Delay distribution estimation at a signalized intersection

Susilawati¹, M I Ramli², and H Yatmar³

¹School of Engineering, Monash University, Malaysia
²Civil Engineering Department Universitas Hasanuddin, Makassar, Indonesia

Email: susilawati@monash.edu

Abstract. Previous studies found that delay at signalized intersection accounted for more than one-third of total travel time. Different level of services tends to have different delay probability. An early assumption was made that multimodality occurs because of two traffic states experienced by the traveler; delay and non-delay at the signalized intersection. This study proposed a method to quantify delay distribution at a signalized intersection by analyzing the patterns in the speed time and speed distance profiles when passing two consecutive intersections and redrawing vehicle trajectory generated by ‘second by second’ GPS data. The searching algorithm was developed to search for the times when the vehicle enters and leaves the delay section by checking the ‘second by second’ speed data. To differentiate between the queue, move up and stop and go traffic, the algorithm searches for the idle time (i.e. when the speed less than 3.5 km/h). Along with Sydney Coordinates, Adaptive Traffic Systems (SCATS) degree of saturation and signal settings (green and red time) generating from SCATS systems at upstream loop detectors, a realistic delay probability model also was developed. This model used the General Additive Model for Location, Shape, and Scale (GAMLSS) which allows location, scale and shape parameter of selected distribution as a function of the explanatory variable. Signal settings and SCATS degrees of saturation were used as an explanatory variable. This model enables us to estimate delay experiencing by traveler depending on traffic parameters including traffic flow, signal settings and degree of saturation that are readily available on the site.

1. Introduction

The dynamic and stochastic nature of traffic networks and irregular events such as short term incidents (e.g. vehicle breakdowns) and long incidents term (e.g. bridge collapse, road works) have greatly influenced the consistency and predictability of travel time. Almost 40% of the recurrent and non-recurrent congestion in the USA is due to bottlenecks, the incident is the second most important factor accounting for about 25% of congestion, while bad weather contributes about 15 % and poor signal timing and special events share the same percentage, only 5% [1].

Moreover, the schematic figure of the effect of each incident to the travel time variability and in the interrelationship between factor-related demand and capacity fluctuation [2]. The seasonal effect - day to day, time of day traffic fluctuation and the traffic mix are some factors that significantly affect the demand fluctuation, while incidents, road works, accident, weather, and road geometry are likely to affect the capacity fluctuation.

In the urban arterial context particularly in urban arterial road controlled by signalized travel time, both uniform and overflow delay and poor signal timing were a larger contributor for high variability travel time. Due to poor performance of downstream signalized intersections, travelers experience two types of traffic regimes i.e. one group of vehicles experience no delay whereas another group will
experience a delay during the red phase [3]. To differentiate two separate traffic conditions according to a mixture variable and the variability in the means of the component distributions, used the ‘maximising multimodal normal distribution’ and suggested that congested conditions could be seen as those with high variability in the mean values, whereas uncongested conditions had a relatively stable and consistent mean value [4].

A study about the different shapes of the delay distributions after implementing a coordinated signal system along a road corridor in North Carolina have been conducted. This study suggested that the distribution of the control delay has a single and large peak in the level of service A and B, whereas the control delay distribution tends to become bimodal as the mean control delay increases [5]. The benefits of installing signal coordination in the study area were also investigated and found that signal coordination could considerably reduce the mean and standard deviation of the control delay. It also changed the shape of the control delay distribution. At some locations, the control delay distributions clearly showed bimodality with large differences between the two peaks (e.g. 10 and 70 seconds/vehicle). On some parts of the corridor, the control delay distribution clearly showed bimodality with a large difference between two peaks (modal frequencies), 10 and 70 second/veh respectively. The delay distribution at an arterial urban signalized intersection in The Netherlands was box-shaped [6]. In a follow-up study in China, the researchers applied neural networks to actual travel time data from urban arterial roads with predicted signal settings including cycle time and green splits generated from SCATS maximum degree of saturation to estimate delays. As a result, it is suggested that a minimum number of observations of field data resulted in a discrepancy of distribution of estimated and fieldwork travel time. To gain insight into the ways that traffic control settings influence urban arterial road travel times, the Burr regression technique used to estimate the shape parameter of the Burr distribution (c) as a function of SCATS maximum degree of saturation [7]. The higher the SCATS degree of saturation, the higher the travel time variability and vice versa.

