A Data-Driven Approach To Bus Timetable Optimisation Recommendations.

Project Proposal

Jonathan Foot

psyjpf@nottingham.ac.uk

A project proposal for the degree of
MSci Computer Science

University of Nottingham
UK | CHINA | MALAYSIA

Background and Motivation

Bus transit systems remains the UK's leading public transportation method, with the average UK resident making 48 trips annually [5]. Bus networks allow for a relatively low initial infrastructural investment, while offering great adaptability and environmental benefits; particularly with the popularity of hydrogen, electric and bio-gas buses increasing. Although, private transportation methods still remain the UK's preferred method of transport, with the average person taking 602 trips by car [5], over 12 times that of the number of bus trips.

One of the leading deterrents for greater adoption of bus transportation is "a belief that buses cannot be relied on to stick to their timetables" [2]. Transport operators are responsible for the network route design, setting the frequencies, timetables and assigning vehicles and drivers to trips. When creating the timetable operators must consider both external factors, such as traffic, roadworks, severe weather, sporting events and internal factors, such as delays from previous trips affecting the subsequent trips and delays on passenger boarding and alighting [4].

The UK Senior Traffic Commissioner's guidelines are for 95% of buses to be "on-time", which is defined as arriving between 1 minute early to 5 minutes late at a timing point [1]. If bus operators repeatedly fail to adhere to this guideline, they can be fined up to £550 per vehicle [1]. One way to ensure compliance is to add a significant amount of excess wait times at stops to recoup for any delays. But this will increase journey times and unneeded waiting can be considered "dead time" where the bus operator is not making any money. Creating a reliable timetable is therefore important to both the operators and customers alike.

98% of buses in Great Britain are equipped with automatic vehicle location (AVL) devices [6], which is used to monitor the punctuality of buses to a high degree of accuracy. With some operators even storing the exact longitudinal and latitudinal position of every bus, up-to every 30 seconds. There is more data about buses than ever-before and with the United Kingdom Government legislating the Bus Service Act 2017 [3], bus open data is becoming increasingly more accessible. It is my intentions that this open-data can be used to drive an optimisation algorithm to improve average on-time performance and instil greater customer confidence in bus timetables and using buses as a form of transport.

There is already a large amount of research in this area and bus timetable setting can be described as an NP-Hard problem [7]. There are also many external factors to consider when creating a timetable, such as how does it connect with train times, school times, other bus routes and the number of available vehicles and drivers. It is therefore impractical to believe a single algorithm would be capable of generating a perfect new timetable but can instead be used to provide recommendations for bus operators to follow.

Aims and Objectives

The end-user of the program is bus operators, I aim to create a program which will take in their historic timetables and bus times and then output newly optimised timetables; which should increase the buses average on-time performance. The program will try to find weeks where performance patterns were similar and suggest multiple different timetables for different parts of the years, to deal with demands and conditions changing throughout the year; for example, school summer holidays.

I have decided to use the Reading Buses Open Data API, due to it being an exemplary bus open data source, providing a year's worth of historic data. As well, as having previous development experience using it and local knowledge of the routes and destination, which might aid my understanding for any erroneous data received.

The following lists my main objectives I intend to achieve:

1. Interface with the Reading Buses Open Data API, request and download data required from the API.

2. Search through the data for weekly patterns and group them, to create multiple-timetables for the year; one for each group.

3. Generate the performance metrics, such as average on-time performance of the old timetable for the last year.

   (a) Use these metrics, to compare and contrast against the new timetable later.

4. Take in the historic vehicle data and generate a new optimised timetable, to minute accuracy.

   (a) Explore multiple different methods for timetable optimisation.
   (b) The new timetable will try to minimise unneeded idle/dead time at timing points.
   (c) The new timetable will not consider constraints on number of vehicles, or drivers available.
   (d) The new timetable will not consider how these changes might affect connections between other routes timetables.
   (e) The new timetable must bear some resemblance to the original headway times and will be restricted to prevent an overzealous solution.
   (f) The new timetable cannot alter the bus services route design.

