Signalized Intersections

Introduction

Signalized intersections are the important points or nodes within a system of highways and streets. To describe some measure of effectiveness to evaluate a signalized intersection or to describe the quality of operations is a difficult task. There are a number of measures that have been used in capacity analysis and simulation, all of which quantify some aspect of experience of a driver traversing a signalized intersection. The most common measures are average delay per vehicle, average queue length, and number of stops. Delay is a measure that most directly relates driver's experience and it is measure of excess time consumed in traversing the intersection. Length of queue at any time is a useful measure, and is critical in determining when a given intersection will begin to impede the discharge from an adjacent upstream intersection. Number of stops made is an important input parameter, especially in air quality models. Among these three, delay is the most frequently used measure of effectiveness for signalized intersections for it is directly perceived by a driver. The estimation of delay is complex due to random arrival of vehicles, lost time due to stopping of vehicles, over saturated flow conditions etc. This chapter looks are some important models to estimate delays.

Types of delay

The most common measure of operational quality is delay, although queue length is often used as a secondary measure. While it is possible to measure delay in the field, it is a difficult process, and different observers may make judgments that could yield different results. For many purposes, it is, therefore, convenient to have a predictive model for the estimate of delay. Delay, however, can be quantified in many different ways. The most frequently used forms of delay are defined as follows:

- Stopped time delay
- Approach delay
- Travel time delay
- Time-in-queue delay
- Control delay

These delay measures can be quite different, depending on conditions at the signalized intersection. Fig. below shows the differences among stopped time, approach and travel time delay for single vehicle traversing a signalized intersection. The desired path of the vehicle is shown, as well as the actual progress of the vehicle, which includes a stop at a red signal.
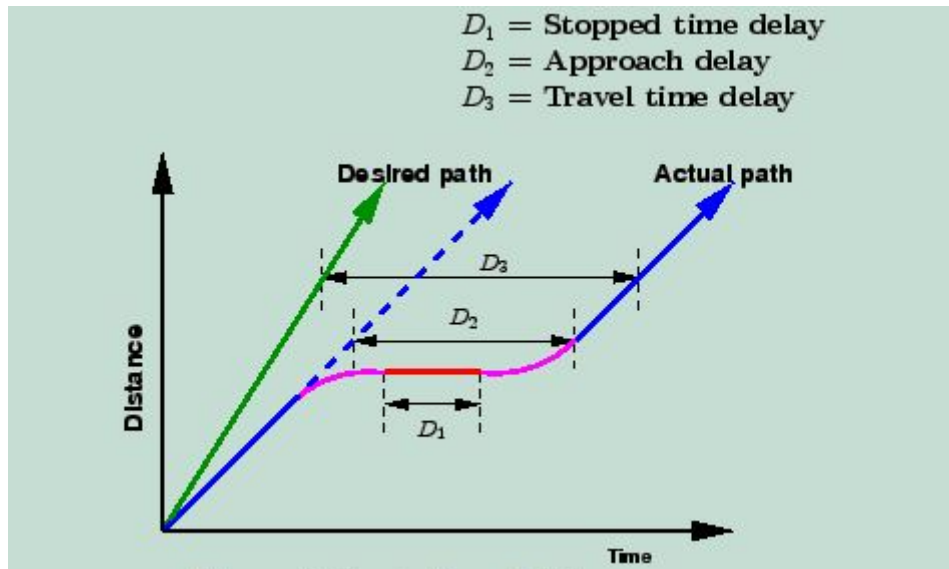
Figure : Illustration of delay measures.

Note that the desired path is the path when vehicles travel with their preferred speed and the actual path is the path accounting for decreased speed, stops and acceleration and deceleration.

Stopped Time Delay

Stopped-time delay is defined as the time a vehicle is stopped in queue while waiting to pass through the intersection. It begins when the vehicle is fully stopped and ends when the vehicle begins to accelerate. Average stopped-time delay is the average for all vehicles during a specified time period.

Approach Delay

Approach delay includes stopped-time delay but adds the time loss due to deceleration from the approach speed to a stop and the time loss due to re-acceleration back to the desired speed. It is found by extending the velocity slope of the approaching vehicle as if no signal existed. Approach delay is the horizontal (time) difference between the hypothetical extension of the approaching velocity slope and the departure slope after full acceleration is achieved. Average approach delay is the average for all vehicles during a specified time period.

