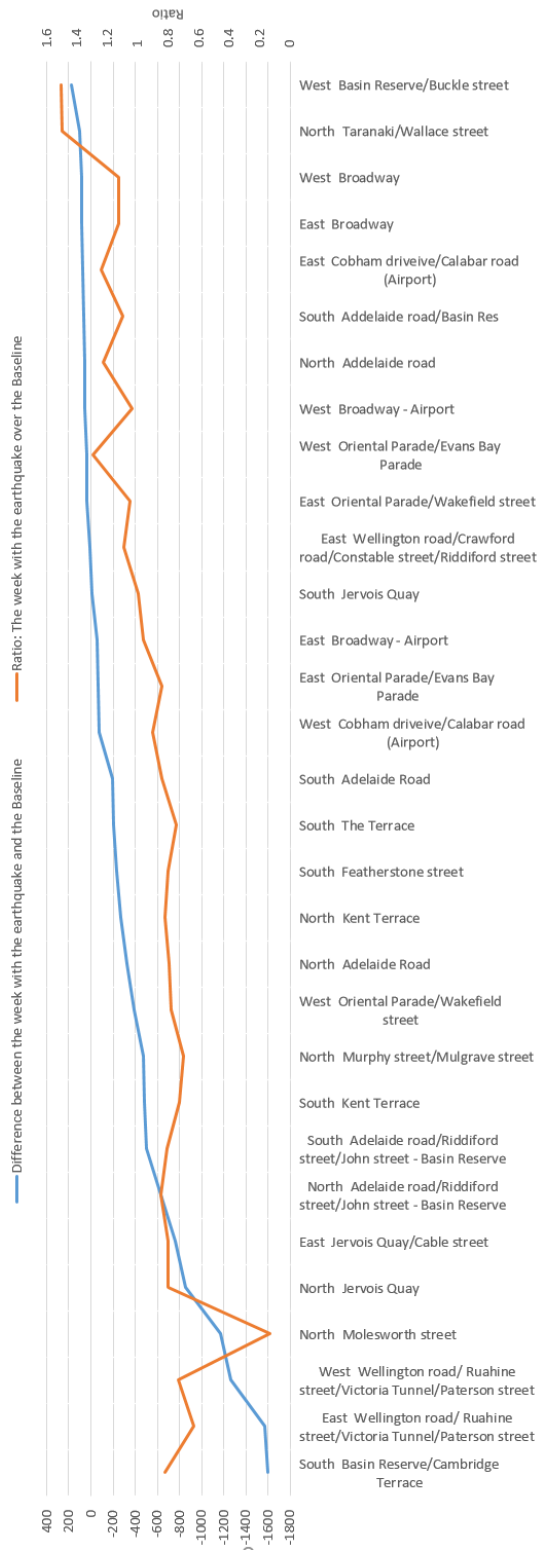
Figure 7.24: Difference and relative difference (ratio) together per micro route based on Fig. 7.21.
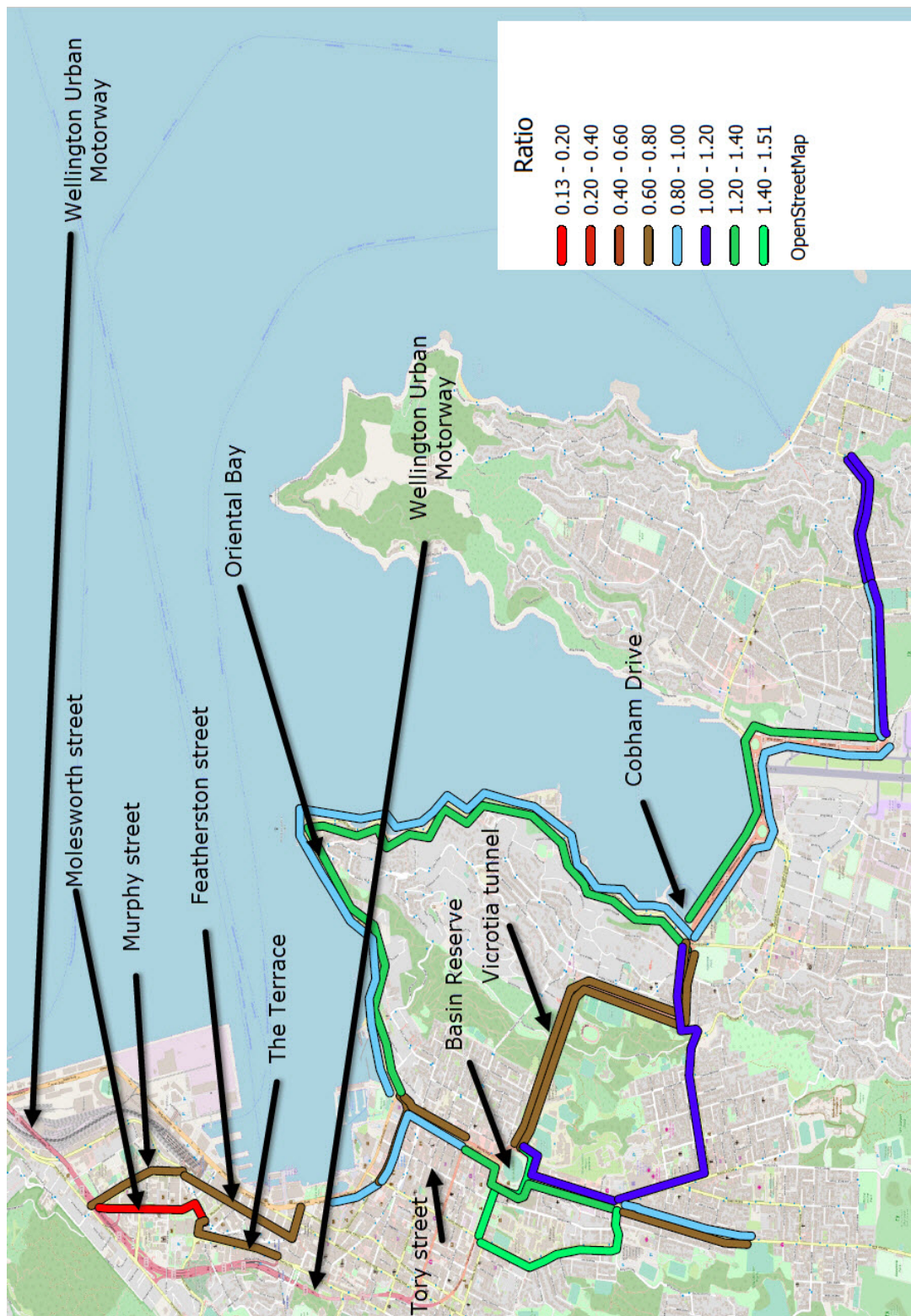
Figure 7.25: Heatmap of the routes based on the relative difference (ratio as shown in Fig. 7.22), Open Street Maps (OSM). Warm colors (e.g. red) shows decrease in the traffic flow after the earthquake and cool colors (e.g. green) shows an increase after the earthquake.

in the traffic patterns, demand, and affected routes due to the Kaikoura earthquake. I accomplished this through investigation of three different methodologies: the examination of individual Bluetooth counters, macro routes, and micro routes. I also visualised the results spatially through various heatmaps to detect the impacted areas more accurately. I identified the affected routes through investigation of the various traffic characteristics such as ratio, volume, and difference of traffic volume before and after the earthquake. Based on the results of this chapter most of the affected routes are located in the CBD, possibly due to damage to the buildings on those streets. Therefore, drivers tried to avoid passing through CBD and selected alternative routes.This conclusion complements with what I concluded in Chapter 2 as the CBD is a hazardous zone.

Although it is possible to automate these detections in the traffic flow due to an incident or a traffic disruption even through connecting the feed of the collected BTMS data to Intelligence dashboards, but it should be noted that these kind of automation systems cannot be fully real time as BTMS needs to work in pair (as explained in Chapter 6) and also there is a need to have enough sample for analysis (possibly 10 minutes intervals, although this depends on the size of the collected data). Therefore thees kind of systems can be considered as almost real-time.

Chapter 8

Simulation environment

8.1 Introduction

Building an accurate picture of a disaster evacuation scenario is a complicated task especially as models, tools, methodologies from several overlapping disciplines such as computer science, transportation engineering, medical science, disaster management, mathematics, statistics, and Geographic Information System (GIS) are needed. Examples include simulation environments, routing algorithms, Car-following model, behavioural model, driving culture, traffic light signalling, and psychological patterns, which are examined in this chapter. Traffic simulation can also be categorized into three sub-groups: microscopic, mesoscopic, and macroscopic. Macroscopic simulation investigates the total traffic flow such as the density of vehicles, while the microscopic simulation tracks each vehicle individually, mesoscopic sits between these two and investigates the traffic fundamentals in smaller groups of vehicles.

In this chapter, various details of traffic simulation are explained in the next section. Section 8.3 briefly presents the behavioural model. Section 8.4 introduces the principals of spatial visualisation, and the last Section, 8.5, reviews the high-level design of the Wellington disaster management plan.

8.2   Traffic simulation

This section examines various components which are required for a large-scale evacuation framework.

8.2.1   Simulation engine

Although SUMO was explained in Chapter 6, this section provides more detailed descriptions of this simulation tool as it is composed of multiple components.

SUMO (Simulation of Urban MObility) is an open-source, library-based, space-continuous and time-discrete microscopic road traffic simulator [162].

SUMO models the movement of each mobile node (e.g. pedestrian, vehicle) in a given path (i.e. collection of edges) with certain movement characteristics (e.g. speed, acceleration) individually. Each road in a real-world can be composed of several edges in SUMO, and the path of a mobile node is a combination of various edges from its source to its destination. In every single simulation step, required values (e.g. location, speed) are updated based on the given settings and/or the interaction between various involved components in the simulation environment (e.g. traffic light, the gap between two vehicles, speed signs).

SUMO is an open-source software tool with great technical support, is used in various projects, and can model each vehicle individually (i.e. is a microscopic simulation).

SUMO is used in many industrial projects [39] for transportation modelling such as Intelligent Driving Model [210], FIFA-World Cup 2006 [45], Berlin [55], MAVEN[30], 3D robots [73], and VANET [51]. Furthermore, it is integrated with network simulators (e.g. NS2, NS3, and OMNET++) for mobility modelling [165], [38]. Along with industrial projects, SUMO is also used in many academic projects such as a car to car communication [227], [251], Intelligent Transportation Systems [254], Software description [163], traffic light scheduling [254], and evacuation modelling [117]. I select this simulation tool for microscopic modelling of traffic, although it can be replaced with any other microscopic simulation which covers the discussed principals in this chapter and also the next one.

Communication with SUMO can be done in two ways, static and dynamic. Static communication needs flat files and dynamic interaction is done through an Application Protocol Interface (API).

8.2.2 Routing algorithms

A routing algorithm in transportation modelling considers various criteria for selecting the most optimal route from a source to a destination(s), such as, travel distance/time, ease of driving, petrol consumption, delays and capacity of the links, travel cost, processing time, or any combination of these factors.

- Obtaining travel distance, link delays, and travel time are common criteria for finding the optimal route [130], [260]. Link delays refer to any kind of delay on a link such as congestion, and link capacity refers to the capacity of the link (i.e. edge) and investigated by [249], and [255]. Haghnia in [130] proposed a dynamic vehicle routing algorithm by considering travel time. The travel time from a source to a destination in cities is highly dynamic and depends on many factors such as congestions of streets. Yu and Dong in [260] studied scheduling of freight vehicles to maximise the business profit by utilising the routing algorithm to find the best routes from a source to several destinations. Nha et. al. in [189] focused on trip planning as part of routing research. In this area, they proposed an approach to update a driver with the most current optimal route which is triggered by approaching at an intersection. In their research, they applied the Dijkstra algorithm on an artificial graph and evaluated this using SUMO. The Dijkstra algorithm is a popular routing algorithm which finds the least cost route from a source to a destination and used by many researchers [90], [122]. A^* is another popular routing algorithm which is used in many papers [223], [133] [104], has a high performance and accuracy, and finds the best route with the smallest cost (e.g. travel time).

- Ease of driving was used by Kanohs and Chakraborty [150], [87]. For example, the brightness of roads can impact the ease of driving at night [207].

- The travel cost considers the cost of a trip such as the number of tolls and tolls prices and is supported by some navigation systems such as Navman [58], although no cost is likely to be associated for using tolls during an emergency event.

- Petrol consumption and cost are also supported by navigation systems and consider petrol consumption of the vehicle carrying the device, emitted carbon, and fuel price for the route selection. As an example, the ecoRoute functionality by Garmin offers this service [41].

- Processing time is also another factor which was used for routing and evacuation algorithms and scrutinized by different researchers such as [249].

8.2.3   Car-following model

A car-following model describes how a vehicle follows another vehicle in transportation modelling. Generally speaking, a response for vehicle *n* to a stimulus is defined as follows [169]:

$$[Resonse]_n \alpha [stimula]_n, \tag{8.1}$$

which is also known as the motion equation of the *nth* vehicle. The response can be braking, acceleration, or deceleration. Generally, a vehicle, f, follows another vehicle, leader or l, with a certain distance (bumper to bumper) of $(x_l^t - x_f^t)$ at time t, where x is the location.

In this group of modelling, there are several subtypes such as:

- **Pipes model:**
  In this model the safe distance D is at least one vehicle length for every $16.1 km/h$, so the distance is :

$$D_f(v) = L_f(1 + v/16.1) \tag{8.2}$$

This equation is also known as the Pipe's rule [208] in which the distance increases linearly with speed [138].

- **Forbes model:**
  This model is similar to the Pipe's rules, but this model also considers the time of braking, and the required time by a driver to perceive a need for a deceleration, break, which increases the safe distance [118], [42].

- **Krau model:**
  This model developed by [164] and calculates the safe speed as:

$$v_{safe} = v_l(t) + \frac{D(t) - v_l(t)\tau}{\frac{v_l(t) + v_f(t)}{2b} + \tau} \tag{8.3}$$

, where t, a, b, v, τ, and $D(t)$ are the duration of simulation step t, acceleration, deceleration, speed, and reaction time of the driver, and gap between the two vehicles at the simulation step t, respectively. Since the v_{safe} can be higher than the speed that a vehicle can obtain or the limited speed of the road, v_{max}, then the desired speed, v_{des} is defined as [9], [228],[210]:

$$v_{des}(t) = min\{v_{safe}(t), v + \alpha t, v_{max}\} \tag{8.4}$$

The Krau model is recommended by SUMO due to its simplicity and high execution speed in [162].

- **Intelligent Driver Model**
  This model in addition to the parameters that Krau considers, also includes the velocity difference, $\Delta v = v_f - v_l$, between the two vehicles (l and f) for a single lane road, and the intelligent acceleration is defined as [238]:

$$\alpha_{IDM} = \frac{dv_f}{dt} = \alpha(1 - (\frac{v_f}{v_{des}})^\delta - (\frac{s^*(v_f, \Delta v)}{s_f})^2) \tag{8.5}$$

$$s^*(v_f, \Delta v) = s_0 + v_f T + \frac{v_f \Delta v}{2\sqrt{\alpha\beta}}, \qquad (8.6)$$

where v_{des}, α, s_0, T, β, $s^*(v_f, \Delta v)$, and δ are desired speed, acceleration, minimum gap distance between the two vehicles, desired time between f and l, comfortable deceleration, desired distance, and the exponent which models how the relation between v and α shall be modelled (1 shows linear, and inf shows constant α) [155]. These parameters in this model reflect real word scenarios and are well-justified [155], [153], and are shown in Table 8.1.

Table 8.1: Parameters of IDM for a light vehicle (length 5m) [154]

Parameter	Typical value (Highway)	Typical value (City)
v_{des}	$120km/h$	$54km/h$
T	$1.0s$	$1.0s$
s_0	2	2
δ	4	4
α	$1.0m/s^2$	$1.0m/s^2$
β	$1.5m/s^2$	$1.5m/s^2$

By considering the length of vehicles as L and with the velocity of v_l^t, the distance is defined in [210] as:

$$s_f(t) = x_l^t - x_f^t - L_l \qquad (8.7)$$

Some other models are available such as General motor (GM), Optimal velocity, Gipps, PWag2009, SmartSK, Wiedemann, and Fuzzy-logic models [210]. Some of these models are available in SUMO such as Wiedemann, IDM, and Krau [13], [153]. The authors in [210] found that IDM can more accurately model the behaviours of drivers compared to Krau and Widemann models.

8.2.4   Lane changing model

A lane changing model manages how, where, and when vehicles in a simulation environment change their lanes. These lane changes can be modelled with different approaches such as Lv [171], DK2008 [162], SL2015 [48], and LC2013 [113]. The last three models are implemented in SUMO. LC2013 is the model which was fully tested and has significant improvements compared to its origin (DK2008). This model, LC2013, is a comprehensive model which distinguishes four types of criteria for lane changing as [113], [162]:

- Strategic change

  This is referred to a type of lane change, which a vehicle needs to change its lane to stay within a predefined route, otherwise, the vehicle cannot follow its path. An example is entering the right ramp in the highway. In this regard, identifying the correct time, location, and the number of lanes, which are required to be changed are critical. Furthermore, the algorithm needs to consider the speed of the current vehicles, predecessor and successor vehicles to avoid being blocked or cause any deadlock. Deadlock is referring to a situation when two vehicles need to change their lanes, but both of these required lanes are occupied by these two vehicles.

- Cooperative change

  In this lane change, which is not mandatory, vehicles cooperate to allow lane changing. A roundabout with multiple lanes can be a good example. A vehicle in a roundabout, which has not reached to the exit points (edges), allows other vehicles to change their lanes. Another example is merging into one lane on an off/on-ramp.

- Tactical change

  In this change, a vehicle changes its speed in order to do a lane change and not to follow a slow leader either by moving to the left or right lane (if the right lane is occupied). To meet this criterion, the expected speed

of the desired lane should be higher than the current lane, and the vehicle should not constantly keep changing lanes.To calculate the efficiency of lane changing, the below formula is used:

$$Gain = \frac{\text{Expeced speed} - \text{Current speed}}{\text{Expeced speed}} \tag{8.8}$$

- Regulatory change

  This change is compulsory due to traffic laws [48] in some countries. For example, vehicles allow other vehicles to use the speed lane(s) if they are not using that lane, or they are on the speed lane for a long time. This can be accomplished by recording the time of being on a speed lane and setting up a threshold to move to slower lanes. It should be noted that other parameters are also required to avoid returning immediately to the speed lane.

8.2.5   Intersection model

Modelling an intersection is a complicated task in traffic modelling as both vehicles and intersections are required to be monitored to avoid any incident. On average, more than $10,000$ peopled died each year, from 1998 to 2007 (total reported deaths:$90,000$) due to accidents at intersections in the US [34]. Research shows bad weather is not a major contributor toward accidents at intersections. As reported by [54], less than 20% of incidents are due to bad weather, while about 80% are due to drivers errors [232], [168]. Vehicles at an intersection can be monitored by radio/light frequencies such as LIDAR, or by technologies based on image processing such as CCTV and stereo vision [102]. For example, the author in [121] proposed an approach for grouping vehicles based on motion segmentation, corner feature grouping, and speed similarity with an accuracy of 56% in 2014. The sensor setup was Monocular vision [102]. Monocular vision uses a single eye (e.g. camera, sensor) and has lower accuracy in measuring spatial distance compared to binocular vision [102]. In traffic modelling, an urban intersection, Fig. 8.1, usually has dedicated lanes for each direction, while at a rural

intersection, (e.g. Fig. 8.2), such directed lanes are missing (flared intersection) [231]. An intersection can have different shapes. Some of these shapes are shown in Fig. 8.3 [221].

Figure 8.1: An urban intersection, Google Maps.

Figure 8.2: A rural intersection, Google Maps.

An intersection in SUMO is modelled by the intersection, incoming lanes, and outgoing lanes. Indexing starts from zero at the rightmost lane and increases by one to the leftmost lane [114] as shown in Fig. 8.4. give-way rules are applied by an intersection to the coming vehicles with the help of a matrix to identify which

**This content is unavailable. Please consult
the print version for access.**

Figure 8.3: Types of an intersection [221]

vehicles need to give way and which do not. Vehicles approaching an intersection
inform the intersection about their speed and possible arrival times. Hence, the
intersection knows which lane will be occupied and for how long, which leads
to more efficient traffic management at an intersection. The speed of vehicles
approaching an intersection should not be higher than their braking capabilities.
If the approaching vehicle can pass the intersection with a safe gap with other
vehicles, it will pass, otherwise, it will stop. Furthermore, SUMO tries to prevent
intersection blocking, which happens by accumulating vehicles in an intersection,
with assessing different criteria such as length of vehicles and measuring the gap
between vehicles [114].

**This content is unavailable. Please consult
the print version for access.**

Figure 8.4: An intersection in SUMO [114]

8.2.6 Traffic Lights

Each traffic light has several phases, depends on the number of links, and each phase defines the duration and colour of the traffic light. These phases are controlled statically or dynamically. An improper traffic light controller leads to urban congestion, pollution, commuters stress, and financial loss. Since traffic behaviour is not linear, a linear and fixed traffic light logic is not an efficient method for traffic flow enhancement. Furthermore, the lack of adaptation to emergency scenarios is another drawback associated with a linear traffic control [202][201]. Intelligent Traffic Systems (ITS) are composed of various components such as microcontrollers, sensors, antennas introduced to overcome the limitations of conventional traffic light controllers. An emergency scenario can be detected by an ITS in different ways such as siren sounds, radio frequencies, or optical sensors. The range of detection at an intersection can be extended by the placement of additional connected sensors at the streets which form this intersection. Furthermore, drivers can be notified by real-time status messages submitted by an ITS. Sensors can either be installed under the road or over the road. Zhang et. al. in [263] proposed a new approach for controlling the traffic light via Infrared and CAN BUS. In the same year, 2009, the authors in [94] proposed a control mechanism based on fuzzy logic to enhance flow at an intersection. In Fuzzy logic, the truth values can be a real number instead of binary values. Fang. et. al. in [94] showed that the minimum cycle of a small intersection should not be less than $P \times 15$ seconds in order to avoid any hazard, where P is the number of phases. These phases are shown in Fig 8.5.

An example of a Traffic light implemented in SUMO is shown in Fig. 8.6, 8.7, and 8.8.

8.3 Behavioural model

Stress can be caused by many factors and affect human behaviours and the way humans interact with their environments. The stress of evacuees must not be ignored in an evacuation scenario. Some researchers have investigated the behaviour

This content is unavailable. Please consult the print version for access.
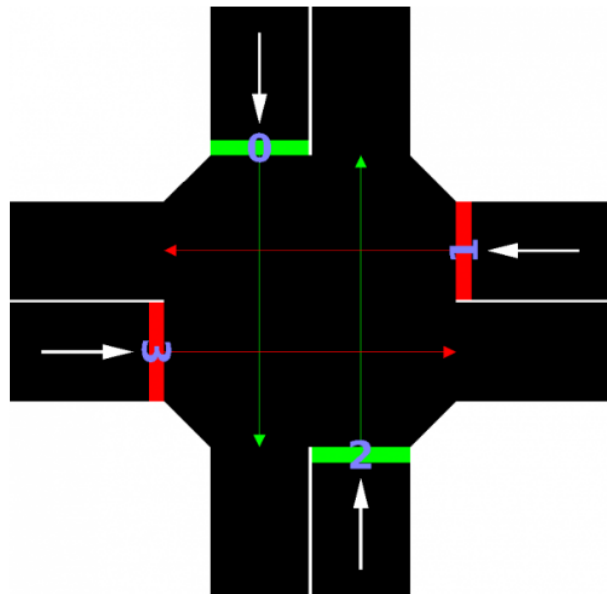
Figure 8.5: Traffic light phases [94]

Figure 8.6: A traffic light in SUMO [123],[123]

This content is unavailable. Please consult the print version for access.

Figure 8.7: A simple traffic light logic in SUMOm [123],[123]

This content is unavailable. Please consult the print version for access.

Figure 8.8: Traffic light colours in SUMO [123],[123]

of humans and tried to model this under different conditions. Helbing et. al. in
[134] studied pedestrian behaviours in a crowded environment affected by panic
(i.e. escape from a room filled by smoke). The authors in [261] utilised the Hel-
bing model to apply panic to their model and considered three levels of panic: not
nervous, slightly, and extremely nervous. The slight nervousness level showed the
best results. These drivers do not make fatal mistakes like drivers with the extreme
nervousness, as extreme nervousness increases the level of irrationals choices (e.g.
very small gap between vehicles, using no indicator during lane changing, ignor-
ing traffic lights). These drivers were also driving better compared to drivers with
no nervousness as an appropriate level of nervousness increase the efficiency of
decisions (e.g. overtaking a slow vehicle at the right time, a safe distance with the
front vehicle). The nervousness value of a driver is modelled by [261] as:

$$\frac{d\mu_i(t)}{dt} = (2\eta_i(t) - 1)\frac{1}{t_r}\mu_i(t)(1 - \mu_i(t)) \tag{8.9}$$

where,

$$\mu_i(0) = \mu_i^{(0)} \tag{8.10}$$

$$\eta_i(t) = \begin{cases} 0 & v_i(t) \geq V_{low}, \\ 1 & otherwise \end{cases} \tag{8.11}$$

and t is time, the nervousness of agent i is μ, $\mu_i^{(0)}$ is the initial value of $\mu_i(0)$, t_r is
the reaction time of agent i to modify its nervousness level, v_{low} is the threshold
that agent i thinks that it is driving slowly at nervousness level $\mu_i(t)$. An agent i is
considered to panic if $\mu_i(t) \geq 0.8$ [261].

The authors in [265] modelled evacuation by considering simple driving be-
haviours such as acceleration, headway ($\Delta(x)$), and reaction time through simula-
tion. Tu et. al in [241] studied the importance of driving behaviour (e.g. accelera-
tion, headway, and speed) on evacuation times and emergency scenarios. Tsai et.
al. in [239] studied the behaviour of evacuees and emotions in the evacuation in a
3D environment for training purposes.

Mubasher et. al. in [184] studied how driving behaviours depend on many

human factors such as age, anger, distraction, gender, mental stress, and response time. They modelled the parameters in IDM for three types of drivers by considering the Five Big Factors (BFF). The BFF are dimensions that describe human personalities and are [244]: Openness to experience (O), Conscientiousness (C), Extroversion (E), Agreeableness (A), Neuroticism (N). These are referring to imagination and creation, organization and structuredness, level of socialization, friendliness, and moodiness, respectively [244], [178], [184]. The three types of driving behaviours adapted in IDM are [244]:

- Resilient who easily adjust themselves with different situations

- Over-controlled who are peaceful and co-operative people

- Under-controlled who take risks and are non-cooperative.

