The provision of real-time information for passengers in metro networks

Case studies: London and Hong Kong

Emily Digges La Touche

Supervised by Professor Nick Tyler

A thesis submitted for the degree of Doctor of Philosophy to the University College London

Centre of Transport Studies, Department of Civil, Environmental and Geomatic Engineering, University College London, September 2014

I, Emily Digges La Touche confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgments

I would like to thank Professor Nick Tyler, my supervisor, for all the guidance, reassurance and inspiration during this period of research.

I would like to thank Professor Benjamin Heydecker, my subsidiary supervisor, for your mathematical and statistical support.

I would like to thank Professor S.C. Wong for giving me the opportunity to spend 3 months living and working in Hong Kong and making the connection to the MTR on my behalf.

Finally I would like to thank the MTR and London Underground for providing data for this project and more specifically, Duncan Horne at London Underground and Henry Lo at MTR for answering my multitude of questions throughout the three years.

Abstract

This study looks at discovering information about the dynamics of a metro network, in real-time, using entry and exit data from the passengers' smart cards. The data shows to be a valuable source of information about the current conditions of the network for both operators and passengers.

An algorithm was developed which used real-time data to determine journey time characteristics, and to determine deviations from normal travel time and the extent to which these constitute a delay.

This study focuses on the London Underground network and the Hong Kong MTR network as case studies to test the algorithm using the data produced by the automated ticketing systems. It aims to mine the data to provide information that can be used by passengers of the network.

This information can lead to passengers knowing optimal routes, a realistic travel time and the number of minutes a delay may cost them; when the delay may be caused by congestion or service problems. Operationally this can allow for delay status reports to be more realistic, dynamic and responsive to crowding and provide information to the operators about the dynamics of the network in real-time.

Contents

Graphs

Tables

1. Introduction

The UN reported in 2009 that the world's population of those living in urban areas had overtaken those living in rural areas with 3.42 and 3.41 billion retrospectively and predicted that by 2050 the urban population of the world will have increased by 84% to 6.3 billion ("United Nations Population Division | Department of Economic and Social Affairs," n.d.). With this expected growth, pressure is put on the transport systems of the cities to keep the city moving.

In many cities expansion of the existing metro network may be complex and progress may be slow. Leading to optimisation of the current network being essential. All metro systems will have to address this issue, but in many cases, especially for systems characterised by old infrastructure, it is important that the operators improve network performance and utilise space because it will not always be possible to build the additional capacity to meet the need in a suitably short timeframe.

However, it is not just the operators that can improve utilisation of space, passengers' routing behaviour can be crucial in optimising the network by maximising flow and minimising delay. To understand how routing may improve performance of a metro network, other types of network can be considered in order to learn how they are being improved by routing. For example, the internet, another network where multiple commodities are being moved around with fixed origins and destinations, uses smart routing to maximise flow.

Optimising flow in the internet is a widely researched. This work falls under the topic of theoretical internet routing. The relevance of this is that the number of people connected to the internet is rapidly growing. In 2002, only approximately 9% of the world's population was connected to the internet whereas ten years later approximately 34% were connected ("Internet Growth Statistics - the Global Village Online," n.d.). In addition, the internet of things is growing at an astonishing rate. In 2003 there were approximately 500 million connected devices but this has grown to approximately 12.5 billion by 2010 (Evans, 2011). Nielsen's law states that bandwidth grows by 50% per year showing that the channel capacity is growing. Nevertheless it is still important to use the capacity as optimally as possible, hence the study of theoretical internet routing.

The internet is set up in a similar way to a transport network with nodes and links but instead of the flow being people, in the internet it is the packets of data. Packets of data are sent in datagrams using Transmission Control Protocol (TCP), Real-time Transport Protocol (RTP) or User Datagram Protocol (UDP) depending on the service needed. The different protocols can optimise either a high throughput or quality of service, trying to minimise loss and delay.

TCP gives accurate delivery of the packets, all packets get through and the rate at which they are transferred depends on the success rate of the delivery of the packets. When a packet is dropped, the success rate falls, but a new route is found so that the rate of packet delivery returns to a satisfactory level. In comparison, UDP is a less reliable form of packet transfer that ensures speed but does not necessarily ensure quality; it floods the data across the network hoping to get as much through as quickly as possible. Finally, RTP is mainly used for audio and video files, here speed of transfer essential to ensure real-time transfer whereas reliability is considered second.

A metro network is currently quite similar to UDP; without real-time information about the system dynamics people are flooding the network and when they realise there is a problem they reroute. With real-time information, a smarter form of routing may take place. People will learn information that will help them to change their travelling behaviour before they incur a delay and thus they could make use of underutilised routes i.e. moving closer to the TCP protocol. The difference between a metro network and the internet being that the packets are routed by these protocols. Whereas passengers have free will to choose their routes. However, with real-time information the hope is that passengers will choose the route that has the least congestion and all routes can be utilised.

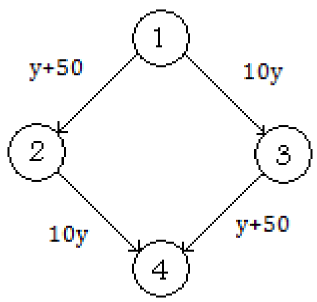
When considering re-routing it is difficult to achieve a stable network. For example, when a packet is dropped in the internet due to a path being congested, flow is moved from the congested path to an uncongested path. This brings instability as there is constant movement of flow. For example, in a simple network of two nodes and two links, if one of the routes experiences a drop then the data is rerouted onto the other link. To regain equilibrium a larger number of drops have to occur on the second link. The level of instability is proportional to the size of the network however; in a small network link failure has a greater impact than in a larger network (Wischik et al., 2009).

Braess' paradox (Braess, 1968) is a well-studied example of where in a road network, the route with the greatest utility to the passenger may not give the minimal travel time due to congestion. Further extending the network may cause a redistribution of traffic that causes passengers to have longer travel times. A famous example of when the network may benefit from an incident in the network is the case of when 42nd street was closed in New York. A normally highly congested street in New York was closed but instead of the network being devastated by the road closure, in fact congestion across the network improved ("What if They Closed 42d Street and Nobody Noticed? - New York Times," n.d.). This shows that it may be possible to improve conditions during incidents. The example below describes this condition.

Example 1

Consider the network below, where y is the number of vehicles.

Figure 1 – A network to demonstrate Braess' paradox



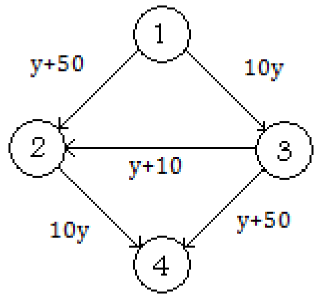
Where 1 is the start node and 4 is the end node. Let's say that we want to move 6 units through this network. The equilibrium is that 3 units go for 1-2-4 and 3 units go from 1-3-4 with overall cost:

$$C(x) = 6(y + 50) + 6(10y) = 66y + 300$$

Equation 1

Now if we add a new arc with cost y +10.

Figure 2 - A network to demonstrate Braess' paradox with added arc



This changes the equilibrium. We now have 2 units going from 1-2, 3-4, and 3-2 and 4 units going from 1-3 and 2-4.

This now gives us the overall cost:

$$C(x) = 4(y + 50) + 2(y + 10) + 8(10y) = 86y + 220$$

So adding this extra arc did not improve overall performance of the network, since the overall cost on the network has increased.

In order to avoid a condition such as Braess' paradox or instability due to sudden changes in the network, smart routing is needed. This is when passengers are continuously updating and adapting routing choices to respond to dynamic information about the network, with the aim to utilise the space in the network. In network routing, smart routing decisions are needed in order to use the space optimally in the network. Smart routing could be an important step in providing service to the growing number of individuals living in cities and assisting passengers of public transport networks to route themselves to their desired destination.

This area has been widely studied for vehicle drivers; in the paper by Aranaout et al. (Aranaout et al., 2010), the authors showed that, using their IntelliDrive system (which is a combination of a vehicle information sharing system and an agent-based model) traffic congestion was significantly reduced with real-time information. Dia used agent-based modelling on a real road network that experiences congestion, and used a behavioural survey to characterise individuals' preferences and choices to examine the impact that real-time congestion information on drivers stuck in traffic (Dia, 2002). Dia found this information had the potential to change passengers' behaviour, alleviate congestion and improve the performance of the network. However, there is a lack of research for how public transport users may respond to this sort of information.

An example of passengers receiving information about congestion and a change in passengers' routing behaviour was observed during the London Olympics 2012. Here, according to Transport for London, 63 per cent reduced their travel, 28 per cent changed the time of their journeys, 21 per cent changed route and 19 per cent changed mode during the London Olympics (Transport for London, 2013). During the Olympics there was an increase of approximately 20% in the number of journeys per day in the London Underground, with the largest ever number of journeys on the Underground, of about 4.6m on Tuesday 7th August, 2012. The network however did not experience any unusual delay due to this high passenger demand. This shows that a metro network can operate more efficiently, if passengers make smarter routing decisions. However, these smarter routing decisions require passengers to have thorough information about the network in order to make their routing decisions.

Within any transport network it is not possible at any given time to gain perfect information about the entire network, because despite the operational dynamics of the system being deterministic, there is a stochastic factor. The stochastic factor in the system is due to the volume and variability of passenger demand. In order to relieve some of the variability, information to passengers can help them self-manage the demand in the system. However, this depends when the information is received.

If the information is received before the passenger's route decision is made, then they have the opportunity to make a smart choice. If it is available after they have made and enacted their choice, it might maximise their regret. Therefore the aim is to improve the information currently available to passengers throughout their journey and to provide a service of information that is as up-to-date as possible.

Therefore, this project looks to answer the question: **Is it possible to give passengers of a metro network real-time information?**

When deciding how it might be possible to obtain dynamic information for passengers it is important that the information available to passengers must be reported in terms of the impact caused to the passengers. To understand what form of information passengers require a questionnaire will be conducted to obtain this information alongside other technical aspects of the work, seen in Section 4.4. However, it is found in this case, this would be the number of minutes of delay. Further, it is important that there are two types of information: congestion information and delay information. Congestion information determines the current state of the network, and describes the perturbations in the network caused by large passenger demand in bottlenecks; this can be relayed to individuals about their specific journeys. Delay information is the result other disruptions to the service, caused operationally. This is found by discovering what delays are incurred by all passengers on the line in question. Together, these two sources provide information to passengers in all situations that can help them avoid delays to their journeys. The necessity of providing information for the difference scenarios is discussed further in Section 3.

Having decided on the criteria of what information should be available, it is then a matter of determining the easiest source of this information. To find how long passengers are delayed there needs to be a method of tracking them to know where they are in the system at what time and therefore when they are delayed. This information can then be used to tell future passengers of the network what the current dynamics are. There are a number of different possibilities of doing this, such as, tracking cameras, tracking Bluetooth and tracking passenger movement through smart card data.

Tracking cameras are an expensive option; to be able to gain perspectives of all bottlenecks in the entire system would be expensive. Further, an algorithm would have to be created to determine what numerical delay corresponds to a queue; this would mean taking a visual image of a queue and determine from it how many passengers may be delayed as a result and by how much. However this would be difficult and potentially inaccurate leading to this option being discarded.

Next tracking Bluetooth has been seen in Rehrl's work on Personal Travel Companions (Rehrl et al., 2007) however it is thought, there would be a limited number of passengers that would have their Bluetooth switched on in any given metro, further a map of all the stations and the lines would need to be created and this is a very complex and lengthy process to create for all of metro tunnels, leading to this option not being possible.

Finally this left the decision that information about the current dynamics of the network should be found through smart card data. This data could then hopefully be used to provide information to passengers.

The rest of the thesis will tackle answering the previously stated question: Is it possible to give passengers of a metro network real-time information? This question will be answered by mining smart card data in such a way that it is possible to extract information about the network dynamics that can be used by passengers to make smart routing decisions. The next section, Section 2, will examine all relevant research that has been completed on closely related topics. Beyond this, a breakdown of the research question will be listed in Section 2 and Section 3 will look at determining a methodology to be used to answer the research questions. Sections 4 and 5 are case studies of the methodology. More specifically, Section 4 fine tunes the discussed methodology and Section 5 determines how transferrable the methodology is to a different metro in a different city. The success of the methodology is discussed in Section 6 and finally Section 7 concludes how well the research questions were answered.

2. Background research

This study aims to discover if it is possible to find information about the current dynamics of a metro network in order to relay such information to the passengers and operators. To gain a thorough understanding of whether it might be possible to discover valuable travel information for passengers though smart card data, it is necessary to know what other researchers have discovered smart card data to be capable of and what passengers find could be useful travel information. Therefore before embarking on analysis of smart card data a review will be completed on current research in the areas of smart card data and travel information.

2.1. Travel information

It is important to understand the information requirements of passengers in a metro system, to ensure the information provided to them may cater to their needs. To fully understand what passengers' prerequisites may be when it comes to travel information, it is useful to consider what information is currently available to them; what is classified as valuable information, how a passenger may use the information about their travel time and finally what influence providing information may have on the network performance. These areas will be investigated to provide a detailed understanding of what really contributes to providing useful information to passengers.

2.1.1. Currently available information

Before ascertaining what information is currently available to passengers it is worth discovering the progression that has brought us to the current information available.

The idea of passenger information is not new. In 1839 the first U.K. timetable was produced by George Bradshaw ("Information resources - London Transport Museum," n.d.). Since then the scheduled timetables for public transport services have been an essential part of planning and undertaking a journey. In the 1930s, Harry Beck's revolutionary map of the London Underground broke the connection with geographical layouts in maps in order to make them easier to understand ("Harry Beck's Tube map - Transport for London," n.d.).

There is then a gap in the evolution of travel information until 1974, when in Paris the European Broadcasting Union launched the Radio Data System. This was the first form of live travel news ("RDS, Radio Data System : Radio-Electronics.com," 2012) and is still available in cars today. In the early 90s the first satellite navigation system (SAT-NAV) was fitted into a BMW 7 Series car and could only be used in Germany. This was developed from the first satellite navigation system that was created by the U.S. military in the 1960's ("A Brief History of SATNAV," 2011).

In 1992 funding was given to the ROad MANagement System for Europe (ROMANSE) project. This project was run by Hampshire County Council with partners in the Public and Private sector. Their aim was to provide efficient management of the network in the hope of reducing congestion. They hoped to achieve this by developing an integrated intelligent control system that would provide real-time information. They provided a gating system that could control flow into and out of congested areas (SCOOT), an online data system that could provide network information of delays (ASTRID) and (STOPWATCH) which would provide a bus location and passenger information service (McDonald and Tarrant, 1994). Within a few years of the project starting they reduced delays by 60,000 vehicle hours per year. This project also produced TRIPlanner which provides information for public transport and private car users. The TRIPlanner was installed in 10 different locations in Hampshire, passengers would enter their origin and destination and a suitable route would be given. This gained 70 usages a day with on average, 47% being public transport users and 53% being private car owners (Wren and Jones, 1996).

Progress in travel information was generally slow until the birth of the internet. Nowadays in the U.K. 76% of adults are connected to the internet in their homes and over a quarter of adults and nearly half of all teens own a smartphone ("Ofcom | Facts & Figures," n.d.), therefore they have access to the internet while travelling. Not only has the internet meant that maps and timetables are more accessible but it has also led to the creation of journey planning services and real-time information.

This has led to the emergence of automated internet based route planning for public transport, for example, Google Transit ("Google Transit," n.d.) and the London Journey Planner ("English - Journey Planner - Transport for London," n.d.). The ever increasing popularity of smart phones has enabled public transport users to access online journey planners to assist their journey anytime and anywhere. Currently these (London Journey Planner and Google Transit) are based on timetabled data and do not respond to live information about the current public transport network status; they fail to report events such as accidents, congestion and service interruptions, leaving passengers uninformed. In addition, the information is inconsistent across different modes, for example, bus times are calculated by maximum journey time and tube times are calculated by average travel time (Transport for London, 2012).

Outside England there are some journey planning services that do provide real-time information, such as the Dutch 9292ov Journey Planner ("9292 reist met je mee," n.d.) and Yahoo, Route Selection in Japan ("Yahoo - Route Selection (路線情報 時刻表)," n.d.) which provides real-time information according to the operational status of the different modes of transport. Within London information about operational status is available on service boards within stations or on the TfL website, but not yet connected to the journey planners. Typically, the information is presented to passengers at the station where they enter the system. However, once a passenger has reached the station and realises there is a problem they may have reduced the set of possible alternative routes, meaning that for maximum benefit to both the passengers and the network, it is essential that this information is available to passenger at any point in their journey, including before they start.

It is apparent that advances in real-time information have developed quickly for modes of transport that operate above ground. This is due to the application of the global positioning system (GPS). Advances in GPS hardware and computer software have led to real-time travel information for buses becoming available. The iBus system, created in January 2006, uses Automatic Vehicle Location and radio data systems to give passengers up-to-date accurate real-time travel information ("iBus | Transport for London," n.d.). Several studies have been conducted that show that passengers' waiting time is reduced as a result of real-time bus information (e.g. Dziekan and Vermeulen, 2006, Schweiger, 2003). Specifically, the OneBusAway system (Watkins et al., 2011), in Seattle, that provides real-time next bus countdown is shown to reduce passengers waiting time by 2 minutes.

In the summer of 2011 Transport for London (TFL), the governing body responsible for most transport in the greater London region, made their journey planning data as well as other travel data available online through an application programming interface (API) data feed from a server which provides the information ("Home | Developers Area | Transport for London," n.d.). This has led to numerous travel planning applications becoming available to assist travellers and commuters in London ("London transport Apps - Android," n.d.).

Konstantinos et al (2010) completed an international survey of internet-based journey planning services and discovered that passengers felt that there was a lack of on-trip information, limited real-time journey data and too few travel alternatives. In general it is found that travellers have a general dislike of the lack of information available (Chorus et al., 2006). Further to this, (Harazeen, 2011), looked into the effect of information during service disruption and the decisions people make at these times. Harazeen found that nearly a quarter of the participants took no action based on travel news that they received via the TfL website as they felt that the information lacked necessary details. This shows that the delivery, relevance and accuracy of information are essential.

Creating real-time information is merely half the challenge. To see benefits to the network it is important to obtain the maximum number of passengers using the service. This might be achieved by considering how passengers want to receive information. It was established by Zografos et al. (2010) that the most important form of information sought by passengers is real-time information available on their phones as well as international journey planning and interurban information.

These results were collected through surveying individuals from 5 European countries as well as people residing in China.

The market has grown rapidly with smartphone applications covering a wide range from tube maps, to exit guidance, to hiring taxis and more, meaning that passengers can be more informed than ever before, providing they are aware of all the different sources of information. Naturally with more travel options available to a passenger the more information they require and the future of travel information will be personalised information that assists the passenger from origin to destination. This is currently available in static time in London via CityMapper ("Citymapper - The Ultimate Transport App - London, New York, Paris, Berlin, Washington DC, Boston," n.d.), yet it is not available in real-time, besides the information available about bus times. It seems that currently the work in this area looks to develop a tool that utilises current information already available and to provide a service that plans your trip from beginning to end using different information sources.

For example, WISETRIP is a project that aims to provide real-time information internationally that connects existing sources of travel information into one multi modal journey planner (Spitadakis and Fostieri, 2012). WISETRIP was expanded to the Enhanced WISETRIP project where more spatial ground was covered and advanced features were added such as re-planning facilities, services for disabled and elderly users, more details on walking segments and information about CO_2 emissions (Solar and Marques, 2012).

The utilisation of multiple information sources has led to multimodal route planning services being a popular area of interest, for example, the PATH2GO service in California (Zhang et al., 2011). The PATH2GO service includes information about real-time public transport information, parking information and information about traffic. Another approach is providing a directory of the different available information sources (Seng et al., 2012). However, due to this being a fairly modern area of research there is much space for improvement. For instance, improving the search algorithms used so that it is focused for the users' needs. Dibbelt et al. (2012) comment that many of these services just provide the shortest path, however this may not provide the passengers choice transport mode.

The work completed for the present thesis looks to improve the information currently available about metro networks. These results could then be taken and used as part of a multimodal journey planning service. Likewise, research has been conducted by (Haicong and Feng, 2012) to improve the information provided to passengers about their pedestrian movements, for the use of a multimodal journey planner.

As well as considering what information is currently available to passengers and what might be available to them in the future, to fully understand travel information it is important to know how it affects passengers perceptions.

2.1.2. The importance of information provision

This project aims to provide real-time information for passengers of a metro network. To deliver the information effectively it is important to consider what is useful to passengers' needs.

The different types of information that are available to passengers are static, dynamic and real-time. These different types of information can assist the passenger at different parts of their journey. Passengers can receive information at three different stages of a journey; pre-trip, wayside and on board. Pre-trip information helps the passenger plan the journey they want to take, choose their route and their departure time. Wayside information is information that travellers may pick up en-route such as announcements, changes to the time tables, directions and information obtained from other passengers. Finally, on board information is information that is gained when inside a public transport vehicle.

It is currently unknown how a journey planner may influence a passenger's journey. Maximising flow on the public transport network is essential for the future of transportation this may be possibly achieved by intelligent routing. However, Liu (1996) shows that humans do not perform well in finding the best route option between two given points on their own. Sun and Winter, (2013) argue that the level of familiarity to the network defines how much an individual will need a journey planning service.

For the purposes of this thesis, literature on travel times will be reviewed to understand how passengers respond when gaining new information about their journey. When thinking of how a passenger might respond to additional information when travelling, it is useful to consider how they currently think about their travel times. Mazloumi et al. (2011) found that it is hard for passengers to estimate their travel times as there is a great level of uncertainty, for example when a passenger uses a bus, the waiting times at the stop gives a lot of variability. It is anticipated that in this thesis some of the uncertainties within travelling in the Metro may be removed. If more accurate times are given, people might have more faith in the information they are gaining and respond more willingly to journey planners.

In general it is thought that an individual estimates their travel times through a process of learning, however the details of this differ in different papers. One assumption is that travellers update their estimation on the basis of different triggers (Chen and Mahmassani, 2004). Whereas another idea is that travellers update their travel times on a day to day basis (Jha and Mahmassani, 1998). A common assumption is that a traveller updates depending on the time difference between what they have perceived their travel time to be and what it actually was the day before (Axhausen et al., 1995), (van Berkum and vabb der Mede, 1998).

There is little literature in the area of how a passenger responds to travel time information given to them in advance or how a travel time estimation from a journey planner might be used to update their perceived travel time as a result. It is complicated to isolate the influence of journey planners on passengers' travel decisions from more general information provision for public transport users and this may be why a gap in the literature exists. It has, however, been well researched how maps influence passengers route choice. For example, when a map contains information about the headways between vehicles, passengers will utilise this information and use it to determine their departure time (Hochmair, 2009). It has been shown that passengers' route choice can be heavily influenced by the map provided to them, for example in the case of the London Underground the map used does not represent real life distances; this leads passengers to choose routes that may appear to be the shortest on the map, yet in real life are longer in travel times and a detour to their journey (Guo, 2011). Further, Cats et al. (2011), discovered by a model being developed in the Stockholm metro system, that real-time information has the potential to change a passenger's route, such that they can save time with their journey.

It is important to discover when and where a traveller will want to receive information about their journey. Rather intuitively a passenger will seek to find external information, that which is not from their memory, when the journey they are planning on taking is less habitual (Verplanken et al., 1998). Valuable information has been shown to reduce the uncertainty a passenger may face when it comes to deciding what route they should take and at what time they should depart (Bhat and Sardesai, 2005). In turn this can help to relieve stress and anxiety for the travellers (Bates et al., 2001). More recently Mendes Caiafa (2010) revealed that older passengers will react to situations that occur in real-time, showing a willingness to discover new information when they are en-route. This is reassuring that different age groups are willing to gain additional information to assist them when choosing a new route in times when things go unexpectedly wrong showing the importance of the provision of information.

Useful information can even help a passenger experience a more pleasurable trip, Balcombe et al. (2004) concluded that this will leave a passenger feeling more satisfied with the public transport they are using. It is currently unclear to what extent travel information would change a traveller's behaviour and route choice. However, Kenyon and Lyons (2003) noted it does have the potential to do so. Passenger information has been shown to have psychological effects on passengers' travelling experiences. For instance, information can provide passengers with a sense of security and reduce the anxiety, especially when travelling late at night (Schweiger, 2003). Smith et al. (1994) discovered that even if operational reliability were to decrease, real-time information can make passengers feel reduced stress which makes them believe the service is more reliable, Dziekan and Vermeulen (2004) found this is only the case however when passengers trust the information they are obtaining. Shah et al. (2001) showed that intelligent information for car users, when used before their trip, has managed to reduce the number of individuals arriving late at their destinations by 62%. Information even has the potential to reduce carbon emissions (Brazil and

Caulfield, 2013) and help with usability of the network for passengers with disabilities (Lamont et al., 2013).

It appears that information is incredibly useful to passengers when provided in a reliable way. Reliable information can help passengers who are new to the system and more familiar passengers. It can reduce passengers' stress, make them enjoy their travel experience, make them arrive at their destinations on time with more certainty and can even reduce their carbon emissions as well as helping those that are disabled. With so many positive outcomes from information provision to passengers it is left to determine if research has been conducted to understand how the network can benefit from travel information.

2.1.3. Influences on the network, case study: The SAT NAV

Intelligent information systems within the public transport network are still relatively new and although there are some studies that discuss the influence of this provision of information on passengers, there is little research about how either the network or passengers responds to this information.

It was established in Section 2.1.2 that there is good evidence that with reliable information passengers show a willingness to change their travel behaviour. Naturally this will lead to some outcome with the dynamics of the network changing. To get a fuller understanding of what are potential influences to the network a case study of the introduction of Advanced Traveller Information Systems (ATIS) within automated vehicles will be examined.

In order to understand how real-time information will affect a given network, the current stable conditions of the network need to be identified and understood. It is assumed that a traveller selects the path they will travel on by trying to minimise their perceived travel times (Daganzo and Sheffi, 1977), this leads to a stochastic user equilibrium. This model has since been extended to include variations to the network from day-to-day and route choice options and the possibility of travellers not making a trip at all (de Palma et al., 1983). The variations that arise day-to-day can be described by a Markovian model and this leads to the network settling in steady-state. Further to this, Cascetta and

Cantrella (1991), used a stochastic approach to model the day-to-day dynamics which then included the potential differences that could occur within a day. Linking the two models a Markovian assignment model will converge to a dynamic deterministic model as the users' perceived cost becomes increasingly deterministic (Watling, 2003).

In the early 90's a number of research projects were conducted in laboratories to understand the how ATIS might change passengers' decision making process. These looked at different aspects of the information available and tried to determine how the information changes passengers' route choice decisions and the network(Koutsopoulos et al., 1994) (Adler et al., 1993) (Bonsall and Parry, 1991). However, the results found in these projects are constrained by being produced from simulation rather than real life experience that might include factors that had not been included in the simulation models but that could have influenced passengers' decisions.

In Japan some results have been obtained outside the laboratory; the CACS project was carried out over 6 years in Japan although this was well before commercial 'SAT NAVs' (the Agency of Industrial Science and Technology bought out a system that helped passengers plan their route in 1973). Fujii (1989) discovered as a result of this information that passengers' travel times could be reduced between 9 and 15%. Kobayashi (1979) learnt that the whole system might benefit from real-time information with a reduction in overall travel times of up to 6%.

When analysing the time saved during regular congestion for drivers using ATIS compared to those who do not use it appears to be negligible. However, when looking at incident reports Al-Deek et al. (1989) discovered that passengers could save up to 25% of their travel times. Al-Deek and Kanafani (1993) showed that ATIS is most useful at off-peak times when passengers can re-route to uncongested alternatives; during peak times it is more important to encourage passengers to depart at different times to spread the congestion.

Most of the research discussed about ATIS for vehicles so far has concentrated on the potential positives for passengers. However, each of these models focuses on a limited number of vehicles having real-time information. It has not been considered how the dynamics of the network as a whole might alter due to the information. When some passengers, with real-time information, choose to change their behaviour this could cause a number of other passengers to follow suit (Halbing et al., 1997). This leads to the discussion that in fact there is a delicate threshold for the number of passengers who should have real-time information as a proportion of the whole number of passengers, because if all passengers were equally well-informed, it could lead to everyone being worse off (Tsuji et al., 1985) (Arnott, 1991).

This review of the ATIS for vehicles has shown that given the information is delivered correctly then the network could see passengers spending less time in the system leading to additional space and a reduction in congestion . Next this review will focus on the current research being produced using smart card data.

2.2. Smart Card Data

In this project it has been decided to look into the data produced by the Oyster card in London and the Octopus card in Hong Kong, this decision will be discussed later in Section 3. It appears that there is very little research completed to date using Octopus data, which could be due to access to the data being very restricted. However, many projects have been completed using the Oyster data in London, as well as other smart card data from around the world which is reviewed below.

The Oyster card is an automated fare collection system in London; the system covers the entire Transport for London (TfL) networks in Greater London, including trains, metro, buses and boats. A direct debit option means that a passenger’s journey could eliminate queuing for tickets and just ‘hop on and hop off’. In April 2012 90% of passengers on the London Underground and 80% of bus passengers were using Oyster Cards (Transport for London, 2012). The ticketing system in the London Underground is such that every passenger must pass a ticket barrier at every entrance and exit, thus the system is closed and provides timed evidence of each passenger’s entry into, and exit from the system. This means rich data about journeys completed in the metro has become available.

In Hong Kong 95% of passengers use the Octopus card to pay for their journeys in the metro system (MTR corporation, 2013). The system in Hong Kong is similar to that seen in London, such that every passenger must pass a ticket barrier at every entrance and exit. The use of the Octopus card has been extended such that it can be used to pay for items in convenience stores, supermarkets, cafes and restaurants as well as all forms of transport in the city ("Get Your Octopus - Octopus Hong Kong," n.d.). For an extensive review about smart cards and their uses see (Blyth, 2004) and (Bagchi and White, 2005).

Oyster data is becoming a popular data source with universities from around the world using the data. Seaborn et al. (2009) used Oyster data to discover complete multi-modal journeys in London. This has led to discovering how many travellers are completing multi-modal journeys in the city and of what type. This discovery leads to TfL knowing 'on an average day' how many passengers use which forms of transport and in what order. If there were to be a delay to a part of the network then, in comparison, it could be seen how the number of passengers choose to reroute themselves. This could lead to a change in the design of stations where common interchanges happen and analysis of rerouting decisions across modes during disruptions. Chan (2007) determined, from Oyster data, that only 46-62% of the time that passengers are in the metro system in London is spent riding on a train, showing that a large proportion of their time in the system is spent walking or waiting within stations, this result is seen later in the analysis through the difference seen when travel times discovered through smart card data are compared to the travel times given by the London journey planner, where the later only accounts for the time the passenger is spent riding the train, Section 4.3.

Other research using Oyster data includes personalising information for passengers (Lathia et al., 2010). Lathia looks at how many trips an individual is taking over a time period, their travel times and similarities between different user groups. This makes it easier to provide information which is based on knowledge of what might be useful to individual passengers. In the future this could be used in the form of information given to passengers, so that the information is personalised to the individual depending on their previous journeys. However, this leaves a gap for passengers that have only just bought a smart card or are completing a new journey and for tourists who have no previous travel history on the system, so base information for all to use is still important. In addition the work on personalised information does not include real-time information, and this leaves the provision of real-time information – a crucial step into passengers gaining more individualised information – uncovered by this thesis, with result seen in Sections 4.5, 4.6, 5.5 and 5.6.

Zhao et al. (2013) looked at the difference between the time it took passengers to complete journeys on the Overground (the above ground train services) in London, by analysing their journey times from Oyster data and estimating which train they boarded. This research aimed to discover the difference between passenger arrival times and scheduled time tables to produce Estimated Journey Time (EJT) as a way of describing the network dynamics that is suitable to passengers and operators.

Guo and Wilson (2011) created a model that predicted what route a passenger most likely took in the London Underground. This was then used alongside analysing Oyster data to produce cost-benefit analysis of changing to another metro line during a passengers' journey.

So far this section has concentrated on the research that has been completed using the Oyster data produced by the smart card ticketing system in London. There are many metro systems around the world that have smart ticketing systems and are producing data that can be analysed to gain information about the metro systems and their passengers' movements.

In Singapore, metro data has been inputted into an agent-based model to determine when there is congestion in the network. This work is aimed to look at bottlenecks over time. This work has the disadvantage that a lot of contextual information about the system is needed in order to use the model; such as train schedules, walking times and station layouts. The thought for this work, in Singapore, would be that it would be useful for special occasions such as New Year's eve, where the data could be taken from the previous year, modelled, and then used to predict what will happen in the network (Othman et al., 2014).

In Seoul, smart card data has been used to understand the travel patterns of the elderly. The research showed that elderly people in Seoul spread their transit use between 9am to 5pm during the weekday which is the opposite pattern to those that are younger. Further, they have noticed that the average elderly passenger tends to transfer across metro lines less than the younger generations. They hope that this research will help transit planners make the city more accessible to the elderly (Eom and Sung, 2011).

In the Netherlands, research using smart card data and agent based modelling also took place (Bourman et al., 2012). They used the model to determine if it would be possible to shift passengers' travel behaviours to reduce congestion. They saw that by offering discounts for passengers travelling off-peak the model showed that congestion in peak times may be reduced. However, the overall revenue decreases. Further, van der Hurk et al. (2012) used smart card data in the Netherlands to forecast demand in the network. Here, van der Hurk used the smart card data to create time series which were then used for the demand modelling. They found that different passenger types have different travel patterns and demand distribution is dependent on the day of the week. The hope for this work was to help inform passengers and operators of how a delay may develop, yet was not completed in this work so may be seen in future work.

In Japan, smart card data has been used to determine what route a passenger may have taken and what train they boarded, the aim for this is to see if trends exist over time that the operators can use to change the current timetable (Kusakabe et al., 2010).

In Santiago, Chile, the smart card data produced does not contain information about where the passenger finished their journey therefore a model has been created that follows passengers travel patterns over time and estimates the time and position of the end of the passengers journey for over 80% of the journeys completed (Munizaga and Palma, 2012).

It is clear that around the world there is much more interest in recent years concerning the data produced from smart card ticketing machines and that the rate of progress in different countries depends on the availability of data sources and the type of information stored by the card. The objective of reviewing current research was to discover what other researchers are studying concerning smart card data and what passengers believe is useful information. Through analysing the literature it appears that although journey planning services for personal navigation on personal transport is gaining much more interest in academia, there is still much more work to be done in this field before passengers can be sure they have taken the route that will leave them with the least regret.

This section has seen a review of the progress of information provision for travellers showing how far information has come since the days that only a map was available. Today passengers can use multiple sources of information teamed with their smart phones and receive static information at any point of their journey, provided there is internet service.

Due to advances in GPS technology it is apparent that information provision is progressing faster on rail lines that it is in metros. This means that there is reduced information about the current conditions of the network in the metro. However, by reviewing the advances in information available to bus, train and car users this means mistakes that were previously made with the advances in information for these modes might be avoided with the metro.

Development of information provision is at many different stages. Some academics look to provide more information such as developing multimodal journey planners whereas others try and improve the information that is currently available. The aim for this thesis is provide passengers of a metro network real-time information. It was shown in the review how important travel information can be to passengers. It can relieve stress or anxiety, help plan unknown routes, reduces emissions and help elderly and disabled people by making the system more accessible.

A case study of the progression of Advanced Traveller Information Systems (ATIS) was examined to understand how travel information may influence the network. It was found with the right information it is possible to improve the experience to passengers as well as improving the network dynamics.

Finally a review of the research currently completed using smart card data was written, as seen in Section 2.2. This showed that a number of researchers are analysing the data for different means. The progress of research is at different stages for different data sets around the world. It is clear however there is little research completed about how passengers may be affected by real-time information in a metro system and the discovery of real-time information in a metro is still not available.

This work plans to provide passengers with relevant information about a metro network by using the data produced by a smart card ticketing machine. If this can be achieved this can lead to the operators gaining insight into the dynamics of the network. In the sense, currently the operators know the operational dynamics of the network but they do not have much information the dynamics of the network in terms of passengers movement and where in the network there are common problems and places that cannot easily handle high frequency of passengers this is discussed in Section 6.4 and 7.3.

The review seen in this section has helped direct an answer to the question: Is it possible to give passengers of a metro network real-time information? This will be broken down into smaller questions in order to be answered:

1. Is there information available about the dynamics of the network in smart card data?
2. Is it possible that this information can be extracted to be useable and reliable to passengers?
3. Is the information found useful to passengers or operators?

The following section will create a methodology for these questions to be answered.

3. Methodology

This project looks to answer the three questions:

1. Is there information available about the dynamics of the network in smart card data?
2. Is it possible that this information can be extracted to be useable and reliable to passengers?
3. Is the information found useful to passengers or operators?

Which will in turn answer the main research question:

Is it possible to give passengers of a metro network real-time information?

Two metro networks will be studied to answer these research questions. This has been decided so that the methodology can be finalised in one city and tested in the second, this will show whether the methodology is transferable to another city.

The two cities chosen for this project are London and Hong Kong. London has been chosen as it is a city in which 90% of passengers in the metro and 80% of bus passengers were using Oyster Cards (in April 2012), (Transport for London, 2012). This means that as such a large proportion of passengers are travelling using the Oyster card, there is comprehensive data produced about the network. In addition the ticketing system in the London Underground is such that every passenger must pass a ticket barrier at every entrance and exit. This has led to rich data about journeys completed in the metro becoming available, and this is vital to the work of this project. Hong Kong has been chosen as the second case study for many of the same reasons as London, in the Hong Kong metro 95% of passengers use an Octopus card, (“MTR: Our pledge for service 2013,” 2013). London will be the first city to be analysed, since this is where the research is based.

To answer the research questions an algorithm will be produced to determine ‘Is there information available about the dynamics of the network in smart card data?’ and ‘Is it possible that this information can be extracted to be useable and reliable to passengers?’

This algorithm, when finalised, will be tested to see what information can be found in the Octopus data in Hong Kong, this will discover how transferable the methodology is across two cities.

Finally after analysis has been completed in both cities and a successful algorithm has been created, the algorithm will be reviewed and the last question: 'Is the information found useful to passengers or operators?' will be answered by concluding on what information has been recovered from the algorithm.

3.1. Criteria needed for the algorithm

To determine what is needed to answer the research questions successfully, each question will be examined to discover what will be need to be considered in the algorithm.

1. Is there information available about the dynamics of the network in smart card data?

To determine what is happening in the network, in real-time, an algorithm will be created that will takes the smart card data from the ticket machines and mines the data to see if it possible to understand what is happening to the passengers. To do this the algorithm will need to take the raw data, convert it to a usable format and determine how live the information obtained is to see if it can provide information about the current dynamics.

2. Is it possible that this information can be extracted to be useable and reliable to passengers?

It was seen in Section 2.1.2 that it is important to provide reliable information to passengers so that the information is trustworthy and therefore used to make routing decisions with. To be able to make the information usable and reliable for passengers it is important that the information reflects the current conditions of the network. Further it is important to minimise the number of reports that may wrongly report the conditions of the network. Therefore the algorithm will need to include a way of making the reportings as smooth as possible and a method that attempts to remove false reportings, so that passengers can trust the information.

3. Is the information found useful to passengers or operators?

Finally, to answer this question the algorithm needs to determine what information can be used by passengers, to know what information passengers want it may be necessary to complete some work on surveying passengers. For the information to be useful to passengers it should be able to provide them with additional information, than they currently have about the network that can help them make travel decisions, when there is congestion and operational delays. For operators to find this information useful it should provide them additional information about dynamics of the network, then they currently have about the network, this may be achieved by providing information about passengers' whereabouts in the network.

3.2. Developing the algorithm

The different criteria, discussed in Section 3.1, needed to answer the research questions have been summarised below into a list that needs to be included in the process of creating the algorithm.

The algorithm must include the following processes:

1. Take the raw data and make it a useable format
2. Determine how quickly the information can be returned and determine operational and congestion delays
3. Smooth the data as much as possible to reduce noise and false reportings of delays should be minimal
4. Provide additional information to passengers regarding their journey and provide additional information to operators about the dynamics of the network

Considering the list above the following algorithm was created that could take the smart card data and determine if real-time information is available. This algorithm will include the list above to answer the research questions.

3.2.1. The Algorithm

The algorithm has been split into six sections, listed below:

1. Data collection
2. Average travel times
3. Regression analysis
4. What is a delay?
5. Congestion reporting
6. Delay reporting

These sections were created to cover the criteria listed in Section 3.1. Below the work that will be completed in each section is detailed.

3.2.1.1. Data collection

Different smart card systems contain different stored information. The London and Hong Kong metro systems have been chosen because they are closed networks; this means a ticket must be used at the beginning and end of each journey. This allows for origin-destination pairs to be determined from the data by matching the card numbers. In addition, to the station code being registered, the time that the passenger enters and exits the network is stored; this means that the journey length and duration can be calculated.

The London Underground has 11 lines and 268 stations, whereas the Hong Kong metro has 10 lines and 84 stations. Everyday approximately 3.5 million and 4.43 million passengers use the London and Hong Kong metros respectively ("Hong Kong: The Facts," 2014) ("London Councils: London Key Facts," 2014). For this project, one line will be studied in each network. This is in order to extract and provide maximum quality information rather than using a large amount of data that will contain a lot of noise. The lines chosen for the analysis are chosen such that they do not contain any loops or splits in the track, this is discussed further in sections 4.1 and 5.1 and the lines chosen stated. A number of different days spanning a few months will be analysed, this is conditional on the data provided by the two supporting operators.

To determine later, if congestion has arisen, it is essential to have some way of deciding what is actually meant by the term “congestion”. It is not possible to know accurately how many passengers there are in each station at any given time as it is not known what route a passenger has taken. This makes it difficult to know in terms of numbers where passengers might be in the network and therefore it is hard to determine if there is congestion due to high demand in a certain place. Therefore, variability in travel times will be used to determine the dynamics of the network. This will look at how passengers’ travel times change at different times of day and try to infer whether a passenger has experienced a delay to their journey due to congestion. Therefore a travel time needs to be defined, which is considered not to be delayed, in order to identify when a travel time is delayed. This leads onto the next section ‘Average Travel Times’; to know what the dynamics of the network are, and determine they are out of the ordinary, what defines ‘ordinary’ conditions need to be determined.

3.2.1.2. Average Travel Times

To understand fully what is happening in the network at any given time, it is necessary to know what the network looks like on an ‘average day’ where an average day is a day with no reported delays. To understand when there is congestion in the network, perturbations to passengers travel times will be analysed. These can be identified as perturbations compared with the average travel time determined. Hence, finding an average time would define a base point with which comparisons could be made. It would therefore be necessary to find all the travel times of the possible OD pairs on the metro line in question, the data will be aggregated to only contain journeys whose origin and destination is on the same line, this decision is discussed in Section 4.1.

The average will be taken rather than the median or mode as this time will take into consideration all passengers travelling on the metro. For example, at a peak time, the average commuter may be able to make their journey faster than someone who is new to the network, since this information will be used to relay information back to all passenger, it is necessary that it can be used by all passengers. With the average, passengers that are familiar to the network and the information can estimate how they compare to the values given, whereas passengers that are new to the network or the information can be provided with a time that caters to the fast and slow and familiar and unfamiliar. This average will also be used in later work as a comparator to determine whether a journey is delayed or not. These mean values will be used rather than the mode since the mode may be overly sensitive to passengers who are new to the network as the mode is likely to be based on those familiar to the network as they are the largest majority using the network. Further a mean will be used rather than the scheduled time or the current journey planner as it is taken from the same data source and therefore reasonable comparisons can be made. With the data provided for the London Underground only the morning peak is analysed, due to the availability of the data, therefore it is only the morning peak that is averaged. The data provided for the Hong Kong metro is throughout the day, therefore, it is questioned whether the average should be taken for different times of day or for the day overall, this is seen in Section 5.2.

3.2.1.3. Regression Analysis

Once average times have been found, these can be compared to the respective journey planners. Comparing the two would not only give validation that the average times provide a good representation of the network but would also lead to understanding what information is currently available to passengers and how realistic it is in relation to the actual network performance.

3.2.1.4. What is a delay?

The average travel times will be compared against journeys completed in real-time. However, in order to determine the current conditions of the network, the algorithm will need a decision variable that classifies a delay. This decision variable can be used to compare the average travel times found with real-time journey times to determine whether or not a passenger is delayed. For this, a numerical value will be needed that can be added to the average travel times to act as a threshold for classifying a delay.

This threshold will define a delay by being added to the average travel times, then compared to real-time data, if these times are over this value they will be classified as delayed. This is explained in Equation 2 and Equation 3.

If $\mu_{ab} + \tau \geq \theta_{ab}$ then the travel time is classified as un-delayed

Equation 2

If $\mu_{ab} + \tau < \theta_{ab}$ then the travel time is classified as delayed

Equation 3

Where μ_{ab}= the average travel time for an o-d pair *a* → *b*

θ_{ab}= an smart card travel time for an o-d pair *a* → *b* with time stamp *t.* The value *t* is found from the time exit *b* was recorded.

Finally, τ= is the delay threshold yet to be determined.

In this thesis the term delay does not necessarily refer to an operational delay. It is used to classify journeys which have breached the threshold described above. The cause of the breach could be congestion, and operational problem or perhaps just due to slow or ill passengers. Determining the cause of the delay is discussed later in section 3.2.1.5.

In London there is no numerical value in place to determine a delay in the network; instead the operators classify a delay as Minor, Major or Severe. These statuses are decided by managers on the basis of the four following factors: the time headways between trains, the speed at which the train is moving, the length of dwell times and the number of trains running on a particular track. However, in Hong Kong, there is a strict numerical value. If the service is delayed by more than 5 minutes it is classified as delayed and a report will be given to passengers. This would make a good threshold of defining a delay since a comparison can then take place between what is happening operationally in comparison to what is happening to the passengers, the decision to take this value is discussed further in Section 5.4.

In London, however, a numerical delay threshold will need to be determined. This provides an opportunity to discover what passengers believe a delay is. For this, a passenger questionnaire will be developed. To obtain information from passengers that is relevant to the specific service in the London Underground the question asked should be related to the service statuses. From this, the time passengers believe represents a delay can be used as the threshold for defining a delay. During this process it could be examined whether the length of journey correlates to the passengers' tolerance to delay. Further, it could question what information passengers would want at different stages in their journey and whether additional information might change their behaviour; this work can be seen in Section 4.4. This will also provide an insight into whether passengers feel they have a lack of information about the system and whether they feel that with improved information their routing choice might change. In order to generalise the findings to all users of the Underground in London a large sample will be needed that contains a representational proportion of the public.

Once a numerical delay threshold has been determined for London, both the threshold in Hong Kong of 5 minutes and the discovered threshold in London can be used to establish when there are delays incurred to passengers in the network, this decision is discussed in Sections 4.4 and 5.4.

3.2.1.5. Congestion reporting

To be able to give dynamic information about the current conditions of the network it is necessary that when looking at the data it should be viewed in the same manner that live data would be. At this stage the Oyster data and Octopus data will need to simulate live data that is returning straight from the ticket barriers in real-time, because, currently, live streams of data from the ticket barriers are not available. The data at present is near to real time but there is a time lag from up to a few minutes (in most cases) to as great as 15 minutes/half hour before a data entry is received by the central system. However the thought is that this could be improved and that real-time data will be available over the next few years. So the data in this project will be simulated to be real-time data. This is done by using the time stamps provided within the Oyster and Octopus data to know when the data was received and to simulate a working metro line. Further, information about a journey should only be considered after the passenger has exited the system since this is when the information would be received and an entrance and exit stamp can be paired.

Congestion reporting will focus on finding delays to passengers that are caused by high passenger demand in places with limited capacity.

On a day that has no reported service problems, congestion can affect passengers as they enter a station; and have to wait for a second train as either the train is too full or the queue is too long on the platform, this shall be known in this work as an entrance delay. Next congestion can affect passengers as they exit a station due to high passenger demand a queue is forming at the ticket barriers making it take longer to exit , this shall be known as an exit delay. High passenger demand may also affect the train scheduling by increasing dwell times which a long side operational delays will be known as line delays. The reporting of service delays on days with not reported service problems is discussed in Section 4.5.1. This leaves the focus of discovering congestion, within the stations, as either delays to passengers trying to enter or exit.

Entrances and exits to the network are logged separately and then paired to make a journey. Information can only be gained after the passenger has exited the system, since the aim is to compare average journey times with journey times in 'real-time', the comparison is completed and a delay is potentially discovered at the time of exit. A passenger can be delayed in the network entering the system, while on board a train and when exiting the station. Since it is not known where a passenger is between entering and exiting the ticket barriers to discover where a passenger is experiencing a delay in the network, information about other passengers travelling in the network is needed. If a number of passengers appear to be experiencing a delay and have a part of their journey in common, it can be deduced that this may be the part of the passenger's journey that is delayed, this is shown visually below in Example 2. The requirement of the number of passengers experiencing the same delay is determined in Section 4.5 for London and Section 5.5 for Hong Kong.

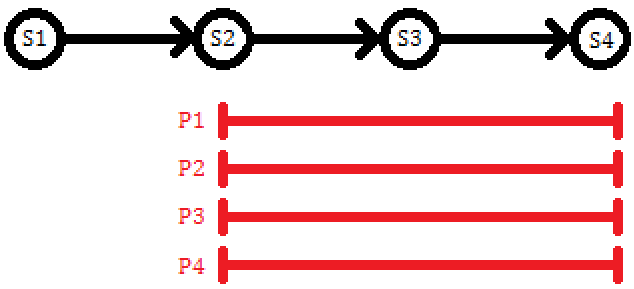
Example 2

Consider a simple network with 4 nodes (S1-S4) and 3 links that only travel in one direction, with 4 passengers (P1-P4) completing journeys on the network. The nodes represent stations and the links are the line the train is travelling along, shown in Figure 3. 4 scenarios are discussed below as examples of how to determine where in a network a delay is taking place.

Figure 3 – An example to determine a delay

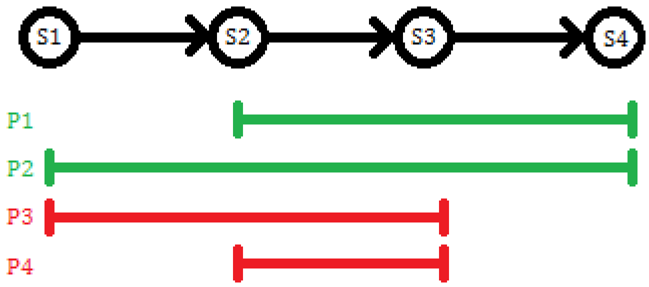
Scenario 1: If all passengers complete the same journey and all are classified as delayed, seen in red, a decision in regards to where the delay took place is unable to be made. Figure 4 shows passengers 1, 2, 3 and 4 they are all starting their journeys at station 2 and ending their journeys at station 4.

Figure 4 - An example to determine a delay: no information



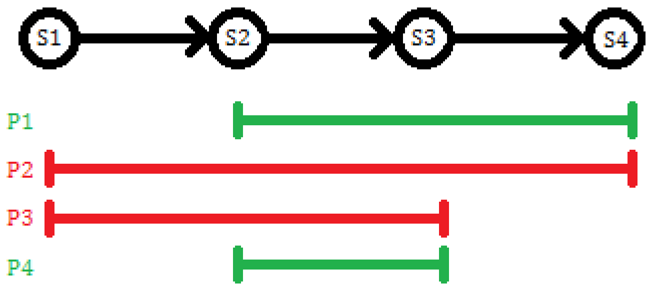
Scenario 2: If the passengers are taking different journeys and two are delayed and two are not delayed, shown in green, for un-delayed. It is possible to determine where the delay is taking place, Figure 5 shows that both of the passengers that are delayed share the same exit station therefore it is deduced they were delayed when exiting the station.

Figure 5 - An example to determine a delay: an exit delay



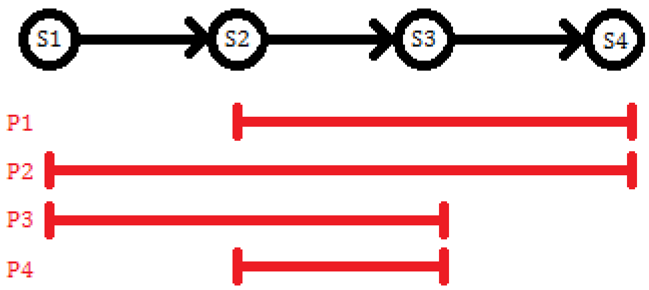
Scenario 3: In Figure 6 it can be seen that the two passengers that are delayed, passenger 2 and passenger 3, share the same entrance station. Therefore, it is decided that the delay is incurred to the passengers when they enter the station.

Figure 6 - An example to determine a delay: an entrance delay



Scenario 4: Finally, Figure 7 shows all passengers are delayed, but since they are completing different journeys it is decided that the line they are travelling on is delayed.

Figure 7 - An example to determine a delay: a line delay



This concept will be used to discover delays to passengers at entrances and exits. When the network is build up and more lines are included in the analysis a more through picture can be built up of the stations that have multiple interchanges. As said, the information about a journey is received only when the passenger has exited the system, entrance congestion can only be seen in hindsight, and this is discussed further in Section 4.5.

3.2.1.6. Delay reporting

Beyond understanding how passengers are affected during peak times with high demand on un-delayed days, analysis can move forward to determining if operational delays to the network can be discovered through the data. This analysis will consist of taking different days which had operational delays taking place which would affect passengers.

This work will start by studying the London Underground network. Once a stable algorithm has been determined for the London case the same mechanisms will be used with the Hong Kong smart card data. The prime objectives in both cases will be to determine how much information about the dynamics of the network is available through the smart card data and how much of this information can be returned to passengers. To be able to inform passengers successfully about the network dynamics, the information returned will need to be stable and consistent and available in a timely manner. This will mean that during the process of analysing the data, steps will be taken to ensure there are few false reports. The data at this stage will also be analysed to determine how quickly information about the network can be returned. This will be achieved by comparing the operational reports of delays with the time at which it is first noticed passengers are delayed.

4. London

In Section 1 the main research question was defined. This asked: Is it possible to give passengers of a metro network real-time information?

This section aims to take the methodology defined in the previous section and use the theoretical algorithm to apply it to the raw Oyster card data produced from the London Underground ticketing system and generate information for passengers and operators about the dynamics of the network.

In order to complete this, a systematic review will be undertaken to determine what an ‘average un-delayed day’ looks like, in the London Underground, on the line in question. Followed by determining what information can be found from data returned from the ticket barriers about congestion and delays and finally understanding how passengers’ travel times are affected when there are delays to the service.

This section will be organised along the lines described in Section 3.2.1– where the algorithm was described in six steps:

1. The data
2. Average travel times
3. Regression analysis
4. What is a delay?
5. Congestion reporting
6. Delay reporting

4.1. Data collection

As discussed in Section 3.2.1, it is essential to know what an un-delayed service looks like. To do this, variability in the network was analysed by looking at passenger travel times and seeing how these change during incidents. To be able to measure variability in the network there will need to be a comparison between an average travel time which has no delay and a travel time with a delay.

A database was obtained from London Underground with all origin – destination pairs on the Victoria Line produced from Oyster card data. Here, London Underground took Oyster data and found travel card pairs, and matched journeys by finding entrance and exit pairs. The data was aggregated such that only journeys with both their origin and destination stations on the line in question were kept. It was chosen to study data on only one metro line as an example of what may be possible using smart card data since the network is large so a large amount of data is produced. The Victoria Line was chosen because it contains no splits or loops, which could give rise to complexities such as conflicted entrance and exit pairs, where the same pair could be reached by different routes (with different travel times). This can be seen to be true for some of the other lines (e.g. Northern Line, Central Line). This analysis can be extended for these cases but for simplicity of the first case study it was chosen to analyse the simplest option.

Journeys that may have either their origin or destination on the line, but not both, were not included in the data set as it is unknown in some cases what route the passenger may have taken. Only those with their origin and destination on the Victoria Line were used.

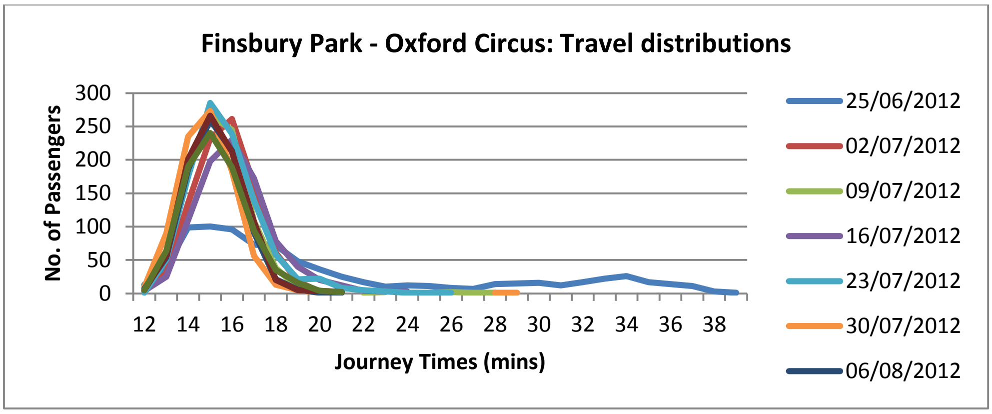
Seen in Table 1 the dataset has a column for date, entry code and entry name, exit code and exit station, the journey times in minutes and the number of people that took the journey in that time. Other data that is also available through Oyster data is the type of ticket used, i.e. freedom pass, 7 day travel card, one month travel card etc. however this information was deemed irrelevant to the purpose of this study and therefore was removed.

Table 1 – Example of Oyster data dataset spanning 8 weeks

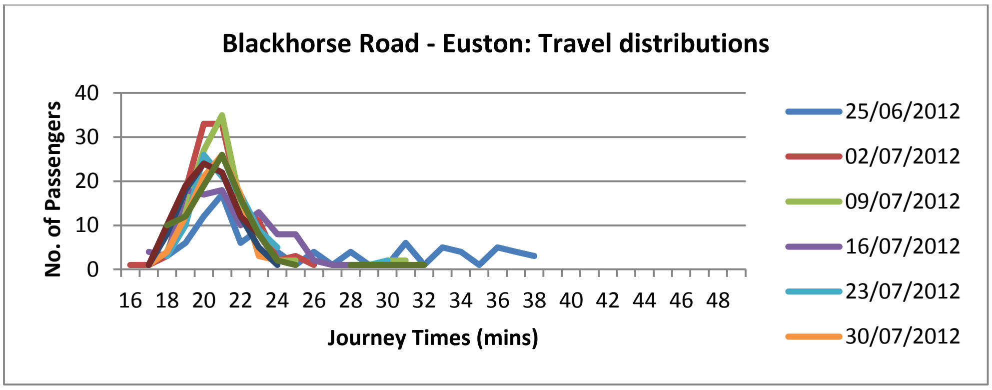
Date	Entry Code	Entry Station	Exit Code	Exit Station	Journey Time	Journeys
25/06/2012	522	Blackhorse Road	574	Euston LU	18	3
25/06/2012	522	Blackhorse Road	574	Euston LU	19	6
25/06/2012	522	Blackhorse Road	574	Euston LU	20	12
25/06/2012	522	Blackhorse Road	574	Euston LU	21	17
25/06/2012	522	Blackhorse Road	574	Euston LU	22	6
25/06/2012	522	Blackhorse Road	574	Euston LU	23	9
25/06/2012	522	Blackhorse Road	574	Euston LU	24	4

The Oyster data file contained 12,007 journeys completed on the Victoria line. The data made available for this work by London Underground in this file consisted of 7 days spanning 2 months. All journeys completed over the 7 days were in the AM peak (6:15-10:30) and on weekdays, due to the availability of the data from London Underground and the regular pattern exhibited during the weekdays (Lathia et al., 2010).

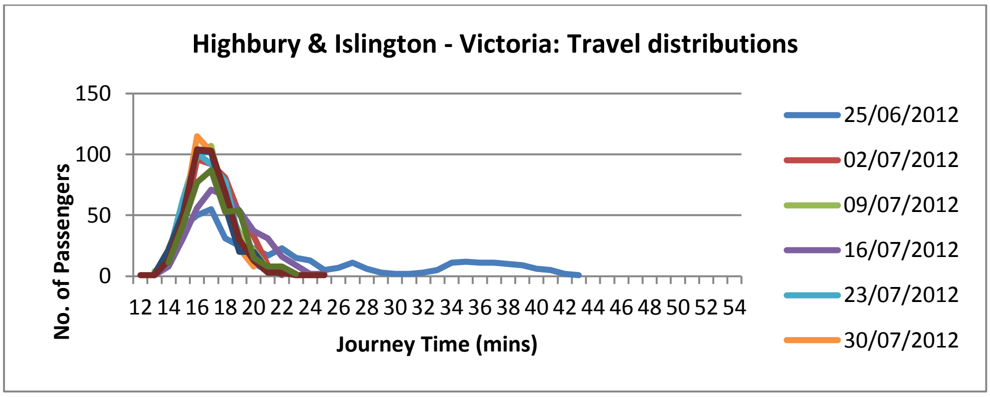
The file was organised by entry and exit station pair followed by journey times then by dates. Graph 1, Graph 2 and Graph 3 are three examples of the different travel time distributions over the 7 different days, they show frequency over time for all passengers completing the journeys over the 7 days.



Graph 1 – Finsbury Park to Oxford Circus: Travel distributions of 7 days



Graph 2 – Blackhorse Road to Euston: Travel distributions of 7 days



Graph 3 – Highbury and Islington to Victoria: Travel distribution of 7 days

The date of the 25th June has an apparent different distribution to the other dates shown in Graph 1, Graph 2 and Graph 3.

The distribution of the 25th June in all cases does not reach as high a frequency at the peak and the tail of the distribution is longer. As this is the same in each example it is reasonable to assume that a delay has occurred on this day, however the means and standard deviations of each day will be review to confirm this. Table 2 shows the mean and standard deviation of each of the distributions for the journeys. For each journey it is clear the mean and the standard deviation is substantially larger on the 25th June. For this reason the data for this date was removed from the dataset because the aim for obtaining this data set was for it to be used to create a database that defines an 'average travel time' for the line in question.

Table 2 – Mean and standard deviation of selected journeys

	Finsbury Park - Oxford Circus		Blackhorse Road - Euston		Highbury & Islington - Victoria	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
25/06/2012	20.18	41.81	26.12	16.25	22.93	35.50
02/07/2012	15.74	13.30	20.95	5.55	17.25	10.53
09/07/2012	15.42	13.53	21.26	5.84	16.98	9.85
16/07/2012	16.22	14.44	21.56	7.09	18.15	11.69
23/07/2012	15.77	15.11	21.34	5.53	16.76	9.57
30/07/2012	14.96	12.44	20.76	4.20	16.64	8.42
06/08/2012	15.24	12.34	20.53	4.78	16.76	8.91

These results are encouraging that within this work it will be possible to find delays incurred to passengers.

4.2. Average travel times

With the database seen in Section 4.1, the plan is to discover average un-delayed travel times for journeys completed on the Victoria Line. To find these times the mean travel times were found from the data. To be sure that any unwanted data that may be an unrealistic journey time was not included in the average it was decided that outliers should be removed from the dataset.

What constitutes an outlier journey and whether or not it should be removed is decided upon by what the data looks like, due to some analysis being incredibly sensitive to outliers. In general, the rule for removing outliers is that points which lie more than three standard deviations above or below the mean should be removed. However, it has been shown that this can produce problems for certain distributions, particularly when the sample is relatively small (Miller, 1991). Therefore it was decided to make a visual appraisal of the data to decide what should be considered as an outlier.

Only the upper outliers were removed as this was felt that they can be caused by people being slow, delays or people in groups. Lower outliers were, however, retained as it was felt that minimum times should remain part of the data set as in nearly all cases the lower bound was characterised by negative numbers, so does not affect the data. A bound of 0 is necessary since it is not possible for passengers' journey times to be negative numbers. In the few cases where the lower bound was a positive number this bound was rejected as removing these entries would be removing passenger time that had managed to complete the journey in free-flow conditions i.e. by traveling between the ticket barriers and platforms as quickly as possible, with no delay to the running of the service, the passenger boarding a train immediately after arriving on the platform and exiting the station as quickly as possible with no delay.

Data beyond two standard deviations from the mean were removed after analysing the data, and then the average was recalculated, as the mean should be as realistic to how long the 'average' passenger would take on the same journey on an un-delayed day without removing too much of the data. An example is given in Table 3 in this example, the original mean for Blackhorse Road to Walthamstow Central was 9 minutes. The standard deviation was found to be 9 minutes, so any values above 27 minutes were removed (shown in red). This gave a new mean of 8 minutes. In each case the value has been rounded to the nearest minute. This is because the data from the ticket barriers are recorded minute by minute (Chan, 2007). This means a journey time can only ever be a recorded as a whole number of minutes in length. Therefore, the average for the journeys should be expressed in minutes also.

Table 3 – Blackhorse Road to Walthamstow Central dataset

Entry Station	Exit Station	Journey Time	Journeys
Blackhorse Rd	Walthamstow Central	4	1
Blackhorse Rd	Walthamstow Central	5	6
Blackhorse Road	Walthamstow Central	6	6
Blackhorse Road	Walthamstow Central	7	7
Blackhorse Road	Walthamstow Central	9	2
Blackhorse Road	Walthamstow Central	12	1
Blackhorse Road	Walthamstow Central	20	1
Blackhorse Road	Walthamstow Central	51	1
Blackhorse Road	Walthamstow Central	4	2
Blackhorse Road	Walthamstow Central	5	4
Blackhorse Road	Walthamstow Central	6	5
Blackhorse Road	Walthamstow Central	7	5
Blackhorse Road	Walthamstow Central	8	2
Blackhorse Road	Walthamstow Central	9	1
Blackhorse Road	Walthamstow Central	10	2
Blackhorse Road	Walthamstow Central	48	1
Blackhorse Road	Walthamstow Central	4	2
Blackhorse Road	Walthamstow Central	5	4
Blackhorse Road	Walthamstow Central	6	8
Blackhorse Road	Walthamstow Central	7	4
Blackhorse Road	Walthamstow Central	8	2
Blackhorse Road	Walthamstow Central	9	1
Blackhorse Road	Walthamstow Central	10	1
Blackhorse Road	Walthamstow Central	47	1
Blackhorse Road	Walthamstow Central	56	1
Blackhorse Road	Walthamstow Central	4	1
Blackhorse Road	Walthamstow Central	6	2
Blackhorse Road	Walthamstow Central	7	2
Blackhorse Road	Walthamstow Central	8	1
Blackhorse Road	Walthamstow Central	9	1
Blackhorse Road	Walthamstow Central	19	1
Blackhorse Road	Walthamstow Central	22	1
Blackhorse Road	Walthamstow Central	24	1
Blackhorse Road	Walthamstow Central	55	1
Blackhorse Road	Walthamstow Central	4	1
Blackhorse Road	Walthamstow Central	5	2
Blackhorse Road	Walthamstow Central	6	5
Blackhorse Road	Walthamstow Central	7	8
Blackhorse Road	Walthamstow Central	8	4
Blackhorse Road	Walthamstow Central	10	2
Blackhorse Road	Walthamstow Central	28	1
Blackhorse Road	Walthamstow Central	50	1
Blackhorse Road	Walthamstow Central	5	8
Blackhorse Road	Walthamstow Central	6	4
Blackhorse Road	Walthamstow Central	7	5

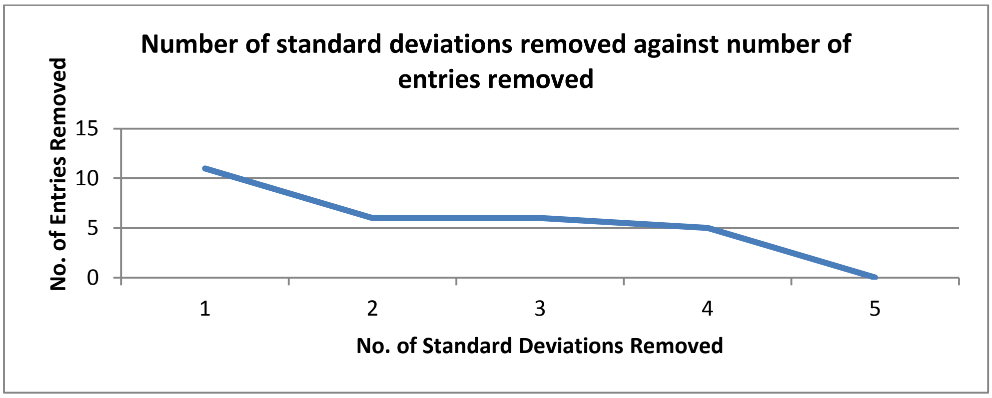
Blackhorse Road	Walthamstow Central	8	5
Blackhorse Road	Walthamstow Central	10	1

Table 4 and Graph 4 show the different numbers of standard deviations that could have been removed for this origin destination pair, with the number and the percentage of the total of journeys that would have been removed in each case.

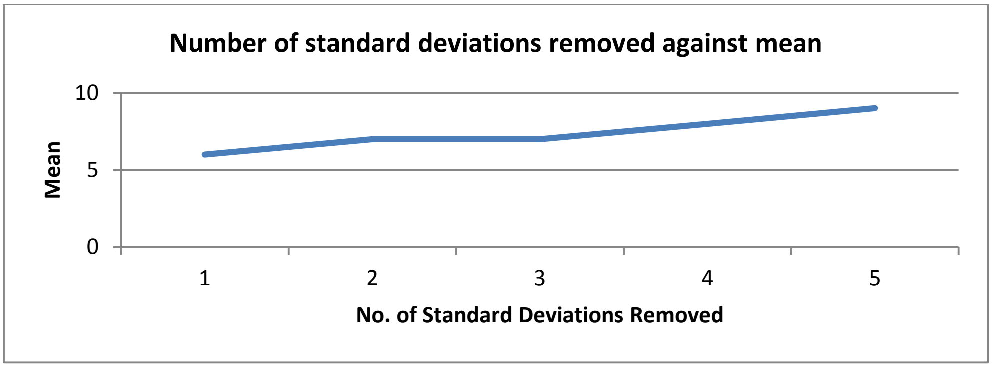
Graph 5 shows the effect on the resulting mean as data points beyond standard deviations thresholds are removed.

Table 4 – Blackhorse Road to Walthamstow Central: Removal of standard deviation and revised means

K	$\mu + \kappa\sigma$			M
No. of standard deviations	**No. of standard deviations + mean**	**No. of entries removed**	**Percentage removed from total data set.**	**New mean**
1	18	11	8.50%	6
2	28	6	4.60%	7
3	38	6	4.60%	7
4	48	5	3.80%	8
5	58	0	0%	9



Graph 4 - Blackhorse Road to Walthamstow Central: Number of standard deviations removed against number of entries removed



Graph 5 - Blackhorse Road to Walthamstow Central: Number of standard deviations removed against mean

On examining the data, it was noted that there was only a small difference between the amount of data affected by choosing 3 rather than 2 standard deviations as the upper bound. Due to this, it was decided to look no further than 3 standard deviations. It was however, interesting to consider 1 standard deviation, which would mean removing all the data points above 18 minutes, leaving a new average of 6 minutes. When looking at data sent from TfL that contains expected journey time distributions, it was found that around only 50% of Oyster card travellers make the journey in less than 7 minutes, whereas 85% of travellers make the journeys in less than 9 minutes (Transport for London, 2012). For this reason it was decided to use 2 standard deviations as the upper bound. This bound was then applied to all journeys on the Victoria Line. The times found for all the Victoria Line journeys are shown in the appendices in Table 73 and Table 74.

4.3. Regression analysis

Regression analysis was then undertaken to see what relationship is found between average journey times gained through Oyster card data and the London Journey Planner times. Some journey times were missing in the Oyster data due to a lack of data. In order to complete the regression analysis without these, a heuristic procedure was used to replace some of the missing values in the Oyster data. For the first regression (Table 5, Table 6, Table 7 and Graph 6), the relevant journey planner values were substituted for the missing Oyster data values.

Northbound journeys for the Victoria Line are as follows, with all values rounded to two decimal places.