Travel Time Delay

It is the difference between the driver's expected travel time through the intersection (or any roadway segment) and the actual time taken. To find the desired travel time to traverse an intersection is very difficult. So this delay concept is is rarely used except in some planning studies.

Time-in-queue Delay

Time-in-queue delay is the total time from a vehicle joining an intersection queue to its discharge across the STOP line on departure. Average time-in-queue delay is the average for all vehicles during a specified time period. Time-in-queue delay cannot be effectively shown using one vehicle, as it involves joining and departing a queue of several vehicles.

Control Delay

Control delay is the delay caused by a control device, either a traffic signal or a STOP-sign. It is approximately equal to time-in-queue delay plus the acceleration-deceleration delay component. Delay measures can be stated for a single vehicle, as an average for all vehicles over a specified time period, or as an aggregate total value for all vehicles over a specified time period. Aggregate delay is measured in total vehicle-seconds, vehicle-minutes, or vehicle-hours for all vehicles in the specified time interval. Average individual delay is generally stated in terms of seconds per vehicle for a specified time interval.

Components of delay

In analytic models for predicting delay, there are three distinct components of delay, namely, uniform delay, random delay, and overflow delay. Before, explaining these, first a delay representation diagram is useful for illustrating these components.

Delay diagram

All analytic models of delay begin with a plot of cumulative vehicles arriving and departing vs. time at a given signal location. Fig. below shows a plot of total vehicle vs time. Two curves are shown: a plot of arriving vehicles and a plot of departing vehicles. The time axis is divided into periods of effective green and effective red. Vehicles are assumed to arrive at a uniform rate of flow (v vehicles per unit time). This is shown by the constant slope of the arrival curve. Uniform arrivals assume that the inter-vehicle arrival time between vehicles is a constant. Assuming no preexisting queue, arriving vehicles depart instantaneously when the signal is green (ie., the departure curve is same as the arrival curve). When the red phase begins, vehicles begin to queue, as none are being discharged. Thus, the departure curve is parallel to the x-axis during the red interval. When the next effective green begins, vehicles queued during the red interval depart from the intersection, at a rate called saturation flow rate (s vehicle per unit time). For stable operations, depicted here, the departure curve catches up with the arrival curve before the next red interval begins (i.e., there is no residual queue left at the end of the effective green).
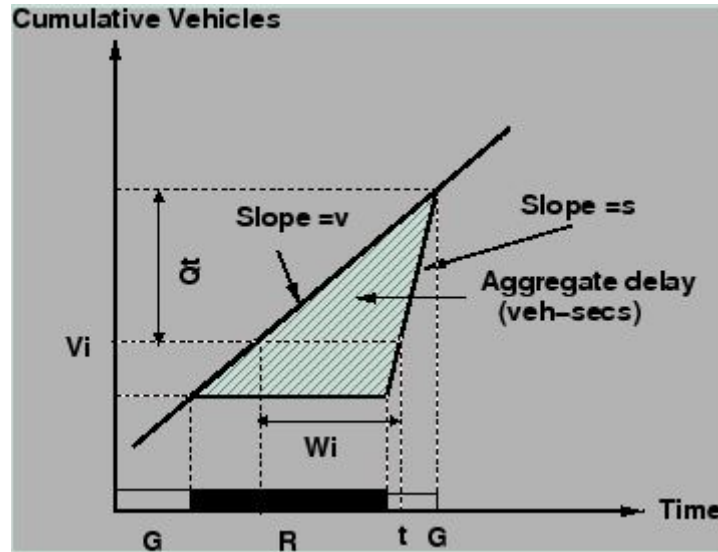
Figure: Illustration of delay, waiting time, and queue length.

In this simple model:

1. The total time that any vehicle (v_i) spends waiting in the queue (W_i) is given by the horizontal time-scale difference between the time of arrival and the time of departure.

