

Transit Explorer:
a bus network analysis tool

T-TRIID Final report

March 2019



**HIGH VOLUME
TRANSPORT**
APPLIED RESEARCH



Transit Explorer: a bus network analysis tool

Version 1.1

March 2019

Produced by:



For:

DFID and IMC Worldwide

Contact:

Mark Dimond

Integrated Transport Planning Ltd.
32a Stoney Street
The Lace Market
Nottingham. NG1 1LL
UNITED KINGDOM

0115 988 6905
dimond@itpworld.net
www.itpworld.net



Project Information Sheet

Client	DfID/IMC Worldwide
Project Code	2732
Project Name	Transit Explorer: a bus network analysis tool
Project Director	Neil Taylor
Project Manager	Dr Mark Dimond
Quality Manager	Ian Stott
Additional Team Members	David Brenig-Jones, James Seery
Sub-Consultants	None
Start Date	20 October 2018
File Location	F:\2700-2799\2732 IMC Worldwide DfID TTRIID Bus Tracker Tool

Document Control Sheet

Ver.	Project Folder	Description	Prep.	Rev.	App.	Date
V1-1	F:\2700-2799\2732	Final report	MD	IS	IS	22/03/2019
V1-0	F:\2700-2799\2732	Draft final report	MD			19/03/2019
V0-5	F:\2700-2799\2732	Draft text	MD			15/02/2019
V0-1	F:\2700-2799\2732	Structure only	MD			09/11/2018

Issue	Status	Author(s)	Reviewed By	IMC Approved By	Issue Date
1	Draft	Mark Dimond	Holger Dalkmann		22/03/2019
2	Final	Mark Dimond	Graham Fiveash	Holger Dalkmann	08/04/2019

Notice

This report has been prepared for DfID and IMC Worldwide in accordance with the terms and conditions of appointment. Integrated Transport Planning Ltd cannot accept any responsibility for any use of or reliance on the contents of this report by any third party.

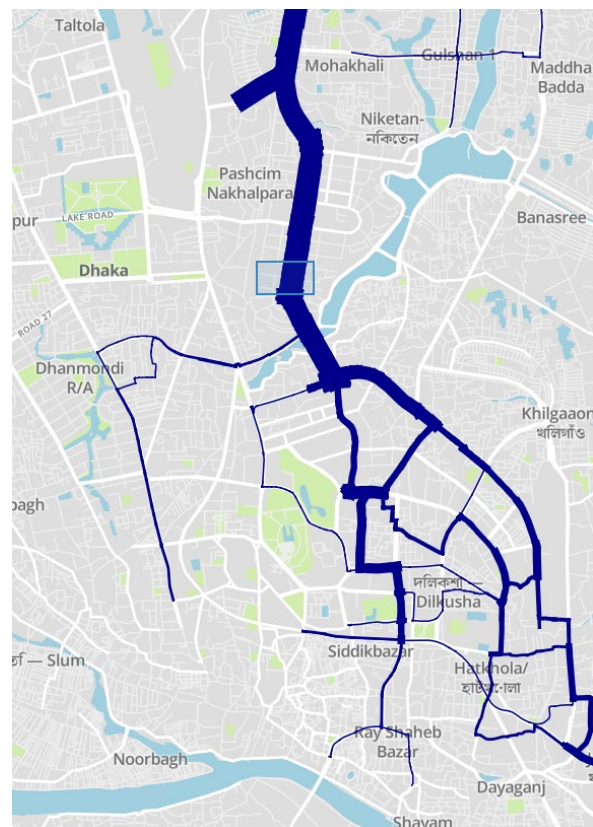
Executive Summary

In October 2018, ITP was awarded one of the first Transport Technology Research and Innovation for International Development (TTRIID) grants by the Department for International Development in the UK. The proposed project envisioned a prototype tool to help in understanding public transport networks in low-income countries, where we have significant experience in helping optimise local transport networks through analysis and monitoring projects.

Cities in developing economies often suffer from unregulated local transport networks, with informal services, high levels of congestion, and poor safety and air quality. Local governing bodies often lack the transport modelling staff and expertise to enforce regulatory controls, meaning that poor transport options continue to have adverse effects upon the local population and economy. Even where local agencies and development bank partners seek to make improvements, understanding the local situation can be challenging without a clear picture of what transport does exist.

The General Transit Feed Specification ([GTFS](#)) has emerged as a popular data format for recording of transport timetables, and is also flexible enough to be used with surveying even where transport is informal. Many cities publish open data GTFS feeds or can use them to power journey planners or analysis applications. [ITP](#) (and its sister company [Conveyal](#)) have pioneered the use of GTFS to analyse city transport networks in different countries, identifying priority areas for investment or intervention. The prototype software tool developed through this project complements existing tools available by matching transport routes to local [OpenStreetMap](#) road networks.

Performing this 'linking' allows aggregation and visualisation across different parts of a city's transport network, identification of bottlenecks and understanding of capacity issues on different routes. This is because unlike other tools, data about routes can be aggregated from a very fine level of detail (the street network 'link' used by each route). The use of more granular data means that the transport situations for individual



corridors or neighbourhoods can be rapidly understood, highlighting congested areas or opportunities for rerouting. The temporal data within GTFS can also be used to filter by time of day and understand different days of the week. Whilst the current prototype does not currently capture demand-side data such as patronage (passenger volume), the software presents this future possibility of better capturing passenger movements through additional use of survey or ticketing data.

The software prototype was developed by ITP over four months, and tested in live project locations including Freetown, Sierra Leone and Dhaka, Bangladesh. GTFS collected through survey work in these locations was loaded and the interactivity of the tool demonstrated, allowing local users to explore public transport at different spatial scales. Testing with local experts (and partners such as development bank teams) helped identify priorities for next steps and the most important further features required to successfully continue development of the tool. This report lists these further features and potential avenues for funding continued development of the tool to maximise the benefits of its use in developing cities, where better understanding of existing infrastructure can help support the case for investment or stronger local regulation.

Table of Contents

1. Introduction.....	7
Outline concept.....	7
Objectives of the project.....	8
2. Project background.....	10
Existing research.....	10
Development at ITP.....	11
3. Implementation	13
Design of the software.....	13
Development priorities	13
Technical implementation	14
Existing technology	14
‘Transit Explorer’	14
User interface.....	15
Limitations	16
4. Demonstration and testing.....	19
Freetown, Sierra Leone.....	19
Dhaka and Chittagong, Bangladesh	20
Other demonstration locations.....	21
Internal testing.....	21
Summary of testing.....	23
5. Exploitation and commercialisation.....	25
Further development: Phase 2 features.....	25
Funding sources and commercial support for further development.....	27
Exploitation next steps.....	28
6. Conclusions.....	30

1. Introduction

- 1.1 Cities in developing countries often suffer from informal, overcrowded and inefficient public transport due to poor infrastructure and lack of regulation. Central to making the case for better transport is proper understanding of existing networks, by taking survey data in different formats or analysing timetables to understand which parts of cities are well served and which are underserved or served inefficiently.
- 1.2 ITP is developing an online tool that will help Local Transport Authorities to better understand both formal and informal public transit networks in their city, and to help provide the information needed to manage and re-plan the network in a way that best suits the city. The 'Transit Explorer' tool is targeted at cities in low-income countries, as city planners in these locations often lack the network knowledge to better coordinate services, and lack the evidence required to design a new regulatory regime that unlocks investment in services and an improvement in standards.
- 1.3 This tool has been developed through funding from the Transport-Technology Research Innovation for International Development (TTRIID) programme, a startup fund provided by DfID in the UK and designed to help develop new technologies at the early stages of deployment. The software prototype has been developed since October 2018. User testing with a number of expert contacts in developing cities has fed into the design of further features, and has identified what additional developments are required to bring the tool to market. In addition, extensive review by ITP's internal project teams has steered development of features, using experience from network optimisation projects in many developing cities.