8.3.1 Driver behaviour questionnaires

Driver Behavior Questionnaires (DBQ) are widely accepted sets of questions that researchers ask drivers to analyze and model their driving behaviours. DBQ were originally proposed by Reason et. al. in 1990 [216] and contained 50 multi-scale questions that were categorized into three different groups: violation, error, and lapses. Violation refers to types of driver behaviours which intentionally break the traffic rules. Errors happen because of mistakes made by drivers and involve a crash risk. Lapses are also due to drivers mistakes, but do not impose a crash risk [216], [177]. This model has since been slightly updated [200], [193], [75], [132], [177], [218]. For example, the authors in [177] added aggressive behavior to the model. To measure the correlation of items within each category, authors usually conducted a test known as Chronbach's alpha, shown in equation 8.12 [98]. Chronbach's alpha measures the accuracy of a statistical test associated with medical science and psychology. The Alpha value varies between 0 to 1 and a higher value means a higher correlation, although homogeneity tests might be required to utilize this value as it is proven that multidimensional data does not necessarily have a lower alpha value [235]. The equation is as follows:

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K}\sigma_{Y_i}^2}{\sigma_X^2}\right), \qquad (8.12)$$

where K is the number of questions, σ_X^2 is variance associated with the observed total test scores, and $\sigma_{Y_i}^2$ is the variance of item i [93],[59], [262].

An implementation of this model in SUMO is discussed in the next chapter (9), section 9.2.1.

8.4 Geographic information system

Geographic information systems (GIS) are widely used in academia and industry to capture, store, process, and visualise spatial data. Spatial data are the type of data that have attached information about their locations and extents. These types of data require special kind of processes, visualisation engine, and storage which are impossible to achieve without GIS tools like Quantum GIS (QGIS), ArcGIS, GeoServer, Leaflet, and PostGIS. This research uses QGIS, Geoserver, PostGIS, and web maps for Extraction, Loading, and Transformation (ETL) of geospatial data. A spatial shape can be represented in different forms such as point, linestring, polygon, and multipolygon. The centre of each shape is called the centroid.

8.5 Disaster management

A disaster is defined as: "A serious disruption of the functioning of a community or a society at any scale due to hazardous events interacting with conditions of exposure, and capacity, leading to one or more of the following: human, material, economic and environmental losses and impacts" by the United Nation Office for Outer Space Affairs [14]. A disaster can be categorised into different groups as follows [191], [56], [125]:

- Natural disasters

Natural disasters are types of natural events such as floods, earthquakes, tsunamis, hurricanes, and volcanic activities, which can impact the main infrastructures of a city, properties, and airlines severely.

- Human-instigated disasters

  Human-instigated disasters can be created by technological hazards such as transportation accidents (on air, land, sea), explosion, fire, chemical and toxic releases, and pandemics diseases (e.g. Swine influenza, Severe Acute Respiratory Syndrome (SARS)).

- Acts of terrorism

  Acts of terrorism are types of disasters that are created due to human violence. Intelligent services are responsible for the identification and elimination of these types of disasters in most countries. This type can also be categorised as human-instigated disasters [191].

A disaster may require a large set of responses such as providing shelters, medical assistance, search and rescue, dealing with embassies for tourists involved in the disaster, traffic management, and dealing with human injury and casualties. These responses require collaboration from different governmental and non-governmental parties such as police and fire departments, emergency services, military, transportation management, Red Cross, and United Nations International Strategy for Disaster Reduction (UNISDR). Due to this complexity, different countries have various approaches and practices for disaster management. In New Zealand, disaster management has a dedicated organisation that fits under the Ministry of Civil Defence and Emergency Management (MCDEM). In terms of control and communication, the New Zealand Civil Defence has published a schema as a natural disaster response guideline [66]. This document also predicted fatalities to vary between 140 to 2000 if the earthquake hits in the day time, and 140 to 560 if it hits at night time. Tertiary and welfare infrastructure and services may be damaged or disrupted due to a severe earthquake.

Wellington City Council (WCC) as a regional response has published a set of responses for mitigating the impact of an earthquake as shown in Fig. 8.10.
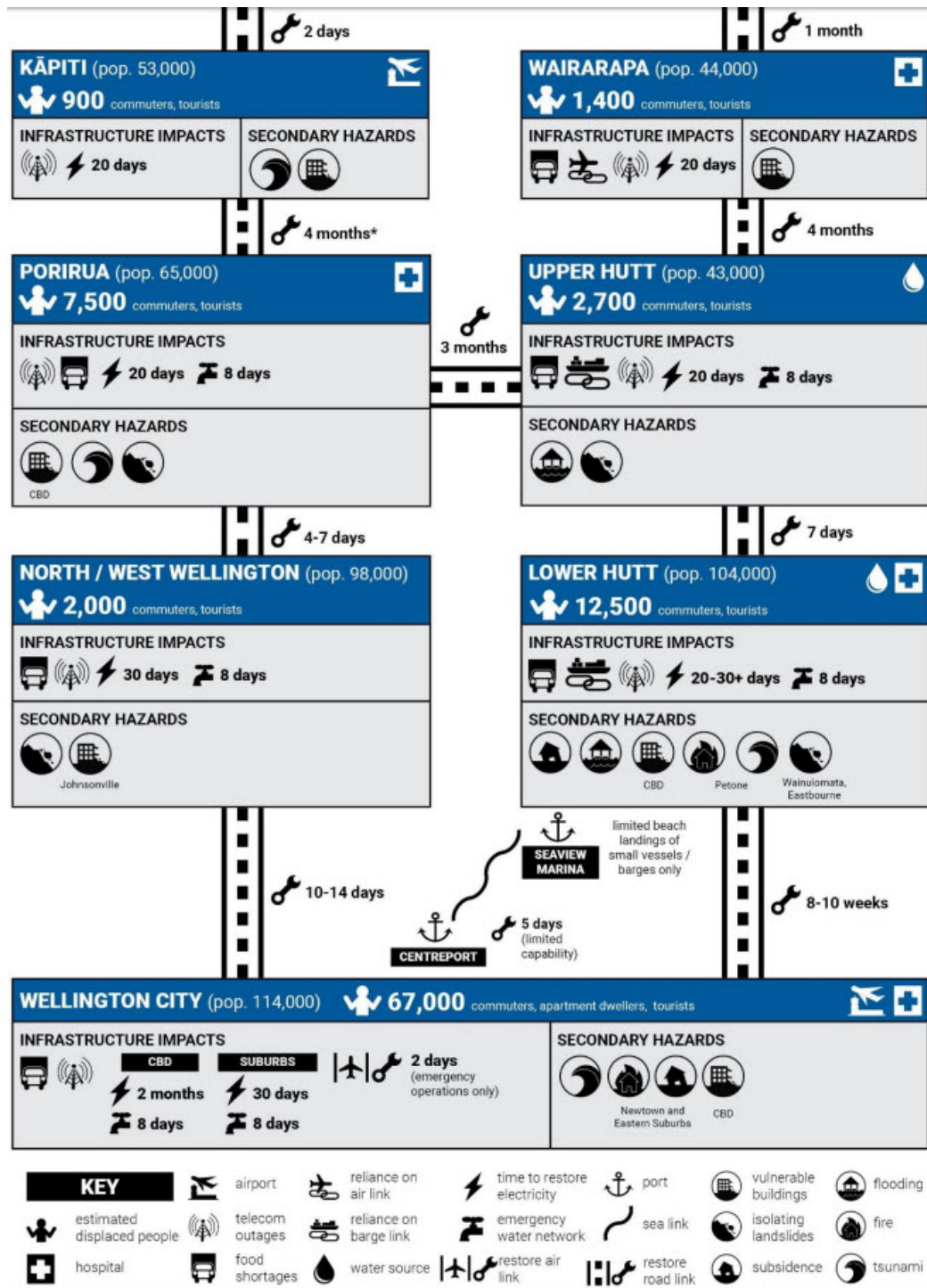
Figure 8.9: Civil defence anticipated restoration plan for Wellington [66]

Hazard	Effect on ground	Effect on facilities	Mitigation options: existing facilities	Mitigation options: planned facilities
Fault movement	Ground disturbances vertically and horizontally over a zone depends on depth to rock below surface. Cracks in land surface.	Upheaval, tearing apart, movement of foundations, severe damage to structures which cross the fault.	1. Verify. 2. Assess impact. 3. Options: a) strengthen to survive b) move facilities from fault zone c) limit damage by providing weak links or isolation	1. Verify. 2. Assess impact. 3. Options: a) construct facilities elsewhere b) incorporate special strengthening c) provide weak links or special isolation to limit damage
Ground shaking	Violent horizontal and vertical motions for up to one minute duration.	Cracking, fracture, collapse of buildings. Breaks in underground services. Deformation of surface infrastructure.	1. Verify. 2. Assess impact. 3. Options: a) strengthen or base isolate b) secure/improve vulnerable parts c) limit damage by providing weak links or isolation.	1. Verify. 2. Assess impact. 3. Options: a) comply with current codes for design and construction b) incorporate strength and resilience c) secure vulnerable parts and contents
Liquefaction	Shaking causes some soils to behave like liquid, causing loss of support to structures above. Such soils may be up to 10m below ground surface. Lateral movement of large soil masses, especially adjacent to rivers. Variable subsidence of ground surface.	Sinking and tilting of structures supported on liquefied material. Severe damage to underground services. Flotation of empty underground tanks and chambers.	1. Verify. 2. Assess impact. 3. Options: a) install piles b) install gravel drains c) drain liquefiable layers d) prepare for quick reinstatement	1. Verify. 2. Assess impact. 3. Options: a) compact ground at site b) install piles and gravel drains c) drain liquefiable layers
Slope failure	A significant soil masses moves bodily down the slope, from few hundred millimetres to many metres. Landslides occur at many different locations.	Ranges from deformation of foundations and structural failures to total destruction of site and all buildings and infrastructure above and below ground.	1. Verify. 2. Assess impact. 3 Options: a) stabilise slope – retaining walls b) stabilise slope – ground anchors c) improve drainage, reduce erosion	1. Verify. 2. Assess impact. 3. Options: a) find a better site b) stabilise slope retaining walls c) stabilise slope – ground anchors d) improve drainage, reduce erosion
Tsunami	Land flooded. Scouring action erodes soil dramatically	Flooding of basements. Undermining/destruction of surface infrastructure. Exposure/damage to underground services. Undermining of foundations. Bodily movement of some structures, equipment, vehicles etc.	1. Verify. 2. Assess impact. 3. Options: a) construct protective sea walls b) shift critical facilities to higher level	1. Verify. 2. Assess impact. 3. Options: a) find a better site b) construct protective sea walls c) design special foundations / dikes d) put critical facilities at high level

Figure 8.10: Earthquake hazard mitigation measures, published by Wellington City Council (WCC)

8.6   Conclusion

This chapter briefly studied the fundamental of traffic modelling and simulation by explanation various involved components such as lane changing models, car following models, intersection models, human behaviour models, and disaster types. The next chapter utilises these principals to offer a multi-disciplinary large-scale evacuation framework.

Chapter 9

Traffic evacuation modelling

My research inspired by [261] proposes a multi-disciplinary evacuation framework for Wellington, New Zealand. In this framework, various components such as GIS, database, car-following model, and behavioural model are integrated together to provide more realistic traffic modelling. More specifically, I propose a Traffic Framework for Decision Assessment, and Modelling (TFDA2M) of land transportation by integrating real-world components such as road networks, intelligent drivers, driving cultures, and driving errors and violations into a simulation environment to investigate how some realistic assumptions (e.g. utilisation of navigation systems during an evacuation process, Intelligent Driving Model, behavioural model) can affect the performance of transportation infrastructure. Furthermore, my study assesses TFDA2M on Wellington city to measure the performance of its land transportation infrastructures in a large-scale evacuation process by applying various functional features such as power and telecommunication failures. In terms of intelligent driving, my research investigates the impact of using navigation systems (commonly known as GPS) in an evacuation and identifies the negative impact of this technology in an evacuation process in a dense urban area with a limited number of safe zones, and propose an approach to address this issue.

Yuan et al. in 2017 [261] proposed a multi-level and multi-agent model for

evacuation. In their framework, some possible functions such as the distribution of demand, vehicles incidents and mechanical failures, power failure, malfunction of communication networks (e.g. processing servers), and use of modern technology (i.e. Global Positioning Systems (GPS)) by drivers were ignored. Furthermore, in their research similar to previous research, the evacuation started at time zero, T_0, of simulation, which means people are ready to start their trips to the evacuation shelters (safe zones) immediately at time zero. This assumption does not hold in reality.

In an emergency evacuation scenario, not all people are ready to evacuate the city immediately and even informing all citizens in a city at once is very challenging. Since some people need to fill the fuel tank of their vehicles, collect their family members, and the evacuation warning needs to propagate through the population, not all evacuees will participate in the evacuation process at T_0. These factors may impact the evacuation scenario as some road closure may be created by incidents or breakdown of vehicles, that only affect part of the evacuation population.

The rest of this chapter is organised as follows. TFDA2M is proposed in the next section. Section 9.2 discusses the methodology, data model, parameter calibration, and other involved components which are used in TFDA2M. Section 9.3 discusses multiple scenarios and later this chapter evaluates how each set of scenarios behave. In addition, this chapter provides an overview of the performance of the road networks at the end of this section, Subsection 9.3.3, by developing a series of Business Intelligence (BI) dashboards and heatmaps with Business Intelligence and GIS as part of TFDA2M. These tools can be used by authorities for traffic enhancement and evacuation planning.

9.1   Evacuation frameworks

Yuan et al. in 2017 [261] proposed an evacuation framework in 2017. Their framework is composed of driver behaviour, agent, geospatial, and mission managers. The driver behaviour manager translates the behaviour of agents, drivers,

to a practical model in the simulation. The agent manager parametrizes the decisions of drivers, monitors and manages vehicles. The geospatial manager deals with geospatial data such as road topology. The mission manager can be considered as the data repository, which communicates with users, imports the data and configurations, and export the results to an appropriate format.

Shahabi et al. in 2017 [223] proposed a framework for large-scale evacuation with a focus on dynamic evacuation routes which uses stored evacuation plan to accelerate rerouting. Their algorithm uses the recorded capacity of the road, which uses A^* and may be dependent on the number of lanes and update the routes during the evacuation process.

Wang et al. in [249], 2016, proposed a statistical scenario-based framework for evacuation planning through considering available link capacity, distance, cost, and travel time. They formed a graph of road networks.

The authors in [241] proposed a framework based on the emotional behaviour of evacuees in an airport evacuation. They demonstrated that panicked evacuees make more mistakes. They also showed that the reduction of the safe gap between vehicles can decrease the evacuation time, although increases in speed and acceleration do not make a dramatic difference in the evacuation time.

Lindel in [170] highlighted that the departure time, demand, routes, number of shelters, time, distance, route costs, in addition to source and destination are important factors in an evacuation framework.

The authors in [166] proposed a multi-agent framework for large-scale pedestrian evacuation, and highlighted that bottlenecks in the evacuation routes can significantly impair the evacuation process.

9.1.1 TFDA2M architecture

The proposed framework, TFDA2M, is composed of several components as shown in Fig. 9.1. The interactions between different components are shown by directed arrows. Each entity in these components and simulations has a unique identifier and every single interaction is logged.
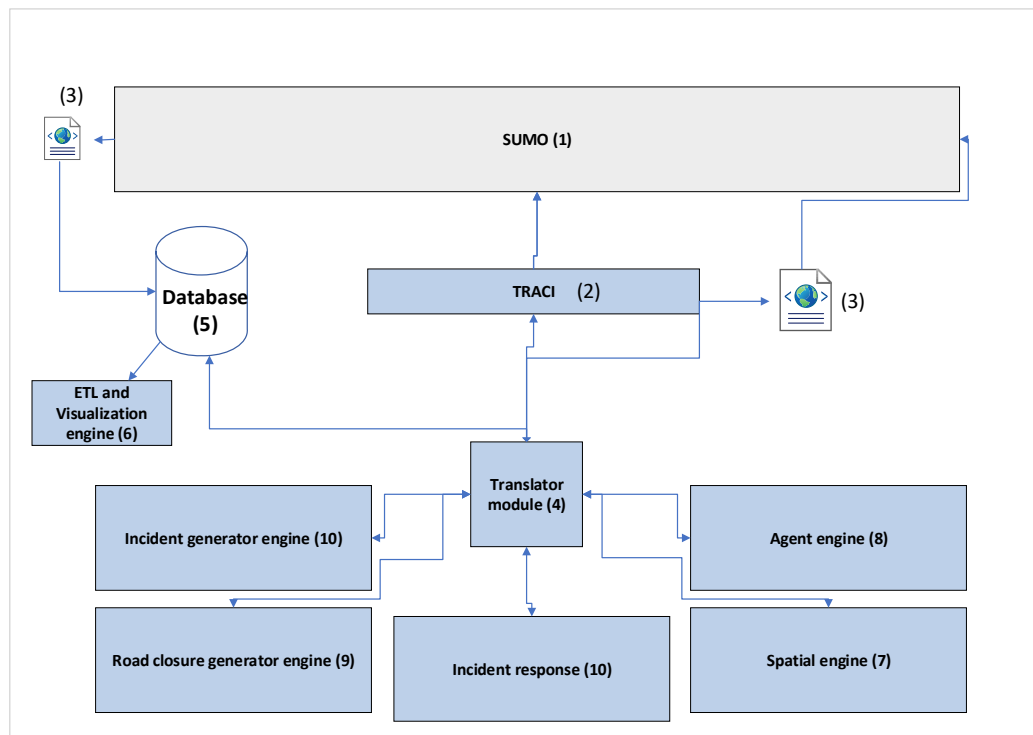
Figure 9.1: Component of TFDA2M

1. Microscopic traffic simulator:

   Microscopic traffic simulator which can model the movement of each mobile object individually. I used SUMO for this purpose.

2. TraCI:

   Traffic Control Interface (TraCI) protocol communicates with the simulation engine, SUMO, during runtime [50].

3. Flat files:

   Extensible Markup Language (XML) is a language that store data in a human and machine readable format by using a set of rules.

4. Translator module:

   This module is responsible for communicating with different engines, database,

and passing parameters to SUMO through Extensible Markup Language (XML) and Traffic Control Interface (TraCI).

5. Database module:

   This research used Microsoft SQL server for storing and querying the collected data, but the logic can be extended to any other data storage format and structure such as PostgreSQL, MySQL, or flat files.

6. ETL and visualisation engine:

   This engine Extract, Load, and Transform (ETL) stored data and then visualise the data based on the request.

7. Spatial engine:

   This engine translate non-spatial queries to spatial queries (if applicable) and vise versa. An example is converting an edge name to a spatial line.

8. Agent engine:

   This engine generates traffic flow based on the given settings (e.g. settings the parameters of IDM, enabling navigation systems).

9. Road closure engine:

   This engine generates the commands for road closures that can be executed in SUMO. Road closure can be done in two different ways: on-demand and pre-defined. On-demand road closures are requested by the translator module through TraCI. Pre-defined road closures are run by an XML file.

   As shown in Fig. 9.2, at the beginning of the simulation the road closure engine sends a request to the translator module by passing the street names to obtain their equivalent edges identifiers (IDs). The translator module passes these received street names to the spatial engine in order to obtain their equivalent edge *ID*s. After receiving the edge *ID*s from the spatial engine, these *ID*s are passed to the road closure engine by the translator module in the same received order. Then, the road closure engine generates the road closure commands along with the beginning and removal of these

road closures and passes these to SUMO for execution through the translator
module.

10. Incident response and generator engines:
    The incident generator engine is responsible for simulating an incident through
    a road closure, and the incident response engine simulates the task of a road
    closure removal as part of the duties of an incident response team. The in-
    cident response module models the consequence of an incident through a
    road closure as suggested by SUMO [19], but it can be modified to send
    an emergency vehicle to the location of the incident. As illustrated in Fig.
    9.3, the incident generator engine sends an edge ID request for street i at
    the time t to the translator module. Then, the translator module retrieves the
    equivalent edge ID through the spatial engine and passes this information to
    the road closure engine. The road closure engine responds by sending the
    closure command for edge i and start time t. Then, the translator module
    passes this received command to SUMO and immediately calls the incident
    response team in order to simulate calling to an emergency number. When
    an incident response team receives this call, it triggers a timer for request-
    ing the road closure removal. This random number timer $t2$, is the road
    closure removal time, at which the road closure should be removed. This is
    done by calling the translator module. Then, the translator module calls the
    road closure module to retrieve the removal command and then passes this
    received command to SUMO for execution.

9.2   Methodology

This section covers the details of the simulation model and the components of the
proposed TFDA2M.
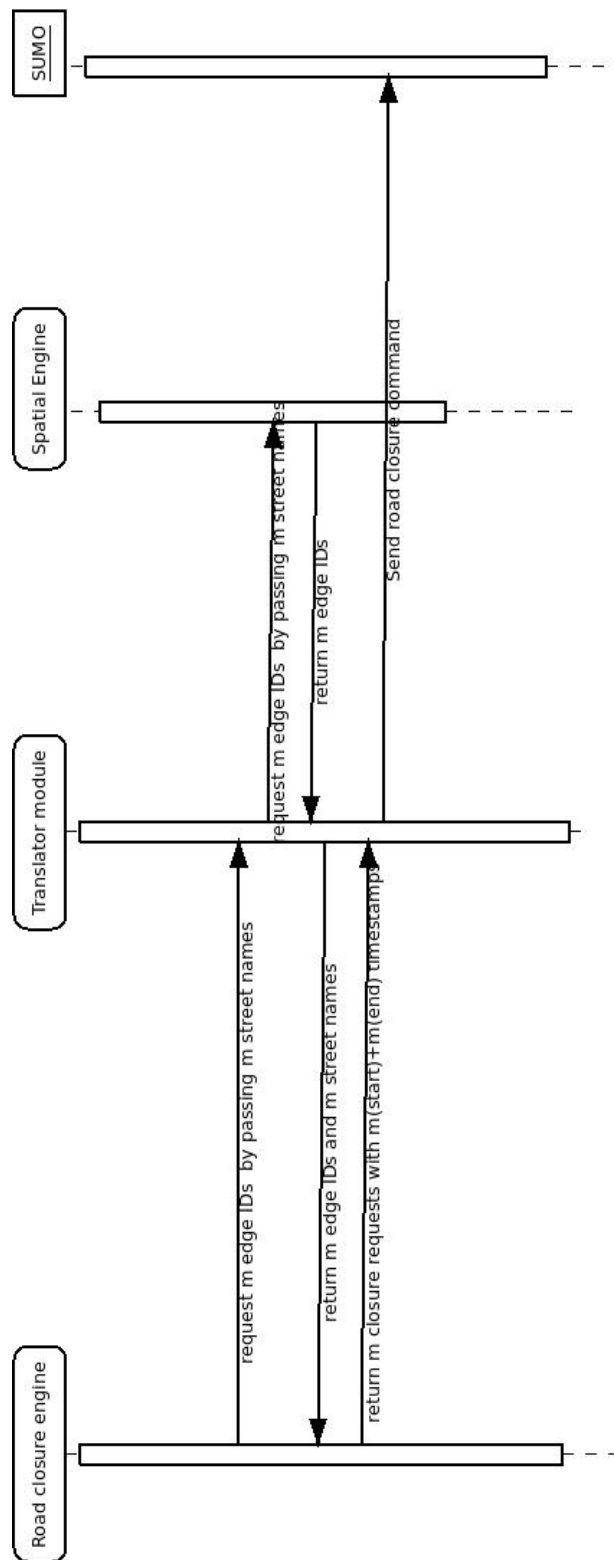
Figure 9.2: Road closure sequence diagram
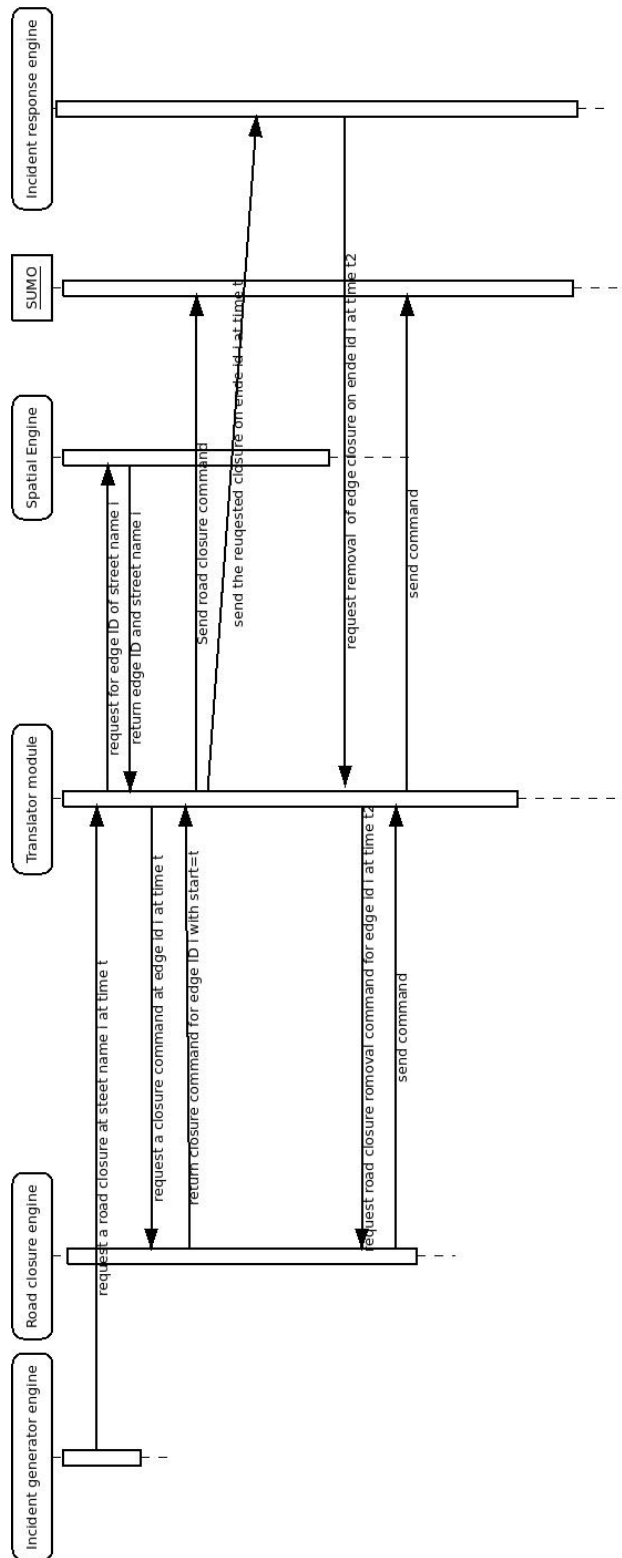
Figure 9.3: Incident generator sequence diagram

9.2.1   Simulation configuration and Model calibration

This subsection describes simulation settings, parameters selection, model calibration, and assumptions made in this novel framework (TFDA2M).

