Performance Analysis of Delay Estimation Models for Signalized Intersection Networks

Hyung Jin Kim[1], Bongsoo Son[2], Soobeom Lee[3]

[1]Dept. of Urban Planning and Eng. Yonsei Univ,, Seoul, Korea
{hyungkim, sbs}@yonsei.ac.kr
[2]Dept. of Urban Planning and Eng. Yonsei Univ., Seoul, Korea
[3]Dept. of Transportation Engineering, Univ. of Seoul, Seoul, Korea
mendota@uos.ac.kr

Abstract. The primary purpose of this study is to examine the models' performance for estimating average delay experienced by the passing vehicles at signalized intersection network, and to improve the models' performance for Intelligent Transportation Systems (ITS) application in terms of actuated signal operation. Two major problems affected the models' performance have been defined by the empirical analyses in this paper. The first problem is related to the time period of delay estimation. The second problem is associated with the fact that the observed arrival flow patterns are so different from those applied for developing the existing models. This paper presents several methods to overcome the problems for estimating the delay by using the existing models.

1. Introduction

Many models have been developed for the purpose of delay estimation at signalized intersection network. It is known that the results of the existing models are very sensitive to the degree of saturation as well as the arrival flow pattern at the intersections during the time period of interest. This implies that the models' reliability seems to be highly dependent on whether the input variables of the model are adequate to describe the real traffic conditions [1, 2, 3, 4]. One main purpose of this study is to evaluate five major models for the feasibility of delay estimation for urban signalized intersection network. The five models are Webster, US Highway Capacity Manual (HCM), Transyt-7F, Akcelik, and Hurdle models. Another main purpose of this study is to improve the models' performances for the purpose of ITS application such as actuated signal operation. To accomplish the study purposes, the input variables of the five models were acquired from the traffic data collected from the field. The models' results were compared with the results obtained from the conventional queuing theory, cumulative arrival and departure technique, by using the field data. Two study sites in Seoul were selected, where traffic states of the two sites were different, one was saturated and another was non-saturated.

2. Related Work

The operation of each intersection approach can be modeled as shown in Figure 1. In the figure, the y-axis is the cumulative vehicle count (*N*), and the x-axis is time (*t*). The curve labeled A(*t*) shows the cumulative number of arrivals by time *t*, and D(*t*) shows the cumulative number of departures. In fact, the A(*t*) curve does not indicate the number of actual arrivals at the stop line, but the number that would have arrived if the signal light had always remained green. The D(*t*) curve shows the actual departures from the stop line. When the signal light is red, there are no departures, so the D(*t*) curve is horizontal. The overall D(*t*) is the stair-step curve outlining the triangles. In reality, it would begin to curve upward as vehicles began to move after the start of green then after a few seconds become nearly straight with a slope equal to the saturation flow [5].

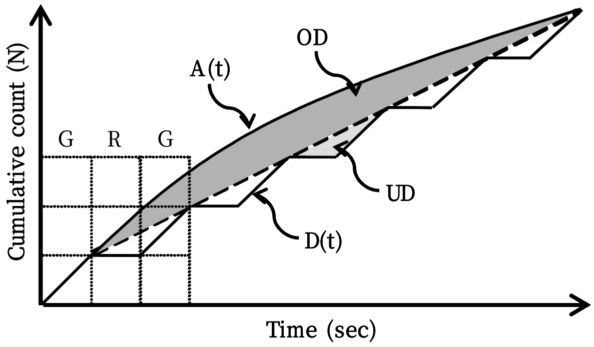

Figure 1. Typical cumulative arrival and departure curves

In Figure 1, the below area of the dashed line is associated with the traffic situation that all the arrivals within one signal phase can pass through the intersection during the same signal phase. This is called as non-overflow situation. In this case, A(t) curve will be the dashed line. The slope of A(t) curve is the arrival rate, and if this rate is constant over several signal cycles, the between the A(t) and D(t) curves is made up of a series of identical triangles. The total delay per cycle can be estimated as the area of any single triangle. Diving this area by the number of arrivals per cycle yields the average delay, that is denoted UD, which stands for average uniform delay, since it was derived under the assumption that vehicles arrive at a uniform rate throughout the signal cycle. It should be noted that, in making this assumption, we ignore both any random effects and any pattern imposed on the arrival stream by upstream intersections. In the same figure, the above area of the dashed line is related to the traffic situation that some of the arrivals within one signal phase cannot get through the intersection during the same signal phase. This is called as over-flow situation or over-saturation. The total overflow delay can be estimated as the area between the A(t) curve and the dashed line. Diving this area by the total number of arrivals during the time period which the arrival flows exceed the capacity yields the average overflow delay, that is denoted OD. The average delay of each signalized intersection approach is expressed by the sum of UD, OD, and a correction term which has a negative value typically. The correction term is generally obtained by simulation, but its value is relatively too small, so it is ignored for the practical purposes.

