An open source framework for GTFS data analytics: case study using the Brisbane TransLink network

Aaron Lim[1], Shubham Sharma[2], Ashish Bhaskar[1], Shriniwas Arkatkar[2]

[1] Queensland University of Technology, 2 George St, Brisbane City QLD 4000

[2] Sardar Vallabhbhai National Institute of Technology, Surat, India - 395007

Email for correspondence: ashish.bhaskar@qut.edu.au

Abstract

This paper presents an open source framework with key logic units design to assist with the acquisition, elaboration, storage and visualization of General Transit Feed Specification (GTFS) real-time data. It bridges the gap for researchers and planners wanting access to GTFS real time through providing the resources to understand and begin extracting GTFS data. To showcase the applicability, the framework is applied on the real data from Translink, Brisbane, Australia and the extracted data is utilised for performance measures and dashboard visualization. The codes are open sourced and are available at https://github.com/darronlim/GTFS_LiveFeed_Extraction.

Keywords: GTFS, transit data, dashboard, public transport, visualization

1 Introduction

Advances in numerous smart data collection technologies have generated much interest in the transportation sector, transit agencies, government organisations and researchers. There is potential in applying this data in assessing network performance, identify prevailing issues, and develop improvement measures. Despite this, there is a gap for researchers and planners wanting to understand and access this data with a lack of open source frameworks and applications available without a high level of programming language knowledge required. Therefore, methods are required to access and store this wealth of data for research and planning.

Within Australia, GTFS is growing in adoption amongst the different state government transport departments. GTFS real time data is available from NSW Transport (NSW Government Open Data. 2019) and in Queensland through TransLink (TransLink, 2019). NSW has many modes of transport being monitored with trains, busses, ferries and light rail and in Queensland there are busses, trains and ferries. The future potential of this data should be seriously considered with the growth of the nation and the infrastructure required. With TransLink continuing to expand and the Cross River Rail project commencing soon (Queensland Government, 2018), the wealth of data being stored now could be used to compare with after in order to access the effectiveness of the new lines.

The main aim of this project is to develop a methodology to extract data from open source GTFS live feeds and tested on the TransLink GTFS live feed. The contributions of the work are:

A. A systematic framework to extract and store GTFS data from live feeds,
B. A framework to apply GTFS data for transit performance indicators,
C. Conceptualize dashboard visualizations with the processed data to perform evaluations of the public transportation network, and
D. Document and provide applications through open-source via GitHub.

The process of developing this framework involved a literature review into existing studies identifying the gap leading to the development of process along with an application to extract, store and process the GTFS real-time feeds. The methodology developed demonstrates the GTFS static and GTFS real-time data flow through various stages.

Along with the creation of a database with historic GTFS real-time data which is beneficial for developing effective planning tools; this information along with the real-time data can be visually interpreted and understood by the end users of a transit network, researchers, app developers as well as transit management agencies.

With the application files made available in a public GitHub (available at https://github.com/darronlim/GTFS_LiveFeed_Extraction), a foundation is provided in which future researchers, planners and application developers can refine and enhance it to meet specific objectives. This work would provide a platform for future transit planning and research to catapult from and would be helpful for those entering the field of public transport data acquisition and visualization to extract the required data and understand the process. This documentation will also highlight areas where further development is required, and more research can be conducted.

2 Literature Review

In the last two decades, advances in field of data collection, storage and analytics like wireless communication, location-aware technologies have helped in mining massive volume of spatiotemporal data available within transport system including real-time transit data. The information in real time and in historic context plays a vital role in implementation of intelligent transportation System (ITS) technologies and managing Public Transport systems (PTS).

Over the last decade since its inception, General Transit Feed Specification (GTFS) has emerged as a world-wide popular open-source industry data standard to describe and publish fixed-route transit operations. Open design of GTFS, clear documentation along with suggestions and feedback from user community has enabled significant development of the GTFS format over time in becoming a popular robust data format. Introduced in 2005 as Google Transit Feed specification as a collective effort of Google and TriMet to create a web-based transit trip planner application, it was made open for general-use by mid-2007 (Antrim, A. and Barbeau, S.J., 2013). In 2011, Google introduced an extension to the static GTFS schedule data, known as GTFS real-time, with potential to standardize real-time information feed, including trip updates (include stop arrival delays), service alerts and vehicle geographic positions. A revised version GTFS-realtime v2.0 was released in 2017 to overcome the limitations of previous version especially lack of proper documentation and open source validation tools (Barbeau, S.J., 2018).

Open data has proved to be a catalyst in numerous innovations in field of designing navigation, warning and planning tools irrespective of whether straight benefits for the

data owning agency are readily apparent or not. GTFS has attracted a lot of attention from industries, transport agencies and researchers from all over the world. It has found numerous applications in transport industry including transit timetable creation (TimeTable Publisher), multimodal trip planning (Google Maps, Opentrip Planner, Bing Maps, etc.), Transit accessibility tools (Travel Assistant Device (TAD Sendero Group BrailleNote GPS,), etc.), transport planning analysis tools (OpenTripPlanner: Analyst Extension, Graphserver, etc.), ridesharing apps, data visualization (Mapnificent), Real-time information to users and development of numerous mobile applications (Antrim, A. and Barbeau, S.J., 2013).