Car-following model, and parameter selection

For assigning values to the parameters of IDM, this chapter utilised the values recommended by Pourabdollah et al. [210]. They monitored and collected information of about 200 vehicles to determine the values of IDM. They applied a Genetic Algorithm for finding the optimal values of IDM.

Traffic light, lane changing, intersection model, and routing algorithm

Modelling an efficient traffic light logic in traffic modelling is a critical task. In Wellington, traffic lights are controlled by Sydney Coordinated Adaptive Traffic System (SCATS) [43]. Due to lack of access to the algorithm/logs of this system, TFDA2M uses the SUMO default traffic light logic, which is 90 seconds and used as an acceptable model in the other research [44]. The lane changing model uses LC2013 as suggested by SUMO [113]. For the intersection model, this research uses the default intersection model available in SUMO [114]. Since A^* is a common routing algorithm in transportation and has a good performance, speed and implemented in SUMO, TFDA2M uses this algorithm for routing.

Model calibration

To measure the accuracy of TFDA2M, after adjusting the parameters of agents by applying IDM and road topology, I conducted 30 independent tests to measure the travel time from different sources to different destinations and compared these results with the Google APIs to measure the accuracy of travel times. In most cases the values were very close although I observed a variation of 5% in travel times. This is possibly due to the difference in the timing of traffic lights. An example is shown in Fig. 9.4. The right side of this screenshot is the simulation environment and the timer (on top) shows that from the start of the trip to the current location

(an Orange thick line shows the route), 435 seconds is required. On the left side, Google Maps show a similar travel time of 7 minutes (420 seconds) for the same route.

Incidents

In New Zealand about 5% of incidents are due to vehicle factors, 12% are due to infrastructure and weather, and the rest, 83%, are due to human errors, based on [31], as shown in Fig. 9.5. There is a direct relationship between human errors and anxiety [196], [95]. Since most of the incidents are because of human errors, the incident rate should be increased during an evacuation scenario due to the increase in anxiety, human errors, and mechanical failure.

There are several categories for incidents such as minor, mild, and severe. Categorizing incidents into the correct group is a challenging task in reality because it mainly depends on the conceptual definition of the severity among operators and response teams in an emergency as some metrics may be unknown at the time that an incident is being logged. Furthermore, in a real-world evacuation scenario, it can hardly be assumed that people will stop and pull-over due to minor or moderate incidents especially when no-body was injured or the vehicles involved in the accident can move. Indeed in some countries (e.g. US, California [20]) vehicles *must* be moved to the shoulder of the road if involved vehicles can move and no one is injured or killed in the collision. Therefore, this study uses the total amount of injuries and fatalities to estimate the number of severe incidents. Based on [32], these numbers are about 267 and 7 (total of 274), respectively, per 100,000 population in New Zealand for 2016 [33]. Since the population of the Wellington region in 2016 was about 207,000 [64], therefore, there would be about 548 ($274 * (207,000/100,000)$) fatalities and injuries in 2016. This value divided by 365 days gives about two injured and fatal incidents per day. This research assumes two people per accident. Hence, in the evacuation scenario, there will be one major incident, although in a real-world scenario it may be more as described earlier (due to anxiety, human error, and mechanical failure). With the incident, the incident response teams clear the incident after about two hours as

Figure 9.4: Travel times for the same route in Google Maps (left, 7 minutes) and SUMO (right, 435 seconds)

**This content is unavailable. Please consult
the print version for access.**

Figure 9.5: Factors involved in incidents [31]

this is the standard clearance time for a major incident in New Zealand[26]. In the
modelling both one and two incidents are considered, which is explained later.

Non-homogeneous drivers and on-demand routing

Driver Behaviour Questionnaires (DBQ) are a set of questions asked of drivers to
categorize their driving style (Chapter 8). The behaviour of a driver is obtained
based on the probabilities in the DBQ and is shown in Table 9.1. This table shows
the probability of each behaviour [177], and the mapping table which converts
these DBQ survey results to numeric values is shown in Table 9.2. The column
Function in Tables 9.1 and 9.2 demonstrates the naming convention associated
with a specific driver behaviour through the asked question (column Question in
Table 9.1). Probability of success if the answer of a driver belongs to a specific
group (Never, N, Hardly, H, Occasionally, O, and Quite often, Q) is shown in
columns 3 to 8 in Table 9.1 (i.e. 6 categories). The column Type in Table 9.2
demonstrates the type of parameter to model the mapped function (F1 to F5) in
SUMO through filling an array of 100 elements with the given Bernoulli distribu-
tion and selecting one element based on the uniform distribution through

$(sample(categories[1 \text{ to } 6], 100, probabilities, replace = TRUE))$. If the se-
lected probability from this 100 element array is less than 0.5 then it is categorised
as 0 and if it is greater than 0.5 it is categorised as 1. If the scaling is required,

then this selected random number from the array is scaled. The methodology of implementation is shown in the last column (Effect and implementation) of Table 9.2.

In addition, the behaviour of each driver in the IDM is obtained from 200 trips based on real-world scenario [210], calibrated for SUMO, as explained in Subsection 9.2.1. By interpolating these factors for generating the vehicles and drivers, the total unique types of vehicles and drivers is 1000 (5 (DBQ) $*$ 200 (IDM)= 1000). Furthermore, drivers are equipped with GPS and the best path is shown in their GPS, similar to the real-world scenario (Fig. 9.6), and drivers select this path if it is faster than the current route. In order to simulate having the GPS and model the intelligent driving, this research uses A^* for routing algorithm and vehicles update their routes on specific intervals based on the fastest route reported by their GPS. A^* is a popular routing algorithm which is used in many papers [223], [133] [104], has a high performance and accuracy, and finds the best route with the smallest cost (e.g. travel time).

In case of a large scale evacuation since the demand is very high and especially when there are a limited number of evacuation destinations (i.e. safe zones/shelters, in this modelling there are two evacuation destinations: Burma road and Motorway), all the vehicles within similar proximity are informed about the faster route simultaneously and change their route instantly (even the new route is just a few minutes faster). This causes the new route returned by the GPS (navigation system) to get congested, and the previous route becomes the new faster route and is selected by GPS again. This is a common problem in computer networks and mostly referred to as the routing loop problem [187], and needs some form of heuristic to resolve. To solve this on-demand rerouting issue, I added a condition that drivers only accept a new route if the travel time of this new route is lower than than the following condition:

$$\text{Travel time of returned route by GPS} \leq$$
$$\text{Remaining travel time} - \text{Routing interval} \tag{9.1}$$

Routing interval is the interval of executing the routing algorithm (in this case

it is every 15 minutes) and remaining travel time is the remaining travel time to the current destination. This equation guarantees that only routes which are faster than the current routes are selected by drivers and reduces the fluctuations between different routes. This is guaranteed because the remaining travel time will always be lower than the addition of reaming travel time and routing period (i.e. Travel time of returned route by GPS + Routing intervals $\leq$ Remaining travel time).

This content is unavailable. Please consult the print version for access.

Figure 9.6: Faster route is available [23]

Table 9.1: Frequencies of responses collected in New Zealand through Driver Behaviour Questionnaires (DBQ) [177] [177]

Function	Question	Never	Hardly ever	Occasionally	Quite often	Frequently	All the time effect
F1	Attempt to drive away from the traffic lights in third gear	82%	14%	3%	1%	0%	0%
F2	Miss "Give Way" signs, and narrowly avoid colliding with traffic having right of way	85%	13%	0%	0%	0%	0%
F3	Disregard speed limit on a residential road	45%	34%	16%	2%	2%	0%
F4	Overtake slow driver on the inside	70%	15%	12%	1%	0%	1%
F5	Disregard the speed limit on the open road	44%	32%	18%	3%	2%	0%

Number of vehicles and trips

An accurate estimation of the number of vehicles provides a more realistic output, therefore this research utilizes the statistics provided by the Ministry of Transport in 2015 (New Zealand). Based on [37], the number of light vehicles in Wellington per 1000 population is about 650 and based on [65] the population of the

Table 9.2: Implementation of Driver Behaviour Questionnaires in SUMO through mapping the type of behaviour (column Function defined in Table. 9.1) to a specific parameter (Type) and the methodology (Affect and implementation).

Function	Type	Effect and implementation
F1	Binary (-1 or 0)	-1:Stop, 0:Driver speed up at the Yellow phase and try to pass if they can pass even at the red phase, see section 8.2.5
F2	Range $[1, 100]$	Ignore other vehicles with the given probability
F3	Binary(0 or 2)	Speed up to two times faster than the speed limit if there is a safe gap
F4	Range $[1, 45[126]]$	Change lane to overtake slow vehicle, see section 8.2.4, tactical lane change
F5	Binary(0 or 2)	0:Max(speed of a vehicle)=speed limit,2:Speed up to two times faster than the speed limit if there is a safe gap

urban area of Wellington is about $207,000$, which leads to about $132,000$ vehicles owned in this city. I assume half of these motor vehicles (one vehicle per four people) will be used for evacuation in the worst-case scenario which is about $50,000$ vehicles, although as part of sensitivity analysis and due to lack of access to the required data, I set the number of vehicles to $10,000$, $25,000$, and $50,000$. In addition, further increasing the number of vehicles is not practical as in the best-case scenario without any traffic disruption event, the results of my modelling shows (discussed later in this chapter) that evacuating $50,000$ takes about 17 hours which is not possible in a real-world scenario (17 hours is too long as evacuees are required to eat, rest, and have access to amenities). Due to this long evacuation time, evacuating more than than $50,000$ vehicles is not practical in real-life, although such a massive evacuation can be done in several phases if the evacuation process can be broken down to multiple stages.

Finding an accurate estimation of required trips by people prior to starting the evacuation process in an evacuation scenario is a challenging task as it requires many datasets, and statistical data about number of times that vehicles fill their gas tanks each day, number of students, number of people who have private vehicles, locations of schools and gas stations, number of infants at kindergartens, time and day of the evacuation and many other factors, which are challenging to achieve. Therefore, for simplicity, this research assumes the total number of trips are approximately equal to the total number of children, who need to be col-

lected by their parents from kindergartens and schools. Based on [36], about 20% of people in New Zealand are between 0 to 14 years, therefore, about $41,400$, $41,400 = 20\% * 207,000$, children needed to be collected, and this research assumes one trip per child, which leads to $41,400$ trips out of $132,000$ vehicles, which is one third of total vehicles. Therefore, the ratio of random trips is one-third of the total vehicles in this research. The source and destination are selected randomly based on the uniform distribution, and these random trips start at T_0, and at the end of these trips, agents move toward the closest evacuation points, although the length of the trips which is required for collecting children may be lower than the length of these random trips.

9.3   Scenario-based simulations and results

As some of the impacts of a natural disaster such as earthquake may be unknown, it is possible to model the potential impacts of a disaster in multiple ways through various settings and still miss some of the actual impacts of a disaster [1]. These settings may be independent of each other and may not occur at the same time. Depending on the nature and severity of a disaster, the proportion of people who are required to evacuate the city and the damage to the infrastructure/properties may vary. For instance, in the case of a tsunami warning, only people who are living in the tsunami hazard zones are required to evacuate, and other people who are living further inland may or may not evacuate the city.

This section models several scenarios with a focus on utilisation of GPS technology for obtaining real-time information about the road networks (denoted by Navigational assistance), variation of demand (denoted by Number of vehicles: $10k$, $15k$, $25k$), power failure which is modelled through Traffic light malfunction (denoted by Traffic light malfunction), divergence of demand (denoted by Extra trips ($15k$)), accidents (denoted by Accidents), and damage to main routes (denoted by Closure), which are explained later in this section. The first subsection, 9.3.1, provides a discussion of the general settings for all the scenarios, an overview of each simulation scenario individually, naming convention, and the

criteria that may impact the modelling of an evacuation scenario. In this chapter, a case study is composed of one or multiple related scenarios.

The second subsection, 9.3.2, provides the results and discussions of each case study.

9.3.1 Case studies descriptions

This subsection explains the case-studies and the settings, and the results are provided in the next subsection.

General settings of all the case-studies

The criteria involved in the modelling of the evacuation of a city must be an accurate reflection of the characteristics of that city. For instance, different people and cars have different driving behaviours and mechanical properties, respectively. Therefore, it is crucial that an evacuation framework models these various real-world criteria. A wide range of driving behaviours exist even in a small random sample set of drivers and depend on many factors such as age and gender [199]. Driving behaviour varies more if the drivers originate from different cities, countries, and continents [200], [250]. To efficiently model this non-homogeneity due to the difference in driving cultures, TFDA2M utilises the DBQ which was obtained from New Zealand by researchers at Massey University [177] (Sections 8.3.1 and 9.2.1) to represent New Zealand driving culture to provide the parameters of the driving model (i.e. IDM). In addition, as Wellington is connected to the wider area through just two main land transportation points (Burma road and Wellington Urban Motorway) as shown in Fig. 9.7, the focus of the scenarios in this chapter is evacuating Wellington through land transportation by using only these two exit points. Hutt Road is another road which is spatially located very close to Wellington Urban Motorway (referred to as Motorway in this chapter) and joins Wellington Urban Motorway in upstream, and therefore, it collapses to a single route (i.e. Wellington Urban Motorway). Non-land transportation in-

frastructures such as Ferry terminals and Wellington airport are located in high hazard zones (e.g. tsunami, liquefaction, and ground shaking, as shown in Fig. 2.11), and have limited capacity which may not be operational efficiently in an evacuation process, therefore this research only focuses on evacuating Wellington through individual vehicles. Furthermore, in case of a disaster, it is likely that not all the evacuees are ready to start their trips toward the safe zones (evacuation points), or the warning cannot reach all evacuees immediately. Therefore, vehicles depart their source within 15 minutes of the beginning of the simulation (time zero) through the Poisson process. The details of the simulation configuration are addressed in Section 9.2.1. The rest of this section covers the details of each case study.

This section explains six case studies (named *A* to *F*) of which the settings are provided in Table. 9.8. The next section demonstrates the results of these case studies in the same order (i.e. *A* to *F*).

- **Case study *A*, optimistic with navigational assistance**
  In this case study, Scenario $A - I$ to $A - III$, I evaluate the impact of utilising GPS for obtaining real-time routing information about the road networks during an evacuation process. In this set of scenarios, there is no damage to transportation infrastructures as evacuees receive in-advance warning, all member of a family are at the same location (e.g. home) and there is no need for an extra trip to collect a family member, there is no immediate threat, and people evacuate the city peacefully and calmly.

  For example, consider evacuating a city because of a fire or hurricane [18], [85]. A big fire in 2018 caused authorities of Nelson city, New Zealand to evacuate the residents [35] with a prior notice at various stages, or a hurricane in Florence, South Carolina, USA, caused authorities to inform evacuees about the evacuation process in advance before the disaster [230], [24]. It can also be assumed an evacuation starts when all the members of a family are at home, and there is no need to pick up any family member during the evacuation process. Therefore, no extra trip is required from the
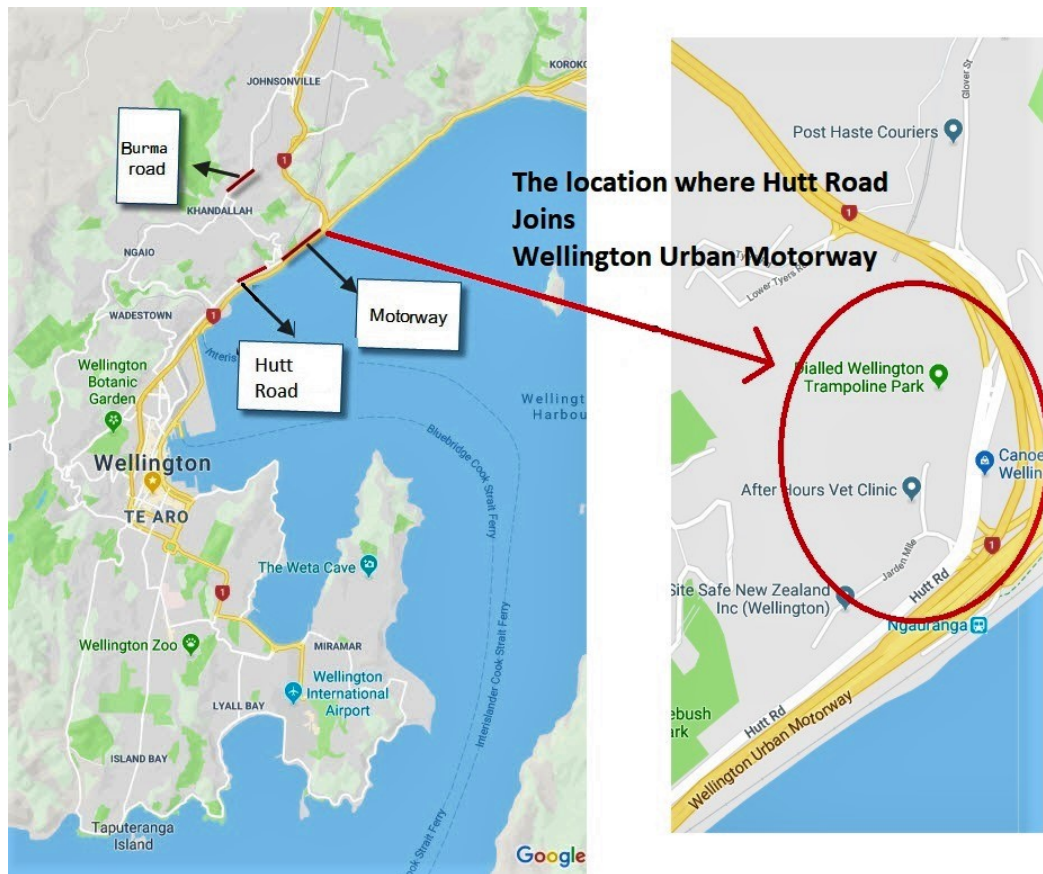
Figure 9.7: Two main exit points of Wellington, Burma road and Motorway. Hutt Road is spatially located very close to Motorway and merges with Motorway, therefore I consider Hutt Road as part of Wellington Urban Motorway, Google Maps.

Case study	Scenario Name	In-advanced warning	Navigational assistance	Number Of vehicles	Traffic light malfunction	Extra trips (15k)	Accidents	Closure
Optimistic with navigational assistance	A-I	Y	N	50 k	N	N	N	N
	A-II	Y	50%	50 k	N	N	N	N
	A-III	Y	Y	50 k	N	N	N	N
Optimistic with various demand	B-I	Y	Y	10 k	N	N	N	N
	B-II	Y	Y	25 k	N	N	N	N
Optimistic with Traffic light malfunction	C-I	Y	Y	50 k	Y	N	N	N
	C-II	Y	N	50 k	Y	N	N	N
Urgent (Extra) trips	D-I	Y	Y	50 k	N	Y	N	N
Vehicle accidents	E-I	Y	Y	50 k	N	Y	Y(1lane)	N
	E-II	Y	Y	50 k	N	Y	Y(2lanes)	N
Main route closures	F-I	N	Y	50 k	N	Y	Y(2lanes)	Y (4Arterials+Burmo)
	F-II	N	Y	50 k	Y	Y	Y(2lanes)	Y (4Arterials+Burmo)
	F-III	N	Y	50 k	N	Y	Y(2lanes)	Y (4Arterils+Motorway)

Figure 9.8: Summary of the simulation settings. Y refers to Yes and N refers to No. $4Arterial + Burmo$ refers to the closure of four arterial roads and Burma Road. $4Arterial + Motorway$ is referring to the closure of the same four arterial roads and Wellington Urban Motorway. The number of vehicles participating in the evacuation process is shown as $10k$, $25k$, and $50k$. 50% in the navigational assistance column demonstrates that only half of the vehicles are obtaining real-time routing information during their evacuation.

source (e.g. evacuees homes) to the evacuation points (i.e. Burma Road and Motorway).

In this case study, $50k$ vehicles participate in the evacuation process. $A - I$ models the circumstances when none of the vehicles are equipped with navigational assistance (also referred to GPS) during the evacuation process. $A - II$ demonstrates occasions when only half of the vehicles use their GPS for finding better routes through receiving regular routing information. $A - III$ demonstrates when all the vehicles receives regular routing information during the evacuation process. If a navigation system is enabled through simulation settings for an agent, then this agent receives regular routing updates (i.e. every 15 minutes) from routing servers and may reroute if Equation 9.1 holds (Subsection 9.2.1). When navigational assistance is disabled for an agent, then the best route is assigned to this agent before departure and will not change (i.e. do not receive regular routing updates).

Scenario $A - III$ is considered as the baseline for the rest of the case studies.

- **Case study B, optimistic with various demand**

  After forming the baseline scenario, $A - III$, this research can examine the impact of the various number of vehicles in an evacuation process through designing Case study B. Scenario $B - I$ and $B - II$ are designed based on $A - III$ and involve optimistic scenarios, where circumstances are similar to Scenario $A - III$ (no transportation infrastructure damage, advanced warning, no immediate threat to evacuees, no extra trip is required, no issue with telecommunication infrastructure) and all evacuees use their navigation systems for finding the optimal route during the evacuation process, and may reroute if Equation 9.1 holds.

  This case study applies a sensitivity analysis to examine the evacuation process by varying the number of vehicles participating in the evacuation process: $10k$ vehicles in Scenario $B - I$, and $25k$ vehicles in Scenario $B - II$. The results of this case study are compared with Scenario $A - III$, where the difference is only in the number of vehicles (i.e. $50k$ vehicles in $A - III$).

- **Case study C, optimistic with traffic light malfunction**

  In a real-world scenario, a power outage can occur and consequently some of the traffic lights may possibly not work. This happened to Wellington in 2018 [29] due to a substation fault. Power outage can happen due to a mechanical failure or damage from a disaster such as a hurricane or an earthquake. In the US, the weather and climate changes such as tropical cyclones, hurricanes, and floods are the most influential factors in power outage [152].

  This Case study, C, is designed to evaluate the impact of traffic light malfunction in an evacuation process. In the first scenario of this case study, $C - I$, all the settings are similar to $A - III$, but due to damage some traffic lights malfunction, as shown in Fig. 9.9 by a blue rectangle. In this blue rectangle drivers need to apply give-way rules at intersections with faulty traffic lights. The second scenario of this case study, $C - II$, models identical parameters to $C - I$ with the only difference that evacuees do not use

navigational assistance during the evacuation process. This is done to investigate the impact of navigational assistance when traffic lights do not function and the results are compared with $A - III$.
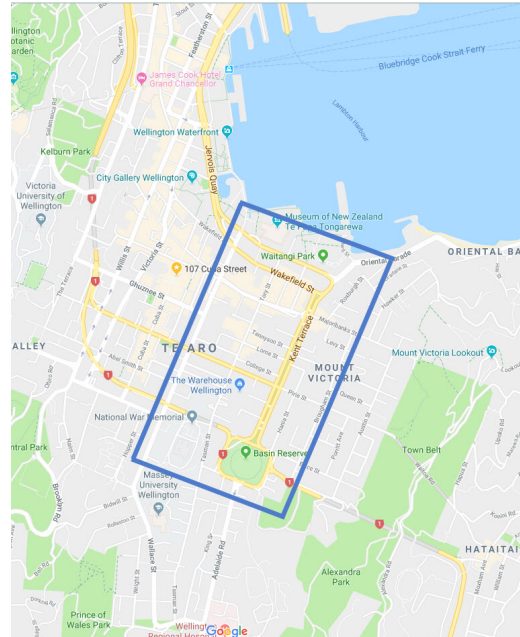
Figure 9.9: Traffic lights in the Blue rectangle are replaced with priority intersections.

- **Case study D, urgent trips**
  Moving away from the optimistic baseline scenario, $A - III$, and after examining various demand and impact of traffic lights malfunction, this research can now consider what happens if the evacuation occurs when not all members of a family are at home. In this Case study, D, there are similar conditions to Scenario $A - III$, but some people are required to make trips during the evacuation process to collect their family members (e.g. adults pick up their children from schools), prepare supplies (e.g. filling petrol tanks of their vehicles), or pick up vulnerable people. As suggested by Center for Preparedness and Response [17], fortunate people who have cars are suggested to take an extra step and assist those who have special needs (e.g.

old or sick people) or do not have cars in an evacuation process.

In this case study, $15k$ trips are made in Scenario $D - I$, and the results are compared with Scenario $A - III$.

- **Case study E, vehicle accidents**

  In order to increase the complexity of the previous scenarios, this case study is structured to evaluate the impact of evacuee mistakes/vehicle breakdowns/incidents in an evacuation process. This case study is built on top of Scenario $D - I$ to add more complexity to the settings of Scenario $D - I$. In this Case study, E, there are similar conditions to Case study $D - I$, but a few drivers make serious mistakes and cause an incident on Tory Street (has two lanes), the North direction. The location of this incident is shown by a red line in Fig. 9.10. Due to this incident, a road closure occurs and it takes about two hours for emergency response teams to clear this incident. This temporary road closure can be due to other reasons such as a vehicle breakdown, or running out of petrol. However, Scenario $E - I$ considers the case when only a single lane is closed and Scenario $E - II$ considers the case when two lanes are closed and the impact of the incident is greater. These results are also compared with Scenario $D - I$ and the baseline ($A - III$).

