Bus bunching and Inter-route interactions: A Case Study of the 430 bus Route on Military Road

CIVL 3704 2020S1 Data Collection Project

SID: 450239612

1 BACKGROUND

The reliability of a bus route is a key factor contributing to the utilisation of public transport and the satisfaction of passengers. It has been shown that passengers value reliability up to four times as much as mean travel time (Hollander and Liu, 2008). The challenge of maintaining service reliability and schedule adherence is controlled by the stochastic environment within which the buses operate (Feng and Figliozzi, 2015). Delays caused by this stochasticity are amplified since commuters accumulate at bus stops which leads to increased boarding time when the bus arrives. The problem is compounded by passengers that would have boarded the following bus but instead become able to board the delayed bus. The bus behind is able to to make time savings due to the decreased number of passengers that board and catches up to the delayed bus. This progressive system breakdown characterises bus bunching (BB).

While much research has been conducted in establishing the interactions between buses and vehicles external to the system, interest in inter-route interaction is emerging. It was hypothesised by Zhou et al. (2019) that buses that run collinearly on sections of their routes have the potential to interact in a way that may aggravate bus bunching. In their research, Zhou et al. (2019) used numerical simulations to study the interactions as well as ways to manage them. However, questions regarding the field generated evidence of inter-route interaction as well as the way that interactions occur remain unanswered. Despite this, it is reasonable to assert that buses that share the same segments, the same lanes in traffic and some of the same stops will have non-negligible interactions.

The objective of this research is to identify the physical occurrence of inter-route interactions and describe the impact that bunching buses of one route have on the interactions with buses that belong to different routes. Specifically, instances of bunching of the 430 bus on Military Road (MTR) are studied in terms of the way they affect the level of delay and the 'bunching status' of nearby buses of different routes. The trips that travel towards Mosman are considered.

MTR is a 2.7 km length of road in Sydney stretching between Mosman and Neutral Bay with between two and three lanes in each direction. MTR serves commuters from a large catchment that includes the Northern Beaches and Eastern portion of Sydney's Lower North Shore. It serves as the primary passage from these areas to major centres such as the Sydney CBD, Chatswood, and North Sydney. To connect these centres forty-six bus routes utilise MTR as express, limited stops and all-stops buses. Major portions of MTR include a bus lane which increases the proportion of interactions traffic to be between buses, rather than private commuter vehicles.

Interactions with the 430 route is the subject of this report as it is a bus with relatively short headways (10 mins) during peak periods and since it is an all stops bus. These characteristics increase the queuing probability at each stop and thus increases number routes that interact (Yang et al., 2018).

2 METHODOLOGY

2.1 Data Description

The study period for data collection ranges from 4 May 2020 to 12 May 2020 and includes all (7) weekdays during that period. Transport for New South Wales (TfNSW), through the Open Data Hub, has made available GTFS-static and GTFS-real-time data sets. GTFS-static contains information related to scheduled spatial and temporal attributes that characterise a routes characteristics and define how trips should operate. GTFS-real-time provides a time-limited window of actual bus behaviour. On each

data collection day, real-time (RT) data was collected during two peak periods: (1) 07:00-10:00, and (2) 16:00-18:30.

2.1.1 GTFS-Static

A file map that describes the available GTFS-Static information as well as the connection between them is available in Appendix A, Figure 6. The important files for this research include:

- 'routes.txt' contains a list of all routes considered and includes agency/operator information and route descriptions
- 'calendar.txt' relates bus services to the days these services are run.
- 'trips.txt' each route in routes.txt is characterised by a list of trips. These trips belong to different services in calendar.txt. Trips.txt connects routes to services. For each route, there is a row for each bus that runs.
- 'stop_times.txt' contains all the stop times (departure and arrival times) as well as stop sequence information
- 'stops.txt' contains spatial information on every stop.

2.1.2 GTFS-Realtime

There are two flavours of GTFS-RT: (1) trip updates, and (2) vehicle positions.

Trip Updates

Vehicle Positions

Each instance that RT data is requested through the Open Data Hub, the response contains RT data for up to two hours in the past and predicted data for one hour beyond the query time. Data is recorded at every stop for buses that deviate from schedule within the time horizon. At each stop, arrival- and departure-time as well as the arrival- and departure-delay are recorded. Information for trip start time and start date is also reported with respect to the bus origin (not necessarily the first stop in the sequence). A sample of raw RT data is produced in Appendix B, Table 2.