Various researchers have developed the framework for GTFS data extraction, storage and management based on the requirements and applications of their research. Hadas, Y. (2013) presented a unified methodology to extract, store and analyze public transport data for spatial analysis on GIS using static transit feeds and transportation network with an aim to develop an easy-to-use tool to help planners to analyze public transport network connectivity and enabling benchmarking. Wong, J. (2013) demonstrated the potential of GTFS data feed as a data source for transit analysis and proposed the methodology for stop, route and system level analysis of transport network and made the source code open. Bast, H et. al (2014) described a framework to create a world-wide live map of public transit using GTFS time-table data using information from GTFS real-time trip-update tab to visualize real-time movements of buses, trains and ferries and addressed the issue with handling and management of big data while creating a live-vehicle tracker.

Since the trip update entity provides stop arrival and departure times along with arrival delays, these studies used linear interpolation of travel time between two stops to locate the live position of a transit vehicle as the vehicle position entity was not introduced into the GTFS real-time structure at that time. This linear interpolation may produce incorrect estimates especially in complex urban transit networks. Since the trip update entity of GTFS real-time now provides arrival and departure information at transit stops only, this study includes in its framework, the time-space information about the location of vehicles for the intermediate points between stops using vehicle position entity data to obtain more precise estimates of transit vehicle positions in visualizing real-time movements.

Further, many studies have focused on calculating transit performance measures using GTFS feeds with supplemental data. Tao, S et al. (2014) used GTFS data and smart card dataset to construct bus passengers' travel trajectories at bus stop level of spatial granularity. The study proposed a GIS platform-based methodology to visually unveil the spatial-temporal travel behavior dynamics of urban public transport (UPT) passengers. Wessel, N. and Farber, S. (2019) in their accessibility study using GTFS and AVL data used PostGIS database to store GTFS data in essentially a similar format to that of individual comma delimited files, allowing PostGIS database to perform certain operations. He noted that the time needed for calculations was a big limitation to his accessibility study concluding with need of purpose-built software for efficient analysis of retrospective GTFS data.

This study focusses on targeting the gap in research by developing a framework for feed extraction, storage, pre-processing and management of GTFS static and GTFS real-time feed for transit data visualization using Tableau including visualization of vehicle movement, vehicle trajectories and trip delay patterns.

3   General Transit Feed Specification

General Transit Feed Specification is a common format for public transportation schedules and geographic information (Google, 2016). Developed by Google, this data includes the following (Google, 2016):

1. GTFS Static,
2. GTFS Real-time.

GTFS Static contains the schedule data of a transit network whereas GTFS Realtime contains the real time information about the transit operations. Hence, the comparison of real time transport performance data from GTFS real-time against the schedule data of GTFS static enables the user to assess transit performance both in real time and in a historic context.

GTFS Static is an open data format for public transportation schedules and associated geographic information. It contains 13 text Comma Separated Values (CSV) files with 6 mandatory files and 7 optional files as shown in Figure 2. Every file contains certain mandatory and optional fields. The data from a transit agency is in GTFS format only if it satisfies the minimum mandatory requisites. The mandatory files in GTFS static are agency.txt (provides information about the transit agency as such, including name, website and contact information), routes.txt (identifies distinct routes), trips.txt (include information of scheduled services), stop_times.txt (contains scheduled arrival and departure time information for sequence of stops for a scheduled transit service), stops.txt (defines the geographic locations of each and every actual stop or station in the transit system as well as, and optionally, some of the amenities associated with those stops) and calendar.txt (defines service patterns that operate recurrently such as, for example, every weekday).

The optional files are Calendar_dates.txt (service patterns that don't repeat such as for a one-time special event will be defined in the calendar_dates table), fare _attributes.txt, fare_rules.txt, shapes.txt (rules for drawing lines on a map to represent a transit organization's routes), frequencies.txt (Headway (time between trips) for routes with variable frequency of service), transfers.txt (rules for making connections at transfer points between routes) and feed_info.txt.

GTFS real-time was designed around ease of implementation, good GTFS interoperability and a focus on passenger information (Google, 2016). GTFS real-time consists of following entities:

- **Trip updates** - delays, cancellations, changed routes
- **Service alerts** - stops moved, unforeseen events affecting a station, route or the entire network
- **Vehicle positions** - information about the vehicles including location and congestion level

4   Framework

The proposed framework is built around a modular architectural schema of four logic units: acquisition, elaboration, storage and visualization as developed by Celestini, A., Di Marco, A. and Totaro, G (2014). This schema has been designed to manage the

extraction of large sources od data and enhance the analysis through visualization (Celestini, A et al.).
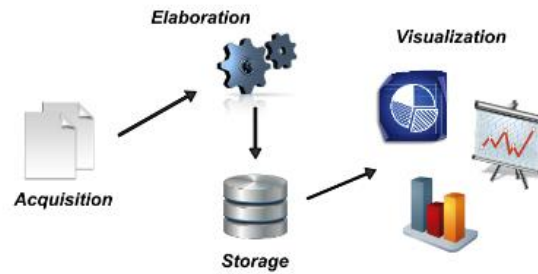

Figure 1: Architectural Schema (Celestini, A et al. 2014)

