General Transit Feed Specification (GTFS) – And GTFS-Realtime Feed for Sydney Light Rail

Consumer Guide

People and Corporate Services

Systems

OFFICIAL

Contents

Author:	David Phillips
Date:	July 2016
Version:	1.0
Reference:	
Division:	Customer Services Division
Review date:	

Change Record and Documentation Control

Version	Date	Name	Change Reference
0.5	26-Nov-2014	David Phillips	Initial version
0.6	16-Mar-2016	David Phillips	Updated for Open Data
0.7	21-July-2016	David Phillips	Enhancements to GTFS bundle
1.0	3-Aug-2016	David Phillips	Minor amendments following review

Reviewed By

Version	Date	Name	Role
0.7	29-July-2016	Stephanie Manning	Manager, Data Enablement, DCIS, CSD
0.7	1-Aug-2016	Ashish Ghimire	Quality Assurance Manager, Group IT

Related Documents

Ref	Document Name	Version / Date	Location
01	GTFS Guides on Google Developers	Accessed 26-Oct-2015	https://developers.google.com/transit/gtfs/?hl=en
02	Realtime Transit on Google Developers	Accessed 26-Oct-2015	https://developers.google.com/transit/gtfs-realtime/

1 Introduction

1.1 Overview

The purpose of this document is to describe the structure and contents of the Sydney Light Rail Timetable data feed supplied by Transdev for consumption and use by Transport for NSW systems and business units and application developers.

The Light Rail Timetable data feed is in the form of a General Transit Feed Specification fileset, known as 'GTFS'.

The intended audience of this document is application developers.

While GTFS is well documented, it provides several optional fields and some flexibility in how to populate some fields. This document provides some general information regarding the specific contents and structure of the GTFS Timetable Feed for Sydney Light Rail

1.2 Purpose of Sydney Light Rail GTFS Fileset

The purpose of the Sydney Light Rail GTFS Timetable Feed is to publish in advance the schedules and route information of Light Rail services on the L1 Dulwich Hill Line.

Consumers of the Sydney Light Rail GTFS Timetable Feed are expected to include:

- TfNSW Transport Info (http://www.transportnsw.info);
- Realtime Transport App providers (http://www.transportnsw.info/en/travelling-with-us/keep-updated/apps/real-time-transport.page);
- Google Maps, Apple Maps, other online services.

General Transit Feed Specification (GTFS) – And GTFS-Realtime Feed for Sydney Light Rail Consumer Guide **Page 5 of 24**
Light Rail GTFS Timetables - Consumer Documentation - V1.0.docx

OFFICIAL

1.3   Definition of Terms Used

Term	Definition
GTFS	The **General Transit Feed Specification** (GTFS) defines a common format for public transportation schedules and associated geographic information.
GTFS-R	**GTFS-realtime** is a feed specification that allows public transportation agencies to provide realtime updates about their fleet to application developers. It is an extension to GTFS.

2   Background

2.1   Lifecycle of GTFS Light Rail Timetable Feed

A brief overview of the process to generate the GTFS Light Rail Timetable feed.

- Sydney Light Rail operates and manages their service timetable in their source of truth system (HASTUS).

- When a change to service timetables is required the relevant updates are managed within HASTUS and a GTFS export process is initiated to generate the required GTFS data feed for delivery to Transport (via the GTFS Server). This process will not require Transport's intervention. Similarly real-time data feeds (GTFS-R) will be generated independently of TfNSW as the static timetable data (GTFS) will be generated in-house.

- Filesets are transferred to the Open Data Portal via internal processes.

3   General Technical Notes

3.1   GTF Specification Compliance

3.1.1   Variations

The fileset includes definitions for certain fields that are different to the GTF Specification:

File	Field	GTFS Definition	TfNSW Definition
routes.txt	route_desc	The route_desc field contains a description of a route. Please provide useful, quality information. Do not simply duplicate the name of the route. For example, "*A trains operate between Inwood-207 St, Manhattan and Far Rockaway-Mott Avenue, Queens at all times. Also from about 6AM until about midnight, additional A trains operate between Inwood-207 St and Lefferts Boulevard (trains typically alternate between Lefferts Blvd and Far Rockaway).*"	Describes the 'Transport Network' that the Route belongs to. Defined by **TfNSW**. For example: • "Sydney Light Rail Network"