The Vehicle Positions data set was not used in this research. Discussion of real time data will solely refer to Trip Updates.

2.1.3 Summary of Data Collected

Since GTFS-RT only contains information for buses that are deviating from schedule. Therefore, to observe the complete behaviour of the system, RT data needed to be merged with static data to regain trip information for trips that run on-time. The resulting data set was used to produce the statistics in Table 1.

	Number of Observations	Time interval of MTR data (hours)
Mean	6419	1.340
Max	7899	1.910
Min	5369	0.986
Standard Deviation	837.4	0.303

Table 1. Description of the data collected. There was significant variation in the time interval in which RT delay information was reported for each sample taken. This is indicative of the significant variability in the number of delays for each day that data was recorded.

2.1.4 Data Processing and Manipulation

Preparing STATIC Data

The stops on MTR that are part of the 'Sydney Bus Network' were determined using the *stops.txt*. A map showing the set-out of each stop along MTR is available in Appendix C, Figure 7. The trip id of each trip that stops on MTR was determine from the *stop_times.txt* which allowed the determination of all the routes along MTR.

The time interval for available RT data was used to determine the time interval for the static data that should be sampled. The nominal headway was calculated for all the buses within this sample time period. A buffer time period was also applied to ensure that all the buses within the start and finish times had

nominal headway values. This is important because the headway of the first trip in a data set is not known because there is no data on the bus before it.

Headways were calculated by combining *trips.txt* and *stop_times.txt* into a data frame. At each stop the headways for each trip were calculated. This was done by subtracting the arrival time of each bus from the previous arrival time of a bus from the same route. This was done so that bus stops that are 'time points' do not mask the level of bunching of buses that arrived at the stop thus allowing a more direct measurement of the system performance. can be calculated independent of dwell time. Prior to merging in RT data, the data frame contained the following columns: route_id, trip_id, stop_sequence, stop_id, direction_number, scheduled_arrival_time_dt and headway_nominal.

It should be noted that the 'direction_number' corresponds to the direction that the buses travels on MTR. It is different to the direction_id column in GTFS-static. The direction number is defined as '0' for buses travelling North-East (towards Mosman) and '1' for buses travelling South-West (away from Mosman). Only data for direction number '0' is considered in this study.

Introducing RT Data

Before RT data was introduced four more columns were added to the data frame: arr_delay, dep_delay, actual_arrival_time_dt, headway_actual. These columns were initialised with values that would correspond to a bus that is running on time. That is, there would be a zero in the delay columns, and, actual arrival time and actual headway was set to scheduled arrival time and nominal headway, respectively. RT data was merged with Static data using a loop that replaced the nominal arrival and delay data for each row where there was a matching (trip_id, stop_id) pair. The actual headway was calculated in a similar way as for the GTFS-static data – that is, for each trip id, calculating the time since the last bus of the same route was at each stop. This allowed for the possibility that buses change their order along a route. The final product had the following columns: route_id, trip_id, stop_sequence, stop_id, direction_number, scheduled_arrival_time_dt, headway_nominal, arr_delay, dep_delay, actual_arrival_time_dt, headway_actual, bus_infront_tripID.

2.2 Approach

All data processing, manipulation and visualisation was completed using python. Python packages that were used include *pandas*, *numpy*, *datetime*, *geopy* and *matplotlib*.

The data processing procedure occurred in three stages. First, a metric for bus bunching was defined and instances of bus bunching were identified for each route. Second, the local level of bunching was calculated at each stop for each bus considering the 430 buses that also stop at that stop within a specified time period. Third, inter-route interactions between the 430 bus and nearby routes were investigated.

2.2.1 Bunching Criteria and Metrics

BB is not defined by a unified standard. One approach to defining BB is to establish minimum threshold headway. For example, if the headway is less than three minutes as was used by Feng and Figliozzi (2011) the bus behind was denoted as bunched. BB can also be defined by the distance between buses as was done by Egge (2017) whereby buses closer than 250 m were denoted as bunched. In this research, bunching is defined using two metrics that depend on the headway. The definition of each metric is as follows:

- Metric 1: BB occurs when the actual headway is less than a quarter of the nominal headway. Such as situation has a bus bunching designation of 1. This is a threshold-based definition with only two possible outcomes either a bus is bunching, or it is not.
- Metric 2: The level of bus bunching is quantified using a continuously differentiable function that varies from 0 to 1 as the ratio of the actual headway to the nominal headway varies from 0.5 to 0. The function used is an ellipse to reflect the importance of buses with small headways.