2. The total number of vehicles queued at any time (q_t) is the vertical vehicle-scale difference between the number of vehicles that have arrived and the number of vehicles that have departed

3. The aggregate delay for all vehicles passing through the signal is the area between the arrival and departure curves (vehicles times the time duration)

Uniform Delay

Uniform delay is the delay based on an assumption of uniform arrivals and stable flow with no individual cycle failures. Fig. below, shows stable flow throughout the period depicted. No signal cycle fails here, i.e., no vehicles are forced to wait for more than one green phase to be discharged. During every green phase, the departure function catches up with the arrival function. Total aggregate delay during this period is the total of all the triangular areas between the arrival and departure curves. This type of delay is known as Uniform delay.
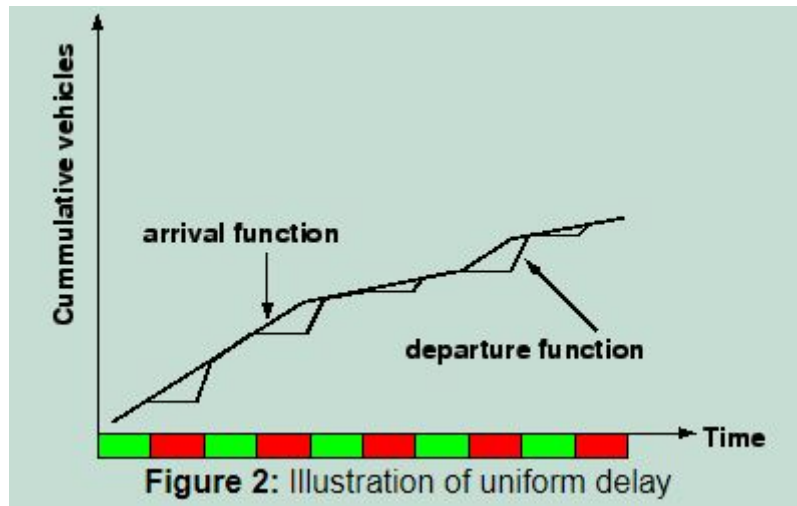
Figure: Illustration of uniform delay.

Random Delay

Random delay is the additional delay, above and beyond uniform delay, because flow is randomly distributed rather than uniform at isolated intersections. In Fig. below some of the signal phases fail. At the end of the second and third green intervals, some vehicles are not served (i.e., they must wait for a second green interval to depart the intersection). By the time the entire period ends, however, the departure function has caught up with the arrival function and there is no residual queue left unserved. This case represents a situation in which the overall period of analysis is stable (ie., total demand does not exceed total capacity). Individual cycle failures within the period, however, have occurred. For these periods, there is a second component of delay in addition to uniform delay. It consists of the area between the arrival function and the dashed line, which represents the capacity of the intersection to discharge vehicles, and has the slope c. This type of delay is referred to as Random delay.
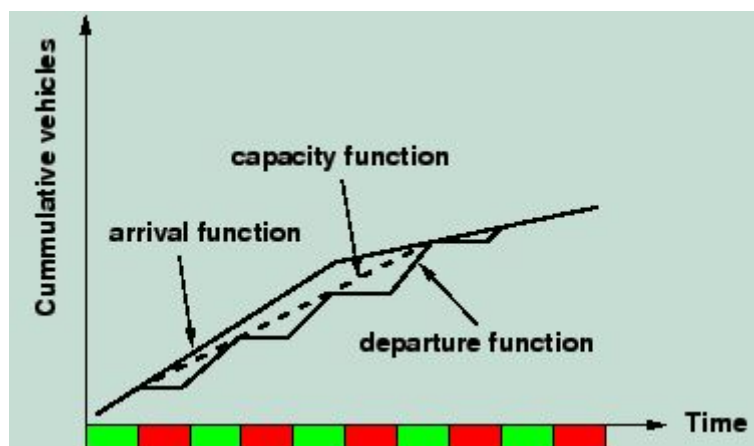

Figure: Illustration of random delay.