- **Case study F, main route closures**

  So far this research structured various scenarios for investigation of the impact of different settings such as demand, urgent trips, and traffic light malfunction. Now, this research can design a case study when evacuees have not received an advanced warning, therefore, during the evacuation process, some of the main parts of the road networks are damaged and unusable. This adds another level of complexity to the evacuation process and may create heavy pressure on the remaining, potentially fragmented, road networks. In addition to this direct damage due to a natural disaster to the road networks, a serious car accident during an evacuation process can cause a permanent (or very long-lasting) road closure in any road. For example, the spill of chemical materials on one of the Wellington highways in 2018,
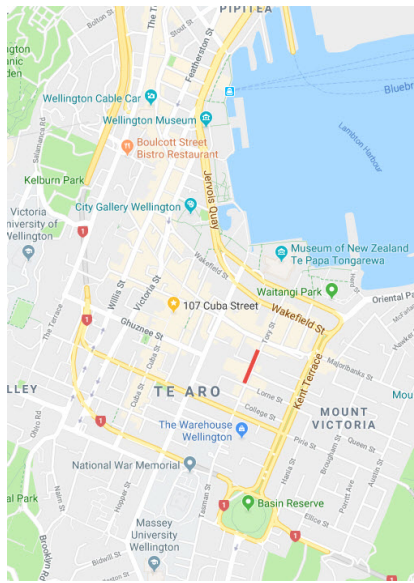
Figure 9.10: Location of road closures due to an incident, shown by a red line (Tory Street), North direction, Google Maps.

due to the overturn of the carrying truck, caused a 17-hour road closure on SH1 [148]. In this Case study, F, there are similar conditions to Case study $E-II$, but there is severe damage to four of the main arterial roads, which leads to a permanent road closure on these four roads. These four arterial road closures (shown in Fig. 9.11) are obtained from Chapter 7 (Fig. 7.25 and 7.22) as these four streets were highly impacted by the 2016 Kaikoura earthquake. Along with these four arterial road closures, this research models the closure of Burma Road and Motorway to investigate the impact of the closure of these two main evacuation routes on the evacuation process in Scenario $F-I$ to $F-III$. The locations of Motorway and Burma Road are shown in Fig. 9.7. Wellington Motorway can easily be damaged due to an earthquake, flood, or tsunami as it located close to the water, on the Wellington fault zone, and in the tsunami hazard area as shown in Chapter 2. Burma Road is also located in a high hazard zone as shown in Chapter 2. The closure of Burma Road and these four arterial roads are denoted by $4Arterials + Burmo$ and the closure of the motorway and these four arterial

roads are denoted by $4 Arterials + Motorway$ in Table 9.8.
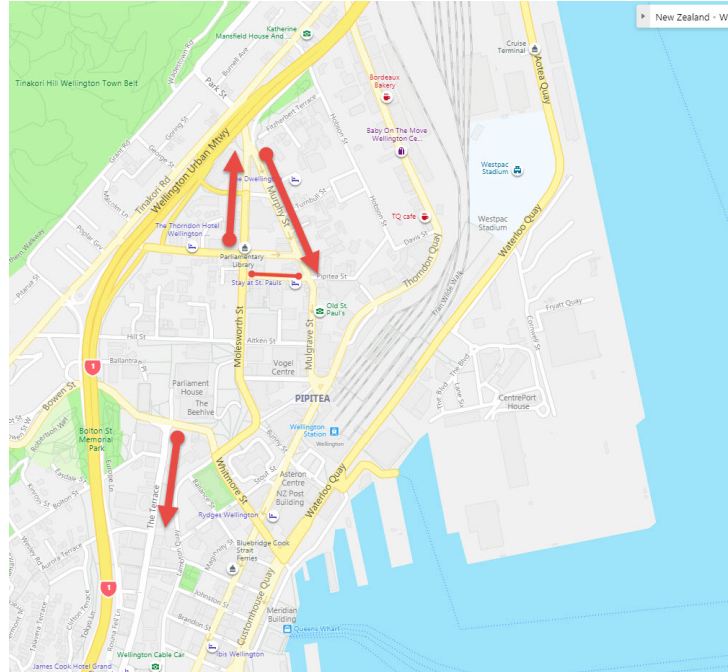
Figure 9.11: Location and direction of the four arterial road closure, Bing basemap.

Scenario $F - I$ and $F - II$ are similar and the only difference is that in Scenario $F - II$ some of the traffic lights malfunction (similar to Case study C). Scenario $F - III$ is similar to Scenario $F - I$ but instead of the permanent closure of Burma Road, Wellington Urban Motorway is permanently closed. In this case study, all vehicles receive regular routing information and may reroute. The results of this case study are compared with Scenarios $E - II$ and $A - III$.

9.3.2   Case study results and discussion

The output of each case study by using three types of graphs is shown in one figure: average speed of all the vehicles, average speed of each individual vehicle, and discharge rate. The average speed of all the vehicles shows the average speed of vehicles (km/h) for the whole simulation time (shown on the top section of the

results as a barplot and also as a boxplot). Boxplots demonstrate the distribution of the average speed of each individual vehicle in the simulation during the evacuation process. Discharge rate is the frequency of vehicles which arrive at the evacuation points in 15-minute intervals, and the last discharge rate (longest time to evacuation completion) used for evaluating the performance of the road networks (shown at the bottom section of the results) as the aim is evacuating the whole population. The plots which demonstrate the discharge rate have sharp increases and decreases. A similar trend was observed in previous research too [92],[185], [204]. These sharp edges happen due to the accumulating a large number of vehicles in queues (e.g. at intersections, ramps, traffic lights) and discharging them.

- **Case study A, optimistic with navigational assistance**

  As shown in Fig. 9.12, the scenario which does not utilise real-time routing information during an evacuation process (i.e. $A - I$) has the lowest performance (lowest average speed and highest discharge rate). The scenario which all cars are equipped with GPS (navigational assistance or navigation systems) and obtain real-time routing information about the condition of the roads (i.e. $A - III$) has the highest performance (highest average speed and lowest discharge rate), and the scenario that half of the vehicles utilise their navigation systems to obtain real-time routing information (i.e. $A - II$) sits in the middle. This is also shown by their boxplots through the position of the interquartile range. As shown by these boxplots, a few vehicles have higher average speed than the maximum speed limit (i.e. 100 km/h). These are vehicles which were closer to the evacuation points and did not experience congestion on their trips. Also these vehicles intentionally disregarded the speed limit as the collected DBQ (Driver Behaviour questionnaires, Function F3 in Table 9.1) shows that a small proportion of drivers disregard the speed limits. However, the average speed of most of the vehicles (interquartile range) is very low during the evacuation. Due to lack of access to the collected traffic data for the Motorway, this research uses the Google API to validate these results. As demonstrated in Fig. 9.13, the typical travel time for part of the Motorway with the length of 230 me-

ters is about 2 minutes. Based on this data, the estimated speed is about 7 km/h, which is close to the simulation results. Therefore, it is most likely that evacuees experience such a low speed during a large-scale evacuation process when the road networks are fragmented under heavy pressure.

- **Case study B, optimistic with various demand**
  Fig.9.14 shows the impact of various demand on the performance of the transportation network during an evacuation process. The scenario with the lowest demand of $10,000$ vehicles, $B-I$, has the highest average speed, and the fastest evacuation time. Boxplots also show that a greater proportion of vehicles in this scenario, $B-1$, have higher average speeds as the interquartile range is wider compared to the other two scenarios. This is because the congestion and utilization of the road networks in the scenario with $10k$ vehicles are the lowest compared to the other two scenarios. As demonstrated by these results, the scenario with $50k$ vehicles, $A-III$, has the worst performance and the scenario with $25k$ vehicles, $B-II$, is in the middle. Based on these results the evacuation time varies from about 11 hours (for $10k$ vehicles) to about 17 hours (for $50k$ vehicles) with the optimistic settings associated with these scenarios.

- **Case study C, optimistic with traffic light malfunction**
  As shown in Fig. 9.15, the scenario with power failure (i.e. traffic light malfunction) and no navigational assistance, $C-II$, has the lowest performance (the longest discharge rate and lowest average speed) among the rest of scenarios. Moreover, $C-II$ has a very long narrow tail. This tail belongs to the vehicles which passed through the area with a traffic light malfunction and applied give-way rules to other vehicles. Since none of the vehicles in the Scenario $C-II$ received regular routing information, they were not able to find a better route (i.e. reroute) during the evacuation process. The evacuation process was more efficient for Scenario $C-I$, where there was traffic light malfunction, but vehicles were able to receive regular routing information. Scenario $A-III$ has the highest performance as there was no traffic
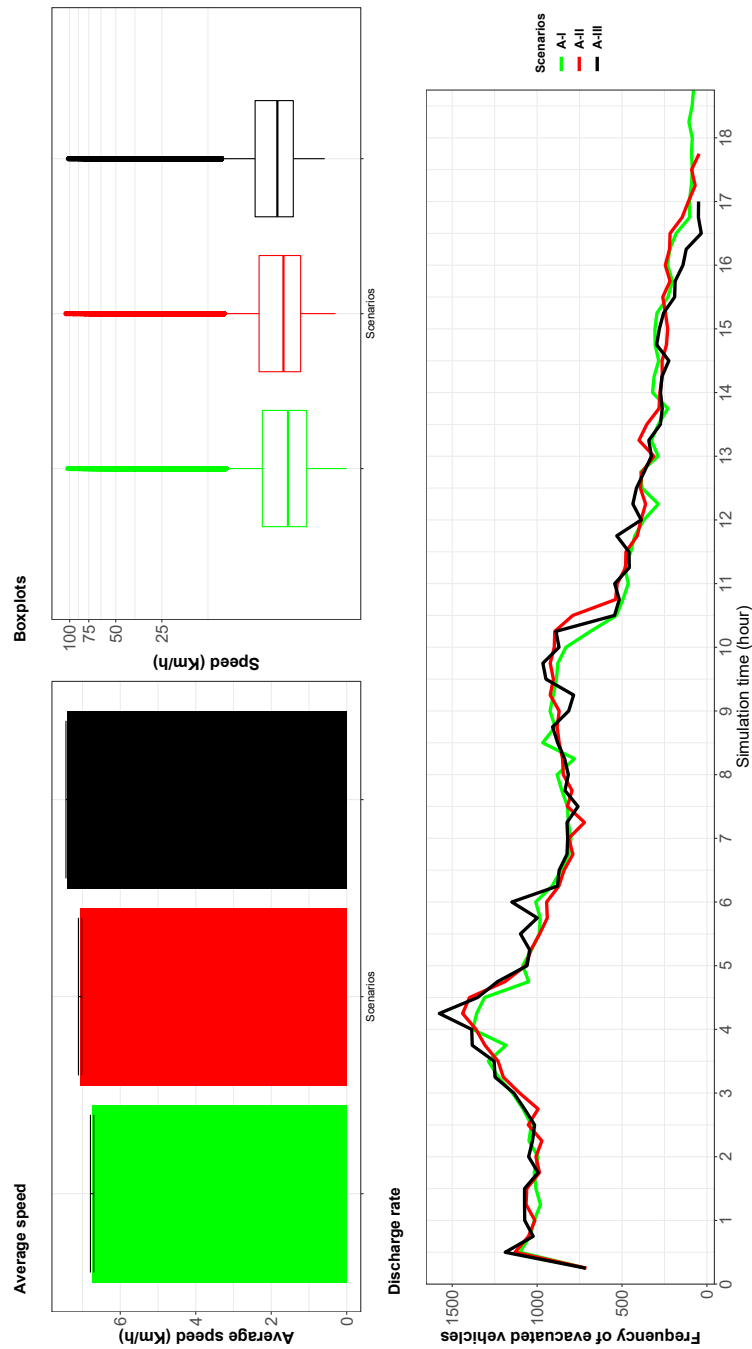
Figure 9.12: Network performance with a focus on utilisation of navigation systems during the evacuation process, Case study A. $A - III$, when all vehicles receive updated routing information, has the highest performance and $A - I$, when none of the vehicles receive routing information during the evacuation process, has the lowest performance. $A - II$ has a lower performance compared to $A - III$, but higher than $A - I$ as only half of the vehicles receive rerouting information.
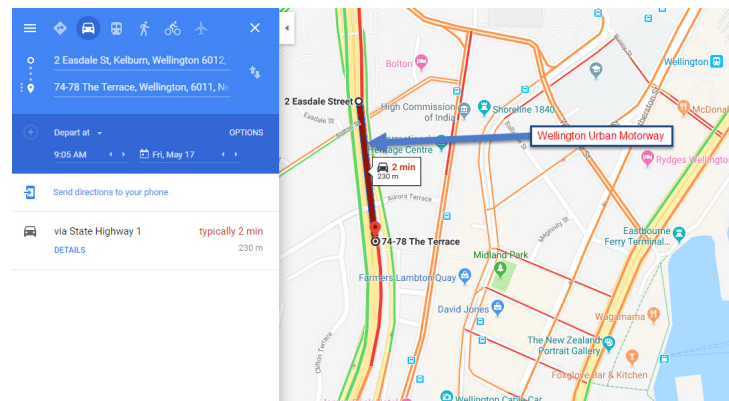
Figure 9.13: The average speed during the morning peak time is about 7 km/h for this part of Wellington Urban Motorway, North direction, Google Maps.

light malfunction and all the vehicles received updated routing information during their evacuation trips. These results highlight that for the settings in these scenarios, traffic light malfunction does not have a severe impact on the evacuation process, if vehicles can receive regular routing information (i.e. reroute), as they can find alternative routes.

- **Case study D, urgent trips**
  In this Case study, $15,000$ evacuees (agents), take one random trip each, before starting their journey to the safe zones (extraction points). As shown in Fig. 9.16, the evacuation time for the scenario with $15k$ trips, $D - I$, is longer than the scenario without extra trips, $A - III$, as vehicles are required to travel longer, but their average speed is not very far from the scenario without $15k$ extra trips, $A - III$. When this research added $15k$ vehicles to the scenario with $10k$ agents, $B - I$, ($25k$ has $15k$ agents more than $10k$) in the Scenario $B - II$, the average speed dropped from 8.99 (km/h) to 8.38 (km/h) (about %7), but when this research added $15k$ trips to the scenario with $50k$ (i.e. $A - III$) in Scenario $D - I$, the average speed dropped from 7.40 (km/h) to 7.31 (km/h), which is only about 1%. In the Case study B, the trips are from source to the evacuation points and in this case study trips are from source to random locations (non-evacuation points). This
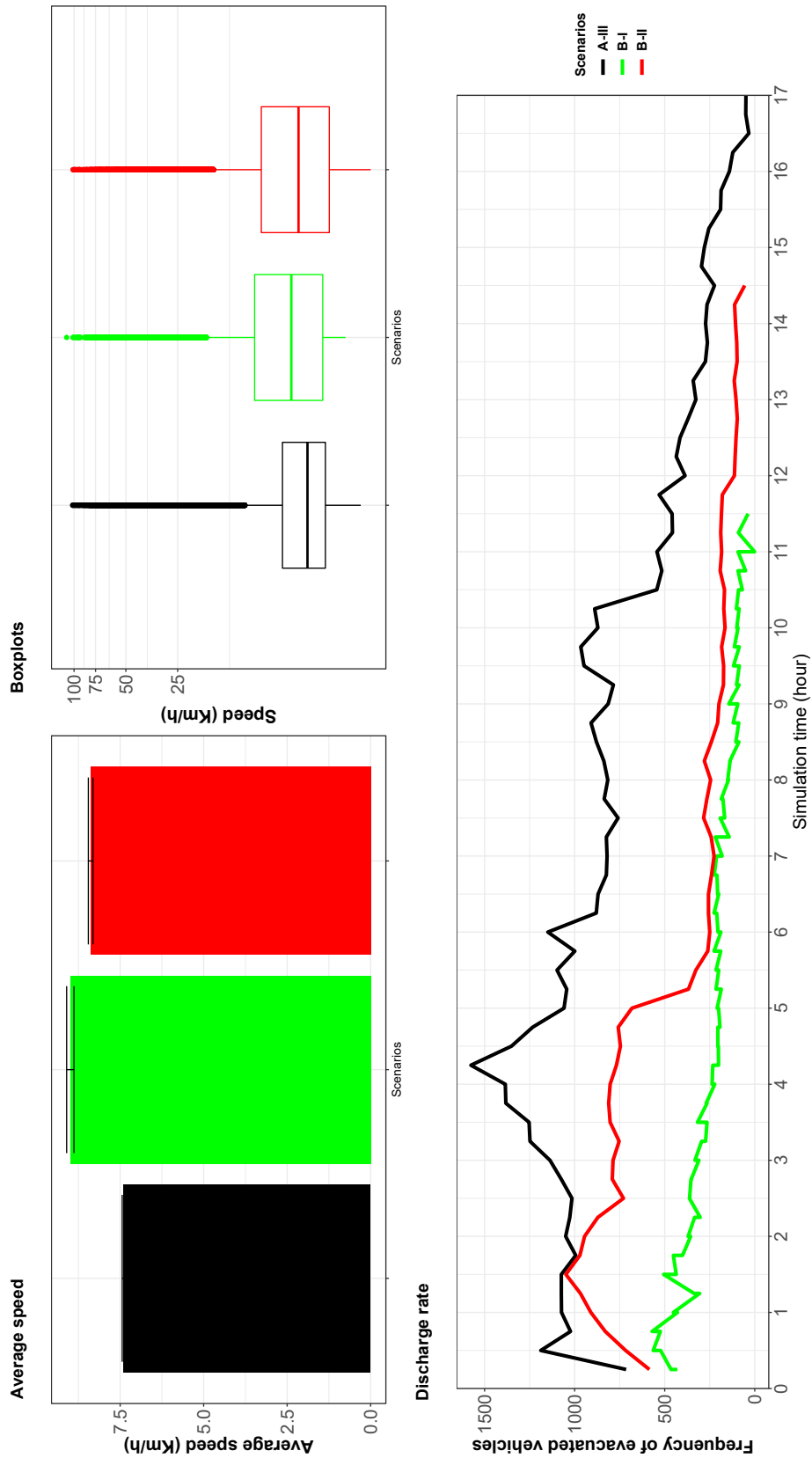
Figure 9.14: Network performance with a focus on various demand, Case study *B*. The optimistic evacuation time for 10*K* vehicles is about 11 hours and this increases to about 17 hours for 50*k* vehicles.

Figure 9.15: Network performance with a focus on power failure and consequently traffic lights malfunction, Case study C. The lack of long tail in Scenario $C - I$ compared to the Scenario $C - II$ demonstrates that using updated routing information can reduce the impact of traffic light malfunction.

difference in the speed degradation (%7 vs %1) highlights that these $15k$ random trips in this case study do not act as an additional $15,000$ trips to the evacuation points, and these extra trips distract the load from the main routes that direct the traffic to the exit points to other parts of the networks. Therefore, some agents drive faster on roads with lower congestion, and consequently, the average speed of all the vehicles does not decrease very much. These vehicles are shown in the tail of the boxplot of the scenario with $15k$ extra trips.

- **Case study E, vehicle accidents**

  An incident can be a vehicle breakdown or a car crash which both result in a closure of either a lane or multiple lanes. The results of the scenarios with an incident, which lead to the closure of one lane, $E - I$, or two lanes, $E - II$, of Tory Street are shown in 9.17. Tory Street has a high demand and it is a critical route during the evacuation process. In terms of impact on the evacuation process, the closure of two lanes made the evacuation about an hour slower and reduced the average speed of vehicles slightly.

- **Case study F, main routes closures**

  Among the scenarios which investigate the impact of road closure on arterial roads and motorway, $F - I$, $F - II$, and $F - II$, the closure of the motorway (i.e. Wellington Urban Motorway) has the highest impact (see Fig. 9.18) on the evacuation process, which is demonstrated by the lowest average speed (also the lowest boxplot interquartile range), and the poorest discharge rate. This is because the motorway has the highest capacity, and the least traffic flow disruption (e.g. intersections, pedestrian crossing).

  The performance of road networks due to the closure of the arterial roads and Burma Road was low, but not as low as the performance of road networks in the case of the closure of the motorway in Scenario $F - III$ and the traffic light malfunction in Scenario $F - II$. In addition, the performance of the road networks with the closure of Burma Road and the four arterial roads was lower than the scenario without any permanent road closure (i.e.
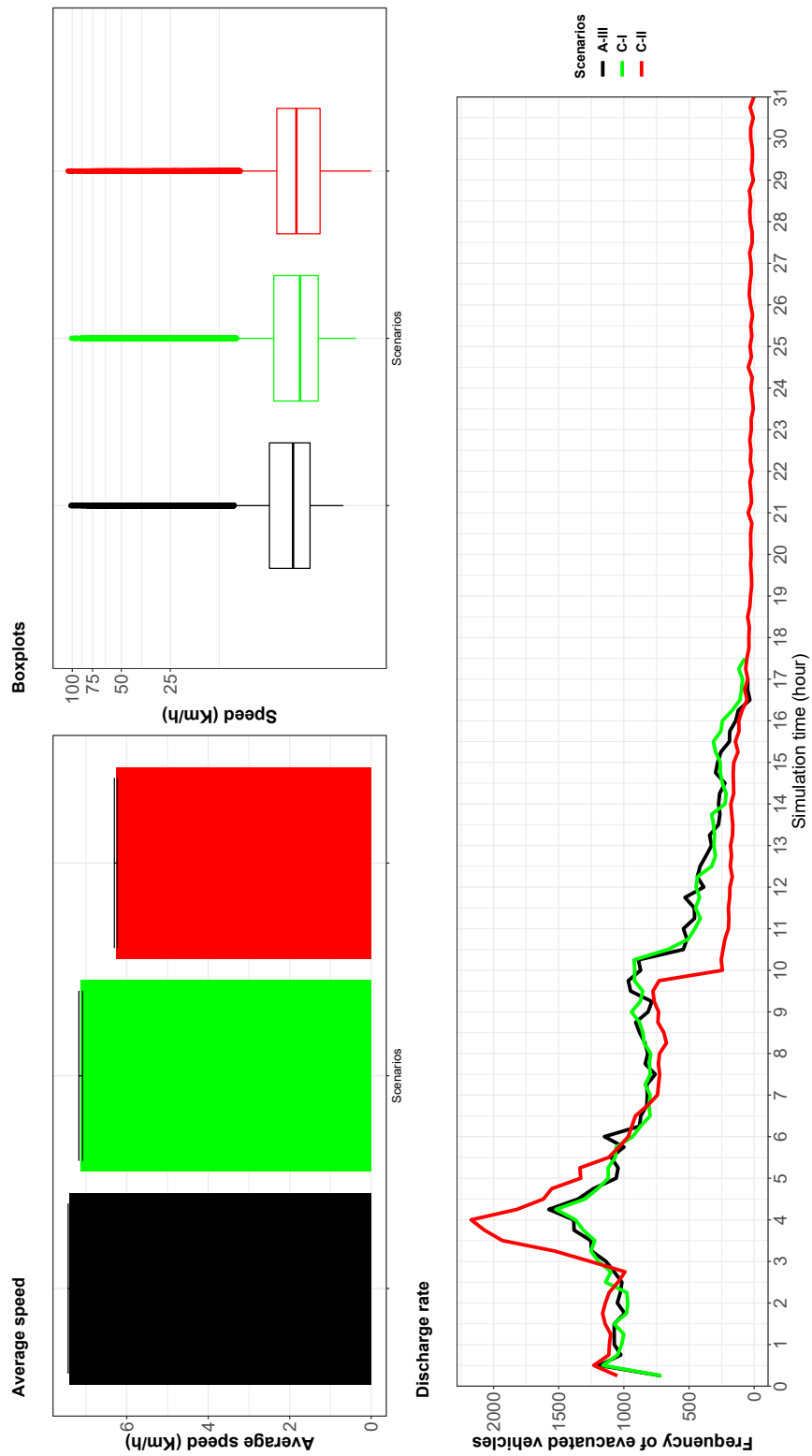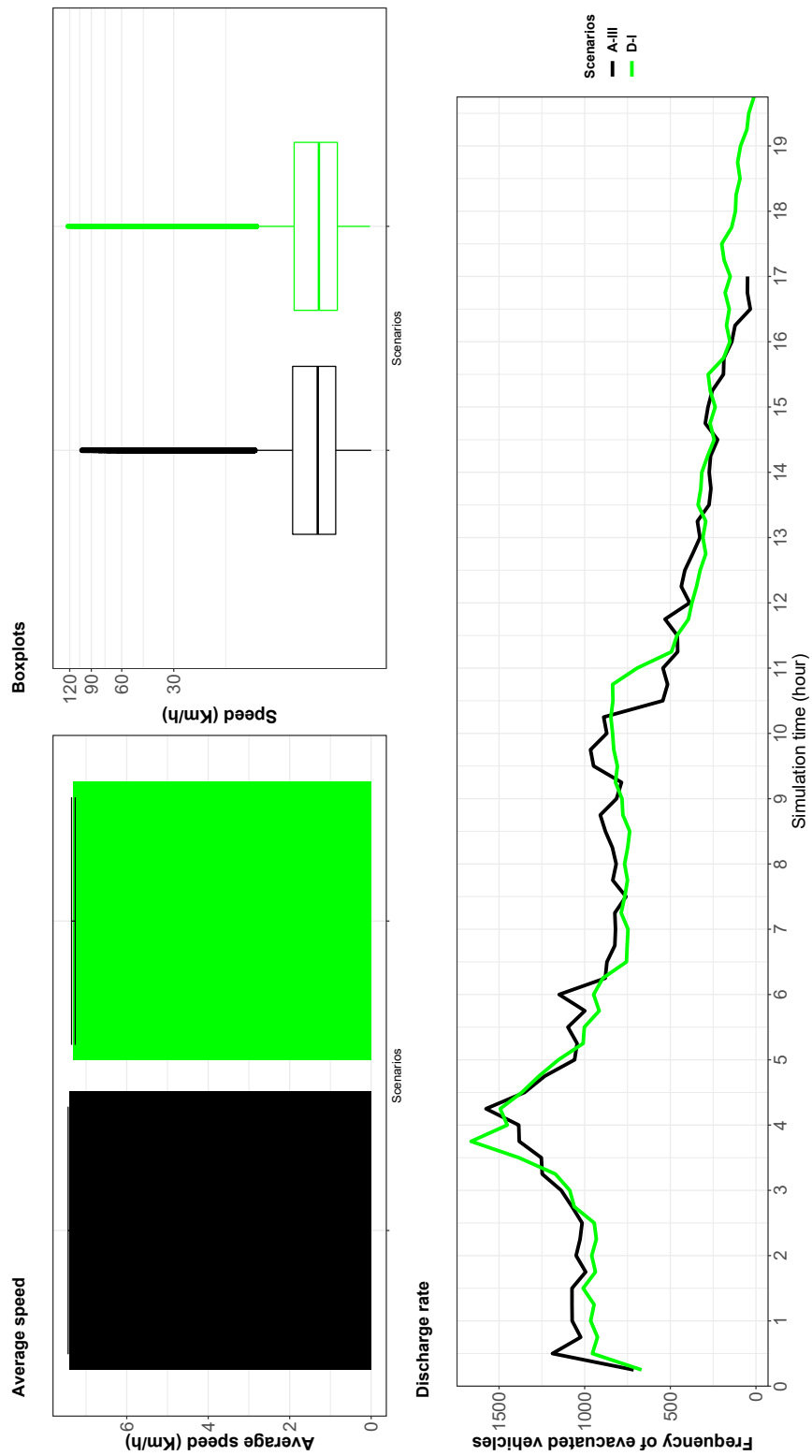
Figure 9.16: Network performance with a focus on the divergence of demand, Case study *D*. These 15*k* extra trips increased the evacuation time but did not reduce the average speed of vehicles during the evacuation process. These 15*k* extra trips acted like a load balancer and distributed the load to the rest of the network.
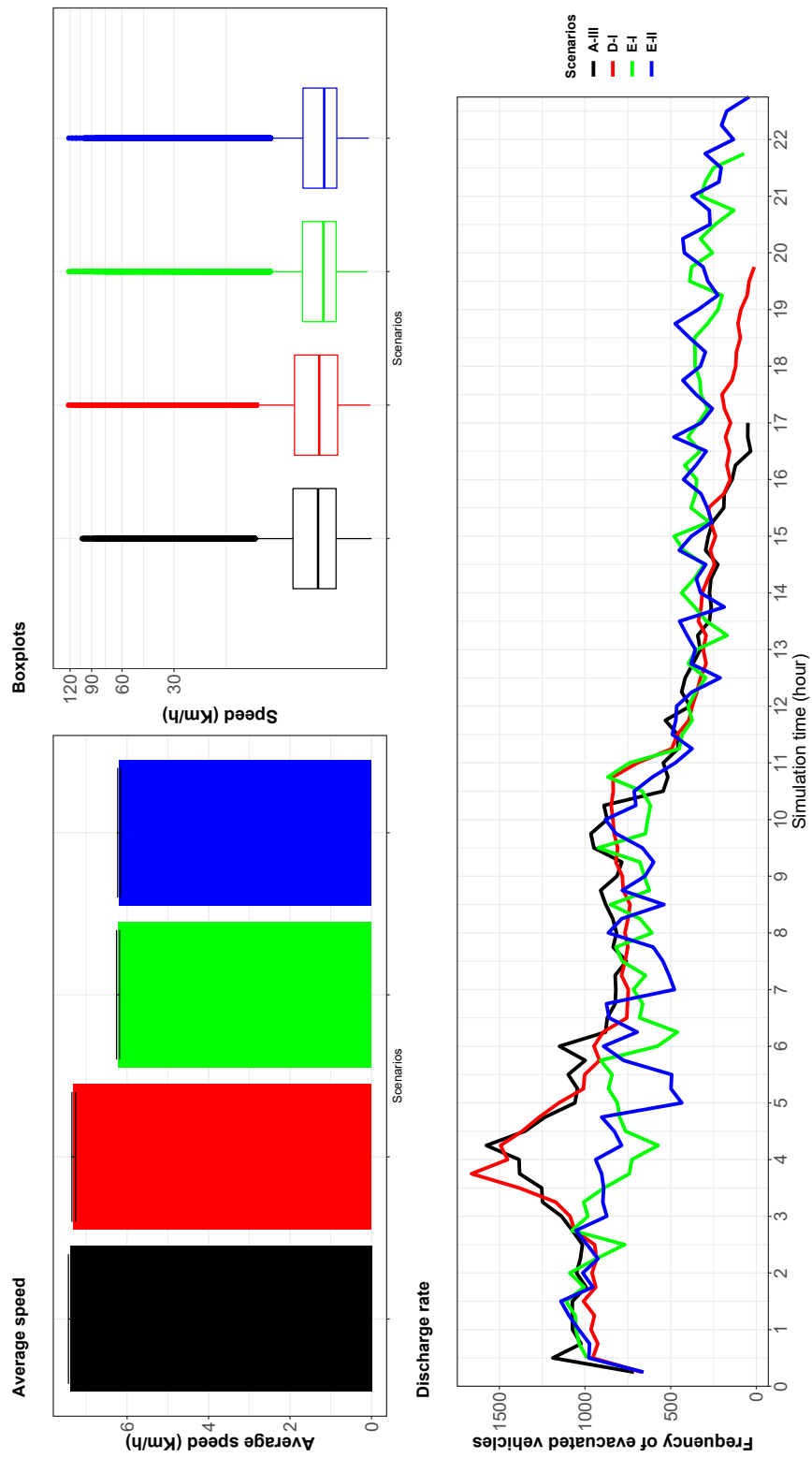
Figure 9.17: Network performance with a focus on human errors, mechanical failure, and vehicle incidents, Case study E. Case Study E has a lower performance compared to Scenario $D - I$ and $A - III$.

infrastructure damage) due to a disaster, $E - II$.