Through adopting following schema, the following framework has been developed to engage with GTFS real-time data.
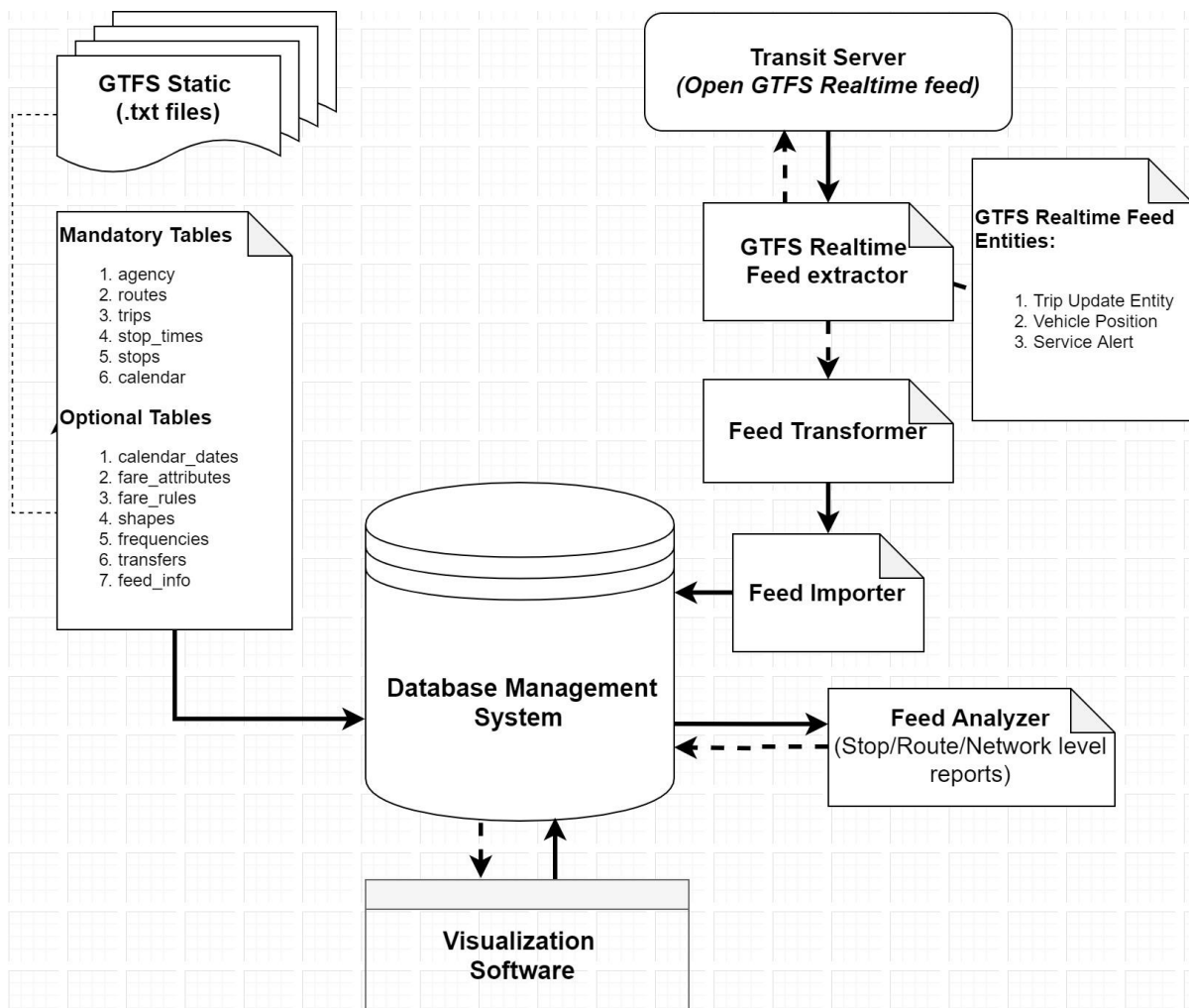

Figure 2: Workflow for GTFS data processing

4.1 GTFS Realtime Feed

GTFS Realtime Feed extractors send requests to the transit server for the GTFS real-time feed and receives it from the Transit server. The frequency of a feed updating

depends on the process of feed generation and can vary from a few seconds to a few minutes and depending on the providing agency. The proposed framework sends the request for a real-time feed every second and compares it with previously received feeds. A GTFS real-time feed contains three entities namely trip update, vehicle position and service alert. If any of the entities within the received feed is distinct compared to the previously received feeds, the feed extractor sends the identified unique entity to Feed Transformer and returns to the process of sending requests to the Transit server for new updates on real time feeds.

4.2 Feed Transformer

The Feed Transformer receives the unique GTFS real-time feed entity and converts the information into tabular form. Separate tables are defined for each entity (Trip Update, Vehicle Position and Service Alert). The timestamp at which the feed was created/received should be stored with each input to the table. A Database Management System (DBMS) is required to store the processed feed. The DBMS should consist of three separate databases for GTFS Static, Transformed GTFS real-time and processed/analyzed data. The GTFS Static database should contain separate tables for each .txt file and the Transformed GTFS real-time database should have separate tables for all three entities (trip update, vehicle position and service alert). The database for processed/analyzed data should have separate tables for stop, route and network levels reports, with the structure depending on user requirements and type of analysis.