Delay can be modeled as a deterministic and stochastic model. The deterministic model tends to assume all factors involving delay, are deterministic including arrival flow and signal phases. However, arrival flow and signal phases can be highly stochastic and time-dependent and hence in such cases the stochastic model is more appropriate. A study examined the effect of different traffic parameters in the overall delay probability and found that green time and saturation flow are the most significant variable on the delay probability while the cycle time does not have a large impact on that probability. This study also found that under oversaturated conditions, the overflow queue affects the probability and the shape of the delay distribution as the queue becomes longer and the distribution becomes flatter and closer in shape to the normal distribution [8].

From the above review, it can be seen that factors, such as cycle time, green splits, the maximum degree of saturation, arrival and departure flow contributed to delay distribution, thus a model which enable to get insight into each factor determining the shape of delay distribution is of interest. This study investigated the delay distribution at Adelaide urban arterial road travel time data by developing a GIS-based search algorithm to differentiate acceleration and deceleration periods to generate delay data. Given delay data generated from the GIS-based search algorithm, this study also developed a realistic delay probability model based on various combinations of traffic parameters generated from SCATS systems. These parameters include the degree of saturation, green time and red time. A model that able to quantify the factors contribute to the shape of delay distribution was also introduced.

A General Additive Model for Location, shape, and scale (GAMLSS) which allows location, scale and shape parameter of selected distribution as a function of the explanatory variable was used [9]. Signal settings (green and red time) and SCATS degree of saturation were used as an explanatory variable. This model enabled us to estimate delays experienced by travelers depending on traffic parameters that are readily available on the site. The next section of this paper reviews delays and queue formation at signalized intersections. Methodology to develop a search algorithm to differentiate stationary and cruise conditions in the speed –time and speed-distance travel time trajectory is then described. The fourth section discusses delay distribution modeling and factors affecting multimodal delay distributions and the fifth section presents and discusses the research results.
1.1 Delay formation

There are two types of delay i.e. (i) stopped delay and (ii) total delay. The stopped delay is the time that the vehicle is in a queue and is stationary on the road section. The total delay is the time difference between the actual time taken to traverse the road segment and the time to traverse the same road segment at the desired cruising speed [10]. Delay formation involves the different stages of vehicle movements such as acceleration, deceleration, and stationary conditions. Acceleration and deceleration behavior can be clearly seen from the vehicle speed-distance and speed-time profiles. From the speed-distance and speed-time profiles of a vehicle, acceleration can be taken as occurring in those time periods when the vehicle speed is generally increasing. Deceleration occurs when the speed is generally decreasing. While the stationary condition is relatively easy to defines because it can be seen as the stop time in the trajectory diagram.

Figures 1 represent the speed-distance profile for four conditions experienced by the test vehicle on one link. The first condition is when the first vehicle did not experience any delay along the road section and traveled freely along with the link (1a), even if its speed was generally below the free-flow speed. This could represent the undersaturated condition when no overflow queuing occurs at the intersection. The second condition in figure 1b is the event when the vehicle experienced a short delay near the stop line while the third condition is when the test vehicle stopped in the mid-section for quite a long time as shown in figure 1c.

![Figure 1. Speed, distance and time profile for cases 1, 2, 3 and 4.](image)

The fourth condition is when the vehicle experienced very long delays, as illustrated by the occurrence of two or more stops along with the link as shown in figure 1d. Condition two and three might be the cases in which the vehicle approaches the stop line during the red phase i.e. the vehicle will stop and join the existing queue until the signal turns to green. There could be another situation where vehicle 2 might join the queue during the red phase when the traffic condition is undersaturated and thereby it is closer to the stop line when compared to vehicle 3. As a consequence vehicle 3 experiences longer delay than vehicle 2. The delay for vehicle 3 increases if the green time for that cycle is not
sufficient to clear all vehicles in the queue, hence, vehicle 3 would then experience overflow queuing; this condition would occur in oversaturated traffic. For the case of vehicle 4, the speed-distance profile clearly shows that this vehicle makes two significant stops in the mid-section before joining the queue at the stop line. Given these examples, the delay experienced by the traveler is greatly influenced by the flow-capacity ratio (degree of saturation) and signal settings.