An example showing the variation of each BB metric is available in Figure 1 for the case where the nominal headway is ten minutes. Each metric is useful in different situations. The application of each metric is described in Section 2.2.2.

2.2.2 Inter-Route Interaction Criteria

The time horizon that was used to look for interactions spanned two minutes ahead and three minutes behind the arrival of each bus at each stop. The time for considering the 430 buses yet to arrive is larger to

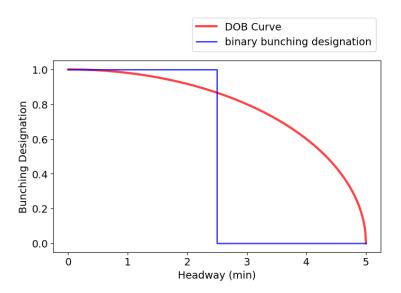


Figure 1. Variation of bus bunching metrics used for a bus with a nominal headway of ten minutes. Metric 1 (threshold bunching criteria) is a step function and Metric 2 is a smooth ellipse. Metric 2 allows the level of bunching to be continuously defined as a variable between 0 and 1.

consider the possibility that passengers will get on the earlier bus (that is not a 430) if the destination is also served by the bus in question. In any case, this criteria is not fundamentally based and it is possible that a more reasonable time horizon should be applied. For each non-430 route, the bunching designation (Metric 1) and DOB (Metric 2) for each of the nearby 430 bus was summed. The result of this summation is denoted as the 'local cumulative BB designation' (Metric 1) or the 'local cumulative DOB' (for Metric 2). It is assumed that the effect of multiple nearby bunching buses is linearly related to the overall local effect of those bunching buses on the system. The local cumulative bunching effect was compared to the change in delay and bunching status of non-430 routes at each stop. Different metrics were used for these comparisons. The metrics used are justified in the Section 3.

3 RESULTS

Information for the characteristics of each route in the study including, route type, number of stops and average headway was collected. A route characteristics summary for the sampling period is presented in Appendix D, Table 3.

3.1 Bus Bunching Observations

Each instance of BB for routes that use MTR was recorded. The level of bunching that occurred for each route along their routes and while on MTR is provided in Figure 2. The distribution of bunching among the MTR stops that experienced bunching is shown in Figure 3.

3.2 Correlating 430-route Bunching Status with Changes in Delay of Nearby Routes

The change in delay at each stop was calculated for each bus by subtracting the departure delay from the arrival delay. Therefore, a positive change in delay corresponds to an increase in overall delay (or less ahead of schedule) and a negative change in delay corresponds to a decrease in overall delay (or more ahead of schedule). It should be noted that decreases in delay were not reported in GTFS-RT data.

The change in delay for each bus, at each stop, could be related to whether a bunched 430 bus was nearby. The result of this comparison produced using Metric 1 is presented in Figure 4.

The largest magnitude delay was recorded in Figure 4(b). It is worth noting that the amount of data that was available for producing Figure 4(b) was significantly larger than the data used in Figure 4(a). As such, the extreme value of delay change in Figure 4(b) is considered as an outlier in the data and is ignored.

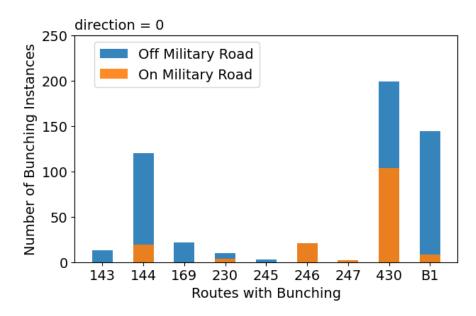


Figure 2. The number of bus bunching instances (Metric 1) for buses with identified bunching travelling in direction 0 (towards Mosman). A significant proportion of 430 BB instances occurs while on MTR and the most prevalent bus type for direction 0 bunching while on MTR is all-stops buses. Thus, all-stops buses will be the focus of the interaction analysis.

