Traffic Management and Control

A. Traffic Signal Control – Introduction

Prof. Nathan H. Gartner
University of Massachusetts
Lowell, MA, USA

© NHGartner, UML, July 2013
Module Objectives:

- Describe the basic operation principles of traffic signal systems
- Present the principal warrants for traffic signals
- Explain fixed-time, and traffic-responsive operation
- Describe performance measures for isolated intersections
- Develop optimal signal timings for isolated intersections

Students should have received prior instruction in the following areas:

- Traffic Flow Principles and Queuing Models

© NHGartner, UML, July 2013
Overview

• Traffic signal control, including sensors, communications, and advanced timing algorithms, is a dynamic, continually evolving form of transportation management.

• While new technology has resulted in improved system components and tools for traffic control signal operations, these components must be properly designed, installed, operated, and maintained if their full potential is to be realized.
Purpose of Traffic Control Signals

The primary function of *Traffic Control Signals* is to assign the right-of-way at intersecting streets or highways where, without such control, a continual flow of vehicles on one roadway would cause excessive delay to vehicles and/or pedestrians waiting on the other roadway.

*Freeway Ramp Control Signal* are a special application of traffic control signals installed on freeway entrance ramps to limit, or “meter,” the amount of traffic entering the freeway.
Required steps in design:

- investigating the need for a traffic signal
- determining the operational requirements
- translating these requirements into traffic control equipment requirements
- determining optimum operation of the traffic signal and
- operating and maintaining the traffic control signal over its expected life.

The information presented in this module builds on procedures documented in the ITE Transportation and Traffic Engineering Handbook as well as the standards and warrants provided in the Manual on Uniform Traffic Control Devices (MUTCD).
Basic Types of Control

Traffic control signals are usually described as either *pre-timed* or *traffic actuated*. Each type may be used in either an independent (isolated) or interconnected (system) application.

**Pre-timed control**: The electronic control circuits provide a repetitive cycle and split (cycle division among the conflicting movements) timing. The timing is repeated over and over regardless of the presence or absence of traffic demand. When operating as part of a system, adjacent intersections operate on the same cycle length and have fixed offsets (relationship of beginning of main street green displays). There may be multiple patterns of cycle lengths, offsets and splits.
Basic Types of Control –(cont'd)

➢ **Actuated control**: The timing is varied for some or all controlled conflicting movements dependent upon vehicular and/or pedestrian demand. Demand is determined from detectors placed in or near the roadway of pedestrian crossing. When all controlled conflicting movements are timed relative to demand, the control is termed “full-actuated.” When only secondary movements vary with demand, the control is termed “semi-actuated.”

➢ **Full-actuated control**: It is generally used for an isolated application. An isolated application is one where the traffic control signal operates independent of any other traffic control signal. Conversely, a system or interconnected application means that a given traffic control signal’s operation is related to (coordinated with) one or more other traffic control signal locations. In system applications, a fully actuated controller must operate as a semi-actuated controller because it must operate on the same cycle length as all the other controllers that are interconnected with it. In systems, pre-timed and actuated control may be mixed.
B. Needs Assessments

Prof. Nathan H. Gartner
University of Massachusetts
Lowell, MA, USA
Section B: Needs Assessment -

Determining the Need for Traffic Signal Control

The first and basic question that must be addressed is whether or not traffic signalization is needed. Since traffic signals are the most restrictive traffic control devices, they should be used only where the less restrictive signs or markings do not provide the necessary level of control.
Required Studies

A comprehensive investigation of traffic conditions and physical characteristics of the location is required to determine the necessity for a signal installation. The engineering study should include:

1. The total number of vehicles entering the intersection in each hour from each approach during 16 consecutive hours of a representative day. The 16 hours selected should contain the greatest percentage of the 24-hour traffic.

2. Vehicular volumes for each traffic movement from each approach, classified by vehicle type during AM and PM 2-hour peak.

3. Pedestrian volume counts on each crosswalk during the same periods as the vehicular counts.
4. Information about nearby facilities and activity centers that serve the young, elderly, and/or persons with disabilities.

5. The 85-percentile speed of all vehicles on the uncontrolled approaches to the location.

6. A conditions diagram showing details of the physical layout, including such features as intersection geometrics, channelization, grades, sight-distance restrictions, bus stops and routings, parking conditions, etc.

7. A collision diagram showing accident experience by type, location, direction of movement, severity, time of day, date and day of week for at least one year.
The following data are also desirable for a more thorough understanding of the operation of the intersection and may be obtained when volume data by approach is being collected (the second step described on the previous page):

1. Vehicle-hours of stopped time delay determined separately for each approach.
2. The number and distribution of acceptable gaps in vehicular traffic on the major street for entrance from the minor street.
3. The posted or statutory speed limit or the 85th-percentile speed on controlled approaches at a point near to the intersection but unaffected by the control.
4. Pedestrian delay time for at least two 30-minute peak pedestrian delay periods of an average weekday or like periods of a Saturday or Sunday.
Adequate roadway capacity at a signalized intersection is desirable. Widening of both the major street and the minor street may be warranted to reduce the delays caused by assignment of right-of-way at intersections controlled by traffic signals. Widening of the minor street is often beneficial to operation on the major street because it reduces the green time that must be assigned to minor street traffic.
Warrants for Traffic Signal Installation

Traffic control signals should not be installed unless one or more of the signal warrants in the **MUTCD** are met:

- **Warrant 1** - Condition A, Minimum Vehicular Volume and Condition B, Interruption of Continuous Traffic
- **Warrant 2** - Four-Hour Vehicular Volume
- **Warrant 3** - Peak Hour
- **Warrant 4** - Pedestrian Volume
- **Warrant 5** - School Crossings
- **Warrant 6** - Coordinated Signal System
- **Warrant 7** - Crash Experience
- **Warrant 8** - Roadway Network
Warrant 1: Eight-Hour

Warrant 1, The Eight-Hour Vehicular Volume includes two components:

- **Condition A** - Minimum Vehicular Volume
- **Condition B** - Interruption of Continuous Traffic

Warrant 1 is meant to be treated as a single warrant. If Condition A is satisfied, then the criteria for Warrant 1 is satisfied and Condition B and the combination of Conditions A and B are not needed.