3.1.2   Extensions

The fileset includes additional fields that do not for part of the GTF Specification:

File	Field	Purpose	Contents
stop_times.txt	stop_note	To reference text contained in notes.txt, being Notes provided by Operators for each Stop, to be available to Customers.	Freetext
trips.txt	trip_note	To reference text contained in notes.txt, being Notes provided by Operators for each Trip, to be available to Customers.	Freetext
trips.txt	route_direction	To reflect the Route Direction Name for each Trip, to be visible to Customers in Timetables.	Freetext
notes.txt	note_id note_text	To reflect note text for Stop Notes and Trip Notes. Entire file is an extension to GTFS.	Freetext

3.2   Publication Cycle & Scope

3.2.1   Validity Period and Timetable Projection

Fileset will contain:

• timetables commencing the day of generation
• a minimum of 100 days of timetables

3.2.2   Generation Triggers

GTFS Filesets are generated to reflect service changes on the Light Rail Network on an ad hoc basis

3.3   Standards Applied

3.3.1   Identifiers

Many of the identifiers used within the feed include dashed (-) to join elements. Care may need to be taken to consider this when using the data or developing applications.

3.3.2   Value Quoting

All values in the text files are double-quoted. This includes numeric values and empty values. For example (from agency.txt):

```
"agency_id","agency_name","agency_url","agency_timezone","agency_lang","agency_phone"
,"agency_fare_url","agency_email"
"LR","Sydney Light
Rail","http://transportnsw.info","Australia/Sydney","EN","131500","",""
```

3.4   File Structure

3.4.1   Single Fileset Contents:

Each fileset will be a 'ZIP' format compressed archive – a .zip file - containing 9 text files.

Each file within the .zip file is a comma delimited / comma separated format file with the ".txt" extension.

For example:

| Name	Type
agency.txt	Text Document
calendar.txt	Text Document
calendar_dates.txt	Text Document
notes.txt	Text Document
routes.txt	Text Document
shapes.txt	Text Document
stops.txt	Text Document
stop_times.txt	Text Document
trips.txt	Text Document

3.5   File Size

Compressed (.zip) filesets range in size from 100KB to 400 KB.
Filesets are unlikely to ever be larger than 1MB.

General Transit Feed Specification (GTFS) – And GTFS-Realtime Feed for Sydney Light Rail Consumer Guide   **Page 9 of 24**
Light Rail GTFS Timetables - Consumer Documentation - V1.0.docx

OFFICIAL

4 GTFS Timetable Feed for NSW Light Rail

4.1 Overview

The following diagram provides a physical data model view of the relationship between the TfNSW GTFS Timetable Fileset files and fields.
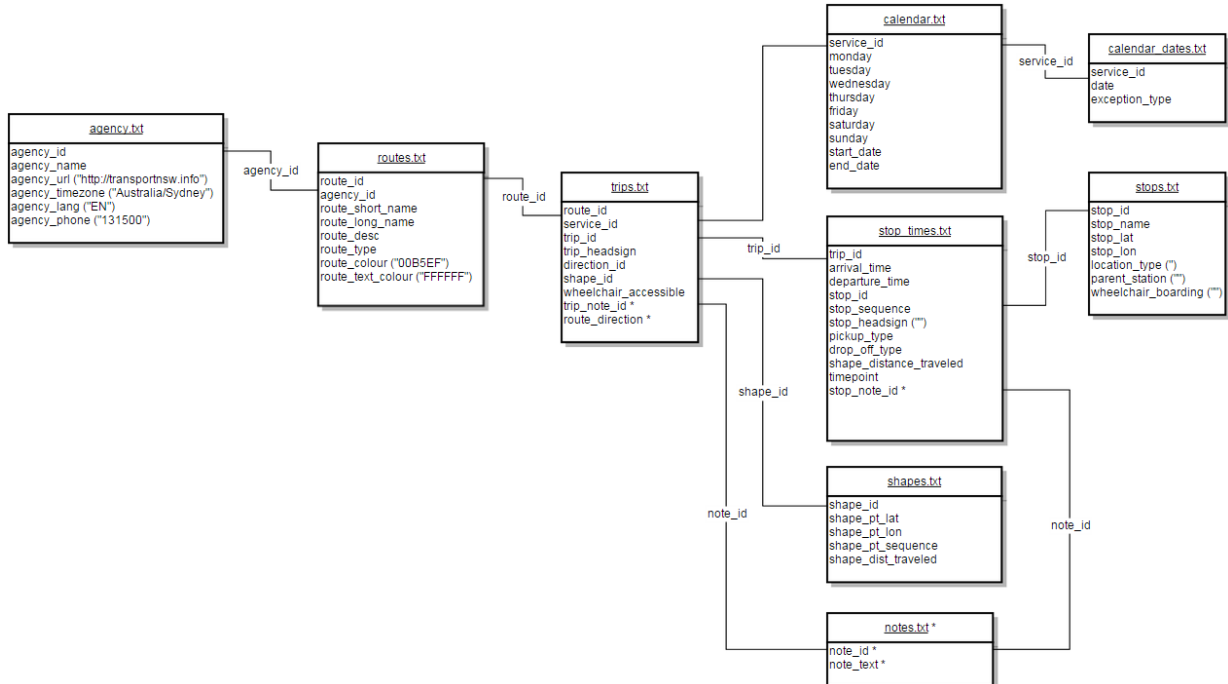
Understanding this diagram:

- This is a representation of the relationships between the components of the GTFS Timetable fileset as implemented by TfNSW.
- "*" denotes a TfNSW implemented extension on the GTFS standard.
- GTFS files and columns not implemented by TfNSW are not shown.
- Columns showing ("text") indicates hardcoded content.

4.2 agency.txt

Defines one or more transit agencies (Operators) that provide the data in this feed.

Field	Included	TfNSW GTFS Definition
agency_id	**Yes**	The Agency ID, allocated by TfNSW. For Example: "**LR**"
agency_name	**Yes**	The Operator's 'Customer Facing Name'. Note that this is not necessarily the same as the legal entity name. For example: "**Sydney Light Rail**"
agency_url	**Yes**	For all: "**http://transportnsw.info**"
agency_timezone	**Yes**	For all: "**Australia/Sydney**"
agency_lang	**Yes**	For all: "**EN**"
agency_phone	**Yes**	For all: "**131500**"
agency_fare_url	**No**	N/A
agency_email	**No**	N/A

4.3   calendar_dates.txt

Defines exceptions for the service IDs defined in the calendar.txt file.

Field	Included	TfNSW GTFS Definition
service_id	**Yes**	The internally generated ID that identifies a set of dates when a service exception is available for one or more routes. This identifier will be unique only within a GTFS Fileset, for example for a single Contract. This number has no meaning outside the Fileset.
date	**Yes**	The date field specifies a particular date when service availability is different than the norm. For Example: "20160315"
exception_type	**Yes**	Indicates whether service is available on the date specified in the date field. A value of 1 indicates that service has been added for the specified date. A value of 2 indicates that service has been removed for the specified date.

4.4   calendar.txt

Dates for service IDs using a weekly schedule. Specify when service starts and ends, as well as days of the week where service is available.

Field	Included	TfNSW GTFS Definition
service_id	Yes	The service_id contains an ID that uniquely identifies a set of dates when service is available for one or more routes. This identifier will be unique only within a GTFS Fileset, for example for a single Contract. This number has no meaning outside the Fileset.
monday	Yes	Each day field contains a binary value that indicates whether the service is valid for all of that day. A value of 1 indicates that service is available for all of that day in the date range. (The date range is specified using the start_date and end_date fields.) A value of 0 indicates that service is not available on that day in the date range. Note: Exceptions for particular dates, such as public holidays and service differences between school term and non-school term, are represented for the service_id in the calendar_dates.txt file.
tuesday	Yes	
wednesday	Yes	
thursday	Yes	
friday	Yes	
saturday	Yes	
sunday	Yes	
start_date	Yes	The start_date field contains the start date for the service.
end_date	Yes	The end_date field contains the end date for the service. This date is included in the service interval.

4.5 routes.txt

Transit routes. A route is a group of trips that are displayed to customers as a single service.