Outline concept

- 1.4 Over the past 10 years ITP have worked with cities in low-income countries to optimise and formalise their complex public transit networks. A majority of inhabitants in these cities depend on public transport services to access employment, education and services, but they are hindered by transport networks that are informal, fragmented, slow, discriminatory and often unsafe. This stifles personal and economic prosperity and causes environmental damage.
- 1.5 We have learned that a thorough understanding of the market through information and data is key to reform. For a city to a) replan its public transport network, b) develop a new regulatory regime, and c) establish a successful new business model for

operation, local planners must confidently understand the supply and demand arrangement for existing services and identify opportunities for improvement.

- 1.6 The use of open source General Transit Feed Specification (GTFS) data to map and describe the supply of transit services is growing in popularity in both low - and high - income countries. Web resources such as [TransitFeeds.com](https://transitfeeds.com) and [TransitLand](https://transitland.com) list freely accessible GTFS 'feeds', often produced by cities as open data or constructed by volunteers. The Transit Explorer tool will fit with this ecosystem to enable city planners to interrogate and visualise routes in GTFS data.
- 1.7 ITP have previously developed novel technology that enables GTFS data to be 'linked' to OpenStreetMap (OSM) road networks. This data mining algorithm estimates the exact street link used at each point of every bus route in a city, by using route planning algorithms and OSM's free local road network data. Transit Explorer harnesses this technology to enable public transport supply information to be aggregated and disaggregated on a street network level, understanding supply issues at even a street-link level of precision.
- 1.8 The functionality developed will ultimately help provide city planners with a tool to both understand a particular city's 'big picture' (how routes are arranged across the city, where capacity is currently provided, and which areas are underserved) in addition to route-by-route analysis (understanding how many operators and vehicles ply each route, and ultimately tracking operating speeds, daily patronage and fare taking). Finally, interactively performing 'select link' analysis allows rapid understanding of routes that traverse particular corridors, which can have huge value for local neighbourhoods within cities and planning accessibility by different transport modes. For the transport user, this better management of the local transport network helps make the case for infrastructure projects, stronger transport regulation, or improved local environments on major corridors. GTFS can contain either estimates of frequency or fixed timetables, and so querying can help analysts ensure levels of service are appropriate.

Objectives of the project

- 1.9 The project team identified four main objectives at the proposal of the project:
 - 1) develop a prototype interface which allows exploration of public transport usage of local street networks;
 - 2) deploy and test the interface within 1 or 2 active ITP projects, with the assistance of local client contacts in Bangladesh or Sierra Leone; and

- 3) identify, based upon feedback and prioritisation with client contacts, the most important aspects of the tool and plan for next steps;
- 4) to start to improve intelligence for city planners, local experts, and development bank partners about local public transport in cities in low-income countries.

1.10 This report will outline how these have been met and the remaining challenges for the exploitation of the tool. It explains the testing process used, as part of two active projects in Sierra Leone and Bangladesh. In particular, completing the TTRIID project with a clear plan for further development tasks and successful deployment of the tool forms an additional aim, and is the focus of Chapter 5.

2. Project background

Existing research

- 2.1 A significant challenge within public transport modelling and optimisation is accurate capture of the services offered across regions. This can be thought of as formalisation of the timetables that are used in planning (e.g.) bus routes and making them available to the general public. The General Transit Feed Specification (GTFS) was developed in the US in 2005 by Google and TriMet (the transport authority for Portland, Oregon), and first used to power an online journey planner.
- 2.2 The simplicity and flexibility of GTFS has led to it becoming a popular specification for many different kinds of transport planning work, including analysis, as well as public information provision. Transport planning work in developed and developing countries around the world uses it as a basis for capturing travel times, route availability, and mode options, often using the format within GIS or transport modelling software to understand many different aspects of local transport provision.
- 2.3 A key part of this analysis is understanding public transport in relation to its usage of local road networks, counting or aggregating services by road *links* and corridors. However, this process is not simple, requiring a 'map matching' task which has been recognised as a major challenge within GIS analysis for many years¹. Map matching estimates which parts of a local road network correspond with sets of geographic points, for example GPS waypoints or bus stops. Breakthroughs such as a milestone work by Newson and Krumm² made significant progress by applying machine learning techniques, but these have yet to be made feasible using analysis software.
- 2.4 Transport modelling packages such as PTV VISUM make basic map matching possible through plugins and extensions. However, use of these requires considerable expertise and high license fees (which would be inappropriate in many low-income countries), and do not use the commonly available mapping dataset OpenStreetMap.
- 2.5 Alongside the development of GTFS tools at ITP, our sister company Conveyal has also helped spearhead development in use of GTFS for analysis. Conveyal's Analysis tool³ was developed following work in the OpenTripPlanner project, and produces interactive 'travel time' mapping from open data sources. Using this tool, planners can

¹ Quddus, M., Ochieng, W., and Noland, R. (2007): *Current map-matching algorithms for transport applications: State of the art and future research directions*. Transportation Research pt. C.

² Newson, p. and Krumm, J. (2009) *Hidden Markov map-matching through noise and sparseness*. Proceedings of ACM SIGSPATIAL 2009.

³ Conveyal Analysis <https://www.conveyal.com/analysis/>

interactivity and 'exploration' of different questions about the network. This rapid combining of two common open data standards (GTFS and OpenStreetMap), both of which are available in many developing cities, represents the unique selling point of the prototype software, whilst also reducing barriers to *interactive* analysis of these datasets. Having an interactive tool to understand city networks not only requires less training, but permits more detailed, exploratory analysis compared to existing approaches in which several 'static' maps must be produced manually and taking considerable time.

- 2.10 This report highlights how the prototype was developed and initial demonstrations given to colleagues and partners in several low-income countries, including Sierra Leone, Bangladesh and Nigeria. Feedback received has informed selection of several next steps, which will ensure longer term deployment of the tool.

3. Implementation

- 3.1 This chapter outlines the implementation process used in developing the Transit Explorer tool. It builds from the Functional Specification document included in Appendix A and explains the development process used and its rationale. Two active projects in Sierra Leone and Bangladesh, both in partnership with the World Bank, contributed to prioritisation of feature development, and the following chapter identifies how experts these locations contributed to the testing of the tool.

Design of the software

- 3.2 The industry standard approach for software development project management is known as *agile* development: meeting of requirements set by internal and external clients through an iterative process of implementation, demonstration and feedback⁵. This enables the development team to ensure that the most needed features of any product or tool are developed first, whilst ensuring through constant communication that the way they can ultimately be used meets the expectations of the target users.
- 3.3 Successful software development nevertheless also requires a specification in order to formalise planning for initial stages of work and ensuing there is clarity on what will be produced⁶. However, in agile development this can be updated following agreement between the different parties involved.

Development priorities

- 3.4 The initial specification highlighted that the 'minimum viable' product to be produced as this prototype should be an interface that supports three different modes of querying: by route, by region, or for whole cities. Following completion of these features for the tool, a further requirement identified was that *direction* of transport services should be visualised by the tool if possible. This followed internal ITP planning, discussion with development bank and local expert partners, and experience of data analysis in past locations.
- 3.5 However, following successful implementation of the minimum viable features, it was identified that current clients (in particular in an active project with the World Bank in Sierra Leone) had a greater need for accurate implementation of route statistics

⁵ Cockburn, A. (2001) Agile Software Development. Addison-Wesley Professional.

⁶ Spolsky, J. (2000) Painless Functional Specifications – Part 1: Why Bother? ('Joel on Software' blog)

(vehicle and seat counts) and so implementing this initially became a higher priority than visualising vehicle direction.

Technical implementation

- 3.6 This section outlines the production of the new tool, the technologies used, and how they work. It should be emphasised that the new tool interacts with existing technology developed by ITP in 2015-16, which is described below for background information. In some cases, the existing technology has had to be extended to accommodate different types of queries within the new tool; where this has happened, it has been documented in the following section.

Existing technology

- 3.7 The existing 'GTFS Linker' tool is built around a PostGIS geodatabase used to store the OpenStreetMap data for a particular city. This must be converted to routable *nodes* and *links* so that the map-matching process can take place. This uses a tool called *osm2pgRouting*, which produces a data structure that can be used by the open-source *pgRouting* route finding framework. An editing and correction API was produced, developed in Python and Django.