In this Scenario, $E - II$, 6035 vehicles used Burma Road and $43,965$ vehicles used the motorway (Wellington Urban Motorway) to exit the city. This shows that the demand for the motorway is about seven times greater than Burma Road. It is noticeable that Wellington Urban Motorway carries more traffic as it is in the center of the city and continues to the outer parts of Wellington. In addition, the motorway has entrances all over the city which direct traffic to the motorway, while Burma Road is located on the outer parts of Wellington and only carries a small proportion ($1/7$ based on the results of $E - II$) of traffic to the outside of the city as shown in Fig. 9.19. This is also shown by boxplots in Fig. 9.18. The interquartile range of the scenario with the motorway closure has the lowest range. The scenario with the closure of Burma Road has a lower average speed and a higher discharge rate compared to Scenario $E - II$, but it is not as low as the scenario with the closure of the motorway ($F - III$). $A - III$ has the best performance in this set as it is the optimistic scenario. Moreover, the traffic light malfunction, Scenario $F - II$ did not have a huge impact on the evacuation process compared to the performance of road networks in Scenario $F - I$, as evacuees can find a better route through utilising their navigation systems. Therefore, it can be concluded that road closures on the main routes (critical routes), which direct agents outside the city, have a higher impact on the traffic flow compared to other routes (non-critical routes). Thus, this is the confirmation that placing incident response teams on critical routes must have a high priority during an evacuation process compared to non-critical routes.

9.3.3 Road performance

Although the earlier results measure the performance of road networks indirectly through the agents, it is also interesting to investigate how an individual road performs during the evacuation through measuring the occupancy. Occupancy, which is part of transportation engineering discipline [88],[86], will help to identify crit-
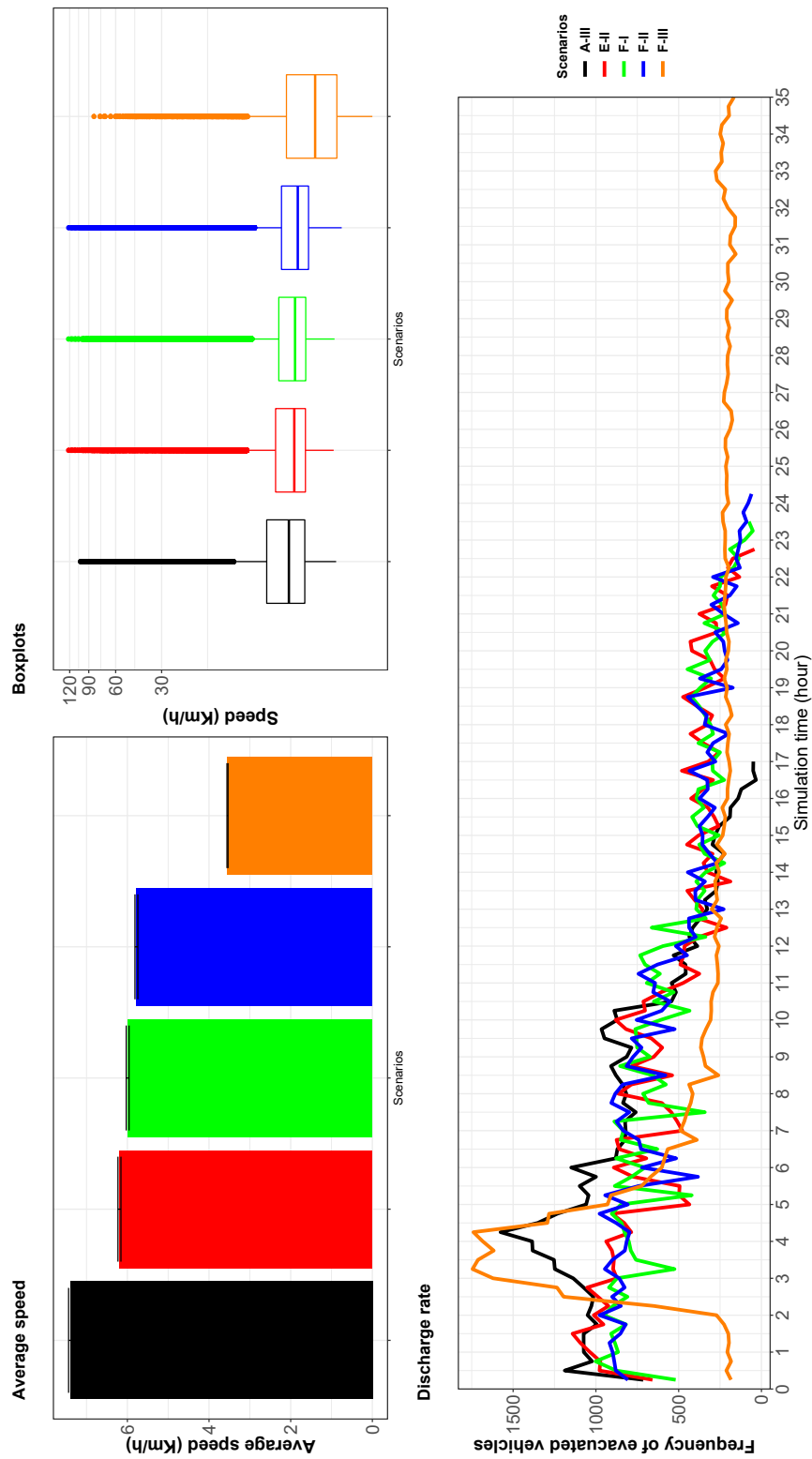
Figure 9.18: Network performance with a focus on damage to some of the arterial routes and one of the main exit points (Burma Road or Wellington Urban Motorway), Case study F. Damage to the motorway has the highest impact on the evacuation process and make the evacuation almost impossible (i.e. about 35 hours), Scenario $F - III$. Rerouting (i.e. receiving regular routing updates) reduced the impact of traffic lights malfunctions ($F - II$).
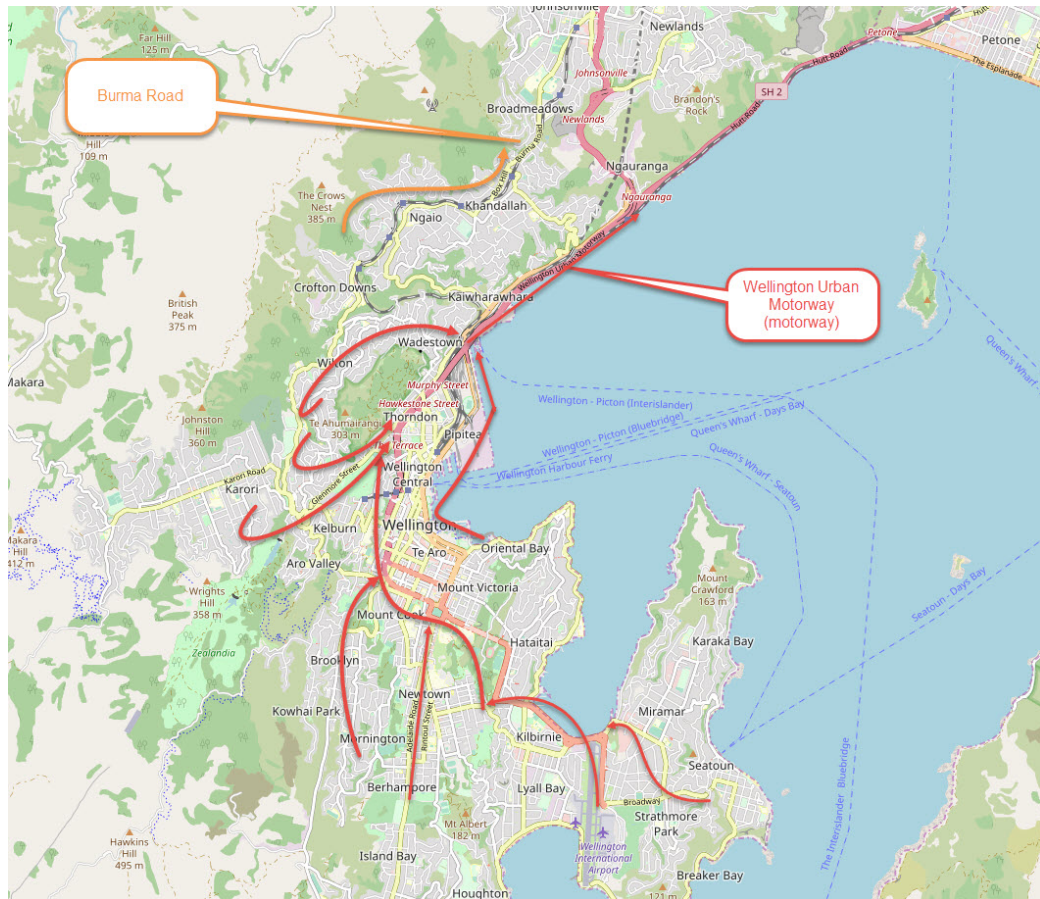
Figure 9.19: Routes which direct traffics to Burma Road and Motorway. Motorway carries 7 times more vehicles compared to Burma Road, Open Street Maps (OSM).

ical routes, which have the highest impact in the evacuation process.

Since the critical routes are the same in all scenarios and collecting and processing this information is a heavy and time-consuming task and also requires a huge amount of storage space, this research applied this analysis for one of the scenarios, $A - III$, as it was the baseline scenario and the rest of the scenarios were built on top of this one. To collect data, a loop detector is installed in every single lane of an edge in the simulation environment. A street may be composed of several edges, so it may have multiple loop detectors. Therefore, the relation between a street and loop detectors is a $1 \rightarrow n$ relation. Due to this fact, an edge ID is attached to the name of each street to make the result more accurate and meaningful. Otherwise, if a street has multiple edges or multiple directions, then the results are combined and provide an inaccurate conclusion. However, this research used occupancy and volume of the edges as two common metrics to investigate the performance of each edge individually. Volume is an important factor as it shows how traffic flows, but since it also depends on the number of lanes of a street, it is not enough to investigate the performance of road networks. On the other hand, occupancy shows the ratio of the time that a loop detector is occupied by a vehicle standing (or passing) over it. In this regard, ratio one, or 100%, shows that a vehicle was stopped on top of the loop detector in the whole reading period and zero, or 0%, means that no car was passing. The average of the occupancy during the whole simulation time is shown on top, and the total flow for the whole simulation time is shown at the bottom of Fig. 9.20. Since there are thousands of edges in the simulation, this figure only shows a proportion of street names (top n edges) sorted in descending order based on the attribute of interest (either volume or occupancy). As shown in this figure, the motorway edges (Wellington Urban Motorway) have the highest traffic flow, but not the highest average occupancy. It can be concluded that many vehicles used the highway, but it was not very congested and was not fully utilised.

In order to make this information easier to understand, a module (in .Net and Python) is developed in this framework to extract the latitudes and longitudes of edges (stored in SUMO and Open Street Maps (OSM)). Then after processing,

this module forms a spatial line to model a street by concatenating the coordinates of edges. Then, a GIS tool (QGIS) is used to visualise the occupancy and traffic volume. The outputs are shown in Fig. 9.21 and 9.22. The former figure is the average occupancy and the latter is the total flow (volume) for the whole simulation time. The aggregation of traffic flow toward the motorway (distribution of warm colours) agree with Fig. 9.19. In addition, Fig. 9.22 shows that there was a high demand on some parts of the network, although the average of their occupancies show that some of these streets still have enough capacity as their occupancies are still low. For example, Wellington Urban Motorway (referred to motorway in this research) is shown by red in Fig.9.22, but as green in Fig. 9.21. This accentuates that these streets are critical routes but still have enough capacity, while the edges which are shown in Fig. 9.21 by warm colour are required to prioritised for future development to accelerate traffic flow (e.g. widening, dynamic road signs, traffic direction, and traffic lights).

Fig. 9.23, shows the occupancy of the road networks at the simulation time of 1000 second as a screenshot. The colour varies from green to red to represent the variation of occupancy from zero to one. Another screenshot with similar colouring schema and a lower zoom is provided in Fig. 9.24 from the simulation time of 7000 second. The roads which are shown by red are heavily congested routes. In order to improve the performance of road networks during an evacuation process, local authorities may change the direction of other streets which are parallel to these streets, which are shown by warm colours, to increase the capacity of the road networks toward the evacuation zones. As an example, Adelaide Road which is located below the Basin Reserve, is a two-way road and is shown by a warm colour in Fig. 9.21. All the lanes of this road can be changed to make this road one-way (from South to North) to increase the capacity and increase the traffic flow in the case of an evacuation. In addition to the evacuees that should be informed of such changes in advanced, navigation systems also must be updated to reflect these changes in the network topology.

Moreover, Business Intelligence (BI) is used to develop a dashboard for showing the central location of each edge (i.e. middle of each line) and colour-coding

the occupancy (similar to time series). A screenshot is shown in Fig. 9.25, which shows the occupancy at three individual simulation periods for every central point of each edge. X axis shows the latitude and Y shows the longitude of each point and the colour represents average occupancy per interval (30 minutes). As demonstrated in this figure, at the beginning of the simulation (left) most roads are very busy but over time the peripheral roads get less congested, but the main routes which direct the traffic toward the exit points remain congested (right). This developed dashboard in Power BI can be connected to the database to visualise the collected data which come from installed loop detectors in real-time to inform authorities about the condition of the road networks. This provided data can be obtained in real-time through various ways such as sensors installed on the roads, from the inductive loop detectors installed in the simulation environment, or the combination of these two, when the real-world is connected to the simulation environment (Chapter 8).

9.4   Conclusion

In this chapter, I presented a novel framework for traffic modelling and use it to measure the impact of different parameters and assumptions on the performance of road networks in transportation modelling by using research from different disciplines such as transportation engineering, psychology, programming, and GIS. As demonstrated, the results can vary by changing assumptions and criteria, which signify the impact of a more realistic approach for traffic modelling.