Field	Included	TfNSW GTFS Definition
route_id	Yes	The route_id field contains an ID that uniquely identifies a route. The route_id is dataset unique. For example: "**IWLR-191**"
agency_id	Yes	The TfNSW agency ID, allocated by TfNSW. For Example: "**LR**"
route_short_name	Yes	The short code identifying the Route to the public. For example: "**L1**"
route_long_name	Yes	The long name identifying the Route to the public. For example: "**Dulwich Hill Line**"
route_desc	Yes	*Note: TfNSW Variation.* Indicates the Transport Network that the Route belongs to, as defined by **TfNSW** For example: "**Sydney Light Rail Network**".
route_type	Yes	Describes the type of transportation used on a route "**0**": Tram, Streetcar, Light Rail
route_url	No	N/A
route_color	Yes	**"EE343F"**
route_text_color	Yes	**"FFFFFF"**

4.6   shapes.txt

Rules for drawing lines on a map to represent a transit organization's routes.

Field	Included	TfNSW GTFS Definition
shape_id	**Yes**	The shape_id field contains an ID that uniquely identifies a shape.
shape_pt_lat	**Yes**	These fields associate a shape point's longitude and latitude with a shape ID. The field values must be valid WGS 84 values from -180 to 180. Each row in shapes.txt represents a shape point in the trip's shape definition.
shape_pt_lon	**Yes**	For example: **"-32.8407","151.3551139"**
shape_pt_sequence	**Yes**	The shape_pt_sequence field associates the latitude and longitude of a shape point with its sequence order along the shape. The values for shape_pt_sequence will be non-negative integers, and they will increase along the trip. These numbers have no meaning outside the trip point sequence.
shape_dist_traveled	**Yes**	The shape_dist_traveled field positions a shape point as a distance traveled along a shape from the first shape point. The shape_dist_traveled field represents a real distance traveled along the route in **meters**. The values used for shape_dist_traveled will increase along with shape_pt_sequence: they cannot be used to show reverse travel along a route.

4.7 stop_times.txt

Times that a vehicle arrives at and departs from individual stops for each trip.

Field	Included	TfNSW GTFS Definition
trip_id	**Yes**	The trip_id field contains an ID that identifies a trip. This value is referenced from the trips.txt file. This ID is internally generated by the Light Rail service provider.
arrival_time	**Yes**	The arrival time at a specific stop for a specific trip on a route. Times for trips starting before 04:00 am will be expressed in '36 hour format'. For example: "**25:07**" (01:07 am)
departure_time	**Yes**	The departure time from a specific stop for a specific trip on a route. Times for trips starting before 04:00 am will be expressed in '36 hour format'. For example: "**25:09**" (01:09 am)
stop_id	**Yes**	The TSN ID for the Stop. This uniquely identifies a stop. This ID is the unique Transit Stop Number that identifies a stop in the **Transit Stop Management System ('TSM')**, the original Transit Stop reference resource maintained by TfNSW. For example: "**2000257**"
stop_sequence	**Yes**	The sequence of the Stop within the Trip. The stop_sequence will be a non-negative integer, and will increase along the trip. For example: "**3**" indicates that the stop is the 3rd in the Trip.
stop_headsign	**Yes**	For all: "" (i.e. null)
pickup_type	**Yes**	Indicates whether the Stop is for Pickup. For example: "**0**" indicates that regularly scheduled pickup occurs at the Stop for the Trip.
drop_off_type	**Yes**	Indicates whether the Stop is for Pickup. For example: "**1**" indicates that no drop off available for the Stop for the Trip.
shape_dist_travel ed	**Yes**	The shape_dist_traveled field positions a shape point as a distance traveled along a shape from the first shape point. The shape_dist_traveled field represents a real distance traveled along the route in **meters**. The values used for shape_dist_traveled will increase along with stop_sequence: they cannot be used to show reverse travel along a trip.
timepoint	**Yes**	Indicates whether the Stop is a timing point for the Trip. TfNSW consider all Light Rail stops to be timing points For example: "**1**" indicates that the Stop is a Timing Point for the Trip.

General Transit Feed Specification (GTFS) – And GTFS-Realtime Feed for Sydney Light Rail Consumer Guide
Light Rail GTFS Timetables - Consumer Documentation - V1.0.docx

OFFICIAL

Page 16 of 24

Field	Included	TfNSW GTFS Definition
stop_note	**Yes**	***Note: TfNSW Extension.***
		To reflect Notes provided by Operators for each Stop, to be available to Customers.
		May be empty.
		The value is referenced from the notes.txt file.
		For example: "**2143**" refers to notes.txt note_txt "**Stops only on request**".