Table 5 – Regression 1, Northbound Oyster Data: Regresssion statistics

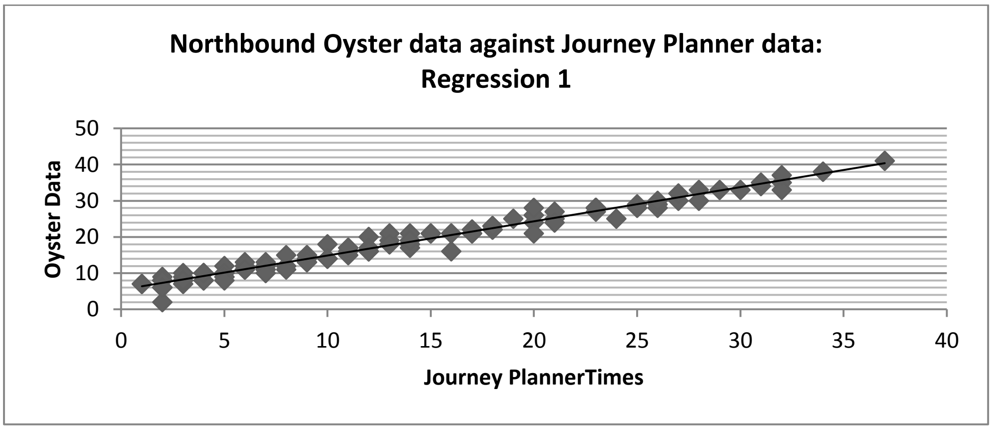
Regression Statistics	
Multiple R	0.99
R Square	0.97
Adjusted R Square	0.97
Standard Error	1.41
Observations	120

Table 6– Regression 1, Northbound Oyster Data: Anova

ANOVA					
	Df	**SS**	**MS**	**F**	**Significance F**
Regression	1	8763.28	8763.28	4410.91	9.80E-95
Residual	117	232.45	1.98673		
Total	118	8995.73			

Table 7– Regression 1, Northbound Oyster Data: Corrolation results

	Coefficients	**Standard Error**	**t Stat**	**P-value**	**Lower 95%**	**Upper 95%**	**Lower 95.0%**	**Upper 95.0%**
Intercept	5.38	0.24	22.63	1.5E-44	4.91	5.85	4.91	5.85
variable	0.95	0.01	66.41	9.8E-95	0.92	0.98	0.92	0.98



Graph 6 – Northbound Oyster data against Journey Planner data: regression 1

From Table 7 the equation of the regression line is $y = 0.95\mathrm{x} + 5.38$. The R squared value in this regression is 0.97 this means that 97% of the Oyster data times can be found from the Journey Planner times. The adjusted R Squared also shows the same result. This is a more accurate value to consider since it takes into account the sample size. It can further be seen that the p-value is very small this means that it is very unlikely that these results occurred at random. Next, using the results of this regression, the missing Oyster values were inserted in place of the Journey Planner values and a second regression was undertaken. The results can be seen in Table 8, Table 9, Table 10 and Graph 7.

Table 8 - Regression 2, Northbound Oyster Data: Regresssion statistics

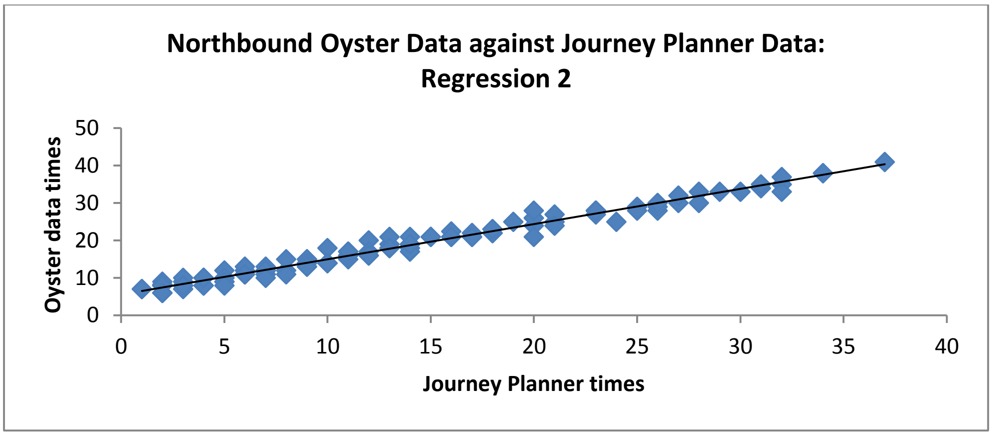
Regression Statistics	
Multiple R	0.99
R Square	0.98
Adjusted R Square	0.98
Standard Error	1.26
Observations	120

Table 9 – Regression 2, Northbound Oyster Data: Anova

ANOVA					
	df	SS	MS	F	Significance F
Regression	1	8646.97	8646.97	5417.09	7.80E-100
Residual	117	186.76	1.60		
Total	118	8833.73			

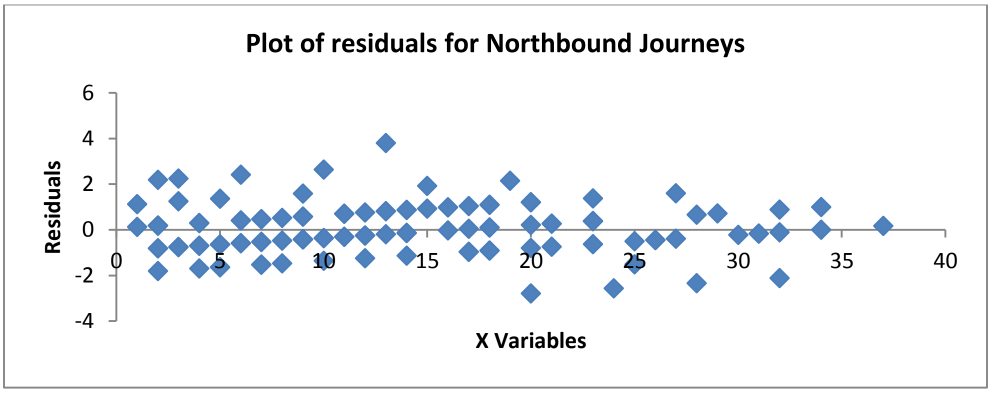
Table 10 – Regression 2, Northbound Oyster Data: Corrolation results

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	4.93	0.18	26.87	0.00	4.57	5.29	4.57	5.29
Variable	0.94	0.01	85.32	0.00	0.92	0.97	0.92	0.97



Graph 7- Northbound Oyster data against Journey Planner data: regression 2

In Table 10 it can be seen that the equation of the line is $y = 0.94x + 4.93$. In this regression the R squared and the adjusted R squared values are both 0.98 this means that 98% of the Oyster data times can be found from the Journey Planner times. This shows that with this second regression and the replaced values the line is an even better fit to the data than before. It can further be seen that the p-value is very small this means that it is very unlikely that these results occurred at random. The significance of F being so small confirms the validity of the regression output.



Graph 8 - Plot of residuals for Northbound journeys

Graph 8 shows the plot of the residuals, it can be seen they are equally distributed above and below the zero-line, with an average of 0; therefore the regression line is a good model for the data.

The results for southbound journeys on the Victoria Line are shown in Table 11, Table 12 , Table 13 and Graph 9. This time, the regression analysis only needed to be completed once as there was no missing data. Again all values were rounded to two decimal places.

Table 11 - Regression 1, Southbound Oyster Data: Regresssion statistics

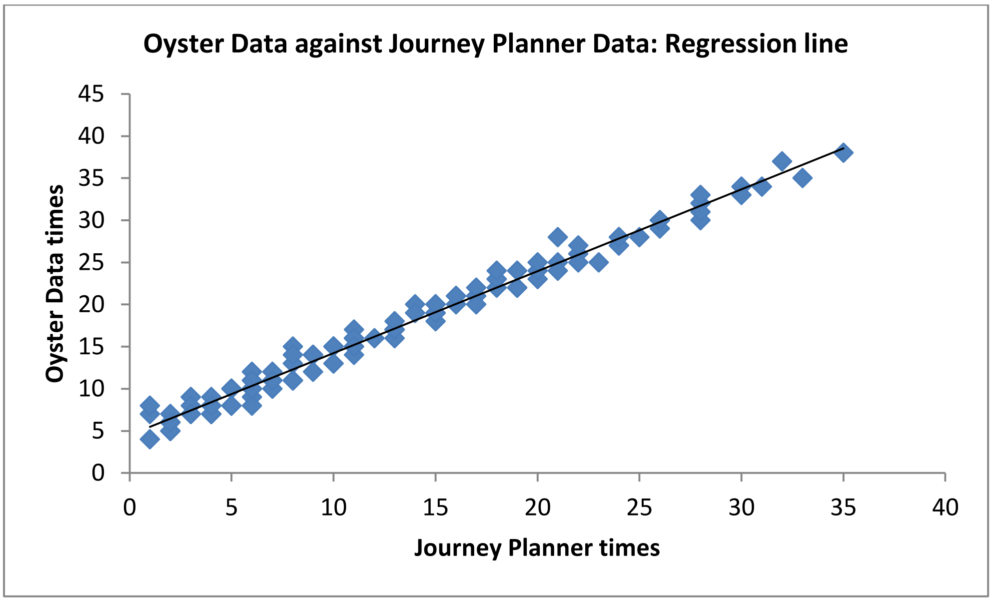
Regression Statistics	
Multiple R	0.99
R Square	0.98
Adjusted R Square	0.98
Standard Error	1.07
Observations	120

Table 12– Regression 1, Southbound Oyster Data: Anova

ANOVA					
	df	**SS**	**MS**	**F**	**Significance F**
Regression	1	8365.22	8365.22	7352.838	3.8E-108
Residual	118	134.2469	1.137686		
Total	119	8499.467			

Table 13 – Regression 1, Southbound Oyster Data: Corrolation results

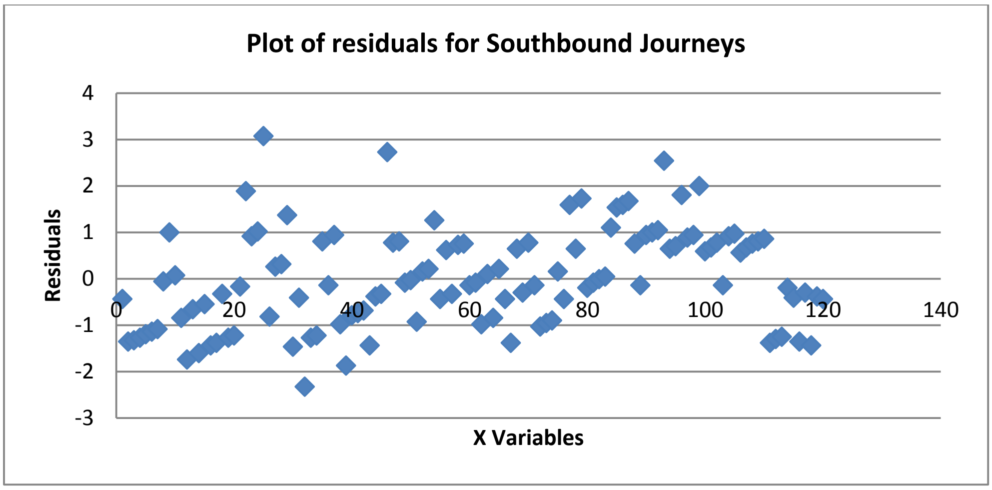
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**Lower 95%**	**Upper 95%**	**Lower 95.0%**	**Upper 95.0%**
Intercept	4.49	0.18	25.49	8.17E-50	4.14	4.84	4.14	4.84
Variable	0.97	0.01	85.75	3.80E-108	0.95	1.00	0.95	1.00



Graph 9 - Oyster data against Journey Planner data: Regression line

In Table 13 it can be seen the equation of the line is $y = 0.97\mathrm{x} + 4.49$. Again for the southbound journeys the R squared value is very close to 1 at 98% showing the same result as the northbound data. This tells us the regression line closely approximates the real data and validates that neither Southbound nor Northbound results are anomalous. So again for southbound journeys there is a strong relationship between Oyster data and the TfL journey planner.

The significance of F being so small confirms the validity of the regression output. Again, the p-value being small for the intercept validates the regression as well. Graph 10 shows an equal scattering of the residuals either side of 0, showing what a good fit the regression line is to the data.



Graph 10 -Plot of residuals for Southbound ourneys: Regression analysis

In this section the average travel times were found for both Southbound and Northbound Victoria Line journeys by taking mean travel times over an 8-week period. Regression analysis was completed to determine the relationship between the mean travel times found through Oyster data and the travel time on the TfL journey planner. In both directions of the line statistical testing shows significance for the strong correlation between the two values, this is encouraging that the times are accurate. In summary, at this stage a dataset has now been created with complete values of all possible journeys with their origin and destination on the Victoria Line, which can be used for further analysis.

4.4. What is a delay?

4.4.1. Passengers perspectives

The aim of this thesis is to determine if it is possible to give passengers of a metro network real-time information. So far in the process of discovering if smart card data does contain information about the dynamics of network, a base point of an ‘average day’ on the Victoria line has been described in terms of travel times. Beyond this, a classification needs to be made to decide whether a passenger, who has just completed their journey has been delayed or not. This will provide information about the current dynamics of the network; as either delayed or a normal service. Finally, a value can be given to the delay as well as information about it. However first a threshold that can be used to define a delay needs to be created.

To decide what value the delay threshold should take it is important to consider what the operators may believe a delay is classified as. In London, a delay is classified as Minor, Major or Severe. Therefore, there is no particular time threshold used to define a delay in the London Underground.

Since this project requires a numerical value for a delay threshold and delays to the network are defined from a passenger's perspective, it seems appropriate to consider what a passenger defines as a delay in the London Underground. Therefore a questionnaire will be taken to determine this value. The aim of conducting a questionnaire will be to determine what passengers think a delay is and what information they would want to know about a delay.

4.4.1.1. Method

When determining the sample size for the questionnaire it was important to take into consideration the number of people that might be using the Underground in London.

A billion journeys are made every year in the Underground ("London Underground | Transport for London," n.d.), in the 2011 Census, London's population was 8.17 million, ("London Key Facts and Statistics," n.d.) and approximately 225,000 people visited London (Kyte, 2012). In addition to this there are people that commute to London for work and visitors from within the UK. To be able to generalise to this large population with a 5% margin of error a sample of 384 was needed, seen below Figure 9 taken from Survey Monkey ("Sample Size for Survey: Calculate Respondent Population | SurveyMonkey," n.d.). This number was rounded up to 400 participants.

Population	Margin of Error			Confidence Level		
	10%	5%	1%	90%	95%	99%
100	50	80	99	74	80	88
500	81	218	476	176	218	286
1,000	88	278	906	215	278	400
10,000	96	370	4,900	264	370	623
100,000	96	383	8,763	270	383	660
1,000,000+	97	384	9,513	271	384	664

Figure 8 – Survey Monkey: How to calculate your sample size

To ensure that the 400 participants represented the demographic of those living and visiting London a number of questions were asked to determine some basic characteristics. These included questions about their age, sex, home location and purpose of journey. The questions concerning purpose of journey and the age categories were chosen to match those of the Office of National Statistics and the categories they use ("Travel Trends, 2012," n.d.).

Further questions were included to determine what format passengers want to receive information about the system; these questions were included to focus the process of determining how to deliver the information.

Finally, a question was included to determine if a passenger's behaviour may change as a result of a change in information provided. A copy of the questionnaire is shown below.

---

1. Time: **Date:** **M / F**

2. What was the station you started your journey at?

What was the station you ended your journey at?

3. What was the purpose of your journey?

Commuting / Business / Education (including escorting) / Shopping / other escorting and personal business / Visiting friends / Other Leisure (including sports)

4. How old are you?

0-4 5-9 10-14 15-19 20-24 25-29 30-34 35-39 40-44 45-49 50-54 55-59 60-64 65-69 70-74 74-79 80-84 85-90 90+

5. Which one are you?

Living and working in London / Living outside London but working in London / Visiting London, living in the UK / Visiting London, living outside the UK

6. For the journey you have just taken, if you could receive information about the current service in the Underground on your phone which would you prefer?

a. Minor / Major / Severe statuses

b. A prediction of how many minutes the line will be delayed based on live information

c. A prediction of your journey time based on live information including delays

d. Not interested in this information

7. For the journey you have just taken, if you could receive information about the current service in the Underground on the service boards in the station which would you prefer?

a. Minor/Major/Severe statuses

b. A prediction of how many minutes the line will be delayed based on live information

c. Not interested in this information

8. How many extra minutes do you think there would be added to your journey if the service was

Minor delay ____minutes Major delay____minutes Severe delay____minutes

9. If you could receive information about the length of time it will take you to queue either on the platform or exiting the station, would you want this information? Yes No

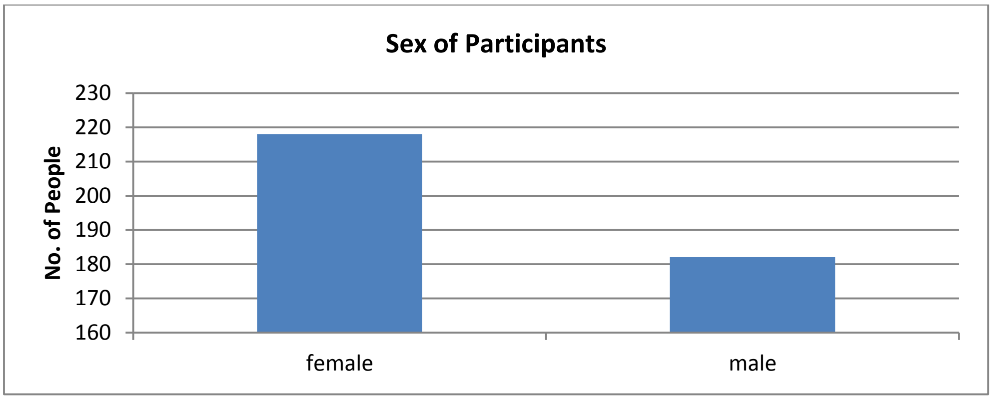
***If yes:* Might this information change your behaviour?** (i.e. leave later, take a difference route)

Yes No

Four stations were chosen for data collection. These were Highbury and Islington, Warren Street, Oxford Circus and Green Park. These stations were chosen because they are on the Victoria Line, due to the accessibility of their exits, as they were central and since there would be a high volume of passengers exiting and entering. However, this choice will not lead any unusual characteristics of the participants that would be any different to any other station on the Victoria line.

The results were collected over 8 days between the 4th to the 12th of July. 33% of the results were collected in the AM peak this is between 07:00 and 10:00, 34% were collected between 10:00 and 16:30 this is off-peak and finally 33% were collected between 16:30 and 19:00 which is the PM peak.

4.4.1.2. Results of the demographic questions



Graph 11 - Sex of participants

Graph 11 shows 45.5% of the respondents were male and 54.5% were female this is slightly out of the demographic of London which estimates that 50.5% of those living in London are females (“Custom Age Tool for ONS Mid-Year Population Estimates | London DataStore,” n.d.). To determine if the results were random or if there is statistical significance, i.e. that the results are not random; this has been tested using a one-way chi-squared test to see the size of variations around the expected value.

For this, the null hypothesis H_0 is: the mean frequency for all responses is equal and individual responses are random.

The alternative hypothesis is $H_1$is: at least one response has a different mean frequency.

To determine which hypothesis is accepted, let the number of responses in each category *i* be n_i and the number of categories *l*.

Then the null mean frequency for each category is estimated using the formula:

$$M_0 = \sum_i \frac{n_i}{l}$$

Equation 4

Then to test for significance, the following formula is used:

$$\chi^2 = \sum_i \frac{(n_i - M_0 - 0.5)^2}{M_0}$$

Equation 5

Where the 0.5 is added for the continuity correction. Then the null hypothesis is tested by calculating the value on the cumulative chi – squared distribution of the test statistic χ^2.

If $1 - p < \alpha$: reject H_0 at level α of statistical significance in favour of H_1.

If $1 - p > \alpha$: cannot reject H_0 at level α of statistical significance, so proceed as if were true.

Where $\alpha = 0.01$.

In this example there are two categories: Male = 182 and Female = 218.

Then using Equation 4 and Equation 5,

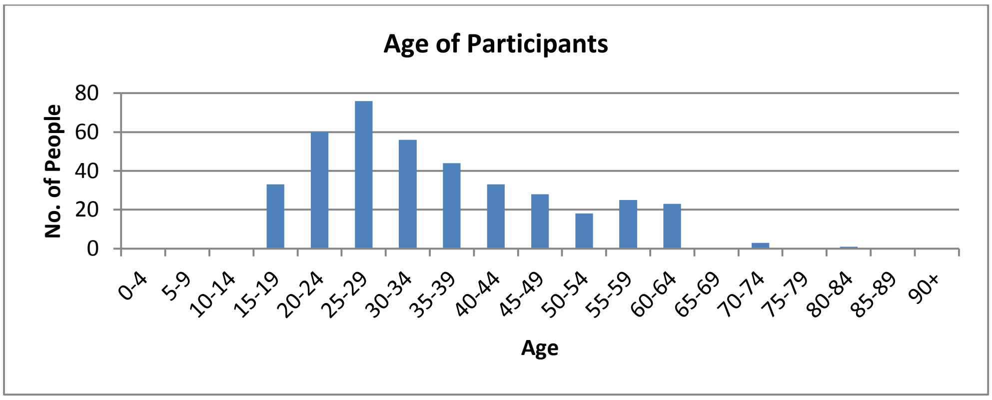
$$M_0 = \frac{218 + 182}{2} = 200$$

Equation 6

$$\chi^2 = \frac{(218 - 200)^2}{200} + \frac{(182 - 200)^2}{200} = 3.25$$

Equation 7

Which gives $1 - p = 0.095$ meaning $H_0$is accepted at level 0.01 (k=1) of statistical significance. Therefore there is a non-bias selection of males and females.



Graph 12 - Age of participants

Graph 12 shows the ages of the 400 participants; again to ensure these results were random Chi Squared was used to test for statistical significance. As a comparator the age structure of those living in London has been taking from the Office of National Statistics , seen in Table 14("Custom Age Tool for ONS Mid-Year Population Estimates | London DataStore," n.d.)

To compare two data sets using chi squared use the formula

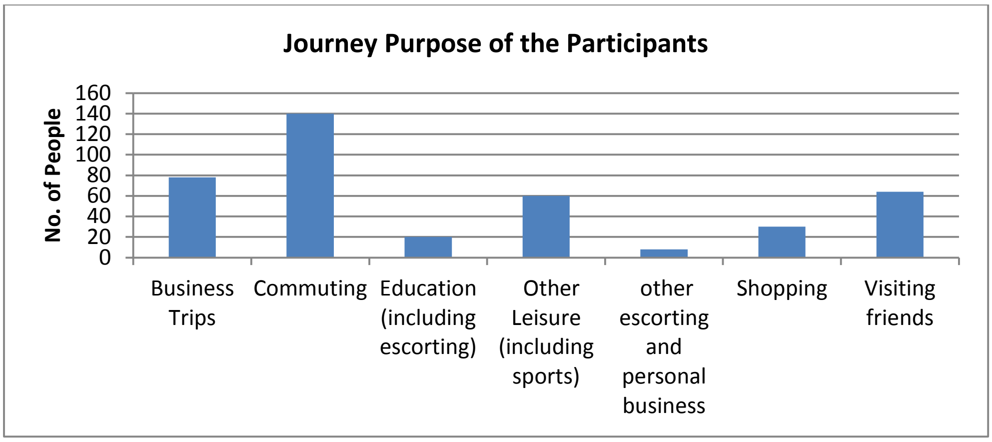
$$\chi^2 = \sum_i \frac{(Observed_i - Expected_i)^2}{Expected_i}$$

Table 14 – Ages of observed and expected

Age	Observed	Expected
0-4	0.00%	7.20%
5-9	0.00%	5.90%
10-14	0.00%	5.60%
15-19	8.00%	5.70%
20-24	15.00%	7.7%
25-29	19.00%	10.2%
30-44	33.00%	25.30%
45-59	18.00%	17%
60-64	6.00%	4.2%
65-74	1.00%	5.8%
75-84	0.00%	3.8%
85-89	0.00%	1%
90+	0.00%	0.5%

$$\chi^2 = 47$$

The two-tailed P value is less than 0.0001. By conventional criteria, this difference is considered to be extremely statistically significant.



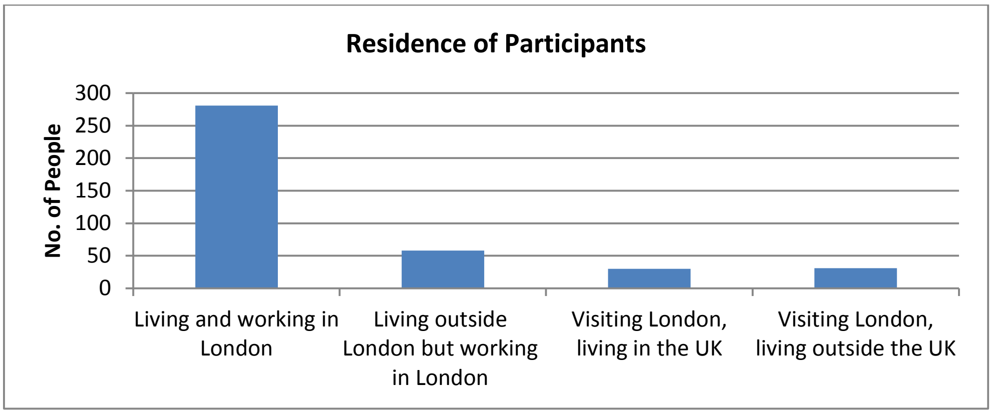
Graph 13 - Journey purpose of the participants

In Graph 13 the purpose of the participants journeys are shown, it be seen that 35% of the respondents were commuting when the questionnaire was taken, this may be due to the times in which the questionnaires were collected. The table below shows the results from the questionnaire and the national average journey purposes (Department for Transport, 2012).

Table 15 – Comparing the national average of journey purposes to participants journey purposes

Questionnaire Respondents (%)	Journey Purpose	National Average (%)
19.5	Business Trips	3
35	Commuting	15
5	Education (including escorting)	11
15	Other Leisure (including sports)	16
2	other escorting and personal business	20
7.5	Shopping	20
16	Visiting friends	15

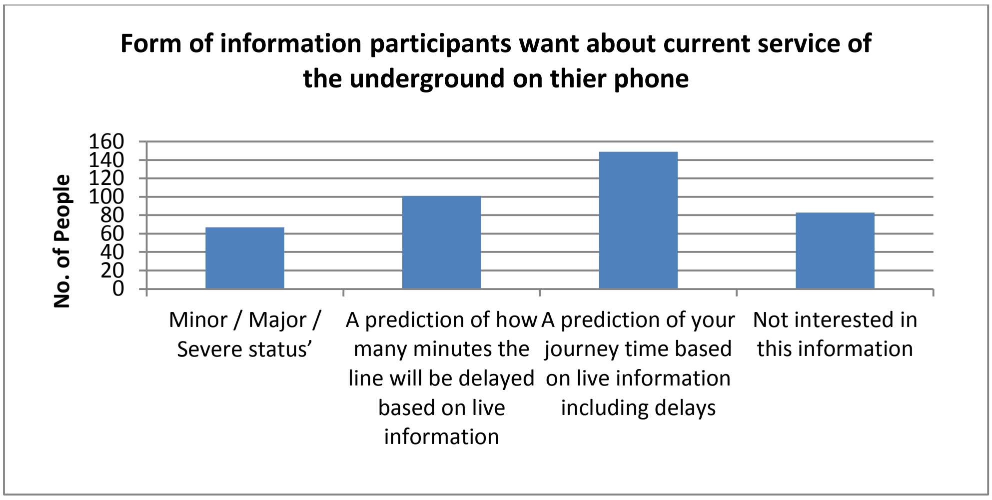
The difference in the two values seen in Table 15 can be put down to the time of day the questionnaires were collected. The times to collect the questionnaire were chosen to ensure large passenger demand however it is assumed that the national average statistics are taken evenly throughout the day, whereas for this experiment there were no entries after 19:00 and 66% of the questionnaires were completed in peak hours. This would explain why there is a higher level of business trips and commuters and a lower amount of shopper and those on personal business. Finally, the questionnaire was taken out of the school term time, this may be the cause of the lack of trip made due to education. Chi squared was used to compare the two sets, significance was found, showing the two sets are statistically similar.



Graph 14 - Residence of participants

Graph 14 shows the residence of the participants. These results cannot be compared to any standard statistics as there does not appear to be any data describing the number of people in London on the basis of different purposes. However, as stated above there are over 225,000 visitors to London every year from outside the UK. This question was asked to ensure that others rather than just those that live and work in London were included in the survey as they do contribute to the passengers in London Underground.

Although the sample of participants does not quite match the demographic of London in general, there is a mixture of ages, a good split between male and female, a wide range of journey purposes and visitors to London have been included. Therefore for the sake of this project the results will be taken and used yet these results cannot be generalised to the population of London due to the slightly skewed sample in some cases.



Graph 15 - Form of information participants want about current service of the underground on thier phone

Graph 15 shows the results to the question:

For the journey you have just taken, if you could receive information about the current service in the Underground on your phone which would you prefer?

a. **Minor / Major / Severe statuses**

b. **A prediction of how many minutes the line will be delayed based on live information**

c. **A prediction of your journey time based on live information including delays**

d. **Not interested in this information**

Here it can be seen a large number of passengers have chosen to receive dynamic information about their journey time. To show that all the responses were not random, chi-squared was used to show statistical significance.

For this question the number of responses in each category are:

67 = Minor / Major / Severe statuses

101 = A prediction of how many minutes the line will be delayed based on live information

149 = A prediction of your journey time based on live information including delays

83 = Not interested in this information

Where $l = 4$.

For this, the null hypothesis H_0 is: the mean frequency for all responses is equal and individual responses are random.

The alternative hypothesis is H_1 is: at least one response has a different mean frequency.

Then using Equation 4 and Equation 5,

$$M_0 = \frac{67 + 101 + 149 + 83}{4} = 100$$

Equation 8

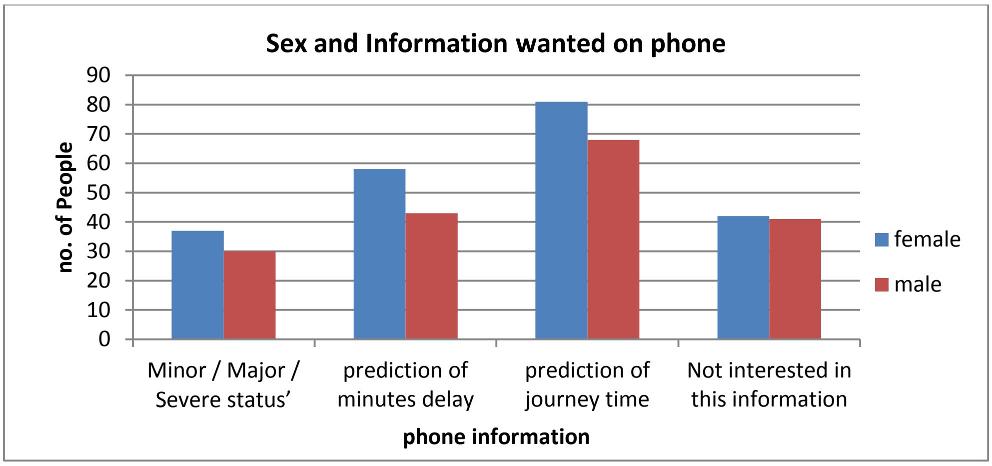
$$\chi^2 = \frac{(67 - 100)^2}{100} + \frac{(101 - 100)^2}{100} + \frac{(149 - 100)^2}{100} + \frac{(83 - 100)^2}{100} = 37.8$$

Equation 9

Which gives $1 - p = 0.00000003$ meaning H_0 is rejected at level 0.01 of statistical significance in favour of H_1.

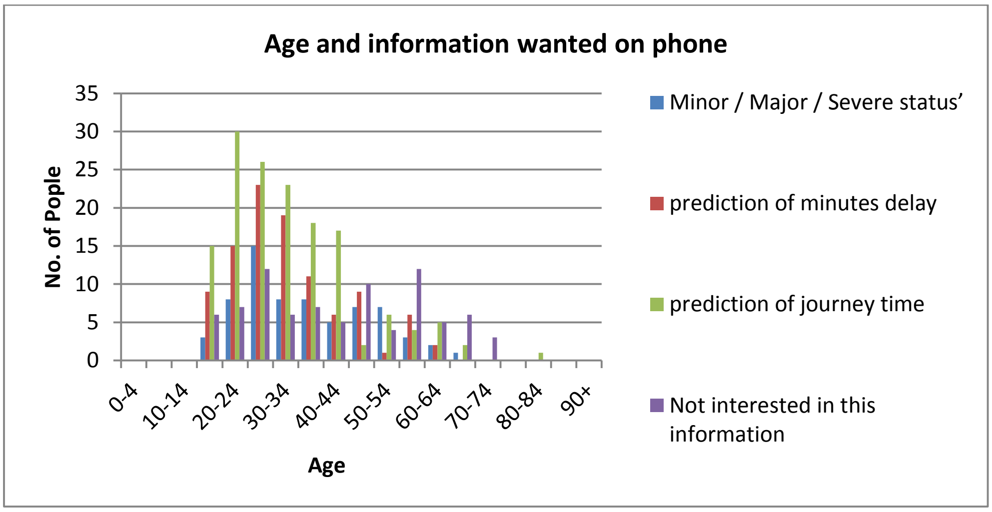
Therefore there is a strong result in favour of passengers of the London Underground wanting a prediction of their journey time based on live information including delays over other forms of information on their phone.

Next it was checked to see whether sex, residency, age of trip purpose made a difference to the answer to this question.



Graph 16 - Sex and information wanted on participants phone

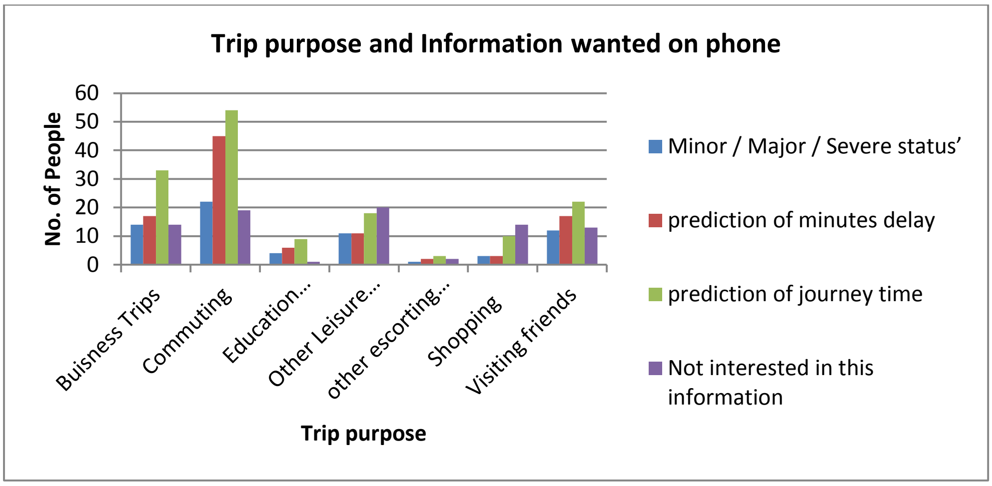
Graph 16 shows that for the types of information wanted there is no difference between males and females, yet for those not wanting information it would appear there are more men interested out of the sample.



Graph 17 - Age and information wanted on participants phone

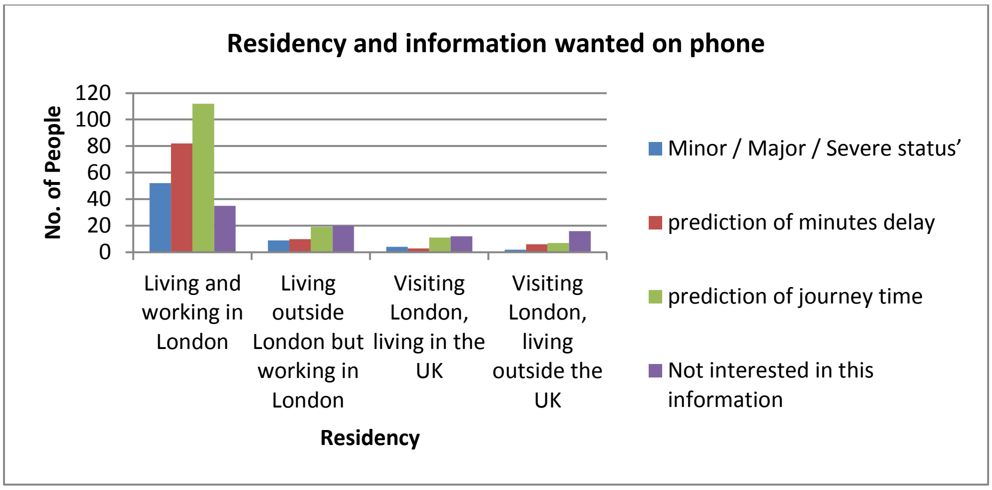
Graph 17 shows that under the age of 44 there is a greater preference for dynamic information, whether it be the number of minutes the line is delayed or a prediction of their journey time. This could be due to the generation gap in usage of smart phones, for example 27% of adults own a smart phone but 47% of teenagers use a smart phone ("Ofcom | A nation addicted to smartphones," n.d.).

Although these statistics don't show the number of people between 15 and 44 being more likely to have a smartphone than those over 45, it indicates that this may be true.



Graph 18 - Trip purpose and Information wanted on participants phone

Graph 18 shows that those taking business trips and commuting are more inclined to want either a prediction of how long the line will be delayed in minutes or a prediction of their journey time. This could be due to the urgency of their travel and therefore a greater need for dynamic information. The same result can be seen for visiting friends and education these also being time dependent activities. This result is also true for those completing a trip for other escorting and personal business however there aren't enough results for anything to be conclusive.



Graph 19 - Residency and information wanted on participants phone

Graph 19 shows there is a strong result showing that those that do not live in London do not have a necessity for dynamic information or information on their phones at all. This is an interesting result for understanding what market the information should cater to.

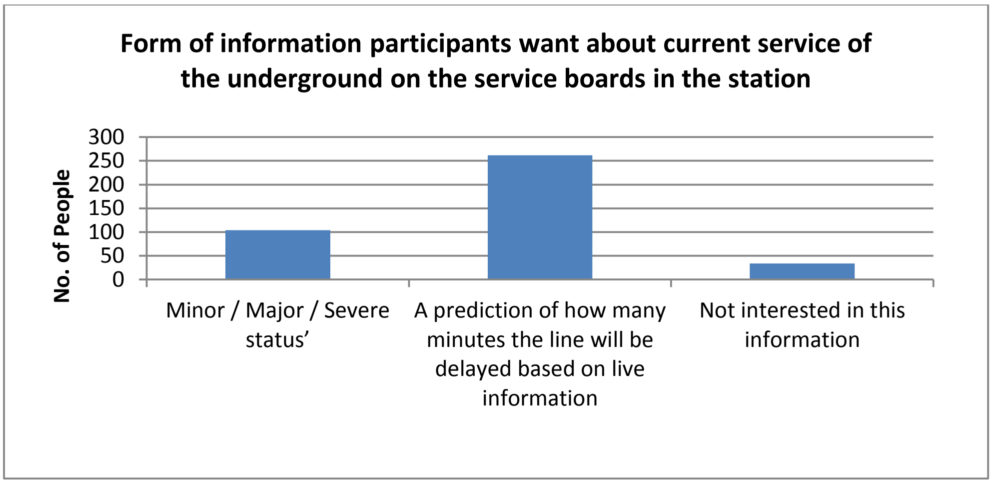
The next question to be analysed is:

For the journey you have just taken, if you could receive information about the current service in the Underground on the service boards in the station which would you prefer?

a. Minor/Major/Severe statuses

b. A prediction of how many minutes the line will be delayed based on live information

c. Not interested in this information



Graph 20 - Form of information participants want about current service of the underground on the service boards in the station

Graph 20 shows the results of the question regarding information on service boards. Again although it looks as if there is a strong result that passengers want a prediction of how many minutes the line is delayed, significance testing will show if the result is random or not.

For this question the number of responses in each category are:

104 = Minor / Major / Severe statuses

262 = A prediction of how many minutes the line will be delayed based on live information

34 = Not interested in this information

Where $l = 3$.

For this, the null hypothesis H_0 is: the mean frequency for all responses is equal and individual responses are random.

The alternative hypothesis is $H_1$is: at least one response has a different mean frequency.

Then using Equation 4 and Equation 5,

$$M_0 = \frac{104 + 262 + 34}{3} = 133.33$$

Equation 10

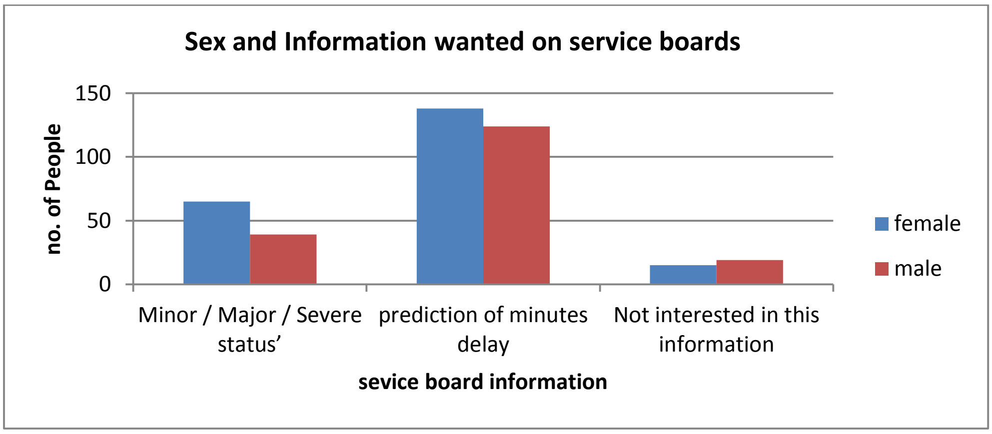
$$\chi^2 = \frac{(104 - 133.33)^2}{133.33} + \frac{(262 - 133.33)^2}{133.33} + \frac{(34 - 133.33)^2}{133.33} = 204.63$$

Equation 11

Which gives $1 - p = 0$ meaning $H_0$is rejected at level 0.01 of statistical significance in favour of H_1.

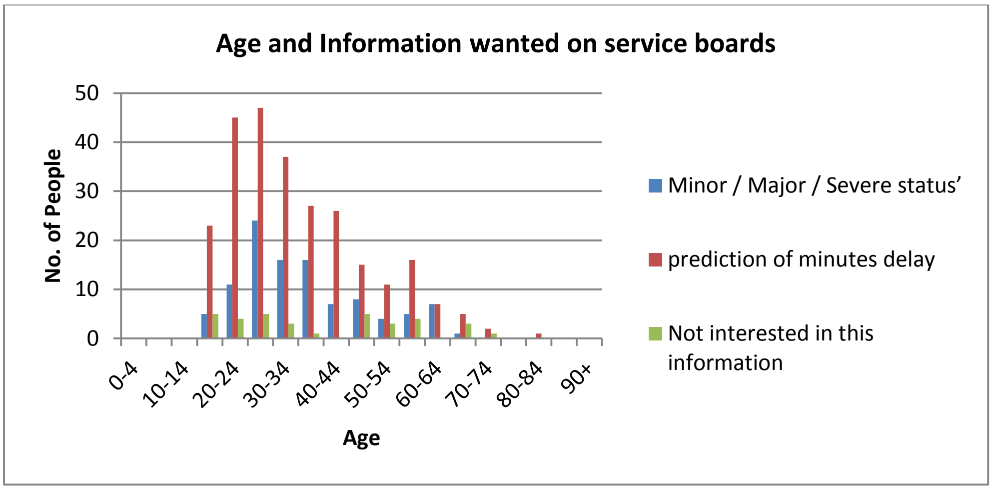
Therefore there is a strong result in favour of passengers of the London Underground wanting a prediction of how many minutes the line is delayed over other forms of information or no information on the service boards within the stations.

Next it was checked to see whether sex, residency, age or trip purpose affected the participants' answers to this question.



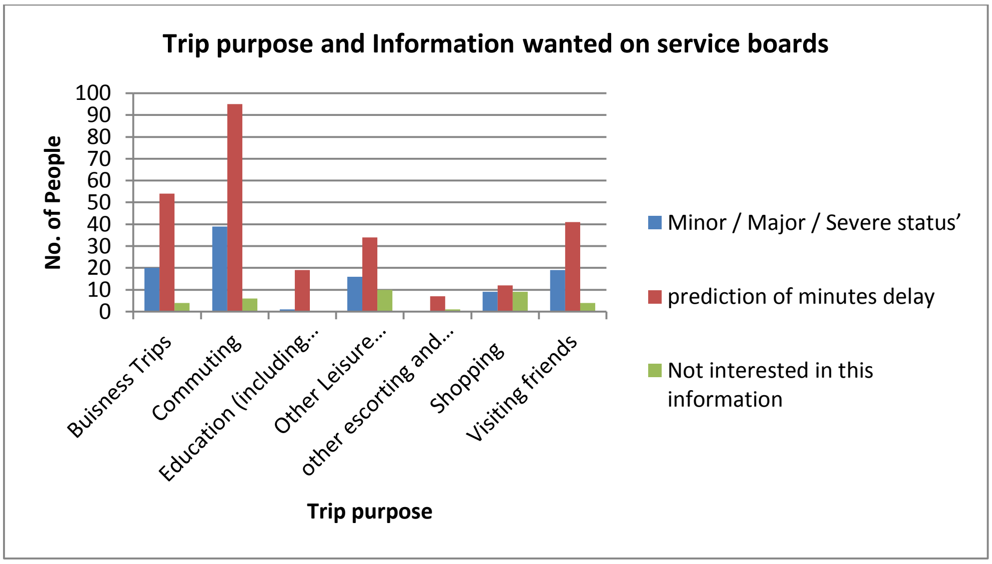
Graph 21- Sex and information wanted on service boards

Graph 21 shows there are small variations between male and female respondents and their results to this question but these variations are negligible.



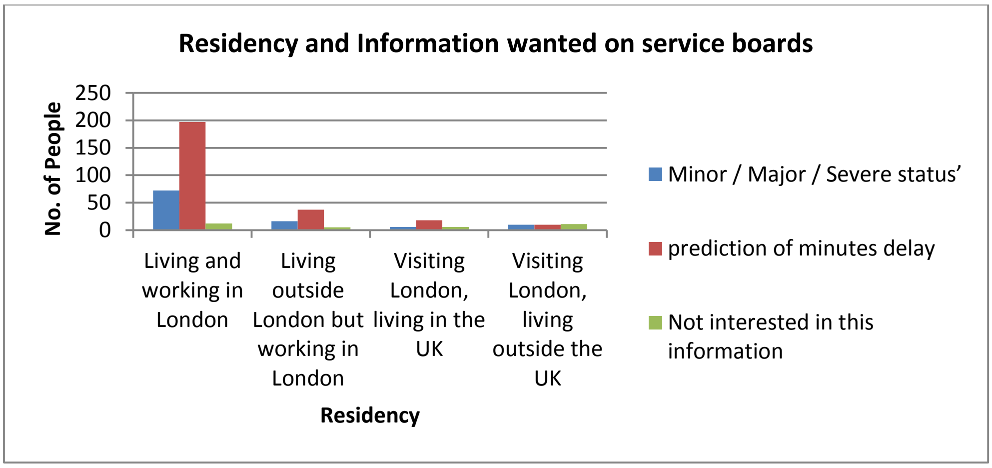
Graph 22- Age and information wanted on service boards

In Graph 22, in comparison to Graph 17, it can be seen this result is slightly different from the question concerning phone information. Whereas with the question regarding information on phones (Graph 17) a clear difference could be seen with the ages, for this question it would appear that throughout the different ages the information preferences stay the same. When looking at this graph it is important to take into consideration the number of respondents in each age category, shown in Graph 12.



Graph 23 - Trip purpose and Information wanted on service boards

Graph 23 shows the preference of information wanted on the service boards in regards to the passengers' trip purposes. In comparison to the question regarding information on the passenger's phone, here there is no noticeable difference in the information preferences for each journey purpose. It would appear that regardless of trip purpose the majority of passengers would want to see a more accurate form of information on the service boards in the Underground.



Graph 24 Residency and information wanted on service boards

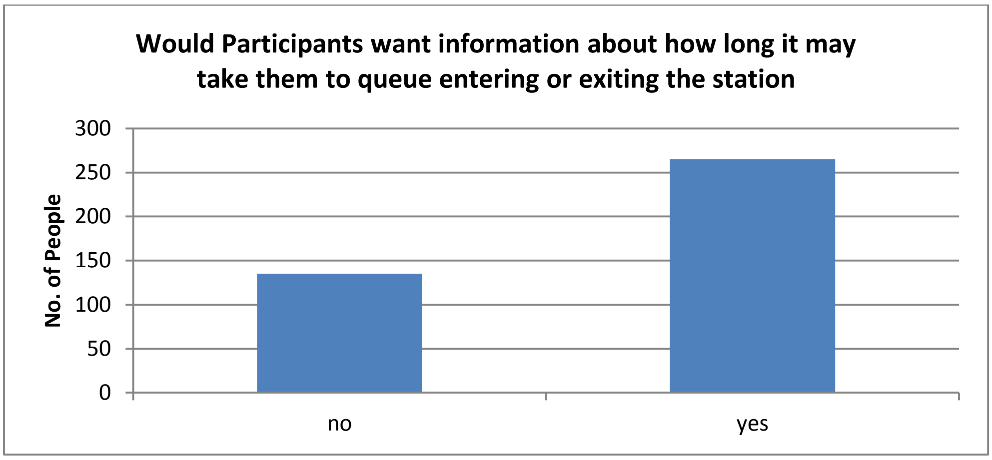
Finally, Graph 24 shows how passengers of the Underground value the information of how many minutes the line would be delayed on the service boards. With passengers that live abroad the least interested in this information and their results being 10 for Minor/Major/Severe statuses, 10 for a prediction of how many minutes the line will be delayed and 11 not interested in the information, this shows that approximately a 1/3 of visitors still want dynamic information.

The next results are for the question regarding congestion information shown below.

If you could receive information about the length of time it will take you to queue either on the platform or exiting the station, would you want this information? Yes No

***If yes:* Might this information change your behaviour?** (i.e. leave later, take a difference route)

Yes **No**



Graph 25 - Would participants want information about how long it may take them to queue entering or exiting the station

Graph 25 shows there is a very strong result in favour of passengers wanting information about how long it may take them to queue, with 2/3 saying yes they do want the information. However, again Chi-squared was used to show statistical significance.

So, for this question the number of responses in each category are:

265 = Yes

135 = No

Where $l = 2$.

For this, the null hypothesis H_0 is: the mean frequency for all responses is equal and individual responses are random.

The alternative hypothesis is $H_1$is: at least one response has a different mean frequency.

Then using Equation 4 and Equation 5,

$$M_0 = \frac{265 + 135}{2} = 200$$

Equation 12

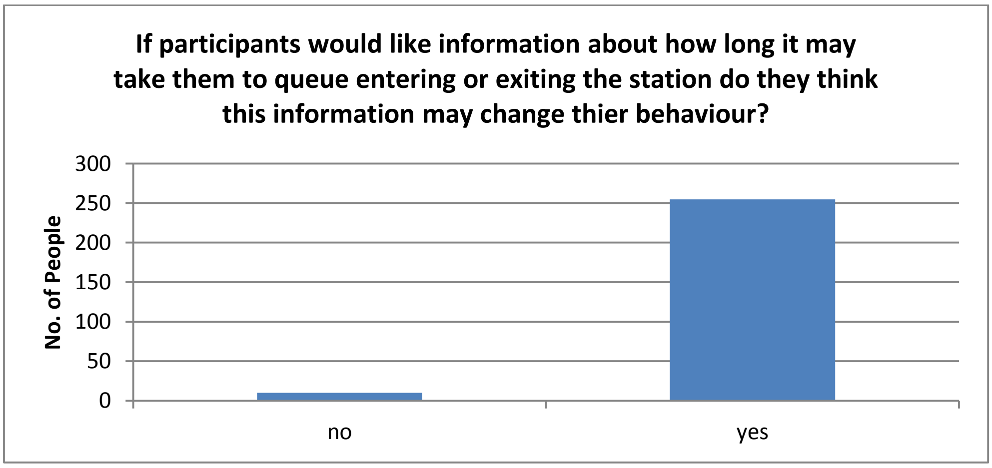
$$\chi^2 = \frac{(265 - 200)^2}{200} + \frac{(135 - 200)^2}{200} = 42.25$$

Equation 13

Which gives $1 - p = 0$ meaning H_0 is rejected at level 0.01 of statistical significance in favour of H_1.

Therefore there is a strong result in favour of passengers of the London Underground wanting information about how long it may take them to queue either to enter or exit the station.

Further to this it was asked whether this information may change their behaviour, which is shown below.



Graph 26 - If participants would like information about how long it may take them to queue entering or exiting the station do they think this information may change thier behaviour?

This was a very strong result with 96% saying yes. Again Chi-squared was used to test the significance of the result.

So, for this question the number of responses in each category are:

255 = Yes

10 = No

Where $l = 2$.

For this, the null hypothesis H_0 is: the mean frequency for all responses is equal and individual responses are random.

The alternative hypothesis is $H_1$is: at least one response has a different mean frequency.

Then using Equation 4 and Equation 5,

$$M_0 = \frac{255 + 10}{2} = 132.5$$

Equation 14

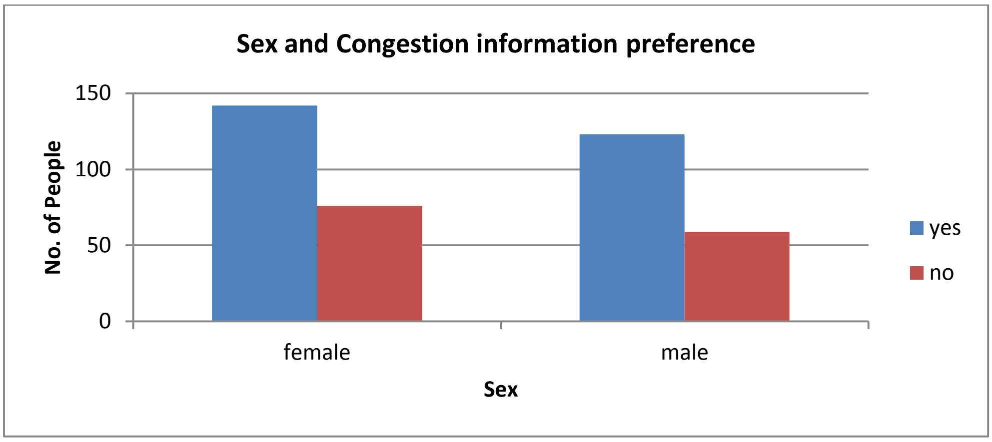
$$\chi^2 = \frac{(255 - 132.5)^2}{132.5} + \frac{(10 - 132.5)^2}{132.5} = 226.51\ (2dp)$$

Equation 15

Which gives $1 - p = 0.00140528$ meaning $H_0$is rejected at level 0.01 of statistical significance in favour of H_1.

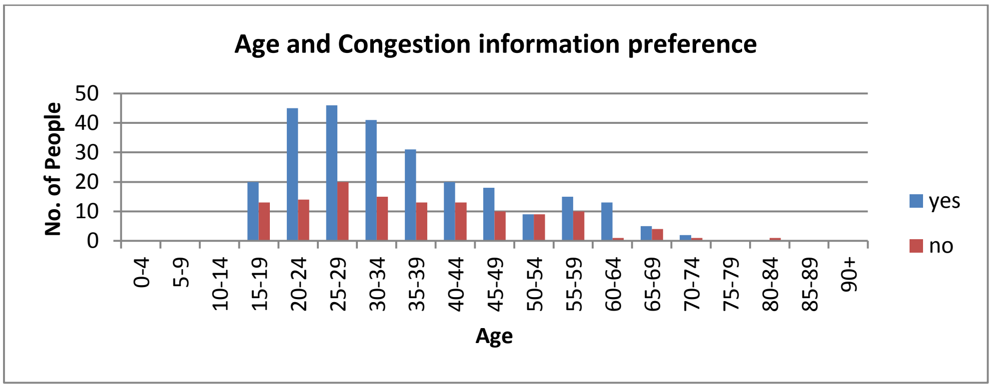
Therefore there is a strong result in favour of those passengers who do want information about how long it may take them to queue believing that the information may change their behaviour.

Analysis was completed to see if there were any connections between those wanting queuing information and their sex, age, trip purpose and residency.



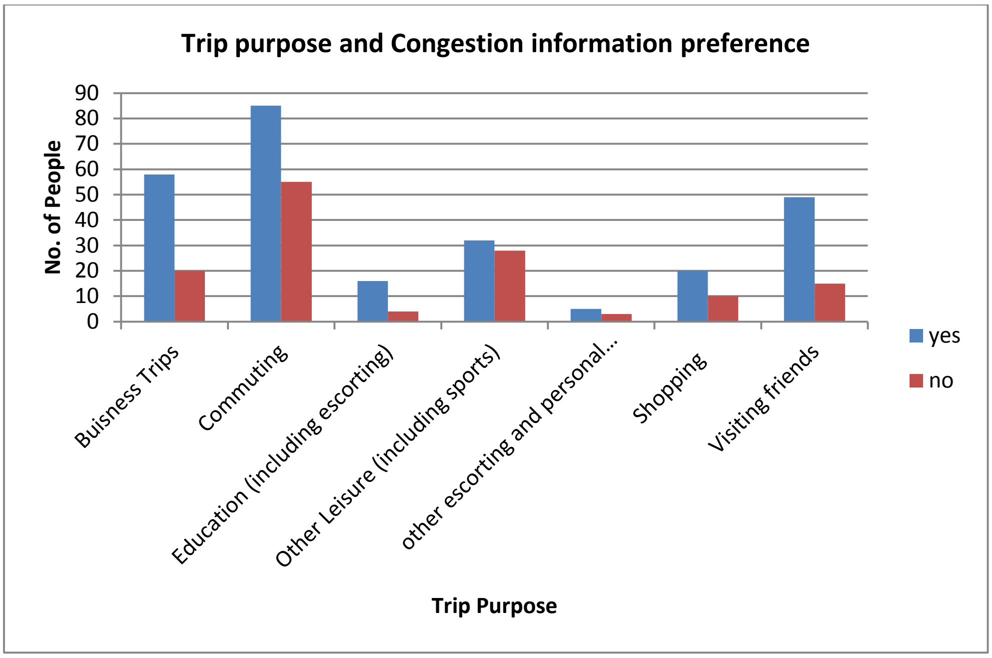
Graph 27- Sex and congestion information preference

Graph 27 shows that there is no difference between male and females and their preference to wanting information about queues.



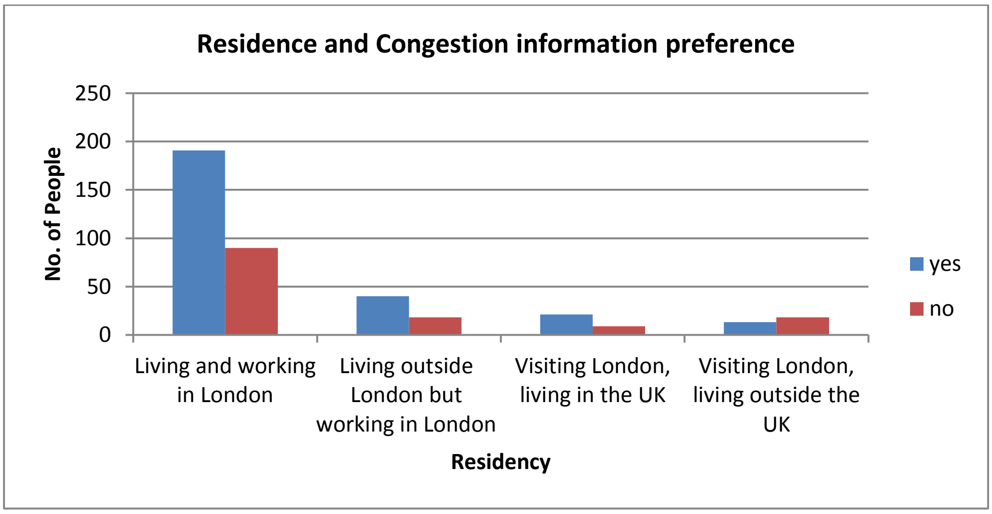
Graph 28 - Age and congestion information preference

Age also does not influence people's preference to congestion information. In all age categories there is an obvious preference for wanting the information.



Graph 29 - Trip purpose and congestion information preference

The only trip purpose that seems to be slightly different from the others is 'other leisure' where there is only a small difference between the two responses. However, most of the respondents that ticked 'other leisure' were tourists or sight seers, this may be the reason they are not that interested in the information as they do not live in the country.



Graph 30 - Residence and congestion information preference

It can be seen in Graph 30 that all those who live in the UK seem to prefer having the congestion information over not having it with the same ratio. However those living outside the UK seem not to be interested.

Finally, the results for the question asking how many extra minutes would be added to your journey if there were minor/major/severe delays gave an average of 6 minutes for a minor delay, 18 minutes for a major delay and 41 minutes for a severe delay.

This questionnaire has given valid results regarding passengers’ want for dynamic travel information. The sample of respondents does not exactly match the population of London but there is a good sample of both males and females, there is a large spread in ages, people completing all types of trips and people from all different sorts of residency were asked.

For the questions regarding passengers’ preferences in relation to information, in all cases it was shown with significance testing that without the chance of randomness passengers wanted dynamic information either about their line, journey or congestion. It was clear that in some cases the respondents were less interested in the information on their phone. However there seemed to be no difference between the sample groups in relation to the information wanted on the service boards. Overall though it was clear that those visiting London from outside the UK were less interested in dynamic information. Finally it can be seen that 64% of those asked would want information about queues and would potentially change their journey as a result, however, significance testing on this result shows that H_0 is accepted implying all results are chosen at random.

Given the results of the questionnaire, what a passenger believes is a delay can now be used as a threshold for defining a delay.

4.4.2. Validating the delay threshold

In order to determine if congestion and operational delays can indeed be discovered in the smart card data, first a delay to a passenger's journey needs to be defined. As the work in this thesis is aimed at passengers and information for them, it is important to take their opinions therefore the results of the questionnaire will be used to define a delay.

More specifically, the question:

How many extra minutes do you think there would be added to your journey if the service was

Minor delay ____minutes Major delay____minutes Severe delay____minutes

An average was taken of all the passengers' answers and an average of 6 minutes was found for a minor delay, 18 minutes for a major delay and 41 minutes for a severe delay. To determine if the passengers' perceptions of a delay were correlated to the length of time their journey took, regression analysis was completed to see if there was statistically significant relationship between the two.

In order to complete this analysis the average times passengers defined as a delay were compared to the predicted average journey times discovered in Section 4.2. The journeys used for this analysis were those that had entry and exit stations on the Victoria line as those are the ones that average travel times have been found for. The results for a minor delay are shown in Table 16, Table 17 and Table 18 (rounded to two decimal places).

Minor Delays

Table 16 – Regression analysis, Minor delays: Regression statistics

Regression Statistics	
Multiple R	0.08
R Square	0.01
Adjusted R Square	0.00
Standard Error	3.86
Observations	139.00

Table 17 – Regression analysis, Minor delays: Anova

ANOVA					
	df	**SS**	**MS**	**F**	**Significance F**
Regression	1.00	13.40	13.40	0.90	0.34
Residual	137.00	2041.83	14.90		
Total	138.00	2055.24			

Table 18 – Regression analysis, Minor delays: Correlation results

	Coefficient	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	6.48	0.98	6.60	0.00	4.54	8.42	4.54	8.42
X Variable	-0.07	0.07	-0.95	0.34	-0.22	0.08	-0.22	0.08

In Table 16 it can be seen that the R squared value is very close to 0. This implies that the regression line does not approximate the data well and that the variance in peoples' expectation of a minor delay is not well explained by their travel time.

Further the significance of F is reasonably large this confirms the result that there is not a strong relationship. This means that regardless of the length of a passenger's journey a minor delay is classified as 6 minutes.

The results for a major delay are shown in Table 19, Table 20 and Table 21 (rounded to two decimal places).

Major Delays

Table 19– Regression analysis, Major delays: Regression statistics

Regression Statistics	
Multiple R	0.10
R Square	0.01
Adjusted R Square	0.00
Standard Error	10.91
Observations	139.00

Table 20– Regression analysis, Major delays: Anova

ANOVA					
	df	**SS**	**MS**	**F**	**Significance F**
Regression	1.00	181.10	181.10	1.52	0.22
Residual	137.00	16297.09	118.96		
Total	138.00	16478.19			

Table 21– Regression analysis, Major delays: Correlation results

	Coefficient	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	19.48	2.77	7.02	0.00	13.99	24.97	13.99	24.97
X Variable	-0.26	0.21	-1.23	0.22	-0.68	0.16	-0.68	0.16

In Table 19 it can be seen by the R squared value and the significance of F that there is not a strong relationship between the predicted time of a passenger's journey and their estimation of a major delay.

Severe Delays

The results for a severe delay are shown in Table 22, Table 23 and Table 24 (rounded to two decimal places).

Table 22– Regression analysis, Severe delays: Regression statistics

Regression Statistics	
Multiple R	0.04
R Square	0.00
Adjusted R Square	-0.01
Standard Error	25.76
Observations	139.00

Table 23– Regression analysis, Severe delays: Anova

ANOVA					
	df	SS	MS	F	Significance F
Regression	1.00	126.08	126.08	0.19	0.66
Residual	137.00	90904.44	663.54		
Total	138.00	91030.52			

Table 24– Regression analysis, Severe delays: Correlation results

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	35.85	6.55	5.47	0.00	22.89	48.81	22.89	48.81
X Variable	-0.22	0.50	-0.44	0.66	-1.20	0.77	-1.20	0.77

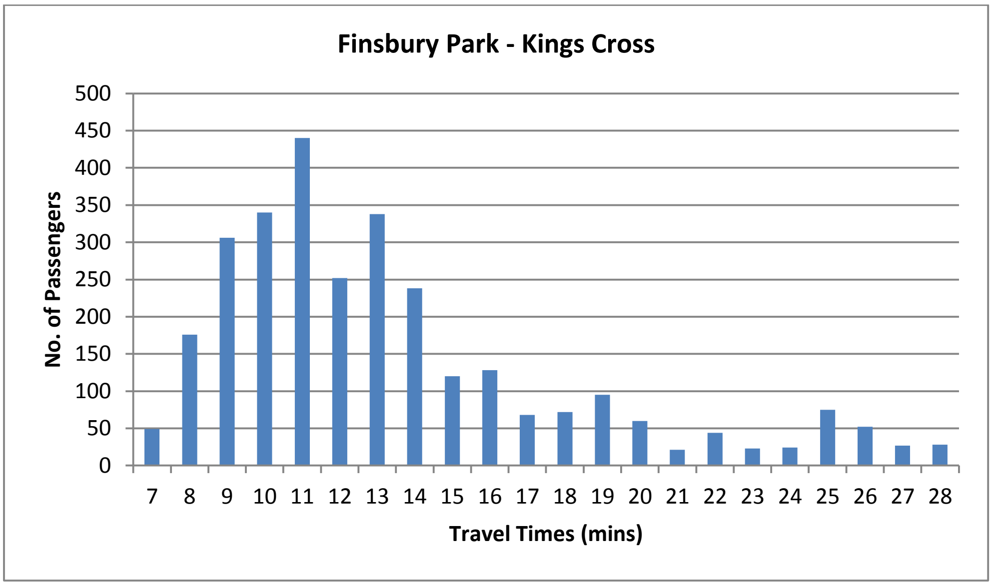
In Table 22 again it can be seen that the regression line does not fit the data as the R squared value again is 0. Therefore there is not a strong relationship between the predicted length of the passengers' journeys and the time they estimated that a severe delay would last.

From these results it will be taken that an individual's perception of a delay does not change with the length of their journey.

4.4.2.1. Historical Oyster Data

Now it is known from the sample of participants that answered the questionnaire that a passenger believes a minor delay is 6 minutes, a major delay is 18 minutes and a severe delay is 41 minutes regardless of the length of their journey. It is now essential to determine whether the minimum delay of 6 minutes is a possible minimum threshold to classify a delay.

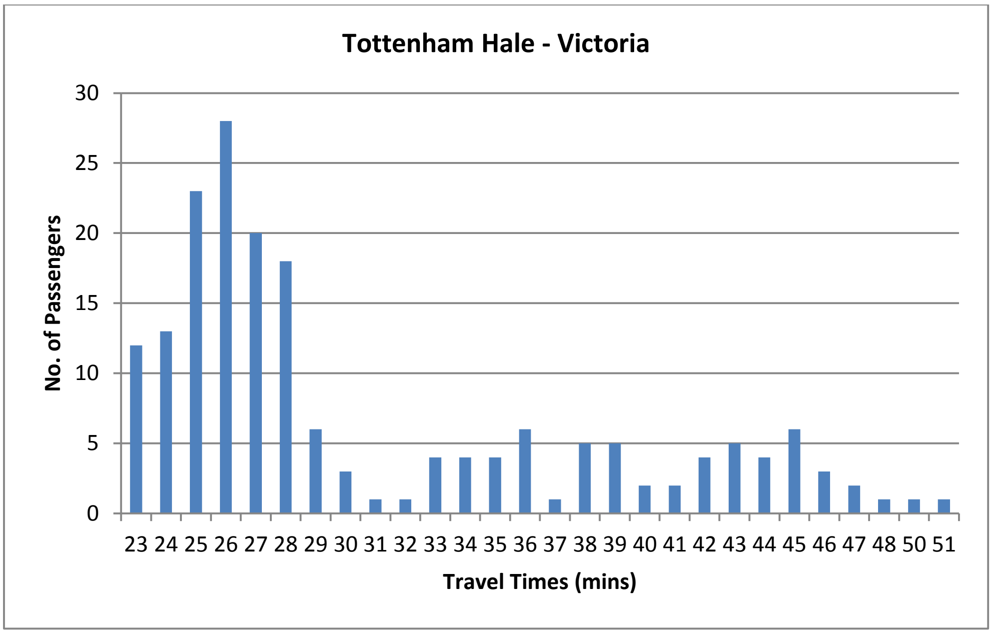
Three origin and destination pairs were used in this analysis. The three journeys chosen were Finsbury Park to King Cross, Tottenham Hale to Victoria and Finsbury Park to Oxford Circus. These three were chosen to gain a mixture of short and long journeys with different passenger demands. The data was taken from the files used to create the average travel times for all journeys on the Victoria Line; this data is described in Section 4.2.



Graph 31 – Frequency of travel times : Finsbury Park to Kings Cross

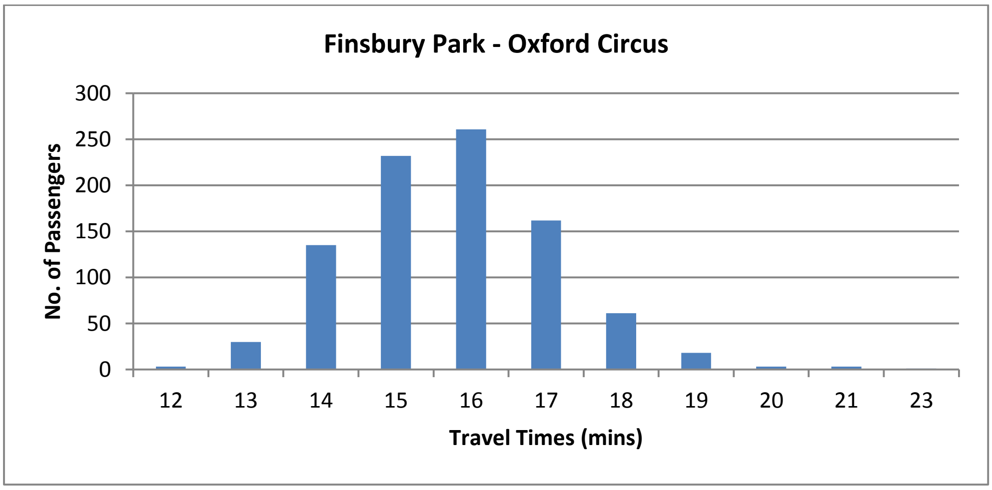
Graph 31 shows the travel times recorded for journeys completed for the origin destination pair Finsbury Park to Kings Cross against the number of people that completed the journey in that time.

This data is taken from 6 AM peaks spanning 8 weeks. For the journey of Finsbury Park to Kings cross the predicted journey time from the above data was 11 minutes. If there was a minimum delay threshold of 6 minutes over the mean time this would mean passengers are classified as delayed if their journey will take them more than 17 minutes. Within the 8 week data set only 10 passengers completed their journey in over 17 minutes. Out of 412 passengers this is 2% of the data.



Graph 32 – Frequency of travel times: Tottenham Hale to Victoria

For the journey of Tottenham Hale to Victoria the average journey time was 26 minutes. If there was a minimum delay of 6 minutes plus the mean travel time as the threshold, all passengers whose journeys were over 32 minutes would be delayed. Within the data set that spans 8 weeks, all passengers but 1 completed their journey in less than 32 minutes. Out of 172 passengers, that is 0.6% of the data.



Graph 33 – Frequency of travel times: Finsbury Park to Oxford Circus

For the journey of Finsbury Park to Oxford Circus the predicted journey time was found to be 15 minutes. This would imply the threshold for a delay is 21 minutes, found from 15 minutes plus the 6 minute threshold. Out 935 passengers' journeys spanning 8 weeks only 8 passengers' journeys were over 21 minutes which accounts for 0.8% of the data.

For each of these origin destination pairs, it is probabilistically unlikely when the London Underground is in a good service that a passenger should be delayed by more than 6 minutes. This implies that 6 minutes is a good base threshold that should minimise false reporting's of delays as stated in the methodology, Section 3.1.

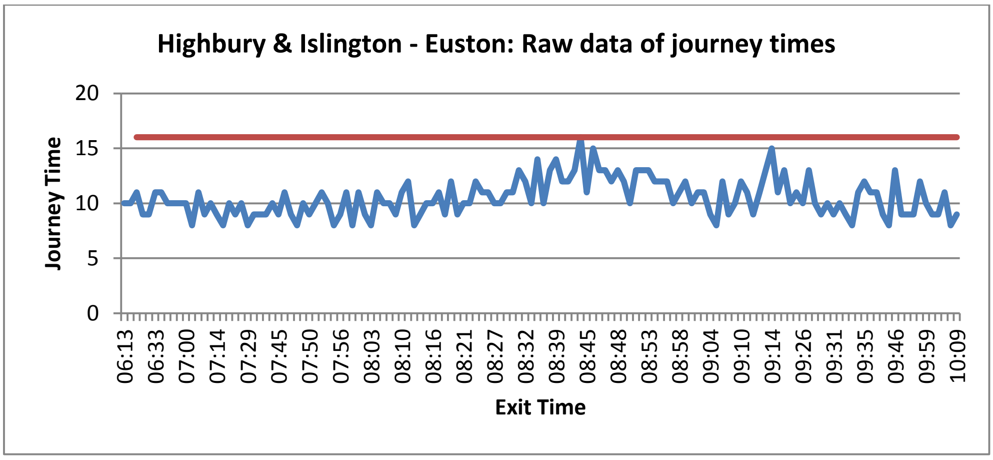
4.4.2.2. Real-Time

Finally to check that the base threshold of 6 minutes is not too sensitive to reporting delays, Oyster data that simulates a day in the network will be analysed to show that on a day with apparently no delays there are not too many passengers being classified as delayed.