'Transit Explorer'

- 3.8 The new analysis tool was developed as a separate 'front-end' user interface, connecting to the *pgRouting* database through Django in order to retrieve aggregated *trip pattern alignments*, which are the most important part of the 'linked GTFS' data structure previously produced by ITP. These capture the exact routes and road alignments of buses and other transport services within GTFS for a particular city.
- 3.9 The tool was developed as a JavaScript application using the Mithril frontend framework, which organises and simplifies user interface software development and is particularly suited to small software teams⁷. Presentation of geographic data was managed through the Mapbox GL JS plugin, which provides a modern, high-performance way of showing large volumes of road network data for a city using technique known as *vector tiles*. An overview of how this approach works and why it is so suitable for 'big' geodata presentation is available through the Mapbox website⁸.

⁷ Mithril JS <https://mithril.js.org/>

⁸ Mapbox (2016) Vector Tiles reference and specification. <https://docs.mapbox.com/vector-tiles/reference/>

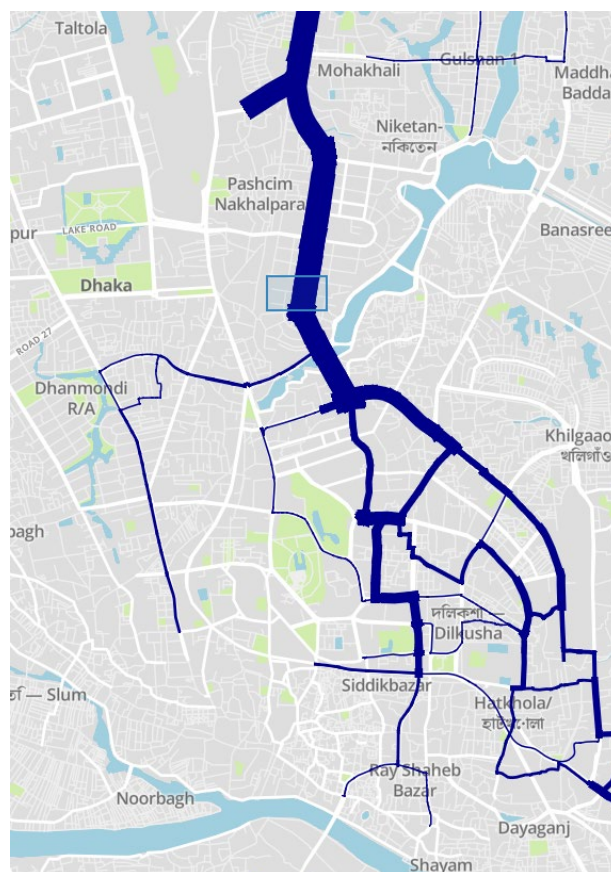
- 3.10 Presenting data using the Mapbox GL plugin allows a significantly volume of city road networks to be drawn and restyled within a web browser interface, making a high-performance interactive tool feasible. This performance is significantly better than other alternatives, in particular the Leaflet JS and OpenLayers frameworks, the former of which had been used for the 'editing' interfaces in ITP's existing GTFS technology. This is because Mapbox GL is able to use computer hardware acceleration (typically used for computer gaming) in presenting large quantities of data, whilst other frameworks rely only on JavaScript technology.

User interface

- 3.11 Presentation of an attractive and easily understandable user interface (UI) is one of the biggest challenges for any project. As highlighted above, the Transit Explorer tool is predominantly an interface to existing 'linked GTFS' technology developed at ITP. Presenting transport network data to users in a fashion that is immediately clear (without need for training) is therefore the central challenge of the project. The UI developed for the tool is divided into a map view, query options and a route list.
- The **map view** shows background mapping produced by Mapbox and based on OpenStreetMap data. This is also used as the basis for the geographic visualisation, with lines of different thickness and colour superimposed to represent route lines within the GTFS feed and different statistics relating to them.
 - Clicking on the map or drawing a 'selection' box around specific corridors returns a summary of routes that travel through that corridor, with width proportionate to a selected statistic at each road segment. For example, if the statistic the user is interested in is 'trip count' (the simplest form of GTFS network analysis), thicker lines will be displayed for routes that have greater numbers of trips from that corridor. An example is shown in Figure 3-1.
 - The **query options** control selection of an appropriate transport route statistic, and are situated alongside other items such as the 'Load all' button which calculates statistics for a whole feed. This set of controls also contains a link which takes a user to a feed selection page, allowing them to navigate between viewing GTFS data for different cities.
 - The statistics available in the current prototype include trip count, vehicle count, and seat capacity, each of which can be visualised on each link in a city's road network. A full list of the possible statistics that could ultimately be calculated by the tool is available in Appendix B

- The options also allow selection of specific times and dates of at which to query the feed, controlling which GTFS trips are used to calculate specific statistics. This is useful because users may be particularly interested in public transport at a particular time of day or day of the week (and a city cannot be assumed to have an 'average' level of service at any particular point).
- The **route list**, below the query options, shows the list of routes in the GTFS feed. Clicking on each route's headsign loads the roads traversed by that route, which are drawn on the map and styled by the appropriate statistic for other routes on each segment. (For example, if a route is particularly prone to congestion at a specific intersection on a corridor, this can be indicated by thicker lines at that point).

Figure 3-1 Interactive selection to view corridor routes



3.12 Figure 3-2 below shows a screenshot of the current prototype of the tool, displaying GTFS data from a recent ITP project in Freetown, Sierra Leone. The tool can also be accessed interactively at <http://transitexplorer.net>.

Limitations

- 3.13 The prototype developed capitalises on the simplicity and the flexibility of the GTFS format to summarise public transport in a city. However, GTFS feeds are intended to operate as timetables, and do not straightforwardly consider a concept of 'average' levels of public transport within a city. Analysis must therefore be limited to the period of validity of a GTFS feed (even if the feed itself is a survey-based approximation, where in the majority of cities in low-income countries transport is arranged informally).
- 3.14 The prototype developed is also exclusively transport supply-oriented: it does not consider levels of use of specific vehicles by passengers. This may be addressed in future iterations, as discussed in Chapter 5. Finally, there is no integration with the

current prototype and 'live' vehicle tracking systems (ITS/AVL), though this may be feasible in future iterations.

- 3.15 The limitations of the current prototype do not prevent local experts and transport planners using the tool to gain a good understanding of public transport supply in a particular city, though the use of GTFS datasets means that the 'picture' of understanding is related to the date of capture for the dataset used. The informal nature of most transport in low-income cities may mean that passenger volume can at least partly be understood without patronage surveys, as the supply of informal transport options is typically very reactive.

4. Demonstration and testing

- 4.1 To maximise the value of the proposed product to low-income countries, thorough user testing was performed within ITP, with partner contacts in funding bodies (in particular the World Bank) and with local experts in project locations including Sierra Leone, Bangladesh, and Nigeria. These contacts included local operational experts, academics, regulatory consultants and prominent city politicians. This chapter will document the initial demonstrations and testing that took place and explain how the feedback received will be incorporated.
- 4.2 Unfortunately, the timescales of parallel ITP projects did not permit demonstration and testing during in-country missions to the envisaged project locations. However this was mitigated by collaboration through teleconferencing and additional internal review. In addition, ITP will continue to perform testing and seek feedback through ongoing projects to maximise the value of the tool to cities in developing economies.
- 4.3 This section will summarise the project locations where demonstrations took place, before giving an overview of the feedback received and the prioritisation of the next steps arising from this.