Overflow Delay

Overflow delay is the additional delay that occurs when the capacity of an individual phase or series of phases is less than the demand or arrival flow rate. Fig. below shows the worst possible case, every green interval fails for a significant period of time, and the residual, or unserved, queue of vehicles continues to grow throughout the analysis period. In this case, the overflow delay component grows over time, quickly dwarfing the uniform delay component. The latter case illustrates an important practical operational characteristic. When demand exceeds capacity ($v/c > 1.0$), the delay depends upon the length of time that the condition exists. In Figure 3, the condition exists for only two phases. Thus, the queue and the resulting overflow delay is limited. In Fig. 4, the condition exists for a long time, and the delay continues to grow throughout the over-saturated period. This type of delay is referred to as Overflow delay.
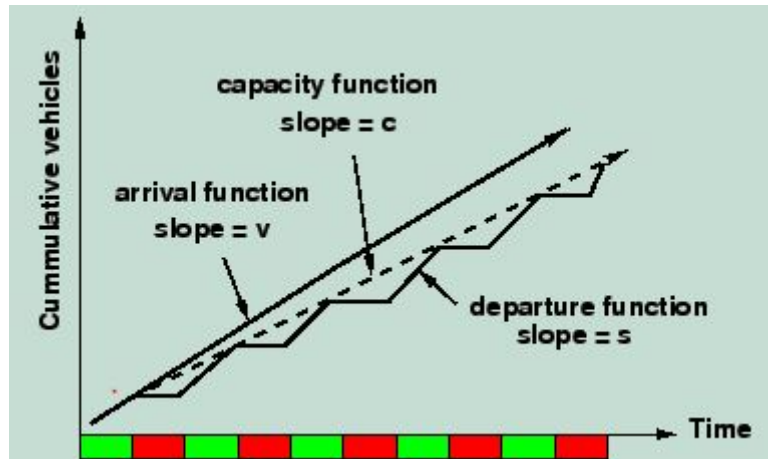

Figure: Illustration of overflow delay.

Webster's Delay Models

Uniform Delay Model

Model is explained based on the assumptions of stable flow and a simple uniform arrival function. As explained in the above section, aggregate delay can be estimated as the area between the arrival and departure curves. Thus, Webster's model for uniform delay is the area of the triangle formed by the arrival and departure functions. From Fig. below the area of the aggregate delay triangle is simply one-half the base times the height. Hence, the total uniform delay (TUD) is given as:
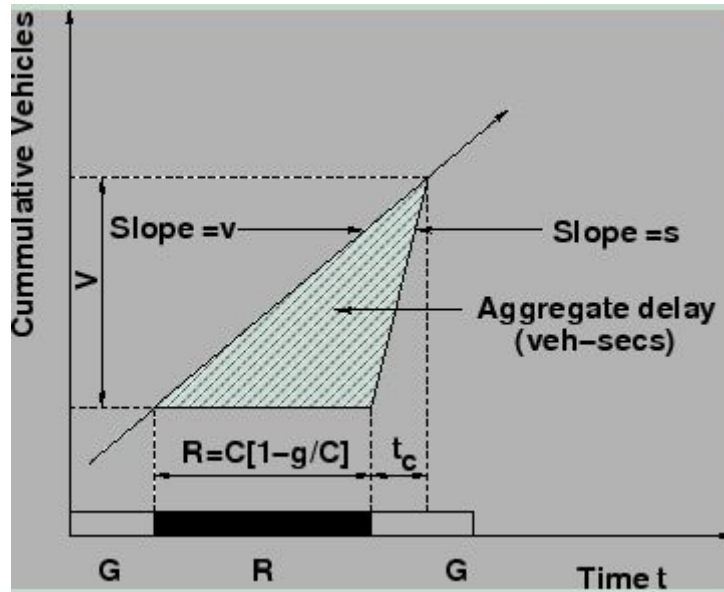
$$TUD = \frac{RV}{2}$$

Figure: Illustration of Webster's uniform delay model.