5. Produce a simulation/scoring system which can take in the proposed timetable and output estimated performance levels, to allow for comparison with original timetable.

6. Create a user friendly GUI, which can graphically highlight the differences between the two timetables, to easily distinguish where the improvements have been made.

Work Plan

My work-plan focuses in the first semester on creating the library to interface with the data-source API and then finding weekly service schedule patterns. In the second semester, I will start to develop the optimisation algorithm, a method for estimating a timetable's performance and an end-user friendly graphical user interface.

Any testing, such as unit tests and integration testing has been included inside of the development times allocated.

The work plan is only an estimation, the timeline will be re-evaluated at the start of the second semester, to add in additional work, or reschedule remaining work where needed.

Code	Title	Duration	Completion
1A	Write Draft Project Proposal	2	100%
1B	Complete Project Proposal	2	100%
1C	Complete Preliminary Research Ethics Checklist	1	100%
Milestone 1: Project Proposal Completed			
2A	Research similar projects	4	0%
2B	Develop the interface with API, download the data and make it usable for input into the program.	4	0%
Milestone 2: Data sourced			
3A	Investigate weekly patterns in a services schedule.	4	0%
Milestone 3: Found weekly patterns			
4A	Draft and write Interim Report	4	0%
Milestone 4: Interim Report Completed			
5A	Research optimisation algorithms and try out/prototype with some.	7	0%
5B	Added slack to catch up and for Holiday and Exam Period	4	0%
5C	Plan & Design Final Optimisation Algorithm Strategy	2	0%
Milestone 5: Planned Optimisation Algorithm Design			
6A	Develop optimisation algorithm and method of testing its simulated performance.	6	0%
6B	Design and Develop a user-friendly GUI.	3	0%
Milestone 6: Optimisation Algorithm Completed			
7A	Test all parts of the system for failures.	2	0%
7B	Evaluate and compare performance of the new timetable(s) generated.	2	0%
7C	Draft and Write Dissertation	9	0%
Milestone 7: Dissertation Completed			

Timeline columns span October through April (weeks 1–4 each). Key annotated dates: 28th Project Proposal (October), 9th Interim Report (December), 30th Final Dissertation (April).

COVID-19 Considerations and Mitigations

The historic bus data is retained for one year, after which it is removed from the API source. This means the majority of available data will be affected by COVID-19, which could result in usual or anomalous data. Particularly, where the official expected timetables were not updated with any temporary COVID-19 reduced timetable. Wherever possible, I shall aim to detect and remove anomalous data entries I suspect are a result of COVID-19.

It is possible I may become ill during the year and this may affect my ability to work. If so the work-plan will have to be adjusted to reflect for any lost time.

Going forwards, the project can be worked on remotely and I do not envisage any further issues, other than the continued aforementioned data quality issue.

Conflicts of Interest

I have previously been employed by Reading Buses as a freelance software-developer, the main data source provider for this project, but do not believe this shall have any effect on the overall project.

There are no other known conflicts of interest.

Bibliography

[1] Senior Traffic Commissioner. Traffic commissioners: local bus services in england (outside london) and wales guidance. page 15, 2020.

[2] Susan McConville Fiona Dobbie and Rachel Ormston. Understanding why some people do not use buses. *Scottish Government Social Research*, page 106, 2010.

[3] UK Goverment. Bus services act 2017. *UK Public General Acts*, 2017.

[4] The voice of urban transport PTEG. Bus punctuality. page 28, 2014.

[5] Department for Transport Statistics. Tsgb0103 (nts0303): Average number of trips, stages, miles and time spent travelling by main mode (ods). 2018.

[6] Department for Transport statistics. Percentage of buses with an automatic vehicle location (avl) device where the device is used to monitor punctuality by metropolitan area status and country, local bus operators only: Great britain. page 3, 2019.

[7] Yinghui Wu, Hai Yang, Jiafu Tang, and Yang Yu. Multi-objective re-synchronizing of bus timetable: Model, complexity and solution. *Transportation Research Part C: Emerging Technologies*, 67:149–168, 2016.