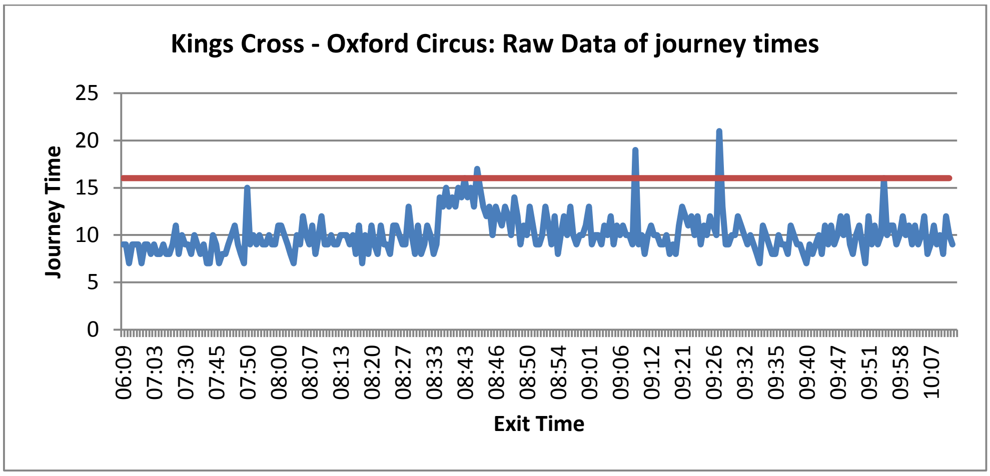
It is important that the threshold of reporting a delay is not too sensitive otherwise too many passengers will be reported as being delayed when they may not be. It will become untrustworthy information for the passengers if delays are reported when the service is fine. It is likely that some journeys will breach the threshold of the predicted journey time plus six minutes since anomalies may exist in the data, but the hope is that on an un-delayed day the number of cases will be minimal. An anomaly in this case would be a journey time that breaches the threshold and therefore is defined as delayed but may in fact not be caused by a congestion or service, delay. In this case there are multiple reasons why a passenger may be delayed when others are not such as: they may be new to the system and might have taken a roundabout route or they may have fallen unwell while in the system or they might have a large amount of luggage that is slowing them down.

In order to complete this work, a number of origin destination pairs were chosen; Highbury & Islington to Euston, Kings Cross to Oxford Street and Highbury and

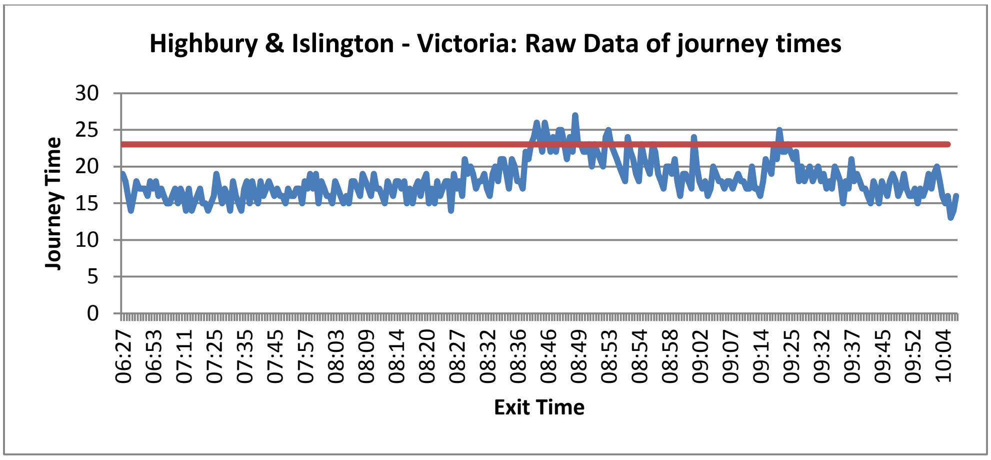
Islington to Victoria. This data is taken from the 29th October when there were no reported delays, Graph 34, Graph 35 and Graph 36 show the raw data of passengers' journey times for the stated journeys.



Graph 34 - Highbury & Islington to Euston: raw data



Graph 35 -Kings Cross to Oxford Circus: raw data



Graph 36 - Highbury & Islington to Victoria: raw data

In each of the graphs (Graph 34, Graph 35 and Graph 36) the respective threshold for classifying a delay is drawn in with a straight line. These values are the average travel time for the journey in question found from the Oyster data (Section 4.2) plus 6 minutes.

It can be seen the amount the raw data crosses the threshold line is variable for each journey. With Highbury & Islington to Euston being the least with it only touching the line once and Highbury & Islington to Victoria crossing in on multiple occasions.

With the threshold being breached inconsistently and on multiple occasions in Graph 36 this may lead to irregular information being returned to passengers. However given that the occurrences of the line being breached is around 08:45-09:00, this may be a good indication that passengers are becoming delayed when aiming to get into work on time for 09:00. In this case the delay threshold of 6 minutes would appear to be suitable at identifying delays. However, in this case some effort may need to be made to smooth the results such that the threshold of identifying a delay is breached on fewer occasions but perhaps for longer time periods. For the journey of Kings Cross to Oxford Circus it would appear that two of the incidents where the line is breached may be due to anomalies in the data. This would count as false reportings and need to be removed. This again highlights the need for smoothing of the data but does not indicate that this threshold may be too sensitive.

It was found in Section 4.4.1 through the questionnaire that passengers of the London Underground found on average 6 minutes to be a minor delay. This was used as a threshold in order to report a delay.

It was found by testing for significance that there was no correlation between the time it is estimated the passenger's journey took to complete and the length of time they believed was a minor delay. This led to the conclusion that regardless of the length of a journey on the Victoria Line a delay should be reported when passengers are delayed over 6 minutes over the expected time for that journey.

It was next shown that when looking at Oyster data over an 8-week period, in particular three origin – destination pairs, a maximum of only 2% of passengers travel times were over the threshold of the respective mean travel time plus 6 minutes. This would show that this threshold for defining a delay is not too sensitive.

It is important that a low number of total passengers will be experiencing a delay on a daily basis since for a day with no reported delay as it would be confusing to passenger if delays were continuously being reported. Over 8 weeks in can be seen that the majority of passengers are able to complete their journey in this time therefore the threshold of 6 minutes is an acceptable level for not being too sensitive. This result further shows that the average travel times are realistic to the variance gained from day-to-day perturbations in passengers' travel times. Finally the threshold of 6 minutes was tested with 'real-time' data. This shows during peak times there were cases of the threshold being breached which could indicate congestion in busy stations. However, there were a few cases of anomalies, that should the data be smoothed, could be removed. This, and the identification of congestion will be discussed in Section 4.5.

4.5. Congestion reporting

4.5.1. Real-Time data

In this section data will be analysed that simulates live Oyster data that has come directly from the ticket barriers. Currently it is not possible to obtain a data feed from the ticket barriers therefore data from previous dates will be used but will be simulated to represent real-time data. This means that a data entry can only be used from their time stamp onwards; i.e. a journey cannot be known until the time stamp of the exit station, as the case would be in real-time.

It is necessary to have data that simulates real-time data in order to understand if it possible to find information about the service as the data is being created. Currently there is a delay to receiving data for analysis from the central server of about a day (Transport for London, 2012). The thought is though that real-time data will be available in the future, therefore, it is important to understand what information is retrievable from the data about the network for future use.

Four data files were acquired from TfL for this analysis; the data was taken from four days each with different service statuses. These files were chosen to get a better understanding of passenger travel times during incidents and when there is a normal service. All of the delays on these dates took place on the Victoria Line during the AM peak[1].

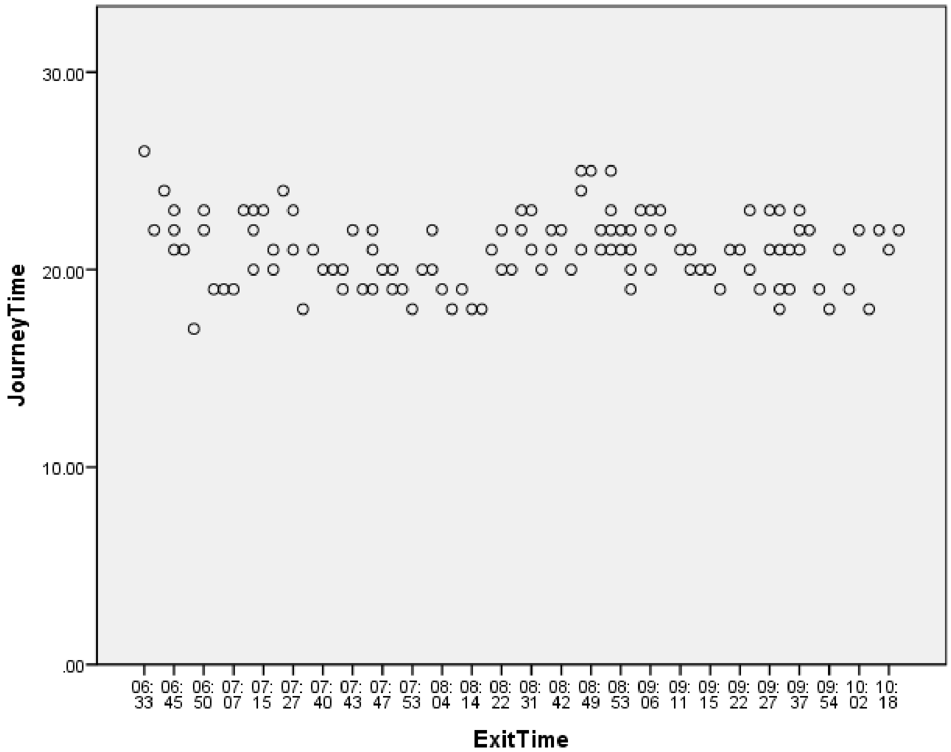
The four dates were the 2nd, 4th, 26th and 29th October 2012. As seen in Section 4.1 the data was paired by Oyster cards and journeys were determined since all of the journeys had both their origin and destination stations on the Victoria Line. As stated in Section 3.1 this was to ensure there were no ambiguous journeys taken in which it would not be possible to determine which route a passenger had used. These files were then ordered by exit time to simulate real-time Oyster data exiting the ticket machines. The speed of pairing entrance and exit pairs would, theoretically, be negligible therefore the journey is given the time stamp of the exit. Table 24 gives an example of the data and Graph 37 shows the data for all the journeys from Blackhorse Road to Euston.

[1] Unfortunately these data files excluded journeys exiting at Brixton due to the file being corrupted.

Table 25 : Example of raw data from 29th October dataset: Blackhorse Road to Euston

Entry Code	Entry Station	Exit Code	Exit Station	Entry Time	Journey Time	Exit Time	Journeys
522	Blackhorse Road	574	Euston LU	06:13	20	06:33	1
522	Blackhorse Road	574	Euston LU	06:30	27	06:57	1
522	Blackhorse Road	574	Euston LU	06:33	24	06:57	1
522	Blackhorse Road	574	Euston LU	06:20	37	06:57	1
522	Blackhorse Road	574	Euston LU	06:41	21	07:02	1
522	Blackhorse Road	574	Euston LU	06:42	20	07:02	1
522	Blackhorse Road	574	Euston LU	06:44	18	07:02	1

Blackhorse Road – Euston: Raw Data



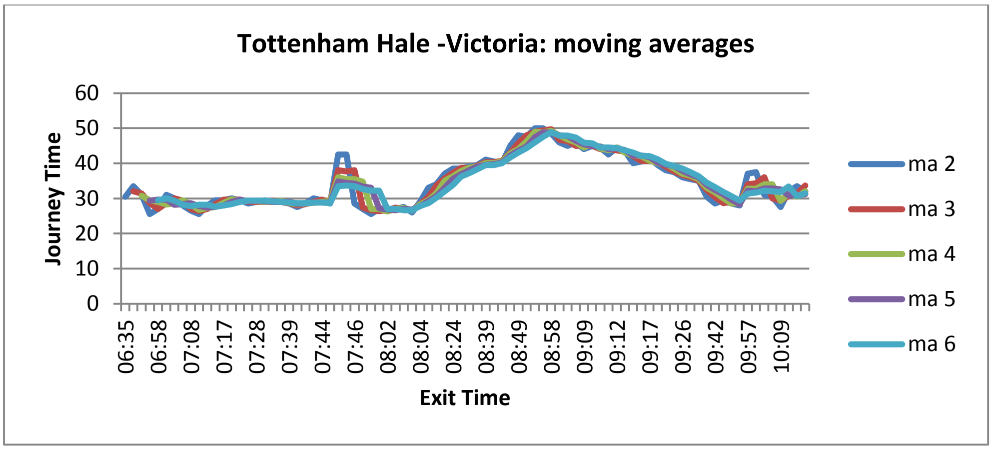
Graph 37 – Blackhorse Road – Euston: Raw data

The range of the data for the journey from Blackhorse Road to Euston is from 17 minutes to 26 minutes with an average travel time of 21 minutes, which is the same time found for the average travel time from the 8-week span data set (Section 4.1).

To determine when there is a delay to passengers, it was decided that a moving average should be calculated to smooth the data to remove noise in the data that could lead to inconsistent delay reporting. The intention of this is to remove any extremities found in the data and gain a more useable data set. This is an important considereation within this methodology to ensure the return of information is as smooth as possible, with reduced noise and improved incident reporting as stated in Section 3.1. It was seen in Section 4.4.2.2 that without a moving average more anamolous journeys are classified as delayed.

Further, although neither the moving average data or raw data is continuous it will be made to be continuous since it is necissary that passengers can receive information at any time.

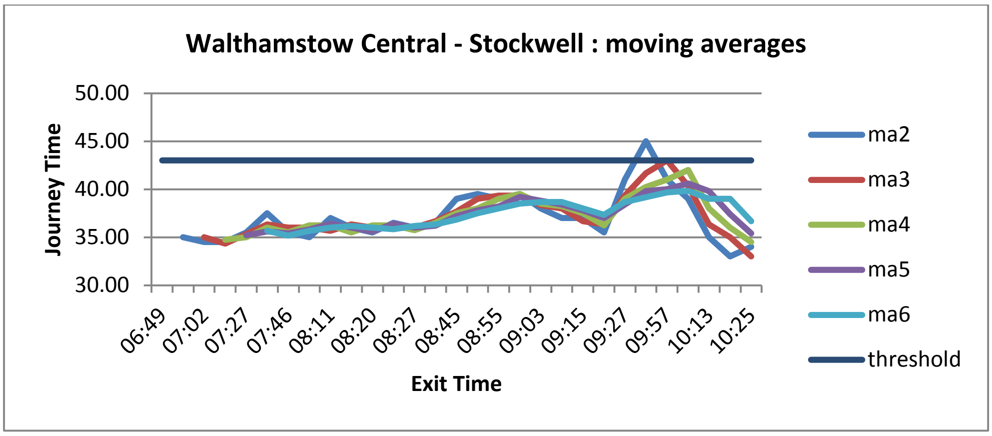
With the process of taking a moving average there is a trade off between time delay of the data and the smoothness. It was decided that the data should be smoothed as much as possible, this was to ensure the information returned to passengers was as reliable as possible. If there are abrupt changes in the information provided passenger’s may be inclined not to trust the information as discussed in Section 2.1.2 . In the process of determining what moving average should be used, it is important that the delay in the return of information isn’t too substantial. Graph 38 shows the different possibilities of moving averages for a mornings journeys between Tottenham Hale and Vicoria, where ma2 means two journey times were averaged, ma3 means three journey times averaged etc. In all cases the time stamp given to the moving average value of the time of the last entry to contribute to the average, as this is the earliest it could be calculated.



Graph 38: Tottenham Hale to Victoria: moving average possibilities

It can be seen in Graph 38, as more values contribute to the moving average the time stamp of the point moves further along. Yet the more values included in the average the smoother the data is.

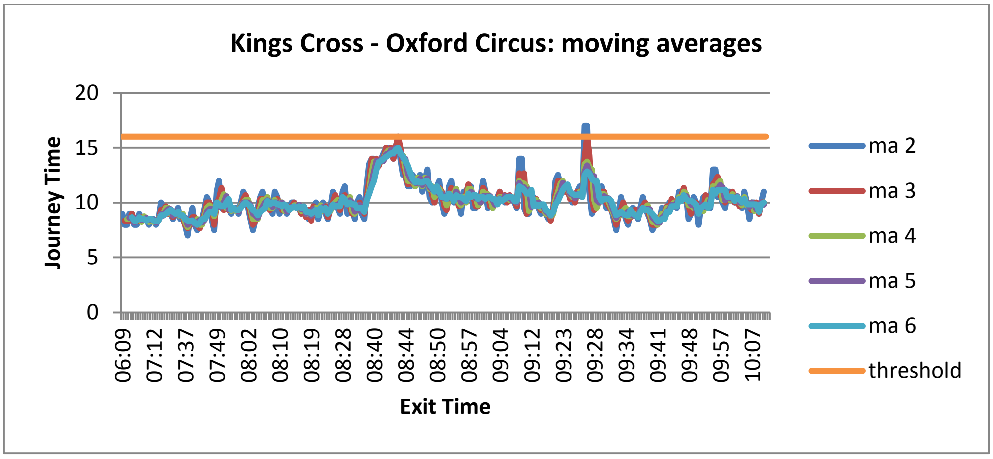
To determine which values of moving averages smooths the data sufficiently two further journeys are to be studied. These are shown in Graph 39 and Graph 40.



Graph 39 - Walthamstow Central to Stockwell : moving average possibilities

The average journey time for Walthamstow Central to Stockwell is found to be 37 minutes. The delay thresehold of 6 minutes, described in Section 4.4, is breached in one instance at 9:27 with a passenger taking 47 minutes to complete their journey.

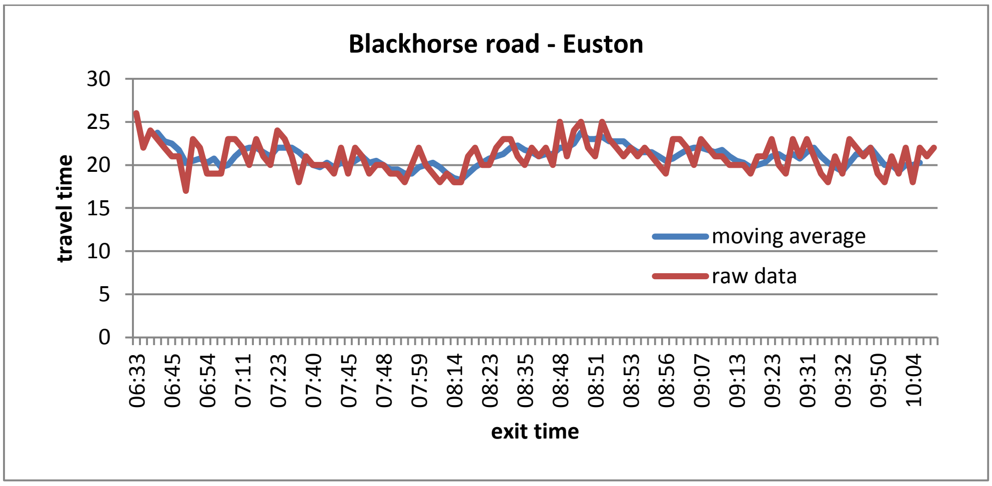
Seeing as this is the only passenger that has this value it should be seen as an anomaly. It can be seen that when two or three moving average points are used this anomaly shows as a delay to the data by being over the threshold of 43 minutes, however, when four, five or six moving average points are considered no delay is found.



Graph 40 - Kings Cross to Oxford Circus: moving average possibilities

In Graph 40 it can be seen again that two and three moving average points find the anomalies, this was seen to be a repetitive theme across different journeys across the line. At this point it is determined that four moving average points successfully remove anomalies from the data making the travel time information smoother. It was decided to look no further at moving averages above 5, as 4 moving average points remove anomalies yet return the data quicker than 5 and 6.

To calculate the moving average four successive values were added together and divided by four. This new value was given the time stamp of the last entry Graph 41 shows this new data.



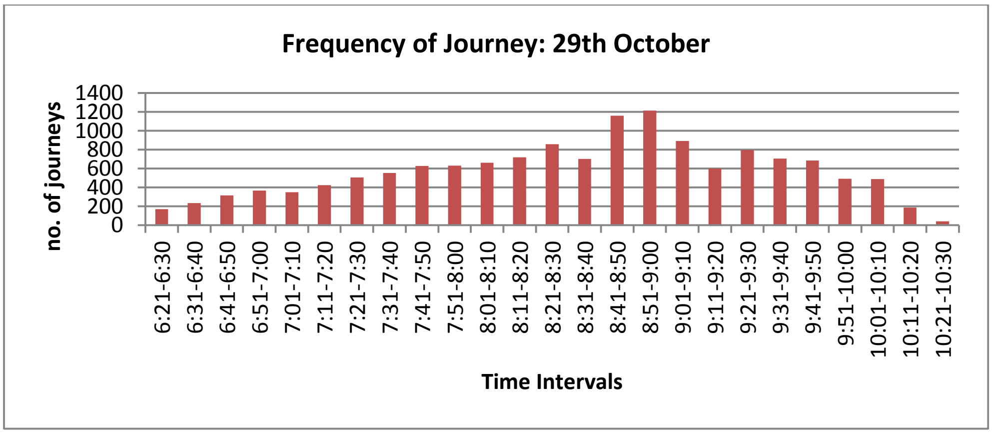
Graph 41 - Blackhorse road to Euston: raw data and moving average

It can be seen in Graph 41 that there is less noise and any rise in travel time is gradual. This will help reduce the sensitivity of the algorithm in finding false delays. The moving average fits the data and smooths out the peaks and troughs.

To determine if congestion can be seen in the data, the morning of the 29th October was analysed. No service problems or any other kind of delays to the train operation were reported by TfL on this day (Transport for London, 2012). Therefore should there be any delay to passengers' travel times it will be as a result of congestion or an unreported delay.

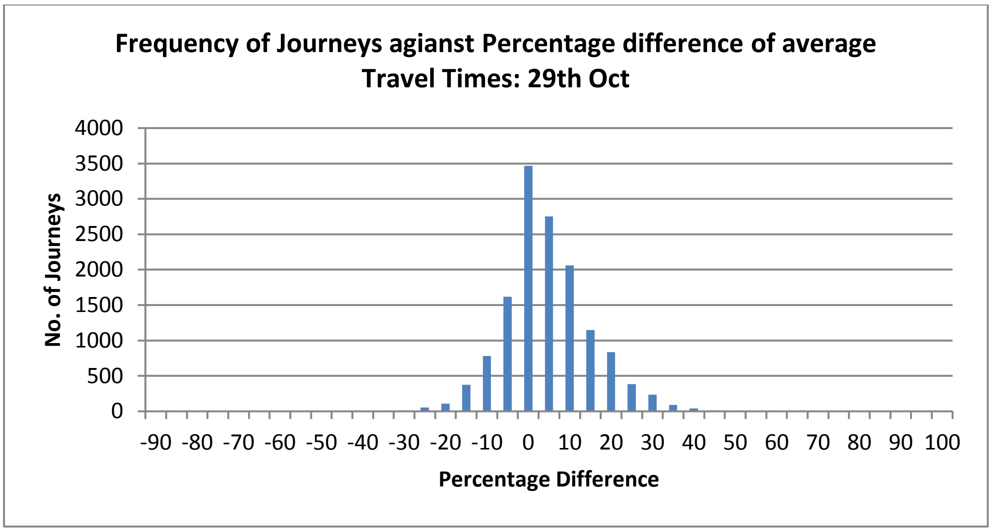
Passengers can experience delay in different places during their journey in the underground. This can be when they enter the system and a queue forms to board a train, or during their journey when they are on the train; this delay can be increased when there is a large passenger demand due, to possessions being stuck in the doors and thus increasing the dwell time, or finally, when a queue forms at the ticket barriers when leaving the station due to the exit being a bottleneck, as discussed in Section 3.2.1.5.

Graph 42 shows the number of passengers on the Victoria Line over the morning peak on the 29th October that have their origin and destination on the Victoria Line and are traveling Southbound.



Graph 42- Frequency of journey: 29th October

In Graph 42 it can be seen that there is a clear increase in the number of passengers travelling between 08:40 and 09:00. There is also a slight rise again between 09:20 – 09:30 but not as high. This would suggest that there is a higher chance of passengers being delayed within these times as with higher demand there is more chance a passenger may need to wait for a second train or a delays may be incurred to the train or a queue may form exiting. Further, to see if congestion is forming on this day. Graph 43 shows the distribution of journey times in comparison to the average travel times found.



Graph 43 - Frequency of journeys agianst percentage difference of average travel times: 29th Oct

To make a scale that is appropriate for all journeys the x co-ordinate in Graph 43 is the percentage difference between the passengers' travel times on the day of the 29th October and the average travel time found from the historical data found in Section 4.2.

If there were to be no delay to any passenger on this morning it would be expected that the distribution would be a distribution close to being centred on 0. However, it can be seen that the distribution is skewed slightly to the right, with a skew value of 0.3. The skew of a distribution is calculated by the following formula:

$$\frac{n}{(n-1)(n-2)}\sum\left(\frac{x_i-\bar{x}}{s}\right)^3$$

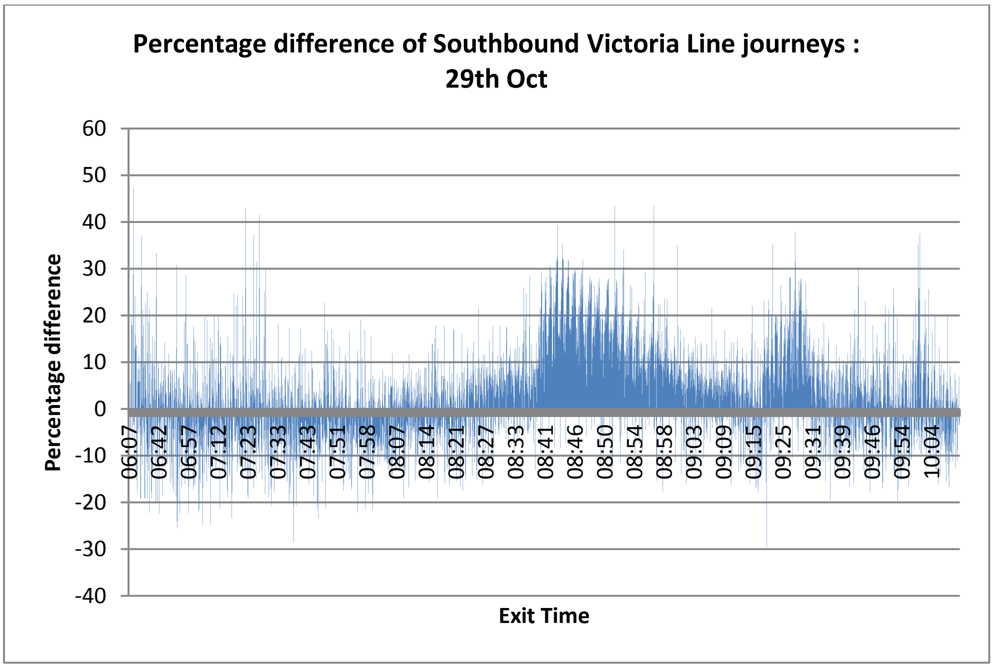
Equation 16

Where n is the number of data points, $\bar{x}$ is the mean and s is the standard deviation, which gives a value to discuss skew that is dimensionless.

To understand how to classify the different skew values a threshold of greater than one or less than minus one seems to be popular within the literature (Garver, 1932)(Hotelling and Solomons, 1932) ("CBU statistics Wiki," n.d.)("GraphPad Statistics Guide," n.d.). If values should fall outside of this range then it is acknowledged as a prominent skew.

The lack of a prominent skew would lead to the hypothesis that there are not in fact any detectable delays to passengers' times due to congestion. It is expected that a day with a delay may well be seen to be skewed since the days (in real-time) are compared to the average taken over 7 days. Therefore it is expected that should the day (in real-time) just be 'average' with no problems it should have a normal distribution like the average travel times found.

Finally the percentage difference of the passengers' journeys is plotted against the exit times of the journeys and this is shown in Graph 44.



Graph 44 - Percentage difference of Southbound Victoria Line journeys : 29th Oct

Graph 44 shows a clear rise in journey times between 08:40-09:00 and between 09:20-09:30 naturally some of these would be due to anomalies though. In order to determine which are genuine delays, a further constraint is needed to ensure a number of passengers are experiencing a delay rather than just one moving average point showing a higher time due to one anomaly pushing up the mean, this idea is first discussed in Example 2.

This extra constraint will look at having a minimum number of passengers classified as delayed before a decided there is a delay. Since a moving average is taken of the data, this analysis will look at how many moving average points are delayed within the same minute to classify a delay. As discussed in Section 3.2.1.5 the different places a passenger could experience a delay are determined as delayed by whether other passengers have experienced a similar delay. Therefore station delays (entrances and exits) and line delays have been examined separately to determine how many delays would appear with different numbers of moving average points showing delays.

For the exit stations, 2, 3, 4 and 5 moving average points being delayed in a minute have been considered shown in Table 26. When deciding which will be the most appropriate result it is important to find a balance of not having too many delays registered or having too few. If too many delays are registered, passengers may be inclined not to take the advice of the information since the service would always appear to have delays and in that case a delayed service would then become the norm. On the other hand if too few delay results are found this may lead to passengers believing the information is unreliable as the delay reports are patchy and inconstant.

For the line delays again 2, 3, 4 and 5 moving average points have been considered shown in Table 27. On a day such as the 29th October when there were no reported delays, it is important that no line delays are discovered, since there were no operational delays, unless it looks likely that there is an unreported delay. If there are reports of delays, every day, on days when there are delays passengers may not understand the severity of the change in the system.

Table 26: Exit station delays, comparsion of common delays in a minute.

	Exit Station	Number of delays in common in the same minute			
		2	3	4	5
Time of Delays	Euston	07:23	07:23		
		07:27			
		07:28			
		07:30	07:30		
		07:41			
	Green Park	08:47			
		08:49	08:49		
	Highbury & Islington	07:29			
		07:57	07:57		
		08:21			
		08:51			
		09:00			
	Oxford Circus	08:40	08:40	08:40	08:40
		08:41	08:41	08:41	08:41
		08:42	08:42	08:42	08:42
		08:43	08:43	08:43	08:43
		08:47	08:47	08:47	
		08:48	08:48		
	Pimlico	09:26	09:26	09:26	09:26
		09:27	09:27	09:27	09:27
		09:28	09:28	09:28	09:28
	Vauxhall	08:48	08:48	08:48	08:50
		08:50	08:50	08:50	09:28
		09:28	09:28	09:28	
		09:30			
	Victoria	08:44			
		08:45	08:45		
		08:46	08:46	08:46	08:46
		08:47	08:47	08:47	08:47
		08:48	08:48	08:48	08:48
		08:49	08:49	08:49	08:49
		08:50	08:50	08:50	08:50
		08:51	08:51	08:51	08:51
		08:52	08:52	08:52	08:52
		09:27	09:27	09:27	09:27
		09:28	09:28	09:28	09:28
		10:01			
	Warren Street	08:43	08:43	08:43	08:43
		08:44	08:44	08:44	08:44
		08:45	08:45	08:45	08:45
		08:46			

Table 27: Line delays, comparsion of common delays in a minute.

	Number of OD pairs with delays in the same minute			
	2	3	4	5
Time of reported Line Delays	07:23:00	08:42:00	08:42:00	
	07:27:00	08:43:00	08:44:00	
	07:29:00	08:44:00	08:47:00	
	08:40:00	08:45:00	08:52:00	
	08:41:00	08:47:00	09:28:00	
	08:42:00	08:48:00		
	08:43:00	08:52:00		
	08:44:00	08:53:00		
	08:45:00	09:28:00		
	08:46:00			
	08:47:00			
	08:48:00			
	08:49:00			
	08:50:00			
	08:51:00			
	08:52:00			
	08:53:00			
	08:57:00			
	08:58:00			
	09:27:00			
	09:28:00			
	09:29:00			

For a day with no reported delays and no obvious un-reported delays it is important that a delay should not be reported through Oyster data as this will lead passengers to not understand the difference between an un-delayed day and a delayed day since this would be a false positive.

As seen in Table 27 there are no reported delays when 5 moving average points are found in the same minute and it would appear there are no un-reported delays, therefore this seems to be a good initial threshold for defining a line delay since when looking at the other possible values no indication of an un-reported delay is seen. In Table 26 it can be seen there is little difference between taking 4 moving average points and 5 moving average points. Whereas when 3 moving average points in the same minute is considered the delay statuses become more patchy, therefore 4 moving average points shall be taken as the initial threshold for defining an exit delay.

For exit delays the threshold has now been set that if there are 4 moving average points in the same minute with a delay over 6 minutes, a delay is to be reported and for a line delay the threshold has been set that if there are 5 moving average points in the same minute with a delay over 6 minutes, a delay is to be reported. These delay reports are shown in Table 28. There is no table showing line delays as no line delays were found for this day.

Table 28 – Results: Exit delays, 29th October

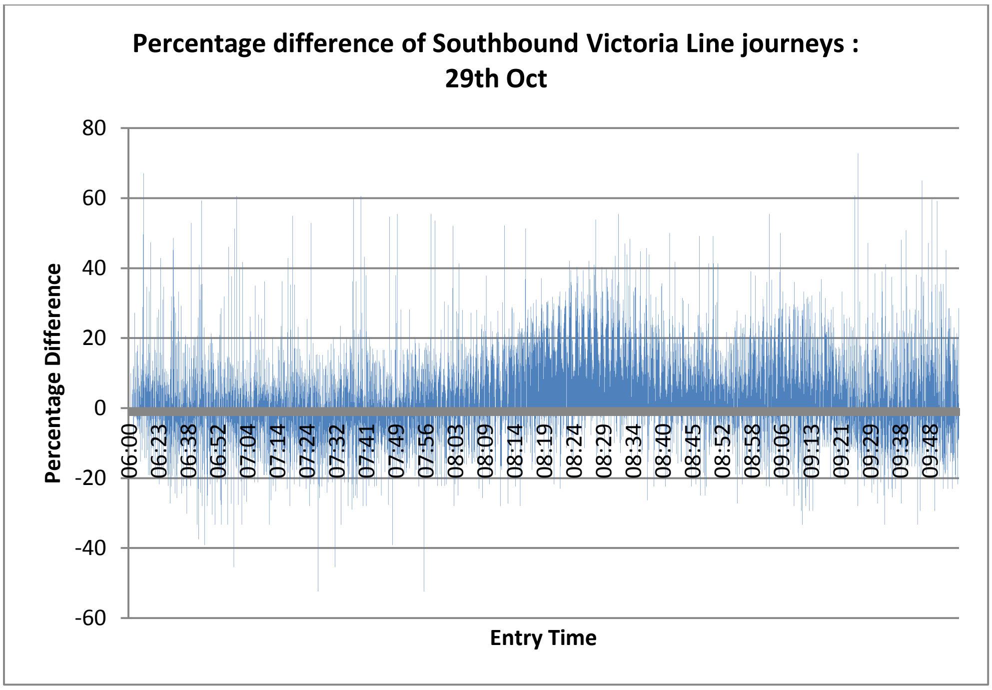
Time	Exits with delay	Average Delay (mins)
08:40	Oxford Circus	1
08:41	Oxford Circus	2
08:42	Oxford Circus	1
08:43	Oxford Circus	1
	Warren Street	2
08:44	Warren Street	2
08:45	Warren Street	2
08:46	Victoria	2
	Warren Street	1
08:47	Oxford Circus	1
	Victoria	1
08:48	Victoria	1
08:49	Victoria	1
08:50	Victoria	1
	Vauxhall	1
08:51	Victoria	1
08:52	Victoria	1
09:26	Pimlico	1
09:27	Victoria	1
	Pimlico	2
09:28	Victoria	1
	Vauxhall	1
	Pimlico	1

Table 28 shows that there are a number of exit stations that are congested in the peak times. These results coincide with the rises in passengers travel times seen in Graph 44, which shows clear peaks in passengers' travel times between 08:40 and

08:55 and again a peak around 09:28. The delays in Table 28 are only marginally greater than the 6 minute threshold and do not last very long in most cases.

The way the passengers and operators receive delay information is an important concern. This information is showing that some stations are clearly more susceptible to delays at certain times. On a daily basis there may not be much option for passengers to re-route when receiving news that a station is experiencing delays, especially when the delay is only for a few minutes. However, if the results were the same over a number of days or weeks this may lead passengers to changing their patterns over a longer period.

For the operators of the system this information can help to understand where the bottlenecks are in the system. In some cases these can be prevented by changing the direction of the ticket barriers such that more are in one direction. Further, these results show it is possible to know where there is high passenger demand within the network. This information can be valuable to organise in emergency conditions or to know when stations should be closed for safety due to the large number of passengers. Finally, to understand the pattern forming on an un-delayed day, Graph 45 shows the difference in travel times that passengers have completed on the 29th October with the average travel times as a percentage. This is plotted against the time the passengers entered the network.



Graph 45 - Percentage difference of southbound Victoria Line journeys : 29th Oct, sorted by entry time

It can be seen that the passengers who enter the system between 08:15 and 08:40 are those that are experiencing the largest delays. In particular those entering the system between 08:25 and 08:35 are almost certainly going to be delayed as it can be seen that very few passengers' times are less than the 0 mark.

In order to see where the delays are occurring at entrances, the same algorithm was used that found exit delays. This was; to define a delay if 4 moving average points were found in the same minute that have an increase in their travel time of over 6 minutes over the average travel time for that journey. Table 29 shows the results of the algorithm, showing the congested entry station.

Table 29 – Results: Entrance delays, 29th October

Time	Entrances with delay	Average Delay (mins)
08:08:00	Blackhorse Road	1
08:09:00	Blackhorse Road	1
	Walthamstow Central	1
08:12:00	Tottenham Hale	2
	Walthamstow Central	1
08:13:00	Walthamstow Central	1
08:14:00	Blackhorse Road	2
08:15:00	Blackhorse Road	1
	Walthamstow Central	1
08:16:00	Tottenham Hale	2
08:17:00	Tottenham Hale	2
	Walthamstow Central	1
08:18:00	Finsbury Park LU	1
08:19:00	Finsbury Park LU	1
	Tottenham Hale	1
08:20:00	Finsbury Park LU	1
	Highbury & Islington	1
	Tottenham Hale	1
08:21:00	Finsbury Park LU	1
	Highbury & Islington	1
08:25:00	Highbury & Islington	1
08:27:00	Highbury & Islington	1
	Euston LU	1

Table 29 shows there are delays to entrances between 08:08 and 08:27. Further, that the delays seem to be migrating along the line as time progresses.

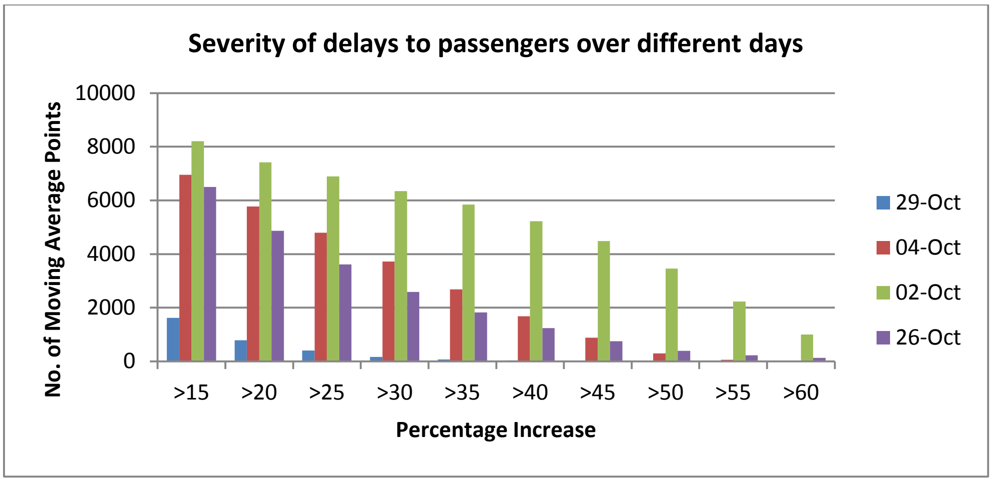
In conclusion, an algorithm shows that no line delays have been defined. Yet it can be seen there are delays to entrances and exits at different times. It was shown that the delays to the entrances were between 08:08 and 08:27 and the delays to the exits are between 08:40 and 08:52 with a smaller delay around 09:28.

4.6. Delay reporting

As seen in Section 4.5 it is possible to use Oyster data to see where congestion is arising in the network and how it affects passengers' travel times. It is now important to determine what delays to passenger's travel times can be seen through Oyster data when the delay is caused by a service fault or an incident on the network. To understand what information is available about the current network, when there are service problems, data for three days with operational delays (2nd, 4th and 26th of October 2012) will be analysed, the official reports from TfL about these dates are as follows.

i. 2nd October: a person went under a train at 08:40; there was a partial line suspension between Victoria and Brixton between 09:00 - 10:30; this led to severe delays until 11:15 along the whole Victoria Line.
ii. 4th October: there was a signal failure at Vauxhall at 08:20 this led to minor delays across the entire line until 14:00.
iii. 26th October: there was a faulty train at 07:30 causing severe delays between Walthamstow Central and King's Cross between 07:30 - 07:45; which lead to minor delays for the rest of the line these minor delays continued across line until 09:30.

Graph 46 shows the number of moving average points lying in each band of different percentage increases, on 2nd, 4th and 26th October, with the no-delay case of 29 October shown for comparison. To make a fair comparison the *x* axis shows the percentage increase from the average travel times, discovered in Section 4.2, to the moving average points on each of the dates. These percentage increases have been collected together in increments of 5% at a time.



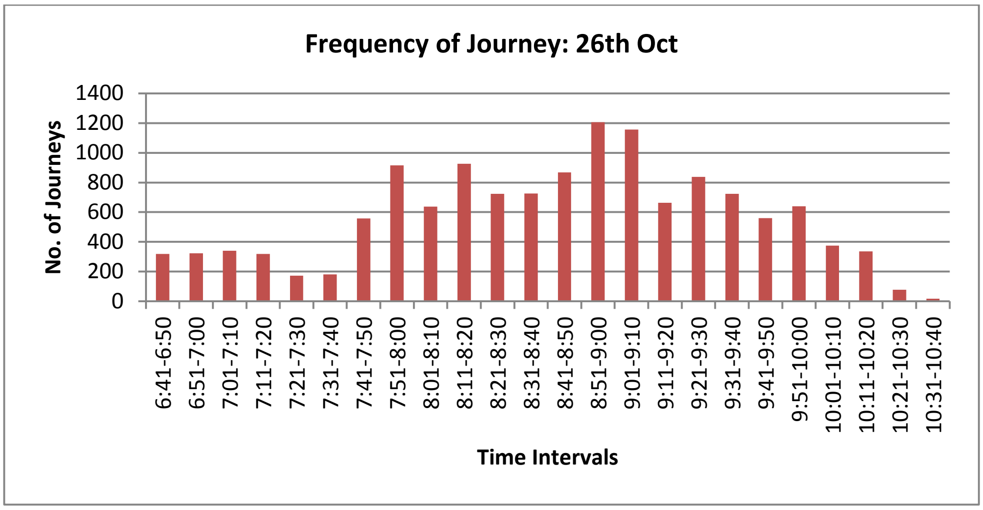
Graph 46 - Severity of delays to passengers over different days

Graph 46 shows that there are a very small number of passengers reaching higher than a 15% increase in their travel time on the 29th October, the day with no reported delay. However for the other dates it is clear how much the different service problems affect the passengers' travel times. Yet it can be seen that although there is not a substantial difference between the 4th and the 26th of October the travel times and passengers on the 26th are affected slightly less on this date. This date will be analysed first.

4.6.1. 26th October 2012

To initiate analysis for the day, the frequency of passengers will be studied. It is not possible to determine whether the service status reduced the number of passengers, firstly due to the data being used is only a sample of passengers who will have used the Victoria Line on this date, since all those that had either their entrance or exit on other lines have been removed and it is unknown whether passengers may have used the line but had their origin and destination on other lines. Secondly, there will be day-to-day perturbations between the numbers of passengers for whom there is data.

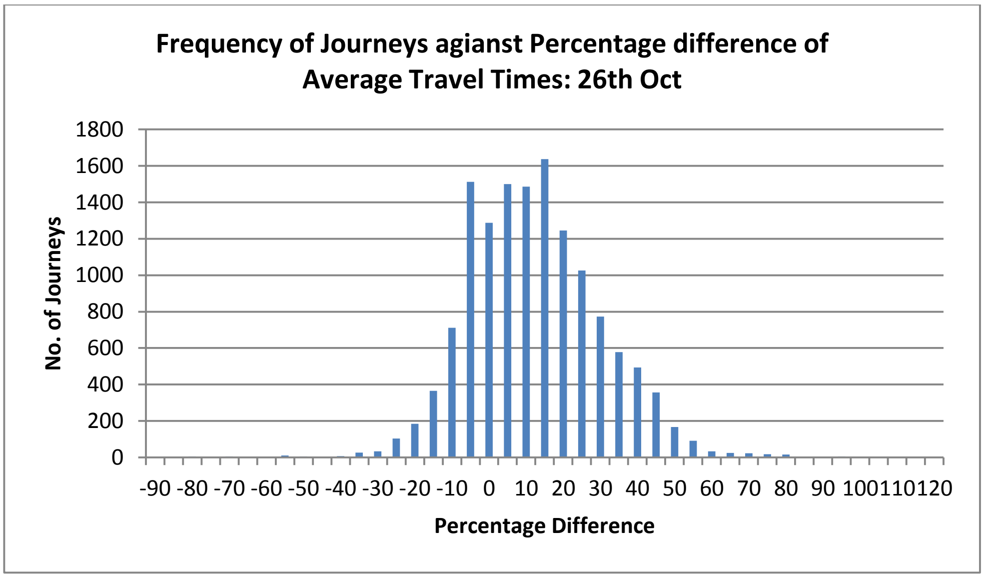
Graph 47 shows the frequency of passengers at different time intervals throughout the morning of the 26th October.



Graph 47 - Frequency of journey: 26th Oct

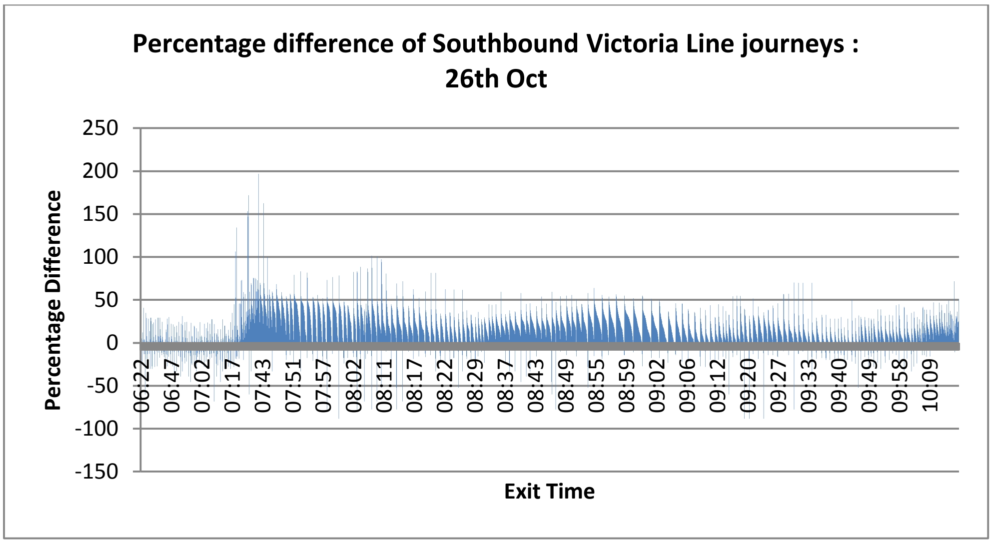
Graph 47, in comparison to Graph 42, shows that there is a clear difference in the pattern between 7:21 and 8:20. For the 29th of October there was a gradual increase as the time intervals increases between 08:30 and 09:00 and around 09:30, however Graph 47 shows that this time period is unstable on comparison.

Graph 48 shows the distribution of journey times in comparison to the average travel times found.



Graph 48 - Frequency of Journeys agianst Percentage difference of Average Travel Times: 26th Oct

The distribution here differs to that on the non-delayed day shown in Graph 43. For a non-delayed day it was centred on 0 but skewed slightly to the right, with value 0.3, the value of this skew, however, is 1.5, over the threshold to classify as a prominent skew. Here it can be seen that the peak of the graph never reaches as high as the peak on the un-delayed day and there is more of a spread between 0 and a 20% increase. It can also be seen that the tail of the graph continues for longer. It was previously seen when there was no delay that the majority of passengers didn't experience increases to their travel times greater than 30% yet here it can be seen that there are a number of passengers experiencing longer delays.



Graph 49- Percentage difference of Southbound Victoria Line journeys : 26th Oct

Graph 49 shows how passengers' travel times are affected by the delays experienced, with the percentage difference between the average travel time and the journey times as they are recorded at the exit ticket barriers. Transport for London (2012) states there were severe delays to passengers between 07:30-7:45 yet here it can be seen that there are clear delays until 8:25 when travel time reaches a minimum delay. However, the passengers' times increase again at this point due to the usual morning peak. This lasts for longer than the congestion seen on the 29th October shown in Graph 44 and passengers are experiencing greater delays. Unusually there seems to be a rise around 09:44 that increases until the end of the data set. For 15 minutes after the 09:30 congestion peak passenger times seem to recover to be free flow again until another rise. There is no information concerning this final rise from TfL and as seen on an un-delayed day there is no congestion at this time to explain the final rise.

In order to see how many of these rises in travel times were classified as delays on the 26th October a simple program was created to see for each origin-destination pair when they go over their respective threshold of the average travel time plus 6 minutes. This program showed that 2836 moving average points were found to be over their respective threshold. This is 11344 journeys out of a possible 22193 journeys (approx. 50%) recorded that morning of the 26th October that had the origin and destination on the Victoria Line. This is many more passengers experiencing delays than on the 29th October when there was no delay.

Table 30 shows the results of all reported line delays on the 26th October in the AM peak for the southbound Victoria Line and their respected average delay to the line.

These results in general concur with the report given by TfL (Transport for London, 2012). The next threshold given by the passengers in the questionnaire showed that a major delay is one that lasts over 18 minutes, seen in Section 4.4.1. Graph 46 shows that this is breached a few times between 07:42 and 07:52 therefore would be classified as a major delay within this time.

Table 30: Results: Line delays, 26th October

Time	Line Delay (mins)
07:35:00	11
07:36:00	10
07:42:00	16
07:43:00	13
07:44:00	14
07:45:00	16
07:46:00	13
07:47:00	14
07:48:00	15
07:49:00	14
07:50:00	13
07:51:00	16
07:52:00	12
07:53:00	11
07:54:00	9
07:55:00	10
07:56:00	11
07:57:00	9
07:58:00	9
07:59:00	12
08:00:00	10
08:01:00	8
08:02:00	5
08:03:00	8
08:04:00	9
08:05:00	9
08:08:00	8
08:09:00	11
08:10:00	11
08:11:00	13
08:12:00	12
08:13:00	7
08:14:00	7
08:15:00	7
08:16:00	5
08:17:00	6
08:18:00	4
08:19:00	5
08:20:00	4
08:21:00	3
08:22:00	3
08:23:00	5
08:26:00	1
08:29:00	3
08:32:00	1
08:33:00	1
08:34:00	1
08:36:00	3
08:37:00	2
08:38:00	2
08:39:00	3
08:40:00	2
08:41:00	3
08:42:00	3
08:43:00	3
08:44:00	2

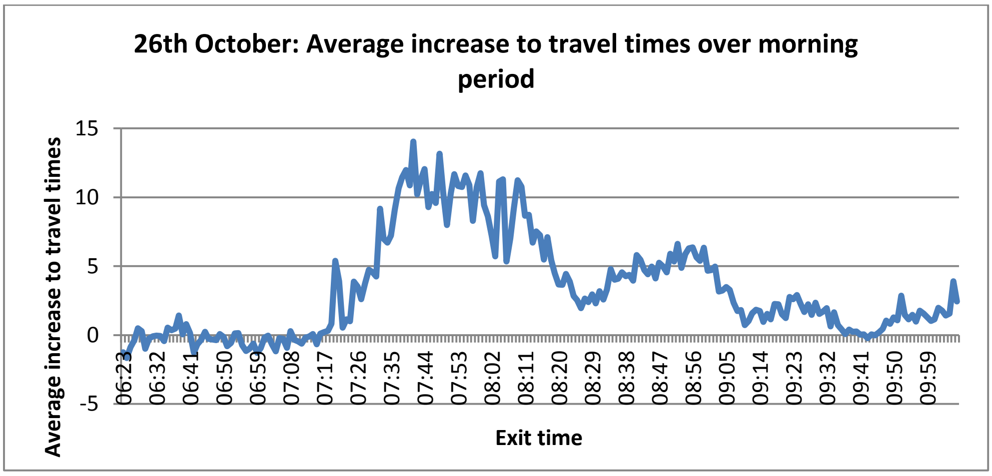
08:45:00	3
08:46:00	3
08:47:00	2
08:48:00	2
08:49:00	3
08:50:00	2
08:51:00	3
08:52:00	2
08:53:00	3
08:54:00	3
08:55:00	3
08:56:00	3
08:57:00	2
08:58:00	4
08:59:00	3
09:00:00	3
09:01:00	3
09:02:00	3
09:03:00	3
09:04:00	2
09:05:00	2
09:06:00	2
09:07:00	3

In Table 30 it can be seen that there is a slight discrepancy between the times that the delays are reported and the official report. TfL reported (Transport for London, 2012) that the delay started at 07:30 at Walthamstow Central. Between 07:30 and 07:35 there are delays that appear at the entrance to Walthamstow followed by the entrance to Blackhorse Road, however there are not enough passengers showing delays to classify the line as delayed until 07:35.

Further, the TfL report says that the delay to the line lasted until 09:30 but these results show that the delay finished at 09:07. There are entrance and exit delays after 09:07 but again not enough passengers were delayed for a conclusion to be drawn that the entire line was delayed. In the appendices Table 75, there is the full list of all entrance and exit delays. What can be clearly seen throughout the morning is that all entrance delays seem to be those that are starting at Walthamstow Central, Blackhorse Road, Tottenham Hale, Seven Sisters and in some cases Highbury & Islington.

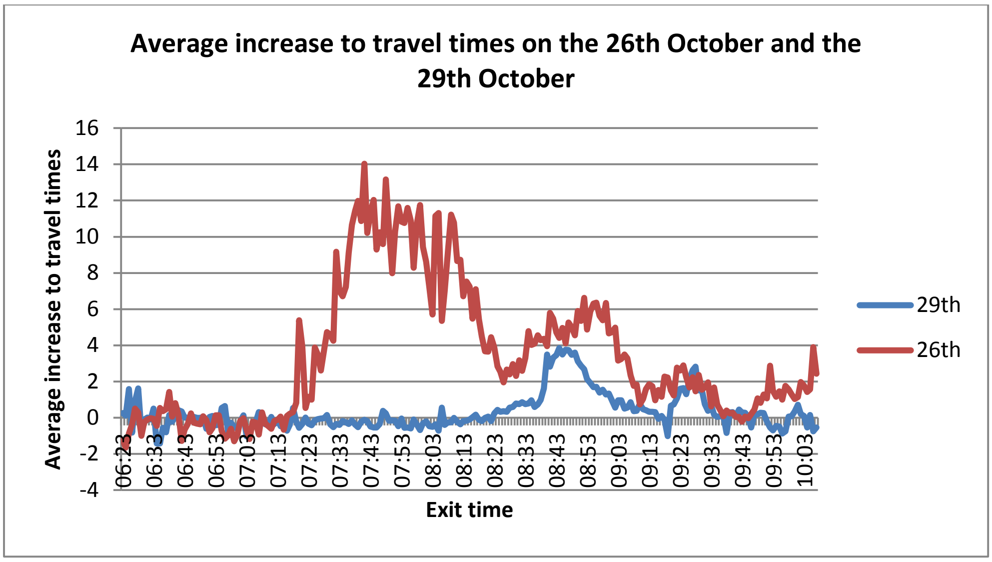
Unfortunately, as stated earlier, Brixton as a destination station is missing from the set. This may indeed affect the results: should these journeys have been in the data set, it would have perhaps made the results smoother with less stop and starts of line delays as there would have been more data contributing to the results and it may have shown the delay lasting longer. However, it would not show the delay starting any earlier since Brixton is the last station on the line therefore passengers that are delayed would not reach there as quickly. The results shown Table 30 to seem to agree with Graph 49, and this suggests that the algorithm seems to be working efficiently.

Finally to understand how congestion and overcrowding of passengers contributes to the delays on the day, the average increase in journey time over the morning has been found. On an un-delayed day it would be expected that this average would be 0, since the times are compared to the average travel times found in Section 4.2. These times are averaged over the morning peak hence the average time on an un-delayed morning should equal the expected travel time, therefore there would be no increase and the average should be 0. Graph 50 shows the average increase to travel times over the morning period on the 26th of October.



Graph 50 - 26th October: Average increase to travel times over morning period

For a comparison to see the effect congestion may have on a delay Graph 51 shows the average increase to journey times of the 29th the un-delayed day, and the 26th.



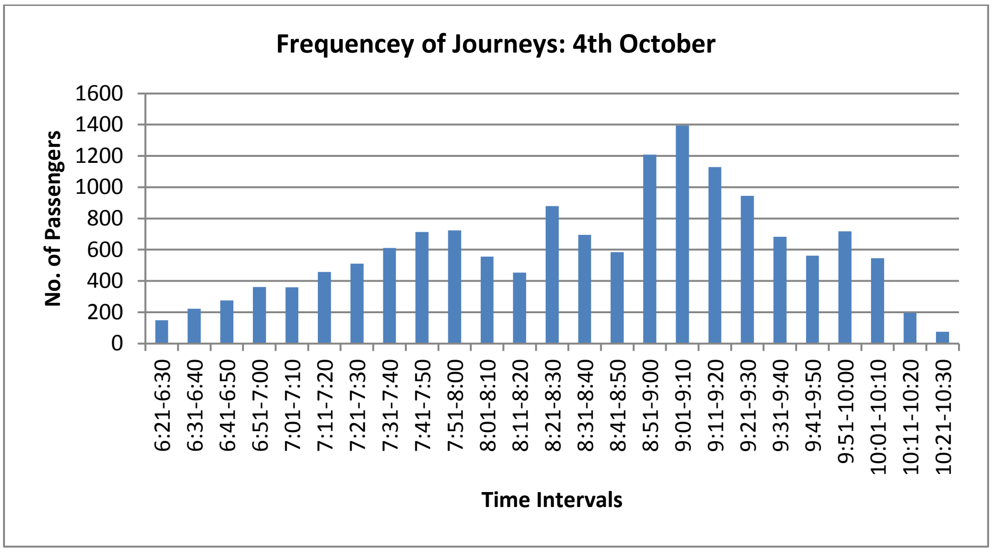
Graph 51 - Average increase to travel times on the 26th October and the 29th October

Graph 51 shows the clear spike in travel times when the faulty train effects passengers' times around 07:30. Beyond that at around 08:25 it can be seen that travel times, on average, drop back to a 2 minute increase of travel times. From 08:25 onwards it would appear that the shape of the travel times on the 26th closely track the shape of the travel times on the 29th. It can be seen that the peaks of one graph match the peaks of the other. Although on the 26th it can be seen the peaks are slightly higher which could be explained by the delay to the service.

In conclusion for this day it can be seen that there is sufficient information for passengers to know how long their journey will be delayed should they choose to take the Victoria Line. Once there is a line delay it is still clear where congestion is affecting the passengers' travel times. This information may make passengers wait a few minutes or decide to take a different route.

4.6.2. 4th October 2012

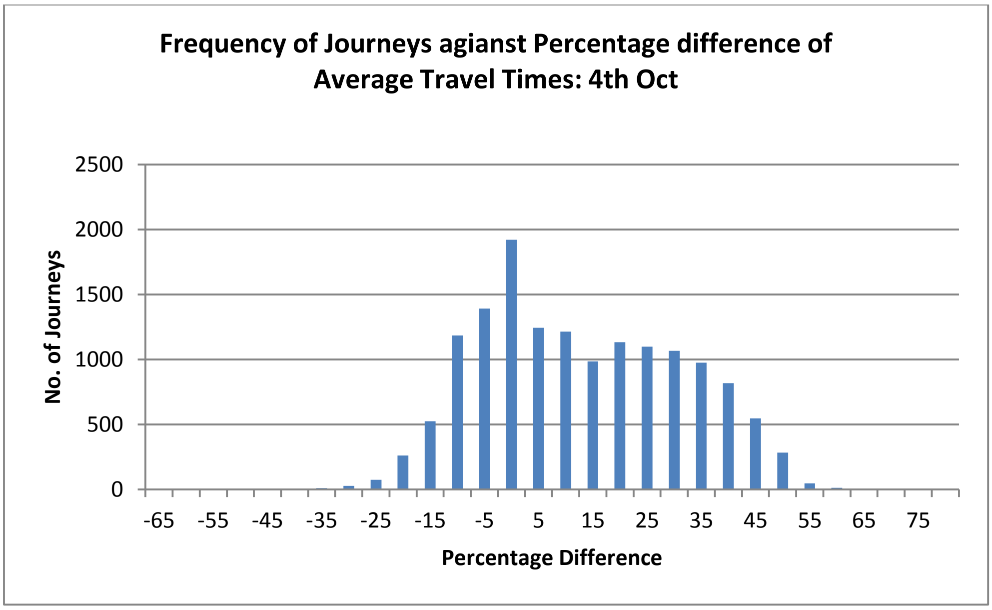
The next day to be analysed is the 4th of October. This day had greater delays than the 26th of October. The official report from TfL states (Transport for London, 2012): on the 4th October there was a signal failure at Vauxhall at 08:20 which led to minor delays across the entire line until 14:00. To begin the analysis for this day the number of passengers at different time intervals will be studied.



Graph 52- Frequencey of journeys: 4th October

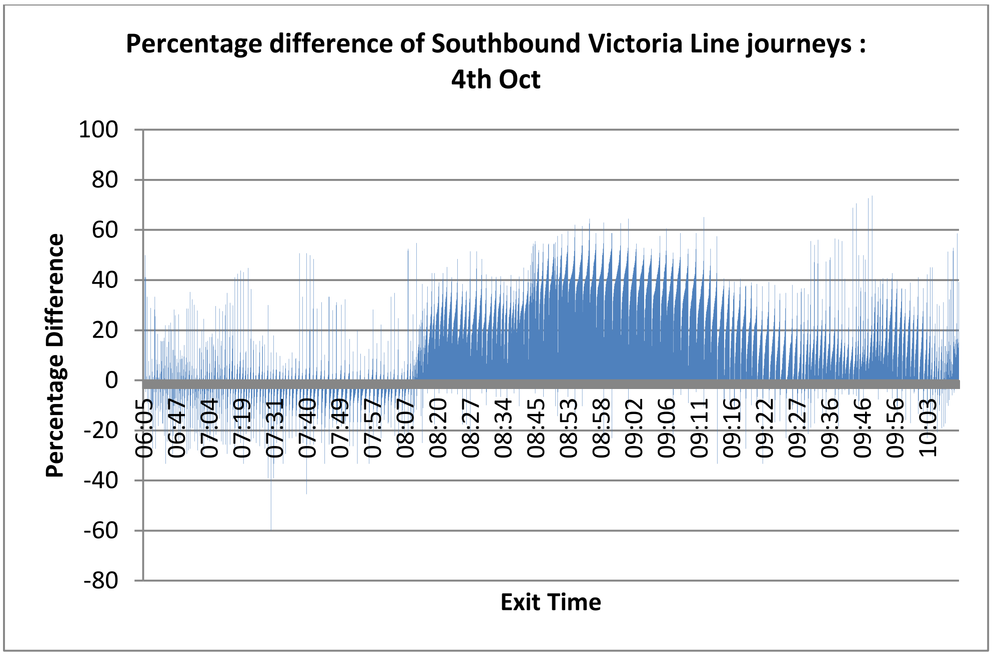
This day exerts an unusual pattern in comparison to the 29th of October, the day with no reported delays. It would be expected that there should be a gradual increase in demand until the peak between 08:40-09:00, and a reduction with a small peak again around 09:30. However, Graph 52 shows that this day does not follow this trend. First, the peak is later than it should be by 10 minutes with the highest demand between 09:01-09:10. This would imply that passengers did not know of the delay and continued with their usual routines. As a result, those passengers who would be expected to be in work at 09:00 would in fact have been late. Further, it can be seen there are two time intervals, 08:00-08:20 and 08:30-08:50, where there are dips in the demand of passengers. This reduction in passenger demand could be an insight into the later problem of problems with the signals but more analysis needs to be completed in order to fully determine this.

Graph 53 shows the distribution of journey times in comparison to the average travel times found.



Graph 53 - Frequency of Journeys agianst Percentage difference of Average Travel Times: 4th Oct

It can be seen in Graph 53 that the distribution is slightly different than expected. It was expected that there should be a distribution centred close to 0. Although the highest point is 0 the rest of the data seems to fall mainly between a 10% increase and 35% with a skew value of 1.92, over the value of classification of a prominent skew. This shows how the delay caused on the day has affected a large number of those travelling.



Graph 54 - Percentage difference of Southbound Victoria Line journeys : 4th Oct

Graph 54 shows the passengers travel times in comparison to the average travel times, expressed as a percentage. It can be seen that the delay starts around 08:20 and affects the passengers throughout the morning period. This graph shows that there are delays to passengers until around 09:30, but the service does not resume normality after this and passengers still experience delays until the end of the data set at 10:05. At 08:40 there is an increase in the delay; this is when on an un-delayed day congestion starts to form. This shows us how congestion worsens the delay due to high passenger demand.

Again an algorithm was used to determine when the delays were taking place on this morning. The same procedure was used as the other days; that an entrance or exit delay was classified when either four or more exits or entrances are delayed more than 6 minutes over the respective average travel time. Further a line delay is classified when 5 or more journeys delayed. The results of the line delays are shown below in Table 30 and all exit and entrance delays appear in the appendices.

Table 31 – Results: Lines delays, 4th October

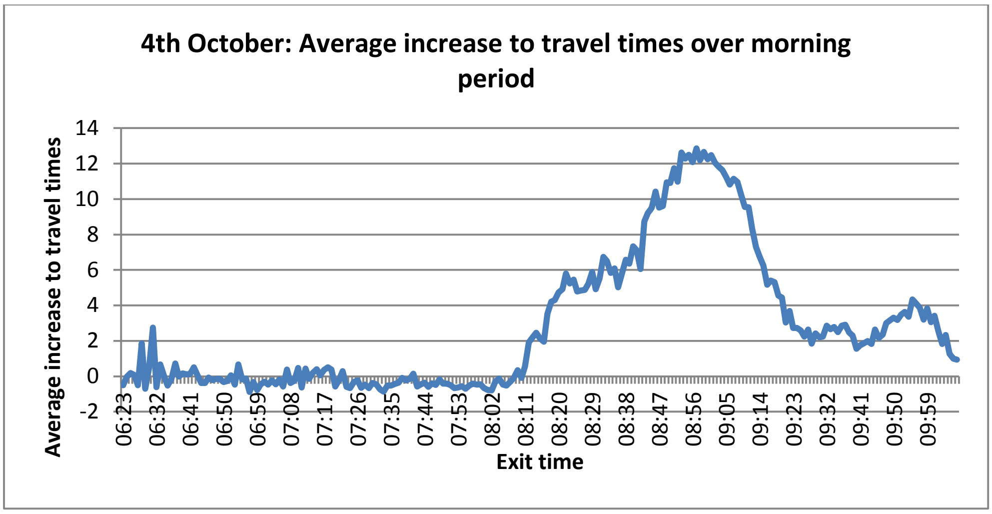
Time	No. of Minutes the Line is Delayed
08:19	1
08:20	1
08:21	1
08:22	1
08:23	1
08:24	1
08:25	1
08:26	1
08:27	1
08:28	2
08:29	2
08:30	1
08:31	2
08:32	3
08:33	3
08:34	2
08:35	2
08:36	3
08:37	2
08:38	2
08:39	3
08:40	3
08:41	3
08:42	4
08:43	4
08:44	4
08:45	5
08:46	6
08:47	5
08:48	5
08:49	6
08:50	5
08:51	6
08:52	6
08:53	7
08:54	7
08:55	7
08:56	6
08:57	8
08:58	7
08:59	7
09:00	7
09:01	8

09:02	7
09:03	7
09:04	6
09:05	6
09:06	6
09:07	6
09:08	6
09:09	7
09:10	5
09:11	5
09:12	5
09:13	4
09:14	4
09:15	4
09:16	4
09:17	4
09:18	4
09:19	3
09:20	3
09:21	3
09:22	2
09:23	3

The original report from TfL (Transport for London, 2012) reported on the 4th October that there was a signal failure at Vauxhall at 08:20 which led to minor delays across the entire line until 14:00.

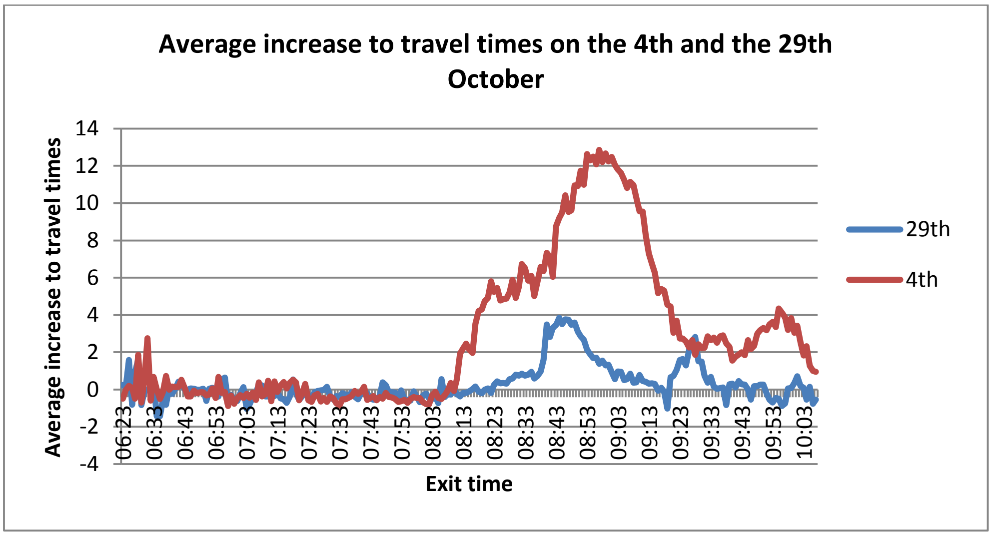
Table 30 shows that there were minor delays from 08:19 – 09:23. There is data beyond 09:23 but results from the algorithm in this thesis indicate that a delay is not classified since not enough passengers are delayed. It can be seen that there are no extra delays occurring to the passengers exiting at Vauxhall. The exit does appear as to be delayed but not noticeably more so than other exit stations.

In order to determine what effects congestion may have on the travel times of passengers during a delay. The average increase to passengers' travel times have been found and plotted against the exit times shown in Graph 55.



Graph 55 - 4th October: Average increase to travel times over morning period

To be able to make a comparison, the travel times of the 29th, the un-delayed day, have be plotted with the travel times on of the 4th, shown in Graph 56.



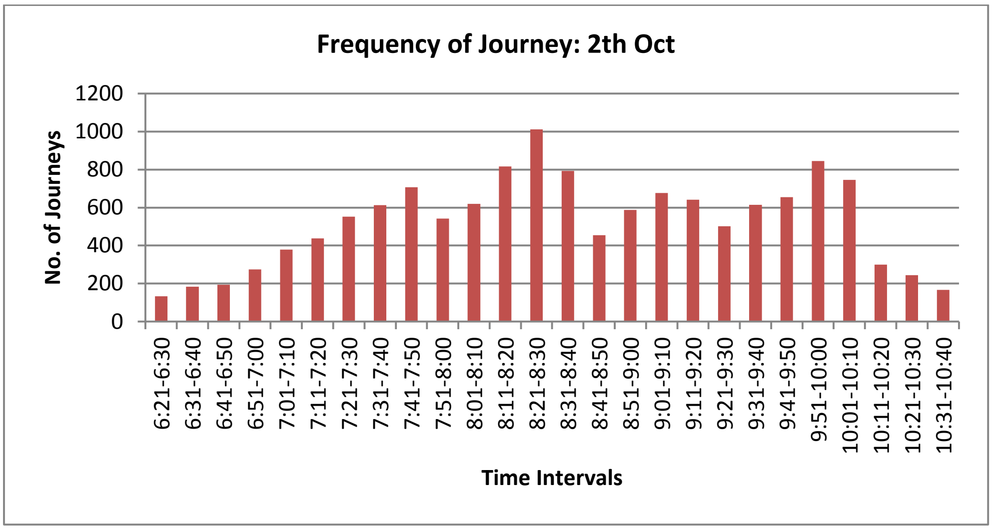
Graph 56 - Average increase to travel times on the 4th and the 29th October

In Graph 56, again it can be seen the shape of the two graphs seem very similar. There is a time delay on the 4th with the highest peak in travel times at 09:00 whereas on the 29th this peak appears at 08:47. This difference seems to increase over time with the second peak in travel times on the 29th being seen at 09:28

which has moved to 09:55 on the 4th. From these similarities it can be concluded that the severity of a delays can be closely linked to passenger demand.

