THE DEFINITIVE GUIDE TO GTFS

Quentin Zervaas

The Definitive Guide to GTFS

Consuming open public transportation data with the General Transit Feed Specification

Quentin Zervaas

About This Book

This book is a comprehensive guide to GTFS – the General Transit Feed Specification. It is comprised of two main sections.

The first section describes what GTFS is and provides details about the specification itself. In addition to this it also provides various discussion points and things to consider for each of the files in the specification.

The second section covers a number of topics that relate to actually using GTFS feeds, such as how to calculate fares, how to search for trips, how to optimize feed data and more.

This book is written for developers that are using transit data for web sites, mobile applications and more. It aims to be as language-agnostic as possible, but uses SQL to demonstrate concepts of extracting data from a GTFS feed.

About The Author

Quentin Zervaas (@HendX on Twitter) is a software developer from Adelaide, Australia.

He is the creator of the iOS & Android app TransitTimes+ (http://transittimesapp.com), which provides public transportation information in Australia, New Zealand, Canada and the United States. It uses many sources of data, including many GTFS and GTFS-RealTime feeds.

Quentin also created TransitFeeds.com, a web site that provides a comprehensive listing of public transportation data available around the world. This site is referenced various times throughout this book.

Credits

Technical Reviewer

Rupert Hanson (@rpy, http://triptasticapp.com)

Copy Editor

Miranda Little

The Definitive Guide to GTFS

http://gtfsbook.com

First Edition. Published in February 2014.

Table of Contents

1. Introduction to GTFS

GTFS (General Transit Feed Specification) is a data standard developed by Google used to describe a public transportation system. Its primary purpose was to enable public transit agencies to upload their schedules to Google Transit so that users of Google Maps could easily figure out which bus, train, ferry or otherwise to catch.

> **A GTFS feed is a ZIP file that contains a series of CSV files that list routes, stops and trips in a public transportation system.**

This book examines GTFS in detail, including which data from a public transportation system can be represented, how to extract data, and explores some more advanced techniques for optimizing and querying data.

The official GTFS specification has been referenced a number of times in this book. It is strongly recommended you are familiar with it. You can view at the following URL:

https://developers.google.com/transit/gtfs/reference

Structure of a GTFS Feed

A GTFS feed is a series of CSV files, which means that it is trivial to include additional files in a feed. Additionally, files required as part of the specification can also include additional columns. For this reason, feeds from different agencies generally include different levels of detail.

Note: The files in a GTFS feed are CSV files, but use a file extension of `.txt`.

A GTFS feed can be described as follows:

> **A GTFS feed has one or more routes. Each route (`routes.txt`) has one or more trips (`trips.txt`). Each trip visits a series of stops (`stops.txt`) at specified times (`stop_times.txt`). Trips and stop times only contain time of day information; the calendar is used to determine on which days a given trip runs (`calendar.txt` and `calendar_dates.txt`).**

The following chapters cover the main files that are included in all GTFS feeds. For each file, the main columns are covered, as well as optional columns that can be included. This book also covers some of the unofficial columns that some agencies choose to include.

2. Agencies (agency.txt)

*This file is **required** to be included in GTFS feeds.*

The `agency.txt` file is used to represent the agencies that provide data for this feed. While its presence is optional, if there are routes from multiple agencies included, then records in `routes.txt` make reference to agencies in this file.

`agency_id`	Optional
An ID that uniquely identifies a single transit agency in the feed. If a feed only contains routes for a single agency then this value is optional.	
`agency_name`	Required
The full name of the transit agency.	
`agency_url`	Required
The URL of the transit agency. Must be a complete URL only, beginning with `http://` or `https://`.	
`agency_timezone`	Required
Time zone of agency. All times in `stop_times.txt` use this time zone, unless overridden by its corresponding stop. All agencies in a single feed must use the same time zone. Example: `America/New_York` (See http://en.wikipedia.org/wiki/List_of_tz_database_time_zones for more examples)	
`agency_lang`	Required
Contains a two-letter ISO-639-1 code (such as `en` or `EN` for English) for the language used in this feed.	
`agency_phone`	Optional
A single voice telephone number for the agency that users can dial if required.	
`agency_fare_url`	Optional
A URL that describes fare information for the agency. Must be a complete URL only, beginning with `http://` or `https://`.	

