# SIGNAL TIMING MANUAL

Prepared for the **Federal Highway Administration**
by **Kittelson & Associates, Inc.**
with the help of **Texas Transportation Institute,** the **University of Maryland, Tom Urbanik**
and with the technical assistance of **Siemens ITS, Purdue University,** & the **Institute of Transportation Engineers**

# TRAFFIC SIGNAL TIMING MANUAL

**Publication Number: FHWA-HOP-08-024**

**Task Order under contract number: DTFH61-98-C-00075**
**Task Order No. B98C75-009**

**JUNE 2008**

# FOREWORD

The purpose of the *Signal Timing Manual* is to provide direction and guidance to managers, supervisors, and practitioners based on sound practice to proactively and comprehensively improve signal timing. The outcome of properly training staff and proactively operating and maintaining traffic signals is signal timing that reduces congestion and fuel consumption ultimately improving our quality of life and the air we breathe.

This manual provides an easy-to-use concise, practical and modular guide on signal timing. The elements of signal timing from policy and funding considerations to timing plan development, assessment, and maintenance are covered in the manual. The manual is the culmination of research into practices across North America and serves as a reference for a range of practitioners, from those involved in the day to day management, operation and maintenance of traffic signals to those that plan, design, operate and maintain these systems.

Regina McElroy
Director
Office of Transportation Management


## NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for the use of the information contained in this document.

This report does not constitute a standard, specification, or regulation. The U.S. Government does not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document.

**Quality Assurance Statement**
The Federal Highway Administration (FHWA) provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. FHWA periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

| 1. Report No.<br>FHWA-HOP-08-024 | 2. Government Accession No. | 3. Recipient's Catalog No. | |
|---|---|---|---|
| 4. Title and Subtitle<br>Traffic Signal Timing Manual | | 5. Report Date<br>June 2008<br>6. Performing Organization Code | |
| 7. Author(s)<br>Principal Investigator: Peter Koonce<br>Co-Authors: Lee Rodegerdts, Kevin Lee, Shaun Quayle, Scott Beaird, Cade Braud, Jim Bonneson, Phil Tarnoff, and Tom Urbanik | | 8. Performing Organization Report No.<br>Project 7257 | |
| 9. Performing Organization Name and Address<br>Kittelson & Associates, Inc.<br>610 SW Alder Street, Suite 700<br>Portland, OR 97205<br>Subconsultants: Texas Transportation Institute, University of Maryland, Purdue University, Siemens ITS, Tom Urbanik II, PE | | 10. Work Unit No. (TRAIS)<br>11. Contract or Grant No.<br>Contract No. DTFH61-98-C-00075,<br>Task Order No. B98C75-009 | |
| 12. Sponsoring Agency Name and Address<br>Federal Highway Administration<br>Turner-Fairbank Highway Research Center<br>6300 Georgetown Pike, HSR 20, Room No. T301<br>McLean, VA 22101 | | 13. Type of Report and Period Covered<br>Final Report<br>March 2006 to June 2008<br>14. Sponsoring Agency Code<br>HOP | |
| 15. Supplementary Notes<br>Eddie Curtis (Eddie.Curtis@fhwa.dot.gov) was the Technical Representative for the Federal Highway Administration. Project focus group members contributed significantly to document organization, content, and exhibits: Michael Kyte, University of Idaho; Ziad Sabra, Sabra, Wang & Associates; James Pond, Monroe County (NY); Jim Poston, Regional Transportation Commission (Reno, NV), Rick Denney, Iteris; Bruce Zvaniga, City of Toronto; Woody Hood, Maryland State Highway Administration; Bill Kloos, City of Portland (OR); Denny Eyler, SRF Consulting; Avery Rhoades, City of Glendale (AZ); Sean Skehan, Los Angeles Department of Transportation (CA); and Susan Langdon, Savant Group. In addition, many staff members of FHWA participated as focus group members and/or provided comments throughout the project, including Ed Fok, John Halkias, Paul Olson, Mike Schauer, Fred Ranck, and Scott Wainwright. | | | |
| 16. Abstract<br>This report serves as a comprehensive guide to traffic signal timing and documents the tasks completed in association with its development. The focus of this document is on traffic signal control principles, practices, and procedures. It describes the relationship between traffic signal timing and transportation policy and addresses maintenance and operations of traffic signals. It represents a synthesis of traffic signal timing concepts and their application and focuses on the use of detection, related timing parameters, and resulting effects to users at the intersection. It discusses advanced topics briefly to raise awareness related to their use and application. | | | |
| 17. Key Words<br>Signalized Intersections, Traffic Signal Timing, Detectors, Policy Based Signal Control, Maintenance and Operation of Traffic Signals | | 18. Distribution Statement<br>No Restrictions. This document is available to the public | |
| 19. Security Classif. (of this report)<br>Unclassified | 20. Security Classif. (of this page)<br>Unclassified | 21. No. of Pages<br>265 | 22. Price<br>N/A |

## TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# GLOSSARY

## TABLE OF CONTENTS

# GLOSSARY

The following section is a collection of terms used in the manual. Sources and references are identified for those terms with multiple definitions.

### Accident modification factors
A means of quantifying crash reductions associated with safety improvements.

### Actuated Signal Control
Phase time based on detection data.

### Adaptive Signal Control
A signal control concept where vehicular traffic in a network is detected at a point upstream and/or downstream and an algorithm is used to predict when and where traffic will be and to make signal adjustments at downstream intersections based on those predictions.

### Added Initial
An interval that times concurrently with the minimum green interval and increases by each vehicle actuation received during the initial period. This time cannot exceed the maximum initial.

### Analysis period
A single time period during which capacity analysis is performed on a transportation facility. If the demand exceeds capacity during an analysis period, consecutive analysis periods can be selected to account for initial queue from the previous analysis period. Also referred to as time interval.

### Analytical Model
A model that relates system components using theoretical considerations tempered, validated, and calibrated by field data.

### Annual average daily traffic
The total volume of traffic passing a point or segment of a highway facility in both directions for one year divided by the number of days in the year.

### Approach
A set of lanes at an intersection that accommodates all left-turn, through, and right-turn movements from a given direction.

### Approach grade
The grade of an intersection approach, expressed as a percentage, with positive values for upgrade and negative for downgrade.

### Area type
A geographic parameter reflecting the variation of saturation flows in different areas.

### Arrival rate
The mean of a statistical distribution of vehicles arriving at a point or uniform segment of a lane or roadway.

### Arrival type
Six assigned categories for determining the quality of progression at a signalized intersection.

### Arterial
A signalized street that primarily serves through traffic and that secondarily provides access to abutting properties, with signal spacings of 2.0 miles or less.

**Arterial LOS**
An arterial- and network-level performance measure associated with the class of arterial and the travel speed of arterial under study.

**Automatic Vehicle Location (AVL) System**
An intelligent transportation system (ITS) technology to track vehicle location, speed and other measures within a system. Most applications are found on transit vehicles and systems.

**Average Speed**
The average distance a vehicle travels within a measured amount of time.

**Average travel speed**
The length of the highway segment divided by the average travel time of all vehicles traversing the segment, including all stopped delay times.

**Back of queue**
The distance between the stop line of a signalized intersection and the farthest reach of an upstream queue, expressed as a number of vehicles. The vehicles previously stopped at the front of the queue are counted even if they begin moving.

**Bandwidth**
The maximum amount of green time for a designated direction as it passes through a corridor at an assumed constant speed, typically measured in seconds.

**Bandwidth attainability**
A measure of how well the bandwidth makes use of the available green time for the coordinated movements at the most critical intersection in the corridor.

**Bandwidth efficiency**
A measure that normalizes bandwidth against the cycle length for the arterial under study.

**Barnes' Dance**
A common term for an exclusive pedestrian phase where pedestrians may cross all intersections legs and sometimes diagonally.

**Barrier**
A separation of intersecting movements in separate rings to prevent operating conflicting phases at the same time.

**Base condition**
The best possible characteristic in terms of capacity for a given type of transportation facility; that is, further improvements would not increase capacity; a condition without hindrances or delays.

**Base saturation flow rate**
The maximum steady flow rate—expressed in passenger cars per hour per lane—at which previously stopped passenger cars can cross the stop line of a signalized intersection under base conditions, assuming that the green signal is available and no lost times are experienced.

**Call**
A term used to describe the presence of vehicle, bicycle, or pedestrian demand in an actuated detection controller system.

**Capacity**
The maximum rate at which vehicles can pass through the intersection under prevailing conditions. It is also the ratio of time during which vehicles may enter the intersection.

**Carryover**
A term commonly used for the “extend” setting in controller manuals. It is another way to describe the time provided for a vehicle to traverse from one detector to the next.

**Change interval**
The yellow plus red clearance interval that occurs between phases of a traffic signal to provide for clearance of the intersection before conflicting movements are released. Also known as the clearance interval.

**Clearance lost time**
The time, in seconds, between signal phases during which an intersection is not used by any traffic.

**Clearance time**
The time loss at a transit stop, not including passenger dwell times. This parameter can be the minimum time between one transit vehicle leaving a stop and the following vehicle entering and can include any delay waiting for a sufficient gap in traffic to allow the transit vehicle to reenter the travel lane.

**Condition Diagram**
An illustration used to highlight the existing characteristics (i.e., number of lanes, signs, adjacent driveways, turn-bay lengths, traffic control, and land uses) of an intersection.

**Concurrent Phases**
Two or more phases in separate rings that are able to operate together without conflicting movements.

**Congested flow**
A traffic flow condition caused by a downstream bottleneck.

**Control Delay**
The amount of additional travel time experienced by a user attributable to a control device.

**Controller Memory**
A term that refers to the controller’s ability to “remember” (i.e., retain) a detector actuation and includes one of two modes (nonlocking or locking).

**Coordinated-Actuated**
Signal operations in coordination with other intersections, and using vehicle, bicycle, and/or pedestrian detection to define signal timing.

**Coordinated Phase(s)**
The phase (or phases) that is provided a fixed minimum amount of time each cycle under a coordinated timing plan. This phase is typically the major through phase on an arterial.

**Coordination**
The ability to synchronize multiple intersections to enhance the operation of one or more directional movements in a system.

**Corridor**
A set of essentially parallel transportation facilities designed for travel between two points. A corridor contains several subsystems, such as freeways, rural (or two-lane) highways, arterials, transit, and pedestrian and bicycle facilities.

**Critical lane group**
The lane groups that have the highest flow ratio for a given signal phase.

**Critical movement analysis**
A simplified technique for estimating phasing needs and signal timing parameters.

### Critical speed
The speed at which capacity occurs for a facility, usually expressed as miles per hour.

### Critical volume-to-capacity ratio
The proportion of available intersection capacity used by vehicles in critical lane groups.

### Crosswalk
A marked area for pedestrians crossing the street at an intersection or designated midblock location.

### Cycle
A complete sequence of signal indications.

### Cycle Length
The time required for a complete sequence of signal indications.

### Cycle Failure
Occasion where all queued vehicular demand cannot be served by a single green indication or signal phase.

### Dallas Display
A type of signal display that attempts to avoid "yellow trap" problem by using louvers on the yellow and green ball indications to restrict visibility of the left-turn display to adjacent lanes while displaying indications based on the opposing through movement.

### Delay
1. The additional travel time experienced by a driver, passenger, or pedestrian.
2. A detector parameter typically used with stop-line, presence mode detection for turn movements from exclusive lanes

### Density
The number of vehicles on a roadway segment averaged over space, usually expressed as vehicles per mile or vehicles per mile per lane. (see also: volume-density, sometimes referred to as density timing)

### Demand
The volume of traffic at an intersection, approach, or movement.

### Detector
A device used to count and/or determine the presence of a vehicle, bicycle, or pedestrian.

### Dilemma Zone
There are two types of dilemma zones.

Type I occurs when yellow and red clearance times are too short for a driver to either stop or clear the intersection before the beginning of a conflicting phase.

Type II, also known as an "Option Zone", ore "Indecision Zone".  This occurs as the result of different drivers making different decision on whether to go or stop, upon the change from a green to yellow indication.

### Double Cycle
A cycle length that allows phases to be serviced twice as often as the other intersections in the coordinated system.

### Downstream
The direction of traffic flow.

**Early Return to Green**

A term used to describe the servicing of a coordinated phase in advance of its programmed begin time as a result of unused time from non-coordinated phases.

**Effective green time**

The time during which a given traffic movement or set of movements may proceed; it is equal to the cycle length minus the effective red time.

**Effective red time**

The time during which a given traffic movement or set of movements is directed to stop; it is equal to the cycle length minus the effective green time.

**Effective walkway width**

The width, in feet, of a walkway usable by pedestrians, or the total walkway width minus the width of unusable buffer zones along the curb and building line.

**Exclusive pedestrian phase**

An additional phase that is configured such that no vehicular movements are served concurrently with pedestrian traffic. See also, Barnes Dance.

**Exclusive turn lane**

A designated left- or right-turn lane or lanes used only by vehicles making those turns.

**Extend**

A detector parameter that extends a detector actuation by a setable fixed amount. It is typically used with detection designs that combine multiple advance detectors and stop-line detection for safe phase termination of high-speed intersection approaches.

**Field Implementation**

A term used to describe the installation of new signal timings in the controller and the review of traffic operations at the intersection.

**Fixed Force Off**

A force off mode where force off points cannot move. Under this mode, non-coordinated phases can utilize unused time of previous phases.

**Fixed Time Signal Control**

A preset time is given to each movement every cycle regardless of changes in traffic conditions.

**Flashing Don't Walk**

An indication warning pedestrians that the walk indication has ended and the don't walk indication will begin at the end of the pedestrian clearance interval.

**Flashing Yellow Arrow**

A type of signal head display that attempts to avoid the "yellow trap" problem by providing a permissive indication to the driver that operates concurrent with the opposing through movement rather than the adjacent through movement.

**Floating Force Off**

A force off mode where force off points can move depending on the demand of previous phases. Under this mode, non-coordinated phase times are limited to their defined split amount of time and all unused time is dedicated to the coordinated phase. Essentially, the split time is treated as a maximum amount for the non-coordinated phases.

**Floating car method**

A commonly employed technique for travel time runs which requires the vehicle driver to "float" with the traffic stream while traveling at a speed that is representative of the other vehicles on the roadway and to pass as many vehicles as pass the floating car.

### Flow rate

The equivalent hourly rate at which vehicles, bicycles, or persons pass a point on a lane, roadway, or other trafficway; computed as the number of vehicles, bicycles, or persons passing the point, divided by the time interval (usually less than 1 h) in which they pass; expressed as vehicles, bicycles, or persons per hour.

### Flow ratio

The ratio of the actual flow rate to the saturation flow rate for a lane group at an intersection.

### Force Off

A point within a cycle where a phase must end regardless of continued demand. These points in a coordinated cycle ensure that the coordinated phase returns in time to maintain its designated offset.

### Free flow

A flow of traffic unaffected by upstream or downstream conditions.

### Fully actuated control

A signal operation in which vehicle detectors at each approach to the intersection control the occurrence and length of every phase.

### Gap

The time, in seconds, for the front bumper of the second of two successive vehicles to reach the starting point of the front bumper of the first.

### Gap Reduction

This is a feature that reduces the passage time to a smaller value while the phase is active.

### Green time

The duration, in seconds, of the green indication for a given movement at a signalized intersection.

### Green time ratio

The ratio of the effective green time of a phase to the cycle length.

### Green Extension

A signal priority treatment to extend a current green phase to give priority to a specific movement or vehicle, typically transit.

### Hardware

The devices that physically operate the signal timing controls, including the controller, detectors, signal heads, and conflict monitor.

### Headway

(1) The time, in seconds, between two successive vehicles as they pass a point on the roadway, measured from the same common feature of both vehicles (for example, the front axle or the front bumper);

(2) The time, usually expressed in minutes, between the passing of the front ends of successive transit units (vehicles or trains) moving along the same lane or track (or other guideway) in the same direction.

### Hardware in the Loop (HITL)

A means of providing a direct linkage between simulation models and actual signal controllers.

### Highway Capacity Manual
A National Academies of Science/Transportation Research Board manual containing a collection of state-of-the-art techniques for estimating the capacity and determining the level-of-service for transportation facilities, including intersections and roadways as well as facilities for transit, bicycles, and pedestrians.

### Inhibit Max
A basic timing parameter that removes the Maximum Green input as a phase parameter during coordination and allows the phase to extend beyond its normal maximum green values.

### Interval
The duration of time where a traffic signal indications do not change state (red, yellow, green, flashing don't walk). A traffic signal controller also has timing intervals (min green, passage time) that determine the length of the green interval.

### Intersection Delay - Average
The total additional travel time experienced by users as a result of control measures and interactions with other users divided by the volume departing from the intersection.

### Intersection Level of Service
A qualitative measure describing operational conditions based on average intersection delay.

### Isolated intersection
An intersection at least one mile from the nearest upstream signalized intersection.

### Lagging pedestrian interval
A pedestrian timing option that starts pedestrian walk interval several seconds after the adjacent through movement phase, thus allowing a waiting right-turn queue to clear before the pedestrian walk indication is presented and thereby reducing conflicts with right-turning vehicles.

### Lane group
A set of lanes established at an intersection approach for separate capacity and level-of-service analysis.

### Lane group delay
The control delay for a given lane group.

### Lane utilization
The distribution of vehicles among lanes when two or more lanes are available for a movement; however, as demand approaches capacity, uniform lane utilization develops.

### Leading pedestrian interval
A pedestrian interval option that starts a few seconds before the adjacent through movement phase, thus allowing pedestrians to establish a presence in the crosswalk and thereby reducing conflicts with turning vehicles.

### Lead-Lag Left-Turn Phasing
A left-turn phase sequence where one left-turn movement begins with the adjacent through movement and the opposing left-turn movement begins at the end of the conflicting through movement. This option may create a "yellow trap" with some permissive signal displays.

### Level of service
A qualitative measure describing operational conditions within a traffic stream, based on service measures such as speed and travel time, freedom to maneuver, traffic interruptions, comfort, and convenience.

**Local Controller**
The device used to operate and control the signal displays using signal timing provided by the user, master controller, or central signal system.

**Locking mode**
A controller memory mode used to trigger a call for service for the first actuation received by the controller on a specified channel during the red interval

**Lost Time**
The portion of time at the beginning of each green period and a portion of each yellow change plus red clearance period that is not usable by vehicles.

**Master Clock**
The background timing mechanism within the controller logic to which each controller is .referenced during coordinated operations.

**Master Controller**
An optional component of a signal system that facilitates coordination of a signal system with the local controller.

**Manual on Traffic Control Devices (MUTCD)**
The MUTCD, published by the Federal Highway Administration, provides the standards and guidance for installation and maintenance for traffic control devices on roadways.

**Maximum Allowable Headway (MAH) / Maximum Time Separation**
The maximum time separation between vehicle calls on an approach without gapping out the phase, typically defined by passage time or gap time. Maximum allowable headway refers to spacing between common points of vehicles in a single lane, but the term is commonly used to refer to maximum time separation in single or multi-lane approaches as well.

**Maximum Green**
The maximum length of time that a phase can be green in the presence of a conflicting call.

**Maximum Initial**
The maximum period of time for which the Added Initial can extend the initial green period. This cannot be less than the Minimum Green time.

**Maximum Recall**
A recall mode that places a continuous call on a phase.

**Measure of effectiveness**
A quantitative parameter indicating the performance of a transportation facility or service.

**Minimum Gap**
This volume density parameter that specifies the minimum green extension when gap reduction is used..

**Minimum Green**
The first timed portion of the green interval which may be set in consideration driver expectany and the storage of vehicles between the detectors and the stop line when volume density or presence detection is not used.

**Minimum Recall**
A recall parameter the phase is timed for its minimum green time regardless what the demand is for the movement.

### Movement

A term used to describe the user type (vehicle or pedestrian) and action (turning movement) taken at an intersection. Two different types of movements include those that have the right of way and those that must yield consistent with the rules of the road or the Uniform Vehicle Code.

### Non-locking mode

A controller memory mode that does not retain an actuation received from a detector by the controller after the actuation is dropped by the detection unit.

### Occupancy

The percent of time that a detector indicates a vehicle is present over a total time period.

### Offset

The time relationship between coordinated phases defined reference point and a defined master reference (master clock or sync pulse).

### Offset Reference Point (Coordination Point)

The defined point that creates an association between a signalized intersection and the master clock.

### Overflow queue

Queued vehicles left over from a green phase at a signalized intersection.

### Oversaturation

A traffic condition in which the arrival flow rate exceeds capacity.

### Passage Time (Vehicle Interval, Gap, Passage Gap, Unit Extension)

A phase timer that ends a phase when the time from the last detector output exceeds the timer setting.

### Pattern Sync Reference

The set start of the master clock.

### Peak-hour factor

The hourly volume during the maximum-volume hour of the day divided by four times the peak 15-min flow rate within the peak hour; a measure of traffic demand fluctuation within the peak hour.

### Pedestrian

An individual traveling on foot.

### Pedestrian Recall

A recall mode where there is a continuous call for pedestrian service resulting in the pedestrian walk and clearance phases to occur each time the phase times.

### Pedestrian Clearance Interval

Also known as “Flash Don’t Walk”. The time provided for a pedestrian to cross the entire width of the intersection.