<u>*Additional notes:*</u>

- Arrival and departing times should be used verbatim from the GTFS Timetable Feed. **Refer 5.1 below**.

4.8 stops.txt

Individual locations where vehicles pick up or drop off passengers.

Field	Included	TfNSW GTFS Definition
stop_id	**Yes**	The Transit Stop Number (TSN) ID for the Stop. This uniquely identifies a stop. This ID is the unique TSN that identifies a stop in the **Transit Stop Management System (TSM)**, the original Transit Stop reference resource maintained by TfNSW. For example: "**2000257**"
stop_code	**No**	N/A
stop_name	**Yes**	The correct name of the Stop. This Stop name is the 'official' name for the Stop as recorded in the TSM System. For example: "**Central Station Light Rail**"
stop_desc	**No**	N/A
stop_lat	**Yes**	The WGS 84 longitude ('y' coordinate') for the stop. For example: "**-33.882378**"
stop_lon	**Yes**	The WGS 84 longitude ('x' coordinate) for the stop. For example: "**151.206724**"
zone_id	**No**	N/A
stop_url	**No**	N/A
location_type	**Yes**	For all: "**0**" Indicates that the stop is a 'Stop' as opposed to a 'Station'.
parent_station	**Yes**	For all: "" (i.e. null).
stop_timezone	**No**	N/A
wheelchair_boarding	**Yes**	For all: "**1**" Indicates that the stop is accessible
platform_code	**Yes**	Light Rail stop do not have platform numbers visible to customers. The GTFS bundle contains indicative values to show the direction of travel. "**1**": Towards Central "**2**": Towards Dulwich Hill

4.9 trips.txt

Trips for each route. A trip is a sequence of two or more stops that occurs at specific time.

Field	Included	TfNSW GTFS Definition
route_id	**Yes**	The route_id field contains an ID that uniquely identifies a route. The route_id is dataset unique. The value is referenced from route.txt. For example: "**IWLR-191**"
service_id	**Yes**	The service_id contains an ID that uniquely identifies a set of dates when service is available for one or more routes. This identifier will be unique only within a GTFS Fileset, for example for a single Contract. This number has no meaning outside the Fileset.
trip_id	**Yes**	The trip_id field contains an ID that identifies a trip. This value is referenced from the trips.txt file. This ID is internally generated by the Light Rail service provider
trip_headsign	**Yes**	The trip_headsign field contains the text that appears on a sign that identifies the trip's destination to passengers. For example: "**Central Station**"
trip_short_name	**No**	N/A
direction_id	**Yes**	Indicates the direction (outgoing, inbound) of the Trip. Possible values are: • "0": Outbound • "1": Inbound
block_id	**Yes**	Identifies the block to which the trip belongs. A block consists of two or more sequential trips made using the same vehicle, where a passenger can transfer from one trip to the next just by staying in the vehicle. Eg: "**P-1-729828**"
shape_id	**Yes**	Contains an ID that defines a shape for the trip. This value is referenced from the **shapes.txt** file.
wheelchair_accessible	**Yes**	Indicates wheelchair accessibility for a scheduled trip. All Light Rail trips are considered wheelchair accessible Possible values are: • "**1**": Accessible service.
bikes_allowed	**No**	• "**1**": Indicates that the vehicle being used on this particular trip can accommodate at least one bicycle
trip_note	**Yes**	***Note: TfNSW Extension.*** To reflect Notes provided by Operators for each Trip, to be available to Customers. May be empty.

Field	Included	TfNSW GTFS Definition
route_direction	**Yes**	***Note: TfNSW Extension.*** To describe the Route Direction for each Trip, to be visible to Customers in Timetables. Enables grouping of Trips together in a meaningful way. For example: "**Dulwich Hill to Central**"

General Transit Feed Specification (GTFS) – And GTFS-Realtime Feed for Sydney Light Rail Consumer Guide   **Page 20 of 24**
Light Rail GTFS Timetables - Consumer Documentation - V1.0.docx

OFFICIAL

4.10   notes.txt

This file is an extension on the GTFS Fileset standard.

Contains a list of notes referenced from trips.txt and stop_times.txt. At the time of publication this file is unpopulated.