Freetown, Sierra Leone

- 4.4 An initial demonstration and testing were carried in conjunction with ITP's partner team at the World Bank, local expert Professor Badamasi Savage and a local data collection team. The team used the tool over a number of weeks to explore the different functionality available and understand what was possible using the prototype. The feedback received was collated and is shown in Table 4-1 below. The tool was also demonstrated to Prof. Badamasi's team and to World Bank partners locally and internationally, providing an overview of the possibilities presented by 'linked' GTFS data and highlighting the major corridors in Freetown's different neighbourhoods.
- 4.5 Forthcoming potential work in Freetown, in partnership with the World Bank, will seek to understand public transport supply and demand data in the city and the options for optimisation on specific corridors. The tool prototype would be instrumental in delivering analysis as part of this work, and further development (beyond the prototype stage) may be necessary to fully realise its benefit. Further information about approaches to exploitation is collated in Chapter 5.
- 4.6 The Transit Explorer tool was also briefly demonstrated in person to the Mayor of Freetown and the Sierra Leone government's Minister of Transportation at an urban

transport workshop at University College, London (see Figure 4-1). Both expressed significant interest in the tool and are in continued contact with ITP in order to give additional feedback.

Figure 4-1 Yvonne Aki-Sawyerr, Mayor of Freetown, Sierra Leone, addressing a sustainable urban transport workshop at UCL



Dhaka and Chittagong, Bangladesh

- 4.7 ITP has undertaken capacity building and public transport planning work in both of Bangladesh's largest cities. In mid-2018, the company delivered a strategic urban master plan for Chittagong, and completed pre-feasibility studies for two significant transport projects. ITP's work in Bangladesh is expected to continue, building on strong links with local partners.
- 4.8 One of these partners is a respected independent local transport consultant with longstanding links to the Dhaka Transport Coordination Authority (DTCA). Our partner has also worked with ITP to deliver significant projects in Bangladesh's second city, Chittagong. As part of the testing of the tool, teleconference demonstrations and local testing were performed by the consultant, the results of which are summarised below. The main feedback received from this colleague stressed the importance of ease of use for advisory staff with differing levels of technical expertise, and the possibility of including more statistics and data for helping with management of route franchising (for example route lengths and vehicle types).
- 4.9 Unfortunately it was not possible to arrange demonstration directly with the DTCA in the timescale of this TTRIID work, but this will continue to be sought in order to help

maximise the value of the tool in cities in southern Asia. ITP staff are planning to visit Dhaka in April 2019, which may provide opportunity for demonstration and feedback.

Other demonstration locations

- 4.10 Further demonstrations were conducted in other locations where ITP has potential for future work but does not have recent active projects. In Indonesia, the tool was extensively demonstrated to a World Bank team leader interested in GTFS data tools for projects in major cities in the country. The tool was also briefly demonstrated to senior staff at LAMATA, the transit authority for Lagos, Nigeria. The feedback received from each of these locations is collated in Table 4-1 below. Finally, confirmed ITP work in the Philippines will further test the prototype in April 2019.

Internal testing

- 4.11 As ITP delivers many different types of public transport optimisation project, it was important to use the benefit of internal expertise from the international team to ensure the developed prototype meets the needs of future projects. The use of software tools developed in-house requires careful consideration of both features required and business model. Internal feedback addressing the functionality of the tool is captured below, whilst future business models are subject to further discussion in Chapter 5.
- 4.12 The most important internal feedback received stressed the importance of the ability to calculate different statistics for transport routes and individual road links, helping deliver calculations for potential future work. The need for some of these to be displayed in the 'route information' box (allowing sorting by statistics for routes) was also raised. Finally, an important piece of further development proposed was the ability to make minor changes to GTFS routes within the interface, allowing visualisation of route proposals etc., though this would represent significant further development effort.

Table 4-1: Feedback received from demonstration and testing of the prototype tool

	Sierra Leone	Bangladesh	Indonesia	Nigeria	Internal review
Staff	Local experts, development bank staff	Local expert	Development bank staff	Local experts	International team project managers
What questions do you want to answer about the transport network?	What are the major corridors? Where can people access from specific locations? Type/size of vehicles.	Total route lengths? Service by time periods? What vehicles are used to serve routes?	Peak hour statistics by type of service Type/size of vehicles. Directionality of specific routes	Where vehicles are at any given point. Planning of new routes	Spatial joins between routes + networks to see busiest corridors 'Select link' analysis to understand specific sites e.g. terminals
Are there any parts of the User Interface you do not understand?	What is the timescale of the data? Is it intended to interact with ITS?	What vehicles does this include? Could I filter by mode?	Can't find relevant trips (date setting etc.) Wording should be clarified (what is a 'trip'?)	Much of the transport in Lagos is informal. Is this captured? How?	Statistics are displayed by hover but could be in legend or annotated Scaling of lines to visualise statistics could be clearer
Are there other features that would be useful to you?	Does this show patronage/demand? Could it? (and what data would be necessary?)	Could we view different modes and types of service (intercity/local)? Total distances for route each day	Annotation of statistics instead of 'hover' feature Whole-route statistics such as length, text description? Permit data for franchises	Is it possible to integrate with ITS? Could this show demand or patronage?	Show different types of query in GTFS feed at same time More data in route list box (vehicles, frequencies, capacities, route lengths).
Other feedback and questions	Who is the target user of the tool? Can it be made open source?	Needs a tool for helping to manage franchising. How could this tool help? Locations and status of terminals/depots	Lots of projects really need this but getting GTFS data in cities is often a problem. What is the business model and how would that fit with WB open source reqs.?	Where does the data come from? What time period was it collected? ITS and patronage data would both be very useful	Interactivity is a big bonus over existing tools/techniques Easier upload of base data required to use for projects Editing and comparison of GTFS

Summary of testing

- 4.13 In addition to the most important feedback received, summarised above, more detailed testing data and 'bug' reports were received from the numerous demonstrations of the software. After analysis these were incorporated to the software's issue log (managed on [GitHub](#)) for tracking and prioritisation.
- 4.14 The feedback received was collated and analysed to determine the broad themes and the most important issues facing potential future users of the tool. Table 4-2 below summarises these, including both specific functionality issues and broader questions for consideration in ongoing development. These have been marked where issues have already been addressed ('fixed'), are the subject of current software development effort ('in progress') or require further evaluation and discussion with partners ('ongoing'). Those items falling into the second category are expected to be developed alongside current ITP projects, whilst the latter may require further funding or specific investment.
- 4.15 The following chapter prioritises the remaining feedback to determine the next steps of software development.

Figure 4-2: Viewing transport statistics on specific local road links

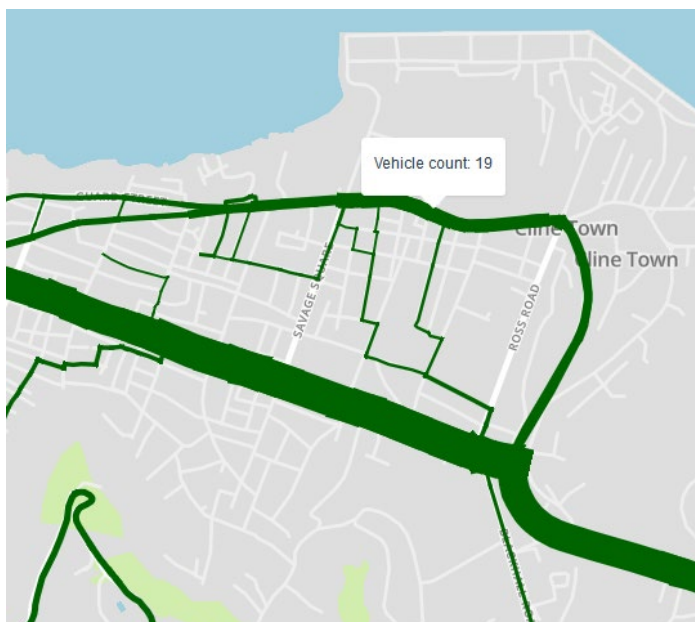


Table 4-2: Prioritised feedback received on the tool interface

Feedback	Status
Show actual statistics: trip count, vehicles, seats etc. in interface (instead of just thickness of line which is relative)	Fixed (see Figure 4-2)
Make the 'settings' box removable to aid production of screenshots showing summary data	Fixed
User interface (UI) tweaks to aid understanding of GTFS feeds for non-experts. (Numerous questions related to the time period of surveying or data validity, some of which can be addressed with better UI design.)	In progress
Better statistics about routes including route length, frequency, operating hours and neighbourhoods served.	In progress
Display of more information on different transport modes and vehicle capacities in the tool interface screen. This includes the ability to 'set' approximate vehicle capacities depending upon local transport characteristics	In progress
Operator details could be included if data available. This could be related to information about different contracts, route lengths etc.	Ongoing
Who is the intended user of the tool? The exact features needed depends significantly upon whether further software to be developed is intended for planners, franchise regulators, or transport operators.	Ongoing
Directionality of routes. Initially proposed as a high-priority feature for ITP's data science staff, intuitive visualisation of inbound and outbound routes on the same corridor was only raised once by testers. However it may become higher priority depending upon future projects.	Ongoing
Which further data sources could be included? In particular some cities have (or are considering) instrumentation and ITS systems; meanwhile patronage data is needed to fully understand usage of the local network and is often collected through ticketing or dedicated tools such as the TransitWand app. Effectively developing the most valuable tool for ITP's clients depends upon answering point 1) above accurately.	Ongoing