Similarly, if Condition B is satisfied, then the criteria for Warrant 1 is satisfied and the combination of Conditions A and B is not needed. We will address each condition in turn.
Warrant 1: Eight-Hour Vehicular Volume - (cont'd)

**Condition A—Minimum Vehicular Volume**

The Minimum Vehicular Volume, Condition A, is intended for application where the volume of intersection traffic is the principal reason to consider installing a traffic control signal.

The warrant is satisfied when, for each of any 8 hours of an average day, the traffic volumes shown in both of the 100 percent columns of Table 1 (on the next screen) exist on the major street and on the higher-volume minor-street approach to the intersection.

When the 85-percentile speed of major-street traffic exceeds 40 mph, or when the intersection lies within the built-up area of an isolated community having a population of less than 10,000, the Minimum Vehicular Volume warrant is 70 percent of the requirements shown in the table.

An "average" day is defined as a weekday representing traffic volumes normally and repeatedly found at the location.
Warrant 1: Eight-Hour Vehicular Volume - (cont'd)

<table>
<thead>
<tr>
<th>Number of lanes for moving traffic on each approach</th>
<th>Vehicular volume on major street (total of both approaches)</th>
<th>Vehicular volume on higher-volume minor-street approach (one direction only)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Street</td>
<td>100% a</td>
<td>80% b</td>
</tr>
<tr>
<td>Minor Street</td>
<td>500</td>
<td>400</td>
</tr>
<tr>
<td>1..........</td>
<td>600</td>
<td>480</td>
</tr>
<tr>
<td>2 or more</td>
<td>2 or more</td>
<td>600</td>
</tr>
<tr>
<td>2 or more</td>
<td>1..........</td>
<td>500</td>
</tr>
</tbody>
</table>

Mouse over the blue note numbers for an explanation.
Warrant 1: Eight-Hour Vehicular Volume - (cont'd)

Warrant 1. Eight-Hour Vehicular Volume
Condition B — Interruption of Continuous Traffic
Table 2

<table>
<thead>
<tr>
<th>Number of lanes for moving traffic on each approach</th>
<th>Vehicles per hour on major street (total of both approaches)</th>
<th>Vehicles per hour on higher-volume minor-street approach (one direction only)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Street</td>
<td>100% a 80% b 70% c 56% d</td>
<td>100% a 80% b 70% c 56% d</td>
</tr>
<tr>
<td>Minor Street</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.........</td>
<td>750 600 525 420</td>
<td>75 60 53 42</td>
</tr>
<tr>
<td>2 or more</td>
<td>900 720 630 504</td>
<td>75 60 53 42</td>
</tr>
<tr>
<td>1.........</td>
<td>900 720 630 504</td>
<td>100 80 70 56</td>
</tr>
<tr>
<td>2 or more</td>
<td>900 720 630 504</td>
<td>100 80 70 56</td>
</tr>
<tr>
<td>1.........</td>
<td>750 600 525 420</td>
<td>100 80 70 56</td>
</tr>
</tbody>
</table>
Warrant 1: Eight-Hour Vehicular Volume - (cont'd)

Combination of Conditions A and B

The combination of Conditions A and B is intended for application at locations where Condition A is not satisfied and Condition B is not satisfied and should be applied only after an adequate trial of other alternatives that could cause less delay and inconvenience to traffic has failed to solve the traffic problems. Combination of Conditions A and B applies when the criteria below are met:

1. The vehicles per hour given in both of the 80 percent columns of Condition A in Table 1 exist on the major-street and the higher-volume minor-street approaches, respectively, to the intersection;

2. The vehicles per hour given in both of the 80 percent columns of Condition B in Table 2 exist on the major-street and the higher-volume minor-street approaches, respectively, to the intersection.
Warrant 2: Four-Hour Vehicular Volume

The Four-Hour Vehicular Volume Warrant is satisfied when, on each of any four hours of an average day, the plotted points representing the vehicles per hour on the major street (total of both approaches) and the corresponding vehicles per hour on the higher volume minor street approach (one direction only) all fall above the curve in Figure 1 below for the existing combination of approach lanes.

Note: 115 VPH applies as the lower threshold volume for a minor-street approach with two or more lanes and 80 VPH applies as the lower threshold volume for a minor-street approach with one lane.
If the posted or statutory speed limit or the 85th-percentile speed on the major street exceeds 70 km/h or exceeds 40 mph or if the intersection lies within the build-up area of an isolated community having a population of less than 10,000, Figure 2 below may be used in place of Figure 1.