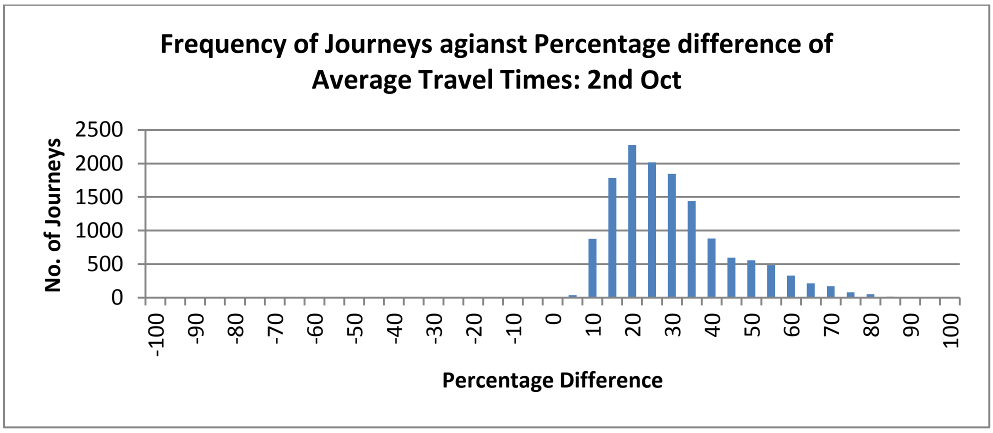
4.6.3. 2nd October 2012

Finally the last delayed day to be analysed is the 2nd October. The report from TfL (Transport for London, 2012) says on the 2nd October a person went under a train at 08:40; there was a partial line suspension between Victoria and Brixton between 09:00 - 10:30; this led to severe delays until 11:15 along the whole Victoria Line. Unfortunately the data to be analysed is the AM peak which ends around 10:30 therefore it is not possible to know when the delay ends from the data. Graph 57 shows the number of passengers at different time intervals.



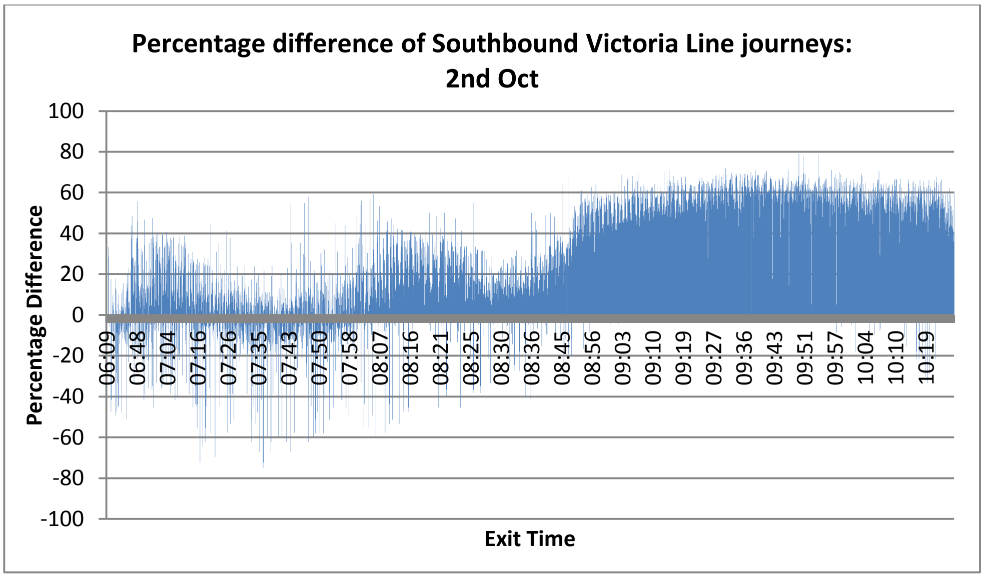
Graph 57 - Frequency of journey: 2th Oct

Graph 57 shows that there is no clear pattern to the frequency of passengers. There is a peak at 08:30 which is earlier than expected (the expected peak in frequency is between 08:40-09:00) for the morning rush hour peak. The highest peak on the other days is usually between 08:50 and 09:00, except when there is a delay and some passengers may be seen to be exiting at 09:10. There is no evidence of this peak on 2 October.



Graph 58 - Frequency of Journeys agianst Percentage difference of Average Travel Times: 2nd Oct

The distribution shown in Graph 58 has dramatically shifted to the right in comparison to the un-delayed day (29th October) with a skew value of 3.1, which is over the value of a prominent skew. On an un-delayed day, it would be expected that the distribution would be centred nearer 0; instead most passengers are above 0 with very few passengers completing their journey in less time than expected. In comparison to the other days analysed this graph shows that the disruption to this day has affected the passenger travel times far more.



Graph 59- Percentage difference of Southbound Victoria Line journeys: 2nd Oct

The TfL report (Transport for London, 2012) states that a passenger went under a train at 08:40, yet it is clear to see there is a disruption to passengers around 07:00 and again at 08:17. The delays incurred by passengers due to the passenger under a train are clearly visible with some passengers reaching a delay of an extra 60% extra time spent in the system on top of their expected travel time. Due to the incident happening just when congestion is at its peak it is unclear to what extent passengers are delayed by congestion and to what extent they are delayed by the incident.

Finally all line delays, exit and entry delays were discovered, the results of the line delays are shown in Table 32 and the entry and exit delays are found in the appendices.

Table 32 - Results: Line Delays, 2nd October

Time	No. of Minutes the Line is Delayed
08:10:00	2
08:16:00	2
08:17:00	2
08:20:00	2
08:21:00	3
08:22:00	2
08:23:00	2
08:24:00	2
08:48:00	3
08:49:00	4
08:50:00	5
08:51:00	4
08:52:00	5
08:53:00	6
08:54:00	7
08:55:00	7
08:56:00	6
08:57:00	7
08:58:00	6
08:59:00	7
09:00:00	9
09:01:00	9
09:02:00	8
09:03:00	9
09:04:00	8
09:05:00	11

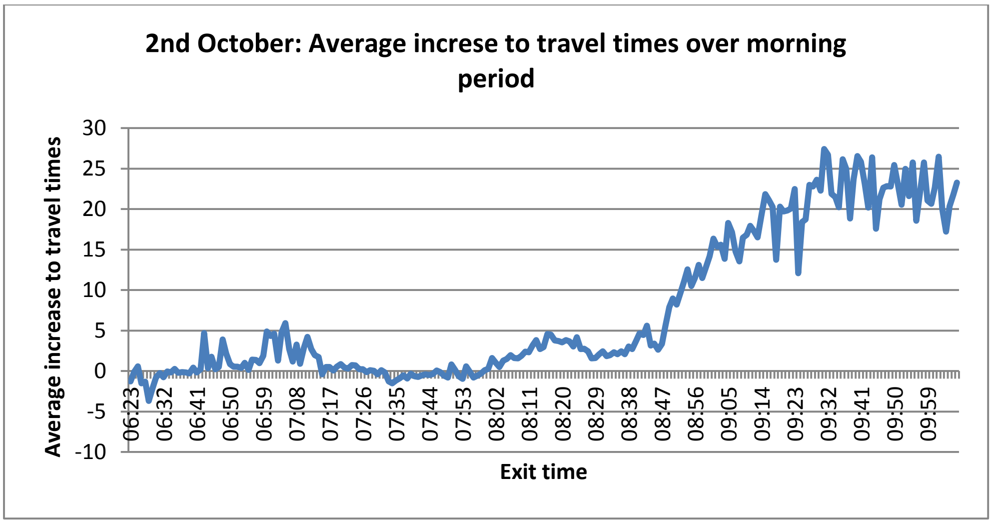
09:06:00	11
09:08:00	7
09:09:00	10
09:10:00	12
09:11:00	11
09:12:00	10
09:13:00	10
09:14:00	14
09:15:00	15
09:16:00	14
09:17:00	14
09:18:00	12
09:19:00	15
09:20:00	14
09:21:00	16
09:22:00	14
09:23:00	18
09:24:00	11
09:25:00	14
09:26:00	17
09:27:00	17
09:28:00	17
09:29:00	19
09:30:00	19
09:31:00	20
09:32:00	20
09:33:00	17
09:34:00	19
09:35:00	14
09:36:00	19
09:37:00	16
09:38:00	15
09:39:00	15
09:40:00	18
09:41:00	18
09:42:00	19
09:43:00	14
09:44:00	24
09:45:00	16
09:46:00	15
09:47:00	17
09:48:00	17
09:49:00	18
09:50:00	18
09:51:00	16
09:52:00	14

09:53:00	19
09:54:00	16
09:55:00	19
09:56:00	12
09:57:00	20
09:58:00	19
09:59:00	13
10:00:00	18
10:01:00	19
10:02:00	20
10:03:00	14
10:04:00	12
10:05:00	20
10:06:00	16
10:07:00	17
10:08:00	16
10:09:00	23
10:10:00	16
10:11:00	13
10:12:00	12
10:14:00	16
10:15:00	17
10:16:00	13
10:17:00	16
10:18:00	15
10:19:00	14
10:20:00	11
10:21:00	15
10:22:00	17
10:23:00	15
10:24:00	20
10:25:00	22
10:26:00	19
10:27:00	14
10:28:00	13
10:29:00	14
10:32:00	11
10:33:00	15
10:34:00	13
10:35:00	13
10:36:00	11
10:37:00	15
10:38:00	13
10:39:00	13

Table 31 shows that the delay to the line lasts longer than the other days. It should also be noted that in response to the questionnaire, according to passengers, a severe line delay is likely to be over 41 minutes. As it can be seen in Table 32 the largest line delay to passengers is 23 minutes at 10:09. The number of minutes the line appears to be delayed seems to be variable. Further there are a few gaps in the report – for example a delay is not registered at 10:30 or 10:31. This is not due to passengers not experiencing delays, but to the lack of data. There may be multiple reasons for this; as (1) Brixton is not included in the dataset (2) the data set ends close to 10:30 therefore not all journeys around this time may be included and finally (3) it could be due to low passenger demand.

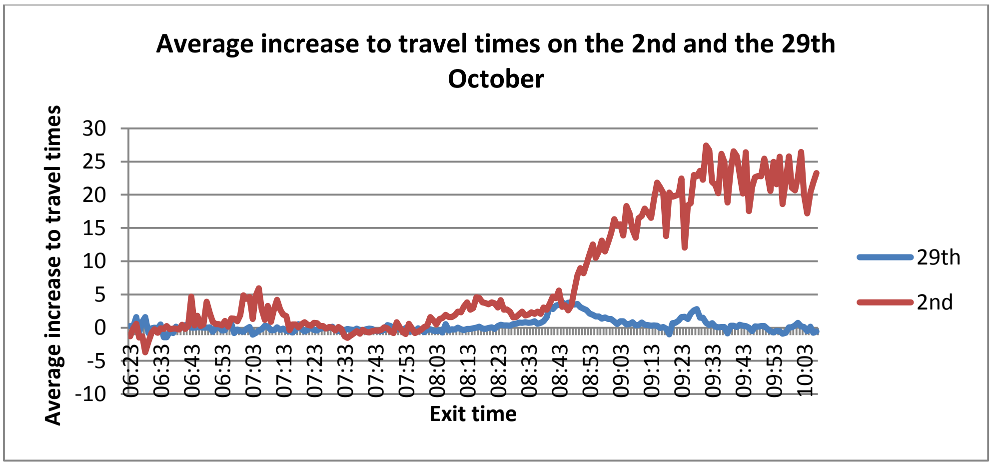
It can be seen also that there are a few line delays reported earlier than the incident. These delays are infrequent and only last a few minutes at a time, but do not appear on uncongested days. Should passengers perhaps see infrequent delay reporting, this may entice them to behave differently. Finally it should be noted that the effect of the incident is not visible through Oyster data until 08:48, 8 minutes after TfL reported it. This is perhaps due to the passengers being stuck in the system and therefore their exit times not being recorded. Oyster data could be used to determine a delay occurring in this case by noting how many passengers are entering and the lack of passengers exiting. However, it should be assumed that other sources of knowledge about the service could be used to gain more information about the current state of the network. A full list of all delays reported in the AM peak can be found in the appendices.

To understand the effects of congestion on the delays found to this day the average increase to travel times over the morning period has been found, shown in Graph 60.



Graph 60 - 2nd October: Average increse to travel times over morning period

Graph 61 shows the average increase to passengers' travel times on the 2nd and the 29th October, where the 29th is the un-delayed day. Here unlike the other delayed days there seems to be no similarities between the two dates. Throughout the analysis of the 2nd October it has become clear the severity a passenger under a train has to passengers travel times.



Graph 61 - Average increase to travel times on the 2nd and the 29th October

Each day analysed above shows a large amount of information available about how passengers are affected during service disruptions. Analysis has been completed showing how the number of passengers are affected on different days when there are delays. The distributions of passengers' travel time were shown for the different days and it was discussed how they skew further to the right depending on how severe the delay is. The algorithm created to find congestion progressed in order to determine if there are line delays. This showed sufficient information for passengers and returned an average delay to passengers using the line. Further, once there are line delays it is still clear where congestion is affecting the passengers' travel times. This information may make passengers wait a few minutes or decide to take a different route. Finally it was seen in some cases that the relay of information was slower in returning results than the reports from TfL. It should be noted that there are multiple sources of information about the current service in the Underground and all should be used to gain a fuller picture to give passengers information. Further, it is thought that a better picture of what state the current service is in will develop when all Oyster data is analysed over all lines. This would give extra information about stations that are served by more than one line and this could lead to gaining a clearer picture of what is happening on each line.

4.7. Conclusion

In this section the aim was to determine if it was possible to gain more relevant information about the service of the London Underground for passengers.

This section started by the average travel times on the Victoria Line being found. The raw Oyster data was aggregated to firstly remove a day that was clearly a delayed day. From this it was sorted by origin and destination pairs on the same line. This was done to remove any ambiguity over what route a passenger may have taken. From this anomalous journey times were removed. Using this 'clean' data a database was created that showed the times it would take a passenger, on average, to complete any journey, that had its origin and destination on the Victoria line.

These results were then compared to the London Journey Planner using regression analysis. In the northbound direction it was found that the regression line was $y = 0.94x + 4.93$, while in the southbound direction the regression line was $y = 0.97x + 4.49$.

In order to successfully classify a delay it was essential to gain the passengers perception of delays in the Underground; this lead to a questionnaire being composed and 400 passengers of the London Underground completing it. Although the sample did not entirely match the excepted demographics, a random mix of difference sex, age, journey purpose and residence was achieved. It was concluded that the difference to the national average may be due to the location of the questionnaires being taken and the time of day they were taken. The results of the questionnaire showed that the sample wanted to see how many minutes their delay may be rather than the traditional minor/major/severe statuses. It was found that the passengers would like information about congestion in the stations and they felt this information may change their behaviour. Finally passengers felt on average a minor delay lasted 6 minutes, a major delay lasted 18 minutes and a severe delay lasted 41 minutes.

The value of 6 minutes for a minor delay was then used as a threshold for classifying a delay. In order to ensure this was not too sensitive this was tested against the Oyster data that spanned 8 weeks. In each case less than 2% of passengers' journey times were over this threshold. Next this threshold was tested against the data that would be considered as real-time data. This indicated that this threshold may be a useful value for determining congestion. However, this test highlighted that there were a number of anomalies in the data, so the data needed to be smoothed.

Focus then turned to data that was being produced on a daily basis. A moving average that took the last few data point was decided to be used to smooth the data. A time dependant average was considered, yet it was decided if there was a slow stream of data in the off peak times no data would be registered potentially. In comparison a moving average over all data points keeps a continuous stream of data in the off peak times. Different values of data points to be included in the average were considered but it was decided that four pointes would be used.

This threshold of 6 minutes was then used to discover congestion in the network. To determine if delays due to congestion can be spotted in the data, a day, at random, was chosen to be studied that contained no reported delays. Journeys that have entrances in common and exits in common were studied together to determine if there are certain stations that cause delays at certain times. When looking at the entrance and exit delays a number of stations were identified as congested. For example: exiting Oxford Street between 08:40 and 08:43, exiting Victoria between 08:46 and 08:52, entering Walthamstow Central between 08:09 and 08:17 and entering at Finsbury Park between 08:18 and 08:21. In addition to these congested stations, on this day it appeared that a number of passengers could be seen to be delayed between 08:40 and 08:57 and again at 09:27 to 09:29. These results would indicate that congestion can be seen in the network and busy stations can be identified. The future for this would be to discover if trends over time occur which could indicate the amount of congestion at particular stations at certain times. This ideally would be research that could follow on from this project and be used to inform passengers of regularly congested stations.

A number of days with different delays were then analysed and it was shown that the Oyster data can spot the operational delays and how much they are delaying passengers' travel times. These days consisted of different severities of delay which took place at different times. From these days it became clear that it was possible to see the delays through the passengers travel times increasing. However, it is not a reliable source of information for showing when the delay starts. Yet it is possible to get numeric values to how much passengers are affected by the delay. Finally, the data shows how much passenger congestion contributes to the delay. This is highlighted in passengers being delayed beyond the operational delay. This result can also be seen when the increase to passengers travel time is starting to reduce then sharply increases as a result of rush hour. Further, in some case it can also be seen that unreported delays are appearing in the data.

In conclusion this section aimed to find if it is possible to obtain information which would help to provide better information for passengers in the London Underground. Using Oyster data it has been shown there is rich information about how long it takes passengers to complete journeys, show congestion and determine how much delays will affect the passengers.

5. Hong Kong

In Section 4, it was seen through analysing Oyster card data in the London Underground that information about the current service in the metro system was available. The algorithm that was created in Section 4 was developed and tailored to the Oyster card data in order to obtain the maximum amount of information about the dynamics of the system.

To understand whether the algorithm created to answer the research questions in London is usable in other countries a second metro system is introduced to the methodology to use as a comparison.

In this Section, data from the Hong Kong MTR metro network is analysed. The smart card used in this system is named the Octopus card.

5.1. Data collection

Data from the automated ticketing system in the Hong Kong's MTR system was received from the operators mid October 2013. The data, produced by the Octopus smart card, contained all journeys completed in the system within the month of September 2012. The data was separated into days with each file containing all journeys completed in the network on that day.

The files contained all information stored from a stamp of the Octopus Card produced when the card is touched on one of the card readers in the stations. The Octopus data is slightly different from the Oyster data in London and includes some data that is not recorded in London, for example, some information obtained is the ticket barrier number and the price of the journey as well as other information such as the ticket gate number that was not needed for this analysis.

Firstly, the unwanted information was removed from the data set. This left the data shown in Table 33. Table 33 contains the individual Octopus card number, the date and time the card was used, whether it was stored as entering (ENT) or exiting the system (USE), the station code of the entry station and the station code of the exit station. If the stamp was for an entrance to the system the entrance and exit station code would be stored as the same value, as seen in Table 33.

Table 33 – Raw Octopus data, 8 day sample

CSC_PHY_ID	BUSINESS_DT	TXN_DT	TXN_TYPE_CO	TRAIN_ENTRY_STN	TXN_LOC
900125532	02/09/2012	02/09/2012 16:58	ENT	29	29
900125532	02/09/2012	02/09/2012 17:17	USE	29	49
900125532	02/09/2012	02/09/2012 22:59	ENT	49	49
900125532	02/09/2012	02/09/2012 23:20	USE	49	29
900125559	02/09/2012	02/09/2012 17:09	ENT	18	18
900125559	02/09/2012	02/09/2012 17:37	USE	18	13

It was essential that the data for each journey was on the same row, so that journey travel times could be determined. In order to do this a program was written, in Matlab R2012b, which first took the time stamps out of the date column and gave them their own column. From this the Octopus card number could be ordered in time such that origin and destination pairs would be together. Although above it shows the times are in order within the original file this was not the case.

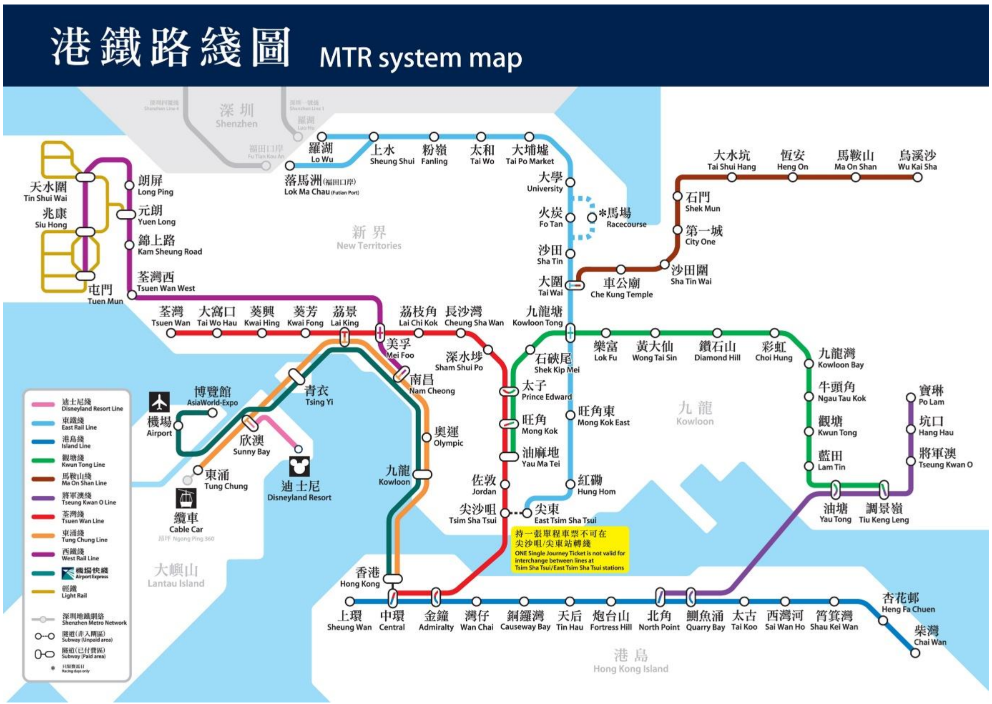
The program then paired Octopus cards with the same number together with the clause that the first value should be an ENT the second should be a USE and that the travel times would be inferred and should not be greater than 120 minutes. It is unlikely that there would be a journey that would last 120 minutes. This clause was introduced to ensure that a stamp that was 'touched in' and stored with Octopus that had no 'touch out' was not paired with a 'touch out' later in the day that had no 'touch in'. This left the data looking as below in Table 33.

Table 34 – Aggregate and sorted Octopus data

Octopus Number	Entry Time	Exit Time	Travel Time	Entry Station	Exit Station
900125559	14:45	14:57	12	3	5
900125613	13:16	13:44	28	45	6
900125682	07:11	07:44	33	2	13
900125682	18:54	19:29	35	13	15
900125860	08:51	09:02	11	3	75
900125860	18:34	18:48	14	65	23

The data was aggregated to find all journeys that start and end on the Island line (ISL), shown in blue at the bottom of Figure 9.

Figure 9 ("Hong Kong MTR Map," 2014)



The Island Line was chosen due to the similarities between this line and the Victoria Line in London. The Island Line has no splits or loops in the track and runs centrally in the city. It also contains 14 stations whereas the Victoria Line in London contains 16 stations. These similarities meant that during the discussion section, Section 6, comparisons could be made.

Further, the data set given by the operators of the MTR was all entry and exit pairs for the whole network, in September, 2012. It was essential that a line was chosen that contained delays within this time period. A spreadsheet was received from the operators of the MTR showing the delays that took place in the network in September 2012, shown in Table 35.

Table 35 – Delay report, MTR, September 2012

No.	Date	Time	Location	Nature of Problems	Delay Minutes	Total No. of Trains Affected	Line
1	01-Sep-12	17:06	TSW DN	APG	5	1	TWL
2	01-Sep-12	21:17	TIK UP	Passenger	7	1	MOS
3	01-Sep-12	23:57	TAK DN	Rolling Stock	5	1	ISL
4	05-Sep-12	20:29	SHW DN	Equipment Failure	13	1	ISL
5	08-Sep-12	21:20	SHW UP	Human Factor (Staff)	9	2	ISL
6	14-Sep-12	15:07	SKM DN	Equipment Failure	17	18	KTL
7	16-Sep-12	06:21	NTK UP	Human Factor (Staff)	6	1	KTL
8	18-Sep-12	08:35	LAT UP	Rolling Stock	5	2	KTL
9	19-Sep-12	08:26	CSW DN	Passenger	7	14	TWL
10	19-Sep-12	18:09	SHW DN	Equipment Failure	6	9	ISL
11	24-Sep-12	08:28	WTS DN	Rolling Stock	7	4	KTL
12	24-Sep-12	08:38	YAT DN	Passenger	7	6	MOS
13	24-Sep-12	21:25	YAT Both	PSD	13	8	MOS
14	25-Sep-12	06:10	WTS DN	Rolling Stock	11	25	KTL
15	25-Sep-12	21:21	TIK DN	Human Factor (Staff)	17	1	MOS
16	27-Sep-12	08:48	TIH DN	Passenger	5	1	ISL
17	20-Sep-12	20:16	TSY UP	Rolling Stock	28	1	TCL
18	07-Sep-12	06:40	MEF UP WRL	Equipment Failure	15	13	TWL
19	12-Sep-12	07:52	SHS-LOW UP	Passenger	14	3	ERL
20	14-Sep-12	18:53	TAW DN (ERL)	Equipment Failure	40	19	ERL
21	29-Sep-12	13:11	SHT DN	Rolling Stock	35	7	ERL
22	30-Sep-12	21:16	TWO DN	Passenger	15	4	ERL

Table 36 shows how many delays in the month were reported for each line and the total number of minutes of delays experienced to the line over the month.

Table 36 – Summary of reported delaps: total reported delayed minute

Line code	No. of days delayed in a month	Total minutes of delays reported in a month
TWL	3	27
ERL	4	104
MOS	4	44
KTL	5	46
ISL	5	38
WRL	0	
AEL	0	
TCL	1	28
TKL	0	

In Table 36 it can be seen that the East Rail Line (ERL) line has experienced substantially more minutes delayed than the other lines, however, this line was not chosen as it is slightly further out of the city as was the Ma On Shan Line (MOS) and the Tung Chung Line (TCL). It was then left a choice between TWL, KTL, and ISL. The Island Line (ISL) was chosen as out of these three lines it had the most similarities with the Victoria Line in London.

5.2. Average travel times

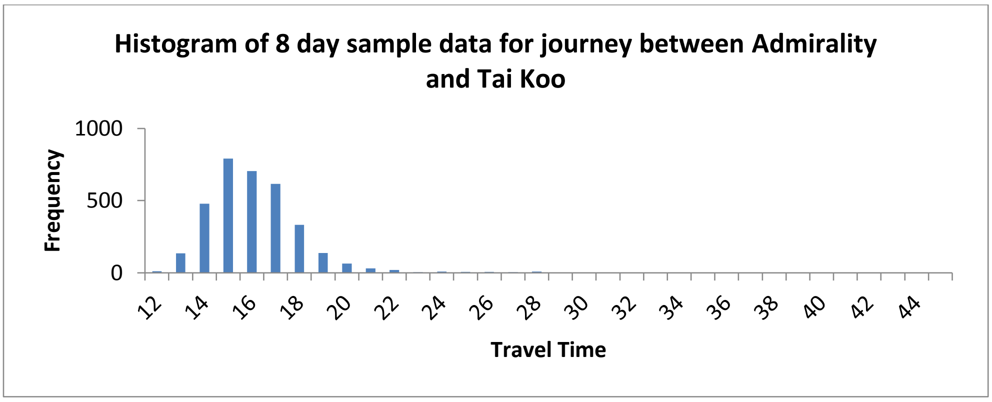
When working with the data from the London Underground, as described in Section 4.1, anomalies were removed from the data before finding how long on average it takes passengers to complete a journey.

To determine what was considered as an anomaly in the MTR data, it was necessary to determine how removing different quantities of standard deviations would affect the mean. An example of this is shown below for the journey between Wan Chai to Tai Koo the average time for this journey was found to be 16.21 minutes. Over an 8 day sample of the data 3400 passengers took this journey. All results in Table 37 the values are rounded to 2 decimal places.

Table 37 – Wan Chai to Tai Koo: Removal of standard deviation and revised means

κ	μ+ κσ	Number of entries removed	Percentage removed from total data set.	M
No. of Standard Deviations	No. of Standard Deviations + mean			New mean
1	16.67	662	19.47	15.52
2	17.04	662	19.47	15.52
3	17.41	662	19.47	15.52
4	17.77	329	9.68	15.87
5	18.14	329	9.68	15.87

It was decided that because all travel times are rounded to the nearest minute and regardless of how many standard deviations are removed the mean stays the same it was decided to not remove any of the data. The code to aggregate the data started with removing all journeys longer than 2 hours, so there were few anomalies in the data. The range for the above journey was 12 minutes to 44 minutes. The histogram, in Graph 62, shows the data.



Graph 62 - Histogram of 8 day sample data for journey between Admirality and Tai Koo

All the mean travel times were then found for all Westbound and Eastbound Journeys on the Island Line. These results are shown in the appendices (Table 78 and Table 79).

Previously when analysing Oyster data from London, the average travel times were found from a data set that only contained the AM peak journeys. This was due to the availability of the data. With the MTR Octopus data a month of journeys was given. This meant the average travel times could be taken from all journeys at all times of day. To show the difference that is made to the average by aggregating the data to different times of day Table 38 was created. It shows 3 different journeys that are travelling eastbound on the Island Line.

Table 38 – Comparison of daily average travel time and timely average travel times

Start Station	End Station	AM peak Average	Off peak Average	PM peak Average	Day Average
Admiralty	Chai Wan	25.92	26.46	26.86	26.61
Tin Hau	Quarry Bay	12.45	12.54	12.81	11.96
North Point	Shau Kei Wan	11.99	12.92	12.40	12.60

The values in Table 38 have been rounded to 2 decimal places to highlight the difference in the times. For the different journeys at the different times of day the average can change by a minute, due to the rounding. This of course would affect the results of the delay analysis by, in some cases, 1 minute, meaning some delays are not counted. However there appears to be no consistency in the times of day that the average appears to have risen.

The day average would take into account the time of day that had the greatest frequency of passengers. This would allow for the travel time to be most accurate when compared with the travel times taken when frequency is high and this would make unexpected congestion more visible in the data. Further, since the frequency of trains throughout the day can allow for variation (as can walking speed), the decision to take an average allows the times found to include these variations. The variations that can occur by these factors allow for greater differences in overall times than seen by time of day variations. For these reasons it was decided that the day average would be used, this means that regardless of time of day one dataset of average journey times can be used as a comparison.

5.3. Regression Analysis

As with the case study of London, Section 4.3, the travel times and the relationship between the times given by the MTR journey planner are to be analysed by regression analysis. This analysis gives information of how relevant the MTR journey planner data is compared to real journeys taken. Further analysis of the residuals provides evidence of any anomalies in the average travel times. Unlike London there was no missing data therefore a heuristic procedure was not needed. For the regression the dependent variable was the average times and the independent variable was the journey planner times. The results for the eastbound Island line results are shown below.

Table 39 - Regression analysis, eastbound Octopus Data: Regresssion statistics

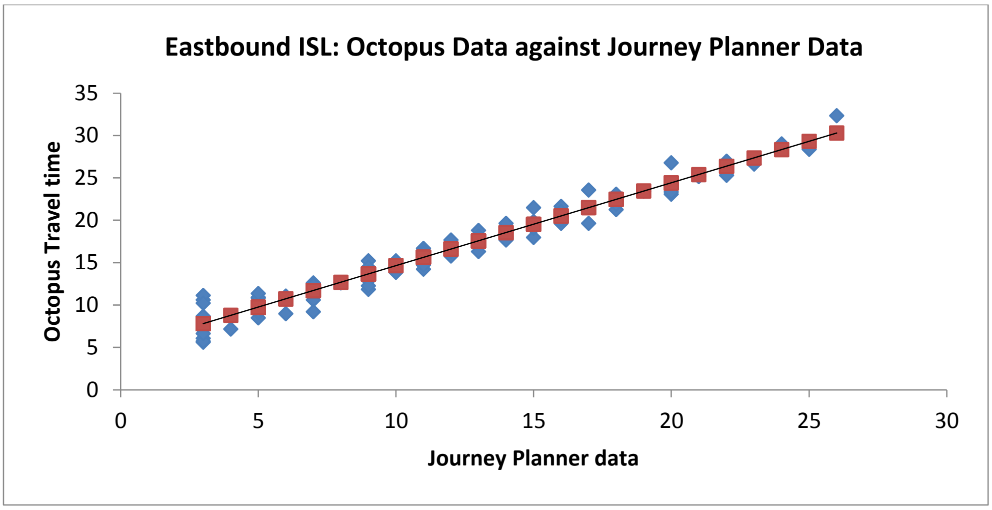
Regression Statistics	
Multiple R	0.98
R Square	0.96
Adjusted R Square	0.96
Standard Error	1.18
Observations	91.00

Table 40- Regression analysis, eastbound Octopus Data: Anova

ANOVA					
	df	**SS**	**MS**	**F**	**Significance F**
Regression	1.00	3205.75	3205.75	2306.23	0.00
Residual	89.00	123.71	1.39		
Total	90.00	3329.46			

Table 41 - Regression analysis, eastbound Octopus Data: Correlation results

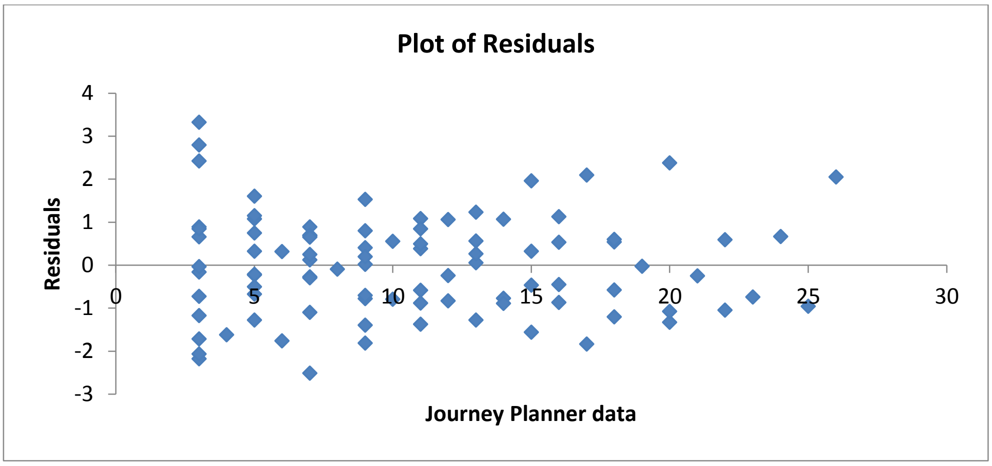
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**Lower 95%**	**Upper 95%**	**Lower 95.0%**	**Upper 95.0%**
Intercept	4.87	0.25	19.41	0.00	4.37	5.37	4.37	5.37
Variable	0.98	0.02	48.02	0.00	0.94	1.02	0.94	1.02



Graph 63 - Eastbound Octopus data against Journey Planner data: regression

It can be seen that the equation for the regression line is $\boldsymbol{y = 4.87 + 0.98x}$ this means that 98% of the Octopus data times are explained by the Journey Planner times. The adjusted R Squared also shows 0.96 which is very close to 1 showing strong correlation. It can further be seen that the p-value is very small this means that it is very unlikely that these results occurred at random.

Finally the results of the residuals can be seen in Graph 64 the values are evenly distributed either side of the 0 line.



Graph 64 – Eastbound regression analysis: Plot of residuals

Graph 64 shows that the greatest difference between journey planner times and average Octopus times occur when the journey times are small. This is due to the journey planner data not including the amount of time it takes for passengers to walk from the trains to the ticket barriers and vice versa. This time represents a larger proportion of a shorter journey and therefore accounts for more variation. The westbound Island line is found below.

Table 42- Regression analysis, westbound Octopus Data: Regresssion statistics

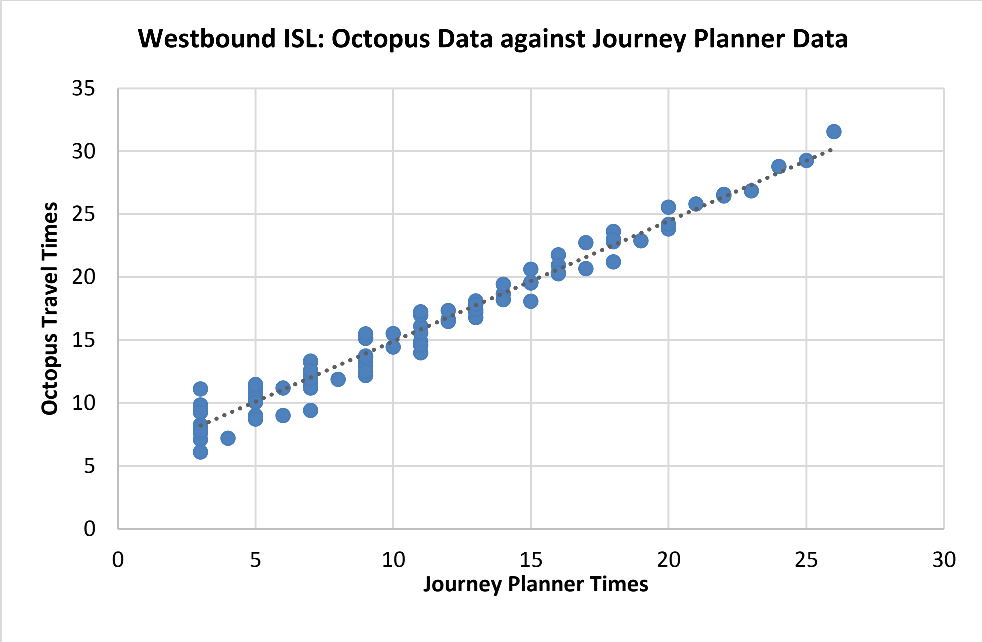
Regression Statistics	
Multiple R	0.98
R Square	0.97
Adjusted R Square	0.97
Standard Error	1.03
Observations	91.00

Table 43- Regression analysis, westbound Octopus Data: Anova

ANOVA					
	df	**SS**	**MS**	**F**	**Significance F**
Regression	1.00	3071.04	3071.04	2882.20	0.00
Residual	89.00	94.83	1.07		
Total	90.00	3165.87			

Table 44- Regression analysis, westbound Octopus Data: Corrolation results

	Coefficients	***Standard Error***	***t Stat***	***P-value***	***Lower 95%***	***Upper 95%***	***Lower 95.0%***	***Upper 95.0%***
Intercept	5.33	0.22	24.25	0.00	4.89	5.76	4.89	5.76
Variable	0.96	0.02	53.69	0.00	0.92	0.99	0.92	0.99

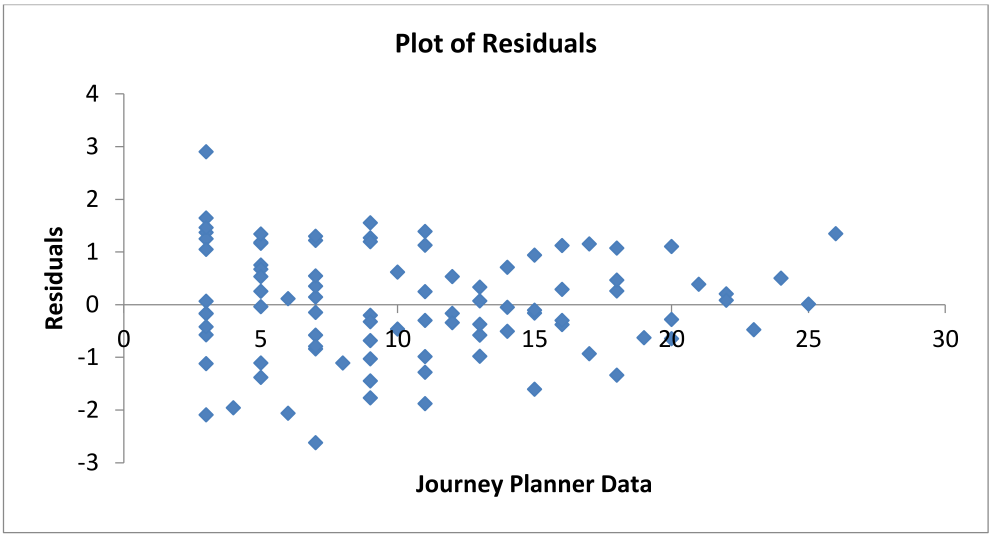


Graph 65 - Westbound Octopus data against Journey Planner data: regression

The equation for the regression line is $y = 5.33 + 0.96x$ this means that 96% of the Octopus data times are explained by the Journey Planner times. The adjusted R Squared also shows 0.97 which is very close to 1 showing strong significance.

There is a small difference between the eastbound times and the westbound times with the eastbound times showing a slightly stronger relationship with the journey planner data.

The residuals have been plotted in Graph 66 it can be seen the values are evenly distributed either side of the 0 line, this shows the journey planner data is closely related to the Octopus data. However, it can be seen that the shorter the journey, the more variation there is.



Graph 66 – Regression analysis: plot of residuals

Finally, regression analysis was completed to determine the relationship between the eastbound and westbound journeys times found. The results are seen below.

Table 45 - Regression analysis, westbound and eastbound Octopus Data: Regresssion statistics

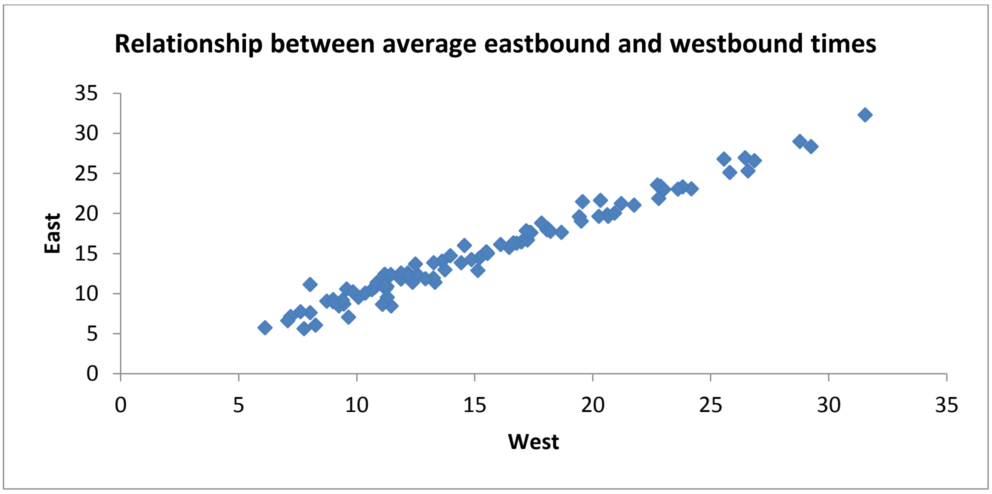
Regression Statistics	
Multiple R	0.99
R Square	0.97
Adjusted R Square	0.97
Standard Error	0.99
Observations	91.00

Table 46 - Regression analysis, westbound and eastbound Octopus Data: Anova

ANOVA					
	df	SS	MS	F	Significance F
Regression	1.00	3243.04	3243.04	3339.99	0.00
Residual	89.00	86.42	0.97		
Total	90.00	3329.46			

Table 47 - Regression analysis, westbound and eastbound Octopus Data: Corrolation results

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-0.42	0.29	-1.45	0.15	-1.00	0.16	-1.00	0.16
Variable	1.01	0.02	57.79	0.00	0.98	1.05	0.98	1.05



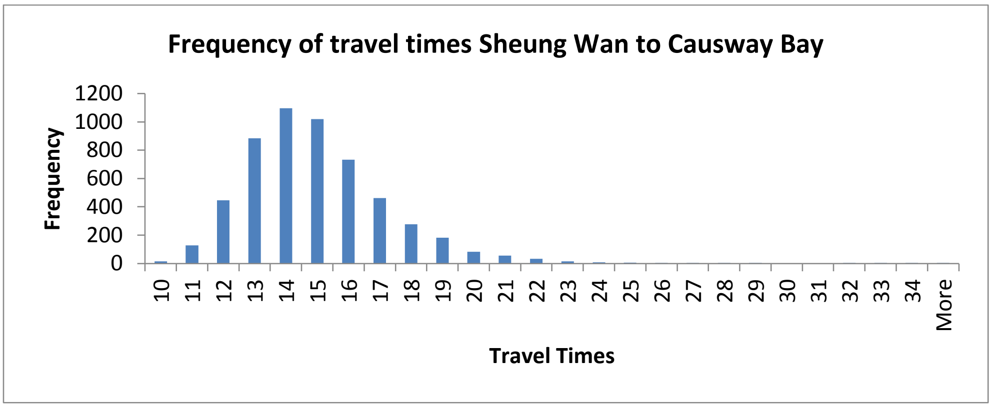
Graph 67 – Regression analysis: elationship between average eastbound and westbound times

Graph 67 shows that with the gradient being 1.01 there is almost no difference between the directions.

5.4. What is a delay?

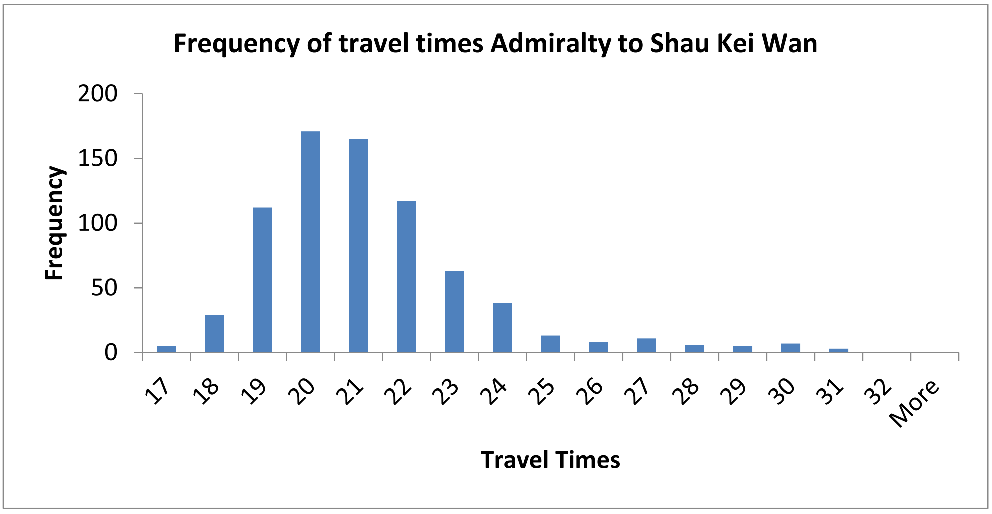
The MTR operators make a pledge to all customers that 99.5% of the passengers journeys will be completed within 5 minutes of the timetabled journey ("MTR: Our pledge for service 2013," 2013). If there is more than 5 minutes delay to the schedule passengers within the station are advised of the delay over the PA systems. If a delay should exceed 20 minutes reports are given on the journey planner, internet and service boards. It can be seen in Table 35 that all the delays to the Island line considered are less than 20 minutes but over 5 minutes.

Now it is known that a delay is classified, operationally, as over 5 minutes. It is now essential to determine whether 5 minutes is a reasonable threshold to use to define a passengers' delay. Three origin and destination pairs have been chosen for analysis of their journey time distributions. The three journeys chosen were Sheung Wan to Causeway Bay, Admirality to Shau Kei Wan and Causway Bay to Tai Koo. These three were chosen to gain a mixture of short and long journeys and those with difference passenger frequencies. The data was taken from 8 days randomly picked from the data set containing all journeys completed in the MTR network in September 2012.



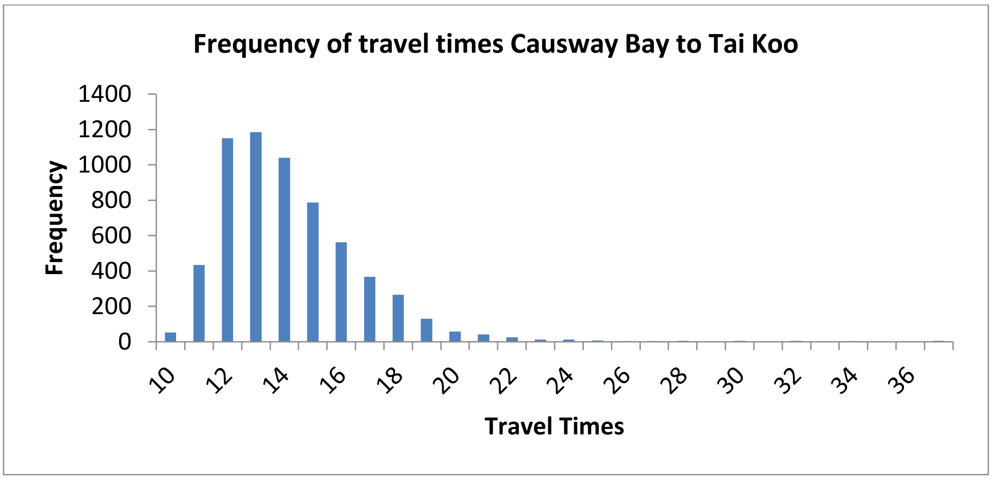
Graph 68 - Frequency of travel times Sheung Wan to Causway Bay

Graph 68 shows the travel times recorded for journeys completed for the origin destination pair Sheung Wan to Causeway Bay against the number of people who completed the journey. This data is taken from 8 days spanning the month of September 2012. For the journey of Sheung Wan to Causway Bay the predicted journey time from the above data was 15 minutes. If there were a minimum delay threshold of 5 minutes over the mean time this would mean passengers are delayed if their journey will take them more than 20 minutes. Within the data set only 150 passengers completed their journey in over 20 minutes. Out of 5476 passengers this is 3%.



Graph 69 - Frequency of travel times Admiralty to Shau Kei Wan

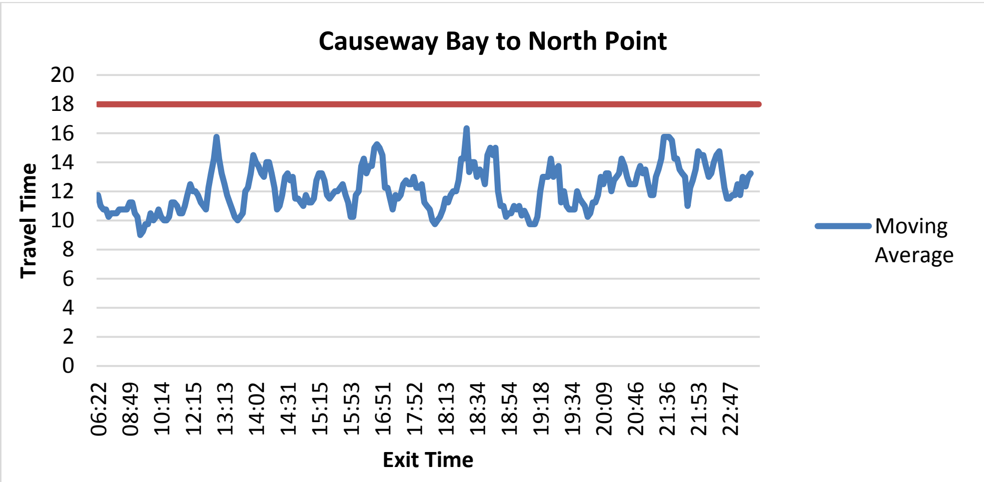
Graph 69 shows the frequency of passengers completing journeys between Admiralty and Shau Kei Wan and their different travel times. The average travel time for this journey was 21 minutes, therefore with a 5 minute delay the threshold for a delay is 26 minutes. Over the 8 days the data was taken from 40 passengers' travel times where over 26 minutes, this was out of 753 passengers which accounts for 5% of the sample.



Graph 70 - F requency of travel times Causway Bay to Tai Koo

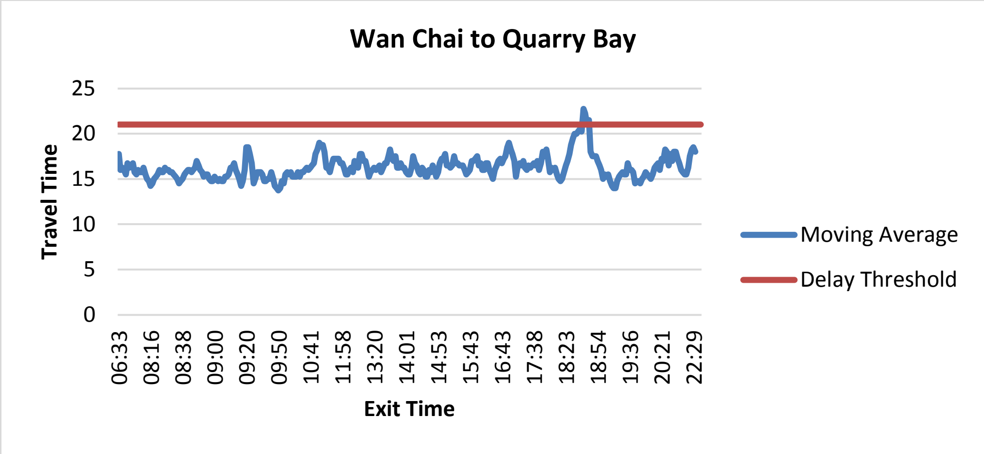
Graph 70 shows the distribution of travel times for the journey of Causway Bay to Tai Koo over 8 days. The average travel time was found to be 14 minutes, therefore with a delay threshold of 5 minutes a delay is classified if a journey should take over 19 minutes. For this sample 193 passengers took over 19 minutes out of 6166 which accounts for 3%.

To further determine that 5 minutes is a reasonable delay threshold, it is important that there are not too many delays appearing in the data on an un-delayed day, this would make the information unreliable as there would be multiple false positive delay statuses. In order to analyse this, a number of different journeys have been chosen for further investigation: Causeway Bay to North Point, Wan Chai to Quarry Bay and Central to Sai Wan Ho. These are shown in Graph 71, Graph 72 and Graph 73 respectively, the moving averages of the data are used for this analysis. Creating the moving average is discussed in Section 5.5.



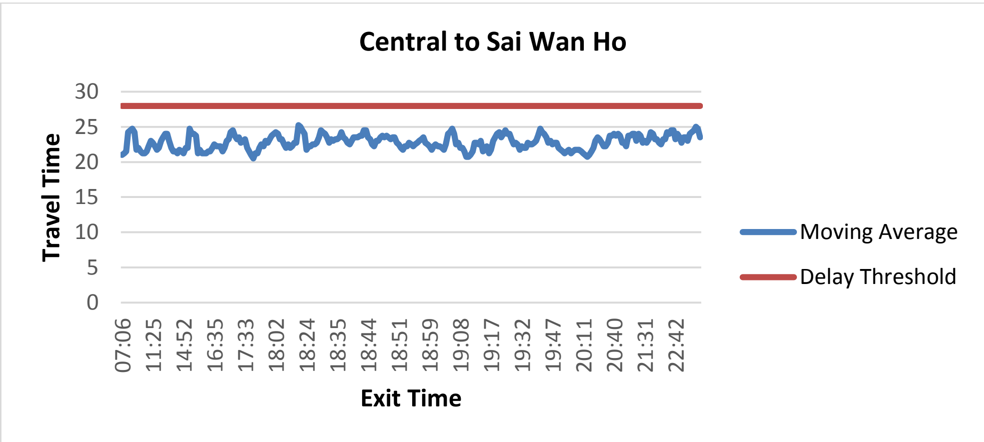
Graph 71 - Causeway Bay to North Point: Moving average

Graph 71 shows all journeys completed on the 18th September between Causway Bay and North Point. The average time that this journey should take was found to be 13 minutes. Therefore the delay threshold would be 18 minutes. It can be seen on this day that although there is a large variation in travel times, none of the values seem to cross the delay threshold of 18 minutes.



Graph 72 - Wan Chai to Quarry Bay: Moving average

Graph 72 shows all journeys on the 18th September between Wan Chai and Quarry Bay. The average time for this journey was found to be 16 minutes, which means the delay threshold is 21 minutes. It can be seen that in one instance the delay threshold is breached during the PM peak period. There are 4 moving average points that breach the threshold out of 326 this accounts for 1% of the data. This is a good indication that when there is a large amount of congestion this threshold will spot it.



Graph 73 - Central to Sai Wan Ho: Moving average

Finally, Graph 73 shows all journeys between Central and Sai Wan Ho. The average time for this journey was found to be 23 minutes, making the delay threshold 28 minutes. It can be seen that no moving average points are above the delay threshold.

Overall in this analysis there are very few instances where there are moving average points are greater than the delay threshold. It is important that the delay threshold is not too close to the average travel time value since it is important that delays are not reported when there are none; since this would be a false positive. However, it also important the value is not too great as it is less likely to spot smaller delays giving false negatives.

5.5. Congestion Reporting

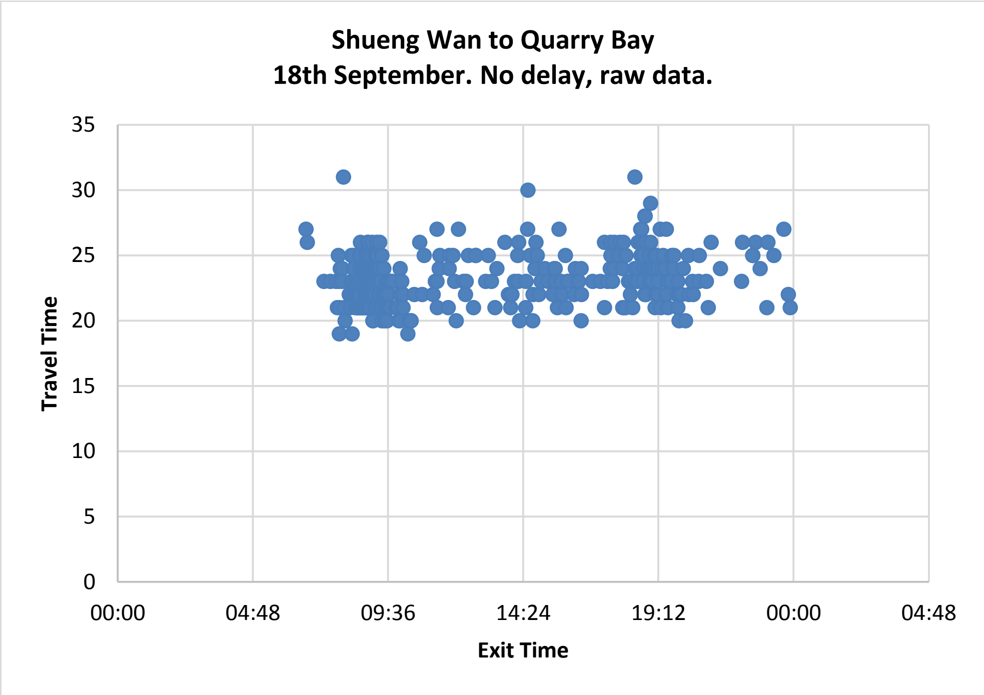
To understand what delays are taking place to the service on the different days, it is important that the data set given by the MTR operators represents real time data. The data from the original files has already been aggregated to find the average travel times described in Section 5.2.

From this the data for a day was sorted by origin-destination pairs then further by exit time to simulate the data being returned from the ticket barriers as someone exits the system, shown in Table 48. In real-time the data would be scrabbled but matching origin-destination pairs by card number can be completed in a negligible amount of time and therefore these journey times can be given the time of the exit station.

Table 48 – Sample of Octopus data simulating real-time

Ticket Number	Entrance Time	Exit Time	Travel Time	Entrance Station Code	Exit Station Code
915717947	07:22	07:34	12	26	27
915652456	07:26	07:38	12	26	27
912846106	07:24	07:38	14	26	27
900744183	07:31	07:42	11	26	27
901354681	07:30	07:42	12	26	27

Graph 74 shows the raw data of one of the journeys completed on the 18th September, the day with no delay to be studied.

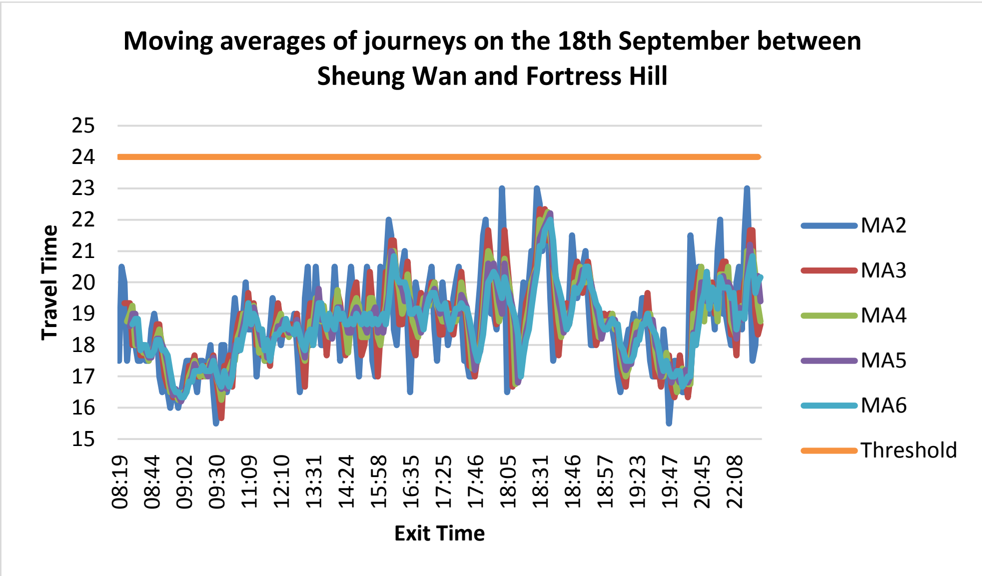


Graph 74 - Shueng Wan to Quarry Bay, 18th September: Raw data

Graph 74 shows clear clustering during the AM and PM peak periods, the average for this journey was found to be 24 minutes, which is visually clear within the data. A number of anomalies can be seen to be within the data. The amount of noise that is visible in the data will make the process of determining a delay harder. The time a passenger enters the station will affect their overall travel time; as the train they catch and the length of time it takes to wait for a train is dependent on what time they entered the system. It is therefore important that times are averaged so that travel times are independent of the time a passenger entered the station. As with the data produced from the Oyster card in London (Section 4.5), the decision was made to smooth the data by taking a moving average.

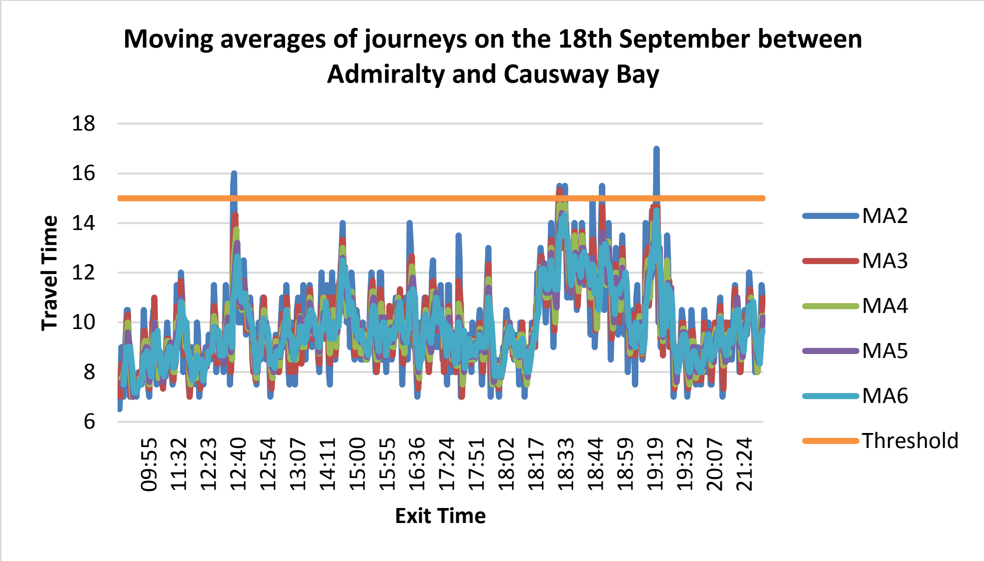
The same criteria as those applied in London were used to determine how many data points should contribute to the average. The number of data points should not be too great a number as this will delay the response of the data, leading to gaps in the real-time data. However, neither should the number be too small, as this could limit the amount the data is smoothed. Graph 75, Graph 76 and Graph 77 respectively show the different possible moving averages for three journeys on the 18th September.

The legends stand for the number of data points used to make 1 moving average point. For example MA2 means two passenger journeys were used to make the average. The journeys considered were Sheung Wan to Fortress Hill, Admiralty to Causway Bay and Tin Hau to Tai Koo. Along with the different moving average possibilities also plotted is the threshold for a delay. This threshold is the average journey time plus 5 minutes. This threshold was decided in Section 5.4.



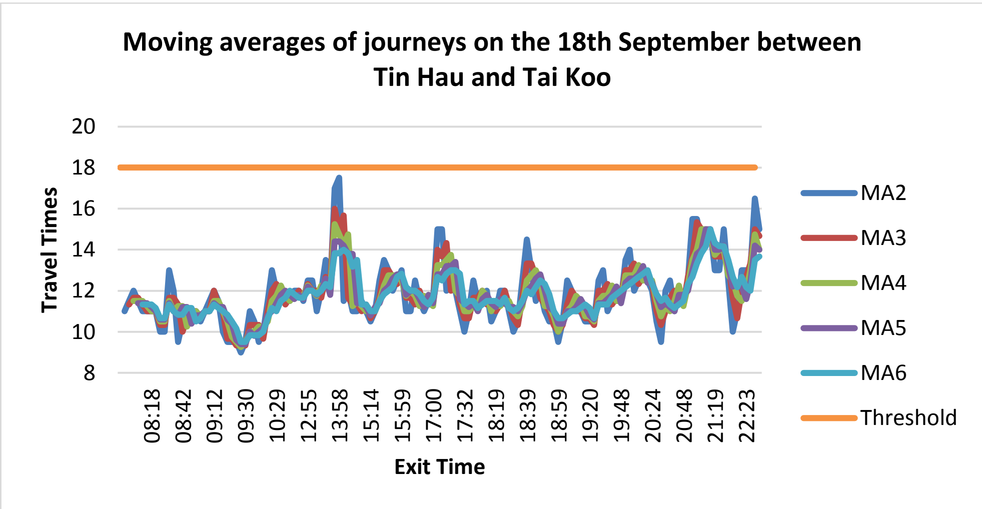
Graph 75 - Sheung Wan to Fortress Hill: Different moving average possibilities

It can be seen in Graph 75 that at no point is the threshold breached. No information is given from this journey which might help in determining how to smooth the data. However it is quite clear that apart from delaying the return of information, a larger number of data points contributing to the moving average is making little difference to the graph. Indeed the data is being smoothed, however all peaks and troughs seem to remain, if only just more spread out.



Graph 76 – Admirality to Causway Bay: Different moving averages

Graph 76 shows that with MA2 and MA3 there are breaches of the threshold. The data for this journey shows that there is a peak in travel times during the evening peak. Although the travel times appear to rise at this time it would appear that in the incidents of the threshold being crossed these are due to anomalies, as although in the peak time the rest of the data is close to the threshold there only seems to be a few occasions that it appears to reach the line.



Graph 77 – Tin Hau to Tai Koo: Different moving average possibilities

Finally once again it can be seen in Graph 77 that the threshold does not appear to be breached. It can however be seen that around 12:30 there appears to be an anomaly in the data. There is a large increase in travel times for MA2 and MA3.

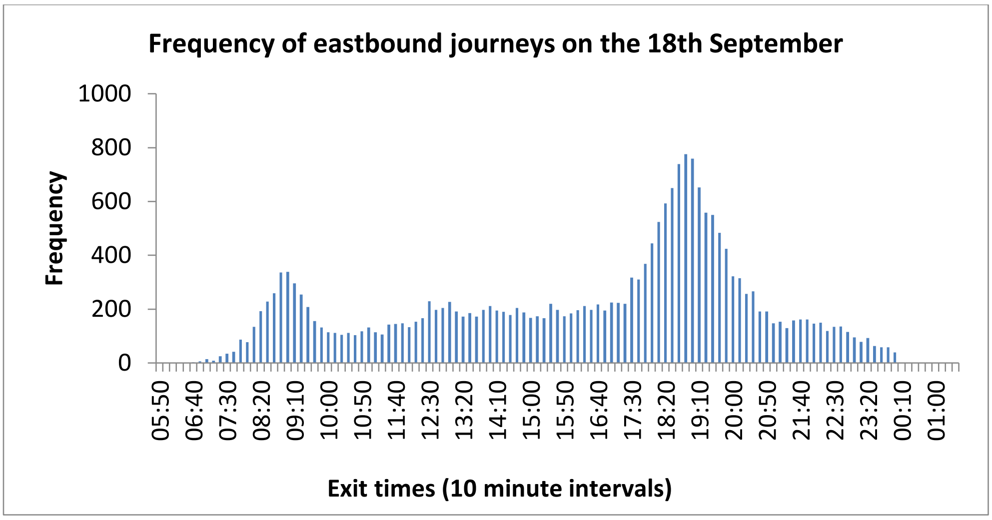
From analysing Graph 75, Graph 76 and Graph 77 it would appear that 2 moving average points and 3 moving average points appear to leave anomalies in the data. It is also important to take a lower number of data points contributing to the moving average so that a large amount of data points still exist, removing 5 and 6 moving average points as possibilities. For this reason it has been decide that 4 points will contribute to the moving average.

5.5.1. Eastbound

5.5.1.1. 18th September

In order to determine if it is possible to see when there is congestion in the network, a day without a delay has been taken to be analysed. This day was the 18th September 2012. The MTR had no reported delays to the Island line on this day (Table 35). Therefore should there be any visible delays incurred on this day it would be reasonable to decide that they are due to passenger congestion rather than operational.

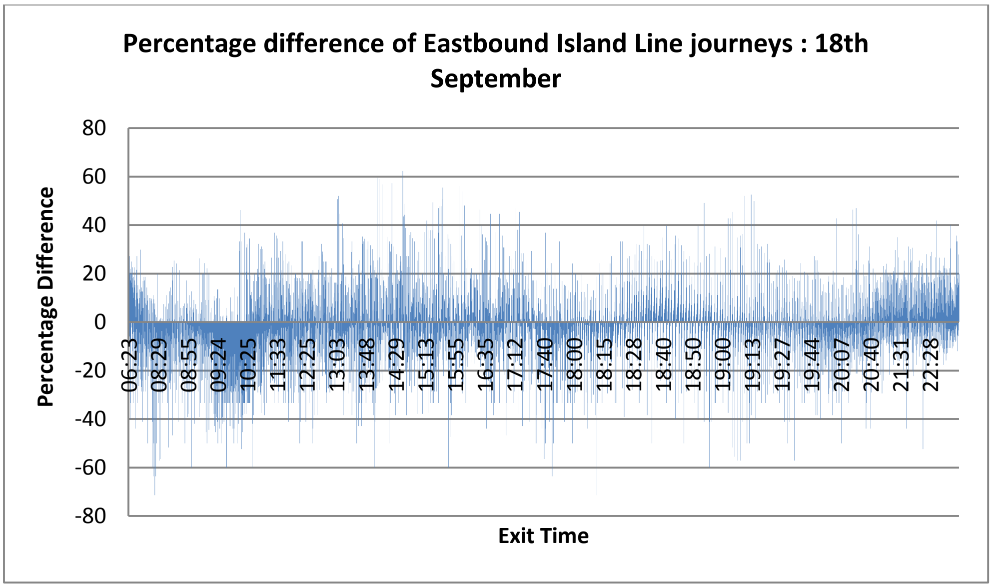
Graph 78 shows the frequency of passengers on the Island line, travelling eastbound, on the 18th September. This is all passengers whose origin and destination were on the Island line.



Graph 78: Frequency of eastbound journeys on the 18th September

Graph 78 shows clear rises in passenger demand during the morning and evening peaks. The AM peak appears to start at around 08:00 and finish around 09:30. The PM peak starts around 18:00 and ends around 20:00. The AM peak appears to reach a maximum of around 300 passengers while the PM peak reaches 800 passengers. This may be due to passengers travelling in one direction in the morning and the other direction in the evening. It may be common to the line that passengers travel east to work but live more in the west.