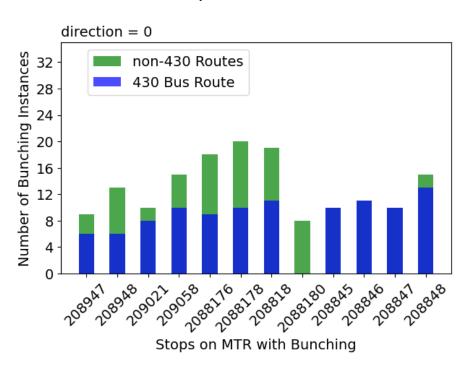


Figure 3. The total level of bus bunching (Metric 1) at each stop on military road is divided into 430 BB and non-430 BB instances. The stops are in order of appearance going from left to right as direction 0 buses travel towards Mosman. While the level of 430 BB instances dominates, BB among other buses is still distributed relatively evenly up to stop 2088178. This stop corresponds to a fork in the road where buses either continue down MTR or turn onto Spit Rd. This suggests that either a change in traffic conditions or decreased discordant interactions between other buses occurs after the fork.

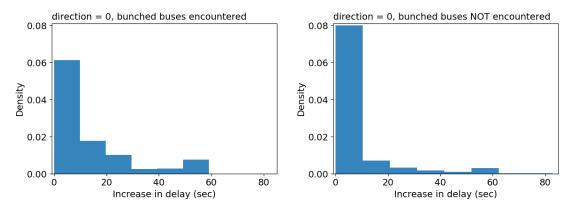


Figure 4. A comparison of the density distribution of delay for (a) buses near bunched 430 buses (Metric 2) and (b) buses near non-bunched 430 buses. Compared to (b), a larger proportion of the delays in (a) are non-zero and the relative magnitude of non-zero delays is larger for each bin in the histogram. This suggests that buses encountering a bunched 430 are more likely to experience an increase in delay. Note that an 'increase in delay' is used to denote a positive change in delay. That is, buses that are early become less early, or even late.

The use of the DOB metric (Metric 2) in Figure 4 means that a bunched 430 is considered to have been encountered if the 430 in question has a headway of less than half the nominal headway. This means that even relatively weak 430 bunching behaviour was recorded.

The change in delay as a result of bus bunching encounters shown in Figure 4(a) is described in terms of the local level of bus bunching that occurred locally in Figure 5.

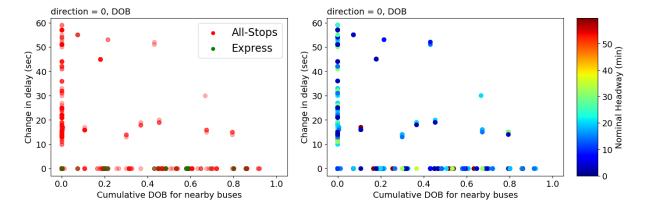


Figure 5. The local cumulative DOB is compared to the change in delay of buses that encounter bunched 430 trips. There are three primary features. There is a horizontal line, a vertical line and evidence of a pattern between a cumulative DOB of 0.3 and 0.8 for changes of delay between 10-20 seconds. The two lines indicate that an interaction has not taken place since the vertical line shows changes in delay in instances where zero BB was encountered and the horizontal line represents the instances where zero delay changes occurred in the presence a of non-zero local cumulative DOB. The subplots describe this relationship in terms of (a) bus type and (b) the nominal headway. This shows that all bus types and nominal headways are included in the two lines and that the third feature represents all-stops buses only with headways between about 10-25 minutes. This suggests that more information than the headway and bus type are required to predict the level of interaction.

The DOB metric was used for two reasons. First, the DOB captures more clearly the level of interaction by giving more weight to buses that have smaller headways for each possible 'actual headway' value, whereas Metric 1 applies a constant bunching designation for all possible 'actual headway' values that are less than a quarter of the nominal headway. Second, by using a bunching designation that is constant

over at least part of the possible actual headway values, as for Metric 1, data in Figure 5 is forced into groups of vertical lines or groups of distinct points at integer x-axis values. An example of this is provided Appendix E, Figure 8.