4.3 Feed Importer and Feed Analyzer

The Feed Importer imports the transformed feed to the database management system and updates the entity specific table in the database with the new transformed feed.

The Feed Analyzer may contain several programs/applications/scripts designed according to the user requirements. It interacts with the database and requests database management system for the data and to perform analysis. It then stores the results back into the database management system in tabular form. The design of the Feed Analyzer depends on the objectives of the analysis.

4.4 Data Visualization

Data visualization is a powerful tool for interpretation of results. Correctly applied data visualization plays a critical role in reiterating the outcome of the study and strengthening the impact on the end user. The GTFS static "shapes.txt" file contains rules of drawing the routes and is used for plotting the routes on map. "Stops.txt" contains the geographic location details of the transit stops along with other information and used to plot location of stop on map. A dual axis map created from shapes.txt and stops.txt with base layer of OpenStreetMap has been used in this study for visualization. The dual axis feature allows for two sets of geographic data to be overlaid on top of one another. The proposed framework suggests an interactive connection between the visualization software and the DBMS where all the required GTFS static, GTFS real-time and processed/analyzed data is stored and updated continuously.

5 Case Study: Queensland Department of Transport and Main Roads Public Transport division – TransLink

5.1 Data acquisition and Pre-processing

This study is primarily built around extracting the open data made available by Queensland's Public Transport Division – TransLink. TransLink provides a live feed of open GTFS data from the vehicles in the fleet equipped with real-time tracking technology (TransLink, 2019). To extract the GTFS live feed data, a protocol buffer is required to deserialize the structured data (Google, 2019). This buffer unscrambles the files extracted from the GTFS data source and present it in a format that is understandable and can be stored. Protocol buffers currently support generated code such as Java, Python, C#, and C++, and therefore can be applied in different platforms depending on requirements.

GTFS Static files can be extracted from the TransLink GTFS Network Data page, the files extracted contain timetable data and stop locations for all available TransLink South East Queensland (SEQ) bus, rail and ferry services. The files (Queensland Government, 2019) and their definitions (Google Transit APIs, 2019a) can be found in the references. Further information regarding types of GTFS files and the definitions of fields can be found on the Google Transit APIs Static Transit Reference Webpage (Google Transit APIs, 2019b).

An application has been developed using the open source language Python that extracts the live feed data and stores it into database tables. The benefits of this application are that it can be run on different operating systems and can be automated to continuously run and extract data. Users will have the ability to adjust the time between the data extraction based on requirements.

The code for these applications is available for download and use at the following link: https://github.com/darronlim/GTFS_LiveFeed_Extraction.

The GTFS real-time data extracted is separated into three entities - Vehicle Entity Data, Trip Update Data and Alert Data. Further documentation on the entities and the fields can be found on the GTFS documentation page (Google Transit APIs, 2019b). If required, GTFS Network static data files can be combined with GTFS Live Feed data extracted as an additional step in pre-processing before storing the data if specific information or calculations are required.

The storage units suggested storing separate GTFS entity data for this project are relational databases and a MySQL database was implemented for the application developed. Different storage solutions can be utilized depending on project requirements.

5.2 GTFS Data Analysis and Visualization

With the data extracted, processed and stored, the most important aspect is how it can then be presented for interpretation and analysis. GTFS data can be utilized in transit performance indicator analysis and visualizations as outlined in the following sections.

5.2.1 Transit Performance Indicators

GTFS provides a wealth of information about the transit and traffic operations which can be utilized for assessing the performance of the transit system. It can assist transit agencies in strategic, tactical and operational level planning and furthermore, in implementation of improvement strategies to enhance the overall performance of the transit system. Some of the performance indicators which can be derived using GTFS Static and GTFS real-time, either as a standalone data source, or in fusion with other data sources, with GTFS as either a primary or a secondary source of data are listed in Tables 1 to 5 (Kittelson et al., 2003, Mistretta et al., 2009).

Table 1 highlights some of the performance indicators for transit service evaluation. Route level performance measures are listed in Table 2 followed by Transit mobility performance measures in Table 3. Performance measures related to dwell time are listed in Table 4. Point based measures are tabulated in Table 5.

Table 1: Performance Indicators for service evaluation

Performance Indicator	Definition	Data Source
Service Coverage	The extent to which a defined service area is being served. Service coverage is generally measured by the percentage of the population residing within quarter mile walking distance of a bus stop.	GTFS Static and Census data
Time and Distance Service Calculations	Is a measure of temporal and spatial availability of a transit service.	GTFS Static
Route and Service Directness	The degree to which a route deviates from the shortest path between the start and end points of the route. It has an inverse relationship to service coverage. The route that deviates provides greater coverage at the expense of directness, on the other hand a more direct route provides less coverage.	GTFS Static
Service Frequency	Represents how often buses arrive at a stop. The term is often used interchangeably with service headway which refer to the interval in minutes between two successive bus departures.	GTFS Static and GTFS Realtime
Span of Service	This measure relates to the capacity of the system. Span of service is the number of hours and days when service operates. The longer that service is available, the greater is the capacity of the system. Generally, high-performing routes will have longer service spans, and weekday spans of service are longer than for Saturday and Sunday	GTFS Static