### Pedestrian Phase

Time allocated to pedestrian traffic that may be concurrent with vehicular phases.

### Pedestrian scramble

See Exclusive Pedestrian Phase

### Pedestrian Walk Interval

An indication to the pedestrian that it allows pedestrians to begin crossing the intersection.

**Pedestrian walking speed**
The average walking speed of pedestrians, in feet per second.

**Percent Runs Stopped**
The percentage of the total number of travel time runs conducted during which a vehicle stops.

**Performance Index**
An arterial- and network-Level performance measure that allows several measures of effectiveness to be mathematically combined.

**Performance Measures**
Signal system related effects on stops, vehicle delay, arterial travel time, or existence of spill back queuing between closely spaced intersections.

**Permissive Movements**
A movement where it is allowed to proceed if there are available gaps in the conflicting flow.

**Permissive Period**
A period of time during the coordinated cycle in which calls on conflicting phases will be result in the coordinated phase transitioning to non-coordinated phase(s)..

**Permitted plus protected**
Compound left-turn protection that displays the permitted phase before the protected phase.

**Permitted turn**
Left or right turn at a signalized intersection that is made against an opposing or conflicting vehicular or pedestrian flow.

**Phase**
A controller timing unit associated with the control of one or more movements. *The MUTCD defines a phase as the right-of-way, yellow change, and red clearance intervals in a cycle that are assigned to an independent traffic movement.*

**Phasing Indication**
The current display for a given phase (green, yellow, red, walk, flashing don't walk, or don't walk).

**Phase Pair**
A combination of two phases allowed within the same ring and between the same barriers such as 1+2. 5+6, 3+4, and 7+8.

**Phase Recall**
A call is placed for a specified phase each time the controller is servicing a conflicting phase. This will ensure that the specified phase will be serviced again. Types of recall include soft, minimum, and maximum. Soft recall only calls the phase back if there is an absence of conflicting calls.

**Phase Sequence**
The order of a series of phases.

**Phasing Diagram**
A graphical representation of a sequence of phases.

**Platoon**
A group of vehicles or pedestrians traveling together as a group, either voluntarily or involuntarily because of signal control, geometrics, or other factors.

### Preemption

Traffic signal preemption is the transfer of normal operation of a traffic control signal to a special control mode of operation.

### Preempt Trap

A condition that can occur when a preemption call is serviced at a signalized intersection near an at-grade train-roadway crossing, where not enough clearance green time is provided to clear a queue of vehicles, and a vehicle could be trapped on the tracks with the railroad crossing lights and gates come down.

### Presence Mode

A detection mode where a signal is sent to the controller for the duration of time a vehicle is inside the detection zone.

### Pretimed control

A signal control in which the cycle length, phase plan, and phase times are preset to repeat continuously.

### Priority

Traffic signal priority (TSP) is an operational strategy communicated between transit vehicles and traffic signals to alter signal timing for the benefit or priority of transit vehicle. Green extension, red truncation, and phase skipping are examples of signal timing alterations under TSP.

### Progression adjustment factor

A factor used to account for the effect of signal progression on traffic flow; applied only to uniform delay.

### Protected Movements

A movement where it has the right-of-way and there are no conflicting movements occurring.

### Protected plus permitted

Compound left-turn protection at a signalized intersection that displays the protected phase before the permitted phase.

### Protected turn

The left or right turns at a signalized intersection that are made with no opposing or conflicting vehicular or pedestrian flow allowed.

### Pulse Mode

A detection mode where vehicle detection is represented by a single “on” pulse to the controller.

### Queue

A line of vehicles, bicycles, or persons waiting to be served by the system in which the flow rate from the front of the queue determines the average speed within the queue. Slowly moving vehicles or people joining the rear of the queue are usually considered part of the queue. The internal queue dynamics can involve starts and stops. A faster-moving line of vehicles is often referred to as a moving queue or a platoon.

### Queue discharge

A flow with high density and low speed, in which queued vehicles start to disperse. Usually denoted as Level of Service F.

### Queue spillback

A term used to describe vehicles stopped at an intersection that exceed the available storage capacity for a particular movement.

**Queue storage ratio**
The parameter that uses three parameters (back of queue, queued vehicle spacing, and available storage space) to determine if blockage will occur.

**Quick-Estimation Method**
A method detined in Chapter 10 of the HCM 2000 that allows an analyst to identify the critical movements at an intersection, estimate whether the intersection is operating below, near, at, or over capacity, and approximate the amount of green time needed for each critical movement.

**Red Change Interval**
The period of time following a yellow period indicating the end of a phase and stopping the flow of traffic.

**Red time**
The period, expressed in seconds, in the signal cycle during which, for a given phase or lane group, the signal is red.

**Red Truncation**
A signal priority treatment to terminate non-priority approach green phasing early in order to more quickly return to green for the priority approach. This treatment is also known as early return to green.

**Ring**
An phases that operate in sequence.

**Ring Barrier Diagram**
A graphical representation of phases within a set of rings and phases within a set of barriers.

**Saturation Flow Rate**
The equivalent hourly rate at which vehicles can traverse an intersection approach under prevailing conditions, assuming a constant green indication at all time and no loss time, in vehicles per hour or vehicles per hour per lane.

**Saturation headway**
The average headway between vehicles occurring after the fourth vehicle in the queue and continuing until the last vehicle in the initial queue clears the intersection.

**Section**
A group of signalized intersections used to analyze traffic operations, develop new signal timings, and operate in the same control mode—manual, time of day, or traffic responsive

**Segment**
A portion of a facility on which a capacity analysis is performed; it is the basic unit for the analysis, a one-directional distance. A segment is defined by two endpoints.

**Semi-Actuated Control**
A type of signal control where detection is provided for the minor movements only and the signal timing returns to the major movement because it has no detection and is placed in recall.

**Signal Head**
An assembly of one or more signal indications.

**Signal Coordination**
An operational mode that synchronizes a series of traffic signals to enhance the operation of one or more directional movements.

### Signal Warrant
A threshold condition to determine whether a traffic signal is justified based on satisfaction of an engineering study. There are eight warrants provided in the MUTCD.

### Signalization condition
A phase diagram illustrating the phase plan, cycle length, green time, change interval, and clearance time interval of a signalized intersection.

### Simple left turn protection
A signal phasing scheme that provides a single protected phase in each cycle for a left turn.

### Simultaneous Gap
This parameter requires all phases to concurrently "gap out" prior to crossing the barrier.

### Software in the loop (SITL)
A means of providing a direct linkage between simulation models and software emulations of controllers,

### Speed
A rate of motion expressed as distance per unit of time.

### Split
The time assigned to a phase (green and the greater of the yellow plus all-red or the pedestrian walk and clearance times) during coordinated operations. May be expressed in seconds or percent.

### Start-up lost time
The additional time, in seconds, consumed by the first few vehicles in a queue at a signalized intersection above and beyond the saturation headway, because of the need to react to the initiation of the green phase and to accelerate.

### Stopped Delay
A measurement of the aggregate sum of stopped vehicles for a particular time interval divided by the total entering volume for that movement.

### Stop time
A portion of control delay when vehicles are at a complete stop.

### Time-Before-Reduction
This volume density timing period begins when the phase is Green and there is a serviceable call on a conflicting phase. When this period is completed, the linear reduction of the Passage Time begins.

### Time of Day Plans
Signal timing plans associated to specific hours of the day associated with fluctuations in demand.

### Time-Space Diagram
A chart that plots the location of signalized intersections along the vertical axis and the signal timing along the horizontal axis. This is a visual tool that illustrates coordination relationships between intersections.

### Time-To-Reduce
This volume density timing period begins when the Time-Before-Reduction ends and controls the linear rate of reduction until the Minimum Gap is achieved.

### Total delay
The sum of all components of delay for any lane group, including control delay, traffic delay, geometric delay, and incident delay .

**Total lost time**

The time per signal cycle during which the intersection is effectively not used by any movement; this occurs during the change and clearance intervals and at the beginning of most phases.

**Track Clearance Green Time**

Associated with rail preemption at signalized intersection near rail crossings. Track clearance green time is the signal timing provided to an approach to ensure queued vehicles can be moved off the rail crossing, prior to the beginning of railroad warning lights and gate lowering.

**Traffic Control Center (Traffic Management Center)**

An optional physical component of a signal system which contains the operational database that stores controller data, allows monitoring of the system, and allows timing and other parameters to be modified.

**Traffic Control Devices**

A device used to control conflicting traffic flows, typically at an intersection. Examples include traffic signals, stop signs, yield signs, and roundabouts.

**Traffic Responsive Operation**

A signal operational method which uses data from traffic detectors, rather than time of day, to automatically select the timing plan best suited to current traffic conditions.

**Traffic Signal**

A device to warn, control, or direct at least one traffic movement at an intersection.

**Traffic Signal Controller**

A device controlling indication changes at a traffic signal.

**Traffic Signal Inventory:**

A database related to the traffic signal including information such as location, signal layout, signal timing, coordinated or uncoordinated signal operation, communication, operating agency, and history of updates.

**Transit Efficiency**

Performance measures for transit vehicles such as percent on-time, ridership, and travel time.

**Travel Time (Average)**

The total elapsed time spent traversing a specified distance. The average travel time represents an average of the runs for a particular link or corridor.

**Travel Time and Delay Study**

This study is used to evaluate the quality of traffic movements along an arterial and determine the locations, types, and extent of traffic delays. Typical measures of effectiveness include travel time, delay, percent runs stopped, and average speed.

**Two-way left-turn lane**

A lane in the median area that extends continuously along a street or highway and is marked to provide a deceleration and storage area, out of the through-traffic stream, for vehicles traveling in either direction to use in making left turns at intersections and driveways.

**Uniform delay**

The first term of the equation for lane group control delay, assuming uniform arrivals.

**Uniform Vehicle Code**
A set of traffic laws prepared by the National Committee on Uniform Traffic Laws and Ordinances. The specific traffic laws within the code vary from state to state and within different jurisdictions.

**Unit extension**
See passage time

**Unmet demand**
The number of vehicles on a signalized lane group that have not been served at any point in time as a result of operation in which demand exceeds capacity, in either the current or previous analysis period. This does not include the normal cyclical queue formation on the red and discharge on the green phase.

**Upstream**
The direction from which traffic is flowing.

**Variable Initial**
A volume density parameter that uses detector activity to determine if the initial green interval will exceed minimum green time.

**Volume**
The number of persons or vehicles passing a point on a lane, roadway, or other traffic-way during some time interval, often 1 h, expressed in vehicles, bicycles, or persons per hour.

**Volume-Density**
A phase timing function that uses parameters (variable initial, min gap, time before reduction, time to reduce) to provide appropriate minimum green time to clear intersection queues when stop bar detectors are not used and/or it is desired to adjust the passage time.

**Volume-to-Capacity Ratio**
Also known as degree of saturation is a ratio of demand volume to the capacity for a subject movement.

**Walk Interval**
An indication providing right-of-way to pedestrians during a phase.

**Yellow Change Interval**
An indication warning users that the green indication has ended and the red indication will begin.

**Yellow Extension**
The portion of the yellow change interval that some vehicles use to pass through the intersection during the yellow change interval.

**Yellow Trap**
A condition that leads the left-turning driver into the intersection when it is possibly unsafe to do so even though the signal displays are correct.

**Yield Point**
A point in a coordinated signal operation that defines where the controller decides to terminate the coordinated phase.

# CHAPTER 1

# INTRODUCTION

## TABLE OF CONTENTS

## TABLE OF FIGURES

# 1.0 INTRODUCTION

This *Traffic Signal Timing Manual* (TSTM) is intended to be a comprehensive guide to the traffic signal timing engineer and technician on traffic signal control logic principles, practices, and procedures. The TSTM represents a synthesis of traffic signal timing concepts, analytical procedures, and applications based on North American practice into a single publication. The manual also presents a framework for evaluating traffic signal timing applications related to maintenance and operations.

This manual is intended to complement policy documents such as the Manual on Uniform Traffic Control Devices, and is not intended to replicate or replace the Highway Capacity Manual, national or local engineering documents on signal timing, nor is it intended to serve as a standard or policy document. Rather, it provides a summary of practices intended to help practitioners in the timing of traffic signals.

## 1.1 BACKGROUND

The origin of traffic control signals can be traced back to the manually operated semaphores first used in London as early as 1868. The first traffic signal in the United States was developed with the objective to prevent accidents by alternatively assigning right of way. The traffic signal has changed significantly since its early development.

Today, there are more than 272,000 traffic signals in the United States (*1*). They play an important role in the transportation network and are a source for significant frustration for the public when not operated efficiently. As the era of freeway building draws to a close, urban arterials are being called upon to carry more users than ever before at a time when the users of these facilities are growing more complex (older drivers, more distractions, larger vehicles, etc) and the demand for such use continues to outpace transportation supply. According to the *2001 Nationwide Personal Transportation Survey,* on average, an individual traveled 40 miles per day, up from approximately 35 in 1990 (*2*). At the same time, the use of traffic signals at a busy intersection in a typical urban area might direct the movement of as many as 100,000 vehicles per day. In fact over ten percent of all intersections in California carry more than 60,000 Average Daily Traffic (ADT) for movements (*3*). It is estimated that many of these signals could be improved by updating equipment or by simply adjusting and updating the timing plans. Outdated or poor traffic signal timing accounts for a significant portion of traffic delay on urban arterials and traffic signal retiming is one of the most cost effective ways to improve traffic flow and is one of the most basic strategies to help mitigate congestion.

Despite their important role in traffic management, traffic signals, once installed, are often not proactively managed. Maintenance activities are frequently delayed or canceled, in reaction to shrinking budgets and staffs. More than half of the signals in North America are in need of repair, replacement, or upgrading. In 2007, the National Traffic Signal Report Card was released by the National Transportation Operations Coalition and consisted of the composite national scores from an agency self-assessment related to traffic signal control and operations, the responses in five sub areas indicate an overall national "grade" of D up from a D- in 2005. (*4*).

FHWA has recognized the critical role that traffic signal timing plays within the overall transportation network. Signal timing offers the opportunity to improve the mobility and safety of the street system and contribute environmental benefits. This document is intended to further increase the awareness of the need for resources devoted to operation of the transportation system.

### 1.1.1 Purpose of Traffic Signals

The Manual on Uniform Traffic Control Devices (MUTCD) defines a traffic control signal as any highway traffic signal by which traffic is alternatively directed to stop and permitted to proceed. Traffic is defined as pedestrians, bicyclists, ridden or herded animals, vehicles, streetcars, and other conveyances either singularly or together while using any highway for purposes of travel. (*5*)

It is with this need to assign the right of way at locations that we consider the dual purpose of traffic signals —efficiency and safety— which in some cases seem to be conflicting. Safety may be seen as an element needed to be sacrificed in order to achieve improvements in efficiency and meet ever-increasing demands. The reality is that traffic signals can, and in fact must, serve both operational efficiency and safety based on the conditions. The MUTCD goes on to describe that traffic control signals can be ill-designed, ineffectively placed, improperly operated, or poorly maintained, with resulting outcomes of excessive delay, disobedience of the indication, avoidance, and increases in the frequency of collisions.

A traffic signal that is properly designed and timed can be expected to provide one or more of the following benefits:

1. Provide for the orderly and efficient movement of people.
2. Effectively maximize the volume movements served at the intersection.
3. Reduce the frequency and severity of certain types of crashes.
4. Provide appropriate levels of accessibility for pedestrians and side street traffic.

The degree to which these benefits are realized is based partly on the design and partly on the need for a signal. A poorly designed signal timing plan or an unneeded signal may make the intersection less efficient, less safe, or both.

### 1.1.2 Intersection Design and its Relationship to Signal Timing

The design of the intersection has a direct influence on its safety and operation from a design and user-ability perspective. Design elements that are particularly relevant include the number of lanes provided on each approach and for each movement, whether there are shared thru-and-turn lanes, the length of turn bays, the turning radii (especially important for pedestrians), the presence of additional through lanes in the vicinity of the intersection, the size and location of detectors, and presence or absence of left-turn phasing. Other geometric features, like additional through or turn lanes, can also have a significant positive impact on intersection capacity, provided that they are sufficiently long. The other aspect of intersection design is the perception and reaction of the end users. Various decisions need to be made as a user approaches the intersection, which makes it important to simplify the decision making process.

Another aspect of the design is detection. Detectors provide the ability to sense vehicle and pedestrian demands at an intersection; enabling modes of operation that may be more efficient than fixed or pre-timed control. It is critical that functional and properly designed detectors communicate with the controller to ensure continued functional signal control at the intersection. Detectors that are improperly located or are an inappropriate length can unnecessarily extend the green indication and increase the frequency of phase termination to the maximum limit (i.e., max out). Conversely, a poorly located detector could cause premature gap-out. A protected left-turn phase provides a time separation for left-turning and opposing traffic streams and may reduce left-turn delays or related crashes. However, the additional phase increases the minimum cycle length and may increase intersection delays and, in the case of a protected-only left-turn, may even increase left-turn delay.

The topics discussed in this section are intended to serve as a reminder of the close relationship between signal timing, intersection design, and traffic control device layout. The quality of the signal timing plan is directly tied to the adequacy of the intersection design and the traffic control device layout. In some situations, achieving safe and efficient intersection operation may require changes to the intersection design or the traffic control device layout. The subsequent chapters of this manual provide more detailed information about the role of these factors in signal timing plan development.

### 1.1.3 Objectives of Basic Signal Timing Parameters and Settings

A primary objective of signal timing settings is to move people through an intersection safely and efficiently. Achieving this objective requires a plan that allocates right-of-way to the various users. This plan should accommodate fluctuations in demand over the course of each day, week, and year.

**Because travel demand patterns change over time, the signal timing plan should be periodically updated to maintain intersection safety and efficiency.**

There are many signal timing parameters that affect intersection efficiency including the cycle length, movement green time, and clearance intervals. Increasing a traffic movement's green time may reduce its delay and the number of vehicles that stop. However, an increase in one movement's green time generally comes at the expense of increased delay and stops to another movement. Thus, a good signal timing plan is one that allocates time appropriately based on the demand at the intersection and keeps cycle lengths to a minimum.

The relationship between signal timing and safety is also addressed with specific timing parameters and the design of the intersection. For instance, the intent of the yellow change interval is to facilitate safe transfer or right-of-way from one movement to another. The safety benefit of this interval is most likely to be realized when its duration is consistent with the needs of drivers approaching the intersection at the onset of the yellow indication. This need relates to the driver's ability to perceive the yellow indication and gauge their ability to stop before the stop line, or to travel through the intersection safely. Their decision to stop, or continue, is influenced by several factors, most notably speed. Appropriately timed yellow change intervals have been shown to reduce intersection crashes (*6*). Signal timing plans that reduce the number of stops and minimize delays may also provide some additional safety benefits.

**The traffic signal controller at an intersection implements timing settings designed for that specific location. The settings are designed to respond to users at the intersection and meet objectives defined by the policies of the responsible agency.**

The policies may include standards defined by the agency with potential guidance from regional or state standards and must consider pedestrians, vehicular traffic conditions, change and clearance intervals, and if actuated, detection layout. These settings may be influenced by adjacent intersections (the concept of coordination is more fully explored in Chapter 6), but are applicable for each intersection considered as an isolated unit.

### 1.1.4 Establishing the Need for Retiming

Traffic professionals have long recognized the value of designing effective signal timing to meet changing travel patterns and characteristics. In 1995, the U.S. General Accounting Office (GAO) reported, "Properly designed, operated and maintained traffic control signal systems yield significant benefits along the corridors and road networks on which they are installed. They mitigate congestion and reduce accidents, fuel consumption, air pollutants and travel times. Resource constraints have prevented the use of traffic signals to their full potential. The Traffic Signal Report Card Technical Report goes on to state:

> "It became clear that for safety and liability reasons, agencies must ensure a basic level of operation of the traffic signal system so that signals continue to turn green, yellow and red. The signals may not function efficiently for traffic or pedestrians, but, technically, the signals are working and that is what people see. However, the uniformly low scores (on the National Report Card) indicate that, for the most part, people consistently experience poor traffic signal performance and, as a consequence, their expectations are low. The pattern, once again, is one where agencies are forced to use their resources to deal with critical maintenance issues when they arise rather than proactively. Signal systems are managed to simply ensure base levels of performance."

The National Transportation Operations Coalition (NTOC) and FHWA continue to work to make the case that additional resources are needed to develop signal timing plans and to modernize equipment. There's an old saying that transportation engineers leave a little of their intelligence on the street when they design an intersection, but these designs are limited by the technology they have to work with. The use of 20-year-old technology and infrastructure may satisfy the requirement for the signal to display green, yellow, and red, but it may not offer the opportunity to efficiently operate the system or provide preferential treatment for a certain type of user to meet the policies and desires of the community. In most cases, upgraded equipment improves the efficiency for staff to manage the system, assuming the staff is properly trained to operate the upgraded equipment.

These efficiencies are observed with updating traffic signal timing plans, developing new strategies to improve transportation, and improving customer service. There have been some great technological advances in the past five years, such as the development of transit signal priority, which seeks to provide preferential treatment to buses as they approach the traffic signal. This new technology allows the engineer to allocate green time that more closely reflects the community's transportation policies.

The MUTCD also speaks to the issue of efficiency and recommends proper design and signal timing to ensure that it satisfies current traffic demands. The Maintenance of signal timing plans will be described in Chapter 8 of this document. This statement does not speak to the need to provide preferential treatment or other policies that will be further described in Chapter 9 of this document.