4 DISCUSSION

In general, BB is a phenomena that develops progressively and as a consequence of a series of delays as opposed to a single event. By relating instances where bunching is encountered to the change in delay Figure 4 is able to establish a clear link between the mechanism causing bunching, that is delay accumulation, and the presence of nearby 430 bunching. While buses that do not encounter bunched 430s also experience delays the relative magnitude of these delays is up to 50% smaller. It is important to note that Figure 4 does not account for larger dwell times to timing points. Therefore, it is likely that this relative difference between Figure 4(a) and (b) is even larger, which would have more severe implications for buses that encountered bunched 430 buses. However, Figure 5 shows that not all buses encountering bunched 430s experienced delays. This implies that there are different levels on which buses interact. This could be a consequence of a local traffic condition, attributes of a bus route or a result of assumptions made in establishing an interaction. It was assumed that an interaction takes place if a non-430 bus arrives at a stop either 2 minutes before or 3 minutes after each bus being considered. For example, it is possible that interactions are insignificant when buses arrive at stops before bunched buses actually arrive. It is also possible that certain attributes of a bus service result in a stronger or weaker interaction with the 430 bus. Further research is required into the parameters governing interaction to understand why some services do not interact at all, while other do. Despite this, some information about on the buses that interact can be drawn from Figure 5 through 'Feature 3' described in the caption, whereby buses that interact are all-stops buses and have headways that are between 10-20 minutes. These characteristics are similar to the 430 bus which has a mean headway of 10 minutes. However, since there are a large number of buses that are also all-stops buses with headways comparable to the 430 that do not interact, there must be other attributes of a bus service which are able to differentiate non-interacting and interacting buses.

While evidence for interactions can be extrapolated from Figure 5, distinct and strong relationship between the delay change the cumulative DOB is not apparent. Even though 'Feature 3' can be interpreted as being a result of inter-route interactions, arguments that assert these values as being due to noise are also plausible. However, this does not rule out the existence of a relationship between a bus' delay change and the cumulative local DOB. The corridor being studied as well as the number of stops for which buses share corridors are important factors. The level of BB for two buses that share a corridor is positively correlated with the length of the collinear section (Zhou et al., 2002). Therefore, it is possible that the 2.7 km section of MTR on which most routes collinear is not long enough to exhibit strong interactions. Further research is required to establish the minimum length of a collinear section that is required for significant interactions to take place.

Significantly, this research was conducted during the implementation of Stage 1 travel restrictions in response to the COVID-19 pandemic. Consequently, this time was characterised a significant decrease in commuter traffic, as well as decreased transit ridership while the number of buses operating remained unchanged. Less traffic means that it is easier for a bus to leave the bus lane to overtake another bus. In circumstances where overtaking is facilitated the level of bunching decreases(Schmöcker et al., 2016). In non-COVID-19 times this option would have been precluded by slow moving traffic in the centre and inside lanes. Moreover, at bus stops, when boarding and alighting is complete, it is easier for buses that are behind bunched buses to pull away into traffic for the same reason that lane changes become more viable – there is less commuter traffic to contend with. The contention here is that the presence of congestion forces 'healthy' buses to interact with bunched buses more directly and without this interaction-behaviour is lessened. As such, the findings of this research must be considered in the context in which data was collected. It is expected that clearer relationships that describe delay changes and interactions with bunched buses can be obtained using a similar methodology to this study in more regular circumstances.

5 ACKNOWLEDGMENTS

The author acknowledges the guidance and support of Professor Emily Moylan throughout this research. The CIVL3704 course tutors, Hema Rayaprolu and Bahman Lahoorpoor, had key supporting roles and their advice and the direction they provided was greatly appreciated.

REFERENCES

- Egge, M. (2017). The when and where of pittsburgh's bunched buses.
- Feng, W. and Figliozzi, M. (2011). Empirical findings of bus bunching distributions and attributes using archived avl/apc bus data.
- Feng, W. and Figliozzi, M. A. (2015). Empirical analysis of bus bunching characteristics based on bus avl/apc data.
- Hollander, Y. and Liu, R. (2008). Estimation of the distribution of travel times by repeated simulation. *Transportation Research Part C: Emerging Technologies*, 16(2):212–231.
- Schmöcker, J.-D., Sun, W., Fonzone, A., and Liu, R. (2016). Bus bunching along a corridor served by two lines. *Transportation Research Part B: Methodological*, 93:300–317.
- Yang, Q., Lin, L., Chen, J., and Wu, W. G. (2018). Study on the threshold of bus arrival rate of shared bus stops based on queuing probability. In *CICTP 2017: Transportation Reform and Change—Equity, Inclusiveness, Sharing, and Innovation*, pages 2773–2780. American Society of Civil Engineers Reston, VA.
- Zhou, X., Fan, W., and Shi, Y. (2002). Effects of collinear section length on bus bunching. *Journal of Transportation Engineering*.
- Zhou, X., Wang, Y., Ji, X., and Cottrill, C. (2019). Coordinated control strategy for multi-line bus bunching in common corridors. *Sustainability*, 11(22):6221.