The utilisation of navigation technologies by agents during a large-scale traffic evacuation can decrease the performance of the road networks if all the individuals follow the routes selected by their GPS (i.e. navigation system). To overcome this issue, I introduced an enhancement to the navigation systems for optimising rerouting criteria, which improved the performance of road networks, although more research is required to identify a correlation between the number of agents, road capacity, and rerouting criteria. It must be noted that the most popular navigation systems are commercial and the lack of access to their source codes restricts
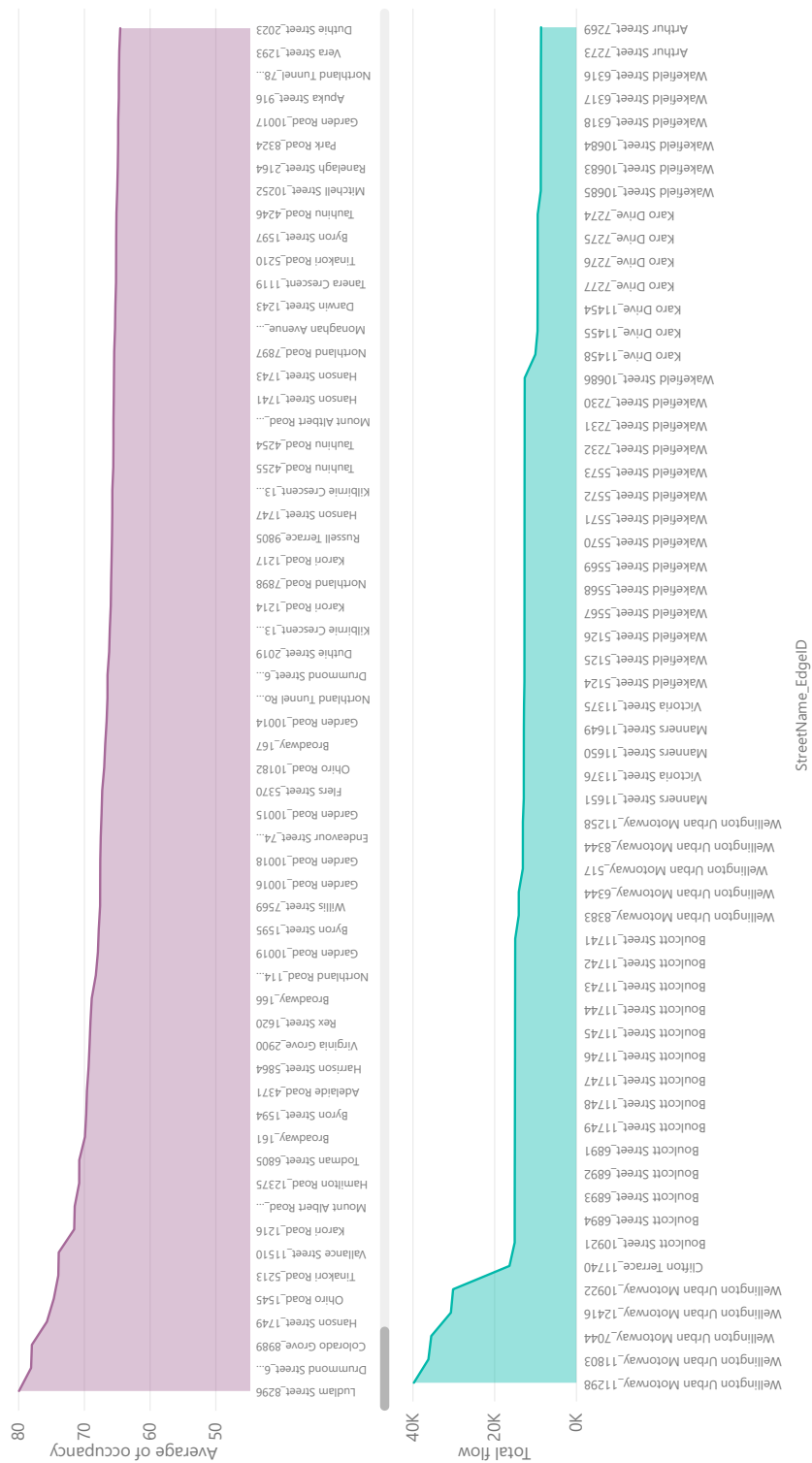
Figure 9.20: Average occupancy (top), total flow (bottom) for top n rows sorted by the attribute of interest (average occupancy or total flow) in a descending order.
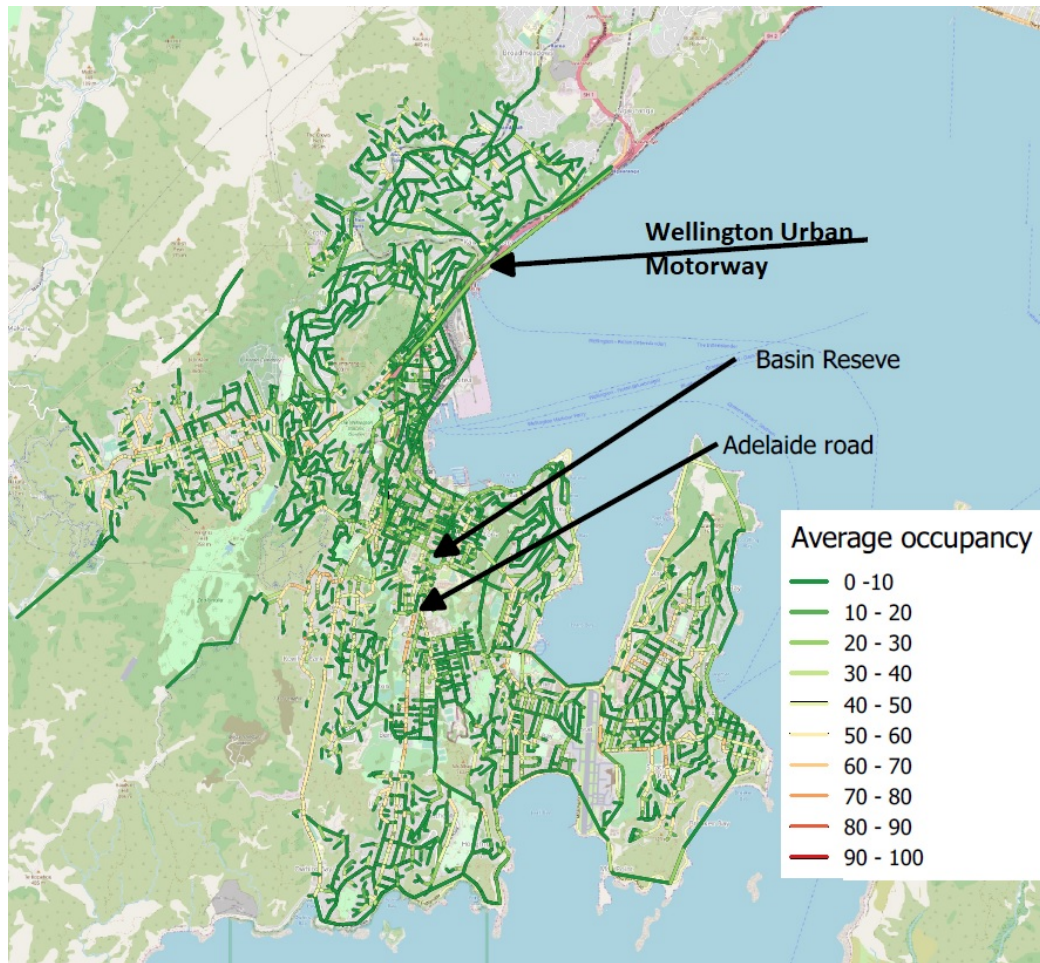
Figure 9.21: Average of the occupancy, spatial presentation, simulation environment (SUMO), Open Street Maps (OSM).
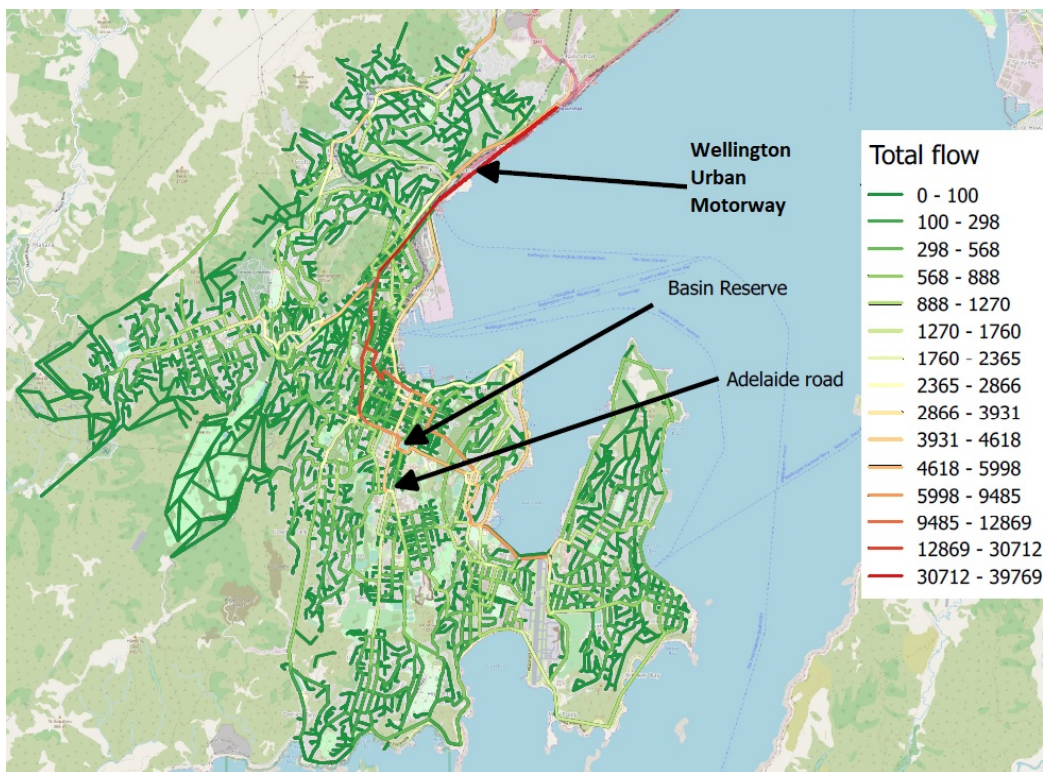
Figure 9.22: Total flow, spatial presentation, simulation environment (SUMO), Open Street Maps (OSM).
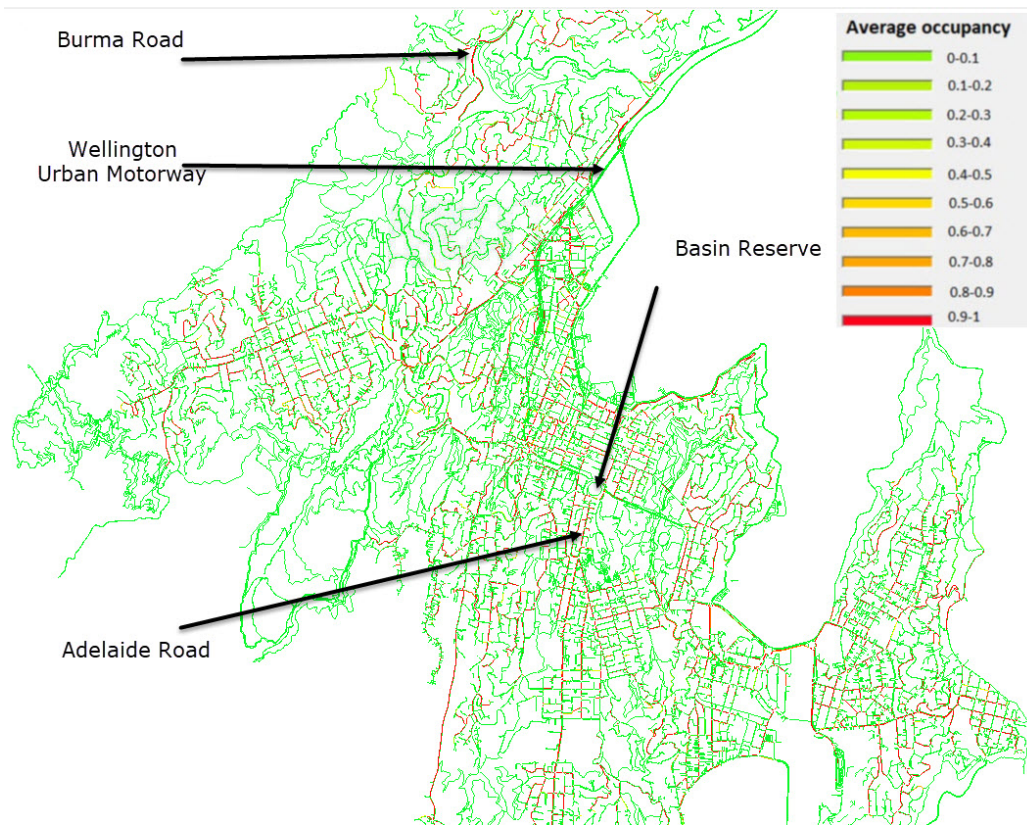
Figure 9.23: Occupancy, a screenshot of the simulation environment at time 1000 second.
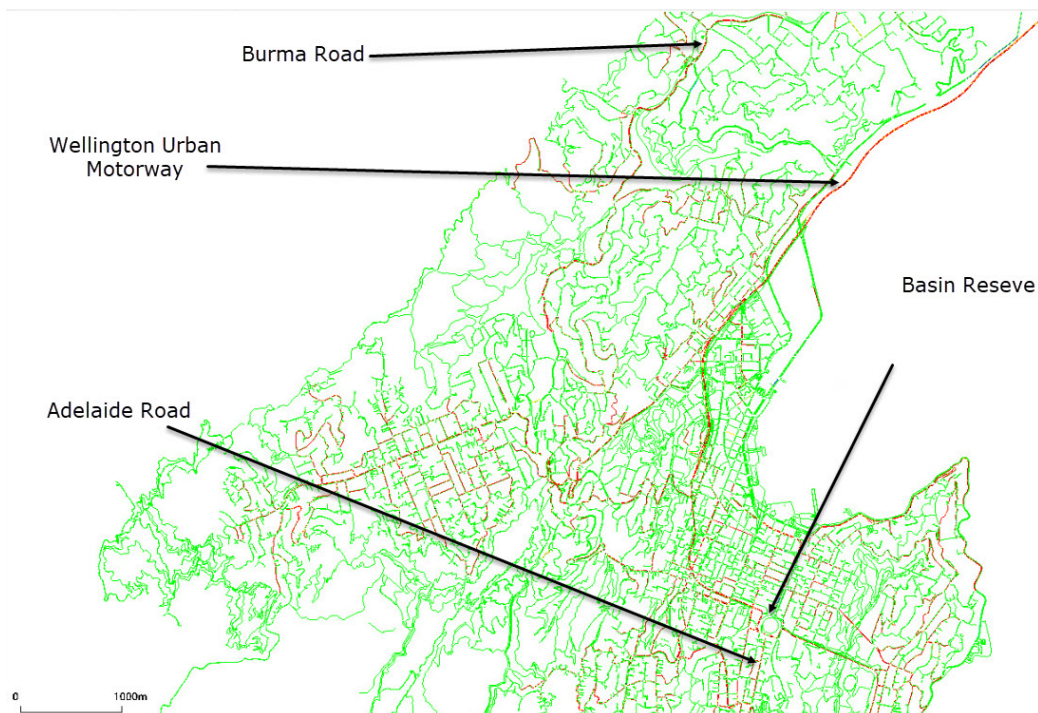
Figure 9.24: Occupancy, an screenshot of the simulation environment at time 7000 second.
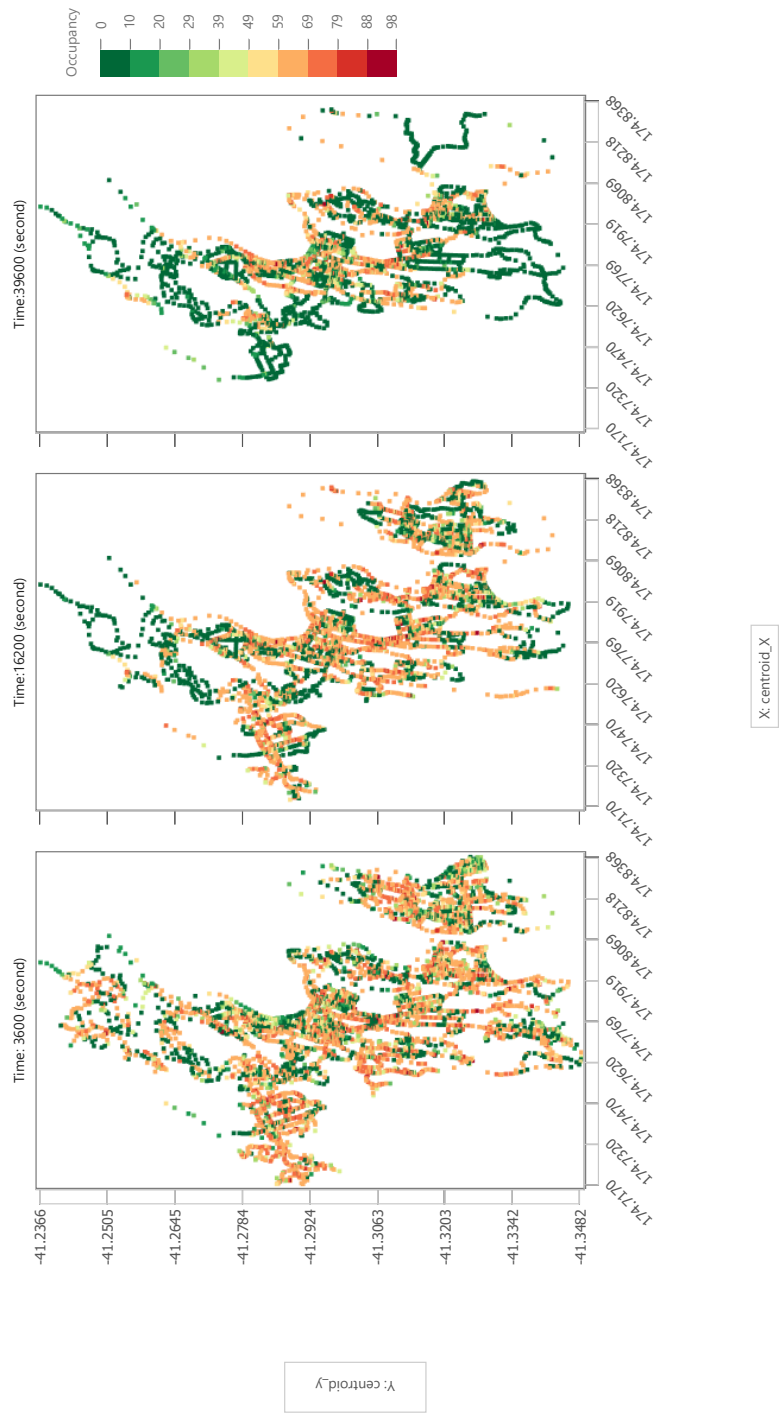
Figure 9.25: Demonstration of the network condition (occupancy) during different simulation times in a BI dashboard.

researchers from evaluating their performance under different criteria.

In addition, this research also measured the evacuation time of Wellington City under various principles, realistic assumptions and scenarios. Results show that the optimistic evacuation time which is required for evacuating $10,000$ vehicles from Wellington is at about 11 hours. This optimistic result is for the case when there is no damage to the infrastructure and no interruption in the traffic flow which can hardly be assumed in an actual disaster. This highlights that this time can be increased significantly under different criteria and authorities in Wellington are required to make serious efforts for planning large-scale evacuations and consider other approaches to enhance the traffic flow during such occasions when the demand is very high and there is the possibility of damage to the infrastructure. This can be accomplished through educating evacuees, traffic operation managers and emergency managers, finding optimal places for the establishment of emergency teams, a collaboration between multiple organisations (e.g. police, fire departments, medical teams), and on-demand route assignments. As another solution for enhancing a large-scale evacuation, some smartphone applications can be developed to form a connected network through Bluetooth technologies or GPRS (4G) to guide evacuees through the evacuation process and assigning them optimal routes. This approach can be invaluable as each smartphone can record and share its location with others. Therefore, it is possible to have an independent and decentralised traffic management system that can manage the traffic in case of damage to the main infrastructure (e.g. communication, transportation) of a city.

Chapter 10

Conclusions and recommendations

10.1 Conclusions

Smartphones and digital portable devices along with their low prices and high popularity, and have become an inseparable part of the routine life of people. It is rare to find a person who is commuting to work in a typical urban area without carrying a smart portable digital device (e.g. tablet, smartphone, headphones), although due to various reasons (e.g. security) the detection rate is low. Moreover, most new vehicles and navigation devices (commonly known as GPS devices) can be interconnected with other digital devices of occupants of a vehicle primarily via Bluetooth technologies. These elements provide an interesting newer emergent avenue for research. This comes with some challenges such as identification of the correct mode of transportation and estimation of the characteristics of all the passing vehicles (i.e. true dataset) from a small portion of detected vehicles (i.e. observed dataset).

This thesis examined Bluetooth data collected from Wellington Bluetooth Traffic Monitoring Systems (BTMS) to identify vehicles that follow regular traffic patterns, estimate the true traffic characteristics from the observed dataset, and

investigate the impact of the 2016 Kaikoura earthquake on the Wellington traffic flow. Moreover, I developed a novel multi-disciplinary evacuation framework for modelling a large-scale evacuation by bringing traffic and computer engineering, statistics, urban planning, and behavioural modelling under one umbrella. This framework assists traffic teams and local authorities in preparing more efficient strategic evacuation plans.

In Chapter 2, I analysed various existing Wellington geospatial layers. This analysis demonstrated that Wellington is located on several fault zones and most of its main transportation infrastructures such as roads networks (land), ferry terminals (sea), and airport terminal are located on extreme hazardous zones, which make these infrastructures highly susceptible to damage during a natural disaster, especially seismic activities such as an earthquake. Also the majority of high buildings are located in CBD, which also is located in a high hazard zone. These findings suggest that the closures of streets within the CBD due to the collapse of buildings during a seismic activity are very likely.

In Chapter 3, I investigated various types of nodes that can be detected by examining BTMS data. This chapter examined various types of mobile nodes that should be excluded from collected Bluetooth data in order to only leave vehicles which follow regular traffic patterns. Moreover, in this chapter, I identified some possible limitations of utilising BTMS in traffic monitoring in harmony with other researchers.

In Chapter 4, I examined this collected data and proposed an approach for data cleaning and identifying trips of vehicles which follow regular traffic patterns. This proposed cleaning algorithm, ExtoVT, does not require a large set of historical data for calibration which makes this algorithm fast and efficient. This algorithm requires to obtain some information about conditions of a route and then combines this information with other computed parameters and Tukey's logic to exclude everything except vehicles which follow regular traffic patterns. Moreover, I identified an almost neglected issue associated with the collected Bluetooth data, multi-tenancy detection, where a vehicle carries multiple detectable Bluetooth devices (multiple tenants in a vehicle). Therefore, a vehicle with sev-

eral devices, may be identified as several vehicles with one Bluetooth device each. If multi-tenant vehicles are unidentified, then this causes inaccurate estimation of the traffic characteristics (e.g. overestimation of the traffic flow). Hence, I proposed an algorithm, MTDiBT, for identification of multi-tenant vehicles, as part of the cleaning algorithm (ExtoVT). This cleaning algorithm has a flexible schema, which makes it adjustable to various traffic conditions in various urban areas.

In Chapter 5, I modelled the behaviours of BTMS in terms of detection of a proportion of passing vehicles through utilising statistical inference and a Monte Carlo Simulation for examining inter-arrival times (IAT). This is because BMTS data identify a proportion of passing vehicles which carry detectable Bluetooth devices, the traffic characteristic (e.g. variance, inter-arrival times) of the true dataset may be different from the observed traffic characteristics. My finding demonstrated that when the detectability ratio is very low (as may be the case in real life), the identified distribution of the observed inter-arrival times (i.e. detected by BTMS) is not necessarily the same as the distribution of the true inter-arrival times (i.e. all passing vehicles). In addition, my results confirmed that if the true inter-arrival times has an Exponential distribution, then the observed dataset will also have an Exponential distribution, and this may not be the case for other distributions. However, a high proportion of the observed inter-arrival times were modelled by Exponential distribution. This finding is broadly in line with the detection of an Exponential distribution for Inter-arrival times modelling by other researchers. Additionally, the finding of this chapter demonstrated that the statistical characteristics of a true dataset can be estimated from an observed dataset if the detection ratio (i.e. calibration ratio) can be determined from other resources such as inductive loops.

After identifying the regular traffic flow, and the analysis and modelling of partial detection of traffic flow by BTMS and the need for determining the detectability ratio in Chapter 5 to estimate the traffic characteristics of the true dataset, I proposed an approach in Chapter 6 to estimate the detection ratio by examining the traffic data collected from other traffic resources (e.g. inductive loops). Moreover, after examining the Bluetooth data and the lessons learnt in the previous chapters

about the key elements of traffic components that may impact a Bluetooth traffic monitoring system in terms of processing and modelling, I developed a digital twin that utilises BTMS data and other traffic data to replicate the physical world in a virtual environment.

Then, I utilised these methodologies to identify the impact of the 2016 Kaikoura earthquake on the traffic flow. I propose several methods in Chapter 7 such as investigation of micro and macro routes to highlight the impacted routes due to this earthquake. In addition to providing quantitative metrics, I use various geospatial tools and Business Intelligence (BI) dashboards to visualise the impact of this earthquake on the traffic flow. These routes mainly were located in the hazardous area in CBD, which is generally compatible with the findings of Chapter 2, where the collapse of a tall building may cause road closures.

In terms of modelling of a large-scale evacuation, I have done an extensive literature review in Chapter 8 across several disciplines: psychology, Geographic Information Systems (GIS), traffic modelling, and traffic simulation to identify various components which impact the performance of an evacuation model. Some of these components were studied by other researchers in evacuation modelling, and some have not received enough attention such as the behavioural model, but the combination is novel and valuable in traffic modelling and particularly disaster modelling.

After examining various critical components that can affect the performance of a model for a large-scale evacuation due to the disaster discussed in Chapter 8, I proposed a multi-disciplinary agent-based framework (TFDA2M). TFDA2M brings various key components such as road networks, road closures, behavioural model, GPS technology, and simulation environment together to model a large-evacuation more accurately. I also used TFDA2M to model several disaster scenarios and demonstrated how some realistic assumptions and criteria can impact the performance of the road networks and the evacuation process.

Aligned with the study of utilisation of navigation systems in a large-scale evacuation process as part of disaster modelling in Chapter 8, I investigated the effect of dynamic rerouting during the evacuation process (i.e. rerouting) in a large-

scale evacuation process and identified new challenges, which already identified in computer networks. These challenges are associated with rapid fluctuations of route costs and formation of routing loops due to utilising the road networks with the maximum capacity during a large-scale evacuation process. Since disaster days are rare and categorising an event as a disastrous event is challenging and requires the collaboration of various teams and authorities, therefore it is highly possible that the commercial routing services (e.g. Google, Navman) for transportation (off-line and on-line GPS) do not differentiate typical days from rare disaster days. Therefore the current version of these technologies may misguide drivers during a large-scale evacuation.

Moreover, in this chapter, I visualised the performance of transportation infrastructures during a large-scale evacuation through a set of Business Intelligence dashboards and GIS tools. These visualisations assist local authorities and traffic operation teams in identifying parts of the transportation networks that exhibit poor performance without installing massive traffic monitoring systems in an urban area. This installation may be impossible due to cost, time and extent of a city and not be the true reflection of the transportation infrastructures in a disaster day. The performance of the road networks of Wellington for a large-scale evacuation was poor, and the results highlight an urgent need for more detailed large-scale evacuation planning. Wellington does not have a fault-tolerance city profile in the context of a large-scale evacuation (as demonstrated by Chapter 2, there are only two land exit routes for this city and all of the main transportation infrastructure are located in high-risk areas). I also proposed a few approaches for increasing the performance of existing transportation infrastructure during an evacuation process for Wellington City. TFDA2M also can be used for other purposes such as maximisation of the utilisation of roads and highways infrastructure, traffic lights optimisation, and evaluation of the impact of different events such as games and concerts on the performance of transportation infrastructure. Moreover, TFDA2M can be used for estimating the remaining capacity of transportation infrastructure in a growth area. Through this exercise, the number of households that the current road infrastructure can support can be estimated. Of course, future roads can be

designed and evaluated in TFDA2M to demonstrate possible future capacity based
on the growth rate.

10.2   Answers to the research questions

1. **How to detect Bluetooth devices in a vehicle which follows the regu-
   lar traffic flow and distinguishing this from multiple vehicles with one
   Bluetooth device in, in a dense urban area?**

   Since in a dense urban area the number of traffic lights, pedestrians, scoot-
   ers, bikes, traffic disruption events, (e.g. pedestrian crossing), and low-
   speed zones (30k zones) are much higher than freeways, the common ap-
   proaches in the literature are not able to efficiently detect vehicles. In a
   dense urban areas vehicles can be converted to outliers. Therefore, a par-
   tially independent approach is required to detect vehicles which follows the
   regular traffic flow and distinguish these from multiple vehicles with one
   Bluetooth device in. This research used Google API, speed, travel time,
   and distance to identify vehicles which follow the regular traffic flow and
   then applied the Tukey's logic to remove any possible outliers. The cleaning
   algorithm was discussed in Chapter 4. Moreover, since BTMS only detect
   a portion of passing vehicles, Chapter 5 estimates some of the characteris-
   tics of all the passing vehicles from a small portion of detected vehicles by
   BTMS.

2. **What are the effects of the 2016 Kaikoura earthquake on traffic flow
   in Wellington City? Are Bluetooth Traffic Monitoring Systems able to
   identify these effects?**
   Chapter 7 investigated the collected data by BTMS at three levels: Sensor,
   Micro-route, and Macro-route on a weekly and daily basis. Due to the lack
   of official emergency traffic management reports after the earthquake, it was
   not possible to cross-reference the data on a more granular approach (e.g.

hourly). The identified routes were in harmony with the hazardous spatial layers discussed in Chapter 2 and collected information available on the Web.

3. **How to model a large-scale evacuation framework? What are the involved parameters/settings? Evacuees will use their navigation systems in a large-scale evacuation process. Is there any unforeseen outcome? How to predict the network performance during a large-scale evacuation modelling? What are the critical routes?**

TFDA2M is a multi-disciplinary large-scale evacuation framework which considers 7 categories (Behavioural model, navigation systems, infrastructure damage, incidents, urgent trips, traffic light, various demand) and implements these on 13 different scenarios in Chapter 9. These 13 scenarios can be used as benchmarks and sensitivity analysis against each other specifically when no-real data is available. The Results show the utilisation of navigation systems during a large-scale evacuation degrades the road network performance during a large-scale evacuation and an approach for enhancing the performance i proposed in Chapter 9. The results also suggest that a special mode for disaster evacuation is also required in Google Maps. Moreover, critical routes were identified in Chapter 9 through the investigation of occupancy and volume and results, presented numerically and spatially. These results show that the Wellington Motorway mostly has enough capacity during a large scale evacuation, but not all the roads which join the Wellington Motorway. In addition, some suggestions for increasing the performance of the road networks during a large-scale evacuation were presented in chapter 9 such as contraflow lane reversal (i.e. changing the traffic direction of a street). Last but not least, the evacuation time for Wellington varies from about 10 hours to about 40 hours and more, depending on the scenarios and settings.

10.3 limitations of the study

For the data cleaning in Chapter 4, I utilised the Google API combined with other computed metrics and other processing techniques to identify only vehicles which follow regular traffic patterns. The positive side of this approach is that the Google API has high accuracy in reporting the traffic conditions, and is very fast and reliable. The negative side is that this approach cannot operate without access to the Google API, although this is a common approach as most of the new services are linked to various service providers such as Amazon web services and Google APIs.

Due to lack of access to emergency records after the earthquake and spatial layers such as building information, consents, materials of buildings of Wellington, it was impossible to identify the exact locations and times of the impacted routes in Chapter 7. Moreover, these datasets are required to analyse the risks of a road closure due to the collapse of a building in future seismic activities.

Due to lack of access to the phasing of traffic lights due to the closed commercial software, I used the default settings of the simulation environment during the traffic light modelling in Chapter 9. Although I have checked the validity of my results through other available resources in a small scale (e.g. the Google APIs), but access to the actual proprietary algorithms would provide a better benchmark for performance evaluation.

The behaviour of a person may change based on the surrounding environmental conditions and cumulative stress, which may lead to more aggressive behaviour over time. In my driver behavioural model in Chapter 9, I formed my static behavioural model based on the Driver Behavior Questionnaires (DBQ), which may not be a true reflection of the driving behaviour during a stressful long evacuation. This model can be expanded in the future to a dynamic behavioural model, which will be a function of the personality of drivers and dynamic environmental conditions.

10.4 Professional recommendations

For future work, I suggest investigating the correlation between the characteristics of a street and the fitted distribution by examining various involved factors and geospatial layers before and after the 2016 Kaikoura earthquake. To achieve this, a wide range of national and local datasets are required. Possibly some factors such as popularity, the criticality of a street in national freights and local transportation, infrastructure in the surrounding neighbourhoods, and capacity of streets may affect this model. Also, several geospatial layers such as fault, tsunami, and parcel (properties) layers are required to support a better understanding of the fluctuation of traffic flow pre and post-earthquake. This seems achievable as by overlaying these layers, the relation between each parcel (houses and assets) and its susceptibility can be estimated and fed as a pre-processed input to a data mining model. The finding can be used by TFDA2M to estimate the performance of the network. Moreover, through overlaying these spatial and non-spatial layers, the changes in traffic patterns after a future disaster may also be predictable with some level of accuracy.

In addition, I suggest examining optimal values for executing a rerouting algorithm (i.e. rerouting interval) and collecting route costs during an evacuation process. Collection of the route costs and propagation of these costs to evacuees during an evacuation process is a challenging task. A lower collection rate of route costs does not reflect the actual condition of the network and a higher rate requires more processing power and data storage. The rerouting interval should also have an optimised value as a lower value causes too much rerouting and needs heavier processing, and a higher value causes waste of available resources in such moments when the demand is very high and time is very critical. Moreover, drivers may not follow the suggested route to them by their navigation system or local authorities, therefore the predicted performance may not be achievable which is another level of complexity in a large-scale evacuation modelling. However, running a reliable and fault tolerant routing algorithm during an evacuation process requires a distributed approach. A centralised system is not efficient as it requires

very high processing power and it is a single point of failure in this particular context. A distributed system is fault tolerant but may not have a global view of the road networks due to intermittent connections between some sensors and servers due to potential damage to some of the main infrastructure (e.g. telecommunication). In this case, an application for smartphones could be developed to assist drivers in the evacuation process and also exchanging information with main processing servers. In this case, smartphones of drivers can be converted to sensors with the ability to form a network with surrounding smartphones for sharing data and gaining more processing power through sharing of resources, when main servers are unreachable. Therefore, this approach could increase the efficiency of an evacuation process through data sharing and localised routing enhancement, when there is damage to telecommunication infrastructure and processing servers.

Bibliography

[1] 3 nuclear reactors melted down after quake, Japan confirms. `http://edition.cnn.com/2011/WORLD/asiapcf/06/06/japan.nuclear.meltdown/index.html`. [Online; accessed 20-Jun-2019].

[2] Active fault layer. `http://data.gns.cri.nz/af/`. [Online; accessed 20-Jun-2019].

[3] Active faults 250K. `https://data.gns.cri.nz/af/`. [Online; accessed 20-Jun-2019].

[4] At the core of everything bluetooth. `https://www.bluetooth.com/specifications/bluetooth-core-specification`. [Online; accessed 20-Jun-2019].