Note: 80 VPH applies as the lower threshold volume for a minor-street approach with two or more lanes and 60 VPH applies as the lower threshold volume for a minor-street approach with one lane.
The Peak Hour Warrant applies where traffic conditions are such that for one hour of the day minor street traffic suffers undue delay in entering or crossing the major street. This signal warrant applies only in unusual cases, such as office complexes, manufacturing plants, industrial complexes, or high-occupancy vehicle facilities that attract or discharge large numbers of vehicles over a short time. The peak hour delay warrant is justified if all three of the following conditions exist for the same 1 hour (any four consecutive 15-minute periods) of an average day:

1. The total delay experienced by the traffic on one minor street approach (on direction only) controlled by a STOP sign equals or exceeds four vehicle-hours for a one-lane approach and five vehicle hours for a two-lane approach;

2. The volume on the same minor street approach (one direction only) equals or exceeds 100 vph for one moving lane of traffic or 150 vph for two moving lanes;

3. The total entering volume serviced during the hour equals or exceeds 800 vph for intersections with four (or more) approaches or 650 vph for intersections with three approaches.
The Peak-Hour Warrant is also satisfied when the plotted point representing the vehicles per hour on the major street (total of both approaches) and the corresponding vehicles per hour on the higher-volume minor-street approach (one direction only) for 1 hour (any four consecutive 15-minute periods) of an average day falls above the applicable curve in Figure 3 below for the existing combination of approach lanes.

*Note: 150 VPH applies as the lower threshold volume for a minor-street approach with two or more lanes and 100 VPH applies as the lower threshold volume for a minor-street approach with one lane.
Warrant 3: Peak Hour – (cont'd)

When the 85th percentile speed of major street traffic exceeds 40 mph or when the intersection lies within a built-up area of an isolated community having a population less than 10,000, the peak hour volume requirements is satisfied when the plotted point referred to above falls above the curve in Figure 4 below for the existing combination of approach lanes.

*Note: 100 VPH applies as the lower threshold volume for a minor-street approach with two or more lanes and 75 VPH applies as the lower threshold volume for a minor-street approach with one lane.
Warrant 4: Pedestrian Volume

The Pedestrian Volume signal warrant is intended for application where the traffic volume on a major street is so heavy that pedestrians experience excessive delay in crossing the major street. Two criteria warrant signal installation:

A. The pedestrian volume crossing the major street at an intersection or midblock location during an average day is 100 or more for each of any 4 hours or 190 or more during any 1 hour.

B. There are fewer than 60 gaps per hour in the traffic stream of adequate length to allow pedestrians to cross during the same period when the pedestrian volume criterion is satisfied. Where there is a divided street having a median of sufficient width for pedestrians to wait, the requirement applies separately to each direction of vehicular traffic.

Continue
The Pedestrian Volume signal warrant shall not be applied at locations where the distance to the nearest traffic control signal along the major street is less than 90 m (300 ft), unless the proposed traffic control signal will not restrict the progressive movement of traffic.

Additionally, the criterion for the pedestrian volume crossing the major roadway may be reduced as much as 50 percent if the average crossing speed of pedestrians is less than 1.2 m/sec (4 ft/sec).

Finally, a traffic control signal may not be needed at the study location if adjacent coordinated traffic control signals consistently provide gaps of adequate length for pedestrians to cross the street, even if the rate of gap occurrence is less than one per minute.
Warrant 5: School Crossing

A traffic control signal may be warranted an engineering study of the frequency and adequacy of gaps in the vehicular traffic stream as related to the number and size of groups of school children at an established school crossing across the major street shows that:

A. The number of adequate gaps in the traffic stream during the period when the children are using the crossing is less than the number of minutes in the same period.

B. There are a minimum of 20 students during the highest crossing hour.

First consideration will be given to the implementation of other remedial measures, such as warning signs and flashers, school speed zones, school crossing guards, or a grade-separated crossing.

Additionally, the warrant shall not be applied at locations where the distance to the nearest traffic control signal along the major street is less than 90 m (300 ft), unless the proposed traffic control signal will not restrict the progressive movement of traffic.
Warrant 6: Coordinated Signal System

Progressive movement in a coordinated signal system sometimes necessitates installing traffic control signals at intersections where they would not otherwise be needed in order to maintain proper platooning of vehicles. Two criteria apply:

A. On a **one-way street** or a street that has traffic predominantly in one direction, the adjacent traffic control signals are so far apart that they do not provide the necessary degree of vehicular platooning.

B. On a **two-way street**, adjacent traffic control signals do not provide the necessary degree of platooning and the proposed and adjacent traffic control signals will collectively provide a progressive operation.

The warrant shall not be applied at locations where the resultant spacing of traffic control signals would be less than 300 m (1,000 ft).
Warrant 7: Crash Experience

This warrant applies when the severity and frequency of crashes are the principal reasons to consider installing a traffic control signal. Three criteria must be met:

A. Adequate trial of alternatives with satisfactory observance and enforcement has failed to reduce the crash frequency.

B. Five or more reported crashes, of types susceptible to correction by a traffic control signal, have occurred within a 12-month period, each crash involving personal injury or property damage apparently exceeding the applicable requirements for a reportable crash.

C. There exists a volume or vehicular or pedestrian traffic not less than 80% of the requirements specified either in Warrants 1 (Minimum Vehicular, Interruption of Continuous Traffic) or 4 (Pedestrian Volume).

When the 85-percentile speed of major-street traffic exceeds 40 mph in either an urban or a rural area or, when the intersection lies within the built-up area of an isolated community having a population of less than 10,000, the traffic volumes in the Minimum Vehicular Table may be reduced to 56%.
Warrant 8: Roadway Network

This warrant applies if the common intersection of two or more major routes meets one or both of the following criteria:

A. The intersection has a total existing, or immediately projected, entering volume of at least 1,000 vehicles per hour during the peak hour of a typical weekday and has 5-year projected traffic volumes, based on an engineering study, that meet one or more of Warrants 1, 2, and 3 during an average weekday;

B. The intersection has a total existing or immediately projected entering volume of at least 1,000 vehicles per hour for each of any 5 hours of a nonnormal business day (Saturday or Sunday).