Table 2: Route level performance measures

Performance Indicator	Definition	Data Source
Scheduled hours of service	It is the number of hours during a day that transit service is provided according to the schedule and can vary by day of the week, by route, and even by stop. Scheduled hours of service impacts the convenience of transit for passengers and can constrain the types of trips that can be made by transit	GTFS Static
Actual hours of service	Actual hours of service are the number of hours during a day that transit service is actually provided.	GTFS Realtime
Number of scheduled trips	Number of trips along a route according to the transit schedule for a specific span of service	GTFS Static
Number of actual trips	Number of actual trips on a particular route for a specific span of service	GTFS Realtime
Number of scheduled km	Distance along a route on which the transit service is supposed to be operating according to the transit schedule for a specific span of service	GTFS Static
Number of actual km operated	Distance along a route on which the transit service was actually operated for a specific span of service	GTFS Realtime
Average scheduled speed	Average travel speed along a route according to the scheduled departure at origin stop and scheduled arrival at the last stop of a route	GTFS Static
Average actual speed	Average travel speed along a route according to the actual departure of a transit vehicle from the origin stop and the actual arrival time at the last stop of the route.	GTFS Realtime

Table 3: Transit Mobility performance measures

Performance Indicator	Definition	Data Source
Percentage of on-time performance	It is the percentage of the services on time to the total services (on network level/route level)	GTFS Realtime and GTFS Static

Percentage of scheduled departures that do not leave within a specified time limit	It is the percentage of transit trips/services that do not depart a stop within a specified time limit.	GTFS Realtime and GTFS Static
Number of public transportation trips	Total number of transit trips (scheduled /actual) between a pair of stops for a specific span of service.	GTFS Realtime and GTFS Static
Average travel time	Average travel time by transit vehicles between two transit stops.	GTFS Realtime and GTFS Static
Average travel speed	Harmonic mean of travel speed of transit vehicles between two transit stops.	GTFS Realtime and GTFS Static
Average travel delay	Deviation of the actual average travel time from the scheduled travel time between two stops.	GTFS Realtime and GTFS Static

Table 4: Dwell Time (system, route or point level) as transit performance indicator

Performance Indicator	Definition	Data Source
Total number of stops	Total number of transit stops along a route.	GTFS Static
Average dwell time	The average time a vehicle or train is allowed to discharge and take on passengers at a stop, including opening and closing doors	GTFS Realtime (Trip update entity)

Table 5: Point-level transit performance measures

Performance Indicator	Definition	Data Source
Number of scheduled trips passing a point	Total number of scheduled transit trips passing a specific point (e.g. stop) during a specific span of service.	GTFS Static
Number of actual trips passing a point	Total number of a transit trips actually passing a particular point (e.g. stop) during a specific span of service.	GTFS Realtime (Trip update entity)

Percentage of actual and scheduled trips	It the percentage of the number of actual trips that passes a specific point (e.g. stop) with respect to the number of scheduled trips passing that point. During a specific span of service	GTFS Static and GTFS Realtime

5.2.2 Dashboard Visualization

It is difficult to understand or quantify how a public transportation system is performing with data presented in rows and columns of a table. For this project, Tableau is utilized to visually present information such as vehicle routes and their travel times, along with stop locations, vehicle movements and speeds along with delays. Tableau was selected as it is an open source platform at an industry standard, being simple to learn and understand but also having the power and capacity to develop more complex interactive visualizations that are compatible on multiple platforms.

5.2.3 Dashboard Visualization of Vehicle Movement

The dashboard presents vehicle movements over a period of time within the South East Queensland Region. This mainly utilises GTFS real-time vehicle position entity data presenting the trip information and vehicle location data. This dashboard would be most effective on an operational scale where the GTFS real-time feed is fed into the visualization software to visualize real-time movement of transit vehicles.

In the example below the data has been paused at 4:16PM on the 15 April 2019 and has been zoomed into the main Brisbane CBD region to visualize the location of transit vehicles.
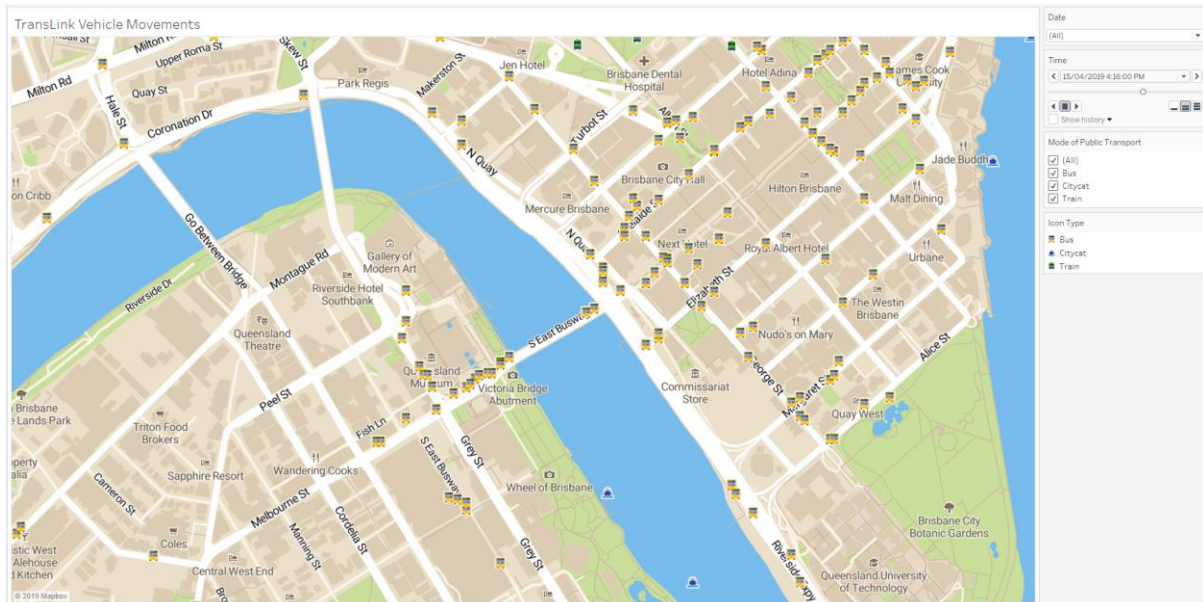