### 1.1.5 Benefits of Up-to-Date Timing

Studies around the country have shown that the benefits of area-wide signal timing outweigh the costs 40:1 (or more). The benefits of up-to-date signal timing include shorter commute times, improved air quality, reduction in certain types and severity of crashes, and reduced driver frustration.(*7*)

The NTOC recently surveyed the quality of traffic signal operations in the United States. The NTOC concluded that the nation scored a D in terms of the overall quality of traffic signal operation. "If the nation supported its signals at an 'A' level, we would see:

- Reductions in traffic delay ranging from 15-40% (*8*); reductions in travel time up to 25%; and reductions in stops ranging from 10-40% (*9*). For example, if you spent two hours in your car commuting to and from work and running errands, you'd save 50 hours per year (or more than a work week) because of improved signal timing.
- Reductions in fuel consumption of up to 10%. Nationwide this would amount to a savings of almost 170 billion gallons of motor fuels per year.(*10*)
- Reduction in harmful emissions (carbon monoxide, nitrogen oxides, volatile organic compounds) up to 22% (*11*). According to the Surface Transportation Policy Project, motor vehicles are the largest source of urban air pollution. (*12*) In addition, the EPA estimates that vehicles generate 3 billion pounds of air pollutants yearly.(*13*",*14*)

Beyond the benefits to vehicular traffic, there are opportunities to improve performance for transit, pedestrians, and freight movement. Chapter 9 summarizes some of the advanced concepts that address some of the broad policies that have gained in popularity since the inception of the Intermodal Surface Transportation Efficiency Act of 1991 (*15*.)

## 1.2 ORGANIZATION OF THE MANUAL

The manual is organized into nine chapters that can be broadly described by four basic parts:

*Part 1 —Policy, Planning, and Funding Considerations (Chapter 2).* This part describes the need for and benefits of signal timing. It will present a discussion of relevant federal, state, regional, and local issues, as well as typical funding needs and options. This section discusses what decision makers consider when determining the effect of signal timing on the regional framework. It also presents how the signal timing process described in Chapter 7 fits into the regional planning framework.

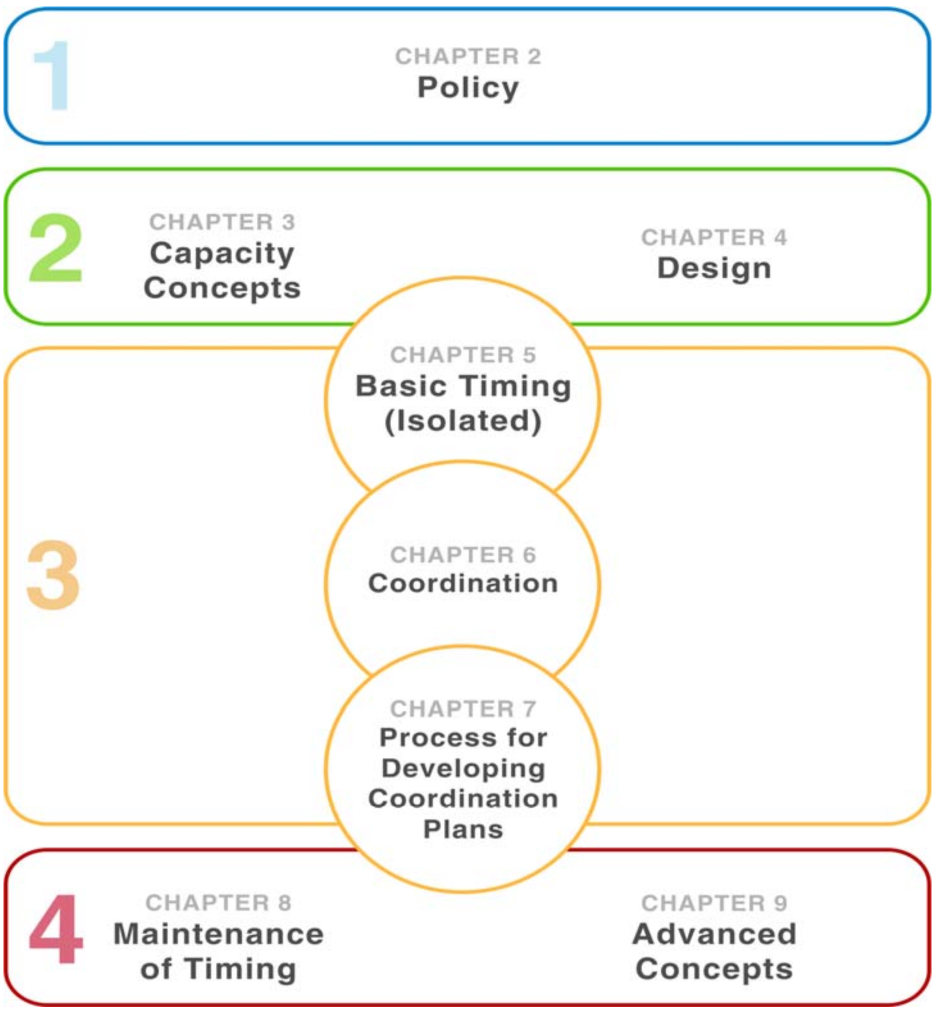
*Part 2— Fundamental Concepts of Capacity (Chapter 3) and Traffic Signal Design (Chapter 4).* This part includes Chapters 3 and 4 which provide key background information needed to understand signal timing. This chapter may not be necessary for experienced engineers, but provides a basic foundation from which to describe more complicated concepts.

*Part 3— Signal Timing Concepts, Guidelines, and Application and Coordination Plan Development (Chapters 5, 6, and 7).* This part describes traffic signal timing from the concepts to application. It provides guidelines where appropriate based on industry practice. This is a chapter that provides several examples from agencies that represent good practice.

*Part 4—Maintenance of Timing (Chapter 8) and Advanced Topics (Chapter 9).* This part presents an overview of a number of advanced topics related to improving signal timing operations that will be especially relevant to sophisticated timing engineers that are implementing innovative strategies (transit signal priority, adaptive signal timing, etc). This section will provide an overview of each topic with references to more detailed documents.

This organization is described in Figure 1-1.

**Figure 1-1 Organization of the Manual**

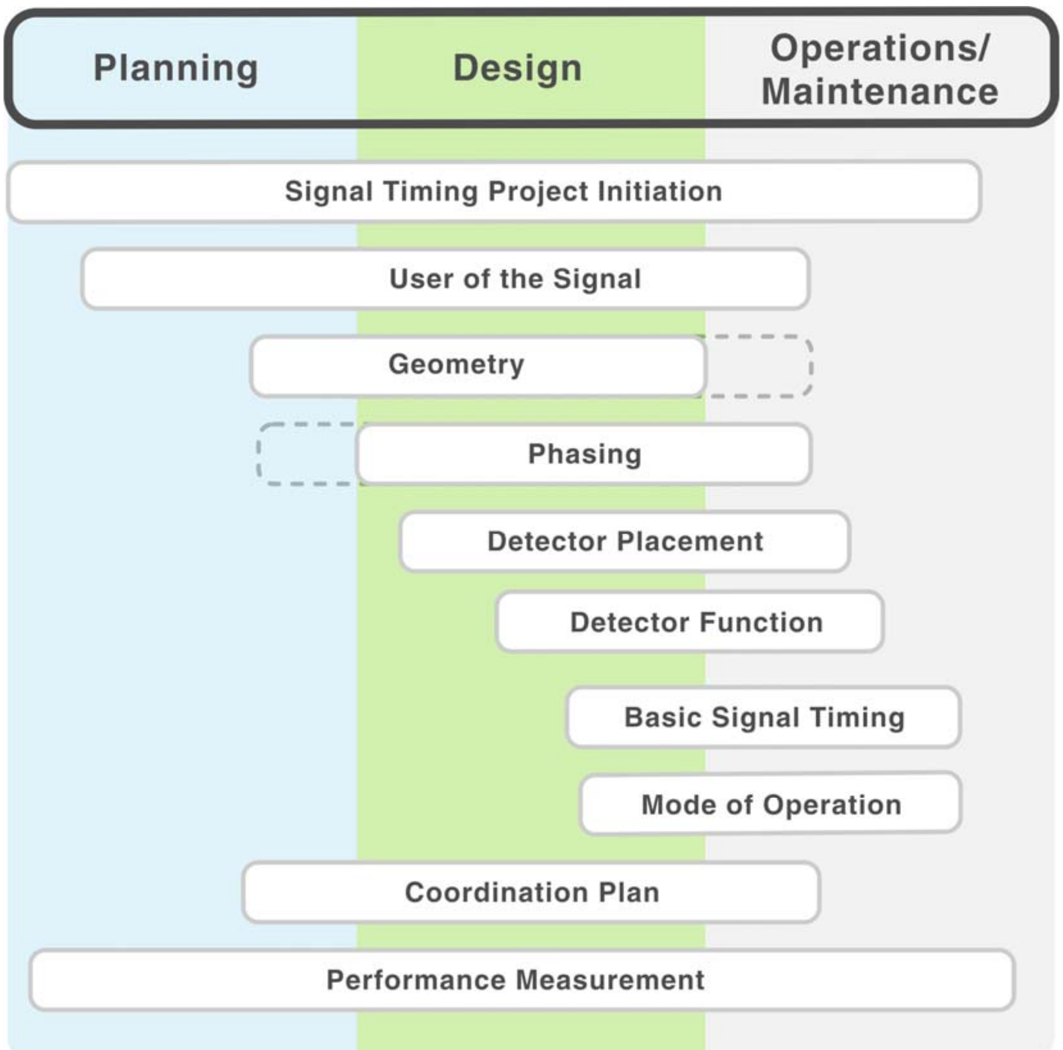



## 1.3 USE OF THE MANUAL

The *Traffic Signal Timing Manual* (TSTM) is intended to be a comprehensive document describing the procedures for generating signal timing plans for North American applications. It is intended for use by a range of practitioners; including traffic engineers, signal technicians, design engineers, planners, transportation managers, teachers, and university students. To use the manual effectively and apply its methodologies, some technical background is desirable, typically technical training either provided as a part of continuing education, or at the university-level.

To help describe the use of this manual, three generalized users of the signal timing manual have been identified, planners, designers, and operators/maintenance staff. Each of these users will interact with different elements of the signal timing process, and should understand where and how the others are involved. Figure 1-2 shows graphically the interaction or relationship between the generalized user function: planning, design, and operations/maintenance, and the key topics of signal timing, shown in white.

**Figure 1-2 User Interaction within the Signal Timing Process**




The general topics shown in Figure 1-2, represent the basic signal timing process, and are described in detail within the manual. The topics are described briefly here:

- **Signal Timing Project Initiation**– Establish the purpose(s) of the signal timing project and how success will be measured. This will encompass all three types of users.

- **User of the Signal** – Determine and prioritize the users on the transportation network that the signal must serve. For example, motor vehicles, pedestrians, bicycles, transit, emergency vehicles, etc.
- **Geometry** – Determine the lane use, dedication, and geometry and roadway infrastructure with which the signal timing will interact. Roadway infrastructure may not change, but lane designation and use can change along with signal timing parameters. For example, multiple turn-lanes, freeway ramp terminal, channelized movements, turn-restrictions, phasing, etc.
- **Phasing** – Set the appropriate signal phasing scheme(s). For example, split phasing, protected left-turns, overlapping movements, etc.
- **Detector Placement** – Determine if and where detectors are needed based on desired signal operations. For example, pedestrian push buttons or bike/vehicle detection locations, or even transit priority or emergency vehicle preemption detection zones.
- **Detector Function** – Set the functionality of the detectors. For example, vehicle detector settings of pulse or presence or count, or a delay setting for movements with heavy right turn on red, or etc.
- **Basic Signal Timing** – Calculate or establish the appropriate basic timing parameters. For example, pedestrian walk and flash don't walk, gap or passage time, yellow and all red vehicle clearance, etc.
- **Mode of Operation** – Will the signal operate in fixed-time, actuated, or adaptive mode?
- **Coordination Plan** – Determine if the signal should be coordinated with other nearby signals as a system or operate in isolation? Coordinating will require interconnection communication lines between traffic signal controller cabinets.
- **Performance Measurement** – What measurements will be used to judge the performance of the signal timing and operations? Typically performance measurements are average delay, travel time, stops, and etc.

Each of these steps may take place in the planning, design or operation of a traffic signal and its signal timing plans. To help understand this interrelationship, let's take the example of detection placement. From Figure 1-2, we see this area involves both design and operations/maintenance. A designer may select a inductive loop detector placement location based on a design manual or agency policy for the design travel speed. What the designer may not realize is once the loop detection is embedded in the pavement, it is difficult to relocate. If the operator or traffic engineer wants to change the posted speed or operate the signal using detection intensive, adaptive control, alternate loop locations maybe necessary. This would force the costly relocation of the designed loops. If the designer and operator/engineer communicate and work together to understand each others needs, a multi-use solution can be developed to benefit the roadway user.

Many of the topics discussed in the Traffic Signal Timing Manual will be applicable to multiple types of users, and an understanding of signal timing holistically will allow the best solutions possible to be developed and implemented, to the gain of the traveling public.

## 1.4 REFERENCES


*1* 2007 National Traffic Signal Report Card Technical Report, National Transportation Operations Coalition.

*2* USDOT. *Nationwide Personal Transportation Survey,* http://nhts.ornl.gov.2001/index.shtml, 2001.

*3* Highway Safety Information System (HSIS). California database, 1994-1998.

*4* National Transportation Operations Coalition, National Traffic Signal Report Card Technical Report, 2007.

*5* Federal Highway Administration (FHWA). *Manual of Uniform Traffic Control Devices.* Washington, DC: USDOT, FHWA, 2003.

*6 Retting, R.A.; Chapline, J.F.; and Williams, A.F., "Changes In Crash Risk Following Re-Timing Of Traffic Signal Change Intervals". Accident Analysis and Prevention 34:215-20, 2002*

*7* Ibid.

*8 "Temporary Losses of Highway Capacity and Impacts on Performance: Phase 2, Oakridge National Laboratory, November 2004, ORNL/TM-2004/209 and "Benefits of Retiming Traffic Signals", ITE draft Informational Report, 2005.*

*9 "Benefits of Retiming Traffic Signals", ITE draft Informational Report, 2005.*

*10 http://www.bts.gov/publications/national_transportation_statistics/2004/html/table_04_09.html*

*11 ITS Benefits and Costs Database, USDOT ITS Joint Program Office, http://www.benefitcost.its.dot.gov/*

*12 Surface Transportation Policy Project. http://www.transact.org/library/factsheets/environment.asp.*

*13 National Toxics Inventory, 1996. http://www.epa.gov/otaq/regs/toxics/d00003.pdf*

*14 Ibid "National Traffic Signal Report Card"*

*15* "ISTEA Reauthorization Policy Statement And Principles", U.S. Department of Transportation, http://www.dot.gov/ost/govtaffairs/istea/isteap&p.html, visited January 6, 2007.

# CHAPTER 2

# SIGNAL TIMING POLICY

## TABLE OF CONTENTS

## LIST OF FIGURES

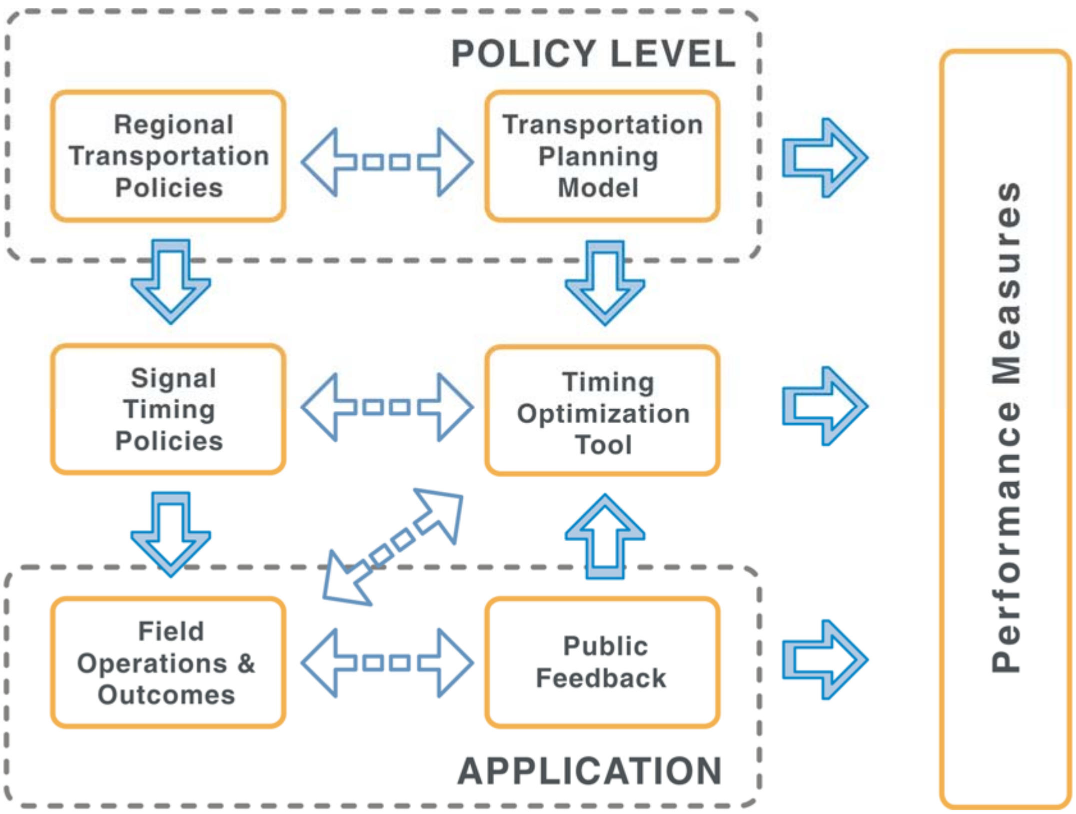
## LIST OF TABLES

# 2. POLICIES, PROCESS, AND FUNDING

Signal timing is important because it directly affects the quality of our transportation system, which affects virtually everything within our communities. Signal timing impacts the time we spend traveling, the quality of the air we breathe, the safety of roadway travel, the costs of our trips, and many aspects of our daily lives. Signal timing policy is important because it is a way to help control and define priorities within the transportation system and how signal timing is applied. A clearly defined signal timing policy should be an extension of a region's transportation policy reflecting a region's values in the operations and safety of their transportation network. The context of the term policy is to support strategic objectives, and do not represent global applications in most cases. Engineering judgment should be used in concert with policy development and applications in the signal timing realm.

This chapter describes the relationship between transportation policies and signal timing applications, summarized in Figure 2-1.

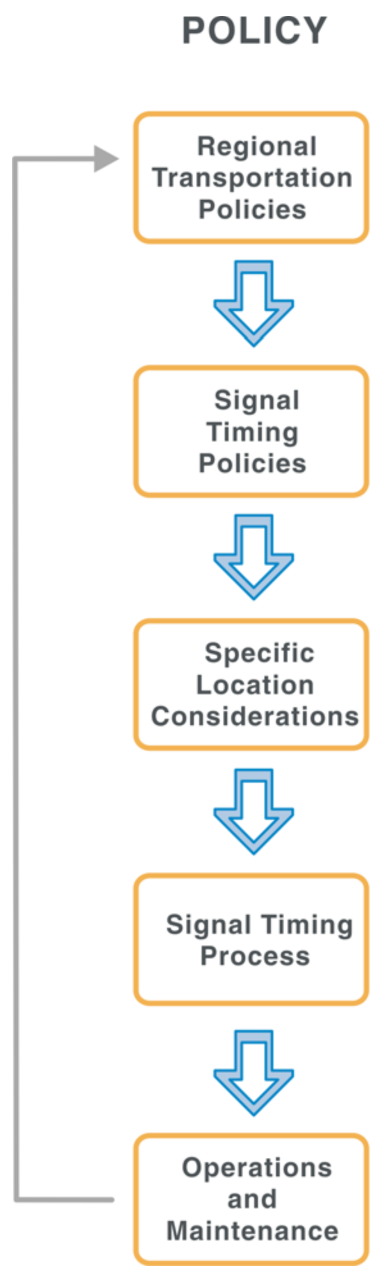
**Figure 2-1 Transportation Policy and Signal Timing Application Relationships**




This chapter contains five sections.

- The first section presents a summary of policy development.
- The second section provides an overview of the signal timing process and considers how policies should affect several decisions made during signal timing. Examples of different signal timing policies are provided as well.
- The third section presents performance measures and a national perspective with regards to signal timing.
- The fourth section presents a discussion of various funding considerations for signal timing.
- The final section presents examples of programs that have been effective in coordinating signal timing, policy, and performance measures.