What is defined as a major route?
A major route as used in this signal warrant shall have one or more of the following characteristics:

1. It is part of the street or highway system that serves as the principal roadway network for through traffic flow; or

2. It includes rural or suburban highways outside, entering, or traversing a City; or

3. It appears as a major route on an official plan, such as a major street plan in an urban area traffic and transportation study.
Traffic Management and Control

C. Operational Requirements

Prof. Nathan H. Gartner
University of Massachusetts
Lowell, MA, USA

© NHGartner, UML, July 2013
Section C: Operational Requirements

After it has been established that a signal is warranted at a particular location (or locations), the next major decision in the design process involves determining the most appropriate method of control. Decision to be made include:

- Controller phasing
- Pre-timed or actuated operation
- Interconnection considerations
Phasing Elements

A signal phase may be defined as that part of the cycle length allocated to a traffic movement receiving the right of way or to any combination of traffic movements receiving the right of way simultaneously.

A traffic movement is a single vehicular movement, a single pedestrian movement, or a combination of vehicular and pedestrian movements.

The sum of all traffic phases is equal to the cycle length.
Example: 3-phase controller sequence
Phasing Elements (cont'd)

Phases are commonly added for protecting left turns. The basic sequences which accommodate left-turn movements include:

**Heaviest left turn protected.** This is a "lead left" in which the left-turning vehicles from only one approach are protected and move on an arrow indication preceding the opposing through movement; or a "lag left" in which the protected left turn follows the trough movement phase.

**Both left turns protected-no overlap.** When the opposing left turns move simultaneously followed by the through movements, it is termed "lead dual left." If the left turns follow the through movement it is called a "lag dual left."
Phasing Elements (cont'd)

Both left turns protected with overlap. In this operation, opposing left turns start simultaneously. When one terminates, the through movement in the same direction is started. When the extending left is terminated, the other through movement is started. When this type of phasing is used on both streets, it is termed "quad left phasing."

Lead lag. This phasing is combined with a leading protected left in one direction, followed by the through movements, followed by a lag left in the opposing direction. It is sometimes used in systems to provide a wider two-way through band.

Directional separation. Each approach obtains exclusive right of way with all opposing traffic stopped.
Example: 8-phase dual ring controller
Types of Control

The principal types of traffic signal control are pre-timed and traffic actuated. Each type of control has its unique advantages and disadvantages.

Pre-timed control assigns the right of way at an intersection according to a predetermined schedule. The length of the time interval for each signal indication in the cycle is fixed, based on historic traffic patterns.
Actuated & Semi-Actuated Control

Actuated control differs from pre-timed in that signal phases are not of fixed length. Through the use of vehicle detectors, this type of control assigns the right of way on the basis of current traffic conditions (demand) within given limitations.

Fully-actuated control requires detectors for all phases with each phase timed according to preset parameters. Fully-actuated control is primarily used at the intersection of streets with approximately equal volumes with sporadic and varying traffic distribution.
Semi-actuated control requires detectors on the minor street approaches and is especially effective at intersections where the major street has a relatively uniform flow and the minor street has low volumes with random peaks. Semi-actuated control, therefore, only takes “away” time from the major street movements when there is actually demand on the minor street.
Choosing a Type of Control

In general practice, the rule of thumb for choosing the type of intersection control is:

• for predictable traffic demand, use pre-timed (Warrant 1);
• for unpredictable traffic demand, use actuated control (Warrant 2).
Interconnection Considerations

The potential benefits to be derived from coordinated operation of two or more signalized intersections are directly related to the “platoon” arrival characteristics at the downstream intersection.

If approaching vehicles arrive at the stop line as a well-defined compact platoon, coordinated operation can provide a significant reduction in stops and delays.

The MUTCD suggests that signals spaced less than ½ miles apart should be coordinated because the cohesion of the platoon can be maintained for this distance.
D. Single Intersection Models

Prof. Nathan H. Gartner
University of Massachusetts
Lowell, MA, USA

© NHGartner, UML, July 2013
Section D: Single Intersection Models

The material in this section is based on the methodology developed by F. V. Webster at the British Road Research Laboratory for estimating the delays to vehicles at fixed-time traffic signals and for computing the optimum settings of such signals.
We will begin our discussion of traffic signal settings by describing two key concepts:

- delay and
- saturation flow rate.
Definition of Delay

The most commonly used measure of effectiveness for signalized intersections is delay.

Delay to the traffic is computed by counting the number of vehicles in the queue at fixed intervals of time and multiplying this number by the value of the time interval.
Saturation Flow

Saturation flow occurs when the queue discharges at a more or less constant rate.

The saturation flow rate is measured in units of vehicles per hour of green time per lane.
Effective Green Time

If there is still a queue at the end of the green period, some vehicles will make use of the yellow interval to cross the intersection. In these circumstances traffic moves on both green and yellow signals. To simplify analysis, it is advantageous to use the concept of “effective green time.” The green and yellow periods together may be considered as an 'effective' green \( (g) \) and a 'lost' time \( (l) \).

The effective green time is determined such that the product of the effective green time and the saturation flow rate is equal to the actual number of vehicles (say, \( b \)) discharged from the queue on the average in a saturated green period (i.e. a green period during which the queue never clears).
Signalized Intersections: Fluid Traffic Model

In this section a continuum (i.e. deterministic) or fluid model is first considered for which basic measures of queue length and delay are developed. This is a relatively simple model that does not take into account the variability of traffic (i.e. differing numbers of vehicles arrive at intersections at different times).