where, R is the duration of red, V is the number of vehicles arriving during the time interval $R + t_c$. Length of red phase is given as the proportion of the cycle length which is not green, or:

$$R = C\left(1 - \frac{g}{C}\right)$$

The height of the triangle is found by setting the number of vehicles arriving during the time ($R + t_c$) equal to the number of vehicles departing in time t_c, or:

$$V = v(R + t_c) = s\, t_c$$

Substituting equation for R and solving for t_c and then for V gives,

$$v\left(C\left(1 - \frac{g}{C}\right) + t_c\right) = s\, t_c$$

$$t_c\,(s - v) = v\, C\left(1 - \frac{g}{C}\right)$$

$$\frac{V}{s}(s - v) = v\, C\left(1 - \frac{g}{C}\right)$$

$$V = C\left(1 - \frac{g}{C}\right)\left(\frac{vs}{s - v}\right)$$

$$TUD = \frac{RV}{2}$$

$$= \frac{1}{2}\left[C\left(1 - \frac{g}{C}\right)\right]^2 \left(\frac{vs}{s-v}\right)$$

where TUD is the aggregate delay, in vehicle seconds. To obtain the average delay per vehicle, the aggregate delay is divided by the number of vehicles processed during the cycle, which is the arrival rate, v, times the full cycle length, C. Hence,

$$UD = \frac{1}{2}\left[C\left(1 - \frac{g}{C}\right)\right]^2 \left(\frac{vs}{s-v}\right)\left(\frac{1}{vC}\right)$$

$$= \frac{C\left(1 - \frac{g}{C}\right)^2}{2\left(1 - \frac{v}{s}\right)}$$

Another form of the equation uses the capacity, c, rather than the saturation flow rate, s. We know,

$$s = \frac{c}{\frac{g}{C}}$$

So, the relation for uniform delay changes to,

$$UD = \frac{C\left(1 - \frac{g}{C}\right)^2}{2\left(1 - \frac{gv}{Cc}\right)}$$

$$= \frac{C\left(1 - \frac{g}{C}\right)^2}{2\left(1 - \frac{g}{C}X\right)}$$

where, UD is the uniform delay (sec/vehicle) C is the cycle length (sec), c is the capacity, v is the vehicle arrival rate, s is the saturation flow rate or departing rate of vehicles, X is the v/c ratio or degree of saturation (ratio of the demand flow rate to saturation flow rate), and g/C is the effective green ratio for the approach.

Numerical Example

Consider the following situation: An intersection approach has an approach flow rate of 1,000 vph, a saturation flow rate of 2,800 vph, a cycle length of 90 s, and effective green ratio of 0.44 for the approach. What average delay per vehicle is expected under these conditions?

Solution:

First, the capacity and v/c ratio for the intersection approach must be computed. Given, s=2800 vphg and g/C=0.55. Hence,

	$c = s \times g/C$
	$= 2800 \times 0.55 = 1540 \text{ vph}$
$v/c =$	$1000/1540 = 0.65$

Since $v/c \leq 1.0$ and is a relatively low value, the uniform delay equation 5 may be applied directly. There is hardly any random delay at such a v/c ratio and overflow delay need not be considered.

$$UD = \frac{C}{2} \frac{\left(1 - \frac{g}{C}\right)^2}{\left(1 - \frac{v}{s}\right)}$$

$$= \frac{90}{2} \frac{(1 - 0.55)^2}{\left(1 - \frac{1000}{2800}\right)}$$

$$= 14.2.$$

Therefore, average delay per vehicle is 14.2 sec/veh. Note that this solution assumes that arrivals at the subject intersection approach are uniform. Random and platooning effects are not taken into account.

Random Delay Model

The uniform delay model assumes that arrivals are uniform and that no signal phases fail (i.e., that arrival flow is less than capacity during every signal cycle of the analysis period). At isolated intersections, vehicle arrivals are more likely to be random. A number of stochastic models have been developed for this case, including those by Newell, Miller and Webster. These models generally assume that arrivals are Poisson distributed, with an underlying average rate of v vehicles per unit time. The models account for random arrivals and the fact that some individual cycles within a demand period with $v/c < 1.0$ could fail due

to this randomness. This additional delay is often referred to as Random delay. The most frequently used model for random delay is Webster's formulation:

$$RD = \frac{X^2}{2\,v\,(1-X)} \tag{1}$$

where, RD is the average random delay per vehicle, s/veh, and X is the degree of saturation (v/c ratio). Webster found that the above delay formula overestimate delay and hence he proposed that total delay is the sum of uniform delay and random delay multiplied by a constant for agreement with field observed values. Accordingly, the total delay is given as:

$$D = 0.90\,(UD + RD) \tag{2}$$

Overflow delay model

Model is explained based on the assumption that arrival function is uniform. In this model a new term called over saturation is used to describe the extended time periods during which arrival vehicles exceeds the capacity of the intersection approach to the discharged vehicles. In such cases queue grows and there will be overflow delay in addition to the uniform delay.