Figure 2 presents a good understanding of the relationship between the five models' performances and the degree of saturation, *v*/*c*, where *v* = arrival flow and *c* = capacity of intersection approach. Although this figure is an example, it provides very useful insights for the features of five models' performances. As the degree of saturation is close to 1.0, the discrepancies of the models' results are drastically increased. The discrepancies are serious in the range of *v*/*c* from 0.9 to 1.10. From the figure we can see that the results of the existing models are very sensitive to the degree of saturation as well as the arrival flow pattern at the intersection during the time period of interest. In the degree of saturation, *c* is a manageable variable, but the variable *v* is not, and thus *v*/*c* cannot be adjusted for the purpose of reducing the discrepancies between the five models' results.

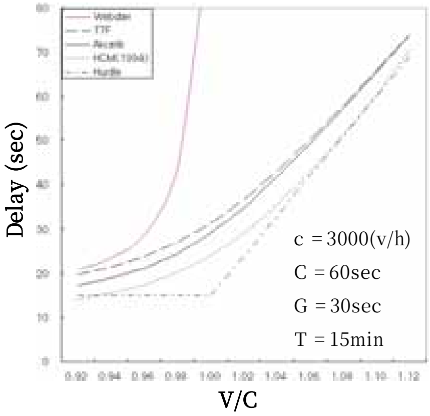

Figure 2. Comparison of the five models' results for the delay

It should be noted that US HCM recently presents new model which has improved the effects of the arrival flow variations by selecting an appropriate type of arrival flow pattern among several predetermined patterns for the analysis [6]. However, this is not the only way to improve the HCM model's performance, so this study has selected the model developed in 1994 for solving another problem involved in the model.

3. Evaluation of Existing Models' Performance

In order to evaluate the five models' performance, two study sites were selected. Table 1 summarizes traffic and signal conditions of the two study sites. Figure 3 shows the signal phases of the analysis intersection and the upstream intersections of the study site #1. The cycle length of the two intersections is 140 sec. and the roadway is 4-lane foe each direction. The travel time between the two intersections was 5-minute during the data collection period. Traffic volume and speed were collected at the 15-minute time interval during morning peak period between 7a.m. and 9a.m. Using the traffic data, the A(*t*) and D(*t*) curves were constructed as shown in Figure 4. In order to match the time of two curves, the travel time between the target and upstream intersections was estimated from the observed speed data.

Table 1. Traffic and signal conditions of two study sites

study sites		target Intersection	upstream intersection	distance
No.1	traffic state	Saturated	Saturated	600m
	number of signal phases	4		
	cycle length	Different		
No.2	traffic state	non-saturated	Saturated	500m
	number of signal phases	4		
	cycle length	Different		