1.2. Delay distribution
There are methods available to differentiate acceleration and deceleration behavior. Colyar and Rouphail analyzed the patterns in the speed time and speed distance using second by second GPS travel time data. A search algorithm was developed to find a delay event and to differentiate between the queue move up and ‘stop and go traffic’. It was time-consuming to run this algorithm, thus some of the previous studies tend to exclude acceleration and deceleration behavior in their delay modeling.

2. Methodology

2.1. GIS-based search algorithm
To obtain insight into every single speed-distance and speed-time profile is a time consuming and complex task. The use of a GIS application eases the process of spatial and temporal data processing, and also allows the analyst to gain insight into all the critical movements in the vehicle profile data. The searching algorithm was developed using tools available in GIS. To conduct that assessment some assumptions have been made.

The (stopped) delay is the time that a vehicle experiences when its speed is less than 3.5 km/h and time while the vehicle is queuing, so delay might not only happen at the approach or at the intersection but also includes any stop time of the vehicle on that specific road section of the link. Speed, distance and time profiles are assessed to obtain the basic information about when and where the studied vehicle stopped on the road section.

To obtain the delay data, a method similar to that of was used. This method not only allowed us to gain insight into the speed data and its variation but also the location where the vehicle first stopped on the link. Model Builder application within ArcGIS was developed to string together sequences of geoprocessing tools and to simplify the long and loop process to ease the process of delay measurement at the signalized intersection. Figure 2 shows a flowchart diagram for delay data extraction.
The first step was to store GPS data into Microsoft Excel and convert it to the ArcGIS layer as point features. The intersect tool was used to segregate GPS data into a different link according to the spatial location of each intersection (coordinate). The select tool was used to select any timestamps with a speed less than 3.5 km. To obtain a delay period in each link, the frequency tool was used to accumulate each point with speed less than 3.5 km in each link and tabulate it into a delay table. The searching algorithm tool was tested and used to obtain delay probability on Adelaide arterial road travel time data.

2.2. Travel time data

Longitudinal travel time data collected by using GPS during the regular journey to work trips started at the same time of day with similar driving behavior at two consecutive intersections in Adelaide arterial roads were used. The set of travel time data are: Link 1 is the link between the Glen Osmond Road (GOR)/Young Street intersection (TCS 0220) and Glen Osmond Road (GOR)/Greenhill Road intersection (TCS 0069). It is 606 m long with a speed limit of 60 km/h. The test vehicle traverses this link in the northbound direction. Figure 3a is the plan for the downstream intersection (TCS 0069). There are three detectors for northbound traffic, detector numbers 14, 15 and 16. Detector 15 is for through traffic while detector number 14 is for left turn traffic only. Detector 16 is for through and right turn traffic during the off-peak period. The analysis focused on detectors 15 and 16.

Similar to Link 1, Link 2 is a link connected between the Glen Osmond Road/Greenhill Road intersection (TCS 0069) and Glen Osmond Road/Hutt Road intersection (TCS 176) (see figure 3b) with 331 m long with a 60 km/h speed limit. Four detectors are for northbound traffic, detector numbers 5, 7 and 8. Detector 5, 6 and 7 are for through traffic while detector number 8 is for right turn traffic only. Detector 5 is for through and left-turn traffic while detector 7 is also for right turn traffic respectively. The analysis focused on detectors 5 and 6.

To get insight into SCATS control parameter and SCATS signal phases, an intersection plan is required. For example, signal phases for intersection Link 1 (TCS 0069) have four phases applying for each cycle: the major phases are phases A and C, phase A for west-east traffic while phasing C for north-south traffic. Phases B and D are for right and left turn and through traffic respectively. Based on the signal phases, the northbound traffic is controlled by phase C.

Figure 3. Glen Osmond Road link 1 and 2 intersection plan.