Figure 3: Dashboard visualization of the vehicle position entity of GTFS real-time to visualize the movement of transit vehicles

When hovering over a vehicle icon data fields are supplied to provide information on the specific vehicle and the trip it is undertaking.

11

For instance, the details of this CityCat can be identified when the cursor hovers over the icon as shown in the figure below.
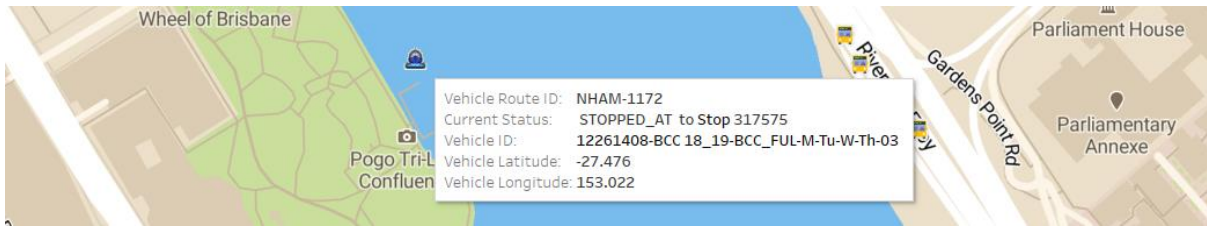

Figure 4: Dashboard visualization of location details of CityCat Ferry at as a specific time

Similarly, the details of busses can be identified when the cursor hovers over the icons too as shown below.
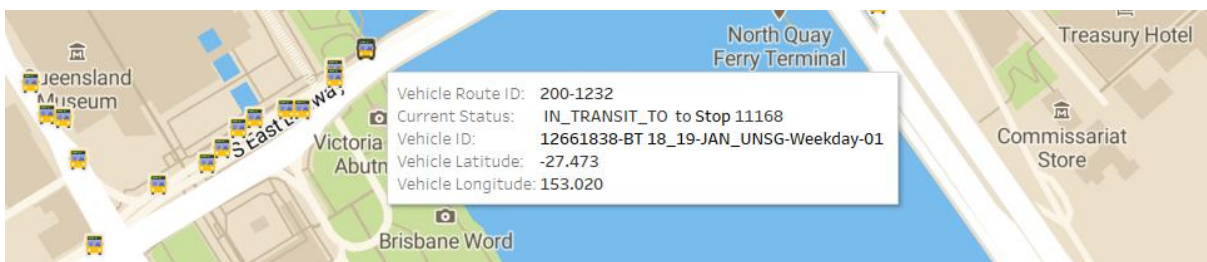

Figure 5: Dashboard visualization of location details of a bus at as a specific time

5.2.4 Dashboard Visualization of Stop Arrival Delays

The transit stop arrival delay dashboard (Figure 6) presents vehicle delays as it travels along a route. For a reliable transit system, it is important to ensure if transit services are performing within the acceptable performance time windows set by transport operators. The following dashboard allows the selection of periods of time along with specific route and trip information that is required. The points that are presented on the map are the stop passed on that specific trip and the corresponding colour coded delay.
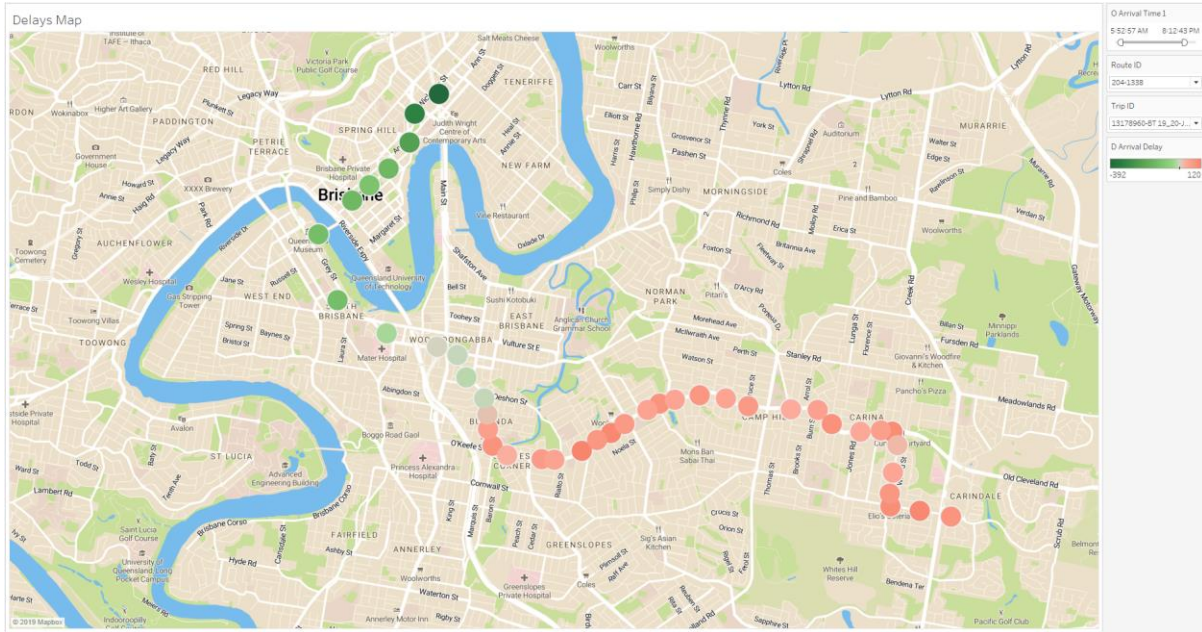
Figure 6: Dashboard visualization of the arrival delays at series of transit stops for specific trip along a specific route