In estimating delay at intersections, traffic flow is considered as consisting of identical passenger car units (pcu's). A truck, for example, may be considered as 1.5 or 2 pcu's and a turning vehicle may also be assigned some value depending on the type of maneuver that is made.

For the discussion of this model, the following variables (defined below) will be used:
Signalized Intersections: Fluid Traffic Model – (cont'd)

Let

\[ c = \text{the cycle time (sec)} \]
\[ g = \text{the effective green time (sec)} \]
\[ r = \text{the effective red time (sec)} \]
\[ q = \text{the average arrival rate of traffic on the approach (pcu/sec)} \]
\[ s = \text{the saturation flow on the approach (pcu/sec)} \]
\[ d = \text{the average delay to a pcu on the approach (sec)} \]
\[ Q_0 = \text{the overflow (pcu)} \]
\[ \lambda = \frac{g}{c} \text{ (the proportion of the cycle that is effectively green)} \]
\[ y = \frac{q}{s} \text{ (the ratio of average arrival rate to saturation flow)} \]
\[ x = \frac{qc}{gs} \text{ (the ratio of average number of arrivals/cycle to the maximum number of departures/cycle)} \]

Thus, \( r + g = c \) and \( \lambda x = y \). The ratio \( x \) is called the degree of saturation of the approach and \( y \) is called the flow ratio of the approach.
Signalized Intersections: Fluid Traffic Model (cont'd)

First, let’s consider an equation to model the cumulative time that is required for a queue of $n$ stopped vehicles at an intersection to pass the signal once the green interval begins. Because of lost time due to starting, the first several vehicles take longer to pass the signal than do vehicles back in the queue.

According to a simple model of traffic flow (in this case, developed by Greenshields) the successive time intervals for vehicles departing from a queue to cross the signal stop line are: 3.8, 3.1, 2.7, 2.4, 2.2, 2.1, 2.1... seconds. Thus, once the fifth car in the queue arrives, all subsequent vehicles will arrive at 2.1 second intervals. The cumulative time for a queue of $n$ stopped vehicles to pass a signal can then be given by

$$Cumulative\ time = 14.2 + 2.1(n-5) \text{ sec for } n \geq 5$$
Signalized Intersections: Fluid Traffic Model (cont'd)

If the lost time due to starting had been ignored, and all of the vehicles departed at the saturation flow rate $s = 1/2.1[\text{veh/sec}] = 1714[\text{veh/hr}]$, the first five vehicles would have required only 10.5 seconds to clear the intersection. Therefore, since these vehicles actually required 14.2 seconds to clear, one will see that the effective green time is the signal green time less 3.7 seconds (i.e. $14.2 - 10.5$).
Before exploring equations to model delay, consider graphically arrivals and departures of vehicles at a signalized intersection. Delay can be calculated by illustrating the meaning of arrival time and departure time for a pcu on an approach. They are demonstrated by reference to Figure D-3, in which distance-time curves are plotted for each of four vehicles.
The line AB on the far left represents the passage of an un-delayed vehicle, where the horizontal line PQ represents the stop line at which the first vehicle waits when there is a queue. CDEF represents the trajectory of the first vehicle that is delayed by a signal. The straight portions of CD and EF are parallel to AB and projected to meet PQ and X and Y so that the length XY is the delay to the first vehicle.

In other words, if the first vehicle had not been delayed by the signal, it would have reached the stop line at time X. However, because it had to slow down/stop (the period between D and E), given the point it is at at E, it would have crossed the stop line at Y if there had been no delay. Therefore, the total delay for the first vehicle is XY. Similarly, X'Y' and X"Y" represent the delays for the next two vehicles.
Fluid Model for Pre-timed Signal

Another representation of a fluid model of traffic flow at a signalized intersection is shown in Figure D-4. The vertical axis represents the number of vehicle arrivals at the stop line and the horizontal axis represents the time $t$. The figure illustrates the behavior when the capacity of the green interval can fully accommodate (i.e. serve) the number of arrivals during the green + red time - i.e., an under-saturated period.

The top figure illustrates the queue length as a function of time ($t_0$ is the queue clearance time, measured from the start of green). The bottom figure illustrates the cumulative arrival and departure functions.
In Figure D-4 the vertical distance ca represents the number of vehicles that have accumulated since the signal entered the red phase at time $c$. At time $r$, the red interval is over, and the queue begins to clear. This is represented by the line forming the “right-hand side” of the triangle. Once the queue clearance line meets the queue formation line (the “left-hand side” of the triangle) the queue is completely cleared.

The horizontal distance $ab$ represents the total time from arrival to departure for any given vehicle. The area of the triangle in the top figure represents the total vehicle delay. This is identical to the shaded area in the bottom figure.
Fluid Model for Pre-timed Signal (cont'd)

The following measures of queue behavior are developed:

1. Time after start of green that queue is dissipated ($t_0$)

2. Proportion of cycle with queue ($P_q$)

3. Proportion of vehicles stopped ($P_s$)

4. Maximum number of vehicles in queue ($Q_m$)

5. Average number of vehicles in queue ($Q$)

6. Total vehicle-time of delay per cycle ($D$)

7. Average individual vehicle delay ($d$)

8. Maximum individual vehicular delay ($d_m$)
Fluid Model for Pre-timed Signal (cont'd)

On the next pages, we will present and describe a number of equations that can be used to analyze delay at an intersection based on the fluid model. The equations have been developed based on the graphs presented on the previous two pages. Students are encouraged to refer back to these graphs to better understand the foundation of the equations:

1. For any given cycle, as with any physical system, “what goes in, must come out.” In other words, once time $t_0$ (the time after green at which the queue is completely cleared) is reached, the total number of vehicles that arrived and waited in the queue must be equal to the total number of vehicles that have left the queue.
Fluid Model for Pre-timed Signal (cont'd)

In the equation below, remember that $q$ is the arrival flow rate, and $s$ is the saturation flow rate (i.e. capacity of the intersection). Thus, the variable $y$ is the ratio of the arrival rate to saturation flow rate. As this ratio becomes closer and closer to 1, the time to clear the queue, $t_0$, becomes larger and larger.