## 2.1 POLICY DEVELOPMENT




*The diagram at right illustrates the policy development cycle discussed in this section. Signal timing policies flow from the overall regional transportation objectives. Local considerations are taken into account prior to initiating the specific timing process. Operation of the system generates user feedback that, in turn, drives overall policy.*

Regional transportation policies, and more specifically signal timing policies, guide the strategy for how to operate signalized intersections, devices, signal timing equipment, and for prioritizing travel modes and/or transportation facilities at a system level branching beyond jurisdictional boundaries. Regional transportation policies should be used to help establish signal timing policies, providing guidance in the development of specific objectives for the implementation of signal timing plans. The development of local signal timing strategies should consider the regional transportation policies to determine if there are regional objectives that will influence the process. These policies may also be consulted to determine their effect on ongoing operations and maintenance activities. The regional transportation policies are a result of a community visioning process that develops a plan that:

- sets the direction and guides the planning for improvements for a region's transportation system during the next 20 years;
- establishes policies and priorities for all forms of travel – motor vehicle, transit, pedestrian, bicycle and freight – and street design and the efficient management of the overall system;
- anticipates the region's current and future travel needs based on forecasts of growth in population, households and jobs as well as future travel patterns and analysis of travel conditions;
- evaluates federal, state and local funding that will be available for transportation improvements; and
- estimates costs of projects and proposes funding strategies to meet these costs.(*1*)

The plan is typically developed with the help of elected leadership and ultimately represents the attributes the community hopes to address. Examples of this type of plan include a Transportation Improvement Program (TIP), a Capital Improvement Program (CIP), a Long-Range Transportation Plan (LRTP) and others.

The signal timing policies should establish the signal timing objectives and performance measures for a jurisdiction all the way down to a particular facility or intersection.

Location-specific considerations must also be identified to determine whether a specific corridor or area has special needs. For instance, a city may have a policy of maintaining a high level of mobility on its arterial street system, but that may not be applicable within a downtown area that may have a policy of maintaining access. An important step in the flow chart to the right is to consider the special needs before moving to the next step of the flowchart.

The signal timing policies can be used to refine the scope of what is to be accomplished during the signal timing process. Signal timing policies should establish answers to the following questions:

- What should be improved?
- What objectives or various user needs should be optimized?
- What performance measures should be tested?
- What standards must our policy follow?
- What data should be collected to develop signal timing?
- What should be measured after implementation?
- What should be monitored as part of maintenance?

Agencies should use their signal timing policies to guide the development, operation, and maintenance of their signal system.

### 2.1.1 Policy Influence on Signal Timing

The primary goal of traffic signal timing is to maintain the safe and efficient transfer of right-of-way between conflicting streams of users; however, a safe and efficient system varies within each community's context as described previously. Thus, local, regional, state, and federal policies must be considered to determine a proper approach. These policies form the foundation from which performance measures are selected. These performance measures should be tracked through the signal timing process.

Policy questions related to signal timing include determining:

- Whether all types of users (transit, freight, emergency respondents, pedestrians, vehicles, bicycles, etc.) will be treated equally or prioritized at the signalized intersection;
- How frequently will signal timing plans be reviewed and updated;
- How approaches with differing street classifications should be treated;
- Whether there will be preferential treatment for certain movements beyond the definition of the coordinated phase (will the coordination timing plan clear all queues during each cycle for left turns and side street through movements);
- How intersections with deficient capacity will be treated; and
- What measures will be used to determine whether the timing plan is effective (vehicle stops, network delay, arterial travel speed, estimated person delay, estimated fuel consumption, transit speed, etc.) and how will they be collected.

Signal timing by its nature is the assignment of right-of-way to competing directions and travel modes. Signal timing often requires trade-offs between various modes at an intersection, such as vehicles versus pedestrians and bicycles. These tradeoffs could result in competing ideas, such as safe pedestrian crossing times versus maximizing automobile capacity.

**For every action, there are following reactions, and these types of trade-offs with signal timing should be made clear in the mind of decisionmakers in order to ensure signal timing policy is developed within the correct context. Examples of these trade-offs can be seen in the following portion of this section.**

Policies help define the objectives for signal timing plans. User expectations for a street network are often the guiding principles for an agency's policies. There are many types of users, including rail, pedestrian, bicyclists, transit, emergency vehicles, and automobiles; however, it is the prioritization of these users on a facility that should reflect an understanding of agency policy and, in turn, provide signal timing objectives. Traditionally, rail is given the highest priority due to its greater momentum and limited braking characteristics relative to other modes of travel. This is seen when there is a rail crossing and all movements are preempted in deference to the rail movement when a train is approaching. Other priority considerations include preemption or priority service for emergency vehicles (preemption overrides normal phasing sequence and coordination; priority service does not). An agency can provide emergency vehicles with devices that give them preference (preemption or priority) when needed. An agency can provide signal priority to transit vehicles to facilitate higher mobility (a bus can transport more people than an automobile) while being less disruptive to other system users. Details regarding signal preemption, priority, and other advanced signal timing topics are covered in Chapter 9.

Each of the signal timing policies and expectations may vary based on the users, roadway facilities, and mode split. For example, the user expectations and objectives for a downtown core are different from those in a suburban setting. Within the Central Business District (CBD) of a medium or large city there are many users to consider. There are typically a higher percentage of pedestrians and transit vehicles in this setting and, as a result, the signal timing should reflect those users. Outside the CBD, the focus may change to one of maximizing mobility for various users. On a major arterial, for example, the focus may be on a policy of decreasing travel times, number of stops, and delays for users on the major street while accommodating transit and emergency vehicles.

CBD environments require substantially different timing strategies than high-speed arterials. No single policy is correct for all situations. Each policy is site-specific and based on agency policy. It is often difficult to have a policy that completely satisfies all interested parties, as some policies inherently shift priorities from one user group to another. Regardless, the traffic engineer should ensure that the measures of effectiveness used to evaluate a system’s performance are appropriate for the policy being implemented. In addition, the traffic engineer is often best served by using a tool which can analyze the desired measures of effectiveness for optimization and evaluation. A signal timing tool that only optimizes for vehicles may need to be adjusted to the unique requirements of transit which are not reflected in many traditional signal optimization packages.

Current local and state agency signal timing policy ranges from very well-defined and well-implemented to non-existent. Based on a review of state and local agency signal timing policies, Table 2-1 shows some generalized signal timing strategies and examples of applicable signal timing policies that would apply on a case by case basis.

**Table 2-1 Signal Timing Strategies, Settings, and Policy Examples**

| Transportation Policy | Setting | Signal Timing Strategy |
|---|---|---|
| Pedestrian/Bicycle-Focused | Downtowns, Schools, Universities, Dense Multi-Use Development, Parks, or any location with high pedestrian/bicycle traffic. | Shorter cycle lengths to reduce wait times<br>Extended Pedestrian crossing timing<br>Bicycle/Pedestrian Detection<br>Exclusive Pedestrian Phasing<br>Leading Pedestrian Interval |
| Transit-Focused | Transit-corridors, along transit routes, near transit stations or crossings. | Signal preemption for high importance transit modes (i.e. rail)<br>Signal priority for strategic transit modes and routes<br>Signal coordination based on transit vehicle speeds<br>Extended Pedestrian crossing timing<br>Exclusive Pedestrian Phasing<br>Leading Pedestrian Interval |

| | | |
|---|---|---|
| Emergency Vehicle-Focused | Key roadways and routes to and from hospitals, fire stations, and police stations. | Signal preemption for high importance vehicles |
| Automobile-Focused / Freight-Focused | Locations with high automobile or truck/freight traffic, facilities of regional importance, freight corridors, ports, or intermodal sites. | Avoid cycle failure (i.e. queued vehicles not making it through the intersection on a single green indication)<br>Maintain progression on coordinated systems as best as possible to avoid unnecessary stops and delay<br>Use appropriate cycle lengths (Shorter cycle lengths will typically result in less delay, but increased “lost time” (time lost in vehicle deceleration, driver reaction time, and vehicle acceleration), while longer cycle lengths may result in more delay, less “lost time”, and potentially more vehicle throughput depending on traffic demand)<br>Ensure appropriate pedestrian signal timing to allow safe multimodal use of the roadway network. |
| Low-Volume Locations or Periods | Locations with low traffic volumes or during off-peak travel periods | Ensure efficient signal timing operations (avoid unnecessary stops and delays)<br>Consider flashing operation (yellow-red or red-red) if conditions allow<br>Use appropriate resting state for the signal with no traffic demand (i.e. rest in red, rest in green and “walk” on major roadway)<br>Allow skipping of unnecessary movements (i.e. uncalled left-turn phases) but assure it does not create “yellow trap”<br>Use half, third, or quarter cycle lengths relative to other coordinated signalized intersections<br>Allowing pedestrian actuations to temporarily lengthen a cycle length, removing an intersection out of coordination, if pedestrian and vehicular volumes are low. |

Each of these generalized signal timing policies reflects the priorities in a transportation system. These priorities should match the regional characteristics of travel in an area, as well as the roadway classification.

### 2.1.2 Challenges to Signal Timing Policy Development

There are distinct challenges associated with interpreting policies and implementing them within traffic signal timing settings. Policies often provide agencies and individuals with considerable room for interpretation. In many cases, an agency’s established policies are not fully understood by the implementing staff, or the implementing staff may not recognize that policy is being established. This is further complicated by the public’s lack of understanding of the signal timing complexities. Implemented signal timing settings that are consistent with the policies (either explicitly or by default) may be changed due to public response or complaint and the resultant changes may not be consistent with policy. A further challenge for the practitioner is that the policymaker may not appreciate the technical complexity and financial implications associated with the actual implementation of the strategy created to address a specific policy. For this reason, it is important to emphasize the relationship between transportation or signal timing policies to performance measures and funding. Details on both are presented later in this chapter.

Many agencies do not have written policies, and there is a poor understanding of the relationship between the settings used in the field and their effects on operations. Historically, many traffic engineers have made decisions based on user expectations and complaints, which may not represent best practice or the agency's policy. For example, an agency's operations standard may dictate the minimum level of service (LOS) of a signalized intersection, which, in turn, determines the capacity and the size of the intersection to achieve this LOS. This policy may result in long pedestrian crossing times, yellows, and all-reds, which can increase the cycle length for the intersection (assuming that pedestrians are present and will try to cross the wide street). The long cycle lengths increase the delay for pedestrians, which may reduce the compliance of pedestrians or increase mid-block crossings (to avoid the signal).

Similar to best design practices, a context sensitive approach can be applied and can benefit signal timing. A context sensitive approach considers the environment of the traffic signal, the local policies, and the unintended consequences of potential changes as a result of the signal timing changes.

Signal timing is also complicated because operators have no clear way to measure the ongoing operations of the entire traffic signal system. Traffic signal operation is an area that is more complicated than freeway operations, and it is difficult to identify effects of changes to the systems without the infrastructure to measure performance. Research related to arterial performance measurement is underway but has been limited in its application. An example of this is freeway management system maps that produce real-time speeds for traveler information. Recent research (NCHRP 3-79) identified that one challenge to producing similar maps for the arterial system is the complexity of measuring performance at each intersection. An additional challenge is the lack of infrastructure to communicate the information back to a location where it can be displayed.

Despite these challenges, the development of signal timing policies that follow regional and local community transportation goals and objectives is a worthwhile undertaking.

**Development of signal timing policies should be a collaborative effort between regional partners and community stakeholders, crossing jurisdictional boundaries, with the service and safety of the customer in mind at all times. Signal timing policies should be clearly documented and thoroughly communicated within an agency to those who operate and maintain the signal system.**

### 2.1.3 Use of Standards

Within all signal timing policies, there should be implicit or explicit direction to use the latest national and local standards available relating to signal timing and operations along with good engineering judgement. Specific standards applying to signal timing can be found in the Manual on Uniform Traffic Control Devices (MUTCD) as adopted by a given state, or other local adopted standards which may follow the uniform vehicle code model.

Standards specific to signal timing in the MUTCD relate to establishing safe and consistent pedestrian crossing times, vehicle clearance intervals, and signal indications, but don't specifically standardize timing parameters. The MUTCD does provide guidance and options for timing parameters of walk, flashing don't walk, solid don't walk, yellow, and all-red indications, which in name are not standards, but many states have interpreted guidance statements as having the force and effect of a standard, thus being mandatory. Still all standards, recommended practices, guidelines, and options should be applied with good engineering judgment. Public safety regarding signal timing and signal timing policy should be of the utmost importance.

A state supplement to the MUTCD may be relevant to provide boundary conditions for the signal timing. Many states produce Manuals of Traffic Signal Design, which may also identify guidance for signal timing settings.

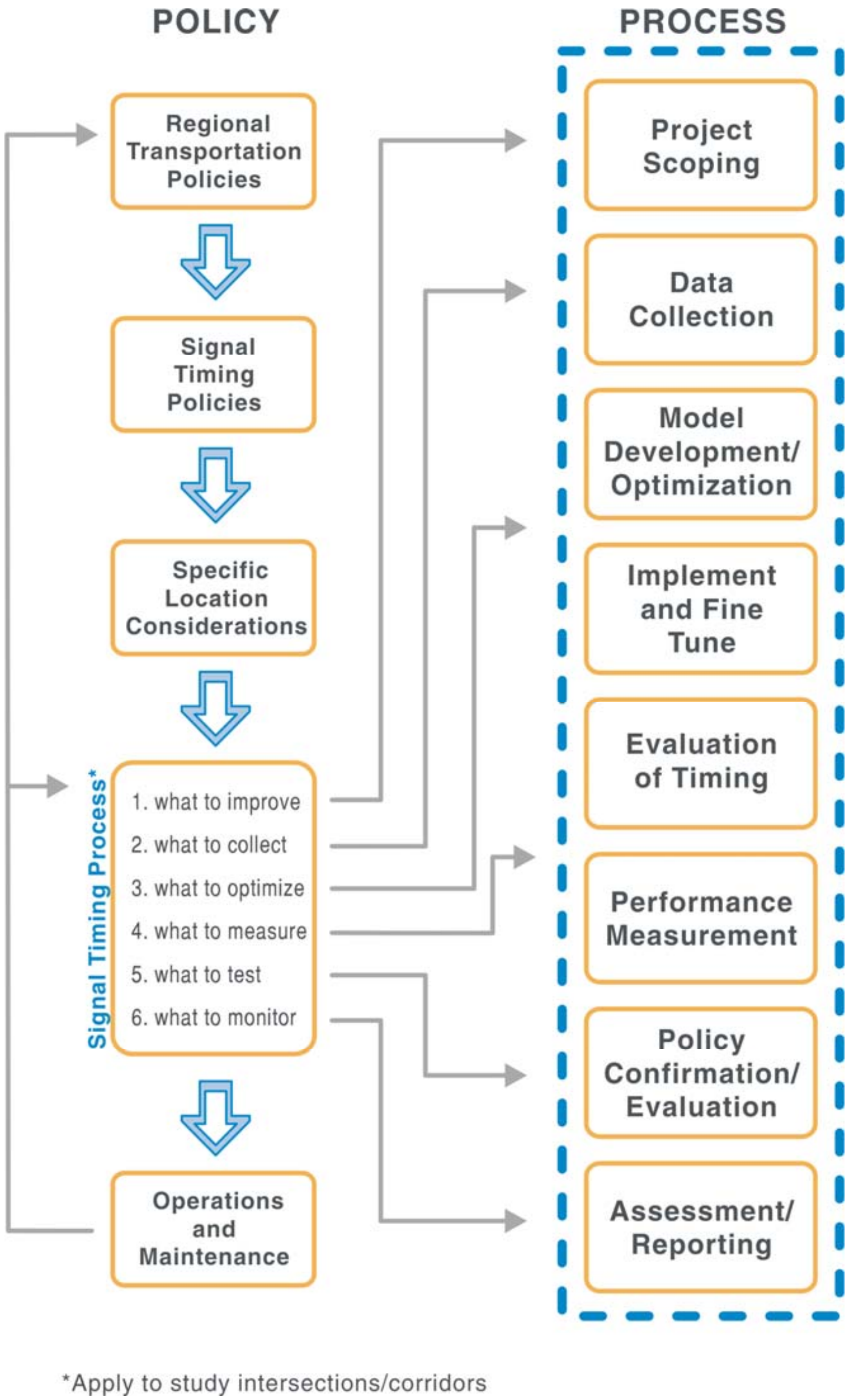
A state's Vehicle Code can also affect vehicle clearance interval timing. A good example of this is laws related to change intervals which vary depending on whether a restrictive or permissive yellow law is in place. Details on how this affects yellow clearance timing can be found in Chapter 5.

It is important to note that this document is not intended to take precedence over federal, state, or local policies. The Signal Timing Manual provides a comprehensive compilation of materials that should be considered when developing timing plans.

## 2.2 SIGNAL TIMING PROCESS

The signal timing environment has two components, the signal timing policy and the signal timing process. The signal timing policy should define how and with what form the signal timing process is implemented, and thus should actually include the signal timing process. As described in the previous section, signal timing policies should be used to refine the signal timing process

**Figure 2-2 Signal Timing Environment**




The process for developing a signal timing process is often well established in many agencies and focused on the mechanics of optimization. In many cases, the process may be affected by the time, funding, or resources available. Signal timing is a process that uses distinct procedures and one interrelated procedure. Data management, signal timing optimization, field deployment, and performance evaluation are the four quadrant procedures described in the FHWA report on Signal Timing Process, but to think beyond the specific activities we have added additional steps to focus the engineer or technician on other concepts that need to be considered. This expanded signal timing environment, shown graphically in Figure 2-2, provides a framework to achieve results consistent with overall regional and agency policies. As shown in the figure, once the policies have been established, a signal timing process can be implemented, which is further described in Chapter 7.

It is important to consider the context and variations within which the signal timing process can be applied, and the level to which specific policy plays. The signal timing process can occur at a basic, isolated intersection timing level or it can occur at a larger, area-wide or corridor-wide scale involving signal timing, a review of signal hardware and software, signal communication technology and a myriad of other signal system applications. The level and detail of policy confirmation/evaluation on a signal timing project is likely to be greater as the scope, size, or importance of the project increases. Table 2-2 shows examples of the difference between signal timing processes focused on a single intersection and a focus on regional signal timing.

**Table 2-2 Signal Timing Process Examples**

| Step | Signal Timing Process Steps | Single Intersection Field Adjustment | Regional Signal Timing Optimization |
|---|---|---|---|
| 1 | Project Scoping | Typical maintenance is performed on a routine basis | A detailed scope of work is developed and the effort is completed every one to five years |
| 2 | Data Collection | Phone call complaints may be the only step necessary, but could include other sources of data | Turning movement counts, existing signal timing sheets, “before” travel time runs, signal timing policy, etc… |
| 3 | Model Development | Field observation | Build signal timing model, typically using a software tool like Synchro or TRANSYT-7F. |
| 4 | Implement New Timing Plans | Change setting and observe | Code signal controller with new timing plans and implement in the field |
| 5 | Fine Tuning / Refinement | Observe and call back person(s) who complained for feedback | Observe and adjust/refine timing plans in the field |
| 6 | Evaluation / Performance Measurement | Customer satisfaction; more complaints? | Conduct “after” travel time and/or delay studies in the field. Compare measures of effectiveness in signal timing software such as Synchro. |
| 7 | Policy Confirmation/ Evaluation | On-site observation | Compare signal timing and traffic operations in the field to prescribed policy. |
| 8 | Assessment/ Reporting | Make change in timing in the office | Develop documentation of adjustments and results of signal timing effort. |

Signal timing optimization at all levels, should ideally begin with a review of established policies to ensure compliance and consistency of the project scope with agency policy. As described in Table 2-2 above, response to routine complaints and maintenance activity may not result in activity in each step of the process, but it is important to start with a framework that could be used during a comprehensive evaluation. Ultimately, consideration of policies must also be relevant to the specific requirements and location of each situation to consider a proper timing plan.

### 2.2.1 Signal Timing Maintenance and Data Management

The monitoring of signal timing operations and maintenance is included as the last step of the signal timing environment and can take place in a variety of ways. Proactive examples would be a signal timing policy for regular timing updates, field inspections, continual maintenance of signal systems, and communications to to identify issues as soon as possible. Reactive examples would be responding to phone calls from the public or notification from outside sources. Where possible, a proactive approach to monitoring, maintenance, and the entire signal timing process is preferred.

**A proactive approach will provide the opportunity for reliable and efficient signal timing and systems operations, particularly in capacity-constrained corridors or regions. The importance of proactive maintenance, monitoring and updates should be made clear in the minds of decisionmakers.**

An important part of the maintenance of signal timing plans and the signal system is good, clear data management. Efficiency and cost-savings will be lost if there is not good documentation or recordkeeping of the process so that it can be recreated and/or followed when future signal timing efforts are needed or when a timing setting is called into question in court proceedings. Organized and simple data management is needed to organize the collected data, track changes in the controller, cabinet or other signal infrastructure, record maintenance efforts and implementations, etc. Ideally data management should be redundant in nature to avoid the loss of data in the future. This may take the form of hard copy notes or data in files or binders, as well as electronic copies stored via computer. A little effort to organize data and efforts related to the signal timing process will have large benefits in the future.

### 2.2.2 Hardware and Software Considerations

While the policy evaluation is undertaken, it is also important to complete an assessment of the hardware and software signal system capabilities and constraints. This effort will provide the agency with an understanding of how certain hardware and/or software capabilities and constraints might enhance or limit their ability to carry out a certain policy. For example, if transit vehicles are a priority in a system or corridor, transit signal priority may be a desired element within a signal timing policy. Yet transit signal priority is a relatively new technique that may require a jurisdiction(s) to upgrade their hardware and software signal controller capabilities to be able to operate transit signal priority and allow this policy to be carried out. Within the evaluation of signal timing policies, the agency needs to consider the tools and resources necessary to collect and evaluate relevant data for optimization and necessary to complete a policy evaluation.