6 APPENDIX

6.1 Appendix A

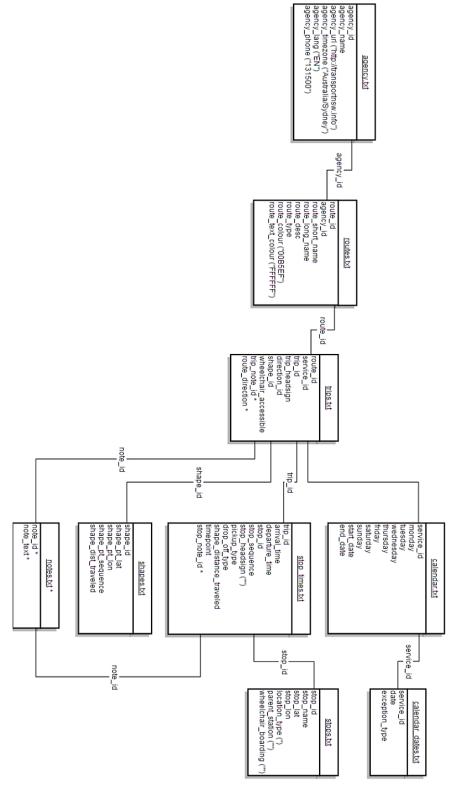


Figure 6. GTFS-STATIC File map and the relationships between files.

6.2 Appendix B

	trip_id	route_id	stop_id	arr_time	arr_delay	dep_delay	route_short_name
0	1060776	2440_144	206073	08:36:21	-279	-232	144
1	1060776	2440_144	2065108	08:37:28	-272	-256	144
2	1060776	2440_144	2065109	08:38:31	-389	-337	144
3	1060776	2440_144	206524	08:40:06	-354	-318	144
4	1060776	2440_144	206512	08:40:45	-375	-360	144
5	1060776	2440_144	206513	08:41:51	-429	-405	144
6	1060776	2440_144	206521	08:42:51	-429	-376	144
7	1060776	2440_144	206527	08:45:13	-527	-418	144
8	1059990	2440_244	2088102	08:34:52	-128	-114	244
9	1059990	2440_244	2088103	08:36:14	-166	-130	244
10	1059990	2440_244	2088240	08:40:10	-170	-170	244
11	899678	2440_B1	210120	08:37:44	-436	-399	B1
12	899678	2440_B1	210323	08:42:14	-466	-416	B1
13	978772	2440_263	200096	08:37:07	-233	-233	263
14	1060364	2440_178	2088185	08:35:27	-153	-137	178
15	1060364	2440_178	208811	08:36:43	-197	-184	178
16	1060364	2440_178	209211	08:38:17	-343	-338	178
17	1060364	2440_178	209311	08:40:14	-346	-317	178
18	1060364	2440_178	209312	08:41:02	-358	-326	178
19	1060364	2440_178	209313	08:41:44	-376	-366	178

Table 2. Sample of raw data in the form of a pandas data frame.

6.3 Appendix C

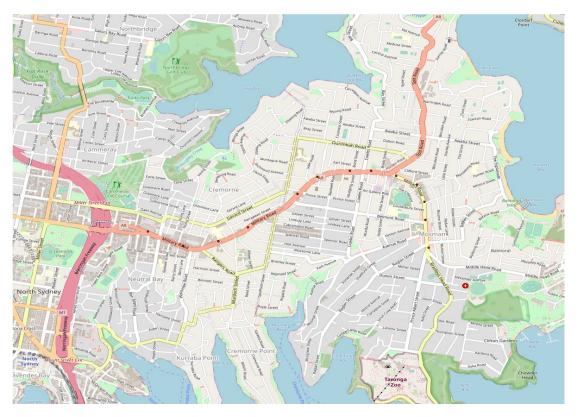


Figure 7. Set out of MTR stops for buses travelling in direction 0. Buses travelling in Direction 0 move left to right in the figure.