Field	Included	TfNSW GTFS Definition
note_id	**Yes**	Unique ID for Notes referenced in trips.txt (trip_note) and stop_times.txt (stop_note).
note_txt	**Yes**	Text for note used in stop_times.txt and trips.txt. A single Note may be re-used multiple times.

5 GTFS Realtime

5.1 Overview

"GTFS-realtime is a feed specification that allows public transportation agencies to provide realtime updates about their fleet to application developers. It is an extension to GTFS (General Transit Feed Specification), an open data format for public transportation schedules and associated geographic information. GTFS-realtime was designed around ease of implementation, good GTFS interoperability and a focus on passenger information."

- 'What is GTFS-realtime', Google Realtime Transit Overview (https://developers.google.com/transit/gtfs-realtime/)

The TfNSW Light Rail GTFS Realtime feed is generated by Transdev based on information from the Supervisory Control and Data Acquisition (SCADA) system used to monitor vehicle activity.

5.2 2. GTFS-R Trip updates

Trip updates will only be present for trips that have already commenced. When a vehicle is detected on the network it will be dynamically linked to the closest matching trip in the static timetable and assigned a trip ID. Therefore trip updates will not be present for Dulwich Hill to Central services, or outbound trips from Central, Capitol Square or Paddy's Markets. Propagation of delay via blocking to these stops may be performed. However it should be noted that services will generally not leave early from Central if they arrive ahead of schedule, they will wait until they are due to depart. Any logic built into propagating delays to subsequent trips should take this dwell time into account.

```
header {
  gtfs_realtime_version: "1.0"
  incrementality: FULL_DATASET
  timestamp: 1469691525
}
entity {
  id: "1"
  trip_update {
    trip {
      trip_id: "3849598-7793-UP-66"
      schedule_relationship: SCHEDULED
      route_id: "IWLR-191"
    }
    stop_time_update {
      stop_sequence: 17
      arrival {
        delay: -108
      }
      departure {
        delay: -108
      }
      stop_id: "200941"
      schedule_relationship: SCHEDULED
    }
  }
}
```

5.3　3. GTFS-R Vehicle Position

5.3.1　3.1 Coverage

Vehicle coverage is based off track readers. Currently no readers are present along the on-street track section which begins between Paddy's Markets and Exhibition Centre and continues to Central. No vehicle positions will be available in this section.

5.3.2　3.3 VehicleID

Vehicle ID will always be the same as trip ID

5.3.3　3.4 Example

```
header {
  gtfs_realtime_version: "1.0"
  incrementality: FULL_DATASET
  timestamp: 1469691997
}
entity {
  id: "1"
  vehicle {
    trip {
      trip_id: "3849599-7793-UP-67"
      schedule_relationship: SCHEDULED
      route_id: "IWLR-191"
    }
    position {
      latitude: -33.8708
      longitude: 151.1922
      bearing: 301.8259
    }
    current_stop_sequence: 16
    current_status: IN_TRANSIT_TO
    timestamp: 1469691997
    congestion_level: UNKNOWN_CONGESTION_LEVEL
    stop_id: "200939"
    vehicle {
      id: "3849599-7793-UP-67"
    }
  }
}
```

5.4　Dynamic Linking

Scenarios and corresponding behaviour of the feed:

a) All vehicles run on, ahead or behind schedule – i.e. no trips are cancelled or added in

When a vehicle activates a track section, the application looks up that track section and the next trip activity that should have taken place on that track section. The trip delay/ahead of time is then calculated.

b) A vehicle is removed from the track with no replacement and the trip gets cancelled.

Cancelled trips are flagged in the backend system to indicate it is out of service. When such a vehicle activates a track section the application looks up the track section and the next trip activity that should have taken place on that track section. This trip is then marked as cancelled.

c) A vehicle is removed from the track and another vehicle is added to the track as a replacement.

When a vehicle is flagged as cancelled as in case b) above, the remainder of the matched trip is marked as cancelled. This will cause the newly added vehicle to trigger the *next* trip's sections, and the following trip will trigger the next trip after that etc.

d) A vehicle is added to the track as an additional trip.

When a vehicle is added as an addition on the track, it will activate the trip track section of the vehicle that follows and therefore each vehicle will activate the following vehicle's trip track sections.