Equation D.1

The total number of arrivals is

$$q(r + t_0) = st_0$$

Letting $y = q/s$, we obtain

$$t_0 = yr/(1 - y)$$
Fluid Model for Pre-timed Signal (cont'd)

2. The proportion of vehicles delayed by the intersection control is equal to queue time (represented by the red interval plus $t_0$) divided by the total cycle length.

Equation D.2

$$P_q = (r + t_0)/c$$
3. The proportion of vehicles stopped is equal to vehicles stopped/total vehicles per cycle.

**Equation D.3**

\[ P_s = \frac{q(r + t_0)}{q(r + g)} = \frac{t_0}{yC} \]
Fluid Model for Pre-timed Signal (cont'd)

4. The maximum number of vehicles in queue can be seen by examining Figure D-4. As seen in the figure, it is the height of the triangle at $t = r$, i.e., at the end of the red period. This is logical in that the maximum queue results at the end of the time that vehicles (at the average arrival rate, $q$) have been arriving during the red phase:

Equation D.4

$$Q_m = qr$$
5. The average number of vehicles in the queue at any given time can be calculated by dividing the total number of vehicle-seconds of delay during the cycle length (represented by the area of the triangle in Figure D.4), divided by the total cycle length, $C \ (or \ r + g)$. Recall that the area of a triangle is half of its base length time its height – this results in the expression for total delay of $[(qr/2)r + (qr/2) t_0]$:

\[ Q = \frac{(qr/2)r + (qr/2)t_0}{r + g} \]

which yields

\[ Q = \left[ \frac{(r + t_0)}{c} \right] (qr / 2) \]
6. As described on the previous page, the total vehicle-time of delay during the cycle is equal to the area of the triangle in Figure D-4 (which, in terms of Calculus, is the integral of the queue-length curve). In equation D.6 below, we simply substitute the variable $Y$ (ratio of average arrival rate to saturation flow rate) to derive an equation that models total delay based on the average arrival rate, red interval, and $Y$.

*Equation D.6*

\[
D = \left( qr / 2 \right) ( r + t_0 ) = \left( qr / 2 \right) \left[ r / (1 - y) \right]
= qr^2 / [2(1 - y)]
\]
7. The average individual delay, \( d \), can be calculated by dividing the total vehicle-time delay (derived on the previous page) by the number of vehicles that arrive during the cycle, \( q_c \).

\[ d = \left[ \frac{qr^2}{2(1 - y)} \right] \frac{1}{q_c} = \frac{r^2}{2c(1 - y)} \]
Fluid Model for Pre-timed Signal (cont'd)

8. Finally, the maximum individual vehicular delay can be seen from Figure D-4 to be simply equal to the total red interval. This corresponds to the vehicle that just arrives at the intersection as the signal turns to red.

*Equation D.8*
If the possible departures during the cycle (i.e., the capacity of the intersection – the saturation flow rate times the total green interval - $sg$) are less than the total number of arriving vehicles during the entire cycle, $qc$, the queue grows with each successive cycle.

In other words, the queue that formed during the red interval is not cleared during the green interval – and some vehicles find themselves delayed by more than one cycle of the intersection. This situation is described as oversaturation. In the case of oversaturation – the previous equations are not applicable and cannot be used for analysis.
Fluid Model for Pre-timed Signal (cont'd)

Example:
Consider the following example of the behavior of a queue at a signalized intersection. Assume that the green phase $g$ is 40 sec, the red phase $r$ is 20 sec, the discharge rate $s$ is 1,200 vehicles/hr, and there are two input flow rates $q_1 = 600$ vph; $q_2 = 800$ vph. In the first case, the capacity of the green interval exceeds the number of arrivals during the green + red time (the undersaturated case). In the second case, the discharge during the green phase is equal to the arrivals during the green + red period (the saturated case). The results are given in Table D-1 (below).

<table>
<thead>
<tr>
<th>Queueing Characteristics at Fixed-Time Signalized Intersection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queueing Characteristic</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>$t_0$</td>
</tr>
<tr>
<td>$P_g$</td>
</tr>
<tr>
<td>$P_s$</td>
</tr>
<tr>
<td>$Q_m$</td>
</tr>
<tr>
<td>$Q$</td>
</tr>
<tr>
<td>$D$</td>
</tr>
<tr>
<td>$d$</td>
</tr>
<tr>
<td>$d_m$</td>
</tr>
</tbody>
</table>
Average Delay Per Vehicle – WEBSTER Model

As stated before, the equations developed and presented up to this point are only applicable for undersaturated deterministic conditions; i.e., when volume is below capacity and the arrival pattern is known apriorily. Clearly, in many situations, this is not appropriate for use in signal analysis.