5. Exploitation and commercialisation

- 5.1 This chapter summarises a plan for how the prototype developed as part of the TTRIID project will be extended and commercialised. This is captured through overview requirements for further functionality, which would make the product more valuable in public transport projects, outlined in the first section. A plan for different avenues to financially support continued development is included in the second section.

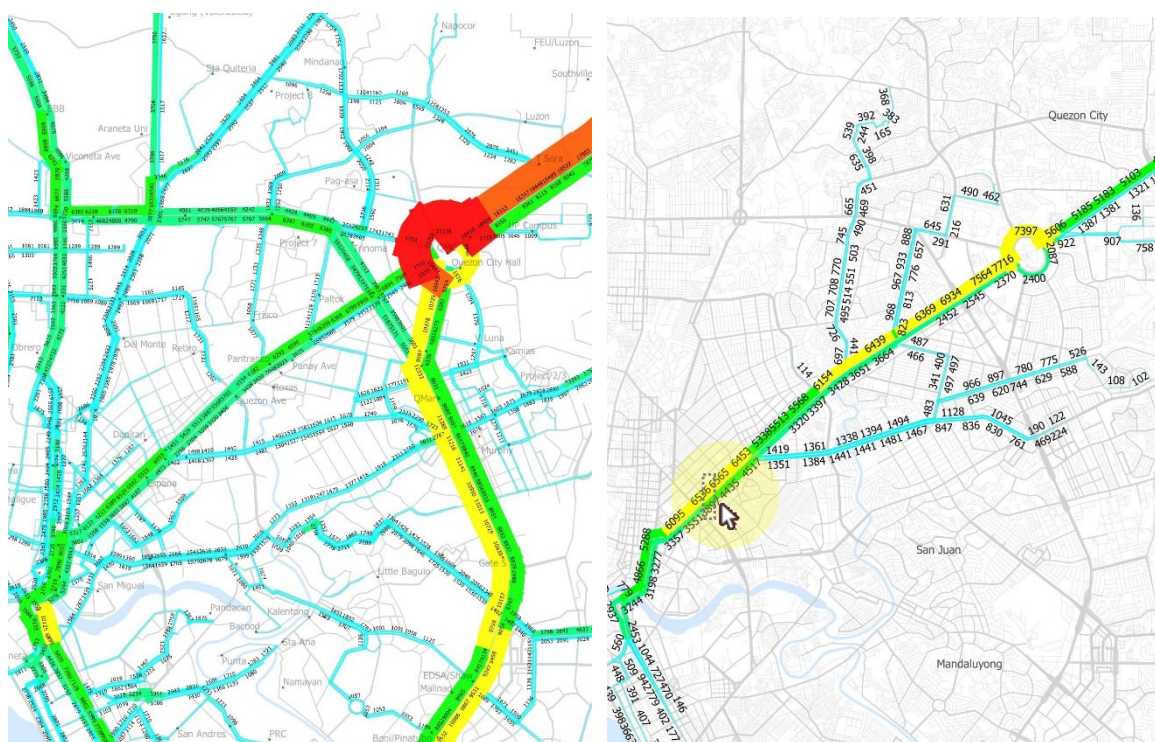
Further development: Phase 2 features

- 5.2 Phase 1 development of the Transit Explorer tool has built the aggregation and visualisation technology to enable the visualisation and exploration of transit supply data. Further development under phase 2 could be of great value for project locations through application of features identified in Chapter 4:
- Incorporate additional route-specific data into the on-screen route list. This might include different statistics depending upon the 'target' user of the prototype tool: number of franchises, number of vehicles, vehicle size, subsidy status and route length would be of great use to regulators and operating authorities, whilst vehicle capacity information, speeds, energy use, and CO2 emissions may be of value to regional governments interested in maximising the quality of the local environment. In particular, understanding production of vehicle-based emissions from vehicles locally could provide significant benefit to cities whilst being straightforward to calculate.
 - Incorporation and visualisation of speed data (from on-board GPS surveys or AVL datasets). Areas of low speed, but high seat count/passenger flow could help identify where bus priority measures will be beneficial, and associated calculation of economic benefits. This can be either 'live' (ITS) or through post-hoc surveys.
 - A simple in-browser GTFS route editing tool, so that either operational or governmental users (or associated partners such as development bank staff) can easily design and view new route proposals within the Transit Explorer tool.
- 5.3 Phase 2 is also intended to build on the same technology to incorporate *passenger load* data. This could enable operational authorities to understand how the transit network is being used by passengers, and is sometimes thought of as the 'demand' for public transport.
- 5.4 Passenger load data can be collected manually (using survey apps such as EpiCollect or TransitWand, or paper based 'on-off' surveys) or automatically from ticket machine

data. In low income countries, manual surveys are the most common source of passenger count data as ticket machines are not common and the adoption of e-ticketing is slow. While this can take many months, surveys are affordable due to lower labour costs. For example, companies such as ITP often undertake manual city-wide passenger load surveys of transit systems through different methods. However processing of this data requires significant expertise and can be very time consuming.

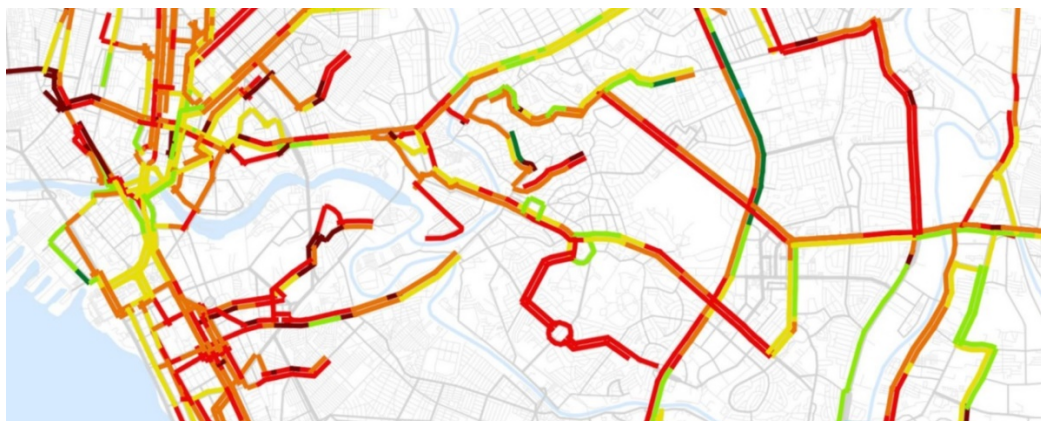
- 5.5 Figure 5-1 gives an example of passenger flow visualisation: GIS outputs produced for passenger load in Manila. As with public transport supply data, producing such outputs is currently a time-consuming statistical analysis and GIS task, and so improving the process with the use of 'linked' GTFS data would greatly improve the speed with which optimisation projects could be delivered.