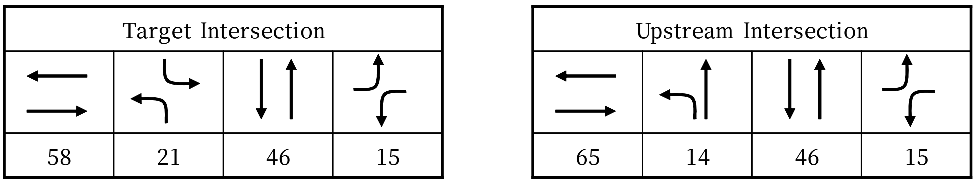

Figure 3. The signal phases of study site #1

In Figure 4, specific traffic counts of y-axis and times of x-axis were not presented, since these values are not important at this stage and things to be discussed in this paper are related to shape of the two curves. The A(*t*) and D(*t*) curves are very similar. This result is quite different from that of Figure 1. The reader may be so confused to figure out which one is the reality, but Figure 4 is the case. At the study site #1, the target and upstream intersections were saturated. Under the traffic situation, all the vehicles passing the upstream intersection traveled at the same speed of the vehicles passing the target intersection, and the arrivals of the target intersection could not exceed the departures of the intersection. The reader should remember that the A(*t*) curve does not indicate the number of actual arrivals at the stop line, but the number that would have arrived if the signal light had always remained green. However, in the congested traffic condition, the A(*t*) curve will not be much changed from the D(*t*) curve of upstream intersection if the signal light had always remained green. As described in Section 1, the field delay can be obtained from Figure 4. From now, we have to review the problem caused by the length of evaluation time period in using the existing models. Figure 5 shows two settings of evaluation time periods, T1 and T2. In practice, it is reasonable that the evaluation time period does match with the congested time period of the intersection interested, but the evaluation time period has been typically defined as 15 minutes or 1 hour. In fact, the length of the evaluation time is not a big problem. The problem is the setting of the time period. T1 and T2 are the same length of evaluation time period, but the starting and ending times of the two time periods are different. The main difference between the two periods is that T1 starts at the beginning time of the first phase of signal and T2 terminates at the ending time of the final phase of signal. Depending on how to set the evaluation time period, the existing models' result for the delay will be changed significantly.

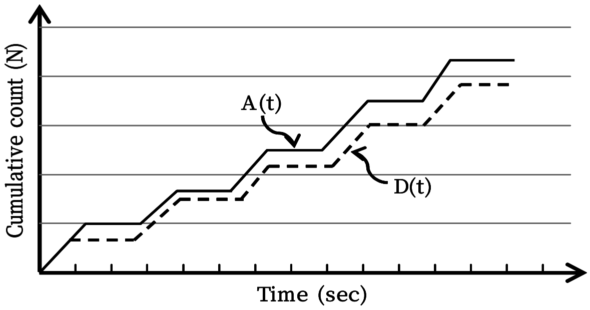

Figure 4. A(t) and D(t) curves of study site #1

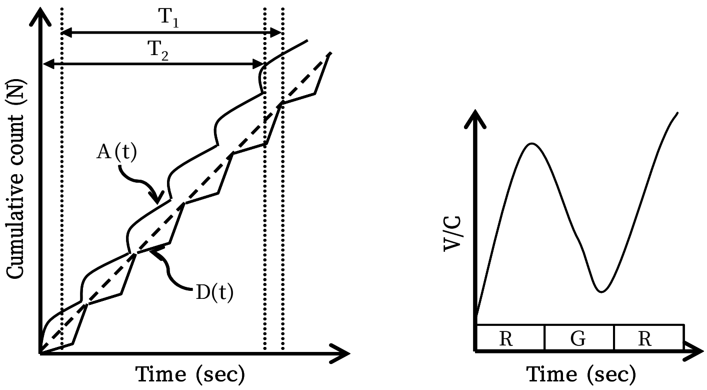

Figure 5. The relationship between delay evaluation period, signal phase, and v/c

As mentioned before, the signal cycle lengths of the intersections at the study site # 1 are 140seconds. If the evaluation time period, 15 minutes, is set as T1, then T1 will be from 0 to 900 seconds and will be terminated before the final phase of signal finished. For setting the evaluation time period as T2, we have to figure out the time not only matched with the ending time of the signal phase but also closed to the 15-minute evaluation period. This time is 980 seconds, so T2 is from 80 to 980 (i.e., 140 seconds x 7 cycles). Table 2 summarizes the models' results and the field delay obtained from the cumulative arrival and departure technique by using T1 and T2. In the table, the field delays obtained from the cumulative arrival and departure technique by using T1 and T2 are very similar. However, the models' results for overflow delay obtained by using the two periods are quite different, while the results for uniform delay of the two periods are identical with the exception of HCM model. Although the time lengths of the two evaluation periods are equal, the overflow delays obtained by using T1 are much greater than those of T2. The evaluation period T1 terminates before the signal cycle finished so that *v/c* is definitely overestimated. From the results of Table 2, it is confirmed that the models' results are very sensitive to *v/c* and it is mainly dependent upon the ending time of evaluation time period. More specifically, the degree of saturation, *v/c*, is mainly determined by the fact whether the ending time of evaluation period agrees with the ending time of signal cycle. The overall models' results for the delay are much greater than those of the field observations. In order to overcome the problems of both T1 and T2, a new evaluation time period that includes both T1 and T2 has been proposed in this study. The new evaluation period starts at the beginning time of T1 and terminates at the ending time of T2, so the new evaluation time period is longer than both T1 and T2.