Webster developed a more complex model of average delay per vehicle that assumes random arrival of vehicles (in this case, based on the Poisson distribution). Although the v/c ratio still must be less than 1.0 for the entire analysis period, there are possible overflows due to the random nature of the arrivals. We will explore this model in more depth on the following pages.
Average Delay Per Vehicle – WEBSTER Model – (cont'd)

\[ d = \frac{c(l - \lambda)^2}{2(l - \lambda)} + \frac{x^2}{2q(l - x)} - 0.65 \left( \frac{c}{q^2} \right)^{\frac{1}{3}} \lambda^{(2+5x)} \]

Where:

- \( d \) = average delay per vehicle on the particular arm of the intersection
- \( c \) = cycle time
- \( \lambda \) = proportion of the cycle which is effectively green for the phase under consideration (i.e. \( g / c \))
- \( q \) = flow
- \( s \) = saturation flow
- \( x \) = the degree of saturation

This is the ratio of the actual flow to the maximum flow which can be passed through the intersection from this arm, and is given by:

\[ x = q / \lambda \]

\[ s = q / K \]

where

\[ K = \lambda \text{ is the capacity flow} \]

(If \( d \) and \( c \) are in seconds, \( q \), \( s \), and \( K \) are in vehicles per second.)
To simplify the equation above, we rewrite the equation as a function of the cycle length, \( c \), and the arrival flow rate, \( q \). Then, the equation uses two terms – \( A \) and \( B \) – these are functions of the green time ratio and the ratio of maximum flow to capacity. Remember, equation D.10 is exactly the same as D.9 presented above – it is just written in a more simple manner.

*Equation D.10*

\[
d = cA + \frac{B}{q} - C
\]

where

\[
A = \frac{(1 - \lambda)^2}{2(1 - \lambda x)} \quad B = \frac{x^2}{2(1 - x)} \quad \text{and } C \text{ is the correction term.}
\]
Average Delay Per Vehicle – WEBSTER Model – (cont'd)

The first term of the delay equation – $cA$ – models delay when vehicles are arriving at a uniform rate (i.e. no randomness in the arrival of vehicles). In fact, this term is identical to the simpler model presented in Equation D.7. At low flow arrival flow rates, this first term does a very good job of modeling delay. However, at higher flow rates when randomness becomes a larger factor, the term alone cannot effectively model delay.
The second term of Equation D.9 is often referred to as the “overflow delay” component of the model. It makes some allowance for the random nature of the arrivals. It is an expression for the delay experienced by vehicles arriving randomly in time at a 'bottleneck', queuing up, and leaving at constant intervals. In queuing theory terminology it is the average delay of a M/D/1 queue. The addition of the second term improves the effectiveness of the delay model – but often overestimates delay. To avoid the need to use an empirical correction factor $C$, a suitable approximation can be obtained by reducing the delay by a percent $P$ in one of the following ways:

Equation D.11

$$d = \left[ cA + \frac{B}{q} \right] \frac{100 - P}{100}$$

where (a) $P = 10$, or (b) $P = 15x$, since the value of the correction term $C$ is generally in the range 5 to 15 percent of $d$ [10]
Average Delay Per Vehicle – WEBSTER Model: Example

The next pages present a numerical example of applying the delay model. In this example, the average flow on a particular approach of an intersection is 600 vehicles per hour, and the signal settings are 30 seconds green, 4 seconds yellow and 60 seconds cycle time. It is observed that on the average 15.0 vehicles are discharged in a fully saturated green period.

If we assume that starting delays, etc. are responsible for 2 seconds of each green-plus-amber period then 15 vehicles are discharged in an effective green time of 30 seconds, i.e. $s = 1800$ vehicles per hour. It is assumed that 2 seconds of the yellow time are also included as part of the effective green. Based on this given information, we can compute the green interval portion of the cycle time ($\lambda$), the capacity flow ($K$) and the degree of saturation ($\chi$).
Average Delay Per Vehicle – WEBSTER Model: Example – (cont'd)

\[ \lambda = \frac{g}{c} = \frac{30}{60} = 0.5 \]

\[ K = \lambda s = 0.5(1800) = 900 \text{ vph} \]

\[ x = \frac{q}{K} = \frac{600}{900} = 0.667 \]
Average Delay Per Vehicle – WEBSTER Model: Example – (cont'd)

From the delay model equations, we can calculate that $A = 0.187$ and $B = 0.667$. Finally, assume that in this case, $C = 9$ percent of the total delay as estimated by first two terms. Using *Equation D.10*, we estimate average vehicular delay to be:

$$d = 60 \times 0.187 + \frac{0.667}{600/3600} - C = 11.2 + 4.0 - C = 15.2 - 1.4 = 13.8 \text{ sec}$$
Alternatively, we can use the approximations according to Equation D.11:

(a) \[ d = \left[ 11.2 + 4.0 \right] \left( \frac{9}{10} \right) = (15.2)0.90 = 13.68 \text{ sec.} \]

(b) \[ d = 15.2 \left[ \frac{100 - 15(0.667)}{100} \right] = 13.68 \text{ sec.} \]
E. Traffic Signal Settings

Prof. Nathan H. Gartner
University of Massachusetts
Lowell, MA, USA

© NHGartner, UML, July 2013
The objective in setting signal timings for a fixed-time signal is to minimize overall vehicular delay. This is in addition to the need to meet the warrants described earlier.
Green Times Determination

The *ITE Transportation and Traffic Engineering Handbook [1]* states that the least delay to traffic at an intersection is obtained when the green periods of the phases are set in proportion to the corresponding ratios of flow to saturation flow, on the dominant (or critical) approaches.

This division of the cycle time makes the capacity of the phases proportional to the average flows of the phases.
Green Times Determination
- Example

Typical results are shown in Fig. D-6, where the best ratio of the $y$ values is approximately 2.0. It can be seen that the best ratios of the effective green times is between 1.88 and 2.17 over a range of cycle times of 35 to 80 seconds.
Optimum Cycle Time

The challenge in this case is to determine the value of cycle time which gives the least delay of all traffic using the intersection.