Figure 5-1: Example diagram of passenger load of all routes in Manila, and a selection of routes made on Quezon Avenue (Data for AM Peak hour)



- 5.6 Passenger flow data can also be compared against supply to create *seat utilisation* maps – helping to identify network inefficiency where buses or vehicles are travelling but not carrying passengers. The data table (route list) will be expanded to include daily patronage, daily passengers per vehicle, seat utilisation, average passenger distance and expected fare taken.

Figure 5-2: Speed data from on-bus GPS survey



- 5.7 Finally, there is scope for further evolution of the tool through integration of socio-economic data such as census data, population, employment, and deprivation indicators. Origin-destination trip lengths can be used to identify journeys that are suitable for public transport trips but currently made by other modes or very slow. This would particularly have benefit for local and city governments in low-income countries looking to improve mobility with optimised routes or new infrastructure.

Funding sources and commercial support for further development

- 5.8 A number of approaches can be considered for developing some or all of the features outlined above, and continuing research into the best way to exploit 'linked' GTFS data and capitalise upon ITP's technology in this area. It is expected that each of the following avenues must be explored concurrently in order to maximise this.
- **DfID's High Volume Transport** programme is expected to provide further funding support for research and development into urban transport optimisation in developing countries. At the time of writing, this is expected to be announced in mid-2019, meaning that there may be a short gap in which development of the technology may have to be paused or funded from other sources.
 - Funding awarded from DfID to date has required that work benefits one of the designated Low-Income Countries (LICs). Though several of ITP's current projects are in these LICs, a significant number of past projects and proposed new work are not, in particular in Ukraine and the Philippines. The need to meet these criteria must therefore be recognised as a moderate challenge aligning research work in partnership with DfID with ITP's project locations.

- **Innovate UK and the Department for Transport** offer several kinds of research and development funding under different topic areas, through schemes such as the Innovation Challenge Fund and the T-TRIG grants. Some of these topic areas are specific to the UK, whilst others are broader. It is common for research funding awarded by Innovate UK, in particular for more market-ready technology, to only provide partial funding of up to around 70% of research costs.
- The European Union's **Horizon 2020** programme offers research funding in a wide range of topic areas for transport, and is available to most types of organisations willing to partner across member states. The UK's exit from the EU in 2019 is not expected to present immediate issues in securing funding as it is committed to ongoing contributions, including for projects awarded in 2019-2020.
- ITP's **existing development bank partners**, including bodies such as the World Bank and the EBRD, have often procured public transport network optimisation using traditional transport planning methods of statistical analysis and GIS. An appropriate partnering arrangement in a 'test' location could provide opportunity for further development of the software, provided agreement could be reached on issues such as the risks associated with technical research and development and long-term ownership of intellectual property (IP). Nevertheless, there may be scope for collaboration to partially fund future work where some parts of any new technology can be made open source⁹ (often a requirement of development bank funding) or permanently licensed.
- A **software product** could ultimately be delivered either on a software-as-a-service (SaaS) basis, or as an open-source tool for which training would be provided through ITP's consultancy business until established enough to sell directly to local authorities in low-income economies at lower cost. It is also possible to combine these options, with open-source software 'managed' through consultancy and training. These approaches present some challenges, in particular including maintenance of software to meet acceptable service-level agreements, and ensuring any user interface is of high enough standard for external use.

Exploitation next steps

- 5.9 Developing a tool that is of high enough quality to sell through affordable SaaS has the additional benefit that it improves ITP's ability to provide high quality analysis

⁹ Open source software is an approach by which the technical implementation of a piece of software is published freely, and hence generally available to both collaborators and competitors free of charge. Though open-source licensing for software can initially incur a modest cost of intellectual property (particularly where the design might be regarded as highly novel) there are potential significant benefits where software is maintained and contributed to by a much larger group of users.

through consultancy. However, it should be noted that the need for Transit Explorer is partly due to the expense and complexity of similar solutions, so providing 'another expensive tool' may be of little value in project locations. ITP may however be able to develop further business in training on a simple-to-use new tool, whilst selling access to analysis software at an affordable price. This should also include further addressing the 'intended user' question, which will in turn help to better prioritise the required functionality.

- 5.10 The UK Knowledge Transfer Network (KTN) also provides advice on improving impact and beginning successful commercialisation for new technology developed through funding such as TTRIID. Planning the next phase of this work will require using the benefit of this expertise as far as possible.

6. Conclusions

- 6.1 This report outlines the development of a prototype visualisation tool for improving understanding of public transport in low income countries. Better understanding of informal and poorly regulated services forms a first step towards making the case for improved infrastructure or stronger local management. The prototype developed uses the GTFS timetable feed format 'linked' to free OpenStreetMap road mapping data. The technology to perform this linking had already been developed through previous projects, but required significant data science skills to understand and analyse. The interactive tool allows visualisation of all road-based public transport for a given town or city, and can be filtered to specific roads or times of day. Answering questions about locations, service levels and timings of buses and other local transport is very important to regulators and management teams locally, with our underlying technology offering significant opportunity for exploitation and future work.
- 6.2 The project objectives were largely met:
- 1) A usable, interactive tool proving the concept and demonstrating the value in two different project locations (Freetown, Sierra Leone and Dhaka, Bangladesh) has been produced. This tool was developed using the Mapbox GL web mapping framework, meaning that the mapping produced is high quality and further functionality can be rapidly built using the framework's flexible set of features.
 - 2) Testing through ITP's existing and former partners was undertaken to understand the potential value of the tool and possible further user needs. Unfortunately, this was not undertaken through 'in-country' missions as initially envisaged, but was instead completed through teleconferencing demonstrations and desktop testing by partners. ITP will continue to pursue formal in-country testing of the tool in active projects, to ensure new functionality developed is of maximum value to clients.
 - 3) Next steps for features have been identified based upon extensive testing of the prototype, with a comprehensive list of potential further features outlined above.
 - 4) Testing has shown that the use of 'linked' GTFS data is of significant interest to both planning staff in low-income countries and development bank partners, improving the ability to rapidly understand local public transport networks.
- 6.3 The choice of forthcoming features for implementation depends to some extent upon the type of commercialisation sought and the different kinds of projects targeted for use of the tool. Commercialisation or next steps may include seeking further research

and development funding; implementation of additional features in partnership with clients, or development of a business strategy around 'training' with the tool. Selecting the most appropriate of these is the focus of ongoing strategic planning within ITP. This will also further evaluate the different models for sale of the visualisation tool, ranging from its release as open source (and provision of related training and management services) to sale as a proprietary product.

- 6.4 This will build from the successful demonstration of a prototype tool using modern web technology to analyse public transport by exploiting free open data in the GTFS and OpenStreetMap formats.

Appendix A: Functional Specification

Title	Network Explorer – Functional Specification
Date	18/03/2019
Author(s)	Mark Dimond
Project Code	2732
Version	1.3-WORKING DOCUMENT



1. Introduction

- 1.1 This document outlines the existing technology, minimum requirements and possible further work for the TTRIID 'bus network explorer' tool. A functional specification is designed to outline the exact behaviour required of a new piece of software or app, as opposed to describing the implementation technology or algorithms used to complete a task¹. In line with agile software development best-practice, this document will be updated continuously through the prototype development process, following iterative feedback from the internal product owner (David Brenig-Jones) and the T-TRIID technical lead (Holger Dalkmann).

2. Existing technology

- 2.1 GTFS Linker – Technology previously developed at ITP allows loading of
- a) OpenStreetMap (OSM) road networks, and
 - b) bus timetables captured in the General Transit Feed Specification (GTFS) static timetable format,
- 2.2 into an interactive Python web application. A data mining process (based upon a routing algorithm) then assigns routes within timetables to the road network in order to calculate statistics about the timetable. As GTFS stop locations and OSM mapping usually have some degree of data inaccuracy, a process of correction is required, where a knowledgeable user sets the exact road links used by each route. A simple interface is provided for this task. This app will form the backbone to the 'Network Explorer' tool, as linked GTFS data is then required to produce the interactive visualisations which will be the main feature of the tool.