When hovering over stop icons data fields are supplied to provide information on the specific stop and the trip and delay information as shown in the figure below.
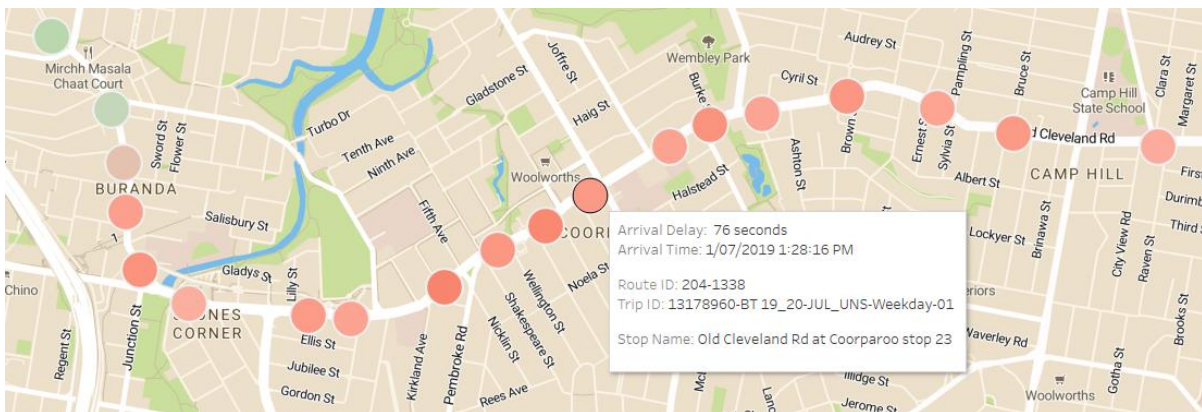

Figure 7: Visualization of trip details along with the arrival delay information at one of the stops along the study route.

To support the representation of delay information using color maps, the graphical representation of arrival delay (Figure 8) represents the delay information at various stops along a route. All trips can be displayed along with the delay information as they progress along the stops on the routes. Data points contain further information on the specific stop and the trip and delay information as shown in Figure 9.

Both the map and graph could be used in a strategic and tactical aspects as it assists in identifying trends, hotspots and areas of improvement for transit networks. It could also be used in an operational capacity if real time data is constantly updated and presented.

Figure 8: Arrival delays at sequence of stops along a transit route (for a number of trips).
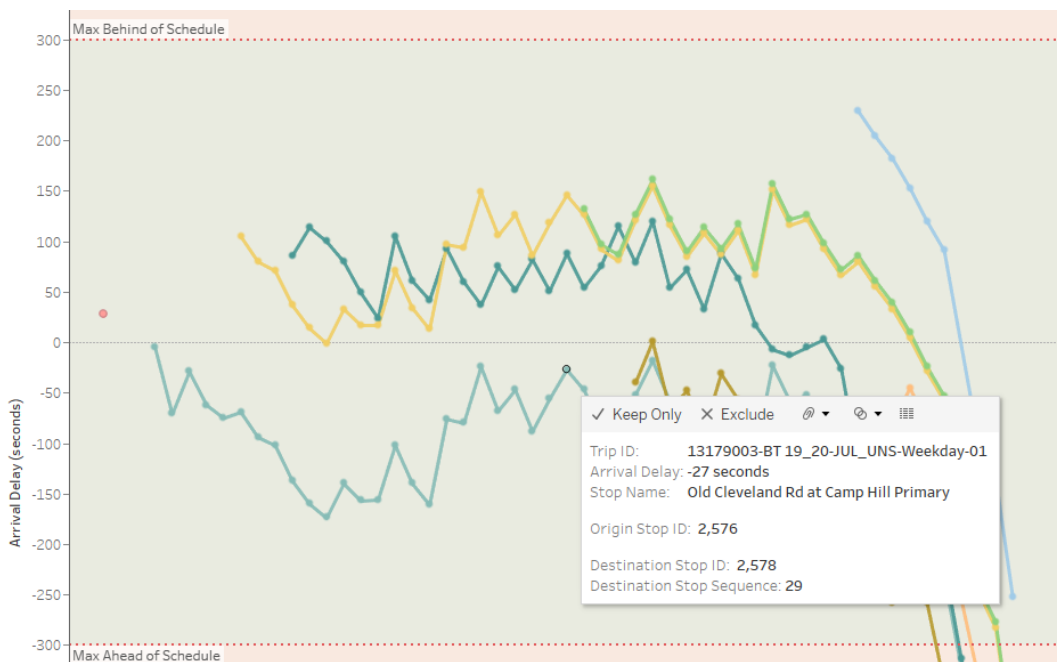

Figure 9: Interactive dashboard visualization of arrival delay information at a selected transit stop.