[5] Avoid this mess. `https://www8.garmin.com/traffic/index.html`. [Online; accessed 20-Jun-2019].

[6] Bosch IoT Things. `https://aws.amazon.com/marketplace/pp / Bosch-Software-Innovations-Bosch-IoT-Things / B07DTJK8MV`. [Online; accessed 20-Jun-2019].

[7] Building on Wellington's Molesworth St being torn down. `https://bit.ly/2KtCqLQ`. [Online; accessed 20-Jun-2019].

[8] Building on Wellington's Molesworth St being torn down. `https://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=11748814`. [Online; accessed 20-Jun-2019].

[9] Car following models. `http : / / sumo . dlr . de / wiki / Car-Following-Models`. [Online; accessed 20-Jun-2019].

[10] Commuting patterns in Wellington: Trends from the census of population and dwellings 2006 and 2013. `http://archive.stats.govt.nz/ Census / 2013-census / profile-and-summary-reports / commuting-patterns-wtn/working-in-wellington.aspx`. [Online; accessed 20-Jun-2019].

[11] Connecting multiple phones with a bluetooth-compatible automotive device. `https : / / support . garmin . com / en-GB / ?faq= IQWVBKIlCm98v5CMATDw3A`. [Online; accessed 20-Jun-2019].

[12] Definition of vehicles, vehicle types, and routes. `https://sumo.dlr. de / wiki / Definition _ of _ Vehicles , _Vehicle _ Types , _and_Routes`. [Online; accessed 20-Jun-2019].

[13] Definition of vehicles, vehicle types, and routes. `https://sumo.dlr. de / wiki / Definition _ of _ Vehicles , _Vehicle _ Types , _and_Routes`. [Online; accessed 20-Jun-2019].

[14] Disaster. `http://www.un-spider.org/node/7661`. [Online; accessed 20-Jun-2019].

[15] Distance Matrix API. `https://developers.google.com/maps/ documentation/distance-matrix/`. [Online; accessed 20-Jun-2019].

[16] Driving difficult on earthquake-damaged roads. `https : / / www . stuff . co . nz / motoring / news / 86434437 / driving-difficult-on-earthquake-damaged-roads`. [Online; accessed 20-Jun-2019].

[17] Emergency preparedness and vulnerable populations: Planning for those most at risk. `https://www.cdc.gov/cpr/whatwedo/ vulnerable.htm`. [Online; accessed 20-Jun-2019].

[18] Evacuate or stay at home. `https://www.cdc.gov/disasters/hurricanes/evacuate.html`. [Online; accessed 20-Jun-2019].

[19] FAQ. `https://sumo.dlr.de/docs/FAQ.html`. [Online; accessed 20-Jun-2019].

[20] Fast facts. `https://www.dmv.ca.gov/portal/dmv/detail/pubs/brochures/fast_facts/ffdl16`. [Online; accessed 10-July-2018].

[21] The forgotten impact Kaikoura earthquake:Wellington still paying the price. `https://www2.deloitte.com/content/dam/Deloitte/nz/Documents/about-deloitte/nz-en-Kaikoura-earthquake-Wellington-impact-report.pdf`. [Online; accessed 20-Jun-2019].

[22] Ghost city: Wellington residents abandon central city after 7.5 earthquak. `https://bit.ly/2XvAKaU`. [Online; accessed 20-Jun-2019].

[23] Google Maps. `https://www.google.com/maps/about/`. [Online; accessed 20-Jun-2019].

[24] Gov. Henry McMaster Orders Mandatory Evacuations for coastal counties effective tomorrow, September 11 at noon. `https://bit.ly/2WV2Fwi`. [Online; accessed 20-Jun-2019].

[25] How often do earthquakes occur along the fault? `https://www.gns.cri.nz/Home/Learning/Science-Topics/Earthquakes/Major-Faults-in-New-Zealand/Wellington-Fault/How-often-do-earthquakes-occur-along-the-fault`. [Online; accessed 20-Jun-2019].

[26] Incident response. `http://www.nzta.govt.nz/assets/resources/maint-specs-incident-response/docs/`

`maint-specs-incident-response.pdf`.   [Online; accessed
20-Jun-2019].

[27] Intelligent   Transport   Systems   (ITS).   `https : / / www .`
`transport . govt . nz / multi-modal / technology /`
`intelligenttransportsystems/`.   [Online; accessed 20-Jun-
2019].

[28] M 7.8 Kaikoura Mon, Nov 14 2016. `https://www.geonet.org.`
`nz/earthquake/story/2016p858000`. [Online; accessed 20-Jun-
2019].

[29] Major   Wellington   power   outage   caused   by   substation   fault.
`https : / / www . stuff . co . nz / national / 105990294 /`
`power-outage-across-parliament-and-wellington-courts`.
[Online; accessed 20-Jun-2019].

[30] Managing   automated   vehicles   enhances   network.   `http : //www .`
`maven-its.eu/`. [Online; accessed 20-Jun-2019].

[31] Monthly   road   crash   statistics   update.   `http : / / www .`
`transport . govt . nz / research / roadcrashstatistics /`
`monthlyoverviewofcrashstatistics                    /`
`monthly-road-crash-statistics-update-february-2018/`.
[Online; accessed 19-July-2018].

[32] Motor   vehicle   casualties.   `http : / / archive .`
`stats . govt . nz / browse _ for _ stats /`
`snapshots-of-nz / nz-social-indicators / Home /`
`Individualsafetyandsecurity / m-v-casualties . aspx`.
[Online; accessed 20-Jun-2019].

[33] Motor   vehicle   crashes   in   New   Zealand   2016.   `https :`
`/ / www . transport . govt . nz / resources /`
`road-safety-resources  /  roadcrashstatistics  /`

motorvehiclecrashesinnewzealand /
motor-vehicle-crashes-in-new-zealand-2016/. [On-
line; accessed 20-Jun-2019].

[34] The national intersection safety problem. https://safety.fhwa.
dot.gov/intersection/other_topics/fhwasa09027/
resources/IntersectionSafetyIssueBrief2.pdf. [Online;
accessed 20-Jun-2019].

[35] Nelson fire: Prepare your evacuation packs - mayor. https:
//www.radionz.co.nz/news/national/382075/
nelson-fire-prepare-your-evacuation-packs-mayor.
[Online; accessed 20-Jun-2019].

[36] New Zealand demographics profile 2018. https://www.
indexmundi.com/new_zealand/demographics_profile.
html. [Online; accessed 20-Jun-2019].

[37] Number of motor vehicles in New Zealand. http://www.
ehinz.ac.nz/assets/Factsheets/Released-2015/
EHI8-9-NumberOfVehiclesInNZ2000-2014.pdf. [Online;
accessed 20-Jun-2019].

[38] The open source vehicular network simulation framework. https://
veins.car2x.org/. [Online; accessed 20-Jun-2019].

[39] Other-projects. https://sumo.dlr.de/wiki/Other/
Projects. [Online; accessed 20-Jun-2019].

[40] Overview. https://people.hofstra.edu/geotrans/eng/
methods/odmatrix.html. [Online; accessed 20-Jun-2019].

[41] Owners manual supplement. https://www8.garmin.com/
software/ecoRoute/pdfs/owners_manual_supplement.
pdf. [Online; accessed 20-Jun-2019].

[42] Pipes and forbes car-following models. `http://www.cerritos.edu/auto/tf-theory-as-applied-its/pfcfm4.htm`. [Online; accessed 20-Jun-2019].

[43] SCATS. `https://www.scats.com.au/`. [Online; accessed 20-Jun-2019].

[44] Simulation-traffic lights. `http://sumo.dlr.de/wiki/Simulation/Traffic_Lights`. [Online; accessed 20-Jun-2019].

[45] Soccer im Einsatz. `http://www.dlr.de/desktopdefault.aspx/tabid-1296/`. [Online; accessed 20-Jun-2019].

[46] Software user manual. `https://www.navman.co.nz/upload/support/guides/MY-SeriesUserManual.pdf`. [Online; accessed 20-Jun-2019].

[47] State highway 1 widening - Hudson Road to P2Wk. `https://www.aucklandcouncil.govt.nz/plans-projects-policies-reports-bylaws/our-plans-strategies/unitary-plan/auckland-unitary-plan-modifications/notices-of-requirement-to-designate-land/docsnorsh1/nor-sh1-transport-assessment-s92.pdf`. [Online; accessed 20-Jun-2019].

[48] Sublane model. `http://www.sumo.dlr.de/userdoc/Simulation/SublaneModel.html`. [Online; accessed 20-Jun-2019].

[49] SUMO OMNET. `http://omnetsimulator.com/sumo-omnet/`. [Online; accessed 20-Jun-2019].

[50] TraCI. `http://sumo.dlr.de/wiki/TraCI`. [Online; accessed 20-Jun-2019].

[51] Traffic and network simulation environment. `http://trans.epfl.ch/`. [Online; accessed 20-Jun-2019].

[52] Traffic model. `https://developers.google.com/maps/documentation/distance-matrix/intro#traffic-model`. [Online; accessed 20-Jun-2019].

[53] Traffic monitoring guide. `https://www.fhwa.dot.gov/policyinformation/tmguide/tmg_2013/traffic-monitoring-methodologies.cfm`. [Online; accessed 20-Jun-2019].

[54] Traffic safety basic facts 2008. `http://erso.swov.nl/safetynet/fixed/WP1/2008/BFS2008_SN-NTUA-1-3-Junctions.pdf`. [Online; accessed 20-Jun-2019].

[55] Transport and the environment. `https://verkehrsforschung.dlr.de/de/projekte/veu`. [Online; accessed 20-Jun-2019].

[56] Types of disasters. `https://www.samhsa.gov/find-help/disaster-distress-helpline/disaster-types`. [Online; accessed 20-Jun-2019].

[57] The tyranny of the peak. `https://www.greaterauckland.org.nz/2017/02/17/the-tyranny-of-the-peak/`. [Online; accessed 20-Jun-2019].

[58] User manual. `https://www.navman.co.nz/upload/support/guides/Navman-EZY260-MY550-MY650-User-Manual.pdf`. [Online; accessed 20-Jun-2019].

[59] Using and interpreting Cronbach's Alpha. `http://data.library.virginia.edu/using-and-interpreting-cronbachs-alpha/`. [Online; accessed 20-Jun-2019].

[60] Vehicle traffic monitoring platform with bluetooth sensors over Zig-Bee. `http://www.libelium.com/vehicle_traffic_monitoring_bluetooth_sensors_over_zigbee/`. [Online; accessed 20-Jun-2019].

[61] Video surveillance for traffic. `https://www.videosurveillance.com/traffic.asp`. [Online; accessed 20-Jun-2019].

[62] Wellington CBD building in danger of collapse. `https://www.radionz.co.nz/news/national/318306/wellington-cbd-building-in-danger-of-collapse`. [Online; accessed 20-Jun-2019].

[63] Wellington City Council. `https://wellington.govt.nz/`. [Online; accessed 20-Jun-2019].

[64] Wellington city current population estimate. `https://profile.idnz.co.nz/wellington/population-estimate`. [Online; accessed 20-Jun-2019].

[65] Wellington city profile. `https://wellington.govt.nz/~/media/about-wellington/profile/files/wellington-city-profile.pdf`. [Online; accessed 20-Jun-2019].

[66] Wellington earthquake national initial response plan. `https://www.civildefence.govt.nz/assets/Uploads/WENIRP-2.0-Final-for-publication.pdf`. [Online; accessed 20-Jun-2019].

[67] Wellington facts and figures. `https://www.wellingtonnz.com/life-in-wellington/facts-and-figures/`. [Online; accessed 20-Jun-2019].

[68] Wellington quakes update 11- Wednesday. `https://wellington.govt.nz/your-council/news/2016/11/wellington-quakes-11`. [Online; accessed 20-Jun-2019].

[69] Wellington shaken awake by deadly earthquake, damage to CBD buildings. https://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=11747451. [Online; accessed 20-Jun-2019].

[70] What can you do if the traffic lights don't change? https://www.drivingtests.co.nz/resources/traffic-lights-change/. [Online; accessed 20-Jun-2019].

[71] What can you say when your P-Value is greater than 0.05? https://bit.ly/2MYWEir. [Online; accessed 20-Jun-2019].

[72] What's the difference between bluetooth versions. https://bit.ly/358s2iM. [Online; accessed 20-Oct-2019].

[73] Why webots? https://www.cyberbotics.com/. [Online; accessed 20-Jun-2019].

[74] ITS innovative project: Arterial travel time monitoring system using bluetooth technology. https://bit.ly/32Vm9Eb, 2011. [Online; accessed 20-Jun-2019].

[75] ABERG, L., AND RIMMO, P.-A. Dimensions of aberrant driver behaviour. *Ergonomics 41*, 1 (1998), 39–56.

[76] AKAIKE, H. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on 19*, 6 (1974), 716–723.

[77] AMOOZADEH, M., DENG, H., CHUAH, C.-N., ZHANG, H. M., AND GHOSAL, D. Platoon management with cooperative adaptive cruise control enabled by {VANET}. *Vehicular Communications 2*, 2 (2015), 110 – 123.

[78] ARAGHI, B. N., CHRISTENSEN, L. T., KRISHNAN, R., AND LAHRMANN, H. Application of bluetooth technology for mode-specific travel time estimation on arterial roads: Potentials and challenges.

[79] ARAGHI, B. N., PEDERSEN, K. S., CHRISTENSEN, L. T., KRISHNAN, R., AND LAHRMANN, H. Accuracy of travel time estimation using bluetooth technology: Case study Limfjord tunnel Aalborg. *International Journal of Intelligent Transportation Systems Research 13*, 3 (2015), 166–191.

[80] BAN, X., HERRING, R., HAO, P., AND BAYEN, A. Delay pattern estimation for signalized intersections using sampled travel times. *Transportation Research Record: Journal of the Transportation Research Board*, 2130 (2009), 109–119.

[81] BAN, X. J., HAO, P., AND SUN, Z. Real time queue length estimation for signalized intersections using travel times from mobile sensors. *Transportation Research Part C: Emerging Technologies 19*, 6 (2011), 1133–1156.

[82] BANERJEE, A., DUFLO, E., AND QIAN, N. On the road: Access to transportation infrastructure and economic growth in China. Tech. rep., National Bureau of Economic Research, 2012.

[83] BEARD, C., AND STALLINGS, W. *Wireless communication networks and systems*. Pearson Education Limited, 2016.

[84] BUSANELLI, S., FERRARI, G., AND GRUPPINI, R. Performance analysis of broadcast protocols in VANETs with Poisson vehicle distribution. In *ITS Telecommunications (ITST), 2011 11th International Conference on* (2011), IEEE, pp. 133–138.

[85] BYRNE, K. Why you should evacuate ahead of a hurricane. https://www.accuweather.com/en/weather-news/why-you-should-evacuate-ahead-of-a-hurricane/70001630. [Online; accessed 20-Jun-2019].

[86] CASCETTA, E. *Transportation systems engineering: theory and methods*, vol. 49. Springer Science & Business Media, 2013.

[87] CHAKRABORTY, B., MAEDA, T., AND CHAKRABORTY, G. Multiobjective route selection for car navigation system using genetic algorithm. In *Soft Computing in Industrial Applications, 2005. SMCia/05. Proceedings of the 2005 IEEE Mid-Summer Workshop on* (2005), IEEE, pp. 190–195.

[88] CHAKROBORTY, P., AND DAS, A. *Principles of transportation engineering*. PHI Learning Pvt. Ltd., 2017.

[89] CHEN, W., CHEN, L., CHEN, Z., AND TU, S. WITS: A wireless sensor network for intelligent transportation system. In *Computer and Computational Sciences, 2006. IMSCCS'06. First International Multi-Symposiums on* (2006), vol. 2, IEEE, pp. 635–641.

[90] CHEN, Y.-Z., SHEN, S.-F., CHEN, T., AND YANG, R. Path optimization study for vehicles evacuation based on Dijkstra algorithm. *Procedia Engineering 71* (2014), 159–165.

[91] CHEN, Z., XIA, J. C., IRAWAN, B., AND CAULFIED, C. Development of location-based services for recommending departure stations to park and ride users. *Transportation Research Part C: Emerging Technologies 48*, Supplement C (2014), 256 – 268.

[92] CHIU, Y.-C. Traffic scheduling simulation and assignment for area-wide evacuation. In *Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on* (2004), IEEE, pp. 537–542.

[93] CHO, E., AND KIM, S. Cronbach's coefficient Alpha: Well known but poorly understood. *Organizational Research Methods 18*, 2 (2015), 207–230.

[94] COLLOTTA, M., BELLO, L. L., AND PAU, G. A novel approach for dynamic traffic lights management based on wireless sensor networks and multiple fuzzy logic controllers. *Expert Systems with Applications 42*, 13 (2015), 5403–5415.

[95] COMPTON, R. J., CARP, J., CHADDOCK, L., FINEMAN, S. L., QUANDT, L. C., AND RATLIFF, J. B. Anxiety and error monitoring: Increased error sensitivity or altered expectations? *Brain and Cognition 64*, 3 (2007), 247–256.

[96] CONRAD, R., AND FIGLIOZZI, M. Algorithms to quantify impact of congestion on time-dependent real-world urban freight distribution networks. *Transportation Research Record: Journal of the Transportation Research Board*, 2168 (2010), 104–113.

[97] COOPER, L. Y., AND STROUP, D. W. ASET-a computer program for calculating available safe egress time. *Fire Safety Journal 9*, 1 (1985), 29–45.

[98] CRONBACH, L. J. Coefficient Alpha and the internal structure of tests. *psychometrika 16*, 3 (1951), 297–334.

[99] CROSS, D., HOECKLE, J., LAVINE, M., RUBIN, J., AND SNOW, K. Detecting non-discoverable bluetooth devices. In *International Conference on Critical Infrastructure Protection* (2007), Springer, pp. 281–293.

[100] CURRIE, J., AND WALKER, R. Traffic congestion and infant health: Evidence from E-ZPass. *American Economic Journal: Applied Economics 3*, 1 (2011), 65–90.

[101] DALL'OSSO, F., AND DOMINEY-HOWES, D. Public assessment of the usefulness of" draft" tsunami evacuation maps from Sydney, Australia-implications for the establishment of formal evacuation plans. *Natural Hazards and Earth System Sciences 10*, 8 (2010), 1739–1750.

[102] DATONDJI, S. R. E., DUPUIS, Y., SUBIRATS, P., AND VASSEUR, P. A survey of vision-based traffic monitoring of road intersections. *IEEE transactions on intelligent transportation systems 17*, 10 (2016), 2681–2698.

[103] DEKKING, F. M. *A Modern Introduction to Probability and Statistics: Understanding why and how*. Springer Science & Business Media, 2005.

[104] DELLING, D., SANDERS, P., SCHULTES, D., AND WAGNER, D. Engineering route planning algorithms. In *Algorithmics of large and complex networks*. Springer, 2009, pp. 117–139.

[105] DEQI, H., XIUMIN, C., AND ZHE, M. A simulation framework for emergency response of highway traffic accident. *Procedia Engineering 29* (2012), 1075 – 1080. 2012 International Workshop on Information and Electronics Engineering.

[106] DION, F., AND RAKHA, H. Estimating spatial travel times using automatic vehicle identification data. `https://rosap.ntl.bts.gov/view/dot/5457`, 2001. [Online; accessed 20-Jun-2019].

[107] DION, F., AND RAKHA, H. Estimating dynamic roadway travel times using automatic vehicle identification data for low sampling rates. *Transportation Research Part B: Methodological 40*, 9 (2006), 745 – 766.

[108] DLM. Coupling SUMO and NS2 via TraCI. `https://sourceforge.net/p/sumo/mailman/message/31690813/`. [Online; accessed 20-Jun-2019].

[109] DRANE, C. R., AND RIZOS, C. *Positioning systems in intelligent transportation systems*. Artech House, Inc., 1998.

[110] DÍAZ, J. J. V., GONZÁLEZ, A. B. R., AND WILBY, M. R. Bluetooth traffic monitoring systems for travel time estimation on freeways. *IEEE Transactions on Intelligent Transportation Systems 17*, 1 (Jan 2016), 123–132.

[111] EAGLE, N., AND (SANDY) PENTLAND, A. Reality mining: Sensing complex social systems. *Personal Ubiquitous Comput. 10*, 4 (Mar. 2006), 255–268.

[112] EL SADDIK, A. Digital twins: the convergence of multimedia technologies. *IEEE MultiMedia 25*, 2 (2018), 87–92.

[113] ERDMANN, J. SUMO's lane-changing model. In *Modeling Mobility with Open Data*. Springer, 2015, pp. 105–123.

[114] ERDMANN, J., AND KRAJZEWICZ, D. SUMO's road intersection model. In *Simulation of Urban MObility User Conference* (2013), Springer, pp. 3–17.

[115] ERKAN, İ., AND HASTEMOGLU, H. Bluetooth as a traffic sensor for stream travel time estimation under Bogazici Bosporus conditions in Turkey. *Journal of Modern Transportation 24*, 3 (Sep 2016), 207–214.

[116] FIALHO, A., CELI, L., CISMONDI, F., VIEIRA, S., RETI, S., SOUSA, J., FINKELSTEIN, S., ET AL. Disease-based modeling to predict fluid response in intensive care units. *Methods Inf Med 52*, 6 (2013), 494–502.

[117] FLÖTTERÖD, Y.-P., AND ERDMANN, J. Dynamic reroute modeling for emergency evacuation: Case study of Brunswick City, Germany. *International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering 12*, 4 (2018), 439 – 449.

[118] FORBES, T., ZAGORSKI, H., HOLSHOUSER, E., AND DETERLINE, W. Measurement of driver reactions to tunnel conditions. In *Highway Research Board Proceedings* (1958), vol. 37.

[119] FRÄMLING, K., HOLMSTRÖM, J., ALA-RISKU, T., AND KÄRKKÄINEN, M. Product agents for handling information about physical objects. *Report of Laboratory of information processing science series B, TKO-B 153*, 03 (2003).

[120] FRIESEN, M. R., AND MCLEOD, R. D. Bluetooth in intelligent transportation systems: A survey. *International Journal of Intelligent Transportation Systems Research 13*, 3 (Sep 2015), 143–153.

[121] FURUYA, T., AND TAYLOR, C. J. *Road intersection monitoring from video with large perspective deformation*. PhD thesis, University of Pennsylvania, 2014.

[122] GALÁN-GARCÍA, J. L., AGUILERA-VENEGAS, G., GALÁN-GARCÍA, M. Á., AND RODRÍGUEZ-CIELOS, P. A new probabilistic extension of Dijkstra's algorithm to simulate more realistic traffic flow in a smart city. *Applied mathematics and Computation 267* (2015), 780–789.

[123] GAO, B. A programmer's note on TraCI_tls, TraCI, and SUMO. https://intelaligent.github.io/tctb/post-learning-traci-tls.html. [Online; accessed 20-Jun-2019].

[124] GOTO, Y., AFFAN, M., NURDIN, Y., YULIANA, D. K., ET AL. Tsunami evacuation simulation for disaster education and city planning. *Journal of Disaster Research 7*, 1 (2012), 92–101.

[125] GREENBERG, M. I. *Encyclopedia of terrorist, natural, and man-made disasters*. Jones & Bartlett Learning, 2006.

[126] GRUMERT, E. F., AND TAPANI, A. Traffic state estimation using connected vehicles and stationary detectors. *Journal of Advanced Transportation 2018* (2018).

[127] HAARTSEN, J. Bluetooth-the universal radio interface for ad hoc, wireless connectivity. *Ericsson review 3*, 1 (1998), 110–117.

[128] HAFEEZ, K. A., ZHAO, L., LIAO, Z., AND MA, B. N. W. Performance analysis of broadcast messages in VANETs safety applications. In *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE* (Dec 2010), pp. 1–5.

[129] HAGHANI, A., AND HAMEDI, M. Application of bluetooth technology in traffic detection, surveillance, and traffic management. *Journal of Intelligent Transportation Systems 17*, 2 (2013), 107–109.

[130] HAGHANI, A., AND JUNG, S. A dynamic vehicle routing problem with time-dependent travel times. *Computers & operations research 32*, 11 (2005), 2959–2986.

[131] HAO, P., SUN, Z., BAN, X. J., GUO, D., AND JI, Q. Vehicle index estimation for signalized intersections using sample travel times. *Procedia - Social and Behavioral Sciences 80* (2013), 473 – 490. 20th International Symposium on Transportation and Traffic Theory (ISTTT 2013).

[132] HARRISON, W. A. Reliability of the driver behaviour questionnaire in a sample of novice drivers. https://bit.ly/2XxhOGl, 2009. [Online; accessed 20-Jun-2019].

[133] HART, P. E., NILSSON, N. J., AND RAPHAEL, B. A formal basis for the heuristic determination of minimum cost paths. *IEEE transactions on Systems Science and Cybernetics 4*, 2 (1968), 100–107.

[134] HELBING, D., FARKAS, I., AND VICSEK, T. Simulating dynamical features of escape panic. *Nature 407*, 6803 (2000), 487.

[135] HERRERA, J. C., WORK, D. B., HERRING, R., BAN, X. J., JACOBSON, Q., AND BAYEN, A. M. Evaluation of traffic data obtained via GPS-enabled mobile phones: The mobile century field experiment. *Transportation Research Part C: Emerging Technologies 18*, 4 (2010), 568–583.

[136] HO, I. W.-H., LEUNG, K. K., AND POLAK, J. W. Stochastic model and connectivity dynamics for VANETs in signalized road systems. *IEEE/ACM Transactions on Networking (TON) 19*, 1 (2011), 195–208.

[137] HO, I. W. H., LEUNG, K. K., AND POLAK, J. W. Stochastic model and connectivity dynamics for VANETs in signalized road systems. *IEEE/ACM Transactions on Networking 19*, 1 (Feb 2011), 195–208.

[138] HOOGENDOORN, S. P., AND BOVY, P. H. State-of-the-art of vehicular traffic flow modelling. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering 215*, 4 (2001), 283–303.

[139] HORN, P. S., FENG, L., LI, Y., AND PESCE, A. J. Effect of outliers and nonhealthy individuals on reference interval estimation. *Clinical Chemistry 47*, 12 (2001), 2137–2145.

[140] HOYOS, M. C., MORALES, R. S., AND AKHAVAN-TABATABAEI, R. Or models with stochastic components in disaster operations management: A literature survey. *Computers & Industrial Engineering 82* (2015), 183–197.

[141] HUI, P., CHAINTREAU, A., SCOTT, J., GASS, R., CROWCROFT, J., AND DIOT, C. Pocket switched networks and human mobility in conference environments. In *Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-tolerant Networking* (New York, NY, USA, 2005), WDTN '05, ACM, pp. 244–251.