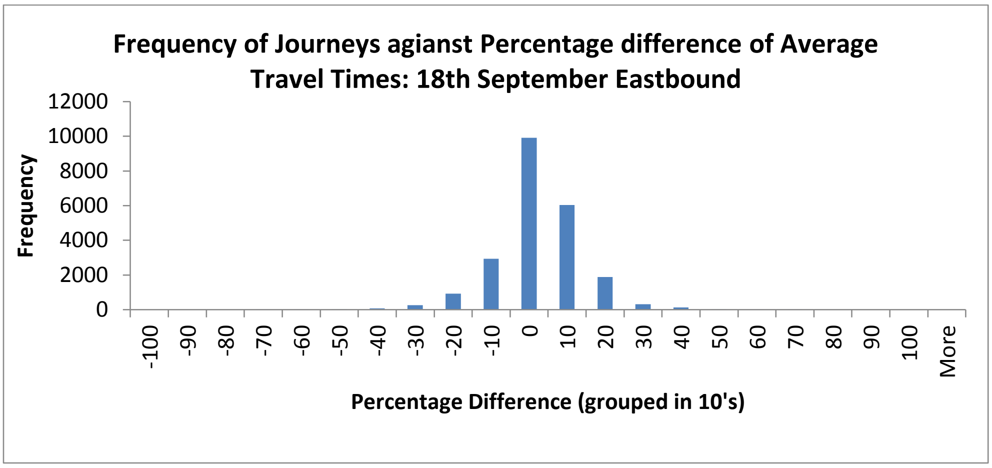
To continue analysis of the 18th September the percentage difference between the average journey time and the moving average journey times recorded on the day has been plotted in Graph 79.



Graph 79 - Percentage difference of Eastbound Island Line journeys : 18th September

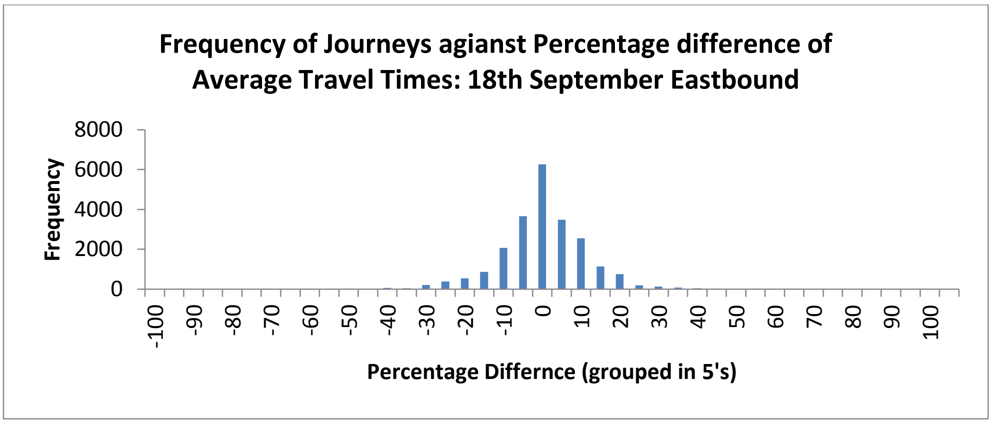
Graph 79 shows there is a large spread in travel times with an approximate range of journey being completed between -20 and +20% of their respective average travel time. There appears to be no clear increase in travel times during the peak travel times. The morning peak appears to have a drop in travel times with more passengers getting though the system closer to the time expected or in less time. The PM peak shows the smallest variation during the day, Graph 78 shows that the highest frequency is during the PM peak; this means that when the average travel times were determined a large number of PM peak travellers would have contributed to the average travel time; this may be the reason why there is less variation at this time. Alternatively, during peak times it is usually found that a lot of passengers move more efficiently through the network, which leads to the entire crowd moving more efficiently. This would mean that only a small number of passengers would take longer than they should do. Further, high passenger demand during these times means it may be impossible for passengers to take their journey any faster due to high quantities of people. During low frequency times and off-peak it is possible for passengers to take their journey much slower or indeed much faster, i.e. they can go at the speed they prefer.

To understand the range of how long the passengers are taking in comparison to the average travel times, the percentage difference of the moving average points and the respective average travel time values have been plotted against frequency to show the distribution, Graph 80.



Graph 80- Frequency of Journeys agianst Percentage difference of Average Travel Times: 18th September Eastbound

Graph 80 shows that when the actual travel times are compared with the average travel times, after smoothing, they produce a skewed distribution centred close to 0. To further understand this distribution the size of the bins has been reduced so that the data fits into intervals of 5 rather than 10 and this is shown in Graph 81.



Graph 81 - Frequency of Journeys agianst Percentage difference of Average Travel Times: 18th September Eastbound – grouped in 5's

Graph 81 shows the distribution of the travel times in comparison to the average travel times.

The skew for Graph 81 is -0.23 (2dp) a negative skew implies that more of the data lies to the left of the mean. Intuitively this means on this day more people were getting through the system in less time than the mean. Since the mean is taken over a number of different days it is expected values can lie either side of it. However this value is less than the threshold discussed in Section 4.5.1 that defines a prominent skew.

To determine if there is congestion causing passengers to be delayed in the network the same algorithm employed with the Oyster data was used to see if there are delays in the Octopus data. This is discussed in detail in Section 4.5 and first discussed in the methodology in Section 3.2.1.5.

As with the London case, it was essential to determine how many moving average points should be delayed within the same minute to classify a delay; this is how many moving average points have greater values that the respective average travel time plus five minutes.

For this analysis three different types of delays were analysed: entrance delays, exit delays and line delays. Since there was no reported delay to the service on the 18th September it is expected that no line delay is found on this date.

Table 48 shows the minutes delays that would be reported with different values of delays found within the same minute, this algorithm was first defined in Section 4.5. It can be seen there are delays when up to 4 delays are found in the same minute.

Table 49 - Line delays: Number of moving average points with delays in common in the same minute. 18th September Eastbound

Line Delays: Number of moving average points with delays in common in the same minute		
2	3	4
13:44	16:07	18:31
13:45	18:42	18:40
14:21	18:55	18:48
14:37	19:04	
14:38	19:08	
15:19		
15:30		
16:07		
16:53		
17:14		
17:17		
18:31		
18:38		
18:40		
18:42		
18:43		
18:45		
18:48		
18:49		
18:50		
18:53		
18:54		
18:55		
18:57		
18:59		
19:03		
19:04		
19:08		
19:15		
19:37		
20:49		
09:27		

For determining delays due to congestion, exit and entrance were considered separately. To determine if congestion can be seen at either entrances or exits, the data was sorted in two ways for this delay analysis, by exit time and by entrance time. The data was sorted by exit time in order to find exit delays and by entrance time to find entrance delays. As pointed out in Section 3.2.1.5, the information regarding entrance time cannot be discovered in ‘real-time’ because it can only be found in hindsight when the passenger exits the network. However, one possible direction of future work could be to determine over time which are crowded stations.

Table 50 – Entrance and exit delays: Number of moving average points with delays within the same minute. 18th September Eastbound

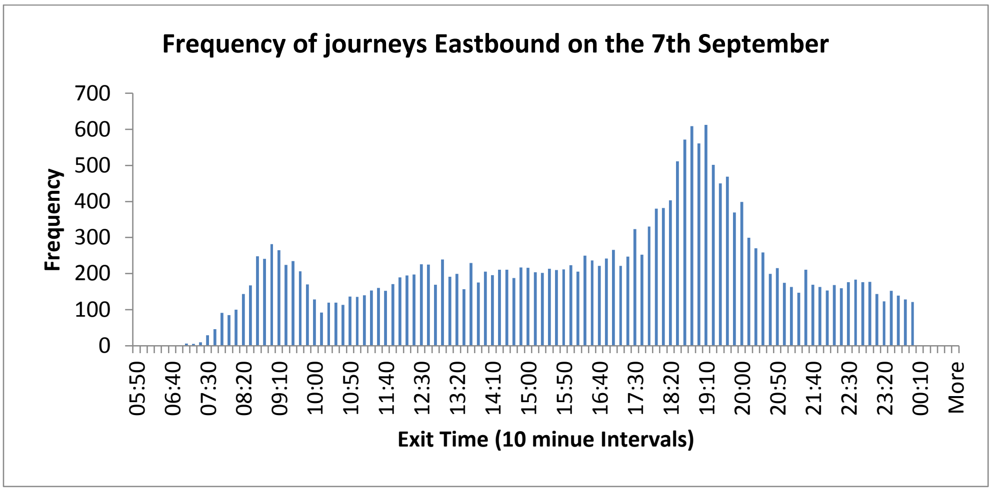
	Number of Delays in common			
	2		3	
	Station Name	Time of Delay	Station Name	Time of Delay
Exit Stations	Quarry Bay	16:07		
	Causeway Bay	17:17		
	North Point	18:31		
	Shau Kei Wan	18:40		
	Tai Koo	18:48		
	Chai Wan	18:49		
	Shau Kei Wan	18:57		
Entrance Stations	Sheung Wan	18:06	Sheung Wan	18:24
	Central	18:11		
	Central	18:13		
	Sheung Wan	18:14		
	Central	18:15		
	Sheung Wan	18:17		
	Central	18:21		
	Sheung Wan	18:24		
	Wan Chai	18:24		

Table 50 shows the results of what congestion delays can be found when different values of moving average pairs are delayed within the same minute. It shows the delays found for the entrance and the exits on the 18th September. It can be seen that when 3 delays are in common over the same minute only 1 entrance delay is found. When there are 2 delays in common sharing the same entrance or exit station it can be seen a number of delays are found. Clearly congestion starts to form between 18:05 and 18:25 at some of the entrances, mainly Sheung Wan and Central.

Further, only 2 moving average points in common showing delays indicates a very low number of delayed passengers. It can be seen that a small amount of congestion takes place entering and exiting stations in the evening peak but this represents only a very small proportion of the passengers travelling. For example 13 people entered at Shueng Wan at 18:14 with 2 reported delays which accounts for 15% of those entering at that time. 30 people are recorded to enter Central Station at 18:15 that are taking journeys that exit on the Island line. Out of these passengers only 2 delays due to congestion have been recorded, accounting for 7% of people that entered at that time. However these delays seem to be short lived and clear quickly suggesting that congestion, on the 'average day', is not a problem to passengers of the MTR network.

5.5.1.2. 7th September

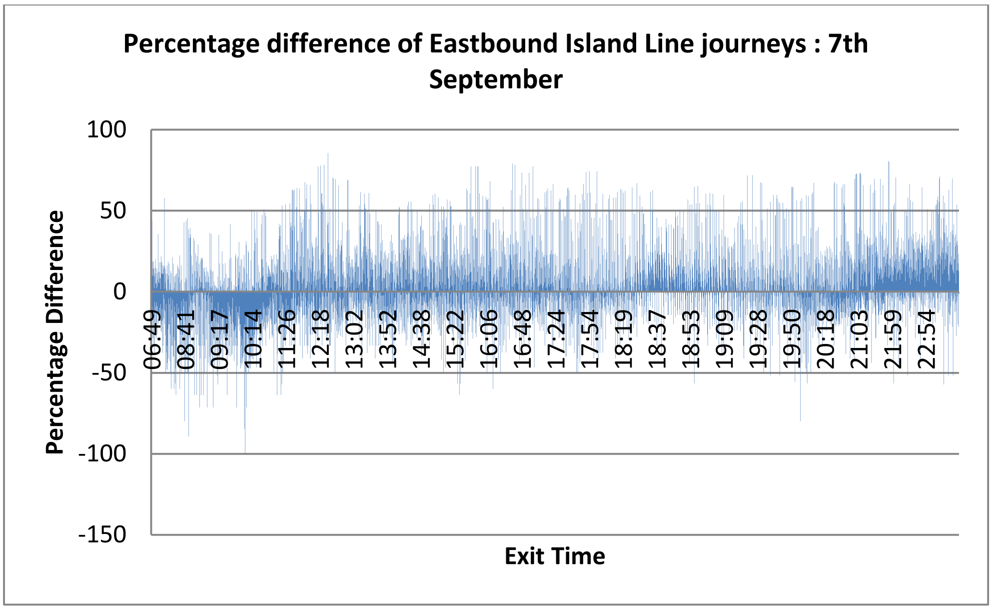
A second day was analysed that had no reported delays, this was the 7th October, in comparison to the 18th September, a Tuesday, the 7th September is a Friday. Again to start the analysis the frequency of journeys over the day has been plotted in Graph 82.



Graph 82 - Frequency of journeys Eastbound on the 7th September

Graph 82 shows a very similar pattern to the frequency of passengers seen in Graph 78, where there were two clear peaks for AM and PM rush hour, where frequency rose to around 300 in the morning, yet there is a slight decline in the evening peak of around 200 passengers, as the peak only reaches around 600.

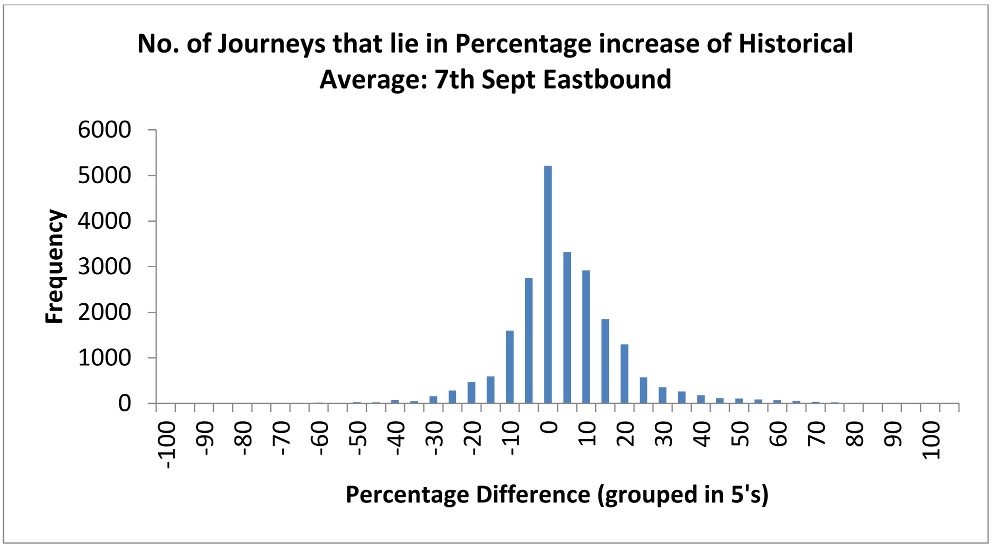
Next the percentage difference between the average travel times the travel times on the morning of the 7th have been plotted in Graph 83.



Graph 83 - Percentage difference of Eastbound Island Line journeys : 7th September

As with Graph 79, Graph 83 shows a large spread in travel times. The AM peak seems to follow a similar pattern to that in Graph 79; however, the PM peak seems to show a slight increase in passengers' travel times. Here it can be seen between 18:20 and 19:20 that very few passengers appear to be travelling through the system in less time than the average, whereas a large number of passengers appear to be taking longer. Since the real-time travel times are being compared to the average travel time, it would be expected that there should be an equal distribution above and below 0, with the greater variation in times above 0 as there is no maximum time, but there is a minimum time a journey can take. However during the PM peak it appears that it is unlikely that a passenger will make it through the network in less time that the average. This is an indication that there is congestion taking place.

For a better understanding of the percentages differences, they have been plotted as a histogram in Graph 84.



Graph 84 - No. of Journeys that lie in Percentage increase of Historical Average: 7th Sept Eastbound

Graph 84 shows that the travel times on the day of the 7th September with a distribution centred close to 0. The skew in Graph 81 is found to be 0.67 (2dp), similarly to Graph 81, so no prominent skew can be seen. However, there appears to be a much greater skew to the right here, which is another indication that there is congestion.

Finally for this day the possible delays have been studied.

The results from an algorithm that discovered how many moving average points are over their respective average travel time plus five minutes first seen in Section 5.4, are seen below. The results of all values that may be over this threshold contain a large number of results; 1001 journeys were delayed out of a possible 22524 journeys completed on the 7th September on the Island Line travelling westbound. To determine if any of these 1001 results are genuine delays a further constraint has been added that two or more delays have to happen in the same minute.

Table 51- Line Delays: Number of moving average points with delays in common in the same minute. 7th September Eastbound

Line Delays: Number of moving average points with delays in common in the same minute				
2	**3**	**4**	**5**	**6**
09:42	12:39	15:14	17:55	19:18
09:57	12:29	18:28	18:29	
10:13	14:25	18:30	18:42	
11:14	14:33	18:46	18:43	
11:19	14:54	18:47	18:57	
11:38	15:10	18:59	19:03	
11:39	15:44	19:31	19:05	
11:44	15:48	19:42	23:47	
11:56	16:07	21:30		
11:57	16:10	23:28		
11:59	16:25			
12:06	16:40			
12:19	17:46			
12:22	17:48			
12:24	17:50			
12:31	18:36			
12:35	18:44			
12:46	18:50			
12:57	18:55			
13:13	18:56			
13:17	18:58			
13:20	19:02			
13:25	19:06			
13:26	19:07			
13:27	19:08			
13:36	19:19			
14:01	19:25			
14:21	19:38			
14:38	19:39			
14:41	19:40			
14:50	19:47			

14:55	20:05				
14:57	20:30				
15:01	21:29				
15:02	22:48				
15:08	23:08				
15:26	23:29				
15:30	23:33				
15:31	23:57				
15:37					
15:39					
15:46					
15:47					
15:50					
15:51					
15:55					
15:56					
16:04					
16:11					
16:46					
16:56					
16:57					
16:58					
17:00					
17:05					
17:16					
17:21					
17:32					
17:34					
17:35					
17:36					
17:37					
17:47					
17:51					
17:52					
17:53					

17:54					
17:56					
17:57					
17:59					
18:00					
18:02					
18:04					
18:05					
18:09					
18:10					
18:12					
18:15					
18:16					
18:19					
18:20					
18:26					
18:31					
18:32					
18:38					
18:39					
18:45					
18:48					
18:49					
18:53					
18:54					
19:01					
19:04					
19:09					
19:10					
19:11					
19:13					
19:15					
19:20					
19:30					
19:33					

19:36				
19:41				
19:44				
19:45				
19:49				
19:50				
19:51				
19:52				
20:07				
20:11				
20:45				
20:59				
21:05				
21:09				
21:27				
21:45				
21:58				
22:02				
22:03				
22:07				
22:29				
22:32				
22:43				
22:54				
23:04				
23:17				
23:18				
23:19				
23:21				
23:27				
23:30				

Table 51 shows that there are many cases of two delays occurring in the same minutes, it is unlikely that there are delays to the service throughout the whole day therefore it is decided that these delays must be caused by anomalies. This data has been split into how many delays there are in common within the same minute. However, the delays seen under 6 delays in common would appear in 5 delays on and so on. Bearing this in mind, column 3 is interesting. It shows around the PM peak that there are fairly consistent delays appearing. This may be an indication of congestion at either an entrances or exits or an unreported operational delay.

Table 52 – Exit and entrance delays: Number of moving average points delayed within the same minute. 7th September Eastbound

	Number of Delays in Common			
	2		3	
	Station Name	**Time of Delay**	**Station Name**	**Time of Delay**
Exit Station Delays	Causeway Bay	11:14	Chai Wan	23:28
	Sai Wan Ho	13:36		
	Tai Koo	14:41		
	Tai Koo	14:54		
	Shau Kei Wan	16:07		
	Causeway Bay	16:25		
	Tai Koo	16:40		
	Tai Koo	17:00		
	Tai Koo	17:46		
	Wan Chai	18:47		
	Fortress Hill	18:56		
	Causeway Bay	18:59		
	Sai Wan Ho	19:02		
	Causeway Bay	19:19		
	Causeway Bay	19:20		
	Causeway Bay	19:51		
	Tai Koo	22:48		
	Chai Wan	23:17		
	Chai Wan	23:18		
	Tai Koo	23:27		
	Tai Koo	23:29		
Entry Station Delays	Central	11:24		
	Causeway Bay	13:03		
	Central	13:30		
	Central	13:44		
	Central	14:45		
	Causeway Bay	16:00		
	Sheung Wan	16:14		
	Causeway Bay	16:38		
	Central	16:55		
	Sheung Wan	17:08		
	Causeway Bay	17:26		

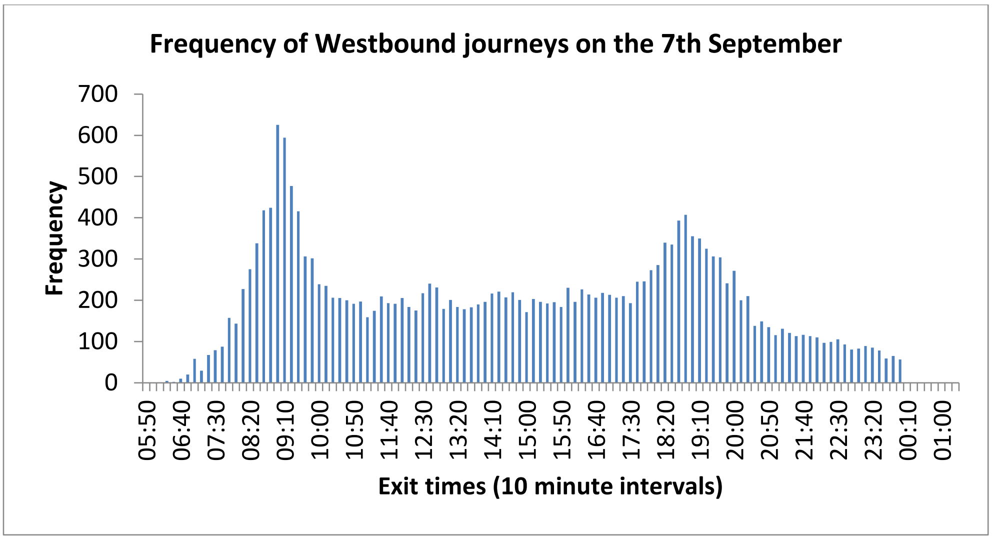
	Sheung Wan	17:38		
	Central	17:49		
	Sheung Wan	17:51		
	Causeway Bay	17:54		
	Admiralty	18:08		
	Sheung Wan	18:13		
	Sheung Wan	18:14		
	Wan Chai	18:16		
	Sheung Wan	18:18		
	Central	18:19		
	Central	18:21		
	Wan Chai	18:22		
	Admiralty	18:24		
	Wan Chai	18:25		
	Admiralty	18:34		
	Admiralty	18:35		
	Admiralty	18:41		
	Causeway Bay	18:42		
	Admiralty	18:44		
	Central	18:47		
	Central	18:48		
	Central	18:50		
	Sheung Wan	18:52		
	Sheung Wan	18:56		
	Admiralty	19:02		
	Admiralty	19:08		
	Sheung Wan	19:26		
	Admiralty	19:41		
	Sheung Wan	19:43		
	Central	19:53		
	Central	21:57		

Table 52 shows that in most cases delays can be seen to passengers in the PM peak, this concurs with the results found in Graph 82, which that shows that most passengers on the Island line appear to travel east for work and west to go home. In comparison to Table 49 it can be seen there is a greater number of delays being reported. Again 2 delays in a minute is a low number so, it would appear not many passengers are delayed. However, there is evidence that passengers are being delayed in the evening peak on this day.

5.5.2. Westbound

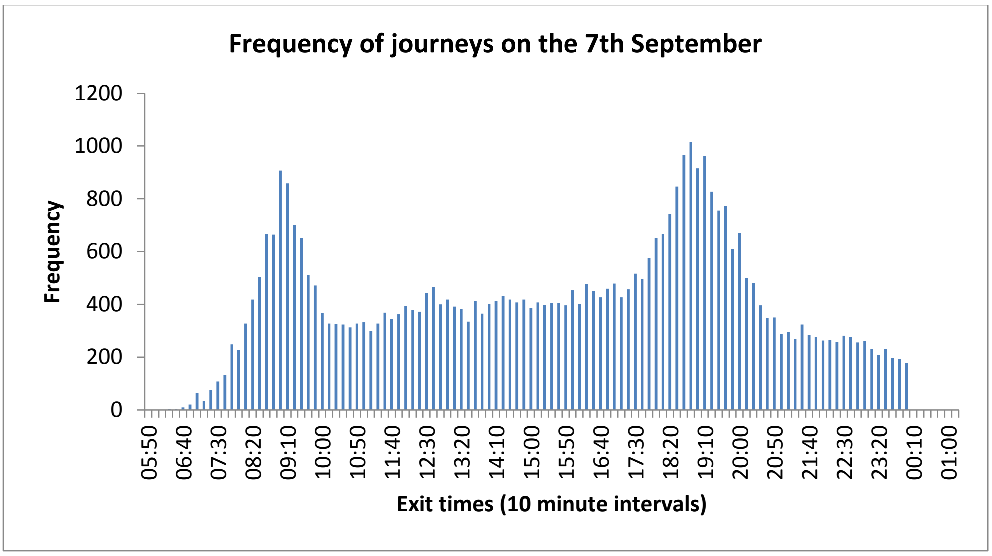
5.5.2.1. 7th September

So far congestion analysis has focused on eastbound journeys. It was seen that the trend would be that passengers travel from east to west for work. This would suggest the westbound analysis would show the AM congestion. For eastbound analysis the 7th September has been chosen to be studied, since it seemed to show more delays than the 18th September. Graph 85 shows the frequency of journeys taken on the 7th September westbound on the Island line.



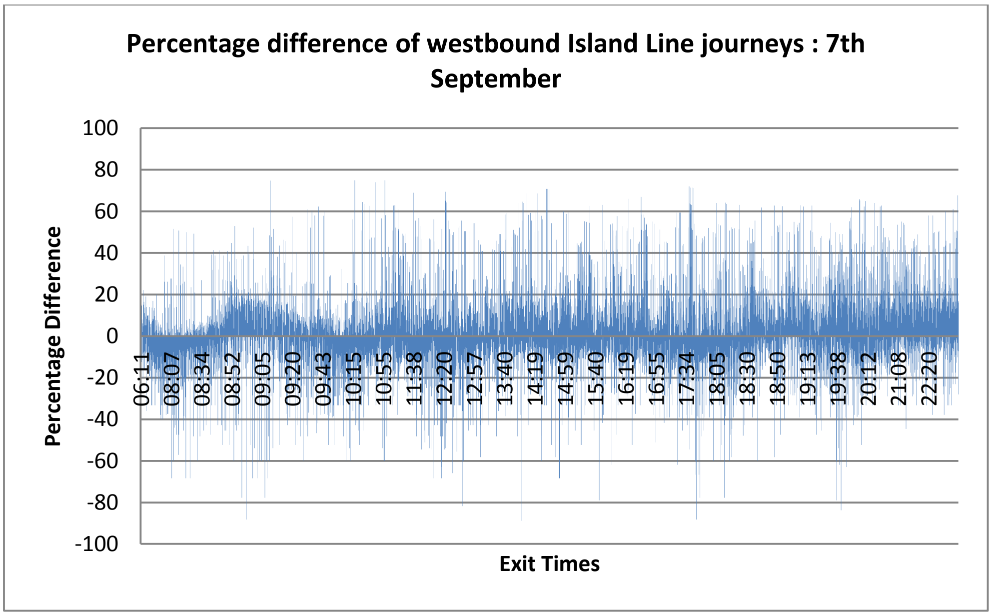
Graph 85 - F requency of Westbound journeys on the 7th September

As predicted it can be seen that passengers tend to travel east for work, with Graph 85 being almost a mirror image of Graph 82. To understand the pattern of a day's travel better, Graph 86 shows the frequency of all journeys on the 7th of September, it can in fact be seen that the pattern of the day is not quite symmetrical with more journeys being completed in the PM peak than the AM peak.



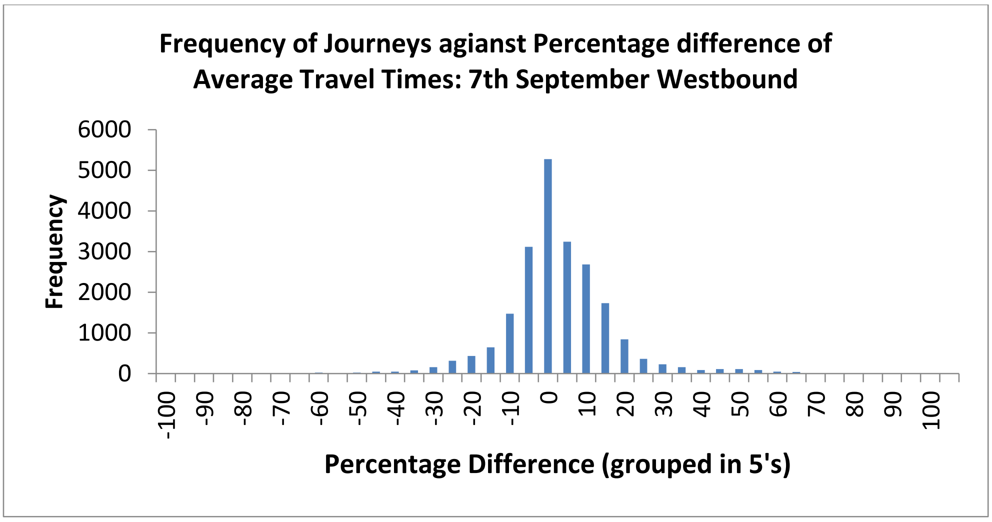
Graph 86 - Frequency of journeys in both directions of the Island line on the 7th September

Graph 87 shows the percentage differences throughout the day between the journey times recorded through the Octopus card and the average travel times.



Graph 87 - Percentage difference of westbound Island Line journeys : 7th September

Graph 87 shows quite a consistent pattern from around 10:00 of journeys lying between -20% and +20% of the average journey time. However there is a clear dip in the time it takes passengers between 06:30 and 08:30, showing very few passengers are taking longer than average at that time. This is followed by a rise in passengers' times around 08:50. This could be a sign of congestion caused by passengers aiming to get to work for 09:00. The rise appears to settle around 09:10, when although it would appear more passengers seem to be making a quicker journey than average, there is more of an even distribution either side of the 0 line.



Graph 88 - Frequency of Journeys agianst Percentage difference of Average Travel Times: 7th September Westbound

Graph 88 shows the distribution of travel times on the 7th September in comparison to the average travel times found. It can be seen that the distribution has a skew to the right; in fact the skew is calculated to be 0.47. In comparison to Graph 81 there is a more prominent skew to the right yet not as great as that seen in Graph 84. A skew to the right does show evidence of congestion, but as there is not a prominent skew this suggests that there is no delay.

Finally, analysis was completed to determine if delays could be found due to congestion, using the methodology first given in Section 4.5, on the 7th September to passengers travelling westbound. The results of entrance and exit delays are shown in Table 52.

Table 53 – Entrance and exit delays: Number of delays in common in the same minute. 7th September Westbound

	Number of delays in common			
	2		3	
	Station name	Time of delay	Station name	Time of delay
Exit station delays	Sheung Wan	09:07	Sheung Wan	09:03
	Central	11:14		
	Central	11:18		
	Central	11:24		
	Causeway Bay	12:16		
	Sheung Wan	13:33		
	Sheung Wan	13:34		
	Tai Koo	14:24		
	Sheung Wan	16:20		
	Sheung Wan	16:21		
	Sheung Wan	16:42		
	Sheung Wan	17:37		
	Sheung Wan	17:38		
	Sheung Wan	17:39		
	Wan Chai	18:10		
	Admiralty	18:11		
	Causeway Bay	19:17		
	Central	19:20		
	Admiralty	19:40		
	Central	19:42		
	Wan Chai	19:51		
	Admiralty	20:08		
	Admiralty	20:09		
	Causeway Bay	22:22		
	North Point	23:35		
Entrance station delays	Chai Wan	08:19	Chai Wan	08:25
	Chai Wan	08:20	Chai Wan	08:32
	Chai Wan	08:23		
	Chai Wan	08:24		
	Heng Fa Chuen	08:25		
	Chai Wan	08:27		
	Chai Wan	08:29		
	Chai Wan	08:35		
	Sai Wan Ho	08:44		
	Chai Wan	10:31		
	Shau Kei Wan	11:41		
	Heng Fa Chuen	11:44		
	Causeway Bay	13:56		
	Wan Chai	15:54		
	Wan Chai	16:30		
	Fortress Hill	17:17		

	Tai Koo	17:19		
	Quarry Bay	17:50		
	Causeway Bay	18:43		
	Heng Fa Chuen	18:47		
	Quarry Bay	18:57		
	Tai Koo	19:28		

The entrance station delays seem to show more delays before 09:00, in particular it can be seen that passengers entering at Chai wan between 08:19 and 08:35 are getting delayed.

Table 54 shows the results of the line delay analysis. The algorithm used here was first discussed in Section 4.5. Line delays found when 2 moving average points were found to be delayed in the same minute were left out of as it was concluded that as they were so regular it was likely they were due to anomalies and therefore are not shown.

Table 54 - Line Delays: Number of moving average points with delays in common in the same minute. 7th September Westbound

Line Delays: Number of moving average points with delays in common in the same minute			
3	**4**	**5**	**7**
08:52	08:53	16:42	09:08
08:54	09:07	18:34	
08:55	17:40		
08:56	18:58		
09:03	19:17		
09:04	20:08		
09:33			
10:36			
11:24			
11:28			
11:46			
11:54			
12:06			
14:04			
14:16			
14:37			
16:38			
18:05			
18:07			
19:03			
19:40			
19:41			

Table 54 shows that there appears to be a number of delays between 08:52 and 08:56; this is what was seen in Graph 87 and evidence that within this small timeframe passengers are experiencing delays.

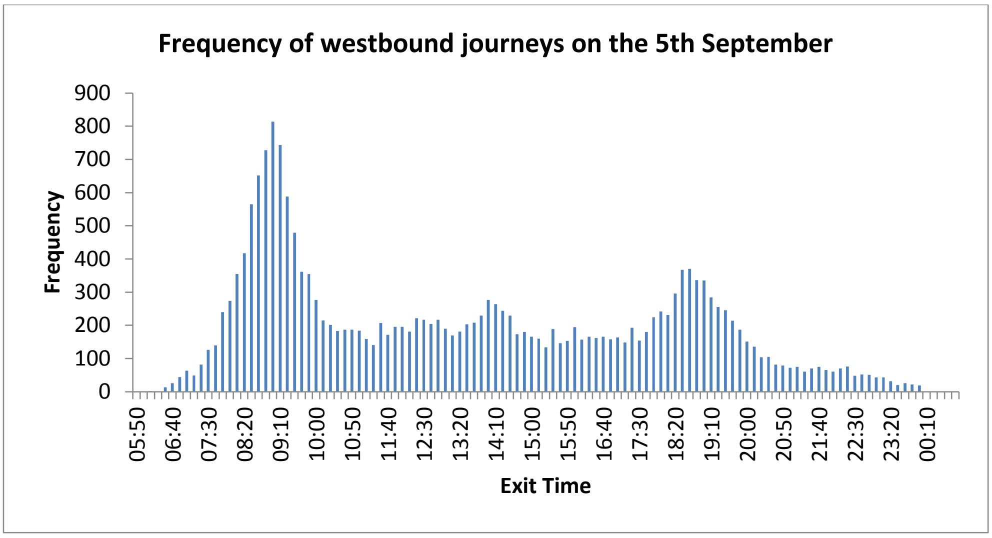
In this section it has been suggested that direction and day of week make a large difference in whether congestion will be experienced. The 18th September, a Tuesday, showed no apparent delays to passengers caused by congestion. However, on the 7th September, a Friday, passengers seemed to be experiencing delays in the evening in the eastbound direction. When analysing the 7th September in the westbound direction, however, there was very little evidence of delays. Beyond this 4 days within the month of September 2012 have been analysed to determine how visible operational delays are in the data.

5.6. Delays reporting

In order, to determine if Octopus data can be used to see how passengers are affected when there are operational delays a number of different days containing operational delays to the Island line will be analysed. These days were chosen from Table 35. The first day to be analysed will be the 5th September according to Table 35 the delay took place at Sheung Wan in the westbound direction. The delay started at 20:29 and lasted 13 minutes effecting 1 train.

5.6.1. 5th September

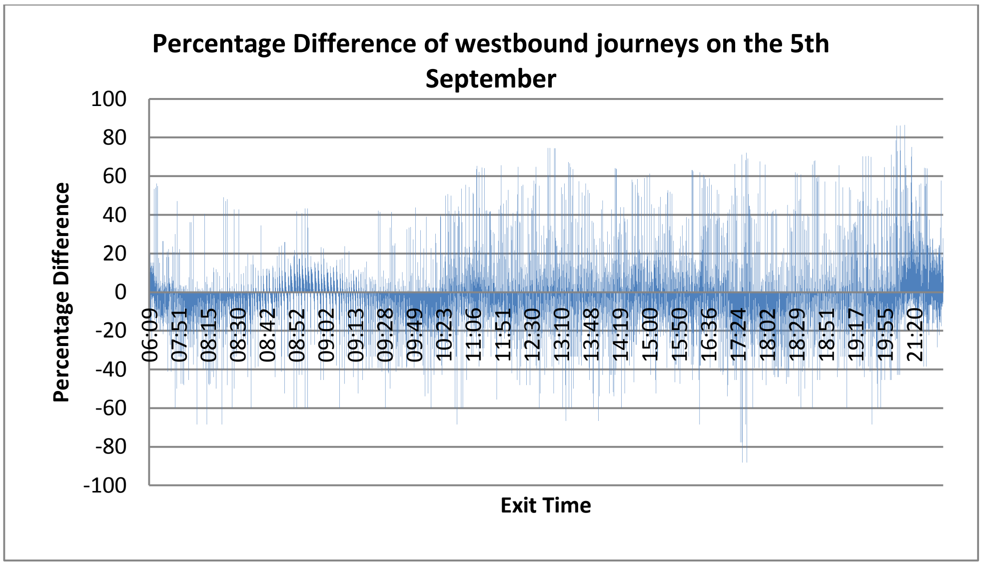
Although it is known that the delay takes place in the westbound direction. Both directions in this case are to be analysed to understand how a delay at the end of the line affects the service. Therefore to start the analysis of this day westbound journeys will be studied. To start, Graph 89 shows the frequency of westbound journeys on the 5th September.



Graph 89 - Frequency of westbound journeys on the 5th September

Graph 89 shows a similar pattern to Graph 85, with more journeys taken in the morning than the evening. The delay takes place at 20:29, at this time it is clear there are very few passengers travelling on the line. This will make analysis harder as with fewer passengers there are fewer results to look at.

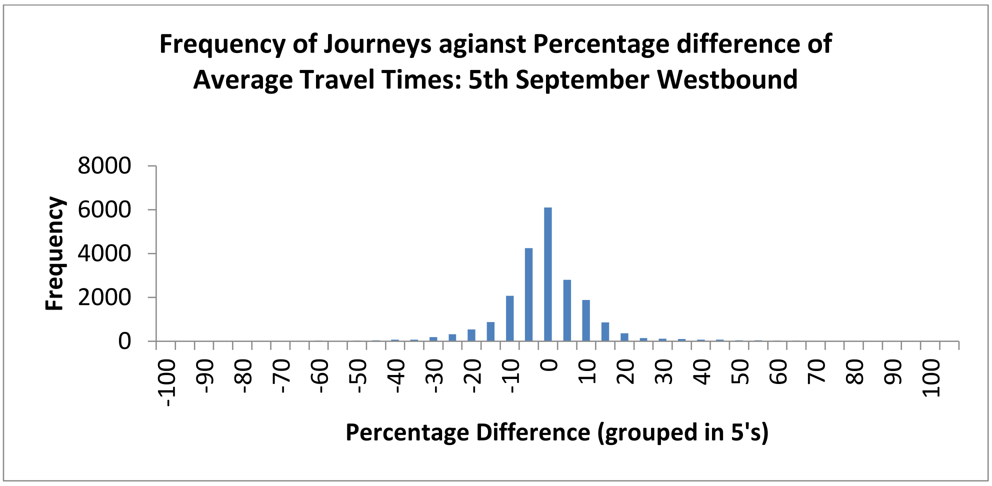
Next Graph 90 shows the percentage difference between the average travel time found in Section 5.2 and the journeys taken on the 5th of September.



Graph 90 - Percentage Difference of westbound journeys on the 5th September

Graph 90 exhibits the same pattern as seen in Graph 87, but the rises and dips in the morning peak are much more visible. Again it is clear that there are delays experienced by passengers between 08:45 and 09:00.

Next the distribution of travel times for the 5th September, westbound, is analysed.



Graph 91 - Frequency of Journeys agianst Percentage difference of Average Travel Times: 5th September Westbound

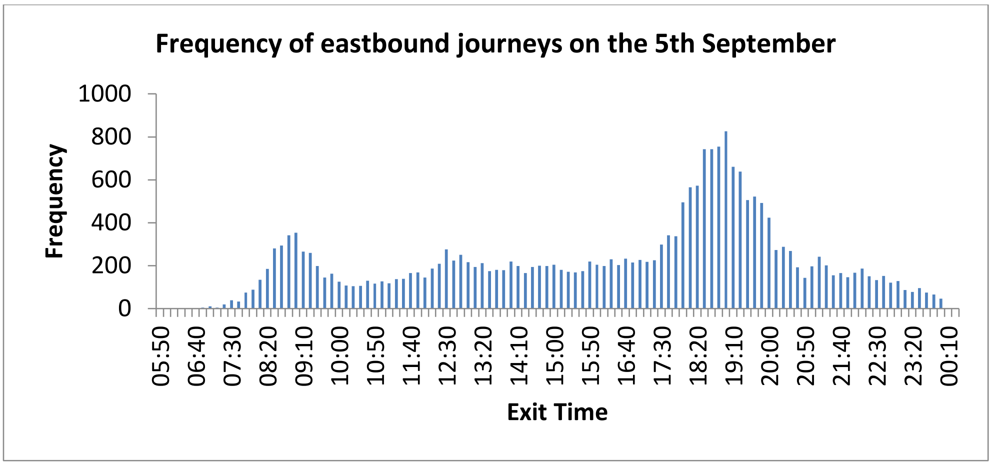
Graph 91 seems to show a different pattern to that seen in Graph 88. Here the data is skewed to the right by 1.01 this is above the threshold of a prominent skew, discussed in Section 5.5 implying that there is indeed a delay taking place. Finally the data was studied to determine if there were any clear line delays.

Table 55 - - Results: Line Delays: 5th September Westbound

Number of delays in a minute	Time	Delay in minutes (number of minutes over journey average + 5 minutes)
3	08:55	3
4	08:56	6
3	08:58	1
3	09:02	1
4	11:19	3
3	11:34	5
3	11:53	3
3	12:19	3
3	13:26	2
3	16:05	2
3	16:32	5
3	16:34	1
3	16:36	6
3	17:32	6
3	17:42	8
3	17:43	4
3	17:44	4
5	18:29	4
3	18:32	4
3	19:07	2
3	19:16	4
3	19:17	6
3	19:26	7
3	21:17	2

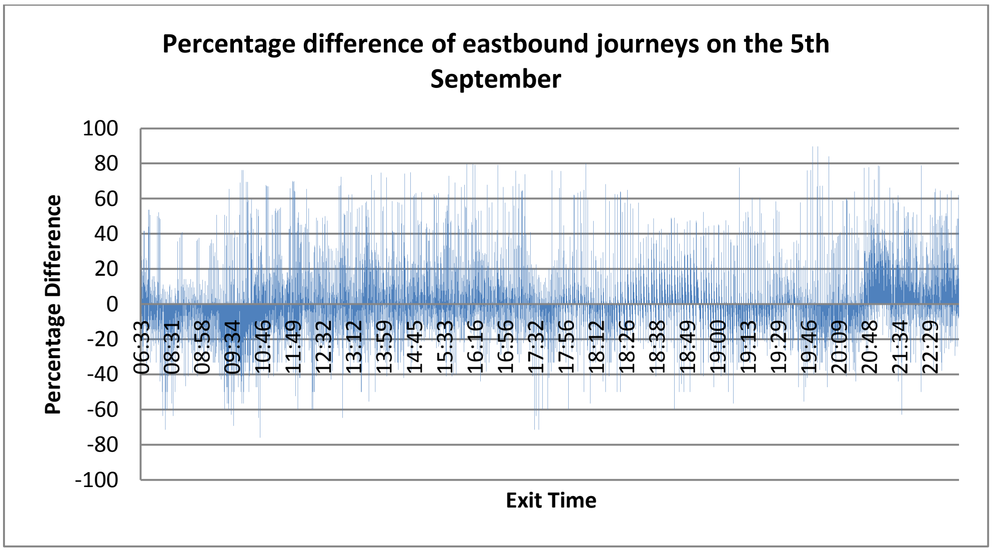
Table 55 shows that the results appear to be sporadic and therefore can be appear to be anomalous rather than any meaningful delay, there is a slight clustering of delays around 08:56 and 17:44, that would suggest congestion. There appears to be no delays seen around 20:30, showing that the delay experienced at the end of the westbound line does not appear to affect the passengers.

Next in analysing the 5th September the eastbound trains will be analysed. Firstly Graph 92 shows the frequency of journeys on the 5th September eastbound.



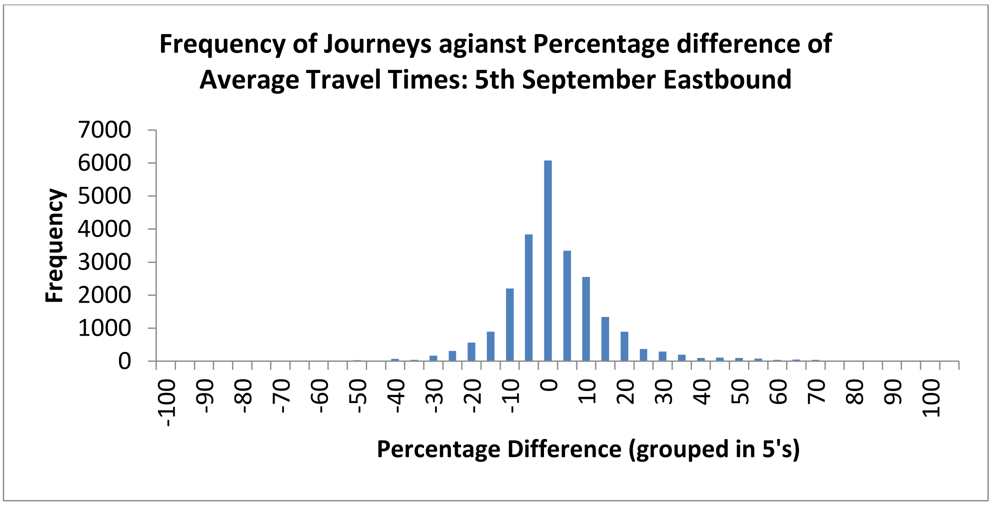
Graph 92 - Frequency of eastbound journeys on the 5th September

Graph 92 appears to follow the same pattern as Graph 78 and Graph 82, with similar numbers shown as on the 18th September. It can be seen that around the time of the delay at 20:30 there is a clear drop in numbers. This suggests a delay, which could be caused for example by the train not being able to start the route back in the eastbound direction as it is delayed at the end of the westbound line. As the evening peak appears to be ending there will be a lower passenger frequency which means it is less likely that delays will be able to be seen in the data as there is less data to look at; however this also means that fewer passengers will be affected.



Graph 93 - Percentage difference of eastbound journeys on the 5th September

Graph 93 shows a similar pattern in travel times as in Graph 79 and Graph 83. It can be seen in the evening peak between 18:00 and 18:30 that passengers seem to be taking their journeys in general quicker than the average. Then between 18:30 and 19:00 it can be seen nearly all passengers seem to take longer. From 19:00 till 19:30 passengers again can be seen to be travelling through the system quickly. This then evens out either side of 0 until 20:40 when interestingly there appears to be a spike in passengers travel times. This is a good sign that the delay is visible in the data; since passengers travel times appear to be higher.



Graph 94 - Frequency of Journeys agianst Percentage difference of Average Travel Times: 5th September Eastbound

Graph 94 shows the distribution of travel times on the 5th of September on the Island Line, eastbound. This distribution has a skew to the right, as to be expected, the skew is calculated to be 0.77 which is not shown to be a meaningful skew. This would suggest there is a delay taking place on this day, either as a result of congestion or the delay to the service.

Finally the data was analysed to determine if a line delay could be found in the data, the results are shown below.

Table 56 – Results: Line delays. 5th September Eastbound

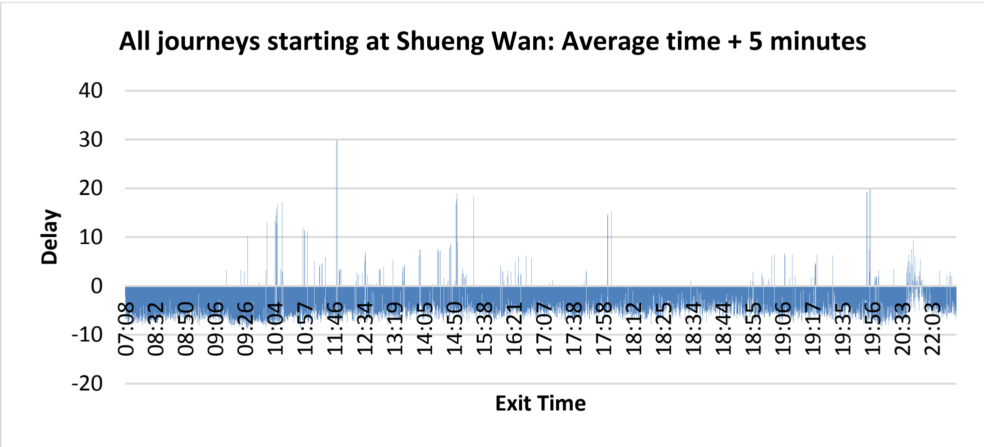
Number of delays in a minute	Time	Delay in minutes (number of minutes over journey average + 5 minutes)
4	12:00	2
3	12:10	1
3	12:14	7
3	12:54	8
3	13:25	6
4	13:34	4
3	13:36	3
3	13:59	7
3	15:40	6
3	15:43	2
5	16:03	6
3	16:17	6
3	17:13	0
3	17:18	4
3	18:01	8
3	18:03	11
3	18:37	2
4	18:38	0
3	18:44	8
3	18:45	4
4	18:46	2
5	18:47	1
4	18:49	3
5	18:50	5
3	18:51	8
5	18:52	2
5	18:53	4
5	18:55	5
4	18:56	4
3	18:57	4
3	18:58	1
3	18:59	2
8	19:01	4
3	19:14	4
3	19:17	5
4	19:18	3
5	19:19	7
5	19:28	3
3	19:30	1
4	19:53	7
5	20:51	2
3	21:00	3
4	21:02	2
5	21:03	8
5	21:04	5
4	21:05	3
5	21:06	3
4	21:07	3
3	21:09	1

6	21:10	5
6	21:11	3
4	21:12	1
3	21:14	1
5	21:16	1
5	21:23	6
4	21:24	5
3	21:30	3
3	22:40	8

The results shows the congestion seen in Graph 93 between 18:37 and 19:00. Further it can be seen there are consistent delays between 21:00 and 21:15. This suggests that the delay lasted 15 minutes and the passengers were experiencing delays between 6 – 8 minutes over their expected travel time.

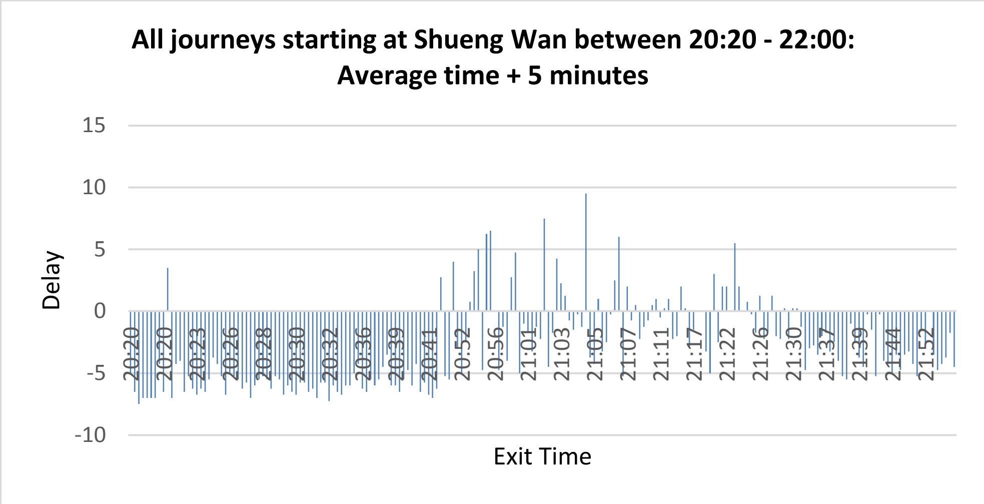
This would suggest that when there is a delay to the last station in the westbound direction, this will affect the start time of a train in the eastbound direction.

Further to understand exactly what is happening at Sheung Wan station, where the delay was reported, Graph 95 and Graph 96 show the difference between the delay threshold; average time plus five minutes and the journeys recorded.



Graph 95 - All journeys starting at Shueng Wan: Average time + 5 minutes

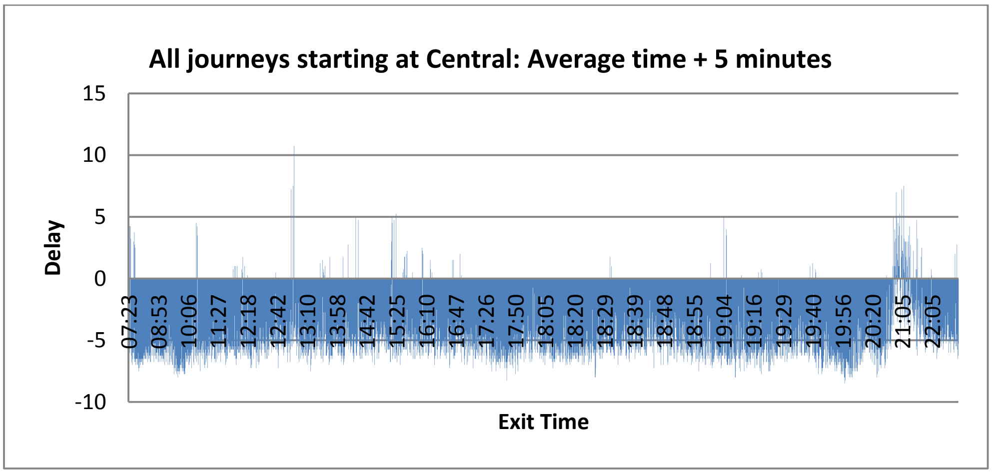
Graph 95 shows all journeys on the 5th September eastbound and their relative delay statuses. This shows us where there are clear anomalies. However around 21:00 there seems to be a denser region with few values falling below 0, for no delay.



Graph 96 - All journeys starting at Shueng Wan between 20:20 - 22:00: Average time + 5 minutes

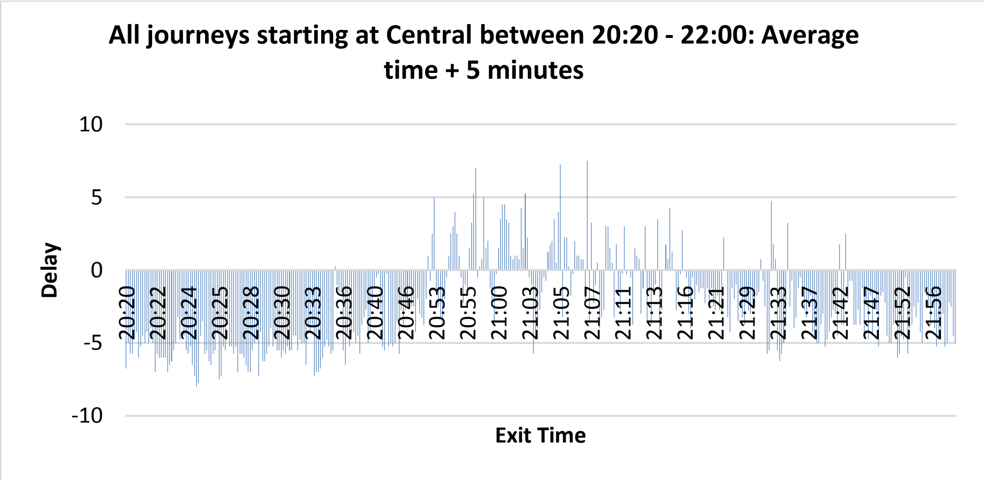
Graph 96 shows only the journeys competed between 20:20 and 22:00. Here the effect of the delay to passengers travelling from Sheung Wan can be seen.

Finally, Graph 97 and Graph 98 show the delays found from the journeys that started at Central station, the next station on from Sheung Wan. Again both of these show delays around 21:00. Showing a number of passengers and trains were delayed.



Graph 97 - All journeys starting at Central: Average time + 5 minutes

Graph 97 shows all journeys throughout the day on the 5th September starting at Central station. Plotting the data like this shows us what are likely to be anomalies in the data and what can be seen as a delay. Here it can be seen there is a clear rise in travel times around 21:00.

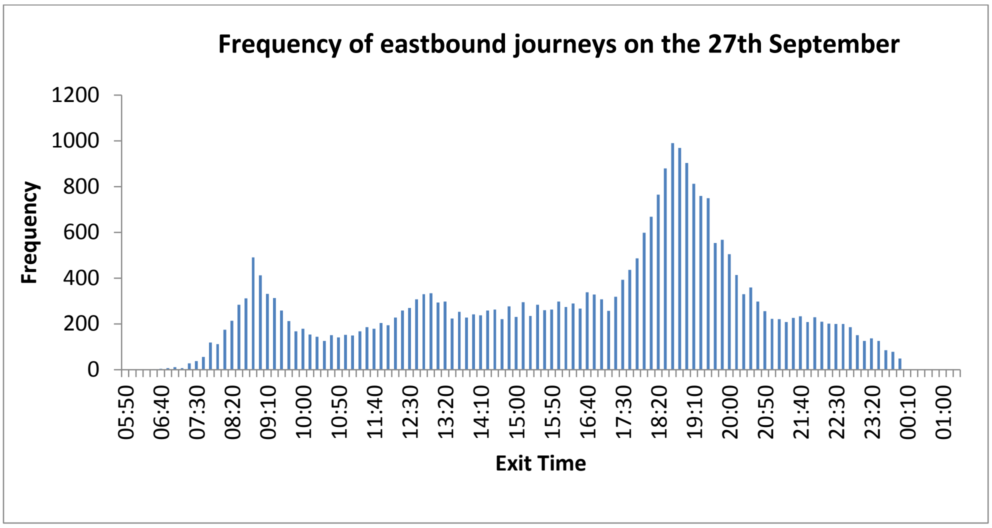


Graph 98 -All journeys starting at Central between 20:20 - 22:00: Average time + 5 minutes

The graph above shows what exactly is happening with the data in the time period between 20:20 and 22:00. Each line represents a passenger's journey. It can be seen that a number of passengers are showing delays in this time period between 20:50 and 21:20. In conclusion for this day it shows in the data that there are delays experienced by passengers between 21:00 and 21:15 in the westbound direction. This is shown graphically and by the delay algorithm.

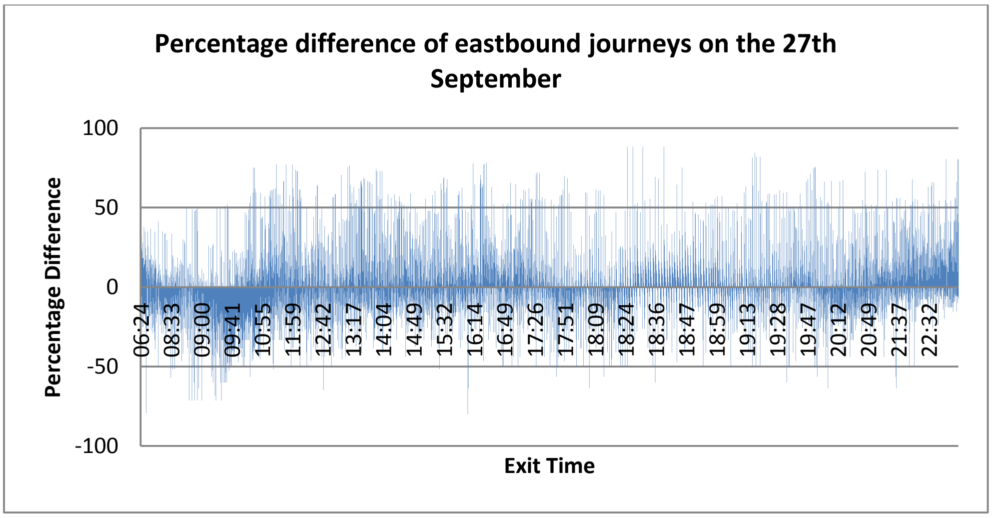
5.6.2. 27th September

The next day to be analysed is the 27th September, according to Table 35, the delay took place at Tin Hau in the westbound direction at 08:48, lasting 5 minutes and affecting one train. To see how the delay has affected passengers both directions have been studied to get a good idea of the delay that is taking place. Firstly the Eastbound direction is studied.



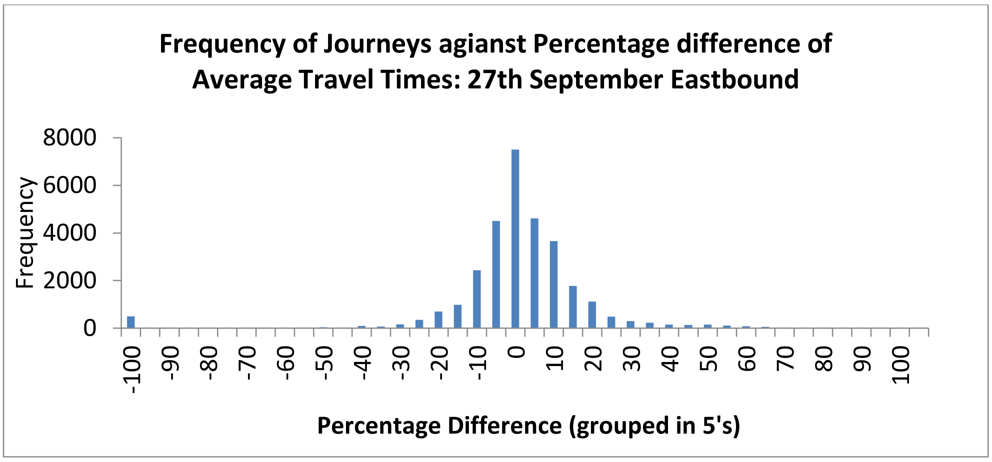
Graph 99 - Frequency of eastbound journeys on the 27th September

Graph 99 shows the frequency of passengers on the 27th September. Similarly to the analysis completed on the 18th and 7th September that shows the frequency of journeys in the eastbound direction (Graph 78 and Graph 82), it can be seen that there is low passenger demand in the morning peak. However the AM peak appears to reach a maximum of just under 600 passengers exiting the system at around 08:50, which should be enough passengers to show a delay should there be one.



Graph 100

Graph 100 shows a similar pattern to that seen in Graph 79 and Graph 83. There appears to be an even distribution either side of 0 at the time of the delay, 08:48, suggesting there was no obvious delay to passengers at this time.



Graph 101 - Frequency of Journeys agianst Percentage difference of Average Travel Times: 27th September Eastbound

Graph 101 shows a distribution centred close to 0 with a very slight skew to the right. The skew was calculated to be 0.95 this is very close to the threshold of a prominent skew which in this case is 1. Implying that it is very likely that there is a large amount of congestion or a delay on this day.

Table 57 shows the result of the algorithm first discussed in Section 4.6. It shows how many passengers are delayed over 5 minutes of their expected travel time, within the same minute. It can be seen here that there appears to be no delay to passengers' travel times during the delay in the AM peak. However, rather unusually there appears to be quite a clear delay taking place in the PM peak that is unreported. Between 18:30 and 20:00 there appears to be a consistent delay experienced by passengers. It can be seen that every couple of minutes a delay is registered that is delaying passengers between one and ten minutes with a large number of passengers showing the delay. Further this delay appears to be taking place throughout the line.

Table 57 – Results: Line delays. 27th September Eastbound

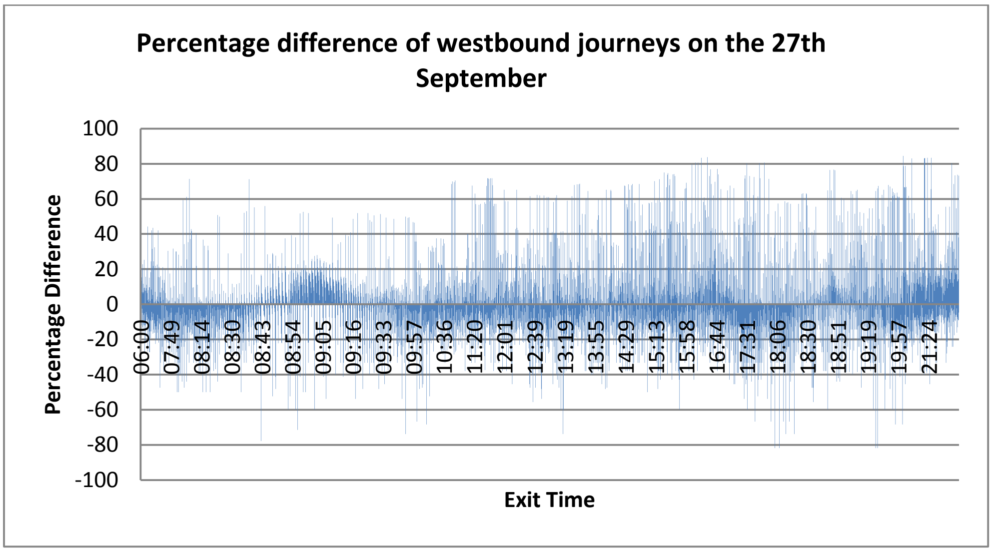
Number of delays in a minute	Time	Delay in minutes (number of minutes over journey average + 5 minutes)
3	08:57	4
3	11:28	4
3	12:55	4
3	13:16	6
3	13:23	2
3	13:25	4
4	13:26	2
3	13:29	7
3	13:35	6
3	13:36	9
3	14:25	3
3	14:30	2
3	14:38	2
3	15:19	1
3	15:34	3
5	15:54	4
3	16:10	4
3	16:12	7
4	16:22	9
4	16:37	7
3	17:14	2
3	17:21	9
3	17:24	6
3	17:35	3
4	17:41	6
3	17:49	2
4	17:53	3
4	17:54	3
3	18:10	7
4	18:11	6
4	18:14	4
3	18:28	7
6	18:29	5
3	18:30	8
4	18:31	4
3	18:34	3
4	18:37	5
5	18:44	4
7	18:46	8
3	18:48	1
6	18:50	3
5	18:51	2
4	18:53	2
4	18:54	4
3	18:55	3
3	18:56	1

	3	18:57	3
	5	18:58	3
	5	19:00	2
	4	19:01	4
	5	19:04	5
	5	19:05	7
	3	19:06	3
	4	19:07	6
	4	19:08	4
	5	19:09	5
	5	19:11	2
	7	19:12	3
	6	19:13	5
	4	19:14	2
	5	19:15	1
	5	19:16	7
	3	19:19	5
	6	19:20	9
	4	19:21	3
	3	19:23	3
	5	19:24	2
	4	19:25	3
	3	19:26	1
	7	19:27	2
	3	19:28	2
	3	19:30	5
	4	19:34	9
	3	19:35	8
	4	19:36	8
	4	19:37	3
	3	19:38	2
	3	19:39	1
	3	19:42	5
	3	19:43	7
	4	19:47	10
	3	19:50	6
	3	19:52	9
	3	19:53	9
	4	19:57	5
	3	19:58	5
	3	20:00	7
	3	20:01	3
	3	20:03	7
	3	20:58	2
	3	21:51	4
	3	21:55	1
	4	22:12	5
	3	22:13	3
	3	22:17	3
	3	22:22	3
	3	22:37	3

3	23:13	4
3	23:18	4

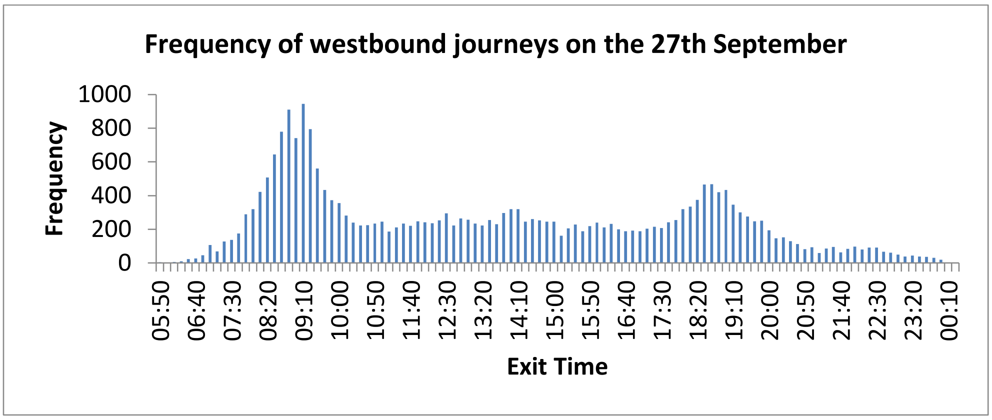
The delay seen in the evening peak seems to start around 18:30 and end around 20:03. In Graph 101 it can be seen there is a rise in passengers' travel times between 18:20 and 19:30 but then seems to settle. This disruptions shows that delays can exist in the network that can be caused just by congestion but affect passengers for longer, with greater delays than the average delay incurred during rush hour. The registered delay in the morning cannot however be seen, in this direction.

Next the westbound journeys will be analysed. The delay was reported to take place in the westbound direction at 08:48. Graph 102 shows the percentage difference between the journeys completed in the westbound direction on the 27th September 2012 and the average travel times found.



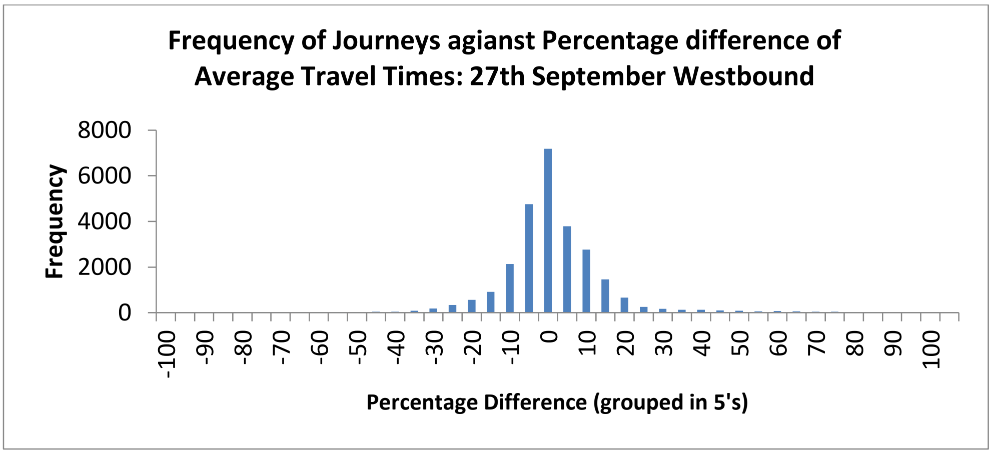
Graph 102 - Percentage difference of westbound journeys on the 27th September

In Graph 102 it would appear that there is a rise to passengers travel times around the time of the delay. The delay was reported at 08:48 and between 08:40 and 09:20 there seems to be an increase in passengers travel times. In Graph 87 it can be seen during the morning peak there is a slight rise to passengers travel times, however the peak seen in Graph 102 appears to affect more passengers, for a longer time.



Graph 103 - Frequency of westbound journeys on the 27th September

Graph 103 shows the frequency of passengers on the 27th September in the westbound direction. At approximately 09:00 there shows to be a dip in frequency. This would be further evidence that passengers may be experiencing a delay in the network. This unusual pattern would suggest that some passengers are leaving the system later than expected causing a gap in the exit frequency.



Graph 104 - Frequency of Journeys agianst Percentage difference of Average Travel Times: 27th September Westbound

Graph 104 shows the distribution of travel times in comparison to the average times found, here the skew is found to be 1.19 this is a prominent skew to the right insinuating that there is a clear delay affecting passengers' times.

Table 58 shows the average delays found on the 27th September in the westbound direction.