Fig. below illustrates a time period for which $v/c > 1.0$.

Figure: Illustration of overflow delay model.

During the period of over-saturation delay consists of both uniform delay (in the triangles between the capacity and departure curves) and overflow delay (in the triangle between arrival and capacity curves). As the maximum value of X is 1.0 for uniform delay, it can be simplified as,

$$UD = \frac{C\,(1-\frac{g}{C})^2}{2\,(1-\frac{g}{C}X)}$$

$$= \frac{C}{2}\left(1 - \frac{g}{C}\right)$$

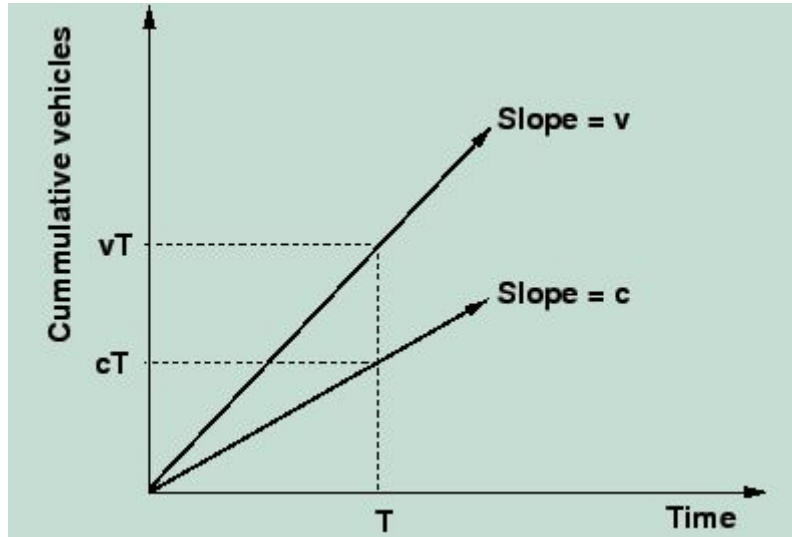From Fig. below total overflow delay can be estimated as,

**Figure:** Derivation of the overflow delay model.

$$TOD = \frac{1}{2}T(vT - cT) = \frac{T^2}{2}(v - c) \tag{4}$$

where, TOD is the total or aggregate overflow delay (in veh-secs), and T is the analysis period in seconds. Average delay is obtained by dividing the aggregate delay by the number of vehicles discharged with in the time T which is cT.

$$OD = \frac{T}{2}\left(\frac{v}{c} - 1\right) \tag{5}$$

The delay includes only the delay accrued by vehicles through time T, and excludes additional delay that vehicles still stuck in the queue will experience after time T. The above said delay equation is time dependent i.e., the longer the period of over-saturation exists, the larger delay becomes. A model for average overflow delay during a time period T1 through T2 may be developed, as illustrated in Fig. below Note that the delay area formed is a trapezoid, not a triangle. The resulting model for average delay per vehicle during the time period T1 through T2 is:

$$OD = \frac{T_1 + T_2}{2}\left(\frac{v}{c} - 1\right) \tag{6}$$
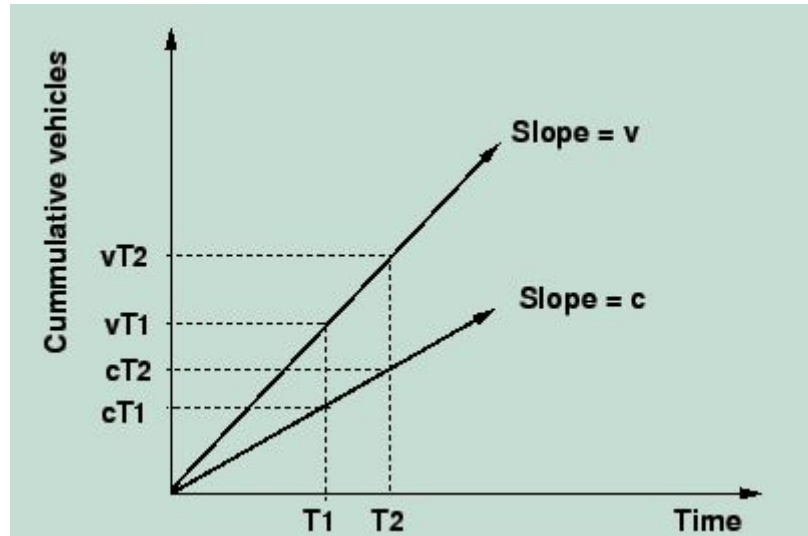
Formulation predicts the average delay per vehicle that occurs during the specified interval, T_1 through T_2. Thus, delays to vehicles arriving before time T_1 but discharging after T_1 are included only to the extent of their delay within the specified times, not any delay they may have experienced in queue before T_1. Similarly, vehicles discharging after T_2 are not included in the formulation.

Numerical Example

Consider the following situation: An intersection approach has an approach flow rate of 1,900 vph, a saturation flow rate of 2,800 vphg, a cycle length of 90s, and effective green ratio for the approach 0.55. What average delay per vehicle is expected under these conditions?

Solution:

To begin, the capacity and v/c ratio for the intersection approach must be computed. This will determine what model(s) are most appropriate for application in this case: Given, s=2800 vphg, g/C=0.55, and v =1900 vph.

$$c = s \times g/C$$

$$= 2800 \times 0.55 = 1540 \text{ vph}$$

$$v/c = 1900/1540 = 1.23.$$

Since v/c is greater than 1.15 for which the overflow delay model is good so it can be used to find the delay.

D $= UD + OD$

UD $= \dfrac{C}{2}\left(1 - \dfrac{g}{C}\right)$

$= 0.5 \times 90[1 - 0.55]$

$= 20.3$

OD $= \dfrac{T}{2}\left(\dfrac{v}{c} - 1\right)$

$= \dfrac{3600}{2}(1.23 - 1)$

$= 414$

D $= 20.3 + 414$

$= 434.3 \text{ sec/veh}$

This is a very large value but represents an average over one full hour period.

Inconsistencies between Random and Overflow delay

As explained earlier random and overflow delay is given as, Random delay,

$$RD = \frac{(X)^2}{2v(1-X)} \tag{1}$$

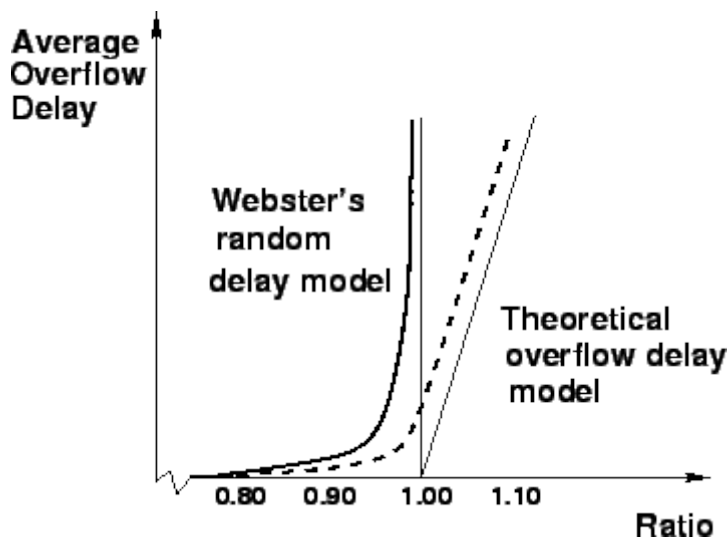Overflow delay,

$$OD = \frac{T}{2}(X-1) \tag{2}$$

Figure : Comparison of overflow & random delay model

The inconsistency occurs when the X is in the vicinity of 1.0. When $X < 1.0$ random delay model is used. As the Webster's random delay contains 1-X term in the denominator, when X approaches to 1.0 random delay increases asymptotically to infinite value. When $X > 1.0$ overflow delay model is used. Overflow delay contains $1-X$ term in the numerator, when X approaches to 1.0 overflow becomes zero and increases uniformly with increasing value of X. Both models are not accurate in the vicinity of $X = 1.0$. Delay does not become infinite at $X = 1.0$. There is no true overflow at $X = 1.0$, although individual cycle failures due to random arrivals do occur. Similarly, the overflow model, with overflow delay of zero seconds per vehicle at $X = 1.0$ is also unrealistic. The additional delay of individual cycle failures due to the randomness of arrivals is not