6.4 Appendix D

	route_short_name	route_long_name	bus_type	mean nominal hw (mins)
0	257	PrePay-Only - Balmoral to Chatswood via Crows	AS	21.66
1	143	PrePay-Only - Manly to Chatswood via Balgowlah	AS	21.41
2	144	PrePay-Only - Manly to Chatswood via Royal Nor	AS	16.61
4	168	PrePay-Only - North Balgowlah to Milsons Point	AS	29.75
5	169	PrePay-Only - Manly to City Wynyard via Narraw	AS	31.57
7	173	PrePay-Only - Narraweena to Milsons Point	AS	22.71
8	178	PrePay-Only - Cromer Heights to City Wynyard	AS	31.94
9	180	PrePay-Only - Collaroy Plateau to City Wynyard	AS	24.79
12	227	PrePay-Only - Mosman Junction to Milsons Point	AS	32.66
13	228	PrePay-Only - Clifton Gardens to Milsons Point	AS	40.63
14	229	PrePay-Only - Beauty Point to Milsons Point vi	AS	129.00
15	230	PrePay-Only - Mosman Wharf to Milsons Point vi	AS	14.53
16	236	PrePay-Only - South Mosman Wharf to Spit Junction	AS	21.92
17	243	PrePay-Only - Spit Junction to City Wynyard	AS	33.94
18	244	PrePay-Only - Chowder Bay Mosman to City Wynyard	AS	33.25
19	245	PrePay-Only - Balmoral to City Wynyard	AS	47.80
20	246	PrePay-Only - Balmoral Heights to City Wynyard	AS	8.43
21	247	PrePay-Only - Taronga Zoo to City Wynyard via	AS	30.05
22	248	PrePay-Only - Seaforth to City Wynyard	AS	18.42
23	249	PrePay-Only - Beauty Point to City Wynyard	AS	28.30
25	B1	PrePay-Only - B-Line Mona Vale to City Wynyard	EX	5.37
26	E50	PrePay-Only - Manly to Milsons Point (Express	EX	12.48
27	E54	PrePay-Only - Mona Vale to Milsons Point (Expr	EX	14.07
28	E65	PrePay-Only - South Curl Curl to City Wynyard	EX	14.96
29	E66	PrePay-Only - Allambie to City Wynyard (Expres	EX	30.63
30	E68	PrePay-Only - Brookvale to City Wynyard via No	EX	25.54
31	E69	PrePay-Only - Manly to City Wynyard via Narraw	EX	13.07
32	E70	PrePay-Only - Manly to City Wynyard (Express S	EX	13.77
33	E71	PrePay-Only - Manly to City Wynyard via Clonta	EX	30.49
34	E75	PrePay-Only - Warringah Mall to City Wynyard (EX	22.56
35	E76	PrePay-Only - Dee Why to City Wynyard via Nort	EX	27.92
36	E77	PrePay-Only - Dee Why to City Wynyard via Wing	EX	25.92
37	E78	PrePay-Only - Cromer Heights to City Wynyard (EX	20.79
38	E79	PrePay-Only - Wheeler Heights to City Wynyard	EX	14.62
39	E80	PrePay-Only - Collaroy Plateau to City Wynyard	EX	27.86
40	E83	PrePay-Only - North Narrabeen to City Wynyard	EX	18.33
42	E88	PrePay-Only - North Avalon Beach to City Wynya	EX	18.40
43	E89	PrePay-Only - Avalon Beach to City Wynyard (Ex	EX	21.33
44	L90	PrePay-Only - Palm Beach to City Wynyard (Limi	EX	4.00
45	430	Opal only - Sydenham to Taronga Zoo	AS	11.71

Table 3. Summary of the characteristic of each for which data was obtained during the study period.

6.5 Appendix E

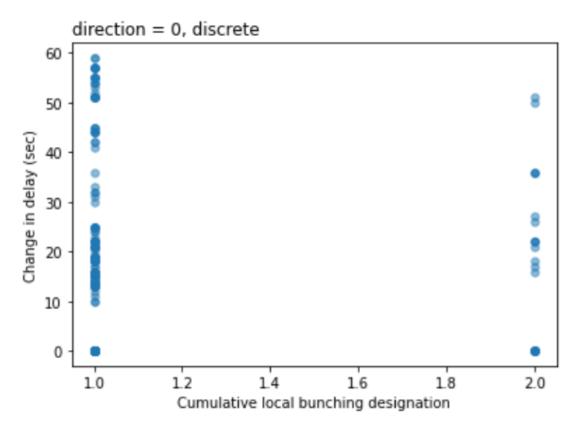


Figure 8. The local cumulative bunching designation is compared to the change in delay of buses that encounter bunched 430 buses. Metric 1 is used which has caused data to be ordered in vertical lines at integer values of the cumulative local bunching designation.