Table 58 – Results: Line Delays. 27th September Westbound

Number of delays in a minute	Time	Delay in minutes (number of minutes over journey average + 5 minutes)
3	08:48	1
3	08:55	1
4	08:56	2
3	08:58	1
5	09:02	1
6	09:03	1
5	09:04	2
8	09:05	1
3	09:06	1
3	09:07	1
3	09:08	2
6	09:09	1
3	09:11	2
3	09:12	2
3	11:35	4
3	13:35	8
3	13:45	6
3	13:58	5
3	14:37	4
3	14:38	6
3	15:00	5
3	15:07	7
3	15:10	6
3	15:25	6
3	15:29	7
5	15:30	8
4	15:52	4
4	16:23	5
5	16:33	3
3	17:26	5
3	17:27	4
3	17:29	8
3	17:34	4
3	17:36	7
3	18:07	3
3	18:13	4

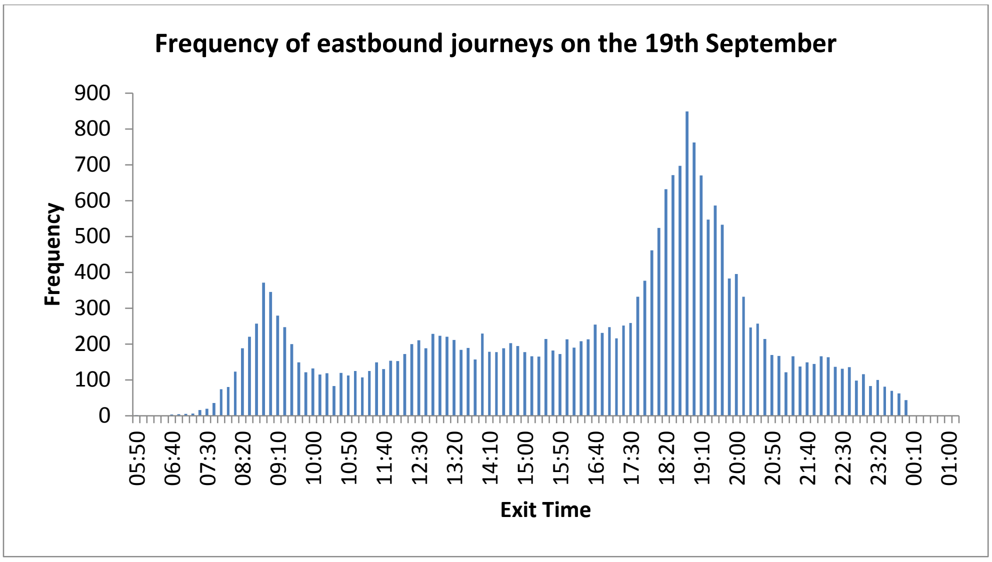
3	18:16	3
4	18:18	3
3	18:22	1
3	18:24	3
4	18:26	9
4	18:27	9
3	18:29	8
3	18:31	2
3	18:35	2
3	18:46	3
5	18:49	3
3	19:06	5
4	19:07	4
4	19:15	5
3	19:26	5
3	19:28	3
3	19:31	4
3	19:33	1
7	19:34	4
3	19:55	8
3	19:56	7

Table 58 shows that passengers are delayed by a minute or two over the threshold of 5 minutes up until 09:12.

In comparison to Graph 89, delays can be seen quite consistently throughout the day to passengers. Further, the average delay seems to be quite substantial. Although there are frequent gaps between the reportings suggesting that many passengers are indeed traveling un-delayed.

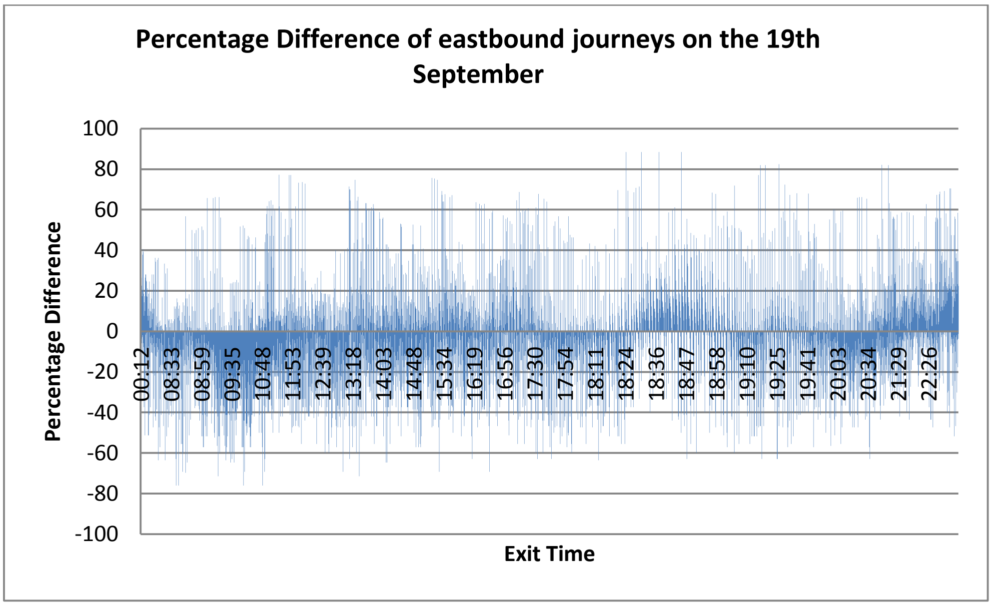
5.6.3. 19th September

The next day to be analysed is the 19th of September as seen in Table 35 the delay stated in Sheung Wan in the westbound direction at 18:09, the delay affected 9 trains and lasted 6 minutes. As seen with the 5th September when there is a delay at the end of the line it affects the departure time of the train heading in the other direction as there is a delay to the train turning around. Therefore for the analysis for this day it will only focus on the eastbound direction.



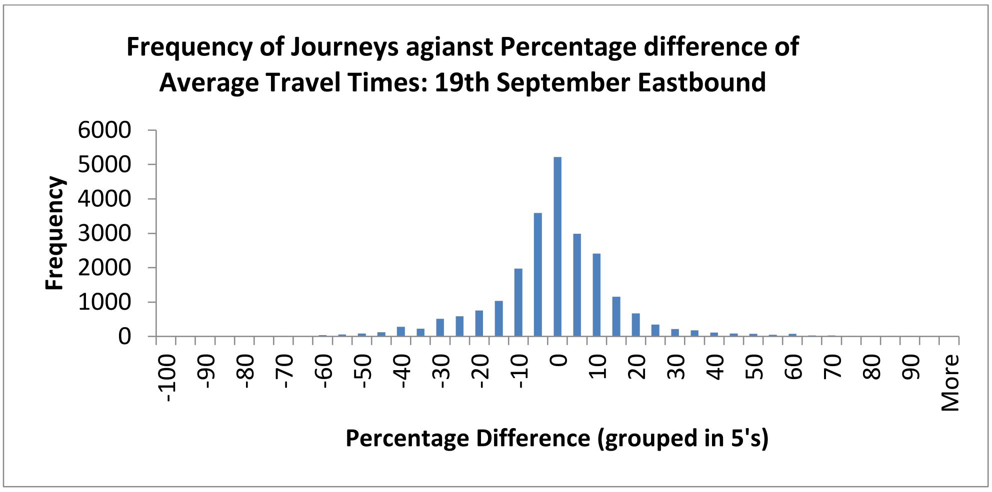
Graph 105 - Frequency of eastbound journeys on the 19th September

Graph 105 shows the frequency of journeys taken on the 19th September; these are all journeys which have an origin and destination on the Island line. There appears to be no noticeable difference between this graph, Graph 78 and Graph 82. Each shows a frequency of around 800/900 during the PM peak. Since the delay is in the PM peak the high frequency will mean that data will give a good indication of what is happening in the system.



Graph 106 -Percentage Difference of eastbound journeys on the 19th September

Graph 106 shows the percentage difference between the average travel times and the travel times on the 19th of September. It can be seen between 18:30 and 19:00 it appears very few passengers appear to be making their journeys in less time than expected. Since the times are averaged over the day it would be expected that the times should show to be evenly distributed either side of the 0 mark. In comparison to the peak seen at 18:30 there seems to be a drop between 17:30 and 18:30.



Graph 107 – Frequency of Journeys: Eastbound 19th September

Graph 107 shows the distribution of travel times on the 19th September, the skew for this day is found to be 0.22. This skew implies that it is very unlikely that there are delays taking place on this day. Finally the number of journeys delayed in the same minute was studied; the results are shown in Table 59.

Table 59 – Results: Line Delays. 19th September Eastbound

Number of delays in a minute	Time	Delay in minutes (number of minutes over journey average + 5 minutes)
3	13:20	2
3	13:24	5
3	13:25	7
3	13:32	3

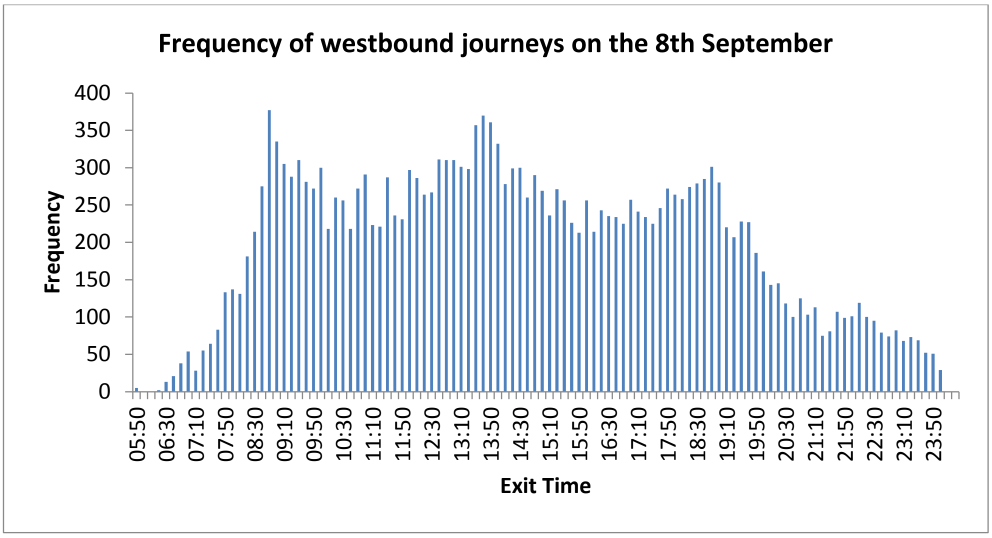
3	13:50	7
3	13:58	7
3	14:47	6
3	15:22	6
3	15:38	2
3	16:56	3
3	16:58	2
3	17:00	6
3	17:16	6
3	17:21	7
4	17:24	5
3	17:41	10
3	17:57	5
3	18:26	7
3	18:29	7
3	18:31	9
4	18:38	8
7	18:39	1
3	18:40	1
6	18:42	2
3	18:43	3
7	18:44	2
8	18:45	1
12	18:46	1
10	18:47	1
10	18:48	2
9	18:49	2
7	18:50	2
6	18:51	1
8	18:52	2
7	18:53	2
5	18:54	3
3	18:55	2
4	18:56	2
8	18:57	4
7	18:58	4
6	18:59	3
5	19:00	4
5	19:01	4
8	19:02	4
3	19:03	2
4	19:04	6
4	19:05	10
3	19:06	1
3	19:07	4
4	19:08	3
3	19:09	2
5	19:10	2
5	19:11	3
3	19:12	3

3	19:13	1
4	19:23	2
5	19:24	3
3	19:27	14
5	19:30	6
6	19:31	4
3	19:34	6
3	19:35	3
3	19:36	9
6	19:37	3
3	19:38	6
3	19:40	2
4	19:42	9
3	19:46	4
3	19:47	3
3	19:48	6
3	22:07	4

The results seen in Table 59 show that there are a large number of passengers being delayed between 18:30 – 19:40 with a variable length in delay. It shows there is a delay in the return of the information with no delay information appearing until 18:30 when the delay was record at 18:09. However it does show that the delay has affected passengers travel times for substantially longer than recorded.

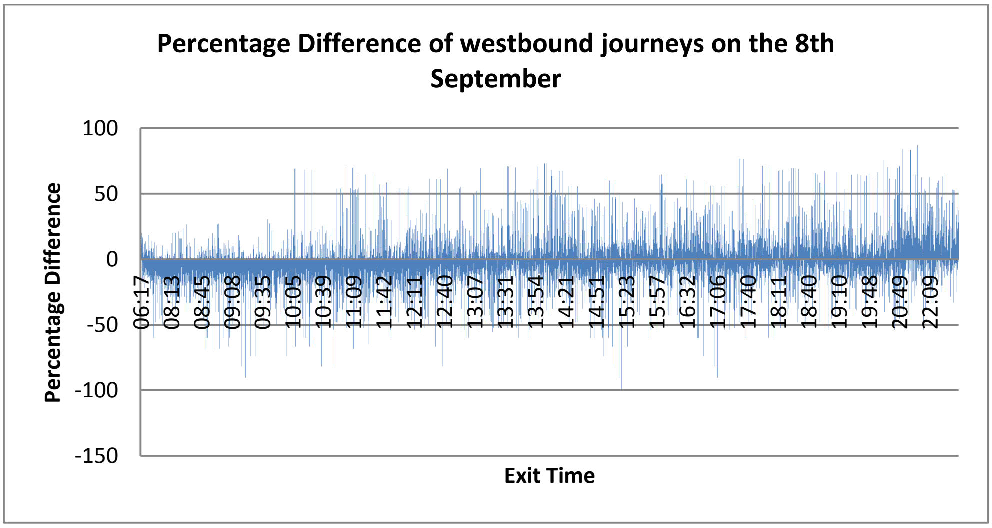
5.6.4. 8th September

Finally the last day to be analysed is the 8th September, the delay took place at 21:20 at Sheung Wan in the eastbound direction. It was unclear if this delay could affect the westbound trains as it is the first station in the eastbound direction. The westbound direction was analysed to see if a delay was caused by trains backing up as they could not depart from Shueng Wan in the eastbound direction.



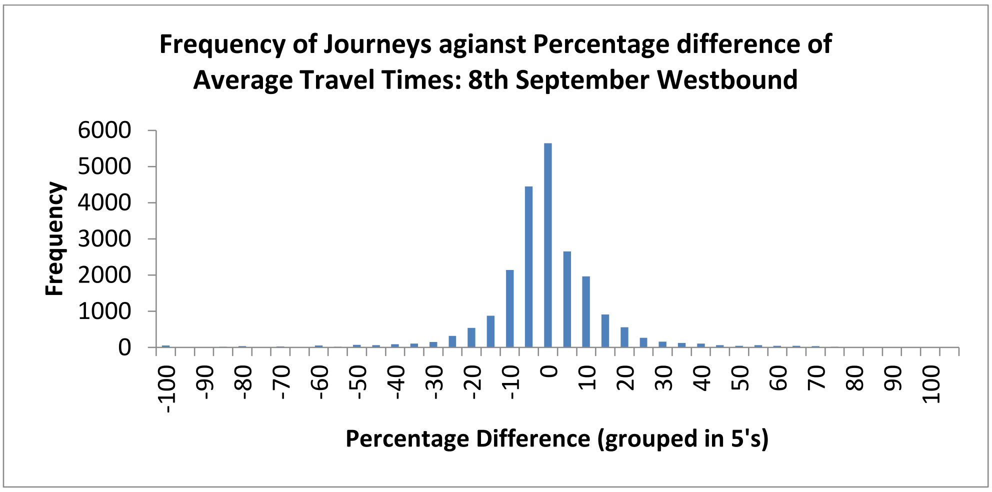
Graph 108 - Frequency of westbound journeys on the 8th September

Graph 108 shows the frequency of westbound journeys completed on the 8th September; this was a Saturday which explains the different pattern. Since a weekend day is yet to be analysed in the data it is difficult to make a comparison. However it can be seen there is very low frequency at 21:20 of around 100 passengers which will make it harder to see a delay.



Graph 109 - Percentage Difference of westbound journeys on the 8th September

Next Graph 109 shows the percentage difference between the expected travel times and the travel times recorded on the 8th September. It can be seen that there appears to be a very slight rise in travel times around 21:20, this may be the sign of a delay.



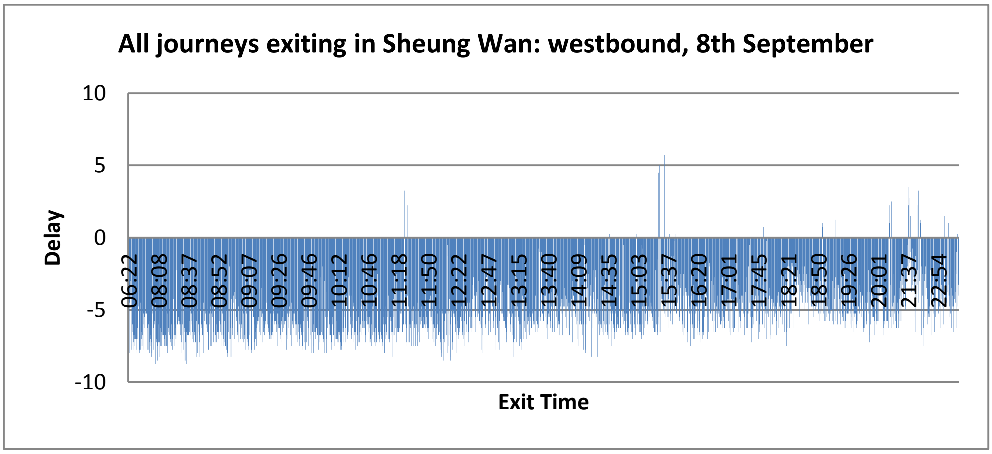
Graph 110 -Frequency of Journeys agianst Percentage difference of Average Travel Times: 8th September Westbound

Graph 110 shows of 0.65 showing that there is a slight skew to the right but nothing prominent, providing no new information in terms of determining if there is a delay or not.

Table 60 – Results: Line delays 8th September Westbound

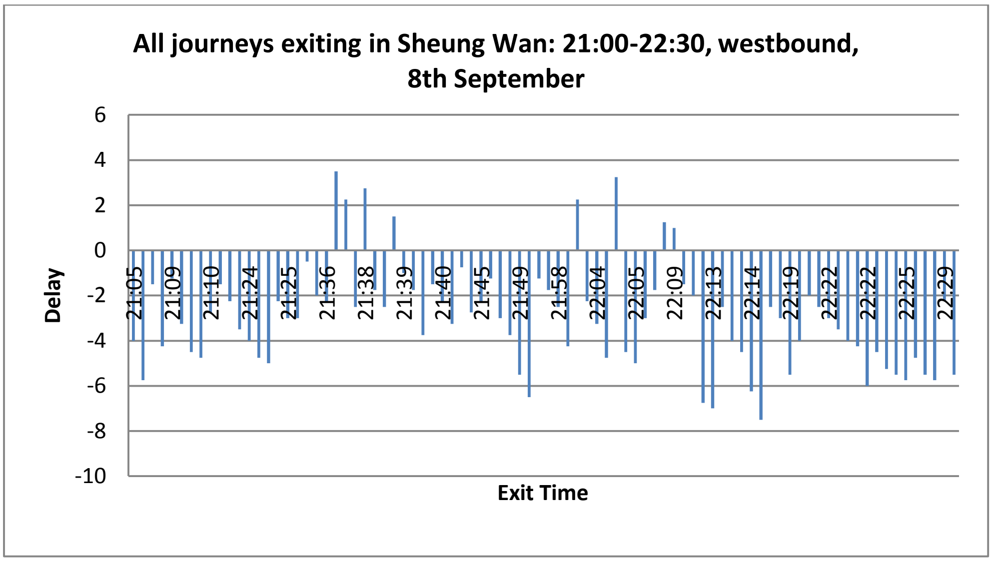
Number of delays in a minute	Time	Delay in minutes (number of minutes over journey average + 5 minutes)
3	11:38	3
3	11:41	6
3	11:47	7
3	13:34	6
3	13:39	7
3	13:56	8
3	13:57	3
3	14:02	8
3	14:12	3
4	14:13	4
3	16:02	8
3	16:24	2
3	16:30	6
4	16:54	1
3	17:14	1
3	18:02	9
4	18:44	3
3	18:55	2
3	19:25	5
3	22:08	5

Table 60 shows the results of what delays were found on 8th September. It can be seen that there is no delay apparent at around 21:20. To understand why Graph 109 showed that there might be a delay Graph 111 and Graph 112 have been plotted to show all journeys exiting at Sheung Wan.



Graph 111 - All journeys exiting in Sheung Wan: westbound, 8th September

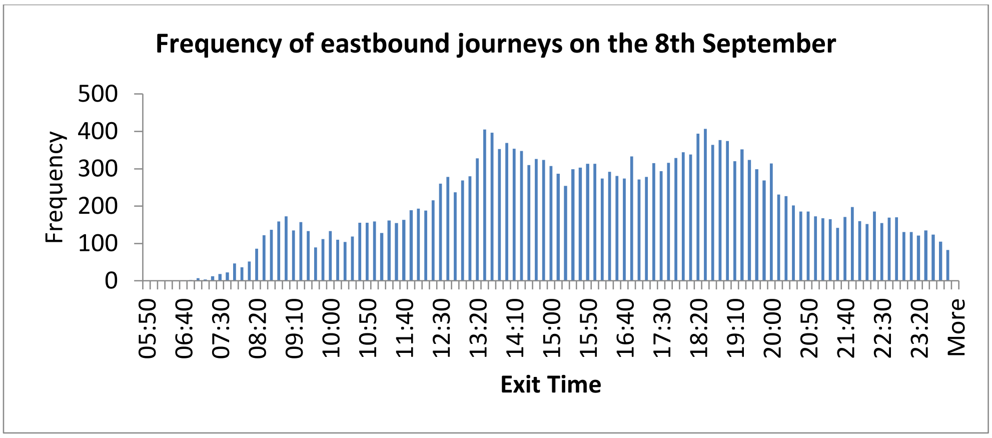
Graph 111 again shows a small clustering of delays around 21:20, hence Graph 112 has been plotted to take a closer look. Graph 112 shows that a very small number of passengers experience a delay. Due to the low frequency this means very few passengers are affected by the delay hence why it didn't appear as a delay in the analysis.



Graph 112 - All journeys exiting in Sheung Wan: 21:00-22:30, westbound, 8th September

This day appears to show a very small number of passengers delayed by the disruption; this is due to low frequency at the time of day that the delay took place. This does show that perhaps the delay in the eastbound direction did prevent trains entering the last station in the westbound direction, Sheung Wan.

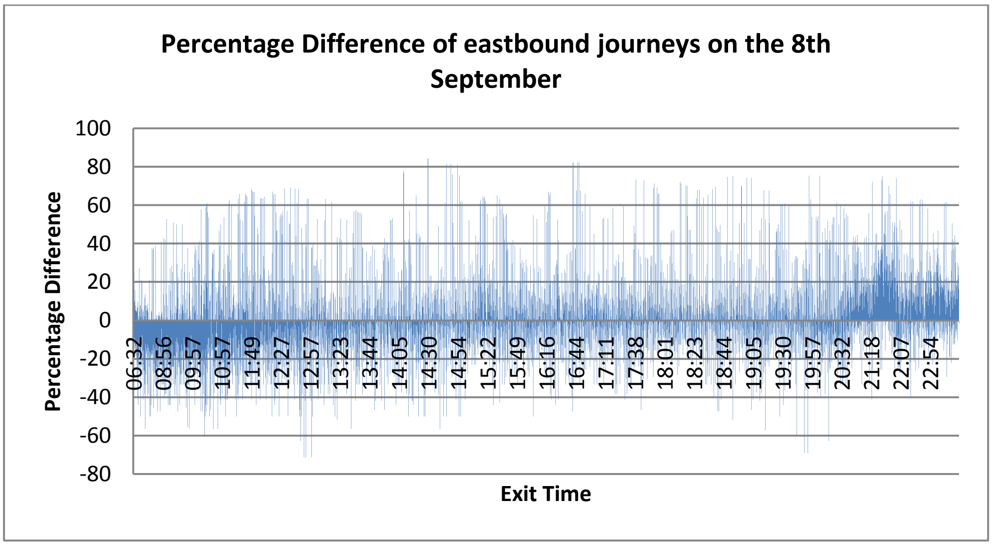
Next analysis of the 8th September will focus on the eastbound direction.



Graph 113 - Frequency of eastbound journeys on the 8th September

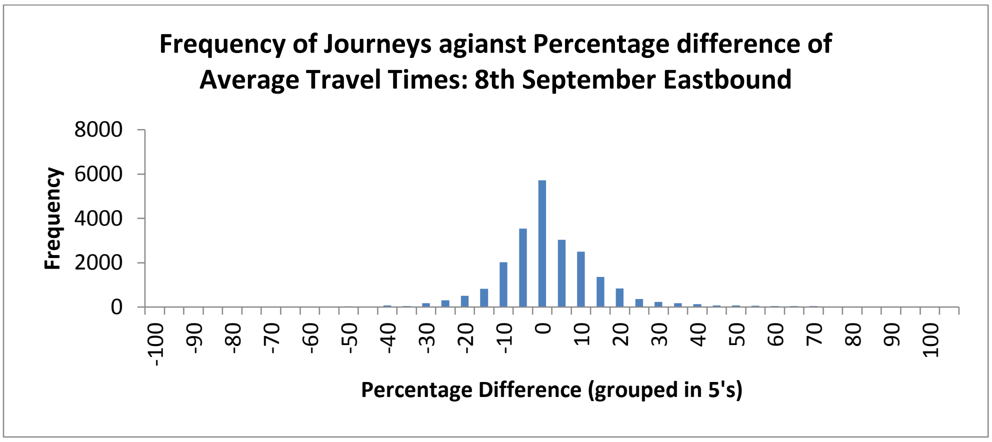
Graph 113 shows the frequency of passengers on the 8th September in the eastbound direction to match Graph 108 there appears to be an evening peak, even though it a weekend day. After this peak the frequency in passengers seems to drop rapidly meaning that there will be less data to show a delay if one can be seen.

Graph 114 shows the percentage difference between the journeys completed in the eastbound direction on the 8th September compared to the average journey times found.



Graph 114 - Percentage Difference of eastbound journeys on the 8th September

Graph 114 shows that around the time of the delay there appears to be a slight rise in travel times. In comparison to the rest of the day, where it appears there is an even distribution of travel times seen either side of the 0 line, at around 21:30 it would seem that few passengers are able to make their journey in less time than the average. Since a weekend day has not been analysed yet, a comparison cannot be made.



Graph 115 - Frequency of Journeys agianst Percentage difference of Average Travel Times: 8th September Eastbound

Graph 115 shows the distribution in travel times on the 8th September in the eastbound direction. Here the skew is calculated to be 1.03 this is over the threshold of a prominent skew implying that there is clearly a delay to passengers on this day.

Finally, Table 61 shows the results of what delays can be seen in the data.

Table 61 – Results: Line delays 8th September Eastbound

Number of delays in a minute	Time	Delay in minutes (number of minutes over journey average + 5 minutes)
3	12:02	4
3	12:03	3
3	13:55	1
3	13:56	3
5	13:58	2
3	13:59	3
3	14:14	5
3	14:15	2
3	14:22	5
3	14:44	6
5	15:16	4
4	15:35	3
4	15:36	3
3	16:18	6
3	17:28	2
3	17:56	9
3	18:09	2
4	18:51	8
3	19:05	5
3	19:09	1
4	19:19	5
4	21:38	5
3	21:39	8
4	21:40	1
3	21:43	1
4	21:45	1
3	21:47	6
3	21:48	2
5	21:49	4
3	21:50	1
3	21:54	5
3	21:58	6

Table 61 shows the delay to passengers seen affecting them until 21:58, there does appear to be gaps in the data, this could be due to the low frequency seen at this time of day. However between 21:38-21:58 passengers seem to be experiencing variable delays.

5.7. Conclusion

In this Section Octopus data was obtained from the MTR cooperation. This data was sorted such that all journeys were matched and journey times were inferred from the time stamps. After considering the data available the Island Line was chosen for analysis as it contained some interesting delays and was similar to the Victoria Line in London.

When working with Oyster data to calculate the mean journey time's anomalous data was removed. However, with Octopus data, a clause was introduced when aggregating the data that stated all journeys over 120 minutes were to be removed. This meant that after studying the data it was clear it did not need data to be removed to find the mean. The means of different times of day were discovered but it was found that the time of day makes little difference to the average.

The average times to complete journeys on the Island line were compared with the MTR journey planner times. In the eastbound direction the regression line equation is $y = 4.87 + 0.98x$ and in the westbound direction the regression line equation was $y = 5.33\ +\ 0.96$.

Attention then turned to data that could simulate real-time data. In order to classify what journey times could be defined as delayed, a threshold was needed. The MTR takes the threshold of 5 minutes to classify an operational delay, it was then decided that this may be a suitable threshold for passenger delays. This threshold was then tested against the Octopus data; a number of journeys were analysed that showed less than 5% of passengers breach this threshold.

Since a moving average has been used to smooth the data in the London case, it was considered for the Octopus data. Comparing different values it again appeared that four data points should be used to calculate the average, this removed anomalies and reduced the noise, yet didn't delay the return of the data dramatically.

At this point analysis could begin to determine if congestion could be seen in the data. As seen with London a day was chosen for this analysis which contained no reported operational delays. To compare the differences with days of the week an extra day was considered in the eastbound direction. The data showed that passengers using the Island line, in general, travelled west for work and east to go home. Studying the eastbound direction showed that different days can have a dramatic effect to the amount of congestion being seen. This would be an interesting area for future research; determining if there are trends in when congestion appears over time and on different days.

In this work it was seen that when analysing a Friday in comparison to a Tuesday there was much more delay to be seen in the data. Both seemed to show a higher frequency of increased travel times in the evening peak. This is further evidence that people are living in the east and working in the west; since, over both days there was no delays to passengers in the AM peak. Further the data showed, in the evening peak, the most congested stations to enter were Central and Sheung Wan whereas Causeway Bay and Tai Koo were the most congested to exit.

Since the Friday seemed to show more congestion this day was then considered for analysis of the westbound direction. This showed many more people were travelling in this direction in the evening; however, there was not a large amount of congestion to be seen in the data. Although, the data did show that Chai Wan was a busy station to enter in the morning.

Attention then turned to discovering how passengers are affected when there are operational delays. Four days were chosen for this analysis. The first day to be analysed is the 5th of September, this had a reported delay lasting 13 minutes in the westbound direction starting at 20:29. Looking at the journey time distribution for the day in the westbound direction it was clear there was a prominent skew to the right which implied a delay was likely. However, after analysing the data no clear delay was found. This lead the analysis to look at the eastbound direction in which a delay can be seen to start around 21:00 this can be seen to be lasting till around 21:16. This suggests the delay to the last train didn't affect passengers exiting but delayed the train in changing direction, delaying passengers travelling in the other direction.

Next the 27th of September was analysed. This had a reported delay starting at 08:48 in the westbound direction lasting 5 minutes. It was decided that both directions would be analysed to see if there was an effect of the delay on both directions. In the eastbound direction it appeared there was a large amount of congestion between 18:28 and 20:03. This could be the signs of perhaps an unreported delay. In the westbound direction there is a prominent skew to be seen to the travel times to the right, this indicates a possible delay. Looking at the data there appears to be frequent delays between 08:48 and 09:12 that are likely to be caused by the delay. However, passengers seem to be only delayed one or two minutes over the five minute threshold.

The next day to be analysed was the 19th of September. On this day a delay started at 18:09 in the westbound direction lasting 6 minutes. Following the analysis on the 5th September as it was the last station it was only to affect the passengers in the eastbound direction, therefore this was the only direction analysed. During the month of September 2012 there was refurbishments taking place at Sheung Wan station, this is why there is a high frequency of delays starting at this station. These refurbishments clearly interfered with the operational running of the Island Line causing delays. For this work, the cause of delay does not affect the analysis.

The eastbound direction showed a number of passengers being delayed between 18:26 and 19:48 due to the time of day it is difficult to know what delays are caused by congestion and what is caused by operational delays. However, the higher number of passengers being delayed indicates there is some effect of the operational delay on the passengers.

Finally, the last day to be analysed is the 8th of September where a delay took place at 21:20 lasting 9 minutes in the eastbound direction. Both directions were analysed for this delay but only the eastbound direction showed a delay. Between 21:38 and 21:58 there seemed to show a delay to the passengers yet it was quite sporadic in reportings, this may be due to the day being a weekend and there being fewer passengers.

6. Discussion

Analysis of the London Underground and the Hong Kong metro networks has taken place to answer the following question: Is it possible to give passengers of a metro network real-time information? This question was broken down into three smaller questions to make answering it more manageable.

1. Is there information available about the dynamics of the network in smart card data?
2. Is it possible that this information can be extracted to be useable and reliable to passengers?
3. Is the information found useful to passengers or operators?

During Section 3 an algorithm was developed to attempt to answer the above questions. This algorithm proceeded to define the structure of analysis in the subsequent sections, London (Section 4) and Hong Kong (Section 5). In each of these sections the following subsections were included.

1. Data collection
2. Average travel times
3. Regression analysis
4. What is a delay?
5. Congestion reporting
6. Delay reporting

These subsections were created to provide a methodology that would include the following criteria:

1. Take the raw data and make it a useable format
2. Determine how quickly the information can be returned and determine operational and congestion delays
3. Smooth the data as much as possible to reduce noise and false reportings of delays should be minimal
4. Provide additional information to passengers regarding their journey and provide additional information to operators about the dynamics of the network

To determine how successful the methodology and analysis were at answering the research questions the above list will be examined to decide whether each criteria has been achieved.

6.1. Take the raw data and make it a useable format

It was decided during this project that in order to determine the dynamics of a metro network the travel times of passengers would need to be analysed. This was due to it being impossible ever to know exactly how many passengers are in one given place in the underground network at any specific time.

Once a passenger has entered the system at a ticket barrier it is unknown where they may be until they have exited the system. It could be estimated, once they have left the system, what route they took, by determining which route has the highest probability to be travelled on given the time it has taken them between entry and exit. It could then be further estimated what train a passenger could have been travelling on by analysing the train scheduling and pairing this with the predicted route with a little guess work about walking speed through the network (Guo and Wilson, 2011)(Zhao et al., 2013). This could be sufficient for post hoc analysis however, in real-time, using this as part of a model would lead to a large amount of uncertainty.

Therefore, to try and remove as much of this uncertainty as possible it was decided to aggregate the data in such a way to be left with reliable data about journeys so that the passengers’ routes could be predicted with more certainty. In the case of the model created in this thesis this involved journeys that have their origin and destination on the same line were kept for analysis and journeys that start and end on different lines were removed, first discussed in Section 4.1.

It was important at every step to try to ensure that the reporting of an incident contains as few false reports as possible. This included making the average travel times as precise as possible to ensure an accurate comparison to the real-time data.

To find the average travel times for passengers to complete journeys within the respective metros, in both cities an 8-day sample of smart card data was used to calculate the time. In London the sample spanned two months and only consisted of morning peak times, whereas in Hong Kong the sample was taken from a month and the whole day was used. The differences here were due to the availability of the data from the different companies.

In the case of the Octopus data in Hong Kong, it was investigated what difference is made to the average when being calculated at different times of day. Three origin-destination pairs were chosen to look at these differences; it was found that the different times of day made little difference to the average travel times. However, if the methodology used in this project were to be used in the future on a further network it would be concluded that the different averages should be determined and the decision made which one should be used depending on the data produced from the specific metro. This is because there are travel patterns visible to the particular lines. For example, on the Victoria Line in London passengers appear to travel south for work and north to go home. In Hong Kong it can be seen that passengers tend to travel west for work and east to go home. These trends affect the average times throughout the day by affecting passenger demand. Although in the case of the analysis completed in Hong Kong this may not make a large difference, with a different city or a different line there is definitely the potential that a difference could be seen.

This leads on to the differences found between the average times found from the smart card data and those used by the journey planners in the respective cities. Table 62 shows the equations for the regression lines found after regression analysis was completed to determine the correlations between the journey planners and the travel times found. The Adjusted R^2 values have been included to show how strong the correlation relationship is.

Table 62 – Summary of regression analysis between smart card data and journey planners

City	Direction	Equation of Regression Line	Adjusted R^2 value
London	North	y = 0.94x + 4.93	0.98
London	South	y = 0.97x + 4.49	0.98
Hong Kong	East	y = 0.98x + 4.87	0.96
Hong Kong	West	y = 0.96x + 5.33	0.97

Table 62 shows that over both cities there is a strong correlation between the journey planner times and the smart card data average times, seen by the adjusted R^2 values. It can be seen that the equations that describe the regression lines are very similar. To understand why this may be the case, Table 63 has been created. It is seen later in this section that 2 participants timed themselves taking journeys on the Southbound London Victoria line. In Hong Kong 2 passengers also timed themselves and Table 63 contains the average times it took the participants to enter, exit, travel between stations and the dwell times for the stations on the Victoria Line and the Island Line.

Table 63 – Summary of average travel times for different journey components in both cities (s)

City	Direction	Average Entrance Time	Average Dwell Time	Average Exit Time	Average Train Traveling Time
London	North	192.53	27.62	108.87	112.21
London	South	180.67	31.43	116.19	100.97
Hong Kong	East	221.08	30.08	120.74	84.55
Hong Kong	West	202.96	25.93	145.65	78.38

It would appear from looking at Table 63 that the time it takes to enter and exit the stations and the dwell times are similar over both lines, with the train travelling time taking longer in London. This is expected since the Victoria Line covers 13 miles whereas the Island line covers 8 miles. The y-intercept seen in the regression line equations could be explained by the time it takes passengers to enter and exit the stations. It is known that neither the Hong Kong nor the London journey planners include these times.

Table 64 – Comparison of average entrance and exit time with regression analysis y -intercept

Line and Direction	Average Entrance Time + Average Exit Time (minutes)	Y Intercept
Victoria Line Northbound	5.02	4.93
Victoria Line Southbound	4.95	4.49
Island Line Eastbound	5.70	4.87
Island Line Westbound	5.81	5.33

Table 64 shows the average entrance times summed with the average exit times. It can be seen that the times appear to be quite similar. Although there is not much data to be used it can be concluded that with such similar times it is likely that the time it takes for a passenger to enter or exit the system could explain the y-intercept.

Further in the process of smoothing the data. A moving average was taken to try and reduce the number of anomalies in the data. The number of data points included to make one moving average point was discussed for the case of London and Hong Kong. It was decided that in both cases a relatively low number could be used. This was due to the fact that in both examples the lines in question had high passenger demand and short train headways. However, in the case that a line should have low passenger demand and greater train headways it may not be able to take a moving average as it may be found that they spread over too large an amount of time. In this case a time dependant moving average may need to be used to ensure the information returned remains relevant.

6.2. Determine how quickly the information can be returned and determine operational and congestion delays

6.2.1. Congestion Information

In this project it was shown that information can be provided to passengers about the dynamics of the network at particular times. It was shown in Section 4.5 and 5.5 that crowding at stations can be found when passengers are entering and exiting the system.

It was made clear that as the methodology laid out in this thesis relies on using journey times to find delays it is not possible to find out information about the entrance station delays until passengers have exited the system. This means that it is unlikely that a passenger would be able to benefit from their 'own' data in relation to their current dynamics. However in the conclusion of this thesis the use of this information for succeeding passengers is discussed.

For information about congestion at the exit stations this is live information. As soon as a passenger exits the system information about their journey can be analysed and this can be used to determine if a delay has incurred. This provides insightful information to passengers about what is taking place at their end station. This information will be particularly useful for a passenger who may have two options of exit stations or when a particular exit to a station is a bottleneck and likely to cause delays this can provide information to passengers within the system to allow them to adapt their route.

In Hong Kong there was a lot of variability in the information provided by the algorithm about congestion. However when the information was available it was clear in the evening there was thorough information about the dynamics of the network and the congestion during the evening peak time.

6.2.2. Delay Information

Through the analysis of the smart card data it has been shown in both cities that it is possible to discover information about operational delays through the data. The algorithm described in Section 4.5.1 showed a method for determining when a delay should be classified. Table 65 shows the results of this algorithm, showing when the first delays were reported during an operational delay. It can be seen that the last row is left blank. This result was taken Table 57, where it can be seen that there is a reported delay at 08:48. However, there appears to be no immediate delays either before or after this event showing that there is no information linked to the operational delay.

Table 65 – Summary of initial delay reportings in comparison to operational delay reportings

City	Date	Time operational delay is recorded	Time algorithm registers delay	Time Difference
London	02/10/2012	08:40	08:48	+ 8
London	04/10/2012	08:20	08:19	-1
London	26/10/2012	07:30	07:35	+5
Hong Kong	05/09/2013	20:29	21:00	+31
Hong Kong	08/09/2013	21:20	21:38	+18
Hong Kong	19/09/2013	18:09	18:26	+17
Hong Kong	27/09/2013	08:48	-	-

Table 65 – Summary of initial delay reportings in comparison to operational delay reportings shows a great variety in the speed of the return of information. Comparing the two cities it is quite clear that the return of information is quicker in London.

In Hong Kong the wait for the information can be up to 31 minutes. This does not provide passengers of the metro sufficient information about the current dynamics of their journeys. In both networks the longest journey times are found to be around 35 minutes and the shortest journeys are found to be 5 minutes. For two un-delayed days the average journey length was found to be in both cities 15 minutes. The results seen in the table above would lead to the conclusion that there are passengers completing longer journeys in this network. However, during the late evening delays on the 08/09/2013 and 19/09/2013 although there are passengers completing short journeys it appears their journeys do not have delays. This suggests that the delay to the network accumulates over the length of longer journeys to delay passengers over 5 minutes, therefore progressing with Octopus data would lead to this being a consideration.

6.3. Smooth the data as much as possible to reduce noise and false reportings of delays should be minimal

It was seen in Section 2.1.2 that it is important to provide passengers with reliable, consistent information. During the development of the methodology in Section 3 it was decided that this could be achieved by taking measures to try and remove noise from the data that could cause false reportings of delays. The hope was that if the data was sufficiently smoothed then consistent results could be delivered to passengers and operators.

While working with the Oyster data in Section 4 a number of different processes took place to remove unwanted anomalies from the data. Firstly it was seen in Section 4.2 that ambiguous journeys were removed; these are journeys for which the route is unknown as their origin and destination are on different lines Table 66 and Table 67 show examples of an ambiguous and an unambiguous journey. Firstly Table 66 shows the journey from Liverpool Street Station to Waterloo. This is a journey that requires an interchange and there are a number of possible routes for passengers to choose, whereas Table 67 shows a journey from Bank to Marble Arch which can be completed on one metro line. The data in these tables comes from a file that contains all journeys completed on the 5th March 2012 in the London Underground.

Table 66 – Averages of Liverpool Street to Waterloo

Journey 1: Liverpool Street - Waterloo	
mean	18.13
range	9 to 49
mode	14.00
median	16.00
standard deviation	6.81

Table 67 – Averages of Bank to Marble Arch

Journey 2: Bank - Marble Arch	
mean	16.31
range	12 to 38
mode	14.00
median	16.00
standard deviation	3.17

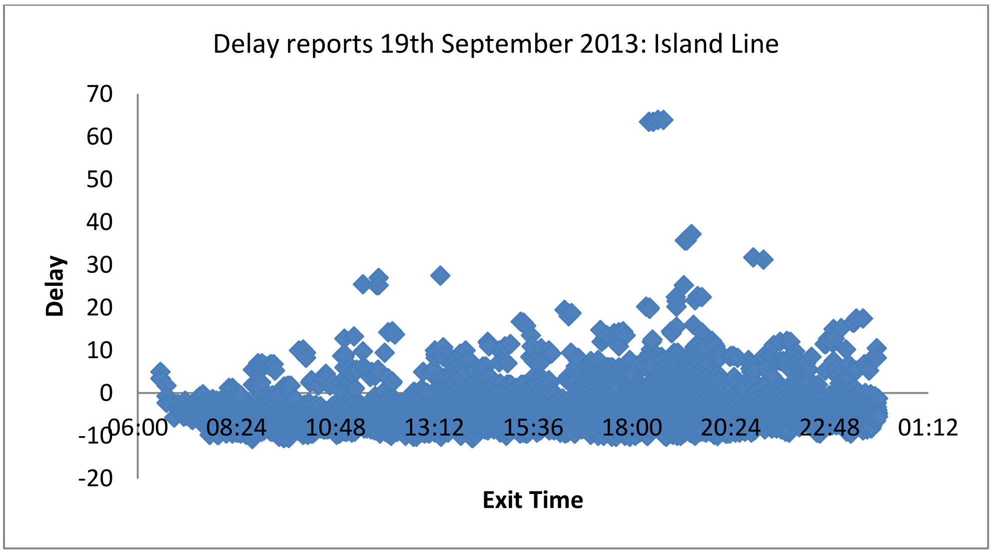
First it can be seen Journey 1 has a much larger range of passenger travel times that Journey 2. Journey 2 also has its mean, mode and median much closer in values. However, the most interesting result would be that there is quite a substantial difference between the two standard deviations. A smaller standard deviation shows that the data is clustered more closely around the mean and therefore is more reliable for analytical purposes.

In the smoothing process in the development of the algorithm in London, it was seen in Section 4.2 that anomalies were removed from the data to ensure a more accurate average travel time. However, with the Octopus data it was seen that this step was not necessary since very extreme travel times had already been removed when aggregating the data. Therefore in conclusion to this step it should be decided if removal of anomalies is necessary depending on the smart card data used and considering how many large anomalies appear in the data.

At this point the 'real-time' data was smoothed to ensure reliable data but also to provide continuous data that passengers can always access. The decision of how many data points (passenger journeys) should contribute to the moving average was decided by aiming to remove anomalies but not to allow large spacing between new data points. For the necessity of providing continuous data in both cities the moving average was decided on depending on the data – in any new city it would be necessary to consider this question according to the available data.

The last step in attempting to smooth the data was to remove anomalies and false reports was to add a constraint that for a delay to be reported, a certain number of moving average point needed to be delayed in the same minute.

It was essential to add an additional constraint in the delay reporting as there were some passengers whose journey times were over the delay threshold yet there was not a delay at the time. Graph 116 shows all journeys completed on the Island line on the 19th of September. This was one of the delayed days seen in Section 5.6.3. The graph shows their delay status plotted against the exit times of the journey. The y-axis shows the delay status in minutes over the threshold; the origin is the average travel time for that journey plus the additional 5 minute threshold. This means any positive value is the number of minutes delayed that journey is over the threshold. It should be expected that most journeys should be under the 0 mark in this case, apart from rush hour and the delay at 18:09. However, it can be seen throughout the delay there are passengers whose journeys are above the 0 mark.



Graph 116 - Delay reports 19th September 2013: Island Line

It can be seen that throughout the day there are passengers that are classified as delayed. When there is a clustering of delayed passengers it is likely the passengers are experiencing genuine delays. However when there are one off delayed passengers it is likely that this is due to another factor other than congestion, such as, elderly, passengers moving in groups or passengers with mobility issues. It was necessary to add an additional clause to try and only report real delays rather than passengers that may be slow.

The graph show a density of delays around the reported operational delay or rush hour showing there is useful information in the data that needs to be extracted.

The more moving average points that showed a delay in a minute showed a greater likelihood that there was in fact a delay to the network. However this also meant that there could be a delay in the return of the information about delays.

The decision was made that 5 moving average points needed to be delayed within the same minute to classify a delay. This to ensure the delay classification was not too sensitive so that only genuine delays to passengers are reported.

The disadvantage to counting delayed moving average points to determine a delay is that it does not consider that if frequency is low this threshold might not be able to be reached. This indicates that during a delay that may occur off peak there may not be sufficient information for this constraint. However as discussed in the next section, this information may benefit from additional information provided by operators to assist in the classification of a delay. This would mean that during an off peak delay the data could be analysed rather than waiting for the constraint to be breached.

6.4. Provide additional information to passengers regarding their journey and provide additional information to operators about the dynamics of the network

This section is broken down into the information that can be provided to operators and passengers as either static information or potential real-time information.

6.4.1. Static

It was seen in Table 65 that there is a delay in some of the information gained about the system dynamics. Congestion information about entrance delays can only be discovered once a passenger has finished their journey and exited the system. This is unhelpful for passengers about to enter the system as is does not provide any useful information about the current dynamics. However this is not redundant information. This information can be used alongside a journey planner to help inform passengers about the expected network dynamics.

Tourists travelling in rush hour are a serious problem. For example, in London, TfL use adjusted fares to discourage unnecessary travel during peak times ("Adult rate Tube, DLR and London Overground fares," n.d.). The results found in Section 4.5, showing the congested stations during peak times in the London Underground and Section 5.5 in the Hong Kong metro can be used to inform passengers through a journey planner of what times are expected to have high congestion. This information can allow passengers the opportunity to reroute or change their departure time when it is expected that there is likely to be high congestion in a particular station.

To be able to provide this information to passengers a journey planner needs to be created that can take the average times found for journeys on the underground lines in question and turn these into a service that passengers can use to determine how long their journey may take them. The complication to this process will be to create a successfully journey planner, it is important that all possible journeys in the underground are provided with information. That means that regardless of how many interchanges a passenger may need or wish to take, information can be provided. This will need to be considered when the journey planner is created.

There are two ways information can be provided to passengers; all passengers collectively can access the same information, or they can do so as individuals. As an example how passengers may be able to use information from smart card data a static journey planner will be created. It was shown in Section 4.3 that there is a strong relationship between the journey planner times and the times found through Oyster data. Part of the aim of this thesis is to see if it is possible to provide better information to passengers and operators. With a new database of passenger times created attention will turn to see if it is possible to create a more accurate static journey planner, London oyster data will be used in this example.

6.4.1.1. London

Using a journey planner, a passenger could enter the origin and destination of the journey they want to complete and have information about their journey time returned to them. It was discussed earlier that only journeys that have their origin and destination on the Victoria Line will be used for this project. However, there are a large number of passengers that make journeys that cross over different lines, therefore it is important to take the information for single line journeys and use it to for all journeys that may use part of that line, so that all passengers of the London Underground are provided with information. In order to be able to take the information found about journeys that have their origin and destination on the same line, and use it for journeys that contain an interchange, the intention is to discover where passengers are spending their time between entering and leaving the system. To discover this, a system of simultaneous equations was set up to break down the different components of a journey. There are 16 stations on the Victoria line, so the maximum number of possible journeys for which there could be data for is 120 journeys (j), this is calculated from

$$j = \frac{n(n-1)}{2}$$

Equation 17

when $n = 16$.

A database was created to store information including where a passenger could be in the system between entering and exiting the ticket barriers. The places where a passenger could be in the system include:

a. walking from ticket barrier to platform (including waiting for a train)
b. the train travel times between one station and another
c. the waiting time at the intermediate stations (dwell time) and
d. the exit from a station.

There are 15 possible entrance stations and 15 possible exit stations on the Victoria Line – you cannot enter the last station or exit from the first. There are 15 train travelling times and 14 intermediate stations for dwell times. This gives 59 unknown times within the system of where a passenger could be, shown in Table 68.

Table 68 – List of unknown variables

Unknown variable name	Type of Variable	Location/Location start	Location end
x1	entrance	Walthamstow Central	
x2	line	Walthamstow Central	Blackhorse Road
x3	exit	Blackhorse Road	
x4	dwell time	Blackhorse Road	
x5	entrance	Blackhorse Road	
x6	line	Blackhorse Road	Tottenham Hale
x7	exit	Tottenham Hale	
x8	dwell time	Tottenham Hale	
x9	entrance	Tottenham Hale	
x10	line	Tottenham Hale	Seven Sisters
x11	exit	Seven Sisters	
x12	dwell time	Seven Sisters	
x13	entrance	Seven Sisters	
x14	line	Seven Sisters	Finsbury Park
x15	exit	Finsbury Park	
x16	dwell time	Finsbury Park	
x17	entrance	Finsbury Park	
x18	line	Finsbury Park	Highbury & Islington
x19	exit	Highbury & Islington	
x20	dwell time	Highbury & Islington	
x21	entrance	Highbury & Islington	
x22	line	Highbury & Islington	Kings Cross
x23	exit	Kings Cross	
x24	dwell time	Kings Cross	
x25	entrance	Kings Cross	
x26	line	Kings Cross	Euston
x27	exit	Euston	
x28	dwell time	Euston	
x29	entrance	Euston	
x30	line	Euston	Warren Street
x31	exit	Warren Street	
x32	dwell time	Warren Street	
x33	entrance	Warren Street	
x34	line	Warren Street	Oxford Circus

x35	exit	Oxford Circus	
x36	dwell time	Oxford Circus	
x37	entrance	Oxford Circus	
x38	line	Oxford Circus	Green Park
x39	exit	Green Park	
x40	dwell time	Green Park	
x41	entrance	Green Park	
x42	line	Green Park	Victoria
x43	exit	Victoria	
x44	dwell time	Victoria	
x45	entrance	Victoria	
x46	line	Victoria	Pimlico
x47	exit	Pimlico	
x48	dwell time	Pimlico	
x49	entrance	Pimlico	
x50	line	Pimlico	Vauxhall
x51	exit	Vauxhall	
x52	dwell time	Vauxhall	
x53	entrance	Vauxhall	
x54	line	Vauxhall	Stockwell
x55	exit	Stockwell	
x56	dwell time	Stockwell	
x57	entrance	Stockwell	
x58	line	Stockwell	Brixton
x59	exit	Brixton	

These values were then used to make a matrix *A* where the rows of *A* are the 120 possible journeys and the columns are the unknowns. Then a cell

$$a_{ij} = \begin{cases} 1\ if\ unkown\ j\ is\ included\ in\ the\ ith\ journey \\ 0\ else \end{cases}$$

Equation 18

A system of equations was then set up to determine a solution to the unknowns:

$$Ax + \varepsilon = b$$

Equation 19

where *A* is the matrix described above and *x* is a vector of the unknowns. *b* is a vector that contains all the predicted mean travel times for the Southbound Victoria Line journeys, shown in seconds and ε is an error term added for the variation of travel times.

This system is an inhomogeneous, singular, over-determined system with the equations being linearly independent. This means that there is not a solution. However, this can be solved using the method of least squares, this method approximates a solution.

This is done by taking the equation

$$Ax = b$$

Equation 20

and multiplying both sides of Equation 20 by the transpose of *A* let's call this *A'* so

$$(A'A)x = A'b$$

Equation 21

This gives a standard square system of linear equations which can be solved.

Equation 21 was solved in Matlab R2012b using the command *lsqlin.* This command solves the system of equations using least squares of the form

$$\min_{x} \frac{1}{2} \|A \cdot x - b\|_2^2$$

Equation 22

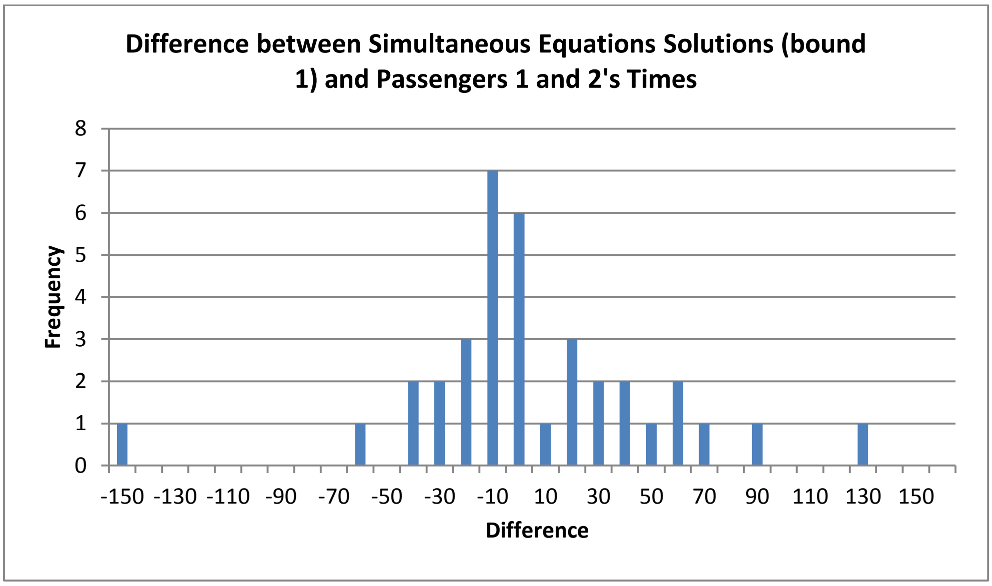
However when the system was first solved it appeared to have a number of results that were zero, implying a travel time of 0 seconds. This makes no sense in the context of the underground that it could take 0 seconds to travel, therefore a bound needs to be introduced to remove these results.

There are infinitely many possible approximate solutions to the simultaneous equations; therefore it is necessary to introduce a bound on the solution so that not only are the solutions all positive, but they also have realistic positive values. Different bounds were introduced to determine what would give the most appropriate solution.

To know which bound would give the most accurate result, it was necessary to know how long it would actually take passengers to travel on the relevant journey in the underground. For preliminary results two passengers took journeys as a pilot study. Timing the length of time it would take to do particular parts of the Southbound Victoria Line all intermediate times were recorded and the overall journey time. The simultaneous equations' solutions were compared to each passenger's completed journey times. Not all 56 unknown times were covered by passengers 1 and 2 but the unknowns they did complete were compared, discussed below.

An iterative process was used to determine what bound would give the best solution. This process started with taking a generic first bound of greater than 0.5 seconds, to make the solutions positive, the difference between the pilot's passengers' times and the solutions were compared and a new bound introduced to minimise total difference between the passengers and the solution. After 15 iterations it was found that a bound of >10 seconds gave the most accurate, in terms of the smallest overall difference between passengers 1 and 2's results and solutions to the equation. The process was stopped after 15 iterations since no improvement was found to be made on the difference after 15 iterations.

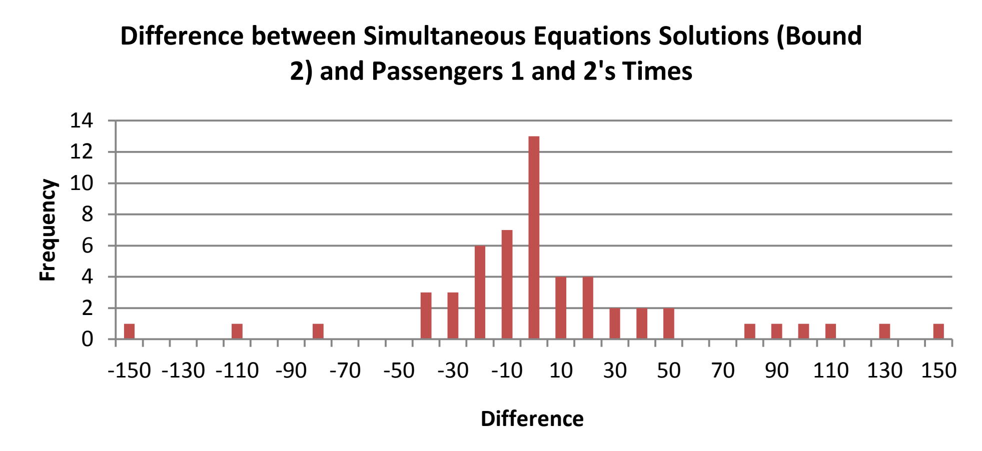
This bound gave a normal distribution of results centred close to 0, with a mean of 0.26 and a standard deviation of 40. Graph 117 shows the difference between the passengers’ travel times and the solution’s calculated travel times. For simplicity we shall call the bound of all results being greater than 10 seconds Bound 1.



Graph 117 - Difference between Simultaneous Equations Solutions (bound 1) and Passengers 1 and 2's Times

Beyond this a second bound was introduced on the individual components to gain a smaller difference between passenger times and the solution. Here the bound of minimum entrance and exit time of 60 seconds, a minimum line time of 90 seconds and a minimum dwell time of 20 seconds was created based on the times the participants took, these will be called Bound 2. These new results were then compared with journey times found by passengers 1 and 2.

In Graph 117 it can be seen that the different bounds do not change the difference between the computed overall journey times and the different passenger travel times and again the solution gives a normal distribution of results centred close to 0.



Graph 118 - Difference between Simultaneous Equations Solutions (Bound 2) and Passengers 1 and 2's Times

To determine further the difference that the bounds make to the solution, the passengers' journeys were analysed in more detail. Passenger 1's journeys were taken on the 2nd May 2013, and there were no reported delays during the afternoon. However the was a problem at the exit of Green Park with one of the escalators being broken. Table 69 shows the time in seconds to complete different journeys, the solutions found, the difference between the two and the average difference in the different components of the journey.

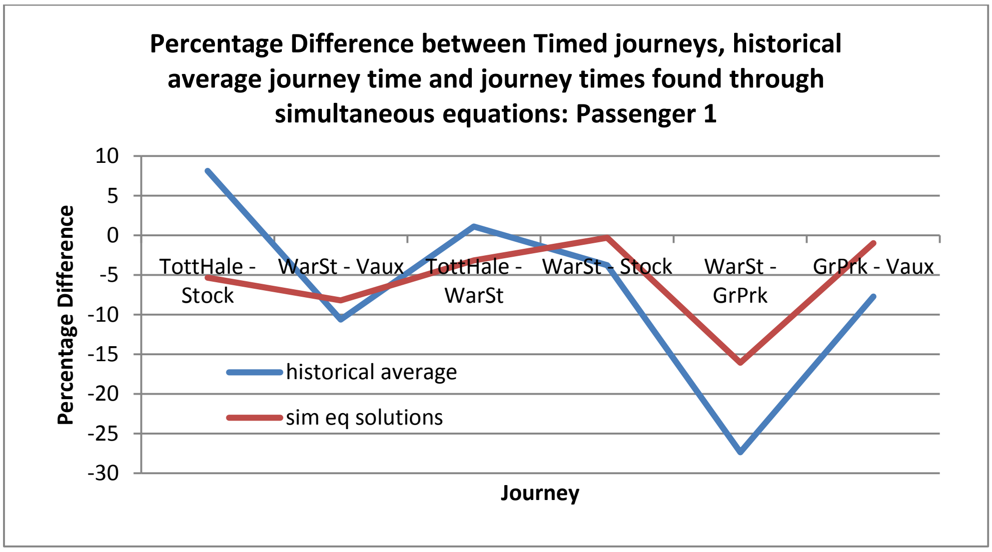
Table 69 – Bound 1 times in comparison to passenger 1's journeys (s)

	Tottenham Hale - Stockwell	Warren Street – Vauxhall	Tottenham Hale – Warren Street	Warren Street - Stockwell	Warren Street – Green Park	Green Park - Vauxhall
Timed Travel Time (TT)	1860	863	1227	996	535	517
Solution Time (Bound 1 (B1))	1831	798	1208	993	461	512
Mean historical Time (MHT)	1920	780	1260	960	420	480
Difference (TT-B1)	-90	59	-27	-3	74	-1
Difference (B1-MHT)	-89	18	-52	33	41	32
Difference (TT-MHT)	-60	83	-33	36	115	37
Average Difference of individual journey components bound 1	-7.5	3.18	-4.15	-6.85	13	-0.25

Table 69 shows the range of differences is -90 seconds to +74 seconds. Given that this model includes waiting times for trains this is a small margin of error for the results as these differences in time can be put down to waiting for trains, differences in dwell times and different walking speeds.

There is an anomaly with the journey from Warren Street to Green Park as the average difference of individual components is a lot higher. This is due to the escalator being broken at the exit of Green Park. This delay can also be seen as the difference between the mean passenger time and the passenger's travel time is nearly 2 minutes. Other than that journey it would appear that passenger 1's travel times are fairly close to the average time found.

Graph 119 shows the percentage difference between the time taken for passenger 1 to take the different journeys against the average travel times and the solution to the simultaneous equations, here we can see that there is an anomaly in the time it took passenger 1 to complete the journey between Warren Street and Green Park.



Graph 119 - Percentage Difference between Timed journeys, historical average journey time and journey times found through simultaneous equations: Passenger 1

Next, Bound 2 was analysed to see if this bound would change the results, the table below shows the journey times completed by Passenger 1, the journey times calculated by the different bounds and the difference between these results.

Further, it shows the average difference between Bound 1 and Bound 2 in the different components of the journeys for Passenger 1, all in seconds.

Table 70 - Bound 2 times in comparison to passenger 1's journeys

	Tottenham Hale-Stockwell	Warren Street-Vauxhall	Tottenham Hale-Warren Street	Warren Street-Stockwell	Warren Street-Green Park	Green Park-Vauxhall
Passenger 1	1929	863	1246	996	535	517
Bound 1	1925.23	797.64	1236.55	992.97	460.92	511.92
Bound 2	1925.23	797.64	1236.55	992.97	460.92	511.92
diff (Passenger 1 - Bound 1)	3.77	65.36	9.45	3.03	74.08	5.08
diff (Passenger 1 – Bound 2)	3.77	65.36	9.45	3.03	74.08	5.08
Average Difference of individual journey components for Bound 2	0.15	5.94	0.73	0.23	14.82	0.73

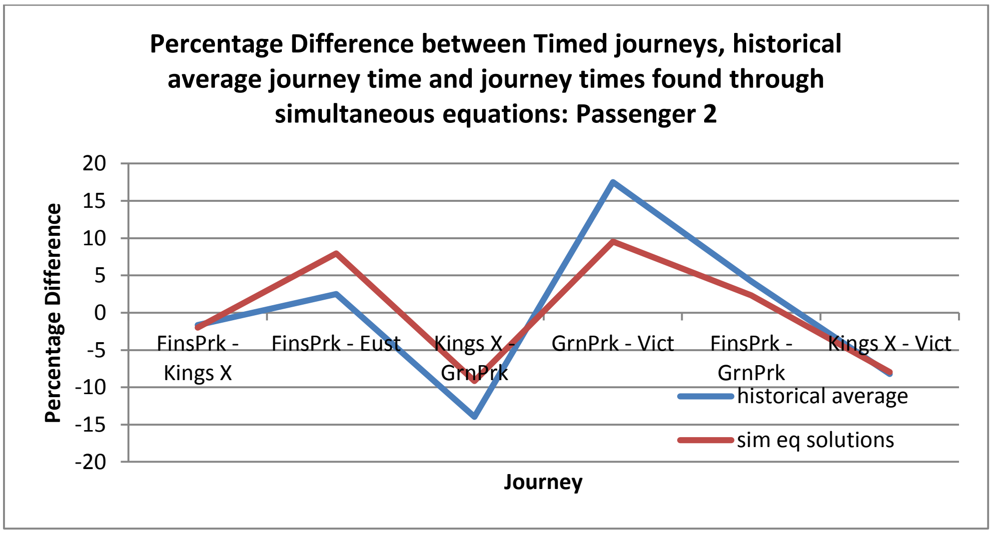
Table 70 shows the different bounds do not change the overall journey time computed by the model. This shows that the solution, regardless of bounds, has minimised the error. Yet what can also be seen is the average difference between the individual component timed by Passenger 1 and those calculated with the Bound 2 are closer, than seen in Table 69.

Passenger 2's journeys were taken on the 13th May 2013, and there were no reported delays during the morning. Below Table 70 shows the time for Passenger 2 to complete the journeys, the solution found with Bound 1, the difference between the two and the average difference in the different components of the journey between passenger 2 and Bound 1, all in seconds.

Table 71 - Bound 1 times in comparison to passenger 2's journeys

	Finsbury Park - King Cross	Finsbury Park - Euston	Kings Cross - Green Park	Green Park - Victoria	Finsbury Park - Green Park	Kings Cross - Victoria
Timed Travel Time (TT)	671	702	752	297	977	909
Solution Time (Bound 1 (B1))	658	763	689	328	1000	842
Mean historical Time (MHT)	660	720	660	360	1020	840
Difference (TT-B1)	13	-61	63	-31	-23	67
Difference (B1-MHT)	-2	43	29	-32	-20	2
Difference (TT-MHT)	11	-18	92	-63	-43	69
Average Difference of individual journey components	2.61	-8.67	7.00	-10.45	-1.79	6.08

Table 70 shows the range of difference between the solution time and the passenger's travel time is -61 seconds and +67 seconds. Again the percentage difference between timed journeys, the average journey time and journey times found through simultaneous equations with Bound 1 were drawn graphically.



Graph 120 - Percentage Difference between Timed journeys, historical average journey time and journey times found through simultaneous equations: Passenger 2

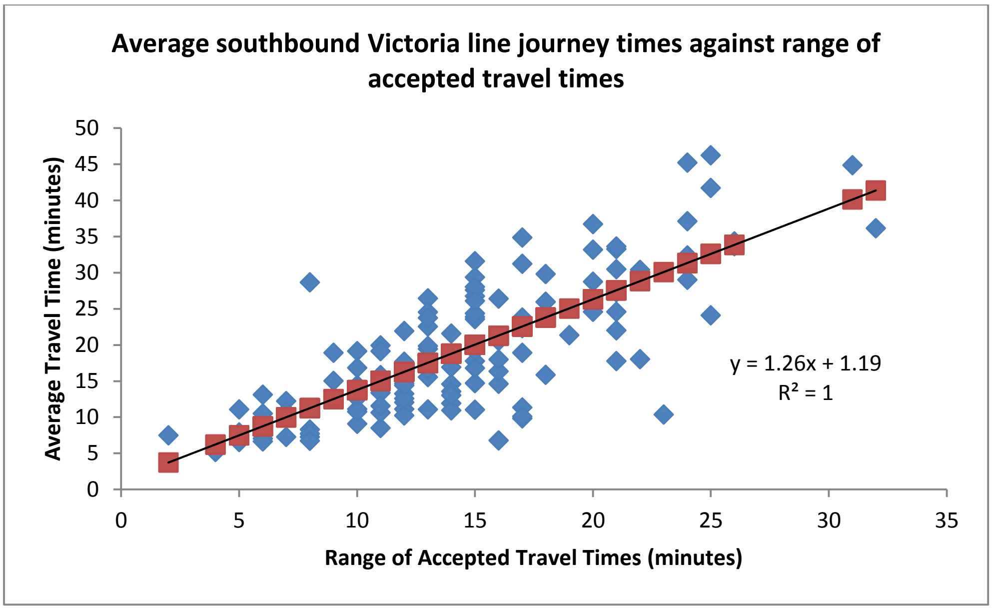
Graph 120 shows that there is a slight rise in times for the journey between Green Park and Victoria but there is no information from Passenger 2 about this journey so it is unknown if there was a delay or not. Next Bound 2 was analysed to see if it changes the result to make them more accurate.

Table 72 - Bound 2 times in comparison to passenger 2's journeys

	Finsbury Park – King Cross	Finsbury Park - Euston	Kings Cross – Green Park	Green Park-Victoria	Finsbury Park-Green Park	Kings Cross-Victoria
Passenger 2	671	702	752	297	977	909
Bound 1	657.93	762.69	688.96	328.36	1000.22	842.12
Bound 2	657.93	762.69	688.96	328.36	1000.22	842.12
diff(Passenger 2-Bound 1)	13.07	-60.69	63.04	-31.36	-23.22	66.88
diff(Passenger 2-Bound 2)	13.07	-60.69	63.04	-31.36	-23.22	66.88
Average Difference of individual journey components for Bound 2	2.61	-8.67	7.00	-7.84	-1.79	6.08

As with the case for Passenger 1, the solutions do not change the overall travel time. In this case it makes little difference to the individual component differences but does reduce the difference for the journey of Green Park – Victoria.

It is clear though with both bounds introduced that the range of accuracy for the different parts of the Victoria line spans a large difference in travel times, with over ±2 minutes (approximately one headway), with the greatest differences in times being found in the entrances and exits. These differences are due to the variation in passengers' walking speeds, choices (e.g. either passengers could choose to walk up and down the escalators or stand on them) this choice could determine if they make the first train or not, explaining a two minute difference. When originally calculating the average travel times two standard deviations above the mean were removed from the data set. Graph 121 shows how the range in accepted travel times increases as the overall journey time increases.



Graph 121 - Average southbound Victoria line journey times against range of accepted travel times

Graph 121 shows that the range in times for passengers to complete journeys is large. Therefore, given this, the initial results found with Passengers 1 and 2 are very helpful in that the model results are close to the actual travel time for a passenger in the underground. This leaves the model in a good position for an initial static journey planner, which could hopefully be expanded on in future work.

6.4.2. Real Time

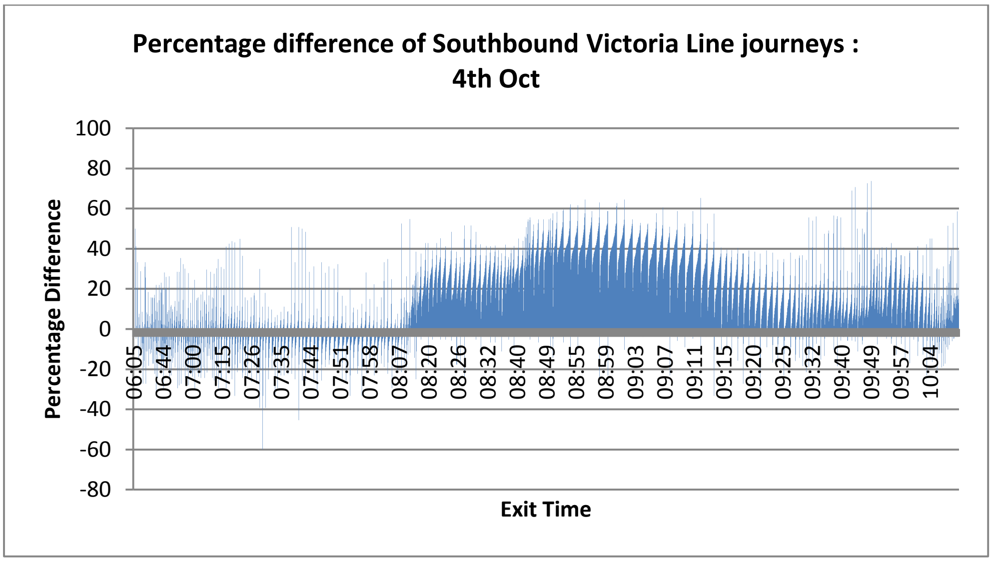
What static information has been made available to passengers through this project has been discussed, therefore it can be analysed what real-time information has been discovered, since the initial research question was: Is it possible to give passengers of a metro network real-time information?

During this project four types of delays have been discussed: operational delays, entrance delays, exit delays and unexpected delays.

It was found when analysing peak time congestion in the morning rush hour in London (Section 4.5) that it was possible to see exit delays in real time as the information is discovered as the passenger exits the system. This information will be very valuable to passengers and operators when it comes to bottleneck stations. For example, when a lengthy queues form, if exit information can be provided to passengers this will allow passengers to make the decision to either exit at a different station or perhaps depart at a later time.

The same analysis in Hong Kong showed that the algorithm was capable of providing information about exit delays, yet the information was patchy and infrequent, suggesting that many fewer passengers were experiencing delays to their journey from exit queues. This provides additional information about the layout of the stations in Hong Kong; that there is much more space for passengers to queue and many more ticket barriers speeding up the process than is the case in London.