Table 2. The comparison of the models' results with the field observation of the study site #1

Evaluation period	Models	UD	OD	UD + OD	Field Delay
Evaluation period is T1 from 0 to 900 sec.	Webster	41.00	-	-	v/c = 1.31 109.09 (sec)
	T-7F	41.00	142.61	183.61	
	Akcelik	41.00	142.65	183.65	
	HCM	39.97	189.59	229.56	
	Hurdle	41.00	140.17	181.17	
	Models	UD	OD	UD + OD	Field Delay
Evaluation period is T2 from 80 to 980 sec.	Webster	41.00	-	-	v/c = 1.20 107.07 (sec)
	T-7F	41.00	93.47	134.47	
	Akcelik	41.00	92.91	133.91	
	HCM	36.30	104.23	140.53	
	Hurdle	41.00	90.06	131.06	

Table 3. The comparison of the models' results with the field observation of the study site #1

Evaluation period	Models	UD	OD	UD + OD	Field Delay
Evaluation period from 0 to 980 sec.	Webster	49.25	-	-	v/c = 1.24 107.55 (sec)
	T-7F	49.25	109.56	158.81	
	Akcelik	49.25	109.26	158.51	
	HCM	37.44	129.27	167.11	
	Hurdle	41.00	106.57	147.57	

Comparing the results in Tables 2 and 3, it is very clear that T1 does overestimate the overflow delay, since the overflow delays obtained by using T1 are still much greater than those of the new period even though the new period is longer than T1. In general, the models' results are very fluctuated by the change of evaluation time period, while the field delays are consistently changed.

The models' overflow delays are persistently greater than the field observations. The reason for this can be found in Figure 6. The A(*t*) obtained from the field observation is stair step curve, while the A(*t*) of the existing models forms a smooth curve. It is interesting that the discrepancy of the overflow delay between the two curves is almost equal to the uniform delay, UD. Thus, if the uniform delay is subtracted from the total delay of the models, the models' results will be matched with the field values reasonably well.

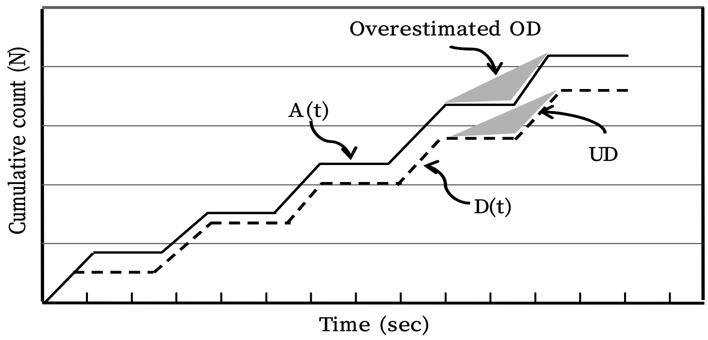

Figure 6. The discrepancy of A(t)

Figure 7 shows the signal phases of study site #2. The study site #2 is not saturated intersections. The travel time between the analysis intersection and the upstream intersection was 5.3-minute during the data collection period. Traffic volume and speed were collected at the 15-minute time interval during morning peak period between 7a.m. and 9a.m. Using the traffic data, the A(t) and D(t) curves were constructed as shown in Figure 8.

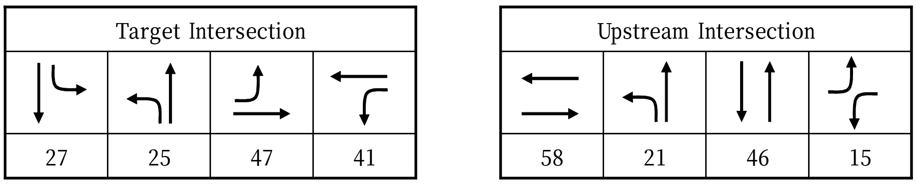

Figure 7. The signal phases of study site #2.

In Figure 8, the shaded area marked by a solid line is the observed delay and the area represented by a dashed line is the estimated delay of the models. The two areas form the diamond shape that is quite different from the triangle as shown in Figure 1. Anyway, the shaded area is larger than the estimated area of model. The difference between two areas is gradually reduced over several signal cycles. Then, as the signal cycle runs over and over again, the two areas will be converged to the same size. The signal cycle lengths of the intersections at the study site # 2 are 140seconds. The delay of study site #2 has been estimated by the same procedures applied for the study site #1. Tables 4 and 5 summarize the models' results and the field delay obtained by using three different evaluation time periods. As shown in Table 4, the evaluation period T1 terminates before the signal cycle finished, so v/c is exceeded to 1.0 although the site is not saturated.