Figure 4 presents the cycle time, green time, red time and the ratio of green time to cycle time for phase C. It can be seen that this phase is a major phase with the ratio exceeding 50% of the cycle time over much of the observation period, to allow the traffic stream to clear during the green time.
According to the SCATS operational manual, the system operation is based on the number of vehicles passing through that intersection over three consecutive signal cycles and sets the next cycle time according to that number. If the number is constant, SCATS will adjust the signal setting. The degree of saturation (DS) obtained from SCATS data is not the actual degree of saturation, however for purposes of simplicity, as a first-order approximation, the SCATS DS data can be treated as equivalent to the theoretical degree of saturation [11, 12].

Figure 5 represents the line graph for maximum and minimum flow and maximum and minimum DS for all detectors in the northbound traffic. It indicates that the patterns of maximum and minimum DS are similar to the pattern of maximum and minimum flow, by which high flow tends to have high DS and vice versa. Even though there are some inconsistencies in this graph, in general, these two data sets deliver a similar pattern.

2.3. Model development
Given a method developed from the second by second GPS data using a GIS-based searching algorithm, the duration of delay was easily identified and delay histogram was drawn respectively. The observed histogram of delay on link 1 and 2 are given in figure 6, which clearly had two or more peaks and long tail. This showed that for some cases the test vehicle experienced more than two queues along with these links.
Since multimodality found in delay histograms, an examination of its occurrence has been conducted in order to get insight into factors to contribute to this phenomenon. The sensitivity analysis GAMLSS method is introduced. GAMLSS is (semi) parametric univariate regression models, where all the parameters of the assumed distribution for the response can be modeled as additive functions of the explanatory variables. GAMLSS model assumes independent observations \( y_i \) for \( i=1,2,3,4,\ldots \) with probability (density) function pdf as follow:

\[
f(y_i | \mu_i, \sigma_i, \nu_i, \tau_i)
\]

As a conditional on a vector of four distribution parameters, each of which can be a function to the explanatory variables. Unlike the bimodality test using DIP Hartigan test which is only able to test the goodness of fit of two mixture normal distribution, GAMLSS is able to test a numerous distribution with many mixtures components (more than two) fitted to delay distribution data [13]. Moreover, unlike Burr regression technique which only allows the scale and location of Burr parameter to be formulated as explanatory variables, this model can also be used to formulate parameters of numerous distribution as explanatory variables. Each GAMLSS model generates Akaike Information Criterion (AIC), a value represents the goodness of fit of each model. To get the best model, AIC should be ranked from the lowest to the highest value. The model in which gave the lowest AIC value will be considered as the best model. To get sufficient proof about the reliability of the developed model, delta AIC and Akaike weights, supply information on the strength of evidence for each model to replace \( \alpha \) (significant level) used in most used goodness of fit test [14].

\[\text{Figure 6. Histogram GOR Link 1 and Link 2.}\]

3. Result and discussion
The degree of saturation and signal settings including green and red time (which was extracted from the SCATS system data files) were used in the model as an explanatory variable of the mean parameter’s Normal distribution \( (\mu) \). The mean parameter which represents the location parameter of Normal distribution has been selected because delay distribution clearly had either two or more peaks. Two and three mixture Normal were fitted to delay data. The results were tabulated in table 1.
Table 1. AIC Values for Mixture normal.

| Explanatory Variables | AIC - 2 Mixture Normal | AIC - 3 Mixture Normal |
|-----------------------|------------------------|------------------------|
|                       | Link 1     | Link 2     | Link 1     | Link 2     |
| None                  | 1020       | 794        | 1012       | 742        |
| Red time              | 1023       | 792        | 1015       | 741        |
| Green time            | 1022       | 804        | 1009       | 759        |
| Degree of Saturation  | 1019       | 795        | 1013       | 797        |

Table 1 shows that two mixture Normal distribution with red time, green time and degree of saturation gave slightly higher AIC value than three mixture Normal distribution for both links 1 and 2. The same results were also found when green time, red time and dos were used as an explanatory variable which means that two mixture Normal distribution with green time, red time and DoS as explanatory variable could not perfectly fit delay distribution, more importantly, those histograms only had one peak.