¹ Joel Spolsky: *Painless Functional Specifications* (available online at <https://www.joelonsoftware.com/2000/10/03/painless-functional-specifications-part-2-whats-a-spec/>)

3. Functional requirements

- 3.1 'Mocked-up' user interfaces for the network explorer tool have been produced by ITP in previous work and as part of preparing for the TTRIID project. These are attached at the end of this document.
- 3.2 There are three 'minimum viable' features required to understand bus network capacity within the interactive interface:
- 1) **Link analysis:** selecting an individual road link within a city region (as displayed in the mockup) should show routes available through that link, or total counts of vehicle movements.
 - This feature can be used to understand where can be reached from a specific location within a city.
 - 2) **Route analysis:** selecting a specific route from a table should display the path of that route, with the map symbology proportionate a specific statistic for all sections of the route. For example, if bus R5 takes a major eastern corridor out of the city, it will have to share roads with up to 12 other services for the first part of its route and 7 other services after approx. 3km. A proportionate-width line is drawn to indicate this.
 - This feature identifies parts of a route that may be responsible for delays or overcrowding along the course of a route.
 - 3) **Region analysis:** selecting an area of the city by drawing a geographic region on the map selects all the bus services that run through that area. Again, map symbology displays road links proportionate to the amount they are used by bus or public transport services: busier links will be drawn wider, or coloured darker.
 - Using this part of the tool will help understand the parts of a city used most frequently by different services.
 - This will also be the first view that the user sees upon loading the tool, with other views accessible through components of the user interface.
- 3.3 For each type of analysis, the exact statistic reported can be one of several. For example, route count, vehicle count, vehicle frequency, seat capacity, or 'mode mix' may all be appropriate to different analysis of a city GTFS network.
- 3.4 The minimum viable feature for the tool is to display *route count* (numbers of distinct GTFS routes²) using each of the above types of analysis. However, it is expected that several of the other statistics listed above will be implemented, though their calculation

² GTFS static transit feeds reference: Routes: <https://developers.google.com/transit/gtfs/reference/#routestxt>

is more complicated due to the need to fully enumerate GTFS timetables and frequencies.

Road link directionality

- 3.5 Previous work with GTFS linking and general transport modelling tools developed by ITP has not always considered an important point in understanding transport networks: the direction of travel of bus or other services when they traverse road links. Directionality is very important in understanding the exact interaction of bus services on a city network. This feature may be regarded as a highest-priority 'next step' upon completion of the minimum viable components of the tool identified above.

User login and workspace functionality

- 3.6 In addition to the key features highlighted above, the tool should use a standard 'user login' approach to separate different GTFS feeds into those used by different people: ITP staff, IMC and DFID staff, ITP clients and partners, etc. This approach has not yet been built into all existing ITP GTFS tools, but is important for long-term usability of the tool and its wider use and commercialisation. Most modern web application frameworks (including Django, used for previous ITP tool R&D) make supporting user login and workspaces relatively straightforward.

4. Features that will not be developed

- 4.1 Possible features that will not be incorporated into the tool as part of the TTRIID work include the following:
- Visualisation of demand or patronage (passenger count) statistics from surveys of buses and public transport.
 - GTFS editing capability, allowing live changing of routes, trips or stops within a public transport feed file.
 - GTFS real-time capability, either receiving or exporting data relative to 'live' updates to GTFS feeds as they occur on local networks, for example referring to service updates.
- 4.2 Each of these may represent valuable long-term features for extension of the tool, but are considered out of scope for current development.
- 4.3 Feedback from project partners, in-country target users of the software, and potential clients will be considered through the course of the project. Where appropriate, the scope of this specification may be updated to reflect feedback regarding the most

important features of the tool, subject to agreement within the project team and with IMC/DfID.

Prototype user interfaces

Aggregation Specification

All modes ▼

Frequency ▼

All Day ▼

Both Direction ▼

Legend

View 1 shows the whole network. The width of the links communicate either the total number of routes, or the frequency of services on that link. Users can then choose to aggregate by all, by mode, time period and direction

← Make a selection of routes to get to View 2. Select by dragging over a link or around a stop

Show selection only

+

-

Routes Table – All modes, All Day, Both Directions

Name	Veh size	No. of vehs	Freq ▼	Total pass	Pass per veh	Max load	Ave pass distance	Route length	No. of round trips	Fare taken
Sandgandaan-Divisori	Small	578	252	13,684	54.3	17.1	4.7	8.1	4.2	205,260
Karuhatan-Ugong (Val	Small	714	218	4,680	21.5				4.8	70,200
Malabon-Monumento	Small	687	192	2,307	12.0				4.4	34,605
Balintawak-Monumen	Small	390	190	1,328	7.0				4.8	19,920
Sta Ana-P Faura	Small	30	173	3,287	19.0	10.0	2.0	4.8	4.6	49,305
L Guinto-Pandacan	Small	262	158	2,101	13.3	11.3	0.6	3.3	4.8	31,515
Kalentong-Boni/Pinat	Small	209	156	3,594	23.0	11.2	1.1	4.0	4.7	53,910

← Routes table lists all of the selected routes. If no routes selected, all routes are listed. The table can be sorted by any column.

Aggregation Type

All modes ▼

Frequency ▼

AM-peak ▼

Inbound ▼

Legend

☞

Show selection only

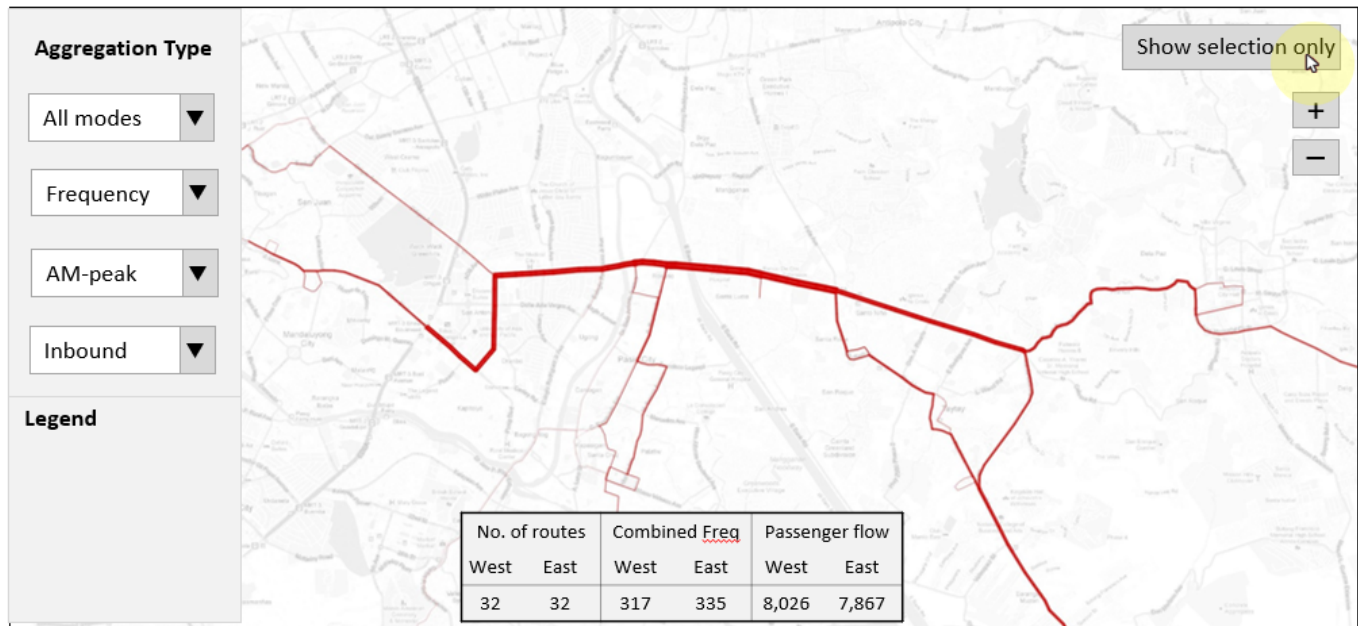
+
-

Summary statistics for selection

No. of routes		Combined Freq		Passenger flow	
West	East	West	East	West	East
32	32	317	335	8,026	7,867