Other than exit information, real-time information was discovered to be available about unreported delays and reported operational delays. It was seen in Section 4.6.2 that an unreported delay was seen to occur on the morning of the 4th October 2012. Graph 122 shows the delay can be seen around 06:30 in the morning and the results from the algorithm show the increase to passengers travel times experienced during that time.



Graph 122 - Percentage difference of Southbound Victoria Line journeys : 4th Oct – Second apperance

It was also clear in Hong Kong that unexpected delays were visible. For example Table 57 shows an unusual increase in passengers' travel times in the evening peak. This is a promising result of this project because it shows that it is possible to discover information about the dynamics of the network at different times of day and in different cities and this can be transformed into information to be provided to passengers. This information can also inform operators of situations they are unaware of so that they can look into the cause of the problem and help to solve it.

Finally, real-time information about operational delays has been discovered through this project. In both cities it is clear that there can be delays to the information provided about the operational delays. Once the information is received it can be converted into information that can be delivered to passengers that can inform them of how long the operational delay may affect their journey. It is clear through this work, however, that this should not be the only source of information about delays that informs a decision about the imposition of a delay status.

7. Conclusion

This project has aimed to answer the main research question: Is it possible to give passengers of a metro network real-time information? This question was broken down into three sub questions:

1. Is there information available about the dynamics of the network in smart card data?
2. Is it possible that this information can be extracted to be useable and reliable to passengers?
3. Is the information found useful to passengers or operators?

In order to conclude whether this thesis has been able to answer the main research question asked, each sub question will be examined and a conclusion will be drawn as to how well this thesis has achieved an answer. Future research will be discussed as well as the future prospects for these findings and from this a conclusion will be drawn to answer the main research question. This Section will be broken into three sub sections; each section focusing on one of the three sub research questions.

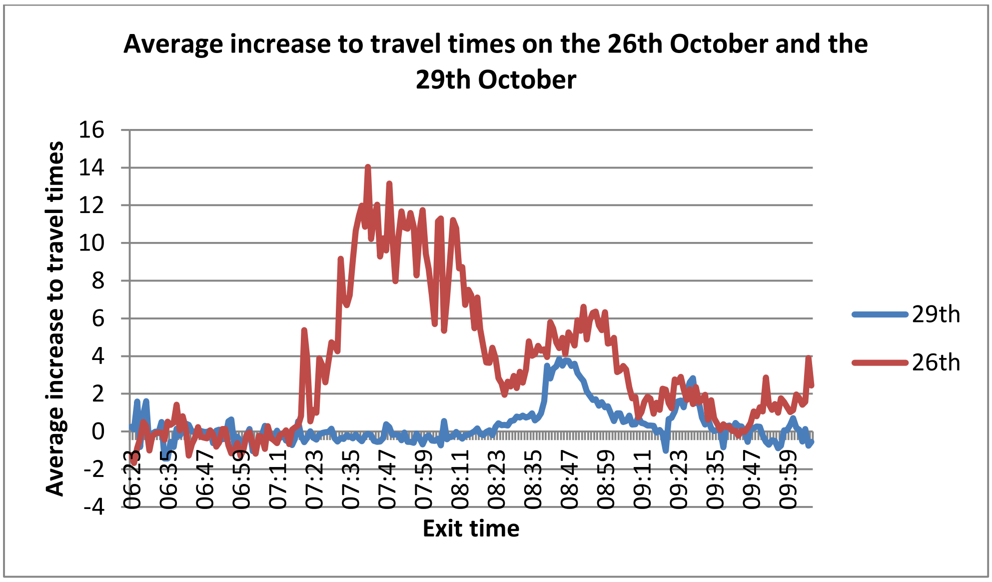
7.1. Is there information available about the dynamics of the network in smart card data?

During this thesis extensive analysis was completed on two data sets from smart card ticketing systems in different cities. The idea to examine two sets was to be able to conclude about the availability of information depending on the city. This analysis of the data has led to previously unknown information about two metros becoming available.

This work has provided in depth information about hotspots for congestion at certain times. It was clear that if a passenger entered a station at a given time in the morning the probability of their journey being delayed was much higher. Further this could be matched to exit stations; that entering at a certain time would cause you to leave at a certain time and in both cases a delay would be incurred. This leads to valuable information about the morning rush hour dynamics, about when there is likely to be full trains and busy stations, where queues may be forming and when it may not be possible for passengers to board the first train.

The algorithm used to determine when congestion was forming then looked for operational delays. With delays to the train services, passengers journeys were delayed which led in some cases to passengers exiting the system much later than expected. This developed into delays in the return of information about the network dynamics during the initial period of the delay. However, although the information in some cases may have been delayed, new information about the operational delays emerged that gave insight into how an operational delay may affect the passengers in the system. This new information provides knowledge about the dynamics of the network during a delay in terms of the length in time the delay may last, how large passenger demand may affect it and the increase to passenger times.

A clear example of the some of the information made available through smart card data is seen in Graph 123. This graph was first seen in Section 4.6.1 as Graph 51.



Graph 123 - Average increase to travel times on the 26th October and the 29th October – Second apperance

This graph highlights the newly discovered information available about the London Underground Victoria Line; about how passengers travel times are effected during delays and how rush hour demand greatly contributes to passenger delays during operational delays. The similarity in the trends of the graphs shows how the demand is similar on both days, regardless of the delay status. Further, the drop seen to passengers' times on the 26th at approximately 08:25 shows how the operational delay had begun to correct itself, then high passenger demand worsened the problem.

In conclusion smart card data is a rich source of information about the dynamics of a network. It was seen however more information was available about the network dynamics in London than Hong Kong, this may be to do with the network layout itself.

7.2. Is it possible that this information can be extracted to be useable and reliable to passengers?

This thesis saw the development of a methodology that aimed to:

1. Take the raw data and make it a useable format
2. Determine how quickly the information can be returned and determine operational and congestion delays
3. Smooth the data as much as possible to reduce noise and false reportings of delays should be minimal
4. Provide additional information to passengers regarding their journey and provide additional information to operators about the dynamics of the network

The aim of having these criteria was to ensure the return of the information would be useable to passengers and as seen in the literature review in Section 2.1.2 for passengers to trust the information it needs to be reliable.

In London it was seen that the delay reporting was fairly consistent with information arriving about the delay every minute with the exception of one or two minutes missing. This constant information source provides reliable information to passengers about the current dynamics of the network. However, in Hong Kong it was seen that there were multiple anomalies still in the data at the end of the algorithm. This was due to the lack of information being found about the network dynamics. When the threshold of the number of delayed moving average points in a minute was low, information could be discovered about congestion and operational delays. This low threshold however allowed anomalous data through into the final results.

This leads to the conclusion and future work that if an algorithm is to provide information about the network dynamics and reliable information for passengers perhaps there needs to be some fluidity in the algorithm. The case may be that depending on the network an algorithm needs to be designed to fit the nature of the data. If work on Octopus data were to continue the algorithm would need to be flexible to the information arriving. For example, if there were a low number of passengers, sporadically being delayed throughout the day, the threshold should remain at a higher number of delays in the same minutes so that anomalous delays are not reported. However, to gain the maximum amount of information about a delay occurring to passengers, once a delay status has been triggered perhaps the threshold could drop to a lower number of delays in the same minute to keep the information consistent and reliable.

The return of the information about congestion was also found to be patchy in some places; this again could lead to unreliable information for passengers. However, in this case to ensure the information is reliable, this information could be teamed with historical data to provide dependable information. A possibility for future research could be to discover trends of congestion occurring over time. This work could be completed by developing a database of congestion activity in a network over a long time period. This could lead to forecasting of congestion given initial conditions by using probabilities to determine the likelihood of congestion occurring in the next few minutes.

It was also seen that in some cases there was a delay in the return of information about the operational delays. It is crucial that passengers are provided information about delays as early as possible to ensure that the delay is not worsened by high passenger demand; prompt information means that passengers have the opportunity to reroute right at the beginning of the delay. This suggests other data sources may be required in some cases to provide information to passengers when there are gaps in the smart card data.

In conclusion, this methodology has removed a large amount of noise from the data, for example by taking only unambiguous journeys and smoothing the data which has led to more reliable information. However, before this information could be given to passengers of a network small changes to the algorithm, that would be specific to each city, would be necessary to ensure the most reliable information is returned.

7.3. Is the information found useful to passengers or operators?

The hope for this project was to discover information that could be found useful for passengers and operators of the network. For the passengers, the aim was that with real-time information they would be able to reroute themselves when there was either congestion or operational delays to the network. For the operators the hope would be that this information could provide them a thorough understanding of what is happening in their network at any particular time. This section will look into determining what useful information has been discovered.

In Section 4.4.1 it was seen that the sample of passengers of the London Underground wanted better information provided on their phones and on the service boards but also 2/3 wanted information about congestion and 96% of those who did want the information felt it may make them change their behaviour. In this project information was discovered about when and where congestion is in the network. This information can be used by passengers to determine what may be the optimal time for them to start their journey and what route they should take. This information also allows passengers to make the choice of rerouting away from congested stations. This information will be particularly useful in the London Underground where there are many stations that have elevators serving the exits. This means at busy times these exits can become very crowded and it can take some time for the queues to clear. At these times information about exit congestion would be extremely useful for passengers, if the information could be provided to the train drivers they could inform the passengers and allow them the option of going to other nearby stations. This idea can be extended to large one off events, for example, New Year's Eve, with this new information about congestion operators can now know when exits are getting overcrowded and becoming increasingly dangerous for passengers safety. The can then inform the driver who could either advise the passengers to reroute or in extreme cases tell the driver not to stop at the station.

Operators can also use congestion information to discover where bottlenecks are in the network. It was seen during this thesis that more passengers' journeys were affected during rush hour in London than Hong Kong. With similar frequency in trains and similar passenger demands there must be differences with the networks. This leads to the conclusion that there are more bottlenecks and a smaller infrastructure in London than Hong Kong. This research could be taken forward to being expanded to the entire network to identify 'hot spots' for congestion, which could provide valuable information to the operators about where expansion is needed within the network.

This project also found useful information about how long passengers may be delayed during operational delays. This information makes it easier for passengers to make the decision to reroute or not by determining how much their journey may be affected and how the delayed journey compares to their other possible options. Operators can use this information to gain a better understanding of how an operational delay affects their passengers. This information can be used by the operators to help passengers move away from the problem and them to try and decrease demand to the line to ease the delay. This new information also provides the operators with a better understanding of the length of the delay beyond the operational problem being solved.

In conclusion this project set out to determine if it was possible to provide information to passengers about the real-time dynamics of a metro network through mining smart card data.

An algorithm was created to take the raw smart card data in London and produce real-time information. This algorithm successfully discovered entrance, exit and line delay that were either reported or unreported. This methodology was expanded to produce the same results to an entirely different city, successfully showing that it is possible to provide passengers of a metro network with real-time information, however further information about the network and fluidity in the algorithm may be needed to maximise the amount of information returned. As aimed this new information will allow passengers to optimise the current network and reduce delays rerouting to underutilised routes.

7.4. Future Research

In the future, work could continue by expanding the analysis to an entire network. Within this project, in both cities metro lines were chosen that contain no splits or loops to the track, however the next step may include more complex lines.

Figure 10 - London Underground Tube map with identification of complexities in the network

Figure 10 highlights certain areas in the London Underground network that should the research be expanded may need extra consideration. The highlighted areas, from left to right, include a section of the track that is served by two lines, the

Northern line which splits into two tracks and the Central line which contains a loop. Each of these sections of the network could still have information discovered about them by removing anomalous routes. For example when two lines cover the same track to determine if something was happening to one line and not the other, journeys that exited further down the line, beyond the overlap, could be used to identify issues; once the track splits and a passenger exits from one of the lines it is clear what line they took. For cases of the line splitting or looping the same conditions would be needed, only journeys that were not ambiguous to what route a passenger took would be analysed.

Expanding the research to cover the entire network would allow the researcher and operators to understand the effect a delay on one line may have on another and how the entire network responds to operational delays. It would also provide a map of all the bottlenecks and congestion in the entire network. This could help with future planning projects by identifying what common traits congested stations share and what in the future should be avoided.

This work could also move forward to editing the algorithm to tailor the information more specifically to the network in question. If larger amounts of data were used, trends could be discovered over time of how passengers are affected in certain peak times and when there are operational delays. This could lead to forecasting of congestion and operational delays.

Finally this work could be used to provide real-time information to passengers; if this were to be completed future research could examine how different information affects passengers' behaviour during delays. This could lead to manipulation of the information source to ensure the network is optimally utilised by the passengers.

The scope of future research in this area is vast and as stated in Section 2.2 research surrounding smart card data is becoming increasingly more popular and this thesis paves the way for other academics and operators to know and utilise the possibilities of available information within the data.

8. Bibliography

9292 reist met je mee, n.d. URL http://9292.nl/ (accessed 6.26.12).
Adler, J.L., Recker, W.W., McNally, M.G., 1993. A conflict model and interactive simulator (FASATCARS) for prediciting enroute driver behaviour in response to real-time traffic conditions. Transportation 20, 83–106.
Adult rate Tube, DLR and London Overground fares, n.d. URL https://www.tfl.gov.uk/cdn/static/cms/documents/tube-dlr-lo-adult-fares-jan-2014.pdf
Al-Deek, H., Kanafani, A., 1993. Modelling the benefits of advanced traveller information systems in corridors with incidents. Transp. Res. Part C 1, 303–324.
Al-Deek, H., Martello, M., May, A.D., Sanders, W., 1989. Potential benefits of in-vehicle information systems in a real life freeway corridor under recurring and incident-induced congestion. Veh. Navig. Inf. Syst. Conf. 288–291.
Aranaout, G., Khasawneh, M., Zhang, J., Bowling, S., 2010. An Intelligent Application for Reducing Traffic Congestions using Agent-Based approach. Proc. 2010 IEEE Syst. Inf. Eng. Des. Symp.
Arnott, R., 1991. Does providing information to drivers reduce traffic congestion? Transp. Res. Part Gen. 25, 309–318.
Axhausen, K., Dimitrakopoulou, E., Dimitropoulou, I., 1995. Adapting to change: Some evidence from a simple learning model. Proceedings 23rd Eur. Transp. Forum.
Bagchi, M., White, P.R., 2005. The potential of public transport smart card data. Transp. Policy 12, 464–474.
Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., White, P., 2004. The demand for public transport: a practical guide. TRL Rep. TRL 593.
Bates, J., Jones, P., Cook, A., Polak, J., 2001. Transportation Research Part E: Logistics and Transportation Review - The valuation of reliability for personal travel, in: Transportation Research Part E: Logistics and Transportation Review, Volume 37. pp. 191–229.
BBC - London - Travel - London Underground Map, n.d. URL http://www.bbc.co.uk/london/travel/downloads/tube_map.html (accessed 6.23.14).
Bhat, C., Sardesai, R., 2005. Transportation Research Part B: Methodological - The impact of stop-making and travel time reliability on commute mode choice. URL http://www.sciencedirect.com/science/article/pii/S0191261505001086 (accessed 7.6.12).
Blyth, P., 2004. Improving public transport ticketing through smart cards. Proc. Inst. Civ. Eng. Munic. Eng. 157, 47–54.
Bonsall, P., Parry, T., 1991. Using an interactive route-choice simulator to investigate drivers' compliance with route guidance advice. Transp. Res. Board 59–68.
Bourman, P., Lovric, M., Li, T., van der Hurk, E., 2012. Recognizing Demand Patterns from Smart Card Data for Agent-Based Micro-simulation of Public Transport. Proc. 7th Workshop Agents Traffic Transp. Valencia.
Braess, D., 1968. Über ein Paradoxon aus der Verkehrsplanung. Unternehmensforschung, 12, 258–268.

Brazil, W., Caulfield, B., 2013. Does green make a difference: The potential role of smartphone technology in transport behaviour. Transp. Res. Part C Emerg. Technol. 37, 93–101.
Cascetta, E., Cantrella, G., 1991. A day-to-day and within day dynamic stochastic assignment model. Transp. Res. 255 277–291.
Cats, O., Koutsopoulos, H.N., Burghout, W., Toledo, T., 2011. Effect of Real-Time Transit Information on Dynamic Path Choice of Passengers. Transp. Res. Rec. J. Transp. Res. Board 2217, 46–54.
CBU statistics Wiki, n.d. URL http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/Simon (accessed 8.26.14).
Chan, J., 2007. DSpace@MIT : Rail transit OD matrix estimation and journey time reliability metrics using automated fare data. URL http://dspace.mit.edu/handle/1721.1/38955 (accessed 6.27.12).
Chen, R., Mahmassani, H., 2004. Travel Time Perception and Learning Mechanisms in Traffic Networks. Transp. Res. Bored Natl. Acad. 1894, 209–221.
Chi-square distribution - Wikipedia, the free encyclopaedia, n.d. URL http://en.wikipedia.org/wiki/Chi-squared_test#mediaviewer/File:Chi-square_distributionCDF-English.png (accessed 8.27.14).
Chorus, C., Molin, E., van Wee, B., 2006. Travel information as an instrument to change car drivers travel choices: a literature review. EJTIR 6, 335–364.
Citymapper - The Ultimate Transport App - London, New York, Paris, Berlin, Washington DC, Boston, n.d. URL https://citymapper.com/london/superrouter? (accessed 6.18.14).
Custom Age Tool for ONS Mid-Year Population Estimates | London DataStore, n.d. URL http://data.london.gov.uk/datastore/applications/custom-age-tool-ons-mid-year-population-estimates (accessed 7.29.13).
Daganzo, C., Sheffi, Y., 1977. On stochastic models of traffic assignment. Transp. Sci. 253–274.
De Palma, A., Ben-Akiva, M., Lefevre, C., Litinas, N., 1983. Stochastic equilibrium model of peak period traffic congestion. Transp. Sci. 17 253–274.
Department for Transport, 2012. National Travel Survey: 2011.
Dia, H., 2002. An agent- based approach to modelling driver route choice behaviour under the influence of real-time information. Transp. Res. Part C Emerg. Technol.
Dibbelt, J., Pajor, T., Dorothea, W., 2012. User-Constrained Multi-Modal Route Planning. 2012 Proc. Fourteenth Workshop Algorithm Eng. Exp. ALENEX.
Dziekan, K., Vermeulen, A.J., 2004. The added value of real-time information and effects to customer behaviour- A before-after study of the implementation of a new real-time information system (project report). KTHHTM Stokholm Hague.
Dziekan, K., Vermeulen, A.J., 2006. Psychological effects of and design preferences for real-time information displays. Journal Public Transp. 9, 71–89.
English - Journey Planner - Transport for London, n.d. URL http://journeyplanner.tfl.gov.uk/user/XSLT_TRIP_REQUEST2?language=en (accessed 6.26.12).
Eom, J.K., Sung, M.J., 2011. Analysis of Travel Patterns of the elderly using transit smart card data. Transp. Res. Board 90th Annual meeting, 1–17.
Evans, D., 2011. The Internet of Things: How the Next Evolution of the Internet Is Changing Everything.

Fujii, H., 1989. The CACS project: How far away are we from the dynamic route guidance system? Transp. Future 145159.
Garver, 1932. Concerning the limits of a measure of skewness. Ann Maths Stats 3, 141–142.
Get Your Octopus - Octopus Hong Kong, n.d. URL http://www.octopus.com.hk/get-your-octopus/en/index.html (accessed 6.10.14).
Google Transit, n.d. URL http://www.google.com/intl/en/landing/transit/#mdy (accessed 6.26.12).
GraphPad Statistics Guide, n.d. URL http://www.graphpad.com/guides/prism/6/statistics/index.htm?stat_skewness_and_kurtosis.htm (accessed 8.26.14).
Guo, Z., 2011. Mind the map! The impact of transit maps on path choice in public transit. Transp. Res. Part Policy Pract. 45, 625–639.
Guo, Z., Wilson, N., 2011. Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground. Transp. Res. Part Policy Pract. 45, 91–104.
Haicong, Y., Feng, L., 2012. Advanced multi-modal routing approach for pedestrians. Consum. Electron. Commun. Netw. 2349, 21–23.
Halbing, D., Keltsch, J., Molnar, P., 1997. Modelling the evolution of human trail systems. Nature 388, 47–50.
Harazeen, S., 2011. Effects of information on decisions of passengers during service disruptions.
Harry Beck's Tube map - Transport for London, n.d. URL http://www.tfl.gov.uk/corporate/about-tfl/culture-and-heritage/art-and-design/harry-becks-tube-map (accessed 6.18.14).
Hochmair, H., 2009. The influence of map design on route choice from public transportation maps in urban areas. Cartogr. J. 46, 242–256.
Home | Developers Area | Transport for London, n.d. URL http://www.tfl.gov.uk/businessandpartners/syndication/default.aspx (accessed 7.6.12).
Hong Kong MTR Map, 2014. URL http://vendyxiao.com/wp-content/uploads/2012/01/hong-kong-MTR-map01.jpg (accessed 2.19.14).
Hong Kong: The Facts, 2014. URL http://www.gov.hk/en/about/abouthk/factsheets/docs/transport.pdf
Hotelling, H., Solomons, L., 1932. The limits of a measure of skewness. Ann. Math Stat 3, 141–114.
iBus | Transport for London, n.d. URL http://www.tfl.gov.uk/corporate/projectsandschemes/2373.aspx (accessed 7.6.12).
Information resources - London Transport Museum, n.d. URL http://www.ltmcollection.org/resources/index.html?IXglossary=Tube+Map (accessed 7.5.12).
Internet Growth Statistics - the Global Village Online, n.d. URL http://www.internetworldstats.com/emarketing.htm (accessed 12.12.12).
Jha, M., Mahmassani, H., 1998. Perception updating and day-to-day travel choice dynamics in traffic networks with information provision. Transp. Res. 6C, 189–212.
Kenyon, S., Lyons, G., 2003. The value of integrated multimodal traveller information and its potential contribution to modal change. Transp. Res. Part F, 1–21.

Kobayashi, F., 1979. Feasibility study of route guidance system. Transp. Res. Board 107–112.
Koutsopoulos, H.N., Lotan, T., Yang, Q., 1994. A driving simulator and its application for modelling route choice in the presence of information. Transp. Res. Part C 2, 91–107.
Kusakabe, T., Iryo, T., Yasuo, A., 2010. Estimation method for railway passengers' train choice behaviour with smart card transaction data. Transportation 37, 731–749.
Kyte, S., 2012. Tourism in London: Greater London Authority.
Lamont, D., Kenyon, S., Lyons, G., 2013. Dyslexia and mobility-related social exclusion: the role of travel information provision. J. Transp. Geogr. 26, 147–157.
Lathia, N., Froehlich, J., Capra, L., 2010. Mining Public Transport Usage for Personalised Intelligent Transport Systems. URL http://dl.acm.org/citation.cfm?id=1934588 (accessed 6.29.12).
Liu, B., 1996. Intelligent Route Finding: Combining Knowledge, Cases and An Efficient Search Algorithm, in: In Proceedings of the 12th European Conference on Artificial Intelligence. Wiley & Sons, Ltd, pp. 380–384.
London Councils: London Key Facts, 2014. URL http://www.londoncouncils.gov.uk/londonfacts/default.htm?category=12
London Key Facts and Statistics, n.d. URL http://www.londoncouncils.gov.uk/londonfacts/default.htm?category=2 (accessed 7.28.13).
London transport Apps - Android, n.d. URL http://www.androidzoom.com/android_applications/london+transport (accessed 7.6.12).
London Underground | Transport for London, n.d. URL http://www.tfl.gov.uk/corporate/modesoftransport/1574.aspx (accessed 7.28.13).
Mazloumi, E., Rose, G., Currie, G., Moridpour, S., 2011. Prediction intervals to account for uncertainties in neural network predictions: Methodology and application in bus travel time prediction. Eng. Appl. Artif. Intelligence 24, 534–542.
McDonald, M., Tarrant, D., 1994. ROMANSE (road management system for Europe) PROJECT. Veh. Navig. Information Syst. Conf. 11–14.
Mendes Caiafa, M., 2010. Information needs along the journey chain: users' perspective about the bus system.
Miller, J., 1991. Reaction time analysis with outlier exclusion: Bias varies with sample size. Q. J. Exp. Psychol. 43, 907–912.
MTR corporation, 2013. Legislative Council Panel on Transport: Adjustment to MTR Fares in 2013.
MTR: Our pledge for service 2013, 2013.
Munizaga, M., Palma, C., 2012. Estimation of disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile. Transp. Res. Part C 24, 9–18.
Ofcom | A nation addicted to smartphones, n.d. URL http://consumers.ofcom.org.uk/2011/08/a-nation-addicted-to-smartphones/ (accessed 7.30.13).
Ofcom | Facts & Figures, n.d. URL http://media.ofcom.org.uk/facts/ (accessed 7.6.12).

Othman, N.B., Legara, E.F., Selvam, V., Monerloa, C., 2014. Simulating congestion dynamics of train rapid transit using smart card data. Procedia Comput. Sci. 29, 1610–1620.

RDS, Radio Data System :: Radio-Electronics.com, n.d. URL http://www.radio-electronics.com/info/broadcast/rds/radio-data-system-basics-tutorial.php (accessed 7.5.12).

Rehrl, K., Salzberg, R., Bruntsch, S., Mentz, H.J., 2007. Assisting Multimodal Travellers: Design and Prototypical Implementation of a Personal Travel Companion... Intell. Transp. Syst. IEEE Trans. On 8, 31–42.

Sample Size for Survey: Calculate Respondent Population | Survey Monkey, n.d. URL http://www.surveymonkey.com/mp/sample-size/ (accessed 7.28.13).

Schweiger, C.L., 2003. Customer and media reactions to real-time bus arrival information systems. Transp. Res. Board 48.

Seaborn, C., Attanucci, J., Wilson, N., 2009. Using Smart Card Fare Payment Data to Analyse Multi-Modal Public Transport Journeys in London. URL http://dspace.mit.edu/handle/1721.1/69577 (accessed 6.29.12).

Seng, J., Etienne, M., Raviart, E., Schettini, F., 2012. Towards a web of transport information. 19th ITS World Congr.

Shah, V.P., Wunderlich, K.E., Larkin, J., 2001. Time Management Impacts of Pre-Trip ATIS: Findings from a Washington DC Case Study. 80th Meet. Transp. Res. Board Wash. DC.

Smith, R., Atkins, S., Sheldon, R., 1994. London transport buses: ATT in action and the London countdown route 18 project. Intell. Transp. Syst. Proc. First Wold Congr. Appl. Transp. Telemat. Intell. Veh.-Highw. Syst. 6, 3048–3055.

Solar, A., Marques, A., 2012. ENHANCED WISETRIP: Wide Scale Multimodal and Intelligent Journey Planning. Social Behav. Sci. 48, 2940–2949.

Spitadakis, V., Fostieri, M., 2012. WISETRIP- International multimodal journey planning and delivery of personalized trip information. Social Behav. Sci. 48, 1294–1303.

Sun, X., Winter, S., 2013. An approach to evaluating the usability of public transport journey planners. Proc. Surv. Spat. Sci. Conf.

Transport for London, 2012. Information obtained through discussion with correspondence at TfL.

Transport for London, 2013. Olympic Legacy Monitoring: Personal Travel Behaviour during the Games.

Travel Trends, 2012, n.d. URL http://www.ons.gov.uk/ons/rel/ott/travel-trends/2012/rpt-travel-trends--2012.html (accessed 7.28.13).

Tsuji, H., Takahashi, R., Hironao, K., Yamamoto, Y., 1985. A stochastic approach for estimating the effectiveness of a route guidance system and its related parameters. Transp. Sci. 19, 333–351.

United Nations Population Division | Department of Economic and Social Affairs, n.d. URL http://www.un.org/en/development/desa/population/publications/urbanization/urban-rural.shtml (accessed 7.18.14).

Van Berkum, E., vabb der Mede, P., 1998. The impact of dynamic traffic information: Modelling approach and empirical results, in: Travel Behaviour Research: Updating the State of Play. pp. 401–421.

Van der Hurk, E., Kroon, L., Maroti, G., Verest, P., 2012. Dynamic forecast model of time dependent passenger flows for disruption management.

Verplanken, B., Aarts, H., Van Knippenberg, A., 1998. Habit, information acquisition, and the process of making travel mode choices. Eur. J. Social Psychol. 27, 539–560.

Watkins, K.E., Ferris, B., Borning, A., Rutherford, G.S., Leyton, D., 2011. Where Is My Bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. Transp. Res. Part Policy Pract. 45, 839–848.

Watling, D., 2003. The Dynamics and Equilibria of Day-to-Day Assignment Models. Netw. Spat. Econ. 3, 349–370.

What if They Closed 42d Street and Nobody Noticed? - New York Times, n.d. URL http://www.nytimes.com/1990/12/25/health/what-if-they-closed-42d-street-and-nobody-noticed.html (accessed 6.16.14).

Wischik, D., Handley, M., Raiciu, C., 2009. Control of multipath TCP and optimisation of multipath routing in the Internet. NET - COOP Proc. 3rd Euro-NF Conf. Netw. Control Optim. 204–218.

Wren, A.C., Jones, P.M., 1996. ROMANSE - Road Management System for Europe. Eight Int. Conf. Road Traffic Monit. Control 1996, 23–27.

Yahoo - Route Selection (路線情報 時刻表), n.d. URL http://transit.loco.yahoo.co.jp/ (accessed 6.26.12).

Zhang, L., Jing-Quan, L., Kun, Z., Smok Datta, G., Meng, L., Wei-Bin, Z., Miller, M., Misener, J., 2011. Traveller Information Tool with Integrated Real-Time Transit Information and Multimodal Trip Planning. Transp. Res. Rec. J. Transp. Res. Board 2215, 1–10.

Zhao, J., Frumin, M., Wilson, N., Zhao, Z., 2013. Unified estimator for excess journey time under heterogeneous passenger incidence behaviour using smart card data. Transp. Res. Part C 34, 70–88.

Zografos, K.G., Androutsopoulos, K.N., Nelson, J.D., 2010. Identifying travellers' information needs and services for an integrated international real time journey planning system. Intell. Transp. Syst. ITSC 2010 13th Int. IEEE Conf. 998–1004.

9. Appendices

Table 73 – Southbound Victoria Line average travel times

From	To	Average Oyster journey times	TfL journey planner times
Stockwell	Brixton	6	2
Vauxhall	Brixton	8	5
Pimlico	Brixton	9	6
Victoria	Brixton	11	8
Green Park	Brixton	14	11
Oxford Circus	Brixton	16	13
Warren Street	Brixton	18	15
Euston	Brixton	20	16
Kings Cross	Brixton	23	18
Highbury and Islington	Brixton	25	21
Finsbury Park	Brixton	27	24
Seven Sisters	Brixton	30	28
Tottenham Hale	Brixton	34	31
Blackhorse Road	Brixton	35	33
Walthamstow Central	Brixton	38	35
Vauxhall	Stockwell	5	2
Pimlico	Stockwell	7	4
Victoria	Stockwell	10	6
Green Park	Stockwell	11	8
Oxford Circus	Stockwell	13	10
Warren Street	Stockwell	16	12
Euston	Stockwell	20	14
Kings Cross	Stockwell	20	15
Highbury and Islington	Stockwell	24	19
Finsbury Park	Stockwell	28	21
Seven Sisters	Stockwell	28	25
Tottenham Hale	Stockwell	32	28
Blackhorse Road	Stockwell	34	30
Walthamstow Central	Stockwell	37	32
Pimlico	Vauxhall	4	1
Victoria	Vauxhall	7	3
Green Park	Vauxhall	8	6
Oxford Circus	Vauxhall	11	8

Warren Street	Vauxhall	13	10
Euston	Vauxhall	16	11
Kings Cross	Vauxhall	17	13
Highbury and Islington	Vauxhall	21	16
Finsbury Park	Vauxhall	22	19
Seven Sisters	Vauxhall	25	23
Tottenham Hale	Vauxhall	29	26
Blackhorse Road	Vauxhall	31	28
Walthamstow Central	Vauxhall	33	30
Victoria	Pimlico	5	2
Green Park	Pimlico	8	4
Oxford Circus	Pimlico	10	6
Warren Street	Pimlico	15	8
Euston	Pimlico	15	10
Kings Cross	Pimlico	16	11
Highbury and Islington	Pimlico	19	15
Finsbury Park	Pimlico	21	17
Seven Sisters	Pimlico	24	21
Tottenham Hale	Pimlico	28	24
Blackhorse Road	Pimlico	30	26
Walthamstow Central	Pimlico	33	28
Green Park	Victoria	6	2
Oxford Circus	Victoria	9	4
Warren Street	Victoria	10	6
Euston	Victoria	13	8
Kings Cross	Victoria	14	9
Highbury and Islington	Victoria	17	13
Finsbury Park	Victoria	19	15
Seven Sisters	Victoria	22	19
Tottenham Hale	Victoria	26	22
Blackhorse Road	Victoria	27	24
Walthamstow Central	Victoria	30	26
Oxford Circus	Green Park	6	2
Warren Street	Green Park	7	4
Euston	Green Park	10	5
Kings Cross	Green Park	11	7
Highbury and Islington	Green Park	15	10
Finsbury Park	Green Park	17	13
Seven Sisters	Green Park	20	17
Tottenham Hale	Green Park	23	20
Blackhorse Road	Green Park	25	22
Walthamstow Central	Green Park	28	24

Warren Street	Oxford Circus	6	2
Euston	Oxford Circus	9	3
Kings Cross	Oxford Circus	10	5
Highbury and Islington	Oxford Circus	14	8
Finsbury Park	Oxford Circus	15	11
Seven Sisters	Oxford Circus	19	15
Tottenham Hale	Oxford Circus	22	18
Blackhorse Road	Oxford Circus	24	20
Walthamstow Central	Oxford Circus	27	22
Euston	Warren Street	7	1
Kings Cross	Warren Street	9	3
Highbury and Islington	Warren Street	12	6
Finsbury Park	Warren Street	14	9
Seven Sisters	Warren Street	17	13
Tottenham Hale	Warren Street	21	16
Blackhorse Road	Warren Street	23	18
Walthamstow Central	Warren Street	25	20
Kings Cross	Euston	8	1
Highbury and Islington	Euston	10	5
Finsbury Park	Euston	12	7
Seven Sisters	Euston	17	11
Tottenham Hale	Euston	19	14
Blackhorse Road	Euston	21	16
Walthamstow Central	Euston	24	18
Highbury and Islington	Kings Cross	8	3
Finsbury Park	Kings Cross	11	6
Seven Sisters	Kings Cross	15	10
Tottenham Hale	Kings Cross	17	13
Blackhorse Road	Kings Cross	20	15
Walthamstow Central	Kings Cross	22	17
Finsbury Park	Highbury and Islington	7	2
Seven Sisters	Highbury and Islington	11	6
Tottenham Hale	Highbury and Islington	14	9
Blackhorse Road	Highbury and Islington	16	11
Walthamstow Central	Highbury and Islington	18	13
Seven Sisters	Finsbury Park	7	4
Tottenham Hale	Finsbury Park	10	7
Blackhorse Road	Finsbury Park	12	9
Walthamstow Central	Finsbury Park	15	11
Tottenham Hale	Seven Sisters	7	3
Blackhorse Road	Seven Sisters	8	5
Walthamstow Central	Seven Sisters	11	7

Blackhorse Road	Tottenham Hale	5	2
Walthamstow Central	Tottenham Hale	8	4
Walthamstow Central	Blackhorse Road	6	2

Table 74 – Victoria Line Northbound average travel times

From	To	Average Oyster journey times	TfL journey planner times
Blackhorse Road	Walthamstow Central	9	2
Tottenham Hale	Walthamstow Central	10	5
Seven Sisters	Walthamstow Central	11	7
Finsbury Park	Walthamstow Central	15	11
Highbury and Islington	Walthamstow Central	18	14
Kings Cross	Walthamstow Central	22	17
Euston	Walthamstow Central	25	19
Warren Street	Walthamstow Central	25	21
Oxford Circus	Walthamstow Central	27	23
Green Park	Walthamstow Central	28	26
Victoria	Walthamstow Central	33	28
Pimlico	Walthamstow Central	34	31
Vauxhall	Walthamstow Central	35	32
Stockwell	Walthamstow Central	38	34
Brixton	Walthamstow Central	41	37
Tottenham Hale	Blackhorse Road	6	2
Seven Sisters	Blackhorse Road	8	4
Finsbury Park	Blackhorse Road	12	8
Highbury and Islington	Blackhorse Road	17	11
Kings Cross	Blackhorse Road	21	14
Euston	Blackhorse Road		16
Warren Street	Blackhorse Road	23	18
Oxford Circus	Blackhorse Road	24	20
Green Park	Blackhorse Road	28	23
Victoria	Blackhorse Road	29	25
Pimlico	Blackhorse Road	30	28
Vauxhall	Blackhorse Road	33	29
Stockwell	Blackhorse Road	33	32
Brixton	Blackhorse Road	38	34
Seven Sisters	Tottenham Hale	7	1
Finsbury Park	Tottenham Hale	11	6
Highbury and Islington	Tottenham Hale	15	9

Kings Cross	Tottenham Hale	20	12
Euston	Tottenham Hale	19	14
Warren Street	Tottenham Hale	21	16
Oxford Circus	Tottenham Hale	23	18
Green Park	Tottenham Hale	27	21
Victoria	Tottenham Hale	28	23
Pimlico	Tottenham Hale	29	26
Vauxhall	Tottenham Hale	30	27
Stockwell	Tottenham Hale	35	31
Brixton	Tottenham Hale	37	32
Finsbury Park	Seven Sisters	8	4
Highbury and Islington	Seven Sisters	12	7
Kings Cross	Seven Sisters	14	10
Euston	Seven Sisters	17	12
Warren Street	Seven Sisters	18	13
Oxford Circus	Seven Sisters	21	16
Green Park	Seven Sisters	22	18
Victoria	Seven Sisters	26	20
Pimlico	Seven Sisters	25	24
Vauxhall	Seven Sisters	28	25
Stockwell	Seven Sisters	32	27
Brixton	Seven Sisters	33	30
Highbury and Islington	Finsbury Park	8	3
Kings Cross	Finsbury Park	11	6
Euston	Finsbury Park	11	8
Warren Street	Finsbury Park	13	9
Oxford Circus	Finsbury Park	16	12
Green Park	Finsbury Park	17	14
Victoria	Finsbury Park	22	18
Pimlico	Finsbury Park	21	20
Vauxhall	Finsbury Park	24	21
Stockwell	Finsbury Park	28	23
Brixton	Finsbury Park	30	26
Kings Cross	Highbury and Islington	9	3
Euston	Highbury and Islington	12	5
Warren Street	Highbury and Islington	12	6
Oxford Circus	Highbury and Islington	15	9
Green Park	Highbury and Islington	16	11
Victoria	Highbury and Islington	19	13
Pimlico	Highbury and Islington	21	17
Vauxhall	Highbury and Islington	22	18
Stockwell	Highbury and Islington	26	20

Brixton	Highbury and Islington	28	23
Euston	Kings Cross		2
Warren Street	Kings Cross	10	3
Oxford Circus	Kings Cross	13	6
Green Park	Kings Cross	14	9
Victoria	Kings Cross	18	10
Pimlico	Kings Cross	21	13
Vauxhall	Kings Cross	21	15
Stockwell	Kings Cross	22	17
Brixton	Kings Cross	28	20
Warren Street	Euston		1
Oxford Circus	Euston	10	4
Green Park	Euston	11	6
Victoria	Euston	15	8
Pimlico	Euston	16	11
Vauxhall	Euston	18	13
Stockwell	Euston		15
Brixton	Euston	23	18
Oxford Circus	Warren Street	8	2
Green Park	Warren Street	9	5
Victoria	Warren Street	13	7
Pimlico	Warren Street	14	10
Vauxhall	Warren Street	16	12
Stockwell	Warren Street		14
Brixton	Warren Street	21	17
Green Park	Oxford Circus	6	2
Victoria	Oxford Circus	10	4
Pimlico	Oxford Circus	12	7
Vauxhall	Oxford Circus	14	9
Stockwell	Oxford Circus	17	11
Brixton	Oxford Circus	19	14
Victoria	Green Park	8	2
Pimlico	Green Park	10	5
Vauxhall	Green Park	11	7
Stockwell	Green Park	13	9
Brixton	Green Park	16	12
Pimlico	Victoria	7	3
Vauxhall	Victoria	9	5
Stockwell	Victoria	12	7
Brixton	Victoria	14	10
Vauxhall	Pimlico	8	2
Stockwell	Pimlico	10	4

Brixton	Pimlico	10	7
Stockwell	Vauxhall	6	2
Brixton	Vauxhall	8	5
Brixton	Stockwell	8	3

Table 75 – 26th October Entrance and Exit delays

Time	Entrances with delays	Exits with delays	Average Delays (mins)
07:20:00		Seven Sisters	10
07:29:00	Walthamstow Central		8
07:32:00		Seven Sisters	11
07:33:00	Walthamstow Central		13
07:33:00		Seven Sisters	15
07:34:00		Highbury & Islington	10
07:34:00	Blackhorse Road		11
07:34:00	Walthamstow Central		11
07:35:00		Victoria	9
07:35:00		Highbury & Islington	14
07:35:00	Blackhorse Road		10
07:35:00	Walthamstow Central		11
07:35:00	Seven Sisters		10
07:36:00		Oxford Circus	9
07:36:00		Victoria	11
07:36:00	Euston		7
07:36:00	Seven Sisters		8
07:36:00	Walthamstow Central		13
07:36:00	Highbury & Islington		9
07:37:00	Walthamstow Central		18
07:38:00		Highbury & Islington	11
07:38:00	Tottenham Hale		9
07:38:00	Seven Sisters		18
07:39:00		Tottenham Hale	16
07:39:00	Walthamstow Central		9
07:40:00		Highbury & Islington	12
07:40:00		Green Park	21
07:40:00	Seven Sisters		21
07:41:00		Highbury & Islington	14
07:41:00	Walthamstow Central		8
07:42:00		Finsbury Park	20
07:42:00		Victoria	18
07:42:00		Highbury & Islington	13
07:42:00	Blackhorse Road		17

07:42:00	Seven Sisters		19
07:42:00	Walthamstow Central		10
07:43:00		Finsbury Park	12
07:43:00		Vauxhall	16
07:43:00		Victoria	15
07:43:00		Highbury & Islington	9
07:43:00	Blackhorse Road		12
07:43:00	Euston		17
07:43:00	Seven Sisters		16
07:43:00	Tottenham Hale		3
07:43:00	Walthamstow Central		14
07:44:00		Finsbury Park	13
07:44:00		Seven Sisters	12
07:44:00		Vauxhall	15
07:44:00		Victoria	11
07:44:00		Kings Cross	16
07:44:00	Blackhorse Road		21
07:44:00	Seven Sisters		7
07:44:00	Tottenham Hale		5
07:44:00	Walthamstow Central		22
07:45:00		Highbury & Islington	10
07:45:00		Green Park	22
07:45:00		Kings Cross	20
07:45:00	Blackhorse Road		16
07:45:00	Tottenham Hale		14
07:45:00	Walthamstow Central		21
07:46:00		Finsbury Park	7
07:46:00		Kings Cross	8
07:46:00		Highbury & Islington	18
07:46:00		Victoria	14
07:46:00		Euston	20
07:46:00	Blackhorse Road		14
07:46:00	Walthamstow Central		7
07:47:00		Finsbury Park	16
07:47:00		Kings Cross	14
07:47:00		Victoria	7
07:47:00		Warren Street	9
07:47:00		Euston	15
07:47:00	Euston		15
07:47:00	Seven Sisters		17
07:47:00	Tottenham Hale		10
07:47:00	Walthamstow Central		22
07:48:00		Euston	14
07:48:00		Kings Cross	13
07:48:00		Warren Street	22
07:48:00		Finsbury Park	5

07:48:00	Tottenham Hale		12
07:48:00	Blackhorse Road		13
07:48:00	Seven Sisters		22
07:48:00	Walthamstow Central		19
07:49:00		Finsbury Park	13
07:49:00		Kings Cross	16
07:49:00		Highbury & Islington	17
07:49:00		Oxford Circus	14
07:49:00		Warren Street	13
07:49:00		Vauxhall	17
07:49:00	Tottenham Hale		9
07:49:00		Euston	21
07:49:00	Blackhorse Road		10
07:49:00	Seven Sisters		11
07:49:00	Walthamstow Central		16
07:49:00	Highbury & Islington		12
07:50:00		Vauxhall	16
07:50:00		Oxford Circus	13
07:50:00	Blackhorse Road		10
07:50:00	Seven Sisters		16
07:50:00	Walthamstow Central		15
07:50:00	Tottenham Hale		6
07:50:00	Highbury & Islington		18
07:51:00		Finsbury Park	16
07:51:00		Green Park	15
07:51:00		Highbury & Islington	8
07:51:00		Oxford Circus	18
07:51:00		Warren Street	14
07:51:00		Vauxhall	14
07:51:00		Euston	19
07:51:00	Blackhorse Road		14
07:51:00	Seven Sisters		22
07:51:00	Walthamstow Central		16
07:52:00		Euston	14
07:52:00		Finsbury Park	8
07:52:00		Kings Cross	16
07:52:00		Highbury & Islington	10
07:52:00		Oxford Circus	4
07:52:00		Warren Street	12
07:52:00		Victoria	13
07:52:00	Blackhorse Road		8
07:52:00	Seven Sisters		6
07:52:00	Walthamstow Central		13
07:52:00	Tottenham Hale		28
07:53:00		Warren Street	8
07:53:00		Green Park	13

07:53:00		Kings Cross	13
07:53:00		Victoria	12
07:53:00	Blackhorse Road		12
07:53:00	Seven Sisters		10
07:53:00	Walthamstow Central		12
07:53:00	Highbury & Islington		6
07:53:00	Tottenham Hale		11
07:54:00		Euston	12
07:54:00		Kings Cross	7
07:54:00		Oxford Circus	9
07:54:00		Victoria	10
07:54:00		Green Park	9
07:54:00		Warren Street	5
07:54:00	Blackhorse Road		10
07:54:00	Seven Sisters		7
07:54:00	Walthamstow Central		16
07:54:00	Highbury & Islington		4
07:54:00	Tottenham Hale		12
07:55:00		Oxford Circus	13
07:55:00		Euston	11
07:55:00		Green Park	9
07:55:00		Warren Street	9
07:55:00		Pimlico	10
07:55:00	Blackhorse Road		14
07:55:00	Seven Sisters		6
07:55:00	Walthamstow Central		14
07:55:00	Tottenham Hale		4
07:56:00		Green Park	12
07:56:00		Warren Street	13
07:56:00		Vauxhall	12
07:56:00	Blackhorse Road		12
07:56:00	Seven Sisters		14
07:56:00	Walthamstow Central		11
07:56:00	Tottenham Hale		6
07:57:00		Warren Street	3
07:57:00		Green Park	4
07:57:00		Euston	15
07:57:00		Warren Street	11
07:57:00		Vauxhall	14
07:57:00	Blackhorse Road		11
07:57:00	Seven Sisters		5
07:57:00	Walthamstow Central		13
07:57:00	Tottenham Hale		5
07:58:00		Oxford Circus	11
07:58:00		Green Park	11
07:58:00		Warren Street	9

07:58:00		Victoria	8
07:58:00	Blackhorse Road		9
07:58:00	Seven Sisters		5
07:58:00	Walthamstow Central		11
07:58:00	Highbury & Islington		7
07:58:00	Tottenham Hale		11
07:59:00		Green Park	18
07:59:00		Victoria	13
07:59:00		Warren Street	10
07:59:00		Vauxhall	18
07:59:00		Kings Cross	9
07:59:00	Blackhorse Road		10
07:59:00	Seven Sisters		10
07:59:00	Walthamstow Central		13
07:59:00	Tottenham Hale		9
08:00:00		Vauxhall	17
08:00:00		Pimlico	10
08:00:00	Blackhorse Road		10
08:00:00	Seven Sisters		8
08:00:00	Walthamstow Central		10
08:00:00	Highbury & Islington		8
08:01:00		Vauxhall	12
08:01:00		Green Park	9
08:01:00	Blackhorse Road		4
08:01:00	Seven Sisters		2
08:01:00	Walthamstow Central		12
08:02:00		Victoria	1
08:02:00		Vauxhall	8
08:02:00	Seven Sisters		5
08:02:00	Walthamstow Central		6
08:02:00	Highbury & Islington		3
08:03:00		Victoria	12
08:03:00		Pimlico	13
08:03:00		Warren Street	6
08:03:00	Blackhorse Road		10
08:03:00	Walthamstow Central		9
08:03:00	Highbury & Islington		2
08:03:00	Tottenham Hale		4
08:04:00		Euston	3
08:04:00		Victoria	17
08:04:00		Vauxhall	10
08:04:00	Blackhorse Road		13
08:04:00	Seven Sisters		2
08:04:00	Walthamstow Central		10
08:04:00	Tottenham Hale		7
08:05:00		Highbury & Islington	1

08:05:00		Euston	5
08:05:00		Green Park	10
08:05:00		Pimlico	9
08:05:00		Vauxhall	19
08:05:00	Blackhorse Road		14
08:05:00	Seven Sisters		6
08:05:00	Walthamstow Central		12
08:05:00	Tottenham Hale		5
08:06:00		Green Park	7
08:06:00	Blackhorse Road		8
08:06:00	Highbury & Islington		3
08:07:00	Seven Sisters		4
08:08:00		Green Park	6
08:08:00		Victoria	12
08:08:00	Seven Sisters		7
08:08:00	Walthamstow Central		11
08:08:00	Highbury & Islington		5
08:08:00	Tottenham Hale		4
08:09:00		Vauxhall	5
08:09:00		Pimlico	13
08:09:00		Victoria	16
08:09:00		Green Park	16
08:09:00		Warren Street	10
08:09:00	Blackhorse Road		10
08:09:00	Seven Sisters		3
08:09:00	Walthamstow Central		14
08:09:00	Tottenham Hale		4
08:10:00		Euston	4
08:10:00		Vauxhall	17
08:10:00		Victoria	7
08:10:00		Green Park	20
08:10:00	Blackhorse Road		15
08:10:00	Walthamstow Central		18
08:10:00	Highbury & Islington		3
08:10:00	Tottenham Hale		4
08:11:00		Pimlico	13
08:11:00		Victoria	14
08:11:00	Blackhorse Road		12
08:11:00	Walthamstow Central		13
08:12:00		Green Park	10
08:12:00		Victoria	12
08:12:00		Vauxhall	15
08:12:00		Warren Street	10
08:12:00	Blackhorse Road		10
08:12:00	Walthamstow Central		14
08:13:00		Vauxhall	11

08:13:00		Oxford Circus	2
08:13:00		Green Park	12
08:13:00		Warren Street	1
08:13:00		Pimlico	14
08:13:00	Blackhorse Road		9
08:13:00	Seven Sisters		6
08:13:00	Walthamstow Central		9
08:13:00	Tottenham Hale		2
08:14:00		Vauxhall	8
08:14:00		Pimlico	14
08:14:00		Oxford Circus	1
08:14:00		Green Park	7
08:14:00		Victoria	8
08:14:00	Blackhorse Road		6
08:14:00	Walthamstow Central		9
08:14:00	Tottenham Hale		7
08:15:00		Oxford Circus	2
08:15:00		Green Park	5
08:15:00		Vauxhall	16
08:15:00		Victoria	11
08:15:00	Blackhorse Road		3
08:15:00	Walthamstow Central		10
08:15:00	Tottenham Hale		2
08:16:00		Green Park	1
08:16:00		Vauxhall	12
08:16:00	Blackhorse Road		2
08:16:00	Seven Sisters		2
08:16:00	Walthamstow Central		9
08:16:00	Tottenham Hale		3
08:17:00		green park	2
08:17:00		Victoria	6
08:17:00		Pimlico	10
08:17:00	Blackhorse Road		3
08:17:00	Seven Sisters		1
08:17:00	Walthamstow Central		8
08:17:00	Tottenham Hale		12
08:18:00		Green Park	3
08:18:00		Victoria	5
08:18:00	Blackhorse Road		4
08:18:00	Walthamstow Central		7
08:18:00	Tottenham Hale		1
08:19:00		Victoria	2
08:19:00		Vauxhall	9
08:19:00	Blackhorse Road		4
08:19:00	Walthamstow Central		5
08:20:00		Green Park	1

08:20:00		Pimlico	9
08:20:00	Blackhorse Road		6
08:20:00	Walthamstow Central		6
08:20:00	Tottenham Hale		0
08:21:00		Vauxhall	3
08:21:00		Seven Sisters	1
08:21:00	Blackhorse Road		4
08:21:00	Walthamstow Central		3
08:22:00		Highbury & Islington	1
08:22:00		Vauxhall	4
08:22:00		Victoria	3
08:22:00	Blackhorse Road		2
08:22:00	Walthamstow Central		4
08:23:00		Green Park	3
08:23:00		Vauxhall	4
08:23:00	Blackhorse Road		4
08:23:00	Walthamstow Central		4
08:23:00	Tottenham Hale		10
08:24:00		Vauxhall	5
08:24:00	Walthamstow Central		1
08:25:00		Highbury & Islington	2
08:25:00	Blackhorse Road		6
08:26:00		Vauxhall	0
08:26:00		Highbury & Islington	2
08:26:00	Walthamstow Central		2
08:26:00	Tottenham Hale		0
08:27:00	Blackhorse Road		8
08:27:00	Walthamstow Central		7
08:28:00	Blackhorse Road		0
08:28:00	Walthamstow Central		0
08:29:00		Vauxhall	1
08:29:00		Green Park	6
08:29:00	Blackhorse Road		5
08:29:00	Walthamstow Central		1
08:29:00	Tottenham Hale		0
08:31:00		kings cross	2
08:31:00		Oxford Circus	2
08:31:00		Victoria	0
08:31:00	Blackhorse Road		3
08:31:00	Walthamstow Central		2
08:31:00	Tottenham Hale		2
08:32:00		Green Park	2
08:32:00		Victoria	1
08:32:00	Blackhorse Road		1
08:32:00	Walthamstow Central		1
08:33:00		Green Park	2

08:33:00		kings cross	1
08:33:00		Highbury and Islington	0
08:33:00	Blackhorse Road		1
08:33:00	Walthamstow Central		1
08:33:00	Tottenham Hale		1
08:34:00		Vauxhall	2
08:34:00		Euston	2
08:34:00		kings cross	2
08:34:00		Highbury & Islington	0
08:34:00	Blackhorse Road		2
08:34:00	Walthamstow Central		2
08:34:00	Tottenham Hale		1
08:35:00		Victoria	2
08:35:00	Walthamstow Central		2
08:36:00		Victoria	3
08:36:00		Euston	3
08:36:00		Warren Street	4
08:36:00	Walthamstow Central		3
08:36:00	Tottenham Hale		1
08:37:00		Warren Street	3
08:37:00		Oxford Circus	3
08:37:00		Euston	1
08:37:00		Vauxhall	3
08:37:00	Blackhorse Road		2
08:37:00	Walthamstow Central		3
08:37:00	Tottenham Hale		2
08:38:00		Green Park	2
08:38:00		Oxford Circus	2
08:38:00		kings cross	1
08:38:00		Warren Street	3
08:38:00	Blackhorse Road		2
08:38:00	Walthamstow Central		2
08:38:00	Tottenham Hale		1
08:39:00		Pimlico	2
08:39:00		Warren Street	2
08:39:00		Oxford Circus	2
08:39:00		Green Park	4
08:39:00	Blackhorse Road		4
08:39:00	Walthamstow Central		2
08:39:00	Tottenham Hale		1
08:40:00		Euston	2
08:40:00		Victoria	2
08:40:00		Oxford Circus	2
08:40:00	Walthamstow Central		4
08:40:00	Tottenham Hale		1
08:41:00		Victoria	4

08:41:00		Green Park	3
08:41:00		Warren Street	2
08:41:00		Vauxhall	4
08:41:00	Blackhorse Road		2
08:41:00	Walthamstow Central		3
08:41:00	Tottenham Hale		2
08:42:00		Green Park	3
08:42:00		Victoria	4
08:42:00		Warren Street	1
08:42:00	Blackhorse Road		2
08:42:00	Walthamstow Central		3
08:43:00		Pimlico	4
08:43:00		Victoria	4
08:43:00		oxford circus	1
08:43:00	Blackhorse Road		3
08:43:00	Walthamstow Central		4
08:43:00	Tottenham Hale		2
08:44:00		Oxford Circus	0
08:44:00		Victoria	3
08:44:00	Blackhorse Road		2
08:44:00	Walthamstow Central		3
08:44:00	Tottenham Hale		1
08:45:00		Pimlico	3
08:45:00		Green Park	4
08:45:00		Victoria	3
08:45:00	Blackhorse Road		4
08:45:00	Walthamstow Central		3
08:45:00	Tottenham Hale		2
08:46:00		Green Park	2
08:46:00		Victoria	3
08:46:00		Pimlico	5
08:46:00		Vauxhall	5
08:46:00	Blackhorse Road		4
08:46:00	Walthamstow Central		3
08:46:00	Tottenham Hale		1
08:47:00		Green Park	2
08:47:00		Victoria	2
08:47:00	Blackhorse Road		3
08:47:00	Walthamstow Central		2
08:48:00		Vauxhall	4
08:48:00		Green Park	1
08:48:00		Victoria	1
08:48:00	Blackhorse Road		1
08:48:00	Walthamstow Central		3
08:48:00	Highbury & Islington		1
08:48:00	Tottenham Hale		1

08:49:00		Victoria	1
08:49:00		Vauxhall	5
08:49:00		Pimlico	4
08:49:00	Blackhorse Road		3
08:49:00	Seven Sisters		1
08:49:00	Walthamstow Central		3
08:49:00	Tottenham Hale		1
08:50:00		Victoria	1
08:50:00		Warren Street	4
08:50:00	Seven Sisters		4
08:50:00	Walthamstow Central		2
08:50:00	Highbury & Islington		1
08:50:00	Tottenham Hale		1
08:51:00		Oxford Circus	1
08:51:00		Pimlico	2
08:51:00		Vauxhall	5
08:51:00	Blackhorse Road		4
08:51:00	Seven Sisters		3
08:51:00	Walthamstow Central		4
08:52:00		kings cross	1
08:52:00		Vauxhall	4
08:52:00		Oxford Circus	4
08:52:00	Walthamstow Central		3
08:52:00	Highbury & Islington		2
08:52:00	Tottenham Hale		1
08:53:00		Stockwell	5
08:53:00		Warren Street	2
08:53:00		Oxford Circus	4
08:53:00	Blackhorse Road		3
08:53:00	Seven Sisters		3
08:53:00	Walthamstow Central		2
08:53:00	Highbury & Islington		3
08:54:00		Warren Street	4
08:54:00		Euston	1
08:54:00	Blackhorse Road		2
08:54:00	Seven Sisters		3
08:54:00	Walthamstow Central		3
08:54:00	Highbury & Islington		4
08:55:00		Oxford Circus	4
08:55:00		Stockwell	4
08:55:00		Victoria	1
08:55:00		Warren Street	1
08:55:00	Blackhorse Road		3
08:55:00	Seven Sisters		3
08:55:00	Walthamstow Central		2
08:55:00	Highbury & Islington		4

08:55:00	Tottenham Hale		2
08:56:00		Euston	4
08:56:00		Oxford Circus	4
08:56:00		Victoria	2
08:56:00		Warren Street	2
08:56:00	Blackhorse Road		6
08:56:00	Seven Sisters		2
08:56:00	Walthamstow Central		1
08:56:00	Highbury & Islington		3
08:56:00	Tottenham Hale		3
08:57:00		Euston	3
08:57:00		Green Park	3
08:57:00		Oxford Circus	2
08:57:00		Victoria	4
08:57:00		Warren Street	2
08:57:00		kings cross	1
08:57:00	Blackhorse Road		1
08:57:00	Seven Sisters		2
08:57:00	Walthamstow Central		2
08:57:00	Highbury & Islington		3
08:57:00	Tottenham Hale		3
08:58:00		Victoria	7
08:58:00		Green Park	3
08:58:00		Oxford Circus	4
08:58:00	Seven Sisters		2
08:58:00	Highbury & Islington		5
08:58:00	Tottenham Hale		3
08:59:00		Oxford Circus	2
08:59:00		Victoria	4
08:59:00		Pimlico	3
08:59:00		Green Park	3
08:59:00	Seven Sisters		4
08:59:00	Walthamstow Central		2
08:59:00	Highbury & Islington		3
08:59:00	Tottenham Hale		2
09:00:00		Green Park	2
09:00:00		Victoria	3
09:00:00		Vauxhall	4
09:00:00	Blackhorse Road		3
09:00:00	Seven Sisters		4
09:00:00	Walthamstow Central		1
09:00:00	Highbury & Islington		4
09:00:00	Tottenham Hale		2
09:01:00		Victoria	4
09:01:00		Vauxhall	3
09:01:00		Pimlico	3

09:01:00		Green Park	2
09:01:00	Blackhorse Road		5
09:01:00	Seven Sisters		4
09:01:00	Highbury & Islington		2
09:02:00		Victoria	3
09:02:00		Vauxhall	3
09:02:00		Green Park	1
09:02:00		Pimlico	4
09:02:00	Blackhorse Road		4
09:02:00	Seven Sisters		3
09:02:00	Walthamstow Central		2
09:02:00	Highbury & Islington		2
09:02:00	Tottenham Hale		2
09:03:00		Vauxhall	3
09:03:00		Pimlico	3
09:03:00	Seven Sisters		3
09:03:00	Highbury & Islington		2
09:04:00		Victoria	1
09:04:00		Pimlico	3
09:04:00		Green Park	1
09:04:00		Vauxhall	4
09:04:00	Seven Sisters		3
09:04:00	Walthamstow Central		2
09:04:00	Highbury & Islington		3
09:04:00	Tottenham Hale		1
09:05:00		Pimlico	2
09:05:00		Green Park	1
09:05:00		Victoria	2
09:05:00		Vauxhall	3
09:05:00	Seven Sisters		1
09:05:00	Walthamstow Central		2
09:05:00	Highbury & Islington		3
09:05:00	Tottenham Hale		2
09:06:00		Pimlico	0
09:06:00		Victoria	3
09:06:00		Vauxhall	2
09:06:00	Seven Sisters		1
09:06:00	Walthamstow Central		2
09:06:00	Highbury & Islington		2
09:07:00		Victoria	4
09:07:00		Vauxhall	1
09:07:00	Walthamstow Central		4
09:07:00	Highbury & Islington		3
09:08:00		Pimlico	8
09:08:00	Highbury & Islington		15
09:09:00		Vauxhall	7

09:09:00	Highbury & Islington		8
09:18:00	Walthamstow Central		7
09:19:00	Highbury & Islington		8
09:20:00		Vauxhall	8
09:20:00	Highbury & Islington		9
09:22:00		Oxford Circus	8
09:22:00	Highbury & Islington		7
09:23:00		Oxford Circus	7
09:23:00	Highbury & Islington		7
09:24:00		Green Park	7
09:27:00		Victoria	7

Table 76 – 4th October entrance and exit delays

Time	Entry Stations with Delays	Exit Stations with Delays	Minutes Delayed
07:41	Walthamstow		1
07:43	Tottenham Hale		3
07:45	Tottenham Hale		14
07:47	Tottenham Hale		11
07:48	Tottenham Hale		14
07:49	Tottenham Hale		9
07:50	Tottenham Hale		9
07:52	Tottenham Hale		8
07:53	Tottenham Hale		11
07:54	Tottenham Hale		13
07:55	Tottenham Hale		6
07:57	Tottenham Hale		4
07:58	Tottenham Hale		9
07:59	Tottenham Hale		9
08:03	Tottenham Hale		4
08:04	Tottenham Hale		5
08:05	Tottenham Hale		5
08:09	Tottenham Hale		4
08:10	Tottenham Hale		3
08:12	Tottenham Hale		2
08:14	Tottenham Hale		6
08:15	Tottenham Hale		2
08:16	Tottenham Hale		3
08:17		Green Park	1
08:17	Tottenham Hale		12
08:18		Vauxhall	0
08:18		Victoria	1
08:18	Highbury & Islington		1
08:18	Seven Sisters		1

08:18	Tottenham Hale		2
08:19		Victoria	1
08:19		Vauxhall	2
08:19	Finsbury Park		1
08:19	Seven Sisters		1
08:20		Oxford Circus	1
08:20	Walthamstow Central		1
08:21		Warren Street	1
08:21		Victoria	1
08:21		Green Park	1
08:21	Finsbury Park		1
08:21	Highbury & Islington		1
08:21	Walthamstow Central		1
08:22		Victoria	1
08:22		Oxford Circus	1
08:22		Green Park	1
08:22	Finsbury Park		0
08:22	Highbury & Islington		1
08:22	Euston		1
08:22	Walthamstow Central		1
08:23		Vauxhall	1
08:23		Green Park	1
08:23		Oxford Circus	1
08:23	Finsbury Park		1
08:23	Highbury & Islington		1
08:23	Kings Cross		1
08:23	Euston		1
08:23	Seven Sisters		2
08:23	Tottenham Hale		9
08:24		Victoria	1
08:24		Vauxhall	1
08:24		Warren Street	1
08:24		Oxford Circus	1
08:24		Green Park	1
08:24	Blackhorse Road		1
08:24	Finsbury Park		1
08:24	Highbury & Islington		1
08:24	Kings Cross		1
08:24	Seven Sisters		1
08:24	Walthamstow Central		1
08:25		Victoria	3
08:25		Pimlico	2
08:25		Oxford Circus	2
08:25		Vauxhall	1
08:25	Finsbury Park		1
08:25	Seven Sisters		1

08:25	Walthamstow Central		1
08:26		Victoria	1
08:26		Oxford Circus	1
08:26		Green Park	1
08:26	Finsbury Park		1
08:26	Walthamstow Central		3
08:27		Green Park	1
08:27		Highbury & Islington	1
08:27		Vauxhall	1
08:27		Oxford Circus	2
08:27		Victoria	2
08:27	Blackhorse Road		3
08:27	Finsbury Park		2
08:27	Euston		2
08:27	Seven Sisters		3
08:27	Walthamstow Central		1
08:28		Highbury & Islington	1
08:28		Vauxhall	4
08:28		Victoria	2
08:28		Oxford Circus	1
08:28		Green Park	1
08:28	Finsbury Park		3
08:28	Highbury & Islington		3
08:28	Kings Cross		1
08:28	Euston		2
08:28	Seven Sisters		2
08:28	Walthamstow Central		1
08:29		Warren Street	1
08:29		Vauxhall	4
08:29		Victoria	3
08:29		Green Park	1
08:29	Blackhorse Road		2
08:29	Finsbury Park		3
08:29	Highbury & Islington		3
08:29	Seven Sisters		3
08:29	Walthamstow Central		2
08:30		Green Park	1
08:30		Victoria	2
08:30		Warren Street	1
08:30	Blackhorse Road		2
08:30	Finsbury Park		3
08:30	Euston		1
08:30	Seven Sisters		1
08:30	Walthamstow Central		1
08:31		Victoria	2
08:31		Pimlico	2

08:31		Oxford Circus	1
08:31	Blackhorse Road		2
08:31	Finsbury Park		3
08:31	Highbury & Islington		3
08:31	Euston		2
08:31	Seven Sisters		1
08:31	Walthamstow Central		2
08:31	Tottenham Hale		2
08:32		Vauxhall	4
08:32		Pimlico	3
08:32		Victoria	2
08:32	Finsbury Park		3
08:32	Highbury & Islington		3
08:32	Euston		2
08:32	Seven Sisters		4
08:32	Walthamstow Central		3
08:33		Vauxhall	4
08:33		Victoria	2
08:33		Green Park	1
08:33		Pimlico	3
08:33	Finsbury Park		5
08:33	Highbury & Islington		3
08:33	Euston		1
08:33	Seven Sisters		3
08:33	Walthamstow Central		5
08:33	Tottenham Hale		1
08:34		Vauxhall	3
08:34		Victoria	2
08:34		Green Park	2
08:34	Finsbury Park		2
08:34	Kings Cross		2
08:34	Euston		2
08:34	Seven Sisters		4
08:34	Walthamstow Central		4
08:34	Tottenham Hale		1
08:35		Warren Street	1
08:35		Vauxhall	4
08:35		Victoria	3
08:35		Green Park	1
08:35	Finsbury Park		2
08:35	Kings Cross		3
08:35	Euston		2
08:35	Seven Sisters		2
08:35	Walthamstow Central		4
08:35	Highbury & Islington		1
08:36		Oxford Circus	2

08:36		Victoria	3
08:36		Green Park	3
08:36	Finsbury Park		1
08:36	Blackhorse Road		5
08:36	Euston		2
08:36	Seven Sisters		2
08:36	Walthamstow Central		5
08:36	Tottenham Hale		1
08:37		Pimlico	1
08:37		Oxford Circus	2
08:37		Victoria	2
08:37		Green Park	2
08:37	Finsbury Park		1
08:37	Blackhorse Road		3
08:37	Seven Sisters		2
08:37	Tottenham Hale		2
08:38		Vauxhall	3
08:38		Pimlico	3
08:38		Kings Cross	1
08:38		Victoria	1
08:38		Green Park	3
08:38	Finsbury Park		2
08:38	Blackhorse Road		2
08:38	Euston		1
08:38	Seven Sisters		4
08:38	Highbury & Islington		2
08:38	Tottenham Hale		1
08:39		Vauxhall	4
08:39		Pimlico	2
08:39		Euston	2
08:39		Victoria	2
08:39	Finsbury Park		3
08:39	Blackhorse Road		5
08:39	Euston		3
08:39	Walthamstow Central		4
08:39	Tottenham Hale		1
08:40		Kings Cross	3
08:40		Green Park	2
08:40		Vauxhall	2
08:40		Euston	4
08:40		Victoria	3
08:40	Finsbury Park		3
08:40	Blackhorse Road		7
08:40	Walthamstow Central		5
08:40	Tottenham Hale		1
08:40	Highbury & Islington		1

08:40	Seven Sisters		4
08:40	Euston		1
08:41		Kings Cross	3
08:41		Vauxhall	3
08:41		Pimlico	0
08:41		Euston	2
08:41		Victoria	3
08:41	Finsbury Park		2
08:41	Blackhorse Road		4
08:41	Euston		2
08:41	Walthamstow Central		5
08:41	Tottenham Hale		2
08:41	Highbury & Islington		1
08:42		Highbury & Islington	1
08:42		Kings Cross	3
08:42		Euston	7
08:42		Victoria	7
08:42	Finsbury Park		1
08:42	Walthamstow Central		6
08:42	Tottenham Hale		2
08:43		Highbury & Islington	1
08:43		Kings Cross	3
08:43		Euston	7
08:43		Victoria	7
08:43	Finsbury Park		1
08:43	Walthamstow Central		6
08:43	Tottenham Hale		2
08:44		Highbury & Islington	1
08:44		Oxford Circus	4
08:44		Warren Street	5
08:44		Vauxhall	7
08:44		Pimlico	1
08:44		Euston	4
08:44		Victoria	4
08:44	Finsbury Park		9
08:44	Blackhorse Road		2
08:44	Euston		5
08:44	Tottenham Hale		7
08:45		Kings Cross	5
08:45		Warren Street	7
08:45		Oxford Circus	9
08:45		Pimlico	2
08:45		Euston	4
08:45		Victoria	7
08:45	Blackhorse Road		7
08:45	Tottenham Hale		1

08:45	Highbury & Islington		7
08:46		Kings Cross	7
08:46		Warren Street	8
08:46		Vauxhall	10
08:46		Pimlico	1
08:46		Euston	3
08:46		Victoria	7
08:46	Blackhorse Road		7
08:46	Euston		3
08:46	Tottenham Hale		4
08:46	Highbury & Islington		7
08:47		Oxford Circus	2
08:47		Kings Cross	6
08:47		Warren Street	9
08:47		Green Park	2
08:47		Vauxhall	9
08:47		Euston	5
08:47		Victoria	2
08:47	Finsbury Park		3
08:47	Blackhorse Road		7
08:47	Euston		2
08:47	Walthamstow Central		8
08:47	Tottenham Hale		7
08:47	Highbury & Islington		7
08:47	Seven Sisters		5
08:48		Oxford Circus	4
08:48		Kings Cross	7
08:48		Warren Street	8
08:48		Green Park	2
08:48		Vauxhall	9
08:48		Pimlico	1
08:48		Euston	4
08:48		Victoria	3
08:48	Finsbury Park		7
08:48	Blackhorse Road		8
08:48	Euston		2
08:48	Walthamstow Central		9
08:48	Tottenham Hale		1
08:48	Highbury & Islington		4
08:48	Kings Cross		3
08:48	Seven Sisters		4
08:49		Oxford Street	8
08:49		Kings Cross	4
08:49		Warren Street	8
08:49		Green Park	9
08:49		Vauxhall	4

08:49		Victoria	5
08:49	Finsbury Park		6
08:49	Blackhorse Road		9
08:49	Euston		3
08:49	Walthamstow Central		11
08:49	Tottenham Hale		1
08:49	Highbury & Islington		5
08:49	Seven Sisters		6
08:49	Kings Cross		3
08:50		Oxford Street	4
08:50		Kings Cross	4
08:50		Warren Street	6
08:50		Euston	7
08:50		Victoria	8
08:50	Finsbury Park		6
08:50	Walthamstow Central		8
08:50	Tottenham Hale		1
08:50	Highbury & Islington		6
08:50	Seven Sisters		6
08:50	Kings Cross		2
08:51		Green Park	6
08:51		Victoria	7
08:51	Finsbury Park		7
08:51	Euston		4
08:51	Walthamstow Central		10
08:51	Highbury & Islington		7
08:51	Seven Sisters		4
08:52		Oxford Circus	8
08:52		Kings Cross	3
08:52		Warren Street	6
08:52		Green Park	7
08:52		Pimlico	3
08:52		Victoria	9
08:52	Finsbury Park		7
08:52	Seven Sisters		5
08:52	Euston		4
08:52	Walthamstow Central		13
08:52	Highbury & Islington		7
08:52	Kings Cross		4
08:52	Warren Street		3
08:53		Kings Cross	5
08:53		Warren Street	8
08:53		Oxford Circus	9
08:53		Vauxhall	9
08:53		Pimlico	10
08:53		Euston	7