6 Conclusion

Public Transport GTFS data has much potential and can be used in powerful analysis and interpretation. It allows insight into different aspects of transit at a strategic, tactical and operational level along with the impact that both the suppliers and users of this service have on it. This project develops an understanding the process of how GTFS data can be extracted, stored and analyzed. That the large amounts of data can be accessed and utilized with the correct applications and platforms developed around it.

With this, the significance of this work is that the framework developed outlines the extraction and storage of GTFS live feed data using the applications built, along with demonstrating the use of this data through transit performance indicators and visualizations that can be applied for different strategic, tactical and operational levels. Ultimately this study provides the foundations to the start of a process in which enthusiast researchers and practitioners can access GTFS data and utilise it for their requirements. The applications and relevant documentation developed from this project are open source and available to all and further development is encouraged.

An area of improvement for this project is the refining of the framework to a higher level, which can be accomplished but requires more time applied to analyze and clearly describe the process for each stage of the framework. Also, the framework and applications require further testing on different GTFS real time sources from within Australia and even on international sources in order to access its performance and apply any refinements as required.

It is planned that further development will continue over the remainder of the year on both the program and visualizations for a more stable and better structured high-level framework process. For others that build upon this project, it is hoped that a more refined product and methodology can be developed from the work that has been conducted.

Once enough data has been collected there is potential for machine learning to be incorporated into the flow. Where the trends of historical data can be analyzed and presented within the visualizations to inform users of predicted speeds, travel distance and even delays based on the historical data collected from transit services. This can be applied across the strategic, tactical and operational levels of transit performance analysis and bring it to a higher level.

Acknowledgements

References

Antrim, A. and Barbeau, S.J., 2013. The many uses of GTFS data–opening the door to transit and multimodal applications. Location-Aware Information Systems Laboratory at the University of South Florida, 4.

Bast, H., Brosi, P. and Storandt, S., 2014, November. Real-time movement visualization of public transit data. In Proceedings of the 22nd ACM SIGSPATIAL international conference on advances in geographic information systems (pp. 331-340). ACM.

Barbeau, S.J., 2018. Overcoming Barriers to Real-time Transit Information.

Celestini, A., Di Marco, A. and Totaro, G., 2014. A Data Extraction and Visualization Framework for Information Retrieval Systems Accessed August 2, 2019. http://ceur-ws.org/Vol-1210/datawiz2014_02.pdf

Google. 2016. "GTFS Static Overview." Accessed May 26, 2019. https://developers.google.com/transit/gtfs/.

Google Transit APIs. 2019a. Static Transit Field Definitions. Accessed May 26, 2019. https://developers.google.com/transit/gtfs/reference/#field-definitions

Google Transit APIs. 2019b. Static Transit Reference. Accessed June 13, 2019. https://developers.google.com/transit/gtfs/reference/

Google. 2019. "Protocol Buffers." Accessed May 26, 2019. https://developers.google.com/protocol-buffers/.

Hadas, Y., 2013. Assessing public transport systems connectivity based on Google Transit data. *Journal of Transport Geography*, *33*, pp.105-116.

Kittelson, Associates, Administration, U. S. F. T., Program, T. C. R. & Corporation, T. D. 2003. *A guidebook for developing a transit performance-measurement system*, Transportation Research Board.

Mistretta, M., Goodwill, J. A., Gregg, R. & Deannuntis, C. 2009. Best practices in transit service planning.

NSW Government Open Data. 2019. Public Transport - Realtime Vehicle Positions API. Accessed August 2, 2019 https://opendata.transport.nsw.gov.au/node/330/exploreapi

Queensland Government. 2018. Cross River Rail. Accessed August 1, 2019. https://crossriverrail.qld.gov.au/

Queensland Government. 2019. "General transit feed specification (GTFS)—South East Queensland." Accessed June 12, 2019 https://data.qld.gov.au/dataset/general-transit-feed-specification-gtfs-seq

Tao, S., Rohde, D. and Corcoran, J., 2014. Examining the spatial–temporal dynamics of bus passenger travel behaviour using smart card data and the flow-comap. Journal of Transport Geography, 41, pp.21-36.

TransLink. 2019. "TransLink GTFS Real-Time Feed". Accessed June 10, 2019. https://gtfsrt.api.translink.com.au/.

Wong, J., 2013. Leveraging the general transit feed specification for efficient transit analysis. Transportation Research Record, 2338(1), pp.11-19.

Wessel, N. and Farber, S., 2019. On the Accuracy of Schedule-Based GTFS for Measuring Accessibility (No. 19-04587).