[142] INSTITUTE, S. R. Automatic vehicle identification model deployment initiative-system design document. Tech. rep., (Texas Department of Transportation, 1998.

[143] ITS KOREA, H. S., AND UNIVERSITY, A. Development of practical technology for DSRC traffic information system, (In Korean). Tech. rep., (Korea Expressway Corporation, 2008.

[144] JACOB, B. Assessment of the accuracy and classification of weigh-in-motion systems part 1: Statistical background. *International Journal of Heavy Vehicle Systems 7*, 2-3 (2000), 136–152.

[145] JANG, J. Outlier filtering algorithm for travel time estimation using dedicated short-range communications probes on rural highways. *IET Intelligent Transport Systems 10*, 6 (2016), 453–460.

[146] JÄRV, O., AHAS, R., SALUVEER, E., DERUDDER, B., AND WITLOX, F. Mobile phones in a traffic flow: a geographical perspective to evening rush hour traffic analysis using call detail records. *PloS one 7*, 11 (2012), e49171.

[147] JIN, X., ZHANG, Y., WANG, F., LI, L., YAO, D., SU, Y., AND WEI, Z. Departure headways at signalized intersections: A Log-normal distribution model approach. *Transportation Research Part C: Emerging Technologies 17*, 3 (2009), 318 – 327.

[148] JULIE ILES, D. G., AND WILLIAMS, K. SH1 north of wellington re-opens after epic closure caused by truck crash. https://bit.ly/2Xh7K6D. [Online; accessed 20-Jun-2019].

[149] KAEWUNRUEN, S., SUSSMAN, J. M., AND MATSUMOTO, A. Grand challenges in transportation and transit systems. *Frontiers in built environment 2* (2016), 4.

[150] KANOH, H., AND HARA, K. Hybrid genetic algorithm for dynamic multi-objective route planning with predicted traffic in a real-world road network. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation* (2008), ACM, pp. 657–664.

[151] KECMAN, P., AND GOVERDE, R. M. Predictive modelling of running and dwell times in railway traffic. *Public Transport 7*, 3 (2015), 295–319.

[152] KENWARD, A., AND RAJA, U. Blackout: Extreme weather climate change and power outages. *Climate central 10* (2014).

[153] KESTING, A., AND TREIBER, M. Calibrating car-following models by using trajectory data: Methodological study. *Transportation Research Record: Journal of the Transportation Research Board*, 2088 (2008), 148–156.

[154] KESTING, A., AND TREIBER, M. Traffic flow dynamics: data, models and simulation. https://www.springer.com/gp/book/9783642324598, 2013. [Online; accessed 20-Jun-2019].

[155] KESTING, A., TREIBER, M., AND HELBING, D. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity.

Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 368, 1928 (2010), 4585–4605.

[156] KHABAZIAN, M., AND ALI, M. K. M. A performance modeling of connectivity in vehicular ad hoc networks. *Vehicular Technology, IEEE Transactions on 57*, 4 (2008), 2440–2450.

[157] KHANAFER, M., GUENNOUN, M., AND MOUFTAH, H. T. WSN architectures for intelligent transportation systems. In *New Technologies, Mobility and Security (NTMS), 2009 3rd International Conference on* (2009), IEEE, pp. 1–8.

[158] KIM, J., AND MAHMASSANI, H. S. Compound Gamma representation for modeling vehicle-to-vehicle and day-to-day travel time variability in a traffic network. Tech. rep., 2014.

[159] KINOSHITA, A., TAKASU, A., AND ADACHI, J. Real-time traffic incident detection using a probabilistic topic model. *Information Systems 54* (2015), 169–188.

[160] KOELLE, A. R. Single antenna location and direction finding system, Apr. 23 1996. US Patent 5,510,795.

[161] KOTSIALOS, A., AND PAPAGEORGIOU, M. The importance of traffic flow modeling for motorway traffic control. *Networks and Spatial Economics 1*, 1-2 (2001), 179–203.

[162] KRAJZEWICZ, D., ERDMANN, J., BEHRISCH, M., AND BIEKER, L. Recent development and applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements 5*, 3&4 (December 2012), 128–138.

[163] KRAJZEWICZ, D., HERTKORN, G., RÖSSEL, C., AND WAGNER, P. SUMO (Simulation of Urban MObility)-an open-source traffic simulation.

In *Proceedings of the 4th Middle East Symposium on Simulation and Modelling (MESM20002)* (2002), pp. 183–187.

[164] KRAUSS, S. *Microscopic modeling of traffic flow: Investigation of collision free vehicle dynamics*. PhD thesis, 1998.

[165] KUMAR, T. S. P. SUMO, Open Street Maps and NS2 - a real traffic simulation. https://www.nsnam.com/2016/06/sumo-open-street-maps-and-ns2-real.html. [Online; accessed 20-Jun-2019].

[166] LÄMMEL, G., GRETHER, D., AND NAGEL, K. The representation and implementation of time-dependent inundation in large-scale microscopic evacuation simulations. *Transportation Research Part C: Emerging Technologies 18*, 1 (2010), 84–98.

[167] LAURIKKALA, J., JUHOLA, M., KENTALA, E., LAVRAC, N., MIKSCH, S., AND KAVSEK, B. Informal identification of outliers in medical data. In *Fifth International Workshop on Intelligent Data Analysis in Medicine and Pharmacology* (2000), vol. 1, pp. 20–24.

[168] LEFÈVRE, S., LAUGIER, C., AND IBAÑEZ-GUZMÁN, J. Risk assessment at road intersections: Comparing intention and expectation. In *Intelligent Vehicles Symposium (IV), 2012 IEEE* (2012), IEEE, pp. 165–171.

[169] LI, Y., AND SUN, D. Microscopic car-following model for the traffic flow: the state of the art. *Journal of Control Theory and Applications 10*, 2 (2012), 133–143.

[170] LINDELL, M. K. EMBLEM2: An empirically based large scale evacuation time estimate model. *Transportation research part A: policy and practice 42*, 1 (2008), 140–154.

[171] LV, W., SONG, W.-G., LIU, X.-D., AND MA, J. A microscopic lane changing process model for multilane traffic. *Physica A: Statistical Mechanics and its Applications 392*, 5 (2013), 1142–1152.

[172] MA, X., YU, H., WANG, Y., AND WANG, Y. Large-scale transportation network congestion evolution prediction using deep learning theory. *PloS one 10*, 3 (2015), e0119044.

[173] MALINOVSKIY, Y., LEE, U.-K., WU, Y.-J., AND WANG, Y. Investigation of bluetooth-based travel time estimation error on a short corridor. In *Transportation research board 90th annual meeting* (2011), no. 11-3056.

[174] MANASKASEMSAK, B., CHINTHANET, B., AND RUNGSAWANG, A. Graph clustering-based emerging event detection from Twitter data stream. In *Proceedings of the Fifth International Conference on Network, Communication and Computing* (2016), ACM, pp. 37–41.

[175] MARTCHOUK, M., AND MANNERING, F. L. Analysis of travel time reliability on Indiana interstates. Tech. rep., NEXTRANS, 2009.

[176] MASHER, D. P., ROSS, D., WONG, P., TUAN, P., ZEIDLER, H., AND PETRACEK, S. Guidelines for design and operation of ramp control systems. https://trid.trb.org/view.aspx?id=33060, 1975. [Online; accessed 20-Jun-2019].

[177] MATTSSON, M. Investigating the factorial invariance of the 28-item DBQ across genders and age groups: an exploratory structural equation modeling study. *Accident Analysis & Prevention 48* (2012), 379–396.

[178] MCCRAE, R. R., AND JOHN, O. P. An introduction to the five-factor model and its applications. *Journal of personality 60*, 2 (1992), 175–215.

[179] MCGOWEN, P., AND SANDERSON, M. Accuracy of pneumatic road tube counters. In *Proceedings of the 2011 Western District Annual Meeting, Anchorage, AK, USA* (2011), vol. 1013, p. 2.

[180] MICHAU, G., NANTES, A., BHASKAR, A., CHUNG, E., ABRY, P., AND BORGNAT, P. Bluetooth data in an urban context: Retrieving vehicle trajectories. *IEEE Transactions on Intelligent Transportation Systems 18*, 9 (Sept 2017), 2377–2386.

[181] MIDDLETON, D., PARKER, R., ET AL. Initial evaluation of selected detectors to replace inductive loops on freeways. Tech. rep., United States. Federal Highway Administration, 2000.

[182] MITSAKIS, E., GRAU, J. M. S., CHRYSOHOOU, E., AND AIFADOPOULOU, G. A robust method for real time estimation of travel times for dense urban road networks using point-to-point detectors. *Transport 30*, 3 (2015), 264–272.

[183] MOHAMMADI, N., AND TAYLOR, J. E. Smart city digital twins. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)* (2017), IEEE, pp. 1–5.

[184] MUBASHER, M. M., JAFFRY, S. W., AND JAHANGIR, R. Modeling of individual differences in car-following behaviour of drivers. In *Multi-topic Conference (INMIC), 2017 International* (2017), IEEE, pp. 1–7.

[185] MUJADDAD AFZAL, PRAKASH RANJITKAR, S. C. N. J., AND SITHIKA. Emergency evacuation modelling using traffic simulation: A case study of Auckland. Australasian Transport Research Forum.

[186] MURRAY-TUITE, P., AND WOLSHON, B. Evacuation transportation modeling: An overview of research, development, and practice. *Transportation Research Part C: Emerging Technologies 27* (2013), 25 – 45. Selected papers from the Seventh Triennial Symposium on Transportation Analysis (TRISTAN VII).

[187] MURTHY, S., AND GARCIA-LUNA-ACEVES, J. J. An efficient routing protocol for wireless networks. *Mobile Networks and applications 1*, 2 (1996), 183–197.

[188] NELSON, P. R., COFFIN, M., AND COPELAND, K. A. *Introductory statistics for engineering experimentation.* Academic Press, 2003.

[189] NHA, V. T. N., DJAHEL, S., AND MURPHY, J. A comparative study of vehicles' routing algorithms for route planning in smart cities. In *Vehicular Traffic Management for Smart Cities (VTM), 2012 First International Workshop on* (2012), IEEE, pp. 1–6.

[190] NICOLAI, T., AND KENN, H. About the relationship between people and discoverable bluetooth devices in urban environments. In *Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st International Symposium on Computer Human Interaction in Mobile Technology* (New York, NY, USA, 2007), Mobility '07, ACM, pp. 72–78.

[191] NILS-AXEL, M. Natural, man-made and imagined disaster. https://bit.ly/2rTOii1, 2010. [Online; accessed 20-Jun-2019].

[192] NISHIUCHI, H., SHIOMI, Y., KURAUCHI, S., YOSHII, T., AND SUGA, Y. Analysis on directivity of MAC address detection for collecting mobile data. *Journal of Japan Society of Civil Engineers, Ser. F3 (Civil Engineering Informatics) 71* (2015).

[193] NORDFJÆRN, T., HEZAVEH, A. M., AND MAMDOOHI, A. R. An analysis of reported driver behaviour in samples of domestic and expatriate Iranians. *Journal of Risk Research 18*, 5 (2015), 566–580.

[194] NOVACO, R. W., STOKOLS, D., CAMPBELL, J., AND STOKOLS, J. Transportation, stress, and community psychology. *American Journal of Community Psychology 7*, 4 (1979), 361–380.

[195] OH, O., KWON, K. H., AND RAO, H. R. An exploration of social media in extreme events: Rumor Theory and Twitter during the Haiti earthquake 2010. In *Icis* (2010), vol. 231.

[196] OLVET, D. M., AND HAJCAK, G. The error-related negativity (ERN) and psychopathology: Toward an endophenotype. *Clinical psychology review 28*, 8 (2008), 1343–1354.

[197] O'NEILL, E., KOSTAKOS, V., KINDBERG, T., SCHIEK, A. F. G., PENN, A., FRASER, D. S., AND JONES, T. Instrumenting the city: Developing methods for observing and understanding the digital cityscape. In *Proceedings of the 8th International Conference on Ubiquitous Computing* (Berlin, Heidelberg, 2006), UbiComp 06, Springer-Verlag, pp. 315–332.

[198] OZIMEK, A., MILES, D., ET AL. Stata utilities for geocoding and generating travel time and travel distance information. *Stata Journal 11*, 1 (2011), 106.

[199] ÖZKAN, T., AND LAJUNEN, T. What causes the differences in driving between young men and women? the effects of gender roles and sex on young drivers' driving behaviour and self-assessment of skills. *Transportation Research Part F: Traffic Psychology and Behaviour 9*, 4 (2006), 269–277.

[200] ÖZKAN, T., LAJUNEN, T., CHLIAOUTAKIS, J. E., PARKER, D., AND SUMMALA, H. Cross-cultural differences in driving behaviours: A comparison of six countries. *Transportation research part F: traffic psychology and behaviour 9*, 3 (2006), 227–242.

[201] PANG, J. Review of microcontroller based intelligent traffic light control. In *Emerging Technologies for a Smarter World (CEWIT), 2015 12th International Conference & Expo on* (2015), IEEE, pp. 1–5.

[202] PARUCHURI, V., CHELLAPPAN, S., AND LENIN, R. B. Arrival time based traffic signal optimization for intelligent transportation systems. In *Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on* (2013), IEEE, pp. 703–709.

[203] PASIN, M. Performing ANOVA test in R: Results and interpretation. http://www.analyticsforfun.com/2014/06/performing-anova-test-in-r-results-and.html. [Online; accessed 20-Jun-2019].

[204] PEL, A., BLIEMER, M., AND HOOGENDOORN, S. Modelling traveller behaviour under emergency evacuation conditions. *European Journal of Transport and Infrastructure Research (EJTIR), 11 (2), 2011* (2011).

[205] PHAM, H. *Springer handbook of engineering statistics*. Springer Science & Business Media, 2006.

[206] PHAM, P. P. Comprehensive analysis of the IEEE 802.11. *Mobile Networks and Applications 10*, 5 (2005), 691–703.

[207] PILLAI, P. J., YALLA, V., AND OGUCHI, K. "Ease of Driving" road classification for night-time driving conditions. In *SAE Technical Paper* (04 2016), SAE International.

[208] PIPES, L. A. An operational analysis of traffic dynamics. *Journal of applied physics 24*, 3 (1953), 274–281.

[209] PISHRO-NIK, H., GANZ, A., AND NI, D. The capacity of vehicular ad hoc networks. In *45th Annual Allerton Conference on Communication, Control and Computing, Allerton, USA* (2007), vol. 29.

[210] POURABDOLLAH, M., BJÄRKVIK, E., FÜRER, F., LINDENBERG, B., AND BURGDORF, K. Calibration and evaluation of car following models using real-world driving data. In *Intelligent Transportation Systems (ITSC), 2017 IEEE 20th International Conference on* (2017), IEEE, pp. 1–6.

[211] PRADHAN, R. P., AND BAGCHI, T. P. Effect of transportation infrastructure on economic growth in India: the VECM approach. *Research in Transportation Economics 38*, 1 (2013), 139–148.

[212] PROTO, A., ALEXANDRE, L. A., BATISTA, M. L., OLIVEIRA, I. L., AND CANSIAN, A. M. Statistical model applied to netflow for network intrusion detection. In *Transactions on computational science XI*. Springer, 2010, pp. 179–191.

[213] PUCKETT, D. D., AND VICKICH, M. J. Bluetooth®-based travel time/speed measuring systems development. Tech. rep., 2010.

[214] RAJASEKHAR, L. *Microscopic Control Delay Modeling at Signalized Arterials Using Bluetooth Technology*. PhD thesis, Virginia Tech, 2011.

[215] RAMEZANI GHALENOEI, M. Traffic modeling, estimation and control for large-scale congested urban networks.

[216] REASON, J., MANSTEAD, A., STRADLING, S., BAXTER, J., AND CAMPBELL, K. Errors and violations on the roads: a real distinction? *Ergonomics 33*, 10-11 (1990), 1315–1332.

[217] RIBEIRO, S. K., KOBAYASHI, S., BEUTHE, M., GASCA, J., GREENE, D., LEE, D. S., MUROMACHI, Y., NEWTON, P. J., PLOTKIN, S., SPERLING, D., ET AL. Transportation and its infrastructure. https://www.ipcc.ch/site/assets/uploads/2018/02/ar4-wg3-chapter5-1.pdf, 2007. [Online; accessed 20-Jun-2019].

[218] RIMMÖ, P.-A. Aberrant driving behaviour: homogeneity of a four-factor structure in samples differing in age and gender. *Ergonomics 45*, 8 (2002), 569–582.

[219] RUTHERFORD, H. Almost 18 months on, BNZ says it won't return to its quake damaged wellington. https://bit.ly/31HS23D. [Online; accessed 20-Jun-2019].

[220] SANKARAN, J. K., GORE, A., AND COLDWELL, B. The impact of road traffic congestion on supply chains: insights from Auckland, New Zealand. *International Journal of Logistics: Research and Applications 8*, 2 (2005), 159–180.

[221] ŞERBU, C., OPRUŢA, D., AND SOCACIU, L. Ranking the types of intersections for assessing the safety of pedestrians using TOPSIS method.

Leonardo Electronic Journal of Practices and Technologies 13, 25 (2014), 242–253.

[222] SEYMOUR, P., JAMES, E., AND MILLER, B. Use of a Google-based mapping system for traffic management center web navigation. In *Transportation Research Board 86th Annual Meeting* (2007), no. 07-3384.

[223] SHAHABI, K., AND WILSON, J. P. Scalable evacuation routing in a dynamic environment. *Computers, Environment and Urban Systems 67* (2018), 29–40.

[224] SHARIFI, E., HAMEDI, M., HAGHANI, A., AND SADRSADAT, H. Analysis of vehicle detection rate for bluetooth traffic sensors: a case study in Maryland and Delaware. In *18th World Congress on on Intelligent Transport Systems* (2011).

[225] SHARMA, S., PITHORA, A., GUPTA, G., GOEL, M., AND SINHA, M. Traffic light priority control for emergency vehicle using RFID. *Int. J. Innov. Eng. Technol 2*, 2 (2013), 363–366.

[226] SOLBERG, H. E., AND LAHTI, A. Detection of outliers in reference distributions: performance of horn's algorithm. *Clinical chemistry 51*, 12 (2005), 2326–2332.

[227] SOMMER, C., YAO, Z., GERMAN, R., AND DRESSLER, F. On the need for bidirectional coupling of road traffic microsimulation and network simulation. In *Proceedings of the 1st ACM SIGMOBILE workshop on Mobility models* (2008), ACM, pp. 41–48.

[228] SONG, J., WU, Y., XU, Z., AND LIN, X. Research on car-following model based on SUMO. In *Advanced Infocomm Technology (ICAIT), 2014 IEEE 7th International Conference on* (2014), IEEE, pp. 47–55.

[229] SRIVASTAVA, A., AND GEROLIMINIS, N. Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-

order model. *Transportation Research Part C: Emerging Technologies 30* (2013), 161–177.

[230] STEWART, J.   As hurricane Florence looms, drone pilots pre-pare for recovery.   https://www.wired.com/story/hurricane-florence-drone-recovery/. [Online; accessed 20-Jun-2019].

[231] STRÖMGREN, P., OLSTAM, J., AND TAPANI, A. A model for traffic simu-lation of flared rural road intersections. *Transportation Research Procedia 6*, 2015 (2015), 239–258.

[232] SUBIRATS, P., DUPUIS, Y., VIOLETTE, E., DOUCET, D., AND DUPRE, G. A new tool to evaluate safety of crossroad. In *4th International Sympo-sium on Highway Geometric Design. Valence* (2010), pp. 2–5.

[233] TADEUSIAK, M. Traffic flow modelling conceptual model and specific implementations. https://warwick.ac.uk/fac/cross_fac/complexity/study/emmcs/outcomes/studentprojects/tadeusiak.pdf. [Online; accessed 20-Jun-2019].

[234] TAO, F., CHENG, J., QI, Q., ZHANG, M., ZHANG, H., AND SUI, F. Dig-ital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology 94*, 9-12 (2018), 3563–3576.

[235] TAVAKOL, M., AND DENNICK, R. Making sense of Cronbach's Alpha. *International journal of medical education 2* (2011), 53.

[236] THOMPSON, R. R., GARFIN, D. R., AND SILVER, R. C. Evacuation from natural disasters: a systematic review of the literature. *Risk analysis 37*, 4 (2017), 812–839.

[237] TOSHNIWAL, D., AND YADAV, S. Adaptive outlier detection in streaming time series. In *Proceedings of International Conference on Asia Agriculture and Animal, ICAAA, Hong Kong* (2011), vol. 13, pp. 186–192.

[238] TREIBER, M., HENNECKE, A., AND HELBING, D. Congested traffic states in empirical observations and microscopic simulations. *Physical review E 62*, 2 (2000), 1805.

[239] TSAI, J., FRIDMAN, N., BOWRING, E., BROWN, M., EPSTEIN, S., KAMINKA, G., MARSELLA, S., OGDEN, A., RIKA, I., SHEEL, A., ET AL. ESCAPES: evacuation simulation with children, authorities, parents, emotions, and social comparison. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2* (2011), International Foundation for Autonomous Agents and Multiagent Systems, pp. 457–464.

[240] TSUBOTA, T., AND YOSHII, T. An analysis of the detection probability of MAC address from a moving bluetooth device. *Transportation Research Procedia 21* (2017), 251–256.

[241] TU, H., TAMMINGA, G., DROLENGA, H., DE WIT, J., AND VAN DER BERG, W. Evacuation plan of the city of almere: simulating the impact of driving behavior on evacuation clearance time. *Procedia Engineering 3* (2010), 67–75.

[242] VAN BOXEL, D., SCHNEIDER IV, W., AND BAKULA, C. Innovative real-time methodology for detecting travel time outliers on interstate highways and urban arterials. *Transportation Research Record: Journal of the Transportation Research Board*, 2256 (2011), 60–67.

[243] VENT, R. Real traffic flow modelling with SUMO. `http://ds.cs.ut.ee/courses/course-files/FinalReport-Vent.pdf`. [Online; accessed 20-Jun-2019].

[244] VINCIARELLI, A., AND MOHAMMADI, G. A survey of personality computing. *IEEE Transactions on Affective Computing 5*, 3 (2014), 273–291.

[245] VINEL, A., KOUCHERYAVY, Y., ANDREEV, S., AND STAEHLE, D. Estimation of a successful beacon reception probability in vehicular ad-hoc

networks. In *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly* (New York, NY, USA, 2009), IWCMC 09, ACM, pp. 416–420.

[246] VO, T. *An investigation of bluetooth technology for measuring travel times on arterial roads: a case study on Spring street.* PhD thesis, Georgia Institute of Technology, 2011.

[247] VON BLOTTNITZ, H., AND CURRAN, M. A. A review of assessments conducted on bio-ethanol as a transportation fuel from a net energy, greenhouse gas, and environmental life cycle perspective. *Journal of Cleaner Production 15*, 7 (2007), 607 – 619.

[248] WANG, F., AND XU, Y. Estimating O–D travel time matrix by Google Maps API: implementation, advantages, and implications. *Annals of GIS 17*, 4 (2011), 199–209.

[249] WANG, L., YANG, L., GAO, Z., LI, S., AND ZHOU, X. Evacuation planning for disaster responses: A stochastic programming framework. *Transportation research part C: emerging technologies 69* (2016), 150–172.

[250] WARNER, H. W., ÖZKAN, T., AND LAJUNEN, T. Cross-cultural differences in drivers speed choice. *Accident Analysis & Prevention 41*, 4 (2009), 816–819.

[251] WEGENER, A., PIÓRKOWSKI, M., RAYA, M., HELLBRÜCK, H., FISCHER, S., AND HUBAUX, J.-P. TraCI: an interface for coupling road traffic and network simulators. In *Proceedings of the 11th communications and networking simulation symposium* (2008), ACM, pp. 155–163.

[252] WIERING, M., VEENEN, J. V., VREEKEN, J., AND KOOPMAN, A. Intelligent traffic light control, 2004.

[253] WOODS, R. J., MCBRIDE, S. K., WOTHERSPOON, L., BEAVEN, S., POTTER, S. H., JOHNSTON, D. M., WILSON, T., BRUNSDON, D.,

GRACE, E. S., BRACKLEY, H., ET AL. Science to emergency management response: Kaikoura earthquakes 2016. `https://ir.canterbury.ac.nz/handle/10092/14597`, 2017. [Online; accessed 20-Jun-2019].

[254] XIE, X.-F., SMITH, S. F., AND BARLOW, G. J. Schedule-driven coordination for real-time traffic network control. In *ICAPS* (2012).

[255] XU, X., CHEN, A., AND YANG, C. An optimization approach for deriving upper and lower bounds of transportation network vulnerability under simultaneous disruptions of multiple links. *Transportation Research Part C: Emerging Technologies* (2017).

[256] YAO, R.-H., WANG, D.-H., AND LI, L.-L. Revised Weibull revision model of headway distribution for motor-vehicle. *Journal of Jilin University (Engineering and Technology Edition) 39* (2009), 331–335.

[257] YOUNG, R. Managing traffic flow at airports - connecting the data. `https://www.beca.com/ignite-your-thinking/ignite-your-thinking / february-2018 / managing-traffic-flow-at-airports-connecting-the`. [Online; accessed 20-Jun-2019].

[258] YOUNG, S. Bluetooth traffic monitoring technology: Concept of operation deployment guidelines. *Center for Advanced Transportation Technology, University of Maryland, College Park* (2008).

[259] YOUNG, S. E., ET AL. Bluetooth traffic detectors for use as permanently installed travel time instruments. `https://bit.ly/33Uareh`, 2012. [Online; accessed 20-Jun-2019].

[260] YU, J., AND DONG, Y. Maximizing profit for vehicle routing under time and weight constraints. *International Journal of Production Economics 145*, 2 (2013), 573–583.

[261] YUAN, S., CHUN, S. A., SPINELLI, B., LIU, Y., ZHANG, H., AND ADAM, N. R. Traffic evacuation simulation based on multi-level driving decision model. *Transportation Research Part C: Emerging Technologies 78* (2017), 129–149.

[262] ZAIONTZ, C. Cronbach's alpha. http://www.real-statistics.com/reliability/cronbachs-alpha/. [Online; accessed 20-Jun-2019].

[263] ZHANG, Y., AND YAN, W. Research of traffic signal light intelligent control system based on microcontroller. In *Education Technology and Computer Science, 2009. ETCS'09. First International Workshop on* (2009), vol. 2, IEEE, pp. 301–303.

[264] ZHENG, P., JEFFERY, D., AND MCDONALD, M. Traffic operations from open to dedicated e-Lane. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on* (2010), IEEE, pp. 968–973.

[265] ZYRYANOV, V., AND FEOFILOVA, A. Simulation of evacuation route choice. *Transportation Research Procedia 20* (2017), 740–745.