08:53		Victoria	7
08:53	Finsbury Park		8
08:53	Blackhorse Road		7
08:53	Euston		6
08:53	Walthamstow Central		11
08:53	Tottenham Hale		1
08:53	Highbury & Islington		7
08:53	Kings Cross		5
08:53	Seven Sisters		9
08:54		Highbury & Islington	5
08:54		Kings Cross	5
08:54		Warren Street	10
08:54		Green Park	9
08:54		Vauxhall	7
08:54		Pimlico	5
08:54		Victoria	7
08:54	Finsbury Park		7
08:54	Blackhorse Road		7
08:54	Euston		5
08:54	Walthamstow Central		10
08:54	Highbury & Islington		8
08:54	Kings Cross		4
08:54	Seven Sisters		9
08:55		Highbury & Islington	4
08:55		Oxford Circus	9
08:55		Highbury & Islington	4
08:55		Kings Cross	4
08:55		Warren Street	7
08:55		Green Park	7
08:55		Vauxhall	8
08:55		Pimlico	4
08:55		Euston	7
08:55		Victoria	9
08:55	Finsbury Park		8
08:55	Blackhorse Road		10
08:55	Euston		5
08:55	Walthamstow Central		11
08:55	Tottenham Hale		2
08:55	Highbury & Islington		7
08:55	Kings Cross		4
08:55	Seven Sisters		9
08:56		Oxford Circus	6
08:56		Highbury & Islington	5
08:56		Kings Cross	5
08:56		Warren Street	8
08:56		Green Park	5

08:56		Vauxhall	9
08:56		Pimlico	3
08:56		Euston	7
08:56		Victoria	6
08:56	Finsbury Park		7
08:56	Blackhorse Road		9
08:56	Euston		4
08:56	Walthamstow Central		10
08:56	Tottenham Hale		3
08:56	Highbury & Islington		8
08:56	Warren Street		1
08:56	Seven Sisters		7
08:57		Highbury & Islington	5
08:57		Warren Street	8
08:57		Green Park	7
08:57		Vauxhall	9
08:57		Oxford Circus	9
08:57		Euston	6
08:57		Victoria	10
08:57	Finsbury Park		10
08:57	Blackhorse Road		8
08:57	Euston		5
08:57	Walthamstow Central		9
08:57	Tottenham Hale		3
08:57	Highbury & Islington		9
08:57	Seven Sisters		9
08:58		Highbury & Islington	3
08:58		Kings Cross	5
08:58		Oxford Circus	8
08:58		Victoria	10
08:58		Warren Street	9
08:58		Pimlico	5
08:58		Vauxhall	9
08:58		Warren Street	9
08:58		Green Park	5
08:58		Euston	6
08:58	Blackhorse Road		9
08:58	Finsbury Park		8
08:58	Highbury & Islington		8
08:58	Kings Cross		3
08:58	Euston		3
08:58	Seven Sisters		8
08:58	Walthamstow Central		10
08:58	Tottenham Hale		3
08:59		Euston	6
08:59		Highbury & Islington	2

08:59		Kings Cross	6
08:59		Oxford Circus	9
08:59		Victoria	8
08:59		Warren Street	8
08:59		Pimlico	8
08:59		Green Park	9
08:59	Blackhorse Road		9
08:59	Finsbury Park		9
08:59	Highbury & Islington		8
08:59	Kings Cross		3
08:59	Euston		5
08:59	Seven Sisters		10
08:59	Walthamstow Central		11
08:59	Tottenham Hale		2
09:00		Euston	7
09:00		Green Park	9
09:00		Kings Cross	5
09:00		Oxford Circus	9
09:00		Victoria	8
09:00		Vauxhall	9
09:00		Pimlico	7
09:00		Warren Street	8
09:00		Highbury & Islington	3
09:00	Blackhorse Road		9
09:00	Finsbury Park		10
09:00	Highbury & Islington		7
09:00	Euston		6
09:00	Seven Sisters		8
09:00	Walthamstow Central		11
09:00	Tottenham Hale		2
09:01		Green Park	10
09:01		Highbury & Islington	2
09:01		Kings Cross	6
09:01		Oxford Circus	8
09:01		Victoria	10
09:01		Warren Street	9
09:01		Vauxhall	12
09:01		Pimlico	10
09:01	Blackhorse Road		8
09:01	Finsbury Park		10
09:01	Highbury & Islington		8
09:01	Euston		4
09:01	Seven Sisters		10
09:01	Walthamstow Central		9
09:02		Euston	6
09:02		Green Park	7

09:02		Highbury & Islington	3
09:02		Kings Cross	4
09:02		Victoria	9
09:02		Warren Street	8
09:02		Pimlico	8
09:02		Vauxhall	9
09:02		Oxford Circus	8
09:02	Blackhorse Road		7
09:02	Finsbury Park		9
09:02	Highbury & Islington		7
09:02	Seven Sisters		9
09:02	Walthamstow Central		8
09:02	Tottenham Hale		2
09:03		Euston	5
09:03		Highbury & Islington	2
09:03		Kings Cross	2
09:03		Oxford Circus	7
09:03		Victoria	10
09:03		Warren Street	7
09:03		Vauxhall	10
09:03		Green Park	8
09:03	Blackhorse Road		8
09:03	Finsbury Park		7
09:03	Highbury & Islington		7
09:03	Euston		4
09:03	Seven Sisters		8
09:03	Walthamstow Central		7
09:04		Euston	4
09:04		Oxford Circus	6
09:04		Victoria	8
09:04		Warren Street	7
09:04		Vauxhall	10
09:04		Green Park	8
09:04	Blackhorse Road		8
09:04	Finsbury Park		8
09:04	Highbury & Islington		6
09:04	Kings Cross		1
09:04	Seven Sisters		9
09:04	Walthamstow		7
09:04	Tottenham Hale		2
09:05		Green Park	7
09:05		Kings Cross	2
09:05		Oxford Circus	6
09:05		Victoria	10
09:05		Warren Street	7
09:05		Highbury & Islington	1

09:05		Pimlico	9
09:05		Euston	4
09:05	Blackhorse Road		7
09:05	Finsbury Park		7
09:05	Highbury & Islington		5
09:05	Seven Sisters		7
09:05	Walthamstow		8
09:05	Tottenham Hale		2
09:06		Oxford Circus	6
09:06		Warren Street	5
09:06		Vauxhall	8
09:06		Victoria	9
09:06		Green Park	9
09:06		Pimlico	4
09:06	Blackhorse Road		6
09:06	Finsbury Park		8
09:06	Highbury & Islington		6
09:06	Kings Cross		1
09:06	Seven Sisters		6
09:06	Walthamstow		7
09:06	Tottenham Hale		1
09:07		Euston	5
09:07		Kings Cross	3
09:07		Oxford Circus	5
09:07		Warren Street	6
09:07		Pimlico	10
09:07		Victoria	10
09:07	Blackhorse Road		5
09:07	Finsbury Park		7
09:07	Highbury & Islington		6
09:07	Seven Sisters		7
09:07	Walthamstow		6
09:08		Euston	3
09:08		Green Park	5
09:08		Kings Cross	3
09:08		Oxford Circus	6
09:08		Victoria	9
09:08		Vauxhall	10
09:08		Pimlico	8
09:08		Warren Street	6
09:08	Blackhorse Road		6
09:08	Finsbury Park		8
09:08	Highbury & Islington		6
09:08	Euston		5
09:08	Seven Sisters		7
09:08	Walthamstow		7

09:08	Walthamstow		8
09:09		Euston	3
09:09		Green Park	7
09:09		Victoria	8
09:09		Warren Street	6
09:09		Vauxhall	10
09:09		Pimlico	8
09:09		Oxford Circus	5
09:09	Blackhorse Road		7
09:09	Finsbury Park		7
09:09	Highbury & Islington		6
09:09	Seven Sisters		6
09:09	Walthamstow		7
09:10		Euston	3
09:10		Green Park	4
09:10		Oxford Circus	5
09:10		Victoria	7
09:10		Warren Street	3
09:10		Vauxhall	8
09:10		Pimlico	7
09:10		Kings Cross	4
09:10	Blackhorse Road		7
09:10	Finsbury Park		5
09:10	Highbury & Islington		5
09:10	Euston		2
09:10	Seven Sisters		7
09:10	Walthamstow		6
09:11		Euston	1
09:11		Oxford Circus	5
09:11		Victoria	8
09:11		Warren Street	2
09:11		Pimlico	7
09:11		Vauxhall	9
09:11		Green Park	6
09:11	Blackhorse Road		7
09:11	Finsbury Park		7
09:11	Highbury & Islington		6
09:11	Seven Sisters		7
09:11	Walthamstow		5
09:12		Green Park	5
09:12		Oxford Circus	4
09:12		Victoria	7
09:12		Warren Street	1
09:12		Pimlico	5
09:12		Vauxhall	7
09:12	Blackhorse Road		6

09:12	Finsbury Park		7
09:12	Highbury & Islington		6
09:12	Seven Sisters		6
09:12	Walthamstow		5
09:13		Oxford Circus	3
09:13		Warren Street	1
09:13		Vauxhall	7
09:13		Victoria	5
09:13		Kings Cross	2
09:13	Blackhorse Road		5
09:13	Finsbury Park		5
09:13	Highbury & Islington		2
09:13	Warren Street		4
09:13	Euston		3
09:13	Seven Sisters		4
09:13	Walthamstow		4
09:13	Tottenham Hale		4
09:14		Euston	1
09:14		Green park	5
09:14		Oxford Circus	2
09:14		Victoria	5
09:14		Warren Street	1
09:14		Vauxhall	6
09:14		Pimlico	4
09:14	Blackhorse Road		5
09:14	Finsbury Park		5
09:14	Highbury & Islington		3
09:14	Seven Sisters		3
09:14	Walthamstow		4
09:15		Oxford Circus	3
09:15		Victoria	5
09:15		Vauxhall	6
09:15		Warren Street	3
09:15		Green Park	2
09:15	Blackhorse Road		4
09:15	Finsbury Park		5
09:15	Highbury & Islington		3
09:15	Seven Sisters		4
09:15	Walthamstow		4
09:16		Victoria	4
09:16		Vauxhall	6
09:16		Oxford Circus	3
09:16		Kings Cross	1
09:16	Finsbury Park		4
09:16	Seven Sisters		4
09:16	Walthamstow		3

09:17		Green Park	3
09:17		Warren Street	2
09:17		Oxford Circus	2
09:17		Victoria	4
09:17		Vauxhall	8
09:17		Pimlico	5
09:17		Kings Cross	1
09:17	Blackhorse Road		6
09:17	Finsbury Park		2
09:17	Highbury & Islington		3
09:17	Seven Sisters		5
09:17	Walthamstow		3
09:18		Green Park	3
09:18		Oxford Street	1
09:18		Warren Street	1
09:18		Vauxhall	6
09:18		Victoria	4
09:18		Pimlico	4
09:18	Blackhorse Road		6
09:18	Finsbury Park		3
09:18	Highbury & Islington		5
09:18	Seven Sisters		4
09:18	Walthamstow		3
09:19		Oxford Circus	1
09:19		Warren Street	1
09:19		Victoria	4
09:19		Vauxhall	4
09:19		Pimlico	5
09:19	Blackhorse Road		2
09:19	Finsbury Park		2
09:19	Highbury & Islington		4
09:19	Seven Sisters		6
09:19	Walthamstow		3
09:20		Green Park	3
09:20		Oxford Circus	1
09:20		Victoria	3
09:20		Vauxhall	4
09:20	Blackhorse Road		4
09:20	Highbury & Islington		2
09:20	Seven Sisters		5
09:20	Walthamstow		2
09:21		Pimlico	3
09:21		Victoria	3
09:21	Blackhorse Road		4
09:21	Seven Sisters		4
09:21	Walthamstow		3

09:22		Victoria	3
09:22		Oxford Circus	1
09:22		Green Park	2
09:22		Pimlico	1
09:22	Blackhorse Road		3
09:22	Finsbury Park		1
09:22	Seven Sisters		2
09:22	Walthamstow		2
09:23		Victoria	1
09:23		Vauxhall	3
09:23	Blackhorse Road		2
09:23	Seven Sisters		3
09:23	Walthamstow		3
09:24		Green Park	1
09:24		Victoria	2
09:24	Blackhorse Road		2
09:24	Walthamstow		2
09:25		Vauxhall	3
09:25	Blackhorse Road		2
09:25	Highbury & Islington		2
09:25	Seven Sisters		3
09:26		Vauxhall	3
09:26		Stockwell	6
09:26	Blackhorse Road		2
09:26	Finsbury Park		4
09:27		Vauxhall	2
09:27	Blackhorse Road		1
09:27	Walthamstow		1
09:29		Vauxhall	1
09:30		Victoria	2
09:30		Vauxhall	1
09:30	Blackhorse Road		1
09:30	Walthamstow		2
09:31	Highbury & Islington		3
09:32		Victoria	4
09:32	Kings Cross		5
09:33		Victoria	4
09:33	Kings Cross		3
09:33	Walthamstow		1
09:34		Vauxhall	2
09:36		Victoria	6
09:48	Highbury & Islington		12
09:49	Highbury & Islington		6
09:50		Oxford Circus	0
09:51		Green Park	1
09:52		Green Park	1

09:52	Finsbury Park		1
09:53		Warren Street	1
09:54		Warren Street	1
09:54	Seven Sisters		2
09:55		Oxford Circus	2
09:55		Warren Street	2
09:55	Highbury & Islington		2
09:55	Seven Sisters		1
09:56		Warren Street	1
09:56		Oxford Circus	2
09:56	Finsbury Park		2
09:56	Highbury & Islington		1
09:57		Oxford Circus	1
09:57		Green Park	2
09:57	Highbury & Islington		2
09:58	Blackhorse Road		1
09:59		Victoria	2
09:59	Finsbury Park		2
09:59	Seven Sisters		3
10:00		Victoria	1
10:01		Victoria	1
10:01		Pimlico	1
10:02	Finsbury Park		1
10:02		Vauxhall	2
10:02	Finsbury Park		2
10:02	Seven Sisters		2

Table 77 – 2nd October entrance and exit delays

Time	Entry Stations with Delays	Exit Stations with Delays	Minutes Delayed
06:48:00		Finsbury Park	1
06:48:00	Walthamstow Central		2
06:52:00		Finsbury Park	2
06:52:00	Walthamstow Central		1
06:57:00	Blackhorse Road		1
06:57:00	Walthamstow Central		1
07:00:00		Oxford Circus	5
07:00:00	Blackhorse Road		5
07:01:00		Oxford Circus	4
07:01:00	Walthamstow Central		3
07:02:00		Green Park	5
07:02:00	Blackhorse Road		1
07:04:00		Victoria	6

07:04:00	Blackhorse Road		5
07:04:00	Walthamstow Central		6
07:05:00		Oxford Circus	4
07:05:00		Victoria	5
07:05:00	Blackhorse Road		5
07:05:00	Walthamstow		4
07:06:00		Green Park	3
07:06:00	Blackhorse Road		3
07:06:00	Walthamstow		7
07:08:00		Victoria	4
07:08:00	Blackhorse Road		2
07:08:00	Walthamstow		4
07:10:00	Tottenham Hale		5
07:11:00	Blackhorse Road		5
07:14:00	Walthamstow Central		2
07:16:00		Oxford Circus	2
07:16:00	Blackhorse Road		2
08:04:00		Pimlico	2
08:09:00		Warren Street	1
08:09:00	Seven Sisters		1
08:10:00		Victoria	1
08:10:00		Warren Street	1
08:10:00	Finsbury Park		2
08:10:00	Highbury & Islington		2
08:10:00	Seven Sisters		1
08:11:00		Pimlico	4
08:11:00	Seven Sisters		2
08:12:00		Pimlico	2
08:12:00	Seven Sisters		2
08:13:00		Warren Street	2
08:13:00		Green Park	1
08:13:00	Highbury & Islington		1
08:13:00	Seven Sisters		1
08:14:00		Warren Street	1
08:14:00	Highbury & Islington		2
08:14:00	Seven Sisters		1
08:15:00		Oxford Circus	1
08:15:00		Warren Street	2
08:15:00	Finsbury Park		1
08:15:00	Seven Sisters		2
08:16:00		Green Park	1
08:16:00		Victoria	4
08:16:00		Oxford Circus	1
08:16:00		Warren Street	1
08:16:00	Finsbury Park		5
08:16:00	Highbury & Islington		2

08:16:00	Seven Sisters		1
08:17:00		Victoria	3
08:17:00		Warren Street	1
08:17:00		Green Park	3
08:17:00	Finsbury Park		4
08:17:00	Highbury & Islington		1
08:17:00	Seven Sisters		3
08:18:00		Green Park	3
08:18:00	Highbury & Islington		2
08:18:00	Seven Sisters		2
08:19:00		Vauxhall	2
08:19:00		Victoria	3
08:19:00	Finsbury Park		3
08:19:00	Seven Sisters		3
08:20:00		Victoria	2
08:20:00		Pimlico	1
08:20:00		Vauxhall	2
08:20:00	Finsbury Park		2
08:20:00	Highbury & Islington		2
08:20:00	Seven Sisters		2
08:21:00		Victoria	3
08:21:00		Pimlico	3
08:21:00		Vauxhall	4
08:21:00	Finsbury Park		3
08:21:00	Highbury & Islington		3
08:21:00	Seven Sisters		4
08:22:00		Victoria	1
08:22:00		Pimlico	3
08:22:00		Vauxhall	2
08:22:00	Finsbury Park		2
08:22:00	Highbury & Islington		3
08:22:00	Seven Sisters		3
08:23:00		Victoria	2
08:23:00		Vauxhall	2
08:23:00		Pimlico	3
08:23:00	Highbury & Islington		3
08:23:00	Seven Sisters		2
08:24:00		Victoria	1
08:24:00		Pimlico	2
08:24:00		Vauxhall	3
08:24:00	Highbury & Islington		2
08:24:00	Seven Sisters		4
08:25:00		Vauxhall	1
08:26:00		Vauxhall	2
08:34:00		Victoria	1
08:34:00	Walthamstow Central		1

08:35:00		Oxford Circus	1
08:38:00		Victoria	3
08:40:00		Victoria	2
08:40:00	Walthamstow Central		2
08:41:00		Euston	0
08:41:00		Victoria	1
08:41:00	Walthamstow Central		1
08:42:00		Victoria	4
08:42:00		Warren Street	1
08:42:00	Walthamstow Central		6
08:43:00		Warren Street	2
08:43:00		Victoria	8
08:43:00	Walthamstow Central		8
08:45:00		Warren Street	2
08:45:00		Victoria	3
08:45:00	Walthamstow Central		3
08:47:00		Warren Street	3
08:48:00		Warren Street	5
08:48:00		Oxford Circus	1
08:48:00		Victoria	3
08:48:00	Finsbury Park		4
08:48:00	Kings Cross		1
08:48:00	Walthamstow Central		6
08:49:00		Victoria	4
08:49:00		Warren Street	6
08:49:00		Oxford Circus	3
08:49:00	Blackhorse Road		3
08:49:00	Finsbury Park		5
08:49:00	Kings Cross		2
08:50:00		Oxford Circus	4
08:50:00		Victoria	6
08:50:00	Finsbury Park		6
08:50:00	Highbury & Islington		7
08:50:00	Kings Cross		4
08:50:00	Seven Sisters		6
08:50:00	Tottenham Hale		5
08:51:00		Euston	3
08:51:00		Warren Street	7
08:51:00		Victoria	7
08:51:00		Green Paarl	3
08:51:00	Blackhorse Road		3
08:51:00	Highbury & Islington		2
08:52:00		Victoria	8
08:52:00	Highbury & Islington		3
08:52:00	Kings Cross		4
08:52:00	Walthamstow Central		1

08:52:00	Finsbury Park		8
08:53:00		Oxford Circus	7
08:53:00		Seven Sisters	4
08:53:00		Warren Street	3
08:53:00		Victoria	6
08:53:00	Blackhorse Road		7
08:53:00	Finsbury Park		7
08:53:00	Highbury & Islington		7
08:53:00	Seven Sisters		5
08:53:00	Tottenham Hale		2
08:53:00	Walthamstow Central		8
08:54:00		Seven Sisters	4
08:54:00		Warren Street	8
08:54:00		Oxford Circus	9
08:54:00		Victoria	5
08:54:00	Blackhorse Road		5
08:54:00	Highbury & Islington		7
08:54:00	Seven Sisters		6
08:54:00	Walthamstow		10
08:54:00	Tottenham Hale		5
08:55:00		Seven Sisters	3
08:55:00		Warren Street	9
08:55:00		Oxford Circus	8
08:55:00	Blackhorse Road		8
08:55:00	Finsbury Park		9
08:55:00	Highbury & Islington		7
08:55:00	Seven Sisters		5
08:55:00	Tottenham Hale		3
08:55:00	Walthamstow Central		9
08:56:00		Oxford Street	8
08:56:00		Warren Street	9
08:56:00		Highbury & Islington	2
08:56:00		Victoria	3
08:56:00		Seven Sisters	6
08:56:00		Euston	6
08:56:00	Blackhorse Road		7
08:56:00	Euston		2
08:56:00	Highbury & Islington		6
08:56:00	Seven Sisters		8
08:56:00	Kings Cross		2
08:56:00	Tottenham Hale		5
08:56:00	Walthamstow Central		8
08:56:00	Finsbury Park		8
08:57:00		Seven Sisters	6
08:57:00		Warren Street	9
08:57:00		Green Park	7

08:57:00		Oxford Circus	7
08:57:00		Victoria	7
08:57:00	Blackhorse Road		9
08:57:00	Euston		5
08:57:00	Finsbury Park		9
08:57:00	Highbury & Islington		7
08:57:00	Kings Cross		6
08:57:00	Seven Sisters		8
08:57:00	Tottenham Hale		9
08:58:00		Seven Sisters	2
08:58:00		Oxford Circus	6
08:58:00		Warren Street	8
08:58:00		Victoria	7
08:58:00		Kings Cross	5
08:58:00	Euston		7
08:58:00	Highbury & Islington		7
08:58:00	Kings Cross		6
08:58:00	Seven Sisters		8
08:58:00	Walthamstow		3
08:59:00		Seven Sisters	2
08:59:00		Green Park	9
08:59:00		Kings Cross	5
08:59:00		Warren Street	8
08:59:00		Oxford Circus	7
08:59:00		Victoria	8
08:59:00	Blackhorse Road		3
08:59:00	Euston		7
08:59:00	Highbury & Islington		7
08:59:00	Kings Cross		6
08:59:00	Seven Sisters		8
08:59:00	Tottenham Hale		10
08:59:00	Walthamstow Central		6
08:59:00	Finsbury Park		9
09:00:00		Finsbury Park	4
09:00:00		Green Park	13
09:00:00		Kings Cross	13
09:00:00		Victoria	11
09:00:00		Warren Street	8
09:00:00	Blackhorse Road		8
09:00:00	Finsbury Park		10
09:00:00	Highbury & Islington		10
09:00:00	Seven Sisters		6
09:00:00	Tottenham Hale		8
09:00:00	Walthamstow Central		9
09:01:00		Finsbury Park	6
09:01:00		Kings Cross	5

09:01:00		Victoria	13
09:01:00		Vauxhall	7
09:01:00		Warren Street	9
09:01:00		Oxford Street	9
09:01:00		Euston	6
09:01:00	Blackhorse Road		14
09:01:00	Finsbury Park		12
09:01:00	Highbury & Islington		9
09:01:00	Kings Cross		9
09:01:00	Seven Sisters		6
09:01:00	Tottenham Hale		8
09:01:00	Walthamstow Central		12
09:02:00		Euston	7
09:02:00		Finsbury Park	5
09:02:00		Kings Cross	8
09:02:00		Seven Sisters	2
09:02:00		Victoria	11
09:02:00		Warren Street	9
09:02:00		Green Park	6
09:02:00		Oxford Circus	10
09:02:00	Blackhorse Road		6
09:02:00	Euston		9
09:02:00	Finsbury Park		11
09:02:00	Highbury &		10
09:02:00	Kings Cross		6
09:02:00	Seven Sisters		9
09:02:00	Tottenham Hale		7
09:02:00	Walthamstow Central		13
09:03:00		Seven Sisters	1
09:03:00		Warren Street	11
09:03:00		Victoria	10
09:03:00		Euston	9
09:03:00		Green Park	9
09:03:00	Blackhorse Road		4
09:03:00	Finsbury Park		12
09:03:00	Highbury & Islington		10
09:03:00	Seven Sisters		6
09:03:00	Tottenham Hale		10
09:03:00	Walthamstow Central		12
09:04:00		Highbury & Islington	2
09:04:00		Oxford Circus	9
09:04:00		Warren Street	9
09:04:00		Green Park	7
09:04:00		Victoria	9
09:04:00		Kings Cross	8
09:04:00		Euston	9

09:04:00	Blackhorse Road		9
09:04:00	Euston		5
09:04:00	Highbury & Islington		10
09:04:00	Kings Cross		6
09:04:00	Seven Sisters		7
09:04:00	Tottenham Hale		11
09:04:00	Walthamstow Central		9
09:05:00		Kings Cross	5
09:05:00		Victoria	15
09:05:00		Green Park	9
09:05:00		Oxford Circus	11
09:05:00	Blackhorse Road		17
09:05:00	Euston		9
09:05:00	Finsbury Park		15
09:05:00	Highbury & Islington		13
09:05:00	Kings Cross		8
09:05:00	Seven Sister		11
09:05:00	Tottenham Hale		15
09:05:00	Walthamstow Central		12
09:06:00		Green Park	12
09:06:00		Victoria	15
09:06:00		Oxford Circus	10
09:06:00		Warren Street	5
09:06:00		Seven Sisters	4
09:06:00	Blackhorse Road		15
09:06:00	Finsbury Park		16
09:06:00	Highbury & Islington		12
09:06:00	Kings Cross		9
09:06:00	Seven Sisters		14
09:06:00	Tottenham Hale		12
09:06:00	Walthamstow Central		8
09:07:00		Finsbury Park	6
09:07:00		Stockwell	3
09:07:00		Oxford Circus	11
09:07:00		Victoria	16
09:07:00		Green Park	9
09:07:00		Kings Cross	9
09:07:00		Euston	11
09:07:00	Blackhorse Road		6
09:07:00	Euston		6
09:07:00	Highbury & Islington		11
09:07:00	Kings Cross		8
09:07:00	Seven Sisters		10
09:07:00	Tottenham Hale		6
09:07:00	Walthamstow Central		11
09:08:00		Finsbury Park	5

09:08:00		Highbury & Islington	6
09:08:00		Oxford Circus	9
09:08:00		Victoria	10
09:08:00		Green Park	7
09:08:00	Blackhorse Road		6
09:08:00	Euston		7
09:08:00	Highbury & Islington		10
09:08:00	Kings Cross		8
09:08:00	Seven Sisters		9
09:08:00	Tottenham Hale		6
09:08:00	Walthamstow Central		5
09:09:00		Finsbury Park	5
09:09:00		Highbury & Islington	7
09:09:00		Green Park	11
09:09:00		Kings Cross	10
09:09:00		Oxford Circus	9
09:09:00		Victoria	13
09:09:00	Blackhorse Road		9
09:09:00	Euston		6
09:09:00	Finsbury Park		18
09:09:00	Highbury & Islington		13
09:09:00	Kings Cross		9
09:09:00	Seven Sister		13
09:09:00	Tottenham Hale		9
09:09:00	Walthamstow Central		8
09:10:00		Warren Street	13
09:10:00		Oxford Circus	11
09:10:00		Victoria	15
09:10:00		Seven Sisters	5
09:10:00	Finsbury Park		16
09:10:00	Kings Cross		8
09:10:00	Seven Sisters		15
09:10:00	Tottenham Hale		11
09:10:00	Walthamstow Central		12
09:11:00		Kings Cross	8
09:11:00		Highbury & Islington	6
09:11:00		Oxford Circus	12
09:11:00		Seven Sisters	4
09:11:00		Warren Street	14
09:11:00		Victoria	14
09:11:00		Green Park	13
09:11:00	Blackhorse Road		12
09:11:00	Euston		5
09:11:00	Finsbury park		15
09:11:00	Kings Cross		10
09:11:00	Seven Sisters		7

09:11:00	Tottenham Hale		15
09:11:00	Walthamstow Central		16
09:12:00		Green Park	13
09:12:00		Kings Cross	15
09:12:00		Warren Street	13
09:12:00		Oxford Circus	8
09:12:00		Seven Sisters	2
09:12:00	Blackhorse Road		16
09:12:00	Euston		4
09:12:00	Highbury & Islington		14
09:12:00	Kings Cross		8
09:12:00	Seven Sisters		13
09:12:00	Tottenham Hale		7
09:12:00	Walthamstow Central		11
09:13:00		Seven Sisters	3
09:13:00		Green Park	10
09:13:00		Oxford Circus	18
09:13:00	Blackhorse Road		4
09:13:00	Highbury & Islington		17
09:13:00	Kings Cross		8
09:13:00	Tottenham Hale		7
09:13:00	Walthamstow Central		8
09:14:00		Warren Street	14
09:14:00		Victoria	6
09:14:00		Green Park	15
09:14:00		Oxford Circus	19
09:14:00	Seven Sisters		13
09:14:00	Tottenham Hale		16
09:14:00	Walthamstow Central		19
09:15:00		Euston	13
09:15:00		Warren Street	17
09:15:00		Oxford Circus	14
09:15:00		Victoria	17
09:15:00	Blackhorse Road		12
09:15:00	Euston		14
09:15:00	Finsbury Park		20
09:15:00	Highbury & Islington		16
09:15:00	Kings Cross		13
09:15:00	Seven Sisters		17
09:15:00	Tottenham Hale		15
09:15:00	Walthamstow Central		16
09:16:00		Oxford Circus	16
09:16:00		Victoria	17
09:16:00		Warren Street	13
09:16:00		Stockwell	12
09:16:00		Euston	16

09:16:00		Finsbury park	6
09:16:00	Blackhorse Road		20
09:16:00	Euston		15
09:16:00	Finsbury Park		20
09:16:00	Highbury & Islington		14
09:16:00	Kings Cross		10
09:16:00	Seven Sisters		16
09:16:00	Tottenham Hale		11
09:16:00	Walthamstow Central		16
09:17:00		Euston	14
09:17:00		Finsbury Park	5
09:17:00		Kings Cross	12
09:17:00		Oxford Street	18
09:17:00		Victoria	17
09:17:00		Warren Street	18
09:17:00		Green Park	11
09:17:00		Stockwell	9
09:17:00	Blackhorse Road		15
09:17:00	Finsbury Park		19
09:17:00	Highbury & Islington		11
09:17:00	Kings Cross		12
09:17:00	Seven Sisters		15
09:17:00	Tottenham Hale		12
09:17:00	Walthamstow Central		16
09:18:00		Highbury & Islington	6
09:18:00		Warren Street	18
09:18:00		Kings Cross	12
09:18:00	Highbury & Islington		15
09:18:00	Seven Sisters		13
09:18:00	Tottenham Hale		10
09:18:00	Walthamstow Central		10
09:19:00		Highbury & Islington	5
09:19:00		Warren Street	18
09:19:00		Oxford Circus	12
09:19:00		Victoria	18
09:19:00		Green Park	15
09:19:00	Blackhorse Road		11
09:19:00	Euston		9
09:19:00	Finsbury Park		21
09:19:00	Highbury & Islington		16
09:19:00	Kings Cross		16
09:19:00	Seven Sisters		18
09:19:00	Tottenham Hale		22
09:19:00	Walthamstow Central		9
09:20:00		Finsbury Park	4
09:20:00		Kings Cross	17

09:20:00		Oxford Circus	16
09:20:00		Victoria	14
09:20:00		Warren Street	16
09:20:00		Green Park	10
09:20:00		Stockwell	16
09:20:00	Blackhorse Road		11
09:20:00	Finsbury Park		18
09:20:00	Highbury & Islington		18
09:20:00	Kings Cross		13
09:20:00	Seven Sisters		19
09:20:00	Tottenham Hale		6
09:20:00	Walthamstow Central		15
09:21:00		Finsbury Park	4
09:21:00		Green Park	22
09:21:00		Kings Cross	13
09:21:00		Victoria	16
09:21:00		Oxford Circus	21
09:21:00	Blackhorse Road		12
09:21:00	Finsbury Park		20
09:21:00	Highbury & Islington		20
09:21:00	Tottenham Hale		18
09:21:00	Walthamstow Central		11
09:22:00		Green Park	17
09:22:00		Oxford Circus	13
09:22:00		Victoria	12
09:22:00		Kings Cross	9
09:22:00	Blackhorse Road		23
09:22:00	Euston		7
09:22:00	Highbury & Islington		14
09:22:00	Kings Cross		11
09:22:00	Seven Sisters		18
09:22:00	Tottenham Hale		20
09:22:00	Walthamstow Central		11
09:23:00		Euston	17
09:23:00		Warren Street	21
09:23:00		Oxford Circus	19
09:23:00	Blackhorse Road		17
09:23:00	Highbury & Islington		13
09:23:00	Tottenham Hale		20
09:23:00	Walthamstow Central		22
09:24:00		Highbury & Islington	5
09:24:00		Oxford Circus	11
09:24:00		Victoria	14
09:24:00		Kings Cross	15
09:24:00	Euston		6
09:24:00	Highbury & Islington		12

09:24:00	Tottenham Hale		13
09:24:00	Walthamstow Central		10
09:25:00		Green Park	21
09:25:00		Kings Cross	14
09:25:00		Highbury & Islington	6
09:25:00		Warren Street	17
09:25:00	Blackhorse Road		14
09:25:00	Tottenham Hale		16
09:25:00	Walthamstow Central		9
09:26:00		Finsbury Park	3
09:26:00		Green Park	18
09:26:00		Oxford Circus	17
09:26:00		Kings Cross	19
09:26:00		Victoria	24
09:26:00	Blackhorse Road		19
09:26:00	Highbury & Islington		21
09:26:00	Seven Sisters		18
09:26:00	Tottenham Hale		14
09:26:00	Walthamstow Central		21
09:27:00		Victoria	23
09:27:00		Oxford Circus	8
09:27:00		Green Park	16
09:27:00		Warren Street	18
09:27:00		Finsbury Park	3
09:27:00	Euston		19
09:27:00	Finsbury Park		23
09:27:00	Highbury & Islington		27
09:27:00	Kings Cross		11
09:27:00	Seven Sisters		18
09:27:00	Tottenham Hale		16
09:27:00	Walthamstow Central		22
09:28:00		Euston	26
09:28:00		Oxford Circus	13
09:28:00		Victoria	19
09:28:00		Finsbury Park	3
09:28:00	Blackhorse Road		22
09:28:00	Euston		15
09:28:00	Highbury & Islington		25
09:28:00	Tottenham Hale		17
09:28:00	Walthamstow Central		16
09:29:00		Oxford Circus	14
09:29:00		Victoria	19
09:29:00		Green Park	18
09:29:00		Kings Cross	20
09:29:00	Blackhorse Road		20
09:29:00	Highbury & Islington		20

09:29:00	Kings Cross		14
09:29:00	Seven Sisters		24
09:29:00	Tottenham Hale		25
09:29:00	Walthamstow Central		17
09:30:00		Euston	19
09:30:00		Green Park	22
09:30:00		Stockwell	18
09:30:00		Victoria	25
09:30:00	Blackhorse Road		22
09:30:00	Highbury & Islington		18
09:30:00	Kings Cross		13
09:30:00	Seven Sister		24
09:30:00	Tottenham Hale		16
09:30:00	Walthamstow Central		17
09:31:00		Euston	15
09:31:00		Stockwell	17
09:31:00		Victoria	26
09:31:00		Green Park	17
09:31:00	Blackhorse Road		29
09:31:00	Finsbury Park		19
09:31:00	Highbury & Islington		22
09:31:00	Kings Cross		2
09:31:00	Tottenham Hale		28
09:31:00	Walthamstow Central		25
09:32:00		Euston	23
09:32:00		Highbury & Islington	7
09:32:00		Victoria	23
09:32:00		Warren Street	25
09:32:00		Oxford Circus	8
09:32:00		Stockwell	20
09:32:00		Green Park	17
09:32:00	Blackhorse road		22
09:32:00	Euston		11
09:32:00	Finsbury Park		22
09:32:00	Highbury & Islington		27
09:32:00	Kings Cross		11
09:32:00	Seven Sisters		24
09:32:00	Tottenham Hale		28
09:32:00	Walthamstow Central		27
09:33:00		Finsbury Park	4
09:33:00		Highbury & Islington	5
09:33:00		Stockwell	22
09:33:00		Oxford Circus	10
09:33:00		Warren Street	24
09:33:00		Green Park	13
09:33:00		Victoria	31

09:33:00		Euston	26
09:33:00		Kings Cross	15
09:33:00	Blackhorse Road		9
09:33:00	Finsbury Park		22
09:33:00	Highbury & Islington		19
09:33:00	Kings Cross		8
09:33:00	Seven Sisters		24
09:33:00	Tottenham Hale		17
09:33:00	Walthamstow Central		20
09:34:00		Finsbury Park	5
09:34:00		Oxford Circus	23
09:34:00		Victoria	27
09:34:00		Green Park	24
09:34:00		Stockwell	28
09:34:00	Blackhorse Road		5
09:34:00	Finsbury Park		27
09:34:00	Highbury & Islington		27
09:34:00	Kings Cross		13
09:34:00	Seven Sisters		22
09:34:00	Tottenham Hale		15
09:34:00	Walthamstow Central		11
09:35:00		Finsbury Park	4
09:35:00		Oxford Circus	8
09:35:00		Victoria	17
09:35:00		Green Park	7
09:35:00		Stockwell	15
09:35:00		Highbury & Islington	9
09:35:00	Blackhorse Road		9
09:35:00	Euston		12
09:35:00	Highbury & Islington		26
09:35:00	Kings Cross		20
09:35:00	Seven Sisters		22
09:35:00	Tottenham Hale		24
09:35:00	Walthamstow Central		13
09:36:00		Green Park	11
09:36:00		Victoria	29
09:36:00		Warren Street	30
09:36:00		Oxford circus	13
09:36:00	Blackhorse Road		28
09:36:00	Euston		6
09:36:00	Finsbury Park		28
09:36:00	Highbury & Islington		22
09:36:00	Kings Cross LU (Tube)		3
09:36:00	Tottenham Hale		33
09:36:00	Walthamstow Central		22
09:36:00	Warren Street		6

09:37:00		Green Park	8
09:37:00		Highbury & Islington	7
09:37:00		Oxford Circus	19
09:37:00		Warren Street	25
09:37:00		Stockwell	19
09:37:00	Blackhorse Road		26
09:37:00	Euston LU		1
09:37:00	Finsbury Park LU		20
09:37:00	Highbury & Islington		19
09:37:00	Kings Cross LU (Tube)		5
09:37:00	Seven Sisters		21
09:37:00	Tottenham Hale		18
09:37:00	Walthamstow Central		23
09:38:00		Green Park	15
09:38:00		Highbury & Islington	6
09:38:00		Oxford Circus	24
09:38:00		Euston LU	21
09:38:00		Stockwell	23
09:38:00		Kings Cross LU (Tube)	9
09:38:00	Blackhorse Road		14
09:38:00	Highbury & Islington		9
09:38:00	Kings Cross LU (Tube)		15
09:38:00	Seven Sisters		18
09:38:00	Tottenham Hale		13
09:38:00	Walthamstow Central		14
09:39:00		Euston LU	18
09:39:00		Finsbury Park LU	7
09:39:00		Kings Cross LU (North)	11
09:39:00		Victoria LU	22
09:39:00		Highbury & Islington	4
09:39:00		Oxford Circus	11
09:39:00	Blackhorse Road		24
09:39:00	Highbury & Islington		11
09:39:00	Kings Cross LU (Tube)		10
09:39:00	Seven Sisters		18
09:39:00	Walthamstow Central		23
09:40:00		Euston LU	16
09:40:00		Finsbury Park LU	5
09:40:00		Green Park	15
09:40:00		Victoria LU	26
09:40:00		Warren Street	24
09:40:00		Oxford Circus	11
09:40:00		Kings Cross LU (Tube)	15
09:40:00	Blackhorse Road		21
09:40:00	Euston LU		11
09:40:00	Finsbury Park LU		26

09:40:00	Highbury & Islington		19
09:40:00	Seven Sisters		15
09:40:00	Tottenham Hale		23
09:40:00	Walthamstow Central		23
09:41:00		Kings Cross LU (North)	19
09:41:00		Highbury & Islington	13
09:41:00		Oxford Circus	18
09:41:00		Victoria LU	24
09:41:00		Warren Street	21
09:41:00		Stockwell	24
09:41:00		Highbury & Islington	5
09:41:00	Blackhorse Road		19
09:41:00	Euston LU		6
09:41:00	Finsbury Park LU		22
09:41:00	Highbury & Islington		19
09:41:00	Kings Cross LU (Tube)		8
09:41:00	Seven Sisters		19
09:41:00	Tottenham Hale		28
09:41:00	Walthamstow Central		28
09:42:00		Euston LU	20
09:42:00		Green Park	12
09:42:00		Highbury & Islington	7
09:42:00		Oxford Circus	19
09:42:00		Warren Street	22
09:42:00		Victoria LU	29
09:42:00	Blackhorse Road		22
09:42:00	Finsbury Park LU		22
09:42:00	Highbury & Islington		18
09:42:00	Kings Cross LU (Tube)		9
09:42:00	Seven Sisters		24
09:42:00	Tottenham Hale		14
09:42:00	Walthamstow Central		24
09:43:00		Finsbury Park LU	4
09:43:00		Green Park	17
09:43:00		Oxford Circus	18
09:43:00		Kings Cross LU (Tube)	16
09:43:00		Highbury & Islington	2
09:43:00	Blackhorse Road		24
09:43:00	Highbury & Islington		12
09:43:00	Kings Cross LU (Tube)		6
09:43:00	Seven Sisters		7
09:43:00	Tottenham Hale		22
09:43:00	Walthamstow Central		15
09:44:00		Green Park	15
09:44:00		Kings Cross LU (North)	20
09:44:00		Oxford Circus	36

09:44:00	Blackhorse Road		22
09:44:00	Walthamstow Central		29
09:45:00		Green Park	16
09:45:00		Kings Cross LU (North)	18
09:45:00		Victoria LU	13
09:45:00		Warren Street	25
09:45:00		Oxford Circus	24
09:45:00		Euston LU	24
09:45:00		Finsbury Park LU	2
09:45:00	Blackhorse Road		23
09:45:00	Finsbury Park LU		23
09:45:00	Kings Cross LU (Tube)		10
09:45:00	Seven Sisters		3
09:45:00	Tottenham Hale		11
09:45:00	Walthamstow Central		11
09:46:00		Euston LU	15
09:46:00		Finsbury Park LU	3
09:46:00		Highbury & Islington	4
09:46:00		Oxford Circus	25
09:46:00		Stockwell	27
09:46:00		Victoria LU	19
09:46:00		Green Park	6
09:46:00	Blackhorse Road		12
09:46:00	Euston LU		10
09:46:00	Finsbury Park LU		29
09:46:00	Highbury & Islington		17
09:46:00	Kings Cross LU (Tube)		11
09:46:00	Seven Sisters		16
09:46:00	Tottenham Hale		11
09:46:00	Walthamstow Central		19
09:47:00		Euston LU	20
09:47:00		Green Park	14
09:47:00		Highbury & Islington	7
09:47:00		Oxford Circus	19
09:47:00		Victoria LU	18
09:47:00		Warren Street	27
09:47:00	Blackhorse Road		23
09:47:00	Euston LU		9
09:47:00	Finsbury Park LU		21
09:47:00	Highbury & Islington		15
09:47:00	Kings Cross LU (Tube)		8
09:47:00	Seven Sisters		21
09:47:00	Tottenham Hale		19
09:47:00	Walthamstow Central		24
09:48:00		Oxford Circus	18
09:48:00		Warren Street	22

09:48:00		Victoria LU	14
09:48:00		Green Park	10
09:48:00		Stockwell	20
09:48:00	Blackhorse Road		29
09:48:00	Euston LU		4
09:48:00	Highbury & Islington		14
09:48:00	Kings Cross LU (Tube)		5
09:48:00	Seven Sisters		19
09:48:00	Walthamstow Central		31
09:49:00		Kings Cross LU (North)	18
09:49:00		Victoria LU	21
09:49:00		Green Park	13
09:49:00		Warren Street	23
09:49:00	Blackhorse Road		28
09:49:00	Euston LU		12
09:49:00	Finsbury Park LU		21
09:49:00	Highbury & Islington		18
09:49:00	Kings Cross LU (Tube)		11
09:49:00	Tottenham Hale		12
09:49:00	Walthamstow Central		27
09:50:00		Euston LU	15
09:50:00		Finsbury Park LU	2
09:50:00		Green Park	14
09:50:00		Victoria LU	23
09:50:00		Kings Cross LU (North)	11
09:50:00	Blackhorse Road		24
09:50:00	Euston LU		11
09:50:00	Finsbury Park LU		32
09:50:00	Highbury & Islington		22
09:50:00	Kings Cross LU (Tube)		12
09:50:00	Seven Sisters		17
09:50:00	Tottenham Hale		21
09:50:00	Walthamstow Central		27
09:51:00		Euston LU	16
09:51:00		Kings Cross LU (North)	11
09:51:00		Highbury & Islington	2
09:51:00		Oxford Circus	15
09:51:00		Victoria LU	24
09:51:00		Warren Street	19
09:51:00		Stockwell	26
09:51:00		Finsbury Park	1
09:51:00	Blackhorse Road		19
09:51:00	Euston LU		8
09:51:00	Finsbury Park LU		26
09:51:00	Highbury & Islington		18
09:51:00	Kings Cross LU (Tube)		10

09:51:00	Seven Sisters		14
09:51:00	Tottenham Hale		18
09:51:00	Walthamstow Central		26
09:52:00		Green Park	20
09:52:00		Highbury & Islington	1
09:52:00		Oxford Circus	22
09:52:00		Warren Street	22
09:52:00		Victoria LU	14
09:52:00		Finsbury Park LU	2
09:52:00		Highbury & Islington	1
09:52:00		Euston LU	11
09:52:00	Blackhorse Road		19
09:52:00	Finsbury Park LU		14
09:52:00	Highbury & Islington		12
09:52:00	Seven Sisters		4
09:52:00	Tottenham Hale		19
09:52:00	Walthamstow Central		30
09:53:00		Euston LU	11
09:53:00		Green Park	15
09:53:00		Kings Cross LU (North)	11
09:53:00		Warren Street	22
09:53:00		Oxford Circus	17
09:53:00		Victoria LU	25
09:53:00	Blackhorse Road		23
09:53:00	Highbury & Islington		16
09:53:00	Kings Cross LU (Tube)		9
09:53:00	Seven Sisters		16
09:53:00	Tottenham Hale		40
09:53:00	Walthamstow Central		24
09:54:00		Euston LU	14
09:54:00		Green Park	17
09:54:00		Oxford Circus	16
09:54:00		Victoria LU	19
09:54:00		Stockwell	16
09:54:00		Warren Street	16
09:54:00		Kings Cross	10
09:54:00	Blackhorse Road		24
09:54:00	Euston LU		6
09:54:00	Finsbury Park LU		15
09:54:00	Highbury & Islington		16
09:54:00	Kings Cross LU (Tube)		12
09:54:00	Seven Sisters		18
09:54:00	Tottenham Hale		18
09:54:00	Walthamstow Central		29
09:55:00		Euston LU	11
09:55:00		Green Park	22

09:55:00		Oxford Circus	23
09:55:00		Victoria LU	25
09:55:00		Warren Street	16
09:55:00	Blackhorse Road		26
09:55:00	Euston LU		5
09:55:00	Finsbury Park LU		18
09:55:00	Highbury & Islington		10
09:55:00	Kings Cross LU (Tube)		9
09:55:00	Seven Sisters		28
09:55:00	Tottenham Hale		26
09:55:00	Walthamstow Central		26
09:56:00		Euston LU	11
09:56:00		Highbury & Islington	1
09:56:00		Warren Street	17
09:56:00		Oxford Circus	21
09:56:00	Blackhorse Road		8
09:56:00	Euston LU		8
09:56:00	Highbury & Islington		8
09:56:00	Kings Cross LU (Tube)		8
09:56:00	Seven Sisters		14
09:56:00	Tottenham Hale		18
09:56:00	Walthamstow Central		21
09:57:00		Euston LU	12
09:57:00		Oxford Circus	26
09:57:00		Victoria LU	28
09:57:00		Stockwell	26
09:57:00		Green Park	11
09:57:00		Highbury & Islington	0
09:57:00	Blackhorse Road		22
09:57:00	Finsbury Park LU		29
09:57:00	Kings Cross LU (Tube)		9
09:57:00	Seven Sisters		14
09:57:00	Tottenham Hale		19
09:57:00	Walthamstow Central		37
09:58:00		Green Park	13
09:58:00		Kings Cross LU (North)	9
09:58:00		Oxford Circus	15
09:58:00		Victoria LU	23
09:58:00		Stockwell	17
09:58:00		Warren Street	17
09:58:00	Blackhorse Road		33
09:58:00	Euston LU		9
09:58:00	Finsbury Park LU		21
09:58:00	Highbury & Islington		15
09:58:00	Kings Cross LU (Tube)		6
09:58:00	Seven Sisters		21

09:58:00	Tottenham Hale		26
09:58:00	Walthamstow Central		40
09:59:00		Euston LU	10
09:59:00		Green Park	14
09:59:00		Kings Cross LU (North)	9
09:59:00		Oxford Circus	18
09:59:00		Victoria LU	19
09:59:00		Warren Street	17
09:59:00		Finsbury Park LU	1
09:59:00	Blackhorse Road		22
09:59:00	Euston LU		8
09:59:00	Finsbury Park LU		15
09:59:00	Highbury & Islington		12
09:59:00	Kings Cross LU (Tube)		5
09:59:00	Seven Sisters		19
09:59:00	Tottenham Hale		13
09:59:00	Walthamstow Central		18
10:00:00		Euston LU	7
10:00:00		Oxford Circus	18
10:00:00		Victoria LU	40
10:00:00		Warren Street	21
10:00:00		Green Park	12
10:00:00	Blackhorse Road		24
10:00:00	Highbury & Islington		10
10:00:00	Seven Sisters		16
10:00:00	Tottenham Hale		12
10:00:00	Walthamstow Central		22
10:01:00		Victoria LU	27
10:01:00		Warren Street	14
10:01:00		Green Park	11
10:01:00		Stockwell	22
10:01:00		Oxford Circus	26
10:01:00		Kings Cross LU (North)	12
10:01:00	Highbury & Islington		20
10:01:00	Kings Cross LU (Tube)		6
10:01:00	Seven Sisters		18
10:01:00	Tottenham Hale		18
10:01:00	Walthamstow Central		31
10:02:00		Green Park	12
10:02:00		Kings Cross LU (North)	9
10:02:00		Oxford Circus	14
10:02:00		Victoria LU	26
10:02:00		Stockwell	24
10:02:00	Blackhorse Road		26
10:02:00	Finsbury Park LU		25
10:02:00	Highbury & Islington		18

10:02:00	Kings Cross LU (Tube)		4
10:02:00	Tottenham Hale		19
10:02:00	Walthamstow Central		38
10:03:00		Euston LU	6
10:03:00		Green Park	15
10:03:00		Oxford Circus	16
10:03:00		Victoria LU	20
10:03:00		Warren Street	18
10:03:00		Kings Cross LU (North)	6
10:03:00	Blackhorse Road		23
10:03:00	Finsbury Park LU		22
10:03:00	Highbury & Islington		9
10:03:00	Kings Cross LU (Tube)		9
10:03:00	Seven Sisters		16
10:03:00	Tottenham Hale		12
10:03:00	Walthamstow Central		12
10:04:00		Euston LU	5
10:04:00		Oxford Circus	16
10:04:00		Warren Street	5
10:04:00		Green Park	16
10:04:00	Blackhorse Road		24
10:04:00	Finsbury Park LU		9
10:04:00	Highbury & Islington		10
10:04:00	Kings Cross LU (Tube)		3
10:04:00	Seven Sisters		14
10:04:00	Tottenham Hale		17
10:04:00	Walthamstow Central		16
10:05:00		Green Park	14
10:05:00		Oxford Circus	17
10:05:00		Victoria LU	24
10:05:00		Warren Street	19
10:05:00	Blackhorse Road		29
10:05:00	Highbury & Islington		15
10:05:00	Kings Cross LU (Tube)		7
10:05:00	Seven Sisters		24
10:05:00	Tottenham Hale		24
10:05:00	Walthamstow Central		25
10:06:00		Oxford Circus	14
10:06:00		Victoria LU	24
10:06:00		Warren Street	15
10:06:00		Green Park	16
10:06:00		Euston LU	4
10:06:00		Finsbury Park LU	0
10:06:00	Blackhorse Road		32
10:06:00	Euston LU		7
10:06:00	Finsbury Park LU		18

10:06:00	Highbury & Islington		22
10:06:00	Kings Cross LU (Tube)		8
10:06:00	Seven Sisters		22
10:06:00	Tottenham Hale		20
10:06:00	Walthamstow Central		27
10:07:00		Green Park	20
10:07:00		Oxford Circus	13
10:07:00		Victoria LU	27
10:07:00		Warren Street	14
10:07:00		Stockwell	20
10:07:00		Euston LU	5
10:07:00	Blackhorse Road		26
10:07:00	Finsbury Park LU		18
10:07:00	Highbury & Islington		10
10:07:00	Kings Cross LU (Tube)		7
10:07:00	Seven Sisters		21
10:07:00	Tottenham Hale		16
10:07:00	Walthamstow Central		21
10:08:00		Oxford Circus	18
10:08:00		Victoria LU	14
10:08:00		Warren Street	12
10:08:00		Stockwell	18
10:08:00	Blackhorse Road		19
10:08:00	Euston LU		5
10:08:00	Finsbury Park LU		15
10:08:00	Highbury & Islington		10
10:08:00	Kings Cross LU (Tube)		11
10:08:00	Seven Sisters		22
10:08:00	Tottenham Hale		29
10:08:00	Walthamstow Central		20
10:09:00		Green Park	15
10:09:00		Victoria LU	21
10:09:00	Blackhorse Road		31
10:09:00	Highbury & Islington		16
10:09:00	Kings Cross LU (Tube)		9
10:09:00	Seven Sisters		25
10:09:00	Tottenham Hale		36
10:09:00	Walthamstow Central		28
10:10:00		Euston LU	13
10:10:00		Green Park	13
10:10:00		Oxford Circus	10
10:10:00		Victoria LU	27
10:10:00		Warren Street	13
10:10:00	Blackhorse Road		20
10:10:00	Highbury & Islington		12
10:10:00	Kings Cross LU (Tube)		8

10:10:00	Seven Sisters		20
10:10:00	Walthamstow Central		23
10:11:00		Euston LU	5
10:11:00		Oxford Circus	13
10:11:00		Warren Street	8
10:11:00		Victoria LU	25
10:11:00	Blackhorse Road		13
10:11:00	Highbury & Islington		4
10:11:00	Kings Cross LU (Tube)		8
10:11:00	Seven Sisters		17
10:11:00	Tottenham Hale		18
10:11:00	Walthamstow Central		15
10:12:00		Green Park	12
10:12:00		Oxford Circus	11
10:12:00		Warren Street	14
10:12:00	Blackhorse Road		14
10:12:00	Finsbury Park LU		12
10:12:00	Highbury & Islington		4
10:12:00	Seven Sisters		16
10:13:00		Green Park	15
10:13:00		Oxford Circus	12
10:13:00	Seven Sisters		16
10:13:00	Tottenham Hale		17
10:14:00		Green Park	9
10:14:00		Kings Cross LU (North)	2
10:14:00		Oxford Circus	11
10:14:00		Victoria LU	24
10:14:00		Warren Street	11
10:14:00	Blackhorse Road		20
10:14:00	Finsbury Park LU		13
10:14:00	Highbury & Islington		13
10:14:00	Seven Sisters		21
10:14:00	Tottenham Hale		25
10:14:00	Walthamstow Central		22
10:15:00		Oxford Circus	17
10:15:00		Warren Street	12
10:15:00		Victoria LU	25
10:15:00	Blackhorse Road		17
10:15:00	Finsbury Park LU		8
10:15:00	Highbury & Islington		14
10:15:00	Kings Cross LU (Tube)		8
10:15:00	Seven Sisters		21
10:15:00	Tottenham Hale		17
10:15:00	Walthamstow Central		28
10:16:00		Euston LU	5
10:16:00		Victoria LU	17

10:16:00		Oxford Circus	14
10:16:00	Kings Cross LU (Tube)		11
10:16:00	Seven Sisters		16
10:16:00	Tottenham Hale		15
10:16:00	Walthamstow Central		14
10:17:00		Victoria LU	20
10:17:00		Euston LU	3
10:17:00	Blackhorse Road		27
10:17:00	Highbury & Islington		12
10:17:00	Seven Sisters		18
10:17:00	Tottenham Hale		11
10:17:00	Walthamstow Central		17
10:18:00		Green Park	14
10:18:00		Victoria LU	20
10:18:00		Oxford Circus	4
10:18:00		Kings Cross LU (North)	4
10:18:00	Blackhorse Road		21
10:18:00	Euston LU		13
10:18:00	Highbury & Islington		13
10:18:00	Kings Cross LU (Tube)		10
10:18:00	Seven Sisters		20
10:18:00	Tottenham Hale		24
10:18:00	Walthamstow Central		22
10:19:00		Euston LU	3
10:19:00		Green Park	22
10:19:00		Kings Cross LU (North)	3
10:19:00		Oxford Circus	11
10:19:00		Stockwell	25
10:19:00		Victoria LU	24
10:19:00	Blackhorse Road		12
10:19:00	Highbury & Islington		7
10:19:00	Kings Cross LU (Tube)		7
10:19:00	Seven Sisters		19
10:19:00	Tottenham Hale		15
10:19:00	Walthamstow Central		20
10:20:00		Euston LU	4
10:20:00		Oxford Circus	9
10:20:00		Warren Street	9
10:20:00		Victoria LU	29
10:20:00	Blackhorse Road		8
10:20:00	Finsbury Park LU		11
10:20:00	Highbury & Islington		7
10:20:00	Seven Sisters		11
10:20:00	Tottenham Hale		11
10:20:00	Walthamstow Central		13
10:21:00		Victoria LU	23

10:21:00		Warren Street	12
10:21:00	Seven Sisters		23
10:21:00	Finsbury Park LU		6
10:21:00	Walthamstow Central		14
10:22:00		Victoria	21
10:22:00		Kings Cross	7
10:22:00	Blackhorse Road		23
10:22:00	Finsbury Park		14
10:22:00	Highbury & Islington		16
10:22:00	Tottenham Hale		24
10:22:00	Walthamstow Central		15
10:23:00		Victoria	21
10:23:00		Oxford Circus	7
10:23:00	Highbury & Islington		8
10:23:00	Seven Sisters		24
10:23:00	Walthamstow Central		15
10:24:00		Green Park	18
10:24:00		Victoria	22
10:24:00		Stockwell	22
10:24:00	Blackhorse Road		20
10:24:00	Highbury & Islington		11
10:24:00	Seven Sisters		24
10:25:00		Green Park	22
10:25:00		Victoria	26
10:25:00	Blackhorse Road		21
10:25:00	Seven Sisters		30
10:25:00	Tottenham Hale		12
10:26:00		Oxford Circus	9
10:26:00		Victoria	20
10:26:00		Green Park	18
10:26:00		Stockwell	22
10:26:00	Blackhorse Road		20
10:26:00	Kings Cross		11
10:26:00	Seven Sisters		23
10:26:00	Tottenham Hale		20
10:26:00	Walthamstow Central		23
10:27:00		Oxford Circus	11
10:27:00		Victoria	20
10:27:00		Warren Street	12
10:27:00		Euston	3
10:27:00	Finsbury Park		16
10:27:00	Seven Sisters		20
10:27:00	Tottenham Hale		14
10:27:00	Walthamstow Central		17
10:28:00		Oxford Circus	13
10:28:00		Warren Street	9

10:28:00		Euston	6
10:28:00		Victoria	18
10:28:00	Finsbury Park		18
10:28:00	Seven Sisters		14
10:28:00	Tottenham Hale		12
10:28:00	Walthamstow Central		11
10:29:00		Euston	4
10:29:00		Victoria	20
10:29:00	Highbury & Islington		12
10:29:00	Seven Sisters		19
10:29:00	Tottenham Hale		16
10:30:00		Victoria	23
10:30:00	Seven Sisters		23
10:30:00	Walthamstow Central		19
10:31:00		Oxford Circus	11
10:31:00		Warren Street	10
10:31:00	Seven Sisters		10
10:31:00	Walthamstow Central		9
10:32:00		Green Park	12
10:32:00		Oxford Circus	15
10:32:00		Warren Street	8
10:32:00		Pimlico	5
10:32:00	Blackhorse Road		14
10:32:00	Tottenham Hale		13
10:32:00	Walthamstow Central		9
10:33:00		Green Park	12
10:33:00		Victoria	25
10:33:00	Blackhorse Road		17
10:33:00	Seven Sisters		9
10:33:00	Walthamstow Central		13
10:34:00		Oxford Circus	10
10:34:00		Victoria	17
10:34:00		Vauxhall	10
10:34:00	Blackhorse Road		15
10:34:00	Finsbury Park		9
10:34:00	Seven Sisters		15
10:34:00	Tottenham Hale		12
10:34:00	Walthamstow Central		16
10:35:00		Oxford Circus	10
10:35:00		Victoria	16
10:35:00	Blackhorse Road		11
10:35:00	Finsbury Park		14
10:35:00	Seven Sisters		16
10:35:00	Tottenham Hale		14
10:35:00	Walthamstow Central		11
10:36:00		Oxford Circus	8

10:36:00		Victoria	19
10:36:00		Pimlico	8
10:36:00	Blackhorse Road		9
10:36:00	Tottenham Hale		11
10:36:00	Walthamstow Central		11
10:37:00		Victoria	15
10:37:00		Pimlico	15
10:37:00		Vauxhall	21
10:37:00	Blackhorse Road		16
10:37:00	Highbury & Islington		12
10:37:00	Seven Sisters		18
10:37:00	Walthamstow Central		10
10:38:00		Oxford Circus	8
10:38:00		Victoria	14
10:38:00		Vauxhall	12
10:38:00	Blackhorse Road		12
10:38:00	Seven Sisters		19
10:38:00	Tottenham Hale		16
10:38:00	Walthamstow Central		12
10:39:00		Victoria	11
10:39:00		Vauxhall	11
10:39:00	Blackhorse Road		15
10:39:00	Seven Sisters		19
10:39:00	Walthamstow Central		12

Table 78 – Island Line Eastbound average travel times

Entry Station	Entry Station	Exit Station	Exit Station Code	Average Time	Journey Planner Times
Sheung Wan	26	Central	1	8	3
Sheung Wan	26	Admiralty	2	10	5
Sheung Wan	26	Wan Chai	27	12	7
Sheung Wan	26	Causeway Bay	28	15	9
Sheung Wan	26	Tin Hau	29	17	11
Sheung Wan	26	Fortress Hill	30	19	13
Sheung Wan	26	North Point	31	21	15
Sheung Wan	26	Quarry Bay	32	24	17
Sheung Wan	26	Tai Koo	33	23	19

Sheung Wan	26	Sai Wan Ho	34	27	20
Sheung Wan	26	Shau Kei Wan	35	27	22
Sheung Wan	26	Heng Fa Chuen	36	29	24
Sheung Wan	26	Chai Wan	37	32	26
Central	1	Admiralty	2	6	3
Central	1	Wan Chai	27	8	5
Central	1	Causeway Bay	28	11	7
Central	1	Tin Hau	29	13	9
Central	1	Fortress Hill	30	15	11
Central	1	North Point	31	18	13
Central	1	Quarry Bay	32	20	15
Central	1	Tai Koo	33	20	17
Central	1	Sai Wan Ho	34	23	18
Central	1	Shau Kei Wan	35	23	20
Central	1	Heng Fa Chuen	36	25	22
Central	1	Chai Wan	37	28	25
Admiralty	2	Wan Chai	27	7	3
Admiralty	2	Causeway Bay	28	10	5
Admiralty	2	Tin Hau	29	11	7
Admiralty	2	Fortress Hill	30	14	9
Admiralty	2	North Point	31	16	11
Admiralty	2	Quarry Bay	32	18	13
Admiralty	2	Tai Koo	33	18	15
Admiralty	2	Sai Wan Ho	34	22	16
Admiralty	2	Shau Kei Wan	35	21	18
Admiralty	2	Heng Fa Chuen	36	23	20
Admiralty	2	Chai Wan	37	27	23
Wan Chai	27	Causeway Bay	28	9	3

Wan Chai	27	Tin Hau	29	10	5
Wan Chai	27	Fortress Hill	30	12	7
Wan Chai	27	North Point	31	14	9
Wan Chai	27	Quarry Bay	32	16	11
Wan Chai	27	Tai Koo	33	16	13
Wan Chai	27	Sai Wan Ho	34	20	14
Wan Chai	27	Shau Kei Wan	35	20	16
Wan Chai	27	Heng Fa Chuen	36	22	18
Wan Chai	27	Chai Wan	37	25	21
Causeway Bay	28	Tin Hau	29	9	3
Causeway Bay	28	Fortress Hill	30	11	5
Causeway Bay	28	North Point	31	13	7
Causeway Bay	28	Quarry Bay	32	14	9
Causeway Bay	28	Tai Koo	33	14	11
Causeway Bay	28	Sai Wan Ho	34	18	12
Causeway Bay	28	Shau Kei Wan	35	18	14
Causeway Bay	28	Heng Fa Chuen	36	20	16
Causeway Bay	28	Chai Wan	37	23	18
Tin Hau	29	Fortress Hill	30	11	3
Tin Hau	29	North Point	31	11	5
Tin Hau	29	Quarry Bay	32	12	7
Tin Hau	29	Tai Koo	33	12	9
Tin Hau	29	Sai Wan Ho	34	15	10
Tin Hau	29	Shau Kei Wan	35	16	12
Tin Hau	29	Heng Fa Chuen	36	18	14
Tin Hau	29	Chai Wan	37	21	16
Fortress Hill	30	North Point	31	11	3
Fortress Hill	30	Quarry Bay	32	11	5

Fortress Hill	30	Tai Koo	33	11	7
Fortress Hill	30	Sai Wan Ho	34	14	9
Fortress Hill	30	Shau Kei Wan	35	14	10
Fortress Hill	30	Heng Fa Chuen	36	16	12
Fortress Hill	30	Chai Wan	37	19	15
North Point	31	Quarry Bay	32	10	3
North Point	31	Tai Koo	33	9	5
North Point	31	Sai Wan Ho	34	12	7
North Point	31	Shau Kei Wan	35	13	8
North Point	31	Heng Fa Chuen	36	15	11
North Point	31	Chai Wan	37	18	13
Quarry Bay	32	Tai Koo	33	8	3
Quarry Bay	32	Sai Wan Ho	34	11	5
Quarry Bay	32	Shau Kei Wan	35	11	6
Quarry Bay	32	Heng Fa Chuen	36	13	9
Quarry Bay	32	Chai Wan	37	16	11
Tai Koo	33	Sai Wan Ho	34	6	3
Tai Koo	33	Shau Kei Wan	35	7	4
Tai Koo	33	Heng Fa Chuen	36	9	7
Tai Koo	33	Chai Wan	37	12	9
Sai Wan Ho	34	Shau Kei Wan	35	8	3
Sai Wan Ho	34	Heng Fa Chuen	36	9	5
Sai Wan Ho	34	Chai Wan	37	12	7
Shau Kei Wan	35	Heng Fa Chuen	36	6	3
Shau Kei Wan	35	Chai Wan	37	9	6
Heng Fa Chuen	36	Chai Wan	37	7	3

Table 79 – Island Line Westbound average travel times

Entrance Station	Entrance Code	Exit Station	Exit Code	Average Travel Time	Journey Planner Time
Chai Wan	37	Heng Fa Chuen	36	7	3
Chai Wan	37	Shau Kei Wan	35	9	6
Chai Wan	37	Sai Wan Ho	34	11	7
Chai Wan	37	Tai Koo	33	12	9
Chai Wan	37	Quarry Bay	32	16	11
Chai Wan	37	North Point	31	17	13
Chai Wan	37	Fortress Hill	30	20	15
Chai Wan	37	Tin Hau	29	22	16
Chai Wan	37	Causeway Bay	28	24	18
Chai Wan	37	Wan Chai	27	26	21
Chai Wan	37	Admiralty	2	27	23
Chai Wan	37	Central	1	29	25
Chai Wan	37	Sheung Wan	26	32	26
Heng Fa Chuen	36	Shau Kei Wan	35	6	3
Heng Fa Chuen	36	Sai Wan Ho	34	9	5
Heng Fa Chuen	36	Tai Koo	33	9	7
Heng Fa Chuen	36	Quarry Bay	32	14	9
Heng Fa Chuen	36	North Point	31	14	11
Heng Fa Chuen	36	Fortress Hill	30	17	12
Heng Fa Chuen	36	Tin Hau	29	19	14
Heng Fa Chuen	36	Causeway Bay	28	21	16
Heng Fa Chuen	36	Wan Chai	27	23	18
Heng Fa Chuen	36	Admiralty	2	24	20
Heng Fa Chuen	36	Central	1	27	22
Heng Fa Chuen	36	Sheung Wan	26	29	24
Shau Kei Wan	35	Sai Wan Ho	34	8	3
Shau Kei Wan	35	Tai Koo	33	7	4
Shau Kei Wan	35	Quarry Bay	32	11	6
Shau Kei Wan	35	North Point	31	12	8
Shau Kei Wan	35	Fortress Hill	30	14	10
Shau Kei Wan	35	Tin Hau	29	16	12

Shau Kei Wan	35	Causeway Bay	28	18	14
Shau Kei Wan	35	Wan Chai	27	20	16
Shau Kei Wan	35	Admiralty	2	21	18
Shau Kei Wan	35	Central	1	24	20
Shau Kei Wan	35	Sheung Wan	26	26	22
Sai Wan Ho	34	Tai Koo	33	8	3
Sai Wan Ho	34	Quarry Bay	32	11	5
Sai Wan Ho	34	North Point	31	11	7
Sai Wan Ho	34	Fortress Hill	30	13	9
Sai Wan Ho	34	Tin Hau	29	16	10
Sai Wan Ho	34	Causeway Bay	28	17	12
Sai Wan Ho	34	Wan Chai	27	19	14
Sai Wan Ho	34	Admiralty	2	20	16
Sai Wan Ho	34	Central	1	23	18
Sai Wan Ho	34	Sheung Wan	26	26	20
Tai Koo	33	Quarry Bay	32	8	3
Tai Koo	33	North Point	31	9	5
Tai Koo	33	Fortress Hill	30	11	7
Tai Koo	33	Tin Hau	29	13	9
Tai Koo	33	Causeway Bay	28	15	11
Tai Koo	33	Wan Chai	27	17	13
Tai Koo	33	Admiralty	2	18	15
Tai Koo	33	Central	1	21	17
Tai Koo	33	Sheung Wan	26	23	19
Quarry Bay	32	North Point	31	10	3
Quarry Bay	32	Fortress Hill	30	11	5
Quarry Bay	32	Tin Hau	29	13	7
Quarry Bay	32	Causeway Bay	28	15	9
Quarry Bay	32	Wan Chai	27	17	11
Quarry Bay	32	Admiralty	2	18	13
Quarry Bay	32	Central	1	21	15
Quarry Bay	32	Sheung Wan	26	23	17
North Point	31	Fortress Hill	30	8	3
North Point	31	Tin Hau	29	11	5

North Point	31	Causeway Bay	28	12	7
North Point	31	Wan Chai	27	14	9
North Point	31	Admiralty	2	15	11
North Point	31	Central	1	17	13
North Point	31	Sheung Wan	26	20	15
Fortress Hill	30	Tin Hau	29	10	3
Fortress Hill	30	Causeway Bay	28	11	5
Fortress Hill	30	Wan Chai	27	12	7
Fortress Hill	30	Admiralty	2	12	9
Fortress Hill	30	Central	1	16	11
Fortress Hill	30	Sheung Wan	26	18	13
Tin Hau	29	Causeway Bay	28	11	3
Tin Hau	29	Wan Chai	27	11	5
Tin Hau	29	Admiralty	2	12	7
Tin Hau	29	Central	1	15	9
Tin Hau	29	Sheung Wan	26	17	11
Causeway Bay	28	Wan Chai	27	9	3
Causeway Bay	28	Admiralty	2	10	5
Causeway Bay	28	Central	1	13	7
Causeway Bay	28	Sheung Wan	26	15	9
Wan Chai	27	Admiralty	2	10	3
Wan Chai	27	Central	1	11	5
Wan Chai	27	Sheung Wan	26	13	7
Admiralty	2	Central	1	8	3
Admiralty	2	Sheung Wan	26	10	5
Central	1	Sheung Wan	26	9	3