The age of signal equipment is typically related to hardware and software capabilities and constraints. As improved technology is introduced into the transportation industry, signal equipment capabilities increase, and constraints such as lack of memory to store timing plans and other problems decrease. In most cases, traffic controllers purchased today are contemporary electronics with central processors that provide sufficient capabilities for the operation of most strategies that can be developed. The Advanced Transportation Controller (ATC) standard (http://www.ite.org/standards/atc/) provides a hardware platform that anyone can provide software and thus affords opportunities to select from multiple vendors both hardware and software, allowing more upgrade opportunities without changing the hardware. In addition to the signal controller, the hardware at the intersection, the detection, and communication are important factors in the ability to effectively implement the policies described above. Examples of hardware and software constraints include:

- Maximum number of signal timing plans (e.g. an older controller may only be able to store and operate 5 timing plans, while newer controllers can store and operate 20 plans or more);
- How pedestrian crossing requirements will be accommodated in selecting the cycle length (with the primary risk being loss of coordination);
- How planned transitions to and from coordination (e.g., preemptions, plan changes) will be handled;
- How different types of users will be detected and accommodated;
- Capability to implement preferential treatment (preemption or priority);
- Ability of equipment to accommodate regional or corridor communication strategies (i.e. fiber, modem, radio, etc); and
- The hardware and software uses of neighboring agencies and the partnership possibilities between them (i.e. cost-sharing).

### 2.2.3 Selection of Optimization Tools and Evaluation of Policies

The data collected as a part of the signal timing process is entered into an optimization tool to evaluate the adjusted signal timing plan. Part of the evaluation should include a confirmation that signal timing policies are upheld by the new timing plans. There are a number of computer programs that can be used to generate signal timing plans. The FHWA recently published a report that describes the resources available in more detail (Traffic Analysis Toolbox, http://www.ops.fhwa.dot.gov/trafficanalysistools/toolbox.htm). The pertinent questions related to the process for optimization and use of the policy evaluation tool include:

- Is the optimization tool capable of providing an interface that resembles the hardware and software and can it reasonably replicate the performance of the equipment?
- Is the optimization tool capable of producing a plan that addresses the measures established within the evaluation criteria?
- Is there an evaluation tool (other than the optimization tool, potentially simulation) needed to assess whether the timing plan has met the relevant policies?

### 2.2.4 Other Policy Considerations

Additional policy issues that are more detail oriented include:

- The maximum allowable cycle length;
- Whether the agency will allow lagging and leading left turns by intersection or variable by time of day;
- Whether the agency will allow the skipping of left turn phases under low volume conditions;
- Whether maximum green times will operate within the coordination plan;
- Whether transit preferential policies such as transit signal priority will be implemented aggressively;
- The number of signal timing plans (time of day plans) in operation per day to respond to fluctuating traffic demand;
- Will coordination timing plans allow intersections to temporarily leave coordination to accomplish tasks (i.e. serve pedestrian calls); and
- Whether coordination patterns will be selected by time-of-day or by real-time traffic data.

Hardware and software options are becoming more complex relative to implementation of coordination plans, and it is important to understand that there are differences between signal timing plans produced by optimization tools in the office and the plans that are actually implemented in the traffic signal controller specific to a jurisdiction. In some cases, signal controller software requires a timing plan to provide sufficient time for pedestrians to cross the street for a coordination plan to take effect. While software programs have greatly reduced the effort required to install the controller parameters there is still a significant difference between field operations and the optimization packages. Some signal system vendors are integrating the ability to transfer signal timing and traffic volume data into optimization tools into their products. This further reduces the barrier between data collection, optimization, and implementation; however, this may come at the expense of an assessment of the timing policies and the flexibility of establishing the proper evaluation criteria. This integration might also limit the user to the particular program integrated into the product.

## 2.3 PERFORMANCE MEASURES AND NATIONAL PERSPECTIVE IN SIGNAL TIMING

Performance measures are the best way to gauge the effectiveness of signal timing policy and its application. Common performance measures related to signal timing include delay per person or vehicle, travel time, $50^{th}$-percentile and $95^{th}$-percentile queue lengths, and air quality or vehicle emissions measures. Presenting performance measures in a clear and understandable format is important to capturing the attention of policy-makers and decision-makers and reinforcing the importance of good signal timing application within their community(s). Specifics on performance measures are presented in Chapter 3.

### 2.3.1 National Traffic Signal Report Card

The exposure of signal timing practice has greatly increased at the local, state, and federal levels after the release of the first National Traffic Signal Report card was first released by the National Transportation Operations Coalition (NTOC) in April 2005 and most recently in October 2007. The report card was an agency self-assessment based on the five core areas for agencies and decision makers to focus on when striving for excellence in traffic signal management:

- **Program Management** – having "clearly defined goals with measureable objectives and specific milestones for achievement to hopefully resulting in improved operational performance, reliability, asset duration, and resource allocation."
- **Traffic Monitoring and Data Collection** – often underestimated in importance, "having specific, clear knowledge of conditions allows transportation professionals to be creative in signal timing solutions, by minimizing unknown variability."
- **Routine Signal Timing Updates** – "to keep pace with changing travel patterns, traffic signal timing should be actively monitored, reviewed, and updated at least every three years and possibly sooner depending on growth and changes in traffic patterns."
- **Sound Maintenance Practices** – "well-trained technicians are needed to properly maintain traffic signals and preserve the investment in hardware and timing updates."
- **Appropriate Traffic Signal Hardware** – ""to keep from using outdated equipment to operate the signal system, signal controllers (and potentially signal communication network) should be upgraded every 10 years, and possibly more frequently in high-growth areas that require more complex control."

The report card studies site numerous benefits in the form of reduced congestion, less vehicle emissions, and improved operational safety. These policies listed above are fairly universal best practice policy elements for urban and suburban environments, and many are applicable in rural environments. Establishing improved signal timing policies should lead to improved signal timing and signal operations resulting in improved report card grades.

The NTOC 2005 and 2007 National Traffic Signal Report Card concluded that many agencies do not have documented policies on how signals are operated, nor is this information shared with employees. In addition, regular meetings with law enforcement and emergency service providers, and annual reviews of major roadways are rarely conducted (*2*). The report card indicated a strong need for improvement in all five categories on average for the 417 responding agencies, with an overall national grade of "D" in the latest report card, up from a "D-" national average in 2005.

While the NTOC 2007 National Traffic Signal Report Card showed room for improvement nationally, a few agencies showed significant improvement. This improvement in general was characterized by (1) more effective management techniques, (2) a more thoughtful approach to resource allocation, (3) new or improved training for staff, and (4) improved communication between neighboring agencies and internal departments (i.e. engineering, technicians, and law enforcement), among others.

Many of the improved agencies noted the 2005 National Traffic Signal Report Card provided them validation and numbers to back up requests for additional funding and resources for signal timing projects.

> **"The agencies managing our traffic signal systems can and want to do better in the daily management of our systems, but this will be accomplished only through the support of local public sector leadership. Proactive traffic signal management, operation and maintenance are critical – our quality of life and the environment depend on it." – 2007 NTOC National Traffic Signal Report Card**

### 2.3.2 National Signal Timing Findings

Good signal timing will contribute to the safe and efficient movement of goods and persons, not just automobiles, through a region's transportation system. Unlike other transportation improvements, improved signal timing typically requires little or no infrastructure costs and produces a very high benefit to cost ratio by operating the existing system with greater efficiency and reduced congestion. The cost to retime and optimize a signalized intersection is approximately $2,500 to $3,100 per signal per update (*3*). Beyond signal timing optimization, establishing or improving signal coordination or updating signal software and hardware equipment can add further system benefit to the traveling public.

The following are some specific examples of signal timing optimization and signal system improvement benefits experienced in various communities nationwide, as documented in the United States Department of Transportation's *Intelligent Transportation Systems for Traffic Signal Control, Deployment Benefits and Lessons Learned* (*3*):

- The Traffic Light Synchronization program in Texas demonstrated a benefit-cost ratio of 62:1, with reductions of 24.6 percent in delay, 9.1 percent in fuel consumption, and 14.2 percent in stops.
- The Fuel Efficient Traffic Signal Management program in California demonstrated a benefit-cost ratio of 17:1, with reductions of 14 percent in delay, 8 percent in fuel consumption, 13 percent in stops, and 8 percent in travel time.
- Improvements to an eleven-intersection arterial in Saint Augustine, Florida, showed reductions of 36 percent in arterial delay, 49 percent in arterial stops, and 10 percent in travel time, resulting in an annual fuel savings of 26,000 gallons and a cost savings of $1.1 million.
- A project in Syracuse, New York, that connected intersections to a communications network produced reductions in travel time of up to 34 percent. Coordinated signal systems improve operational efficiency and can simplify the signal timing process.
- In areas of rapidly changing or unpredictable traffic volumes, adaptive signal timing control may improve system performance by 5 to 30 percent, but should be applied with caution to intersections or corridors.
- Spending less than 1 percent of the total expenditure on highway transportation would lead to a level of excellence in traffic signal operations. It would be an investment with a 40:1 benefit-cost ratio and would result in benefits of as much as $45 billion per year. This corresponds to a price of less than $3 per U.S. household resulting in savings of $100 per household per year. (*2, 16*)

These national signal timing statistics are meant to show a glimpse of the potential signal timing optimization benefits that can result from either implicit or explicit signal timing policy applications. Communicating these types of benefits to policy-makers and decision-makers is likely one of the best ways to bring attention and funding to signal timing and ultimately, the development of solid signal timing policies.

## 2.4 FUNDING CONSIDERATIONS

This section discusses the various funding considerations and sources for a signal timing program. Potential funding sources can include federal funding, state-local arrangements, public-private partnerships, and direct agency funding sources.

### 2.4.1 Direct Signal Timing Funding

Direct agency funding has been the most common approach to funding improved signal timing plan development and plan implementation. Depending on the local, state, or federal level of agency, direct funding sources may include (*4*):

- General Tax Fund (i.e. gas tax, license tax, property tax, sales tax);
- Tolls;

- Bond Proceeds/Interest Income;
- Federal Aid (i.e. CMAQ, Surface Transportation Program, SAFETEA-LU), and
- Developer Mitigation Funds or Impact Fees.

Funding for signal timing projects is available through federal or other sources. For federal funds, the project sponsor must communicate the need for the project with either the State department of transportation or local metropolitan planning organization to determine whether the program fits into the local long-range transporation plan and whether the current transportation improvement program can be amended, if necessary, to include the project. If not, the project sponsor must work with the local or state representatives to make sure the project makes it into the next transportation improvement program update cycle. The planning partners' responsibility is to help qualifying projects obtain funding from either federal or other sources (*3*).

Federal funds through the FWHA Surface Transportation Program and CMAQ can be used for project, operations, and maintenance costs of a traffic control system and are eligible for Federal reimbursement through the Federal Aid Program. This aid is distributed to project sponsors in the following ways (*3*):

1. Transportation Improvement Program (TIP) and State Transportation Improvement Program (STIP). Program funds are typically available for capital improvement projects requiring new or reconstructed infrastructure. The installation of traffic signal systems and traffic control centers will usually be funded under these programs. These funds are programmed into the State's Long Range Transportation Plan (LRTP).

2. Congestion Mitigation and Air Quality (CMAQ) Improvement Program funds. For projects located in air quality non-attainment and maintenance areas, CMAQ funds may be used for operating costs for a 3-year period as long as those systems measurably demonstrate reductions in air quality emissions or increased air quality benefits. Typical eligible operating costs include labor costs, administrative costs, costs of utilities and rent, and other costs associated with the continuous operation of the system, such as costs for system maintenance.

### 2.4.2 Partnerships for Funding Signal Timing

Beyond a single agency funding source, partnerships play a key role in maximizing the benefits of signal timing projects and fund sharing between agencies. In some cases, public-private entities may offer investments to be used more efficiently and effectively. Good examples of this partnering for funding include Oakland County, Michigan’s Signal Retiming & Maintenance Agreements (*5, 6)* and the City of Portland, Oregon’s partnership with the Climate Trust (*7*).

In 2002, the Road Commission for Oakland County, Michigan, began a program to retime nearly 900 signals, in three manageable phases. These phases crossed jurisdictional and public-private boundaries, allowing for the best possible signal timing for the roadway network as a system. Partnerships in this Oakland County effort were formed between the Michigan Department of Transportation (MDOT); the Southeast Michigan Council of Governments (SEMCOG); Wayne County; the Road Commission of Macomb County; the cities of Ferndale, Pontiac, and Royal Oak; and regional consulting engineers. Notable reductions in delay, emissions, and improved travel times were the greatest user benefits measured through this signal timing and implementation partnership, which benefits all agencies involved. Funding for this signal timing effort came through a cost-sharing effort between MDOT CMAQ funds (1/3 of costs) and SEMCOG CMAQ funds (2/3 of costs).

The City of Portland, Oregon’s funding partnership with the Climate Trust is another example of partnering with multiple agencies to achieve a mutually beneficial goal of retiming 170 metro-area traffic signals. The partnership included the Climate Trust, a non-profit organization to support climate change solutions to offset greenhouse gases, because of their vested interest in the environmental improvements associated with good signal timing; the various agency partners (City of Portland, Washington County, and the Oregon Department of Transportation) have a vested interest in improving the efficiency of their roadway systems through signal retiming. A shortfall of available public funds to support signal retiming and implementation led to a partnership with the Climate Trust to make use of their available carbon dioxide offset funding in this signal retiming effort.

According to the Chicago Area Transportation Study website (*8*), their CMAQ program finances three types of traffic flow improvements—bottleneck elimination, intersection improvements, and signal interconnection—that all impact signal timing plans, but do not directly fund signal timing plan development. Signal timing improvements could qualify as “other projects” for CMAQ funding because improved signal timing reduces stops and idling traffic in a network, which results in emission reductions that could be estimated.

A unique signal timing partnership offer is available in the San Francisco Bay Area through the Metropolitan Transportation Commission’s (MTC) Regional Signal Timing Program. This program invites local cities and counties to apply for “assistance from MTC's consultants for development and implementation of new time-of-day traffic signal coordination plans for weekday peak periods. The budget is $1.5 million in federal funds, with which about 650 traffic signals may be retimed. MTC will provide the local matching funds. In previous cycles, all applications that met eligibility requirements were funded. Other public agencies, such as congestion management agencies or transit agencies, are also eligible to apply if authorized to act on behalf of the agencies that operate traffic signals within the project limits (*9*).”

Places to start for securing funding for signal timing projects could be through a local MPO planning office. Traffic signal retiming projects might be added to their Long Range Transportation Plan (LRTP) as well as their more immediate Transportation Improvement program (TIP). For locations outside of MPO areas you can check with your local state DOT office to possibly add this work to their transportation plan.. For more information on federal level funding, contact the State FHWA Division Office. Reaching out to regional partners, as shown in the examples above, can be a great step towards securing funding for signal timing projects that can be mutually beneficial (*3*).

## 2.5 EXAMPLES OF PROGRAMS

The following section provides several examples of various signal timing policy applications corresponding to the array of modal objectives.

**City of Denver, Colorado.** The fundamental part of the Traffic Signal System Improvement Program (TSSIP) is development of new signal timing plans at a regional level in a three- to five-year cycle for major corridors and for all capital projects implemented (including miscellaneous signal equipment purchases). Since 1994, across 14 operating agencies in the region, the programming of TSSIP projects has been completed with regional cooperation and consideration for equitable distribution of resources from federal, state, and regional entities. Because the TSSIP is funded with federal Congestion Mitigation Air Quality funds, the benefits of every project must be measured and reported. The workgroup also targeted some additional funds for selectively developing timing plans that address weekend traffic patterns. Specific criteria would be developed to guide the selection of which weekend timing plans the TSSIP would prepare. Development and evaluation of timing plans for traffic-responsive control and incident management test bed activities are also included in this activity, as is assessment of the transit signal priority projects. (*15*)

**City of Philadelphia, Pennsylvania**. The City currently has informal agreements with two neighboring townships, Upper Darby and Springfield, to provide arterial signal coordination across jurisdictional boundaries, which results in improved operations for all users. (*14*)

**City of Portland, Oregon.** The City of Portland incorporates many travel modes within their signal systems. The expectations vary based on the surrounding areas and type of users. Within the downtown core, mobility for pedestrians, buses, bicycle, and train modes are the main focus. The coordinated system allows the transit modes to travel with higher priority than automobiles. The signals are timed to keep travel speeds relatively low. The low speeds and short cycle lengths allow a safer environment for pedestrians. Outside of the downtown core area, the coordinated system changes the focus to allow higher mobility for autos while still accommodating transit, emergency vehicles, bicycle, and pedestrian users. Transit signal priority is implemented along a majority of bus routes and emergency pre-emption is installed at all signalized intersections. Bicyclists are often provided with bike lanes in the streets and at some locations a bicycle signal is installed to provide them with additional mobility. Pedestrians are also considered within the signal system with pedestrian recall parameters. The networks set the coordination reference point at the beginning of the “Flashing Don’t Walk” time to ensure sufficient pedestrian service on the coordinated phases and to provide a “rest in Walk” operation that benefits pedestrians. In addition, an exclusive pedestrian phase is provided at high pedestrian areas.

**City of San Francisco, California.** The Bicycle Advisory Committee in San Francisco, California, has brought forth several unique signal timing policy suggestions to aid bicyclists. Because effective signal timing can improve traffic flows and reduce stops and delay, for bicycles, the Bicycle Advisory Committee recommends that the City consider "timing the signals along bike routes for bicycle speeds of approximately 12 to 15 miles per hour." In addition, minimum green times and red clearance intervals should take the bicyclist into consideration because bicyclists need more start-up time to get through an intersection. They are recommending that the City actuate signal timing to include "at least 8 seconds" of minimum green time and more if the grade is uphill. Along major thoroughfares, particularly with widths greater than 75 feet, they say "red clearance intervals should be provided to allow time for the bicyclist to clear the intersection before cross-traffic is given a green indication." (*11, 12, 13*)

**City of Vancouver, British Columbia.** The Vancouver Transportation Plan contains a number of policies that are intended to improve the comfort and convenience of pedestrians in the City. One proposed measure is to reduce pedestrian wait times at traffic signals. Requests to reduce pedestrian wait times at signals were received during the public consultation phase of the Transportation Plan. The Transportation Plan recommended this reduction. The issue is also emerging through City Plan Community Visions. A related issue is that pedestrians (or cyclists) are more likely to ignore the signal if wait times are perceived as too long. This can cause pedestrians to cross against the light and result in a signal changing with no pedestrians (or cyclists) crossing. This also affects transit users trying to catch or transfer between buses, as buses are also less likely to wait for passengers if pedestrian/bike crossing wait times are perceived as too long. (*10*)

## 2.6 REFERENCES


1. Metro: Regional Transportation Plan, http://www.metro-region.org/article.cfm?ArticleID=137, visited January 2, 2007.

2. National Transportation Operations Council, "National Traffic Signal Report Card, Executive Summary", http://www.ite.org/reportcard/NTS_ExecSummary.pdf, visited February 8, 2007

3. United States Department of Transportation, "Intelligent Transportation Systems for Traffic Signal Control, Deployment Benefits and Lessons Learned." Project number FHWA-JPO-07-004, January 2007.

4. "Show me the Money, A Decision-Maker's Funding Compendium for Transportation Systems Management and Operations." Public Technology Institute, National Transportation Operations Coalition, US Department of Transportation, and Federal Highway Administration, December 2005. http://www.pti.org/docs/Trans_Funding_PTI_FHWA_2006.pdf

5. "Regional Concept for Transportation Operations (RCTO) Stakeholders Meeting, Signal Timing Projects, Oakland County, Michigan." SEMCOG, June 21, 2006. http://www.semcog.org/TranPlan/Operations/assets/RCTO_SignalRetimingProjects_Oakland County.ppt#256,1,Regional Concept for Transportation Operations (RCTO) Stakeholders Meeting. Signal Retiming Projects Oakland County, Michigan

6. Halkias, John, and Michael Schauer. "Red Light, Green Light." USDOT, Federal Highway Administration. Public Roads, November/December 2004 Vol. 68, No. 3. http://www.tfhrc.gov/pubrds/04nov/07.htm

7. "Traffic Signals Optimization." Climate Trust Website, http://www.climatetrust.org/offset_traffic.php

8. "Congestion Mitigation and Air Quality Program (CMAQ)." Chicago Area Transportation Study website, http://www.catsmpo.com/prog-cmaq.htm

9. "2007 Regional Signal Timing Program." San Francisco County Transportation Authority website, http://www.sfcta.org/content/view/255/8

10. City of Vancouver, British Columbia, "Policy Report, Traffic and Transit." June 5, 2001. http://www.city.vancouver.bc.ca/ctyclerk/cclerk/010605/tt3.htm

11. "Signal Clearance Timing for Bicyclists", Alan Wachtel, John Forester, Gary Foxen and David Pelz.
12. City of San Francisco, California, "Bicycle Plan - Part 5, Recommended Design Standards." http://www.ci.sf.ca.us/site/bac_page.asp?id=11542#P102_12565

13. "Signal Clearance Timing for Bicyclists", Alan Wachtel, John Forester, Gary Foxen and David Pelz.
14. "Cross-Jurisdictional Signal Coordination, Case Studies." Federal Highway Administration, USDOT, February 2002. http://www.itsdocs.fhwa.dot.gov/jpodocs/repts_te/13613.html#_Toc658840
15. "Traffic Signal System Improvement Program, 2003 Update." Denver Regional Council of Governments, July 16, 2003. http://www.drcog.org/documents/TSSIP_Report.pdf
16. Institute of Transportation Engineers, "2007 National Traffic Signal Report Card", http://www.ite.org/reportcard/, visited 1/13/08.