Routes Table ✕

Name	Veh Size	No. of vehs	Freq	Total Pass	Pass per veh	Max load	Ave pass distance	Route length	No. of round trips	Fare taken
Antipolo-EDSA/Shaw	Small	85	17	639	37.6	25.8	7.7	16.0	3.3	9,585



Routes Table ✕

Name	Mode	No. vehicles	Frequency ▼	Capacity/veh.	Pass. Capacity/Hour	Rte Dist	Rnd Trips
Tanay-EDSA/Shaw (via Binangonan-JRC/JRU (Jeepney	298	44	34	1,500	36.8	1.8
Tanay-EDSA/Shaw (via Taytay-Pasig	Bus	8	35	47	1,962	29.0	1.7
Tanay-EDSA/Shaw (via Antipolo-JRC/JRU	Bus	157	24	41	976	43.0	2.1
Tanay-EDSA/Shaw (via Taytay-Quiapo (via Ortigas Center)	Jeepney	95	23	35.7	808	12.5	3.7
Tanay-EDSA/Shaw (via Antipolo-EDSA/Shaw	Jeepney	69	20	29.5	603	19.9	2.8
Tanay-EDSA/Shaw (via Antipolo-EDSA/Shaw	Bus	98	19	95.5	1,799	24.5	4.5
Tanay-EDSA/Shaw (via Antipolo-EDSA/Shaw	Jeepney	85	17	37.6	639	16.0	3.3

5. Implementation Components

- 5.1 This section plans the major parts of the implementation of the prototype, identifying the technologies used where appropriate and listing the anticipated technical challenges.

Views (Mithril JS and Mapbox GL)

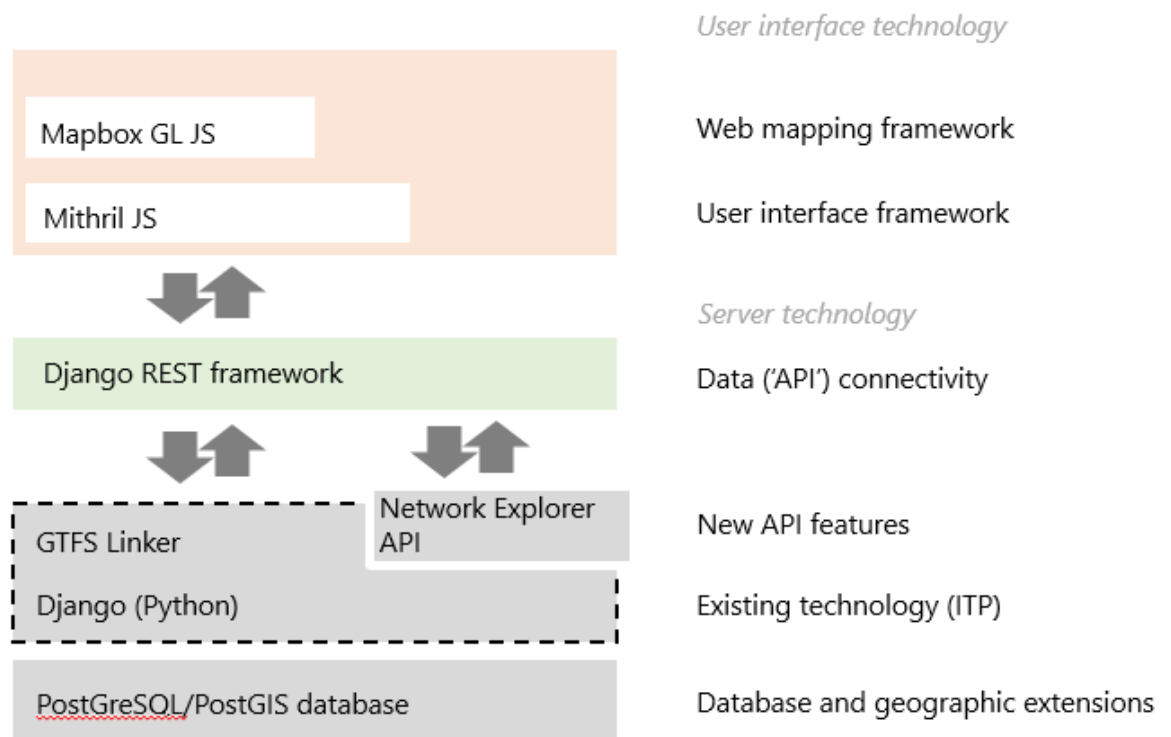
- **Map interface** – ‘Slippy’ map which interactively shows routes, busy links, and a ‘whole feed’ overview summarising the routes of the city. This will be implemented using the Mapbox GL JS plugin (<https://www.mapbox.com/mapbox-gl-js/api/>), a modern framework which supports a relatively new technology known as ‘vector tiles’. These enable displaying large volumes of interactive data in a web map interface, something which had previously represented a technical challenge using other frameworks such as Leaflet JS.
- **Route view** – This is a simple overlaid table which summarises routes selected when a number are shown on the map, for example in a ‘spider’ diagram (see mock-ups). This summarises the names and statistics of routes within the GTFS feed and can be used to select specific routes for viewing on the map.
- **Summary view** – This box will report other statistics regarding a route, including numbers of routes total, numbers selected, and information about timing of services. Though this view is not specified as a ‘minimum viable’ feature (see above), it is expected to be relatively simple to implement and will likely be a valuable addition longer-term.

Data layer (Django REST Framework)

- **Segment API** – This part of the API will accept a geographic coordinate pair (from where the user has clicked the map as a ‘query’ location) and will return a road segment closest to that which has been clicked. This API endpoint is not used directly by the views above but will pass data forward to the ‘route’ and ‘spider’ APIs, both of which are responsible for more detailed data being returned to the Views.
- **Route API** – This API endpoint will build upon a segment query by returning an aggregated set of route segments, each with statistics (e.g. route count or frequency) attached.
- **‘Spider’ API** – The most complicated part of the API will query for all routes which pass through a certain road segment (that selected by the user). It will then return the complete set of segments used by these trips, aggregating to report the desired statistic

Server-side implementation (Django and PostGIS)

- 5.2 The only planned change to the existing technology’s underlying data model is if the ‘network linking’ software already developed by ITP is extended to capture *directionality* of routes and services. Feedback from historic projects has captured that this is a very important part of understanding movement of public transport services, but the current data model does not allow separate understanding of different directions of travel for bus services etc. Note that a particular challenge is determining which side of the road is used for travel, as this is not often captured explicitly in data sources such as OpenStreetMap.
- 5.3 The figure below illustrates the planned architecture of the software, including the API layer and new database features required to support the use of the prototype.



Appendix B: Public transport link statistics

The table below gives an overview of the public transport route statistics that could ultimately be displayed by the tool, including aggregated statistics that specifically require the new 'linked GTFS' technology to calculate.

Please note this does not include passenger patronage data, which would require significant additional development under a following phase of the project.

Public transport data per-route	Public transport data aggregated by highway links
Route Name	Count of routes
Mode	Mode split
Subsidised?	
Franchise start date	
Franchise end date	
Route length, two-way	
Route length, one-way average	
Passenger capacity per vehicle	Combined capacity
Frequency	Combined frequency
	Operating speed (surveyed/estimated)
	Vehicle count



Integrated Transport Planning Ltd
Cornwall Buildings, 45 Newhall Street
Birmingham
B3 3QR UK
+44 (0)121 213 4725

Integrated Transport Planning Ltd
Castlemead
Lower Castle Street
Bristol
BS1 3AG UK
+44 (0)117 917 5155

Integrated Transport Planning Ltd
6 Hay's Lane
London Bridge
London
SE1 2HB UK
+44 (0)203 300 1810

Integrated Transport Planning Ltd
50 North Thirteenth Street
Milton Keynes
MK9 3BP UK
+44 (0)1908 259 718

Integrated Transport Planning Ltd
32a Stoney Street
Nottingham
NG1 1LL UK
+44 (0)115 988 6905

www.itpworld.net

