UCL ACCESSIBILITY RESEARCH GROUP

ARGnote Vol. 1 No. 6 2014 ISSN 2054-4561

# Can we produce better information for passengers of the London Underground?

**Emily Digges La Touche**

## Overview

This note aims to show recent findings produced from studying Oyster data, the automated ticketing system used in London Underground. Large data sets were acquired from Transport for London (TfL), the network operators in London. This data was aggregated in different ways and an algorithm was used to determine in different scenarios how much a passenger is delayed, in terms of minutes. The results show the data is a rich source of information about the current dynamics of the network.

## Key Findings

- There is a lack of information available to passengers of the London Underground
- Information that is available to passengers is ambiguous.
- Better information would help passengers plan their journey more efficiently, which would help to reduce delays in the network.
- Oyster data in London is a rich source of information about the travel patterns of passengers.
- Oyster data also gives us information about how the network responds when there are problems to the service.
- This data can be used to produce information for passengers in terms of delays to their travel times when there are delays due to overcrowding or service problems.

## Aims & Objectives

Discover what information about the current service is available through Oyster data and to use this to produce better information to help passengers make more informed decisions about their journey. This information will need to be produced on a network wide scale by giving information to all passengers about the lines and on a passenger scale though a journey planning system.

## Background

In order to determine whether the London Underground network can operate more optimally, by passengers making smarter routing decisions, it is essential that the passengers gain better information than is currently available to them. In London the most relevant information about the service comes directly from the operator (TfL). This information will tell passengers what ‘state’ the service is currently in on each line. The possible states are: Good Service, Minor Delays, Major Delays, Severe Delays and Closure. These statuses are ambiguous and likely to confuse passengers than inform enlighten them to how their journey may be affected.

Smart card data has been studied more frequently in the last decade as a greater number of researchers are interested in looking at big data sets. In London this has led to discovering how many travellers are doing multi-modal

journeys in the city and of what sort[3]. Other research using smart card data includes personalising information for passengers. This looks at how many trips an individual is taking over a time period, travel times and similarities between different user groups[2]. Further it was found, from Oyster data, that only 46-62% of the time that passengers are in the underground network in London is spent riding on a train[1]. This shows that the data is rich with information about passenger's movements while in the system but to date the data has not been used to create passenger information.

## Methods

Data spanning two months was aggregated to look at one underground line. In order to determine what a delay is to the network; whether it is due to congestion or a problem with the service it is essential to know what an un-delayed service looks like in terms of passenger travel times. A database was created with all southbound origin – destination pairs on the Victoria Line and their average un-delayed travel times.

Passengers can experience delay in different places during their journey in the underground. This can be when they enter the system and a queue can form to alight the train, during their journey when they are on the train; this time can be increased when there is a large passenger demand due to possessions being stuck in the doors and the dwell time increasing and finally a queue may form when exiting the station at the ticket barriers due to the exit being a bottleneck. In order to understand these different delays another data set was acquired that contained days with different service statuses and the morning peak congestion. The journey times found on these days were then compared to the average times found. This gave an insight into how the passengers are delayed when there are problems to the service. From this it could be seen how much extra time would be added to a passenger's journey if they travelled during these service delays.

Beyond this an algorithm was created to take the origin – destination travel times on a line during a delay and be able to establish how journeys that may contain interchanges over two lines may be affected. This was completed by breaking the journeys down to their smallest components, then using these components to build up journey times for journeys that there isn't data for.

## Results

The data shows clear rises in passenger travel times during the morning peak with passengers experiencing and average of a 7 minute delay when entering the network between 08:10 and 08:28 and leaving the network between 08:40 and 08:52. The data has shown that there is no constant time delay to passengers for the different status' with severe delays increasing passengers' travel times anywhere between 10 minutes and 25 minutes. It can clearly be seen how a delay affects a passenger's travel time and this can be relayed into information for passengers as either an average line delay or a more accurate delay to a journey. Finally, the travel times found for journeys with their origin – destination pair on the same line were successfully used to gain travel times for all journeys.

## Future Research Areas

1. Determine how richer information with effect passengers' travel behaviour
2. Use the same methodology to understand what information is available from the Octopus Card in Hong Kong.

## Related ARGnotes

Emily Digges La Touche, What information do passengers of the London Underground want and how will they behave if they gain it?, Year: tbd, Vol 1. No.2

Emily Digges La Touche, Can we produce better information for passengers in the Hong Kong MTR? Year: tbd, Vol 1. No.2

**Endnotes:** **1** Chan J.( 2007) *Rail transit OD matrix estimation and journey time reliability metrics using automated fare data* **2** Lathia N, Froehlich J, Capra L(2010*) Mining Public Transport Usage for Personalised Intelligent Transport Systems* **3** Seaborn C, Attanucci J, Wilson N (2009*) Using Smart Card Fare Payment Data To Analyze Multi-Modal Public Transport Journeys in London*

Contact: pamela@ucl.ac.uk ISSN 2054-4561 
Acknowledgements: I would like to thank Nick Tyler and Ben Heydecker for their supervision to this work.