# CHAPTER 3

# OPERATIONAL AND SAFETY ANALYSIS

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# 3.0 OPERATIONAL AND SAFETY ANALYSIS

The purpose of this chapter is to summarize some of the common techniques used to assess the operational and safety performance of signal timing. The chapter begins by presenting an overview of the characteristics that affect signal timing, including both system and user characteristics. It then presents discussions of operational and safety performance measures and techniques to evaluate those performance measures. Finally, the chapter presents a discussion of signal warrants as presented in the Manual on Uniform Traffic Control Devices and how those warrants relate to signal timing.

## 3.1 TERMINOLOGY

This section identifies and describes basic terminology used within this chapter. Additional terms can be found in the Glossary section of the Manual.

*Capacity*
The maximum rate at which vehicles can pass through the intersection under prevailing conditions.

*Clearance lost time*
The time, in seconds, between signal phases during which an intersection is not used by any critical movements.

*Control Delay*
The amount of additional travel time experienced by a user attributable to a control device.

*Critical movement analysis*
A simplified technique for estimating phasing needs and signal timing parameters.

*Delay*
The additional travel time experienced by a driver, passenger, or pedestrian.

*Effective green time*
The time during which a given traffic movement or set of movements may proceed; it is equal to the cycle length minus the effective red time.

*Flow rate*
The equivalent hourly rate at which vehicles, bicycles, or persons pass a point on a lane, roadway, or other trafficway, computed as the number of vehicles, bicycles, or persons passing the point, divided by the time interval (usually less than 1 hour) in which they pass; expressed as vehicles, bicycles, or persons per hour.

*Level of service*
A qualitative measure describing operational conditions within a traffic stream, based on service measures such as speed and travel time, freedom to maneuver, traffic interruptions, comfort, and convenience.

*Lost Time*
The portion of time at the beginning of each green period and a portion of each yellow change plus red clearance period that is not usable by vehicles.

*Saturation Flow Rate*
The equivalent hourly rate at which vehicles can traverse an intersection approach under prevailing conditions, assuming a constant green indication at all time and no loss time, in vehicles per hour or vehicles per hour per lane.

*Start-up lost time*

The additional time, in seconds, consumed by the first few vehicles in a queue at a signalized intersection above and beyond the saturation headway due to the need to react to the initiation of the green phase and to accelerate to a steady flow condition.

*Stopped Delay*

A measurement of the aggregate sum of stopped vehicles for a particular time interval divided by the total entering volume for that movement.

*Total delay*

The sum of all components of delay for any lane group, including control delay, geometric delay, and incident delay.

*Travel Time (Average)*

The total elapsed time spent traversing a specified distance. The average travel time represents an average of the runs for a particular link or corridor.

## 3.2 CHARACTERISTICS AFFECTING SIGNAL TIMING

Several overall features affect implementation of signal timing including:

- Location
- Transportation network characteristics
- Intersection geometry
- User characteristics

The following sections further describe many characteristics and dynamic nature influencing signal timing.

### 3.2.1 Location

One of the primary factors affecting overall signal timing is the environment in which the intersection or intersections being timed are located. Urban environments are frequently characterized by lower speeds and higher degrees of congestion. In addition, urban environments are frequently characterized by higher pedestrian, cyclist, and transit use that often require priority in consideration. Rural environments, on the other hand, are typically higher speed but with lower levels of traffic volumes and fewer, if any, pedestrians, cyclists, and transit vehicles. As a result, signal timing for rural environments is typically dominated by efforts to safely manage high speed approaches; capacity is seldom a constraint. Suburban environments often present a challenging mix of these characteristics. Suburban environments are often characterized by high speeds during the off-peak periods and capacity-constrained conditions during the peak periods. This requires a careful consideration and balance of both safety aspects and operational efficiency.

### 3.2.2 Transportation Network Characteristics

The configuration of the transportation network under consideration can have a significant impact on the way its traffic signals are timed. Isolated intersections can be timed without the explicit consideration of other traffic signals, allowing the flexibility to either set or target cycle lengths that are optimal for the individual intersection. In these cases, good detection design often yields measurable operational and safety benefits. These are discussed further in Chapters 4 and 5.

For intersections located along arterial streets, isolated operation can often be improved by considering coordination of the major street movements along the arterial. Common cycle lengths are often employed to facilitate this coordination. Coordinated operations are discussed in detail in Chapter 6.

Signalized intersections are often located in grid networks with either crossing arterials or a series of intersecting streets with comparable function and traffic volumes. In these cases, the entire network is often timed together to ensure consistent behavior between intersections. Grid networks, particularly downtown environments with short block spacing, are frequently timed using fixed settings and no detection.

Within these categories, the spacing of signalized intersections affects how one times the signals. For signals that are sufficiently far apart that they can be considered independent of one another, intersections may be operated freely without need for or benefit from coordination, depending on the degree of congestion on the facility. For most arterial streets with signal spacing between 500 feet and 0.5 mile (2,640 feet), coordinated operation can often yield benefits by improving progression between signals. On arterials with higher speeds, it can be beneficial to coordinate signals spaced a mile (5,280 feet) apart or even longer. Signals that are located very close together (less than 500 feet) often require settings that manage queues rather than progression as the dominant policy. It may also be beneficial to operate two intersections with very close spacing with a single controller.

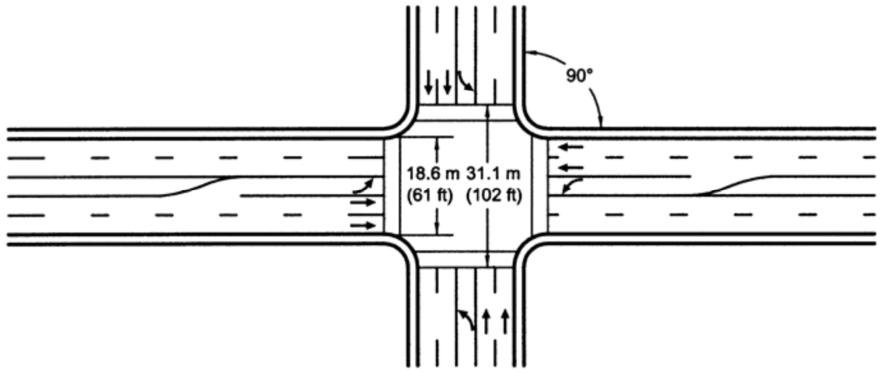
### 3.2.3 Intersection Geometry

The overall geometry of an intersection determines its ability to efficiently and safely serve user demand. Pedestrians are often crossing lanes of traffic, whereas transit, bicycles, and vehicular traffic are using the travel lanes provided at the intersection. The number of lanes provided for each approach has a significant impact on the capacity of the intersection and, therefore, the ability for signal timing to efficiently serve demand. For example, a movement served by two lanes rather than one has a higher capacity and thus requires less green time to serve demand. However, increasing the number of lanes on a particular leg of the intersection also increases the minimum pedestrian crossing time across that leg, which by increasing clearance times will offset some of the increase in capacity.

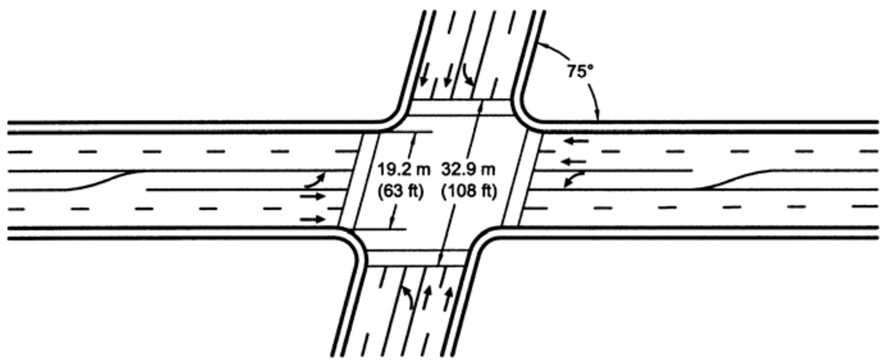
Subtle details of intersection geometry have a significant influence on signal timing. The size and geometry of the intersection, coupled with the speed of approaching vehicles and walking pedestrians, affect vehicular and pedestrian clearance intervals, which in turn have an effect on the efficiency of the intersection's operation. For example, the skew of an intersection (the angle at which two roadways intersect) influences the length of crosswalks and thus pedestrian clearance time. The crosswalk length for an intersection at right angles (90 degrees) with a crosswalk length of 61 feet can extend to 76 feet if the intersection is skewed to 60 degrees, as shown in Figure 3-1 below. If a pedestrian walking speed of 4.0 ft/s, for example, is used for pedestrian clearance interval timing, the skew increases the pedestrian clearance interval from 16 seconds to 19 seconds; for a pedestrian walking speed of 3.5 ft/s, the clearance interval increases from 18 seconds to 22 seconds. Reverse effects can be seen when smaller curb radii and/or curb extensions are used to shorten pedestrian crosswalks, thus reducing pedestrian clearance time requirements (*1*). Although curb extensions can shorten pedestrian crossing times, the resulting narrow width may restrict the ability for through vehicles in shared lanes to bypass right-turning vehicles that are waiting for pedestrians or left-turning vehicles waiting for gaps in opposing traffic.

Regardless of the angle of skew, care should be taken to ensure that good visibility of the signal heads is maintained for approaching vehicles.

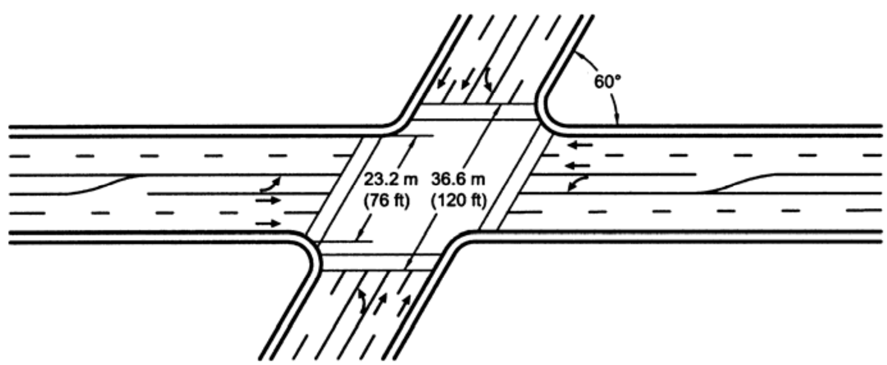
**Figure 3-1 Effect of intersection skew on signalized intersection width**




(a) Intersection skew at 90 degrees.




(b) Intersection skew at 75 degrees




(c) Intersection skew at 60 degrees.
Source: Figure 15, FHWA *Signalized Intersections: Informational Guide*

### 3.2.4 User Characteristics

User characteristics clearly influence the effectiveness of signal timing and should be accounted for early in the planning and analysis process. Some of the important factors include the following:

- *Mix of users:* The mix of users at an intersection has a significant influence on signal timing. Pedestrians with slower walking speeds, persons using wheelchairs, and pedestrians with visual impairments need more time to cross the street; pedestrian walk times and clearance intervals need to be adjusted accordingly. High bicycle use may benefit from special bicycle detection and associated bicycle minimum green timing. Emergency vehicle and/or transit use may justify the use of preemption and/or priority. Truck traffic requires accounting for reduced performance (longer acceleration and deceleration times) and larger size of heavy vehicles.

- *User demand versus measured volume:* Traffic demand represents the arrival pattern of vehicles at an intersection (or the system, if one considers a group of intersections together), while traffic volume is the measured departure rate from the intersection. If more vehicles arrive for a movement than can be served, the movement is considered to be operating over capacity (oversaturated). However, unless the analyst has measured demand arriving at the intersection through either queue observation or through measurement of departure rates from an upstream undersaturated intersection, the true demand at an intersection may be unknown. This can cause problems when developing signal timing plans for a given intersection, as one may add time to a given movement, only to have it used up by the latent demand for that movement. Traffic volume at an intersection may also be less than the traffic demand due to an overcapacity condition at an upstream signal that “starves” demand at the subject intersection. These effects are often best analyzed using microsimulation.

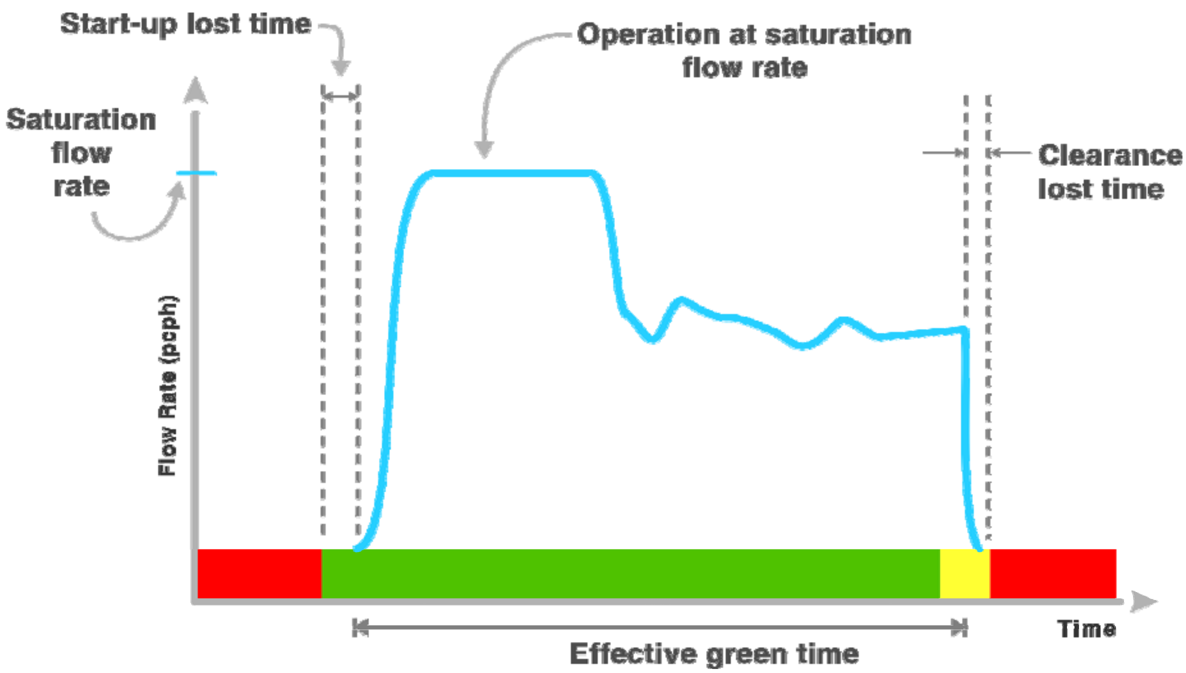
## 3.3 CAPACITY AND CRITICAL MOVEMENT ANALYSIS

An important principle behind effective signal timing is a basic understanding of how signal timing affects the capacity of the intersection. Capacity is discussed in detail in the *Highway Capacity Manual* (*2*), but often only the basic elements are particularly relevant to practical signal timing implementation. This section presents a discussion of the operation of a signalized intersection movement, the definition of capacity and its component elements, and a discussion of techniques to estimate capacity.

### 3.3.1 Basic Operational Principles

The basic operation of vehicular movement through a signalized intersection is presented in Figure 3-2 below. The signal display is presented on the horizontal axis, the instantaneous flow of vehicles on the vertical axis. During the time while the movement is receiving a red indication, vehicles arrive and form a queue, and there is no flow. Upon receiving a green indication, it takes a few seconds for the driver of the first vehicle to recognize that the signal has turned green and to get the vehicle in motion. The next few vehicles also take some time to accelerate. This is defined as the *start-up lost time* or start-up delay and is commonly assumed to be approximately 2 seconds. After approximately the fourth vehicle in the queue, the flow rate tends to stabilize at the maximum flow rate that the conditions will allow, known as the saturation flow rate. This is generally sustained until the last vehicle in the queue departs the intersection. Upon termination of the green indication, some vehicles continue to pass through the intersection during the yellow change interval; this is known as yellow extension. The usable amount of green time, that is, the duration of time between the end of the start-up delay and the end of the yellow extension, is referred to as the effective green time for the movement. The unused portion of the yellow change interval and red clearance interval is called *clearance lost time*.

**Figure 3-2 Typical flow rates at a signalized movement.**




### 3.3.2 Saturation Flow Rate

The saturation flow rate, *s*, is an important parameter for estimating the performance of a particular movement. Saturation flow rate is simply the headway in seconds between vehicles moving from a queued condition, divided into 3600 seconds per hour. For example, vehicles departing from a queue with an average headway of 2.2 seconds have a saturation flow rate of 3600 / 2.2 = 1636 vehicles per hour per lane. Saturation flow rate for a lane group is a direct function of vehicle speed and separation distance. These are in turn functions of a variety of parameters, including the number and width of lanes, lane use (e.g., exclusive versus shared lane use, aggregated in the HCM as lane groups), grades, and factors that constrain vehicle movement such as presence or absence of conflicting vehicle and/or pedestrian traffic, on-street parking, and bus movements. As a result, saturation flow rates vary by movement, time, and location and commonly range from 1,500 to 2,000 passenger cars per hour per lane (*2*). The HCM provides a series of detailed techniques for estimating and measuring saturation flow rate. It should be noted that this is significantly different than the ideal saturation flow rate, which is typically assumed to be 1,900 passenger cars per hour per lane.

The ideal saturation flow rate may not be achieved (observed) or sustained during each signal cycle. There are numerous situations where actual flow rates will not reach the average saturation flow rate on an approach including situations where demand is not able to reach the stop bar, queues are less than five vehicles in a lane, or during cycles with a high proportion of heavy vehicles. To achieve optimal efficiency and maximize vehicular throughput at the signalized intersection, traffic flow must be sustained at or near saturation flow rate on each approach. In most HCM analyses, the value of saturation flow rate is a constant based on the parameters input by the user, but in reality, this is a value that varies depending on the cycle by cycle variation of situations and users.

The HCM provides a standardized technique for measuring saturation flow rate. It is based on measuring the headway between vehicles departing from the stop bar, limited to those vehicles between the fourth position in the queue (to minimize the effect of startup lost time) and the end of the queue. The detailed procedure can be found in Chapter 16 of the HCM. For signal timing work, it is often not necessary to place heavy emphasis on this parameter due to the high degree of fluctuation in this parameter from cycle to cycle.

### 3.3.3 Lost Time

As noted in the previous section, a portion of the beginning of each green period and a portion of each yellow change plus red clearance period is not usable by vehicles. The sum of these two periods comprises the *lost time* (or *loss time*) for the phase. This value is used in estimating the overall capacity of the intersection by deducting the sum of the lost times for each of the critical movements from the overall cycle length. The HCM defines a default value of 4 seconds per phase for total lost time (the sum of start-up lost time and clearance lost time).

The resulting effective green time can therefore be defined as follows in Equation 3-1:

$$g = G + Y + R - (l_1 + l_2) \tag{3-1}$$

where $g$ is the effective green time, $G$ is the actual green interval, $Y$ is the actual yellow change interval, $R$ is the actual red clearance interval, $l_1$ is the start-up lost time, and $l_2$ is the clearance lost time (all values in seconds).

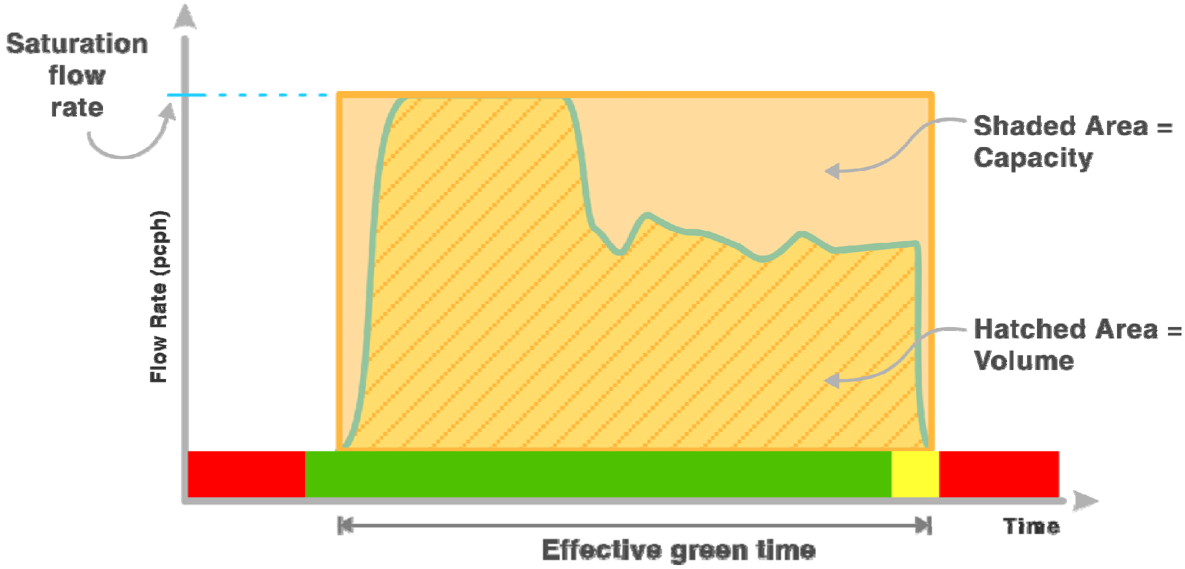
### 3.3.4 Capacity

At signalized intersections, capacity for a particular movement is defined by two elements: the maximum rate at which vehicles can pass through a given point in an hour under prevailing conditions (known as saturation flow rate), and the ratio of time during which vehicles may enter the intersection. These are shown in Equation 3-2 (*2*).

$$c = s\left(\frac{g}{C}\right) \tag{3-2}$$

where $c$ is the capacity, $s$ is the saturation flow rate of the lane group in vehicles per hour, $g$ is the effective green time for the movement in seconds, and $C$ is the cycle length in seconds. Capacity is shown graphically in Figure 3-3 as the area bounded by saturation flow rate and effective green time, and volume is shown as the area under the flow rate curve.