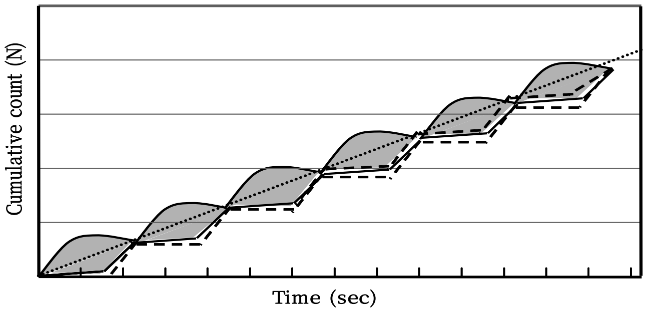

Figure 8. A(t) and D(t) curves of study site #2

Correspondingly, the models have produced the overflow delay. Using the period T2, the models still produce the overflow delay even though *v/c* is not exceeded 1.0, but the delay is very small. However, the models results obtained by T2 are much less than the field observations.

Table 4. The comparison of the models' results with the field observation of the study site #2

	Models	UD	OD	UD + OD	Field Delay
Evaluation period is T1 from 0 to 900 sec.	Webster	49.50	-	49.50	v/c = 1.07 58.99(sec)
	T-7F	49.50	38.10	87.60	
	Akcelik	49.50	34.95	84.45	
	HCM	38.77	33.94	72.71	
	Hurdle	49.50	32.07	81.57	
	Models	UD	OD	UD + OD	Field Delay
Evaluation period is T2 from 80 to 980 sec.	Webster	49.50	-	49.50	v/c = 0.95 62.57(sec)
	T-7F	49.50	6.78	56.28	
	Akcelik	49.50	0.90	50.40	
	HCM	36..83	4.49	41.32	
	Hurdle	49.50	-	49.50	

Table 5. The comparison of the models' results with the field observation of the study site #2

	Models	UD	OD	UD + OD	Field Delay
Evaluation period from 0 to 980 sec.	Webster	49.71	-	49.71	v/c = 1.01 61.67(sec)
	T-7F	49.71	18.03	67.74	
	Akcelik	49.71	11.01	60.72	
	HCM	37.78	13.52	51.30	
	Hurdle	49.50	4.93	54.43	

In Table 5, all the models with the exception of Transty-7F tend to underestimate the uniform delay of non-saturated intersection by using the evaluation time period proposed in this study. However, Transty-7F and Akcelik models produce the reasonable results that are very close to the field observation, so the two models seem to be good to estimate the delay of the non-saturated intersections.

4. Conclusions

The primary purpose of this paper is to examine the models' performance for estimating average delay experienced by the passing vehicles at urban signalized intersection network, and to present the method for improving the models' performance. Two study sites in Seoul were selected, where traffic states of the two sites were different, one was saturated and another was non-saturated. From the empirical analyses, it was reconfirmed that the results of the existing models are very sensitive to the degree of saturation as well as the arrival flow pattern at the intersections during the time period of interest. Depending on how to set the evaluation time period, the existing models' results for the delay have been changed significantly. The field delays obtained from the cumulative arrival and departure technique by using T1 and T2 are very similar. However, the models' results for overflow delay obtained by using the two periods are quite different, while the results for uniform delay of the two periods are identical with the exception of HCM model. Although the time lengths of the two evaluation periods are equal, the overflow delays obtained by using T1 are much greater than those of T2. In order to improve the problem associated with the setting of evaluation time period, a new period that includes both T1 and T2 has been proposed in this study. The models performances have been somewhat improved by using the new period.

References

1. Transportation Research Board., Traffic Flow Theory, Traffic Flow at Signalized Intersection, Chapter 9, 1998
2. W.R. McShane and R. P. Roess, Traffic Engineering, second edition, Prentice Hall, 1998
3. S. Teply, Accuracy of delay Surveys at Signalized Intersections, Transportation Research Record 1225, 1989
4. S. Teply et al., Canadian Capacity Guide for Signalized Intersections, Institute of Transportation Engineers District7-Canada, 1996
5. V. F hurdle, Signalized Intersection Delay Model, TRR 841, 1984
6. Transportation Research Board, Highway Capacity Manual, Chapter16, 2000