Sample Data

The following extract is taken from the GTFS feed of TriMet (Portland, USA), located at http://transitfeeds.com/p/trimet.

agency_name	agency_url	agency_timezone	agency_lang	agency_phone
TriMet	http://trimet.org	America/Los_Angeles	en	(503) 238-7433

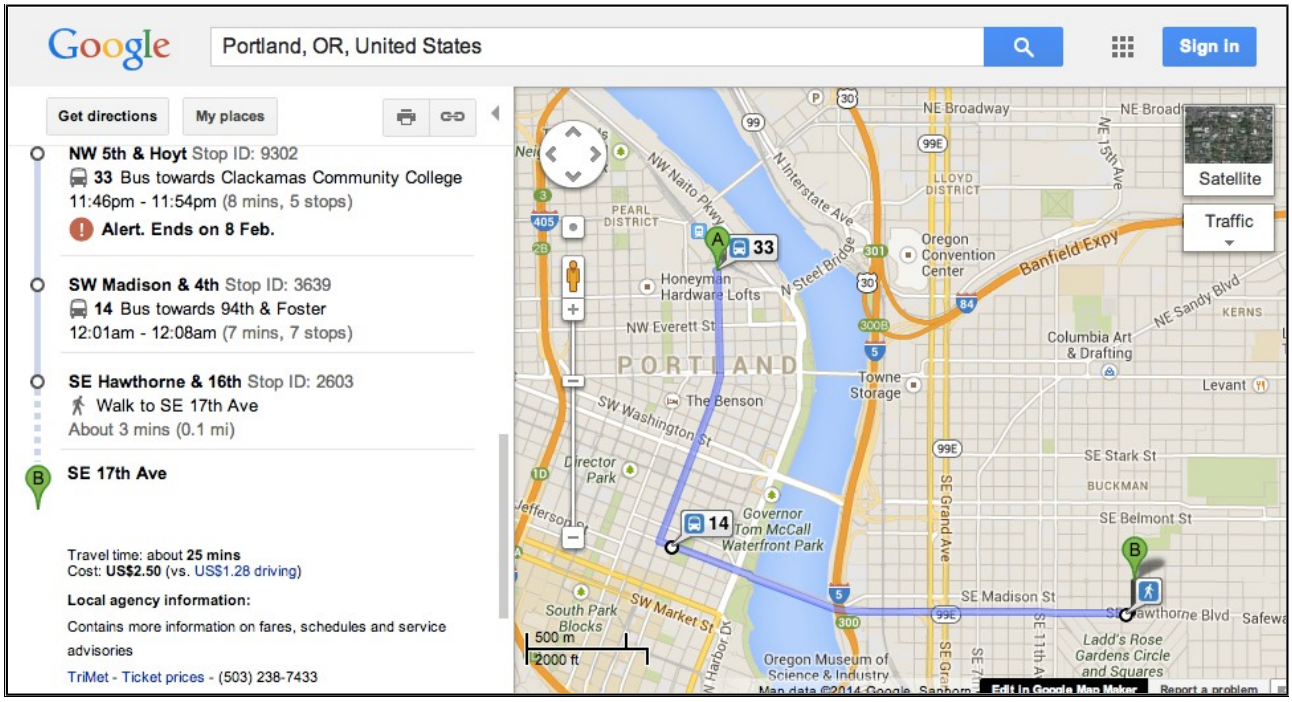
In this example, the `agency_id` column is included, but as there is only a single entry the value can be empty. This means the `agency_id` column in `routes.txt` also is not required.

Discussion

The data in this file is typically used to provide additional information to users of your app or web site in case schedules derived from the rest of this feed are not sufficient (or in the case of

`agency_fare_url`, an easy way to provide a reference point to users if the fare information in the feed is not being used).

If you refer to the following screenshot, taken from Google Maps, you can see the information from `agency.txt` represented in the lower-left corner as an example of how it can be used.



Thanks for reading a sample of

The Definitive Guide To GTFS

Download the full book today:

http://gtfsbook.com

Index

G

H

I

J

L

M

N

O

P

R

S

T

U

W

Z