**Figure 3-3 Illustration of volume and capacity of a signalized movement**




For the purposes of signal timing, the factors within a practitioner's control that directly influence capacity are somewhat limited. In most cases, the number, width, and assignment of lanes are fixed, as are grades and the presence of on-street parking and bus movements. The practitioner has some control over the effect of conflicting vehicle and/or pedestrian traffic through the use of permissive versus protected left-turn and right-turn phasing and/or exclusive pedestrian phases. For signal timing purposes, the most readily accessible parameter is the effective green time for a particular movement.

### 3.3.5 Volume-to-Capacity Ratio

The volume-to-capacity ratio, also known as the v/c ratio or the degree of saturation, is calculated for each movement using Equation 3-3:

$$v/c = \frac{v}{s\left(\frac{g}{C}\right)} = \frac{vC}{sg}, \qquad (3\text{-}3)$$

where $v$ is the demand volume of the subject movement in vehicles per hour and the remaining variables are as defined previously. Using the graphical tool from the previous section, the volume-to-capacity ratio represents the proportion of the area defining capacity that is occupied by volume.

Movements or lane groups with volume-to-capacity ratios less than 0.85 are considered undersaturated and typically have sufficient capacity and stable operations. For movements or lane groups with a volume-to-capacity ratio of 0.85 to 1.00, traffic flow becomes less stable due to natural cycle-to-cycle variations in traffic flow. The closer a movement is to capacity, the more likely that a natural fluctuation in traffic flow (higher demand, large truck, timid driver, etc.) may cause the demand during the cycle to exceed the green time for that cycle. The result would be a queue that is carried over to the next cycle, even though the overall demand over the analysis period is below capacity. In cases where the projected volume-to-capacity ratios exceed 1.00 (demand exceeding capacity) over the entire analysis period, queues of vehicles not served by the signal each cycle are likely to accumulate and either affect adjacent intersections or cause shifts in demand patterns. These conditions are described as *oversaturated* and require significantly different approaches for signal timing.

### 3.3.6 Critical Movement Analysis

A variety of analysis procedures, ranging from simple to complex, are used to evaluate signalized intersection performance. These are summarized in Figure 3-4.

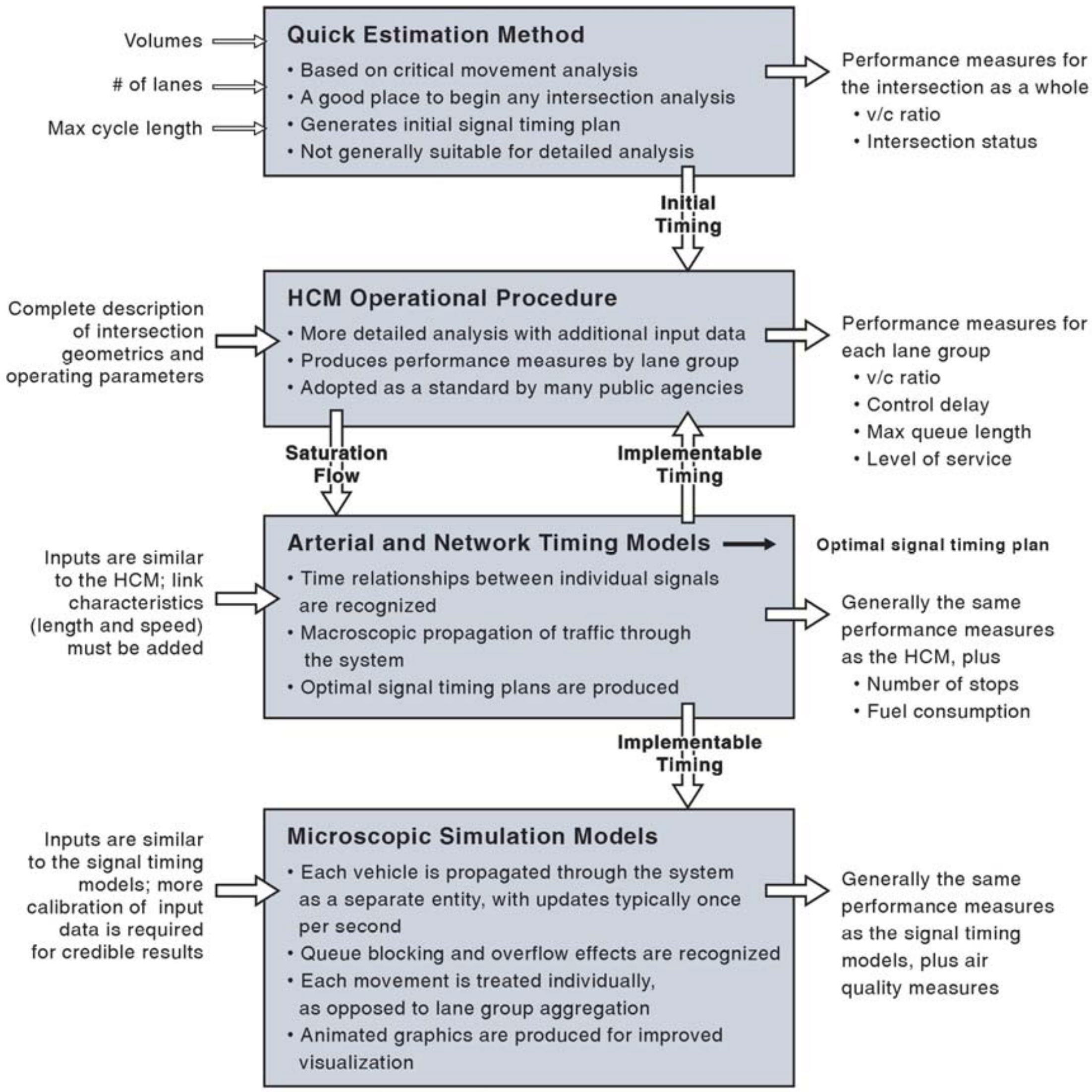
Critical movement analysis is a simplified technique that has broad application for estimating phasing needs and signal timing parameters. The most current implementation of the critical movement analysis method is provided in Chapter 10 of the HCM 2000 as the Quick Estimation Method. This method allows an analyst to identify the critical movements at an intersection; estimate whether the intersection is operating below, near, at, or over capacity; and approximate the amount of green time needed for each critical movement. The method is generally simple enough to be conducted by hand, although some of the more complicated refinements are aided considerably with the use of a simple spreadsheet. In many cases, the lack of precision of future volume forecasts or estimated trip generation of land uses minimizes the precision of more complicated analysis methods; in these cases, critical movement analysis is often as precise as one can achieve.

Critical movement analysis is based on the following fundamental basic principle:

> **The amount of time in an hour is fixed, as is the fact that two vehicles (or a vehicle and a pedestrian) cannot safely occupy the same space at the same time. Critical movement analysis identifies the set of movements that cannot time concurrently and require the most time to serve demand.**

Critical movement analysis is an effective tool to quickly estimate green times for various movements at an intersection and to estimate its overall performance in terms of volume-to-capacity ratios. Appendix A provides step-by-step details for using the procedure, including a sample worksheet. The basic procedure for conducting a critical movement analysis/quick-estimation method analysis is given in Figure 3-5 and Table 3-1. Table 3-2 identifies the various thresholds recommended in the HCM for volume-to-capacity ratios.

**Figure 3-4 Overview of signalized intersection analysis methods**




Source: (*1*)

A simple example is provided in Example 3-1 to illustrate the basic elements of the procedure.

***Application***

From a signal timing perspective, the volume-to-capacity ratio is an important measure that defines how well progression can be achieved. For coordinated through movements, a volume-to-capacity ratio approaching capacity means that virtually all vehicles departing on green will be departing from a queue at the saturation flow rate. This type of operation is efficient in terms of maximizing the throughput of an intersection, but it greatly limits the ability to have vehicles arrive and depart on green without stopping (a typically desirable objective of progression). Therefore, if one desires to minimize stops for through vehicles along an arterial, the volume-to-capacity ratio for through movements must be kept sufficiently low (this has not be documented, but it is likely to be around 0.85 or lower) to allow a portion of the through movement's green time to remain undersaturated. If overall volume levels are too high to permit a large enough undersaturated period for the coordinated through movements to pass through without stopping, the objective of progression through the intersection without stopping may be infeasible.

### *Limitations*

The critical movement analysis procedure is simple and cannot accommodate all real-world conditions encountered when developing signal timing. These include vehicle and pedestrian minimum times (as noted above), assumed constant values of capacity for each lane, and complex signal phasing. For these conditions, more advanced analysis methods are likely to be more accurate. However, the critical movement analysis procedure is still often a good first approximation.

**Figure 3-5 Graphical summary of Critical Movement Analysis/Quick Estimation Method (*1*)**

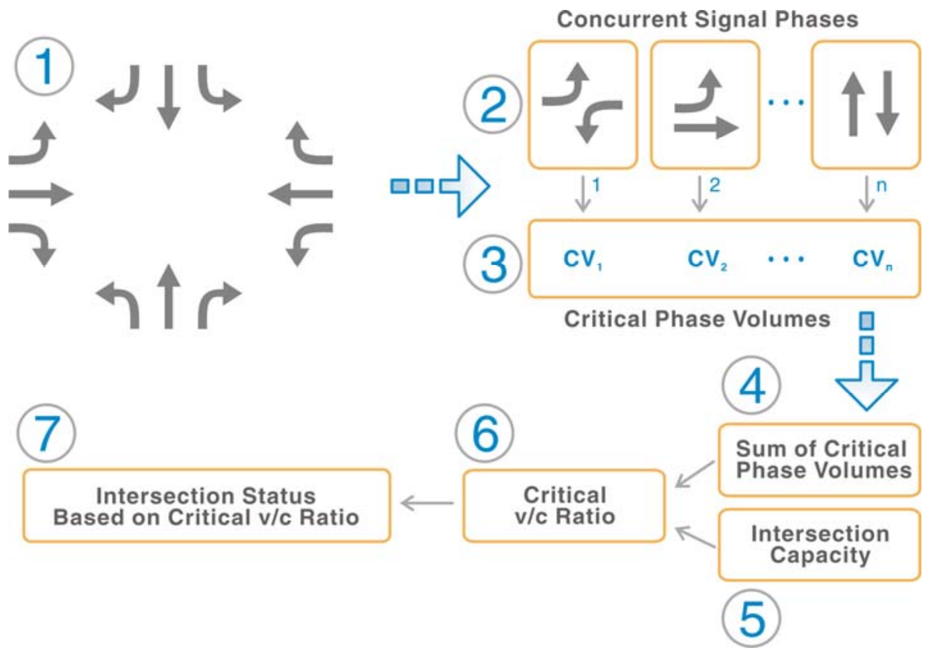


**Table 3-1 Steps of the Quick Estimation Method (QEM) (2, with HCM corrections)**

| Step | Process |
|---|---|
| 1 | Identify movements to be served and assign hourly traffic volumes per lane. This is the only site-specific data that must be provided. The hourly traffic volumes are usually adjusted to represent the peak 15-minute period. The number of lanes must be known to compute the hourly volumes per lane. |
| 2 | Arrange the movements into a desired sequence of phases that can be run concurrently based on the design of the signal. This is based in part on the treatment of each left turn (protected, permissive, etc.). |
| 3 | Determine the critical volume per lane that must be accommodated during each interval. This step determines which movements are critical. The critical movement volume determines the amount of time that must be assigned to the phase on each signal cycle. |
| 4 | Sum the critical phase volumes to determine the overall critical volume that must be accommodated by the intersection. This is a simple mathematical step that produces an estimate of how much traffic the intersection needs to accommodate. |
| 5 | Determine the maximum critical volume that the intersection can accommodate: This represents the overall intersection capacity. The HCM QEM suggests 1,530 vphpl for most purposes. |
| 6 | Determine the critical volume-to-capacity ratio, which is computed by dividing the overall critical volume by the overall intersection capacity, after adjusting the intersection capacity to account for time lost due to starting and stopping traffic on each cycle. The lost time will be a function of the cycle length and the number of protected left turns. |
| 7 | Determine the intersection status from the critical volume-to-capacity ratio. The status thresholds are given in Table 3-2. |

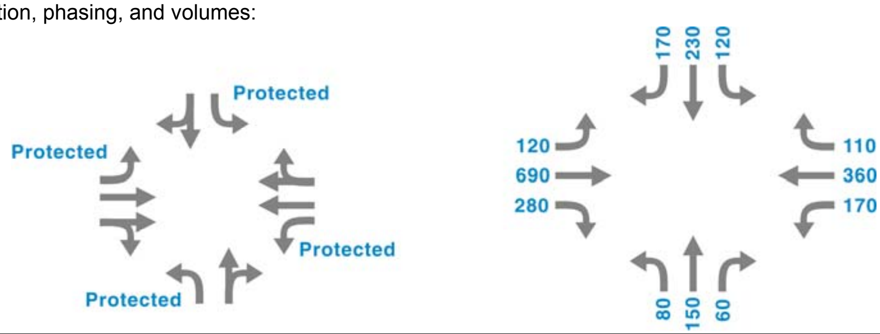
**Table 3-2 Volume-to-capacity ratio threshold descriptions for the Quick Estimation Method**

| Critical Volume-to-Capacity Ratio | Assessment |
|---|---|
| < 0.85 | Intersection is operating under capacity. Excessive delays are not experienced. |
| 0.85–0.95 | Intersection is operating near its capacity. Higher delays may be expected, but continuously increasing queues should not occur. |
| 0.95–1.00 | Unstable flow results in a wide range of delay. Intersection improvements will be required soon to avoid excessive delays. |
| > 1.00 | The demand exceeds the available capacity of the intersection. Excessive delays and queuing are anticipated. |

Source: (*2*)

**Example 3-1: Critical Movement Analysis**

Step 1:
Lane configuration, phasing, and volumes:




Steps 2, 3, and 4:
Check east-west critical lane volumes to determine the pair of consecutive movements that requires the most time:
EB LT + WB TH/RT = 120 + (360 + 110)/2 = 120 + 235 = 355
WB LT + EB TH/RT = 170 + (690 + 280)/2 = 170 + 485 = 655 ← critical

Check north-south critical lane volumes to determine the pair of consecutive movements that requires the most time:
NB LT + SB TH/RT = 80 + 400 = 480 ← critical
SB LT + NB TH/RT = 120 + 210 = 330

Identify critical movements: WB LT, EB TH/RT, NB LT, SB TH/RT
Determine the sum of critical movement volumes: $CS$ = 655 + 480 = 1135

Steps 5, 6, and 7:
Determine the reference sum (capacity): $RS = 1530 \times PHF \times f_a$ = 1530 × 1.00 × 1.00 = 1530
$PHF$ = Peak Hour Factor (assumed to be 1.00 in this example)
Adjustment factor for area type, $f_a$ = 0.90 for Central Business District, 1.00 for other (assumed to be “other” in this example)
Identify cycle length: $C$ = 120 s (assumed)
Determine total lost time: $L$ = 4 critical phases × 4 s per phase = 16 s

Calculate and assess intersection volume-to-capacity ratio:

$$X_{cm} = \frac{CS}{RS\left(1-\frac{L}{C}\right)} = \frac{1135}{1530\left(1-\frac{16}{120}\right)} = 0.86 \rightarrow \text{Near capacity}$$

Additional calculations:
Estimate splits and green times for critical movements, assuming equal volume-to-capacity ratio for all movements:

| Critical Movement | Volume (veh/hr/ln) | Proportion Split | Average Split Time (s) | Assumed Y + R (s) | Average Green Time (s)* |
|---|---|---|---|---|---|
| WB LT | 170 | 15% | 18 | 4 | 14 |
| EB TH/RT | 485 | 43% | 52 | 4 | 48 |
| NB LT | 80 | 7% | 8 | 4 | 4** |
| SB TH/RT | 400 | 35% | 42 | 4 | 38 |
| Total | 1135 | 100% | 120 | 16 | 104 |

* Actual field settings for green times and splits may need to be larger, as these reflect average conditions during the time period being analyzed

** Calculated green time may violate minimum green for phase and need adjustment. All green times need to be checked for compatibility with minimum vehicle and pedestrian timing requirements.

*Additional calculations as noted in the table, the splits calculated according to this method do not account for vehicle or pedestrian minimum times and may need to be adjusted accordingly. In addition, given that the split times reflect average conditions over the time period being analyzed, maximum green times and/or force-offs may need to be set higher to accommodate cycle-to-cycle fluctuations in traffic demand. Changes in demand from cycle to cycle are likely to be larger as a percentage of small-volume movements than on large-volume movements, so a larger buffer above the average split time is often appropriate for the former.*

## 3.4 INTERSECTION-LEVEL PERFORMANCE MEASURES AND ANALYSIS TECHNIQUES

The capacity measures discussed above are essential for determining the sufficiency of the intersection to accommodate existing or projected demand. However, capacity by itself is not easily perceived by the user. This section presents the most common user-perceived operational performance measures and analysis techniques used in timing individual intersections.

### 3.4.1 Performance Measures

The two primary user-perceived performance measures used to evaluate the performance of individual intersections are delays and queues.

#### *Control Delay and Intersection Level of Service*

Delay is defined in *HCM 2000* as “the additional travel time experienced by a driver, passenger, or pedestrian.” Delay can be divided into a number of components, with total delay and control delay being of most interest for signal timing purposes.

The total delay experienced by a road user can be defined as the difference between the travel time actually experienced and the reference travel time that would result in the absence of traffic control, changes in speed due to geometric conditions, any incidents, and the interaction with any other road users (adapted from the HCM definition). Control delay is the portion of delay that is attributable to the control device (the signal, its assignment of right-of-way, and the timing used to transition right-of-way in a safe manner) plus the time decelerating to a queue, waiting in queue, and accelerating from a queue. For typical through movements at a signalized intersection, total delay and control delay are the same in the absence of any incidents.

Chapter 16 of the HCM provides equations for calculating control delay; primary contributing factors are lane group volume, lane group capacity, cycle length, and effective green time. The HCM control delay equation also includes factors that account for elements such as pretimed versus actuated control, the effect of upstream metering, and oversaturated conditions. Control delay is calculated separately for each movement; intersection control delay consists of an average across all movements, weighted by volume. The HCM defines Level of Service for signalized intersections in terms of control delay using delay thresholds given in Table 3-3.

**Table 3-3 Motor vehicle LOS thresholds at signalized intersections**

| LOS | Control Delay per Vehicle (seconds per vehicle) |
|---|---|
| A | ≤ 10 |
| B | > 10-20 |
| C | > 20-35 |
| D | > 35-55 |
| E | > 55-80 |
| F | > 80 |

Source: (*2*)

#### *Queue Length*

Queue length is a measurement of the physical space vehicles will occupy while waiting to proceed through an intersection. It is commonly used to assess the amount of storage required for turn lanes and to determine whether the vehicles from one intersection will physically spill over into an adjacent intersection.

Several queue length estimations are commonly used with signalized intersections. Average queue and 95$^{th}$-percentile queue are commonly estimated for the time period for which the signal is red. However, it is sometimes useful to include the queue formation that occurs during green while the front of the queue is discharging and vehicles are arriving at the back of queue. Queues measured in this way are often noted as *average back of queue* or some percentile of *back of queue*. Appendix G to chapter 16 of the *HCM 2000* provides procedures for calculating back of queue.

### 3.4.2 Evaluation Techniques: The HCM Procedure for Signalized Intersections

To calculate the user-based performance measures described above, the critical movement analysis procedures described previously are insufficient. The most commonly used procedure for estimating intersection-level performance measures is provided by the HCM operational analysis methodology for signalized intersections (Chapter 16 of the HCM).

#### *Capabilities*

The HCM procedure addresses many of the limitations of critical movement analysis, including the assumption of constant values of capacity for each lane and the ability to analyze different types of signal phasing. In addition, some software packages implement procedures that are adequate for many signal timing applications, even though they may or may not be exact replications of the HCM procedure.