Webster developed a relatively simple expression to determine optimal cycle length – based on total lost time and the sum of the \( y \) values at all approaches (recall that \( y \) is the ratio of the arrival flow rate to the saturation flow rate at an approach):

\[
c_o = \frac{1.5L + 5}{1 - Y} \text{ sec}
\]

where \( Y \) is the sum of the \( y \) values for all phases and refers to the intersection as a whole and \( L \) is the total lost time per cycle in seconds.
Cycle Time (cont'd)

The lost time value in the optimal cycle length equation can be expressed by

\[ L = nl + R \]

where:

\( n \) is the number of phases

\( l \) is the average lost time per phase (excluding any all-red periods)

\( R \) is the total time during each cycle when all signals display red simultaneously.

Figure D-7. Possible phase aspect scheduling at a two-phase intersection
Cycle Time (cont'd)

Figure D-8. Effect on delay of variation of the cycle length

2-PHASE, 4-ARM INTERSECTION
Equal flows on all arms
Equal saturation flows: 1800 vph
Equal green times
Total lost time per cycle: 10 seconds

Average delay per vehicle (seconds)

Total flow entering intersection (vehicles per hour)
3000
2800
2400
1600

Cycle Time (seconds)
0 20 40 60 80 100 120 140 160 180

$\frac{3}{4}c_0$, $c_0$, $1\frac{1}{2}c_0$, $1\frac{1}{2}c_0$, $3\frac{1}{4}c_0$, $c_0$, $3\frac{1}{4}c_0$
When traffic is of a truly random character, the minimum cycle time will result in nearly infinite delay.

The optimal cycle length (i.e. the cycle length that results in the minimum amount of total delay at the intersection) is roughly two times the minimum cycle length, therefore:

\[ c_{opt} \approx 2c_{min} = \frac{2L}{1-Y} \]
Guidelines for Selection of Cycle Time for Fixed-Time Signals

A challenge in designing signal settings for a fixed-time signal lies in the fact that, in most cases, arrival volumes vary throughout the day. For example, traffic is often heavier during peak periods than mid-day and at night. Remember that the approach we have discussed so far relies on developing settings for particular arrival volume levels. Ideally, the signal will have different settings that are designed for different times of day (with the corresponding different volume levels). In some cases, vehicle-actuated signals can also be used to account for this.

However, when these approaches are not feasible, another compromise is possible.
Guidelines for Selection of Cycle Time for Fixed-Time Signals – (cont’d)

It can be seen that the delay for cycle times within the range of ¾ to 1½ times the optimum cycle length value is never more than 10 to 20 percent greater than delay for the optimum cycle. Based on this observation, for a single setting of fixed-time signals the simple approximate method outlined below may be used.

i. Calculate the optimum cycle for each hour of the day when the traffic flow is medium or heavy, e.g. between the hours of 8 a.m. and 7 p.m. and average over the day.

ii. Evaluate three-quarters (¾) of the optimum cycle calculated for the heaviest peak hour.

iii. Select whichever is greater for the cycle time.
Guidelines for Selection of Green Times for Fixed-Time Signals –

It is suggested as a reasonable procedure that the division of the available green time during the cycle \((c_0 - L)\) should be in proportion to the average \(y\) values for peak periods.

\[
\frac{g_1}{g_2} = \frac{(y_1)_{PEAK}}{(y_2)_{PEAK}}
\]
Signal Settings: Example

Example: determine settings for a 2-phase, 4-arm intersection. First, the following data is collected in the field:

<table>
<thead>
<tr>
<th></th>
<th>North</th>
<th>South</th>
<th>East</th>
<th>West</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow (q)</td>
<td>600</td>
<td>450</td>
<td>900</td>
<td>750</td>
</tr>
<tr>
<td>Saturation flow (s)</td>
<td>2400</td>
<td>2000</td>
<td>3000</td>
<td>3000</td>
</tr>
<tr>
<td>Ratio (q/s)</td>
<td>0.25</td>
<td>0.225</td>
<td>0.3</td>
<td>0.25</td>
</tr>
<tr>
<td>Y values</td>
<td>0.25</td>
<td></td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Lost time</td>
<td>3 seconds per phase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Starting delays</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All-red periods</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3 seconds at each change of right of way
Signal Settings: Example (cont’d)

Assuming a 4-second yellow interval, the total lost time per phase is $3 + 4(0.5) = 5$ sec. For the two phases it is 10 seconds (see the figure below).

Figure D-7. Possible phase aspect scheduling at a two-phase intersection
Signal Settings: Example (cont’d)

The total lost time for the cycle, \( L \), including all-red periods, is therefore 16 seconds.

The optimal cycle length for the intersection is calculated as follows:

\[
c_0 = \frac{1.5(16) + 5}{1 - 0.250 - 0.300} = \frac{29}{0.450} = 64 \text{ sec}
\]
Signal Settings: Example (cont’d)

Now, given the total cycle time of 64 seconds, it is necessary to set green intervals for both phases. The $y$ value for the North-South phase is 0.25, while it is 0.3 for the East-West phase. We remove the all-red time periods and yellow periods from the total cycle length to determine “available” total green: 64 seconds – 6 seconds – 8 seconds = 50 seconds.

Therefore:

\[
\text{Green Interval for North-South: } 50 \text{ seconds} \left( \frac{0.25}{0.25 + 0.3} \right) = 23 \text{ seconds}
\]

\[
\text{Green Interval for East-West: } 50 \text{ seconds} \left( \frac{0.30}{0.25 + 0.3} \right) = 27 \text{ seconds}
\]