For comparison, figure 7 shows three mixture normal distribution for link 1 with green time, red time and degree of saturation as explanatory variables. These histograms clearly show that the pdf function of estimated data with three mixture Normal distribution fitted delay distribution. In contrast, higher green time creates two tall peaks in the histogram. Since link 2 has three or more peaks, using red time as an explanatory variable gave a more realistic shape of the delay distribution. It also is shown that lower red time tends to shift the shape of the distribution, to have fewer peaks (similar result found in Link 1). Moreover, using green time as explanatory variable gave inconsistent result since the shape of the distribution for various green time value were uneven (both Link 1 and 2) as shown in figure 8. The degree of saturation gave a more realistic shape for all various values of the degree of saturation with less variation. The results are different from a study conducted in the Netherland as the delay distribution tends to have box-shaped and shape of delay distribution were influenced by arrival flow and overflow delay.
Figure 7. GOR Link 1 histogram with 3 mixture normal distribution.
4. Conclusions
This study introduced a new method to generate a delay event from ‘second by second’ GPS data using a GIS-based searching algorithm developed using the ArcGIS model builder application. This new method will enable traffic practitioners to identify the delay event and automatically generate delay data. Given delay data at two signalized intersections in Adelaide arterial roads along with SCATS control parameters, a refined delay distribution model was developed. This model was able to quantify the influences of higher and lower green time, red time and SCATS degree of saturation on the shape of the delay distribution. It was found that the degree of saturation played a significant role in shaping delay distribution showed by lower AIC value and estimated delay density. This result also proved that lower red time tends to shift the shape of the distribution, to have fewer peaks while higher red time resulting in less reliable intersection performance. Having more peaks in delay and travel time distribution significantly reduces travel time reliability and increase travel time variability. This refined model can be extended to get insight into factors to influence multimodality in travel time distribution.

Figure 8. GOR Link 2 histogram with 3 mixture normal distribution.
References

[1] Cambridge Systematics and Texas Transportation Institute 2004 *Traffic Congestion and Reliability: Linking Solutions to Problems* (Federal Highway Administration)

[2] Lint J W C, H J V Zuylen and H Tú 2008 Travel time unreliability on freeways: Why measures based on variance tell only half the story *Transportation Research* 42 258-277

[3] Susilawati, M A P Taylor and S Somenahalli 2010 Travel time reliability and bimodal travel time distribution *Journal of Road and Transport Research* 19(4) 30-43

[4] Guo F, H Rakha and S Park 2010 Multistate model for travel time reliability *Journal of the Transportation Research Board* 2188 46-54

[5] Colyar J D and N M Roupail 2003 Measured distributions of control delay on signalized arterials. *Transportation Research Record* 1852 1-9.

[6] Zheng F, H J V Zuylen and Y Pu 2012 Model for predicting the distribution of link travel times for urban signalized roads *Transportation Research Record: Journal of the Transportation Research Board* 2311 59-72.

[7] Susilawati, M A P Taylor and S Somenahalli 2014 *Urban Arterial Road Travel Time Variability Modeling Using Burr Regression* in 93rd Transportation Research Board Annual Meeting (Washington, DC, USA)

[8] Olszewski P S 1994 Modelling the probability distribution of delay at the signalized intersections *Journal of Advanced Transportation* 28(3) 253-274

[9] Stasinopoulos D M and R A Rigby 2007 Generalized additive models for location scale and shape (GAMLSS) *R. Journal of Statistical Software* 23(7) 1-46

[10] Taylor M A P, P W Bonsal and W Young 2000 *Understanding traffic systems: Data, analysis and presentation, ed. S. Edition* (Ashgate Publishing ltd)

[11] Akcelik R, M Besley and R Roper 1999 *Fundamental Relationship for Traffic Flows at Signalised Intersections, in Fundamental Relationship for Traffic Flows at Signalised Intersections* (New South Wales: Roads and Traffic Authority)

[12] Suksri C, M A P Taylor and W L Yue 2011 *Exploring Stop-line Traffic Flows at Signalised Intersections in the Adelaide CBD* (Australia: Australasian Transport Research Forum)

[13] Susilawati, M A P Taylor and S Somenahalli 2010 Travel time reliability and the bimodal travel time distribution for an arterial road *Road and Transport Research* 19(4)

[14] Hardelid P, et al 2008 Analysis of rubella antibody distribution from newborn dried blood spots using finite mixture models *Epidemiology and Infection* 136(2) 1698-1706