reflected in this model. Most studies show that uniform delay model holds well in the range $X \leq 0.85$. In this range true value of random delay is minimum and there is no overflow delay. Also overflow delay model holds well when $X \geq 1.15$. The inconsistency occurs in the range $0.85 \leq X \leq 1.15$; here both the models are not accurate. Much of the more recent work in delay modeling involves attempts to bridge this gap, creating a model that closely follows the uniform delay model at low X ratios, and approaches the theoretical overflow delay model at high X ratios, producing "reasonable" delay estimates in between. Fig. above illustrates this as the dashed line.

HCM 2000 delay models

The delay model incorporated into the HCM 2000 includes the uniform delay model, a version of Akcelik's overflow delay model, and a term covering delay from an existing or residual queue at the beginning of the analysis period. The delay is given as,

$$d = d_1 \, PF + d_2 + d_3$$

$$d_1 = \frac{c}{2} \frac{(1 - \frac{g}{c})^2}{1 - [min(1, X)(\frac{g}{c})]}$$

$$d_2 = 900T[(X - 1) + \sqrt{(X - 1)^2 + \frac{8klX}{cT}}]$$

$PF = (\frac{(1-P)}{(1-(g/c))}) * f_p$ additional explanation for PF Where, d = control delay, s/veh, d1 = uniform delay

component, s/veh, PF = progression adjustment factor, d2 = overflow delay component, s/veh, d3 = delay due to pre-existing queue, s/veh, T = analysis period, h, X = v/c ratio,

C = cycle length, s,

 k = incremental delay factor for actuated controller settings; 0.50 for all pre-timed controllers, l = upstream filtering/metering adjustment factor; 1.0 for all individual intersection analyses,

c = capacity, veh/h, P = proportion of vehicles arriving during the green interval

and f_p = supplemental adjustment factor for platoon arriving during the green

Numerical problems

Consider the following situation: An intersection approach has an approach flow rate of 1,400 vph, a saturation flow rate of 2,650 vphg, a cycle length of 102 s, and effective green ratio for the approach 0.55. Assume Progression Adjustment Factor 1.25 and delay due to pre-existing queue, 12 sec/veh. What control delay sec per vehicle is expected under these conditions?

Solution:

Saturation flow rate =2650 vphg , g/C=0.55, Approach flow rate v=1700 vph, Cycle length C=102 sec, delay due to pre-existing queue =12 sec/veh and Progression Adjustment Factor PF=1.25. The capacity is given as:

$$c = s \times \frac{g}{C}$$

$$= 2650 \times 0.55$$

$$= 1458 \ vphv/c$$

Degree of saturation X= v/c= 1700/1458 =1.16. So the uniform delay is given as:

$$d_1 = \frac{C}{2} \frac{\left(1 - \frac{g}{C}\right)^2}{\left[1 - min(X, 1)\left(\frac{g}{c}\right)\right]}$$

$$= \frac{102}{2} \frac{(1 - 1.16)^2}{\left[1 - min(1.16, 1)(.55)\right]} = 22.95$$

Uniform delay =22.95 and the over flow delay is given as:

$$d_2 = \frac{C}{2} * (X - 1) + \sqrt{(X - 1)^2 + \frac{8klX}{cT}}$$

$$= \frac{102}{2} * (1.16 - 1) + \sqrt{(1.16 - 1)^2 + \frac{(8 * 5 * 1 * 1.16)}{1458}} = 16.81$$

Overflow delay, d2=16.81. Hence, the total delay is"

$$d = d_1 \ PF + d_2 + d_3$$

$$= 22.95 \times 1.25 + 16.81 + 8 = 53.5 \ sec/veh.$$

Therefore, control delay per vehicle is 53.5 sec.