#### *Known Limitations*

Known limitations of the HCM analysis procedures for signalized intersections exist under the following conditions (adapted from *1*):

- Available software products that perform HCM analyses generally do not accommodate intersections with more than four approaches;
- The analysis may not be appropriate for alternative intersection designs;
- The effect of queues that exceed the available storage bay length is not treated in sufficient detail, nor is the backup of queues that block a stop line during a portion of the green time;
- Driveways located within the influence area of signalized intersections are not recognized;
- The effect of arterial progression in coordinated systems is recognized, but only in terms of a coarse approximation;

- Heterogeneous effects on individual lanes within multilane lane groups (e.g., downstream taper, freeway on-ramp, driveways) are not recognized; and
- The procedure accounts for right turns on red by reducing the right-turn volume without regard to when the turns can actually be made within the signal cycle.

If any of these conditions exist, it may be necessary to proceed to arterial models or to simulation discussed in the next section to obtain a more accurate analysis.

### 3.4.3 Practical Operational Approximations

In many cases, a variety of practical approximations can be used at varying stages in signal timing development. Some practitioners often rely primarily on these practical approximations and then observe and fine-tune their implementation in the field. In all cases, these practical approximations are often simple enough to be calculated in one's head, thus providing a method ready to be used in the field or as a quick check on calculations done using more advanced techniques.

#### *Cycle length*

At a planning level, it is common to assume a cycle length for a given intersection to estimate its capacity performance. If the cycle length for an intersection is unknown, common planning-level assumptions for cycle length based on the complexity of the intersection are given in Table 3-4. These assumptions do not account for cycle length requirements for coordinated operation (see Chapter 6), nor do they account for the ability for actuated intersections to vary the effective cycle length from cycle to cycle. These approximations, however, are useful in the procedures for estimating approximate queue lengths and for identifying whether a movement is substantially under or over capacity.

**Table 3-4 Planning-level cycle length assumptions**

| Signal Complexity | Commonly Assumed Cycle Length (s) |
|---|---|
| Permissive left turns on both streets | 60 |
| Protected left-turns, protected-permissive left turns, or split phasing on one street | 90 |
| Protected left-turns, protected-permissive left turn phasing, and/or split phasing on both streets | 120 |

If an analysis is to be completed for an intersection in an existing coordinated system, the user should use the current cycle length. Choosing a different cycle length would require changing the coordinated plan for adjacent intersections and could have implications throughout the system.

#### *Average and 95$^{th}$-Percentile Green Time*

Green time for individual movements can be approximated in many cases through estimating the number of vehicles per lane expected to be served in a given cycle and then setting an appropriate green time to serve that amount of traffic. This technique is useful for estimating green times for minor movements for which queue clearance is the primary objective and there is no intention to hold green time for vehicles to arrive and depart on green as one might do for coordinated through movements. For movements also serving pedestrians (e.g., minor-street through movements), this technique can provide an estimate of green time needed for cycles in which no pedestrian calls are placed, but it does not account for the additional time that may be required to provide pedestrian walk and clearance time.

The process for estimating the average queue length for each cycle is given in Equation 3-4. It makes the simplifying and conservative assumption that the movement is timed so that all arriving vehicles for the movement arrive on red and that there is no residual queue at the end of the green. This assumption is realistic if the green time for the movement is a small percentage of the cycle length (e.g., less than 10 to 15 percent of the cycle) and thus would not apply to particularly high-volume movements.

$$Queue_{avg} = \frac{v}{3600/C} \quad (3\text{-}4)$$

where $Queue_{avg}$ is the average queue in vehicles per lane; $v$ is the volume of the movement in vehicles per hour per lane, and $C$ is the cycle length in seconds. For example, a volume of 150 vehicles per hour per lane under a cycle length of 90 seconds will result in an average queue length of approximately 150 / (3600 / 90) = 3.75 vehicles. If using this value for timing or design, this queue length should be rounded up to the nearest vehicle, in this case 4 vehicles.

For movements where the green time is a larger proportion of the cycle, the above formula does not account for the proportion of vehicles that arrive on green. In these cases, it is more accurate to account for the green time for the subject movement as shown in Equation 3-5.

$$Queue_{avg} = \frac{v}{3600/(C-g)} \quad (3\text{-}5)$$

where $Queue_{avg}$ is the average queue in vehicles per lane; $v$ is the volume of the movement in vehicles per hour per lane, $C$ is the cycle length in seconds, and $g$ is the effective green time in seconds. For example, a volume of 300 vehicles per hour per lane under a cycle length of 90 seconds and an effective green time of 20 seconds will result in an average queue length of approximately 300 / (3600 / (90-20)) = 5.83 vehicles. If using this value for timing or design, this queue length should be rounded up to the nearest vehicle, in this case 6 vehicles.

**A useful approximation for saturation flow rate is to assume a value of 1800 vehicles per hour for each lane, which results in a numerically simple value of 2 seconds of green time per vehicle. It is also useful to assume a startup lost time of approximately 2 seconds. Therefore, a queue of 4 vehicles can be assumed to take 2 + (4 × 2) = 10 seconds to clear.**

One may use a Poisson distribution to estimate 95$^{th}$-percentile queue from the average queue. In practice, the 95$^{th}$-percentile queue is approximately 1.6 times the average queue for high-volume movements to approximately 2.0 times the average queue for low-volume movements. The following practical and simplifying approximation is useful in many cases for the purposes of signal timing, recognizing that it is overly conservative in some cases:

$$Queue_{95th} \approx 2 \cdot Queue_{avg} \quad (3\text{-}6)$$

Using the example in the previous paragraph, a movement with an average queue length of 3.75 vehicles has a 95th-percentile queue length of approximately 7.5 vehicles, rounded up for timing or design purposes to 8 vehicles. This movement would need a 95th-percentile green time of 2 + (8 × 2) = 18 seconds.

Some practitioners use the 95$^{th}$-percentile queue for estimating green time for protected left turns and the average queue for estimating green time for protected-permissive left turns under the assumption that under protected-permissive operation, vehicles not served during the protected phase can be served during the permissive phase. This depends in part on the availability of gaps in the opposing through vehicle stream and thus may not be applicable in cases where the opposing through movement is approaching capacity. In these cases, a more detailed analysis is advisable.

Another useful planning-level procedure used in design for estimating 95$^{th}$-percentile queue length is to assume a value of 1 foot per vehicle being served during the hour. For example, a left-turn movement with a volume of 150 vehicles per hour would have a 95$^{th}$-percentile queue length of approximately 150 feet. This procedure is most accurate for cycle lengths around 90 seconds. For shorter cycle lengths, the queue length should be shortened (e.g., by 10 to 20 percent for a cycle length of 60 seconds); for longer cycle lengths, the queue length should be extended (e.g., by 10 to 20 percent for a cycle length of 120 seconds). One could then use the value from this estimation method by assuming that each vehicle occupies a space of 25 feet and then using the techniques described above to estimate the 95$^{th}$-percentile green time for the movement.

**One must use caution using these practical approximations where queue length is critical. For example, if the queue from a left-turn lane exceeds available storage, it could block a through lane. If the blocked through lane is a critical movement, the performance of the entire intersection will be adversely affected, and field observations will not match office calculations.**

### 3.4.4 Intersection-Level Field Measurement

Several field techniques can be used to measure some of the key intersection-level operational measures of effectiveness. The most common in practice is an intersection delay study. The FHWA/NTOC performance measures project selected delay (both recurring and non-recurring) as one of the measures to be considered for national standardization. Since non-recurring delay is the delay that occurs in the presence of an incident, recurring delay is the measure that is applied to evaluate signal timing.

Some (but not all) of the more commonly used performance measures can be calculated from each of these techniques. These and others are described in more detail in the ITE *Manual of Transportation Engineering Studies* (*3*) and other references. This section discusses some of the more commonly used methods.

***Stopped Delay***

A common intersection delay study is the method to estimate stopped delay. This procedure surveys a specified movement over a period of time. There are two data points collected during the survey, the total volume and the number of stopped vehicles at a given time interval. The aggregate sum of the stopped vehicles at the time interval is divided by the total entering volume to determine an average stopped delay. This method has been superseded in the HCM by the method to estimate control delay, described in the next section.

***Control Delay***

Control delay can be measured in the field by recording the arrival and departure time of vehicles for a movement or approach. This procedure is described in detail in the HCM in Appendix A of Chapter 16. A detailed description of the methodology and a worksheet are provided. This method does not directly capture all of the deceleration and acceleration associated with control delay but is indicated in the HCM to yield a reasonable estimate of control delay.

Queue formation during oversaturated conditions can make it difficult to use this method, as queues often extend beyond the measurement area and can spill into other intersections, confounding any measurements. In these cases, travel time estimates for selected origin-destination pairs (as described in the following section) may be more useful.

***Delay Weighting for Specific User Types***

Delays can be focused primarily on vehicular traffic or can be weighted by particular vehicle types. To improve freight mobility, data collection could weight trucks more heavily than other traffic. Person delay is sometimes used by weighting the person carrying capacity of the vehicle. In these cases, transit vehicle capacity is calculated into the overall delay on the system.

## 3.5 ARTERIAL- AND NETWORK-LEVEL PERFORMANCE MEASURES AND PREDICTION TECHNIQUES

When considering signal timing among a series of signalized intersections, as for coordinated signal operation, performance measures and models that account for the relative interaction of adjacent intersections become important. This section presents the most common performance measures, prediction techniques, and field measurements used for evaluating signal timing performance along arterial streets and coordinated networks. The measures and methodologies in this section, along with the measures for individual intersections described previously in this chapter, support the signal timing policies, the coordination techniques, and the signal timing process presented in Chapters 2, 6, and 7, respectively.

### 3.5.1 Arterial- and Network-Level Performance Measures

In addition to the estimation of intersection-level performance measures at each intersection along an arterial or within a network, a number of performance measures are used to assess how well the intersections fit together in terms of signal timing. These include stops, travel speed, and bandwidth and are described below. Other arterial- and network-level measures of effectiveness, including transit level of service, bicycle level of service, and pedestrian level of service, are sometimes used in developing and evaluating the effectiveness of signal timing. These are outside the scope of this document but are provided in the HCM.

#### *Stops*

Stops are also used frequently to measure signal system effectiveness. Although this measure has not been identified as a candidate for standardization nationally, it is very important for two reasons:

- Stops have a higher impact on emissions than delay does because an accelerating vehicle emits more pollutants and uses more fuel than an idling vehicle. In other words, an idling vehicle must idle for many minutes before it emits as many pollutants as those emitted for a single stop.
- Stops are a measure of the quality of progression along an arterial. Motorists are often frustrated when they have to make multiple stops. In some of the signal timing software, the user is given the option of defining the relative importance of stops and delays through the use of weighting factors. If stops are assigned a high level of importance, progression on arterials will be improved, even though this may result in higher overall delay.

Motor vehicle stops can often play a larger role than delay in the perception of the effectiveness of a signal timing plan along an arterial street or a network. Stops tend to be used more frequently in arterial applications where progression between intersections (and thus reduction of stops) is a desired objective. Many software packages, including signal timing optimization programs and simulation packages, include estimates of stops.

#### *Travel Time, Travel Speed, and Arterial Level of Service*

Travel time and travel speed is one of the most popular measures used to assess how well arterial traffic progresses. Travel speed accounts for both the delay at intersections and the travel time between intersections.

The HCM can be used to determine arterial level of service (LOS) based on travel speed. The HCM defines arterial LOS as a function of the class of arterial under study and the travel speed along the arterial. This speed is based on intersection spacing, the running time between intersections, and the control delay to through vehicles at each signalized intersection. Because arterial level of service is calculated using delay for through vehicles regardless of origin or destination, the resulting speed estimates may not necessarily correspond to speed measurements made from end-to-end travel time runs that measure a small subset of the possible origin-destination combinations along an arterial.

Table 3-5 presents the HCM's definitions of classes of arterials based on design and functional categories, and 0 presents the threshold speeds for each class of arterial. Further detail can be found in Chapters 10 and 15 in the HCM. Note that these level of service thresholds only reflect the perspective of vehicular travel time; they do not account for the effects on transit, bicycles, and pedestrians. A broader, multimodal perspective on arterial level of service is being considered for the next edition of the HCM.

**Table 3-5 HCM arterial class definitions**

| Design Category | Functional Category | |
|---|---|---|
| | **Principal Arterial** | **Minor Arterial** |
| High-Speed | I | N/A |
| Suburban | II | II |
| Intermediate | II | III or IV |
| Urban | III or IV | IV |

Source: (*2*)

**Table 3-6 Arterial level of service**

| **Urban Street Class** | **I** | **II** | **III** | **IV** |
|---|---|---|---|---|
| Range of free-flow speeds (FFS) | 55 to 45 mph | 45 to 35 mph | 35 to 30 mph | 35 to 25 mph |
| Typical FFS | 50 mph | 40 mph | 35 mph | 30 mph |
| LOS | Average Travel Speed (mph) | | | |
| A | > 42 | > 35 | > 30 | > 25 |
| B | > 34 – 42 | > 28 – 35 | > 24 – 30 | > 19 – 25 |
| C | > 27 – 34 | > 22 – 28 | > 18 – 24 | > 13 – 19 |
| D | > 21 – 27 | > 17 – 22 | > 14 – 18 | > 9 – 13 |
| E | > 16 – 21 | > 13 – 17 | > 10 – 14 | > 7 – 9 |
| F | ≤ 16 | ≤ 13 | ≤ 10 | ≤ 7 |

Source: (*2*)

### *Bandwidth, Bandwidth Efficiency, and Bandwidth Attainability*

Bandwidth, along with its associated measures of efficiency and attainability, are measures that are sometimes used to assess the effectiveness of a coordinated signal timing plan. Unlike the preceding measures, bandwidth is purely a function of the signal timing plan as it is oriented in time and space; the effect of traffic is not explicitly accounted for in the calculation. As a result, bandwidth is not directly tied to actual traffic performance, although it is sometimes used as a surrogate for potential traffic performance. The use of bandwidth in developing coordinated signal timing plans is described in more detail in Chapter 6.

Bandwidth is defined as the maximum amount of green time for a designated movement as it passes through a corridor. It can be defined in terms of two consecutive intersections (sometimes referred to as *link bandwidth*) or in terms of an entire arterial (sometimes referred to as *arterial bandwidth*). Bandwidth is typically measured in seconds.

Two related bandwidth measures, bandwidth efficiency and bandwidth attainability, are sometimes used to normalize bandwidth measures. Bandwidth efficiency is a measure that normalizes bandwidth against the cycle length for the arterial under study. The specific formula is given in Equation 3-7:

$$E = \frac{B_A + B_B}{2C} \tag{3-7}$$

where $E$ is bandwidth efficiency; $B_A$ and $B_B$ are the bandwidths in the forward and reverse directions, respectively, in seconds; and $C$ is the cycle length in seconds.

Efficiency is typically reported as a unitless decimal or percentage. The software program PASSER II provides guidelines to determine the quality of a particular efficiency; these are given in Table 3-7.

Table 3-7 Guidelines for bandwidth efficiency

| Efficiency Range | PASSER II Assessment |
|---|---|
| 0.00 – 0.12 | Poor progression. |
| 0.13 – 0.24 | Fair progression. |
| 0.25 – 0.36 | Good progression. |
| 0.37 – 1.00 | Great progression. |

Source: (*4*)

Bandwidth attainability is a measure of how well the bandwidth makes use of the available green time for the coordinated movements at the most critical intersection in the corridor. It is given in Equation 3-8 as follows:

$$Attainability = \frac{bandwidth}{g_{crit}} \quad (3\text{-}8)$$

where $g_{crit}$ is the green time for the coordinated movements at the intersection with the least amount of green time available for the coordinated movements.

Attainability, as with efficiency, is typically reported as a unitless decimal or percentage. The software program PASSER II provides guidelines to determine the quality of a particular attainability; these are given in Table 3-8. An attainability of 100 percent suggests that all available green time through the most constrained intersection is being used by the green band.

Table 3-8 Guidelines for bandwidth attainability

| Attainability Range | PASSER II Guidance |
|---|---|
| 1.00 – 0.99 | Increase minimum through phase. |
| 0.99 – 0.70 | Fine-tuning needed. |
| 0.69 – 0.00 | Major changes needed. |

Source: (*4*)

Bandwidth is highly dependent on the demands on the non-coordinated phases (actuated intersections) and the offsets along the arterial. Bandwidth importance also varies based on traveler origin and destinations along a coordinated corridor. The more vehicles traveling the length of a coordinated corridor, without making turning movements, the more important bandwidth becomes as a measurement tool, and vice versa. In addition, bandwidth, travel time, and stops are related to one another; a large bandwidth will result in shorter travel times and less stops to vehicles traveling the length of the corridor.

### *Emissions*

Emissions are an important measure because they directly affect air quality. Emissions may be particularly important in the case of programs funded by Congestion Mitigation and Air Quality (CMAQ) funds, which (as the name implies) use funds allocated with the objective of improved air quality. Emissions can be estimated by some of the software packages in use.

### *Fuel Consumption*

Improved fuel efficiency of the transportation system is an important measure for reducing the costs to consumers and use of this energy source. Fuel consumption can be estimated by some of the software packages in use.

### *Performance Index*

Some software packages allow several measures of effectiveness to be mathematically combined into a single performance index, or PI. This allows an optimization routine within the software package to search for combinations of cycle lengths, offsets, and splits that achieve an optimal value of the performance index. The performance index itself is unlikely to be observable in the field, although its component measures may be observable.

### 3.5.2 Evaluation Techniques

Measures of effectiveness can be measured directly in the field or estimated using an analysis tool. Table 3-9 indicates the variety of field data collection and analytical techniques that can be used.

**Table 3-9 Evaluation Techniques for Performance Measures**

| Measure | Primary Evaluation Technique | Supplemental Technique |
|---|---|---|
| Delays | Deterministic traffic analysis tool | Floating car runs (through movement or intersection observation) |
| Stops | Field data collection using manual observers or floating car runs (See text on the next page.) | Floating car runs or simulation |
| Emissions | Microscopic traffic simulation | Environmental or planning models |
| Fuel Efficiency | Microscopic traffic simulation | Environmental or planning models |
| Travel Time | Floating car runs | Microscopic traffic simulation |

#### *Arterial and Network Signal Timing Models*

For many practitioners, arterial and network signal timing models are the models of choice when developing signal timing plans, particularly for coordinated systems. Arterial and network signal timing models are distinguished from the isolated intersection methods described previously by the way they account for the arrival and departure of vehicles from one intersection to the next. Traffic progression is treated explicitly through using time-space diagrams or platoon progression techniques. Some models deterministically estimate the effect of actuated parameters for both vehicles and pedestrians, typically by using a combination of scaling factors for vehicle demand and the presence or absence of vehicle and/or pedestrian calls.

A key feature of arterial and network signal timing models is the explicit effort to “optimize” signal timing to achieve a particular policy (see Chapter 2). To accomplish this, each model uses some type of algorithm to test a variety of combinations of cycle length, splits, and offsets to achieve a calculated value of one or more performance indices and then attempt to find an optimal value for those performance indices. In addition, most arterial and network signal timing models use elements of the HCM procedures to estimate certain parameters, such as saturation flow rate.

Arterial and network signal timing models are frequently the most advanced type of model needed for most signal timing applications. **However, because the models described in this section are deterministic, they lose validity in cases where demand exceeds capacity or where the queues from one intersection interact with the operation of an adjacent intersection.** Therefore, cases with closely spaced intersections or with intersections exceeding capacity may not be well served by these types of deterministic models. In these cases, it may be necessary to use microscopic simulation models to obtain more realistic assessments of these effects.

#### *Microscopic Simulation Models*

In the context of signal timing, microscopic simulation models can be thought of as an advanced evaluation tool of a proposed signal timing plan. Most simulation models do not have any inherent signal timing optimization algorithms and instead depend on other programs to provide a fully specified signal timing plan for evaluation. However, simulation models typically can directly evaluate the effects of interactions between intersections or the effects of oversaturated conditions or demand starvation caused by upstream oversaturated conditions. Simulation models can also be useful to estimate the effects of short lanes where either the queue in the lane affects the operation of an adjacent lane or where the queue in the adjacent lane prevents volume from reaching the subject lane.

Recent advances in technology have allowed direct linkages between simulation models and either actual signal controllers or software emulations of those controllers, known as hardware-in-the-loop (HITL) and software-in-the-loop (SITL), respectively. These in-the-loop simulations allow actual controllers and/or their algorithms to replace the approximation used in simulation models to more accurately reflect how a controller will operate. The simulation model is used to generate traffic flows and send vehicle and pedestrian calls to one or more controllers based on a detection design implemented in the simulation. The controller receives the calls as if it was operating in the field and uses its own internal algorithms to set signal indications based on the calls received and the implemented signal timing. The signal displays are passed back to the simulation model, to which traffic responds.

When running simulation analyses to compare alternatives, care is needed to ensure that the resolution of the model is sufficient to model the type of control anticipated. A simulation model, for example, typically requires precision on the order of a tenth of a second to accurately replicate gap detection. Therefore, the use of coarser resolution settings to speed up simulation run time may yield incorrect results.

### 3.5.3 Arterial and Network Field Measurement Techniques

Data collection of arterial and network measures of effectiveness can be conducted either manually or using automated techniques. The following sections discuss some of the more commonly employed techniques in more detail.

***Travel Time Runs***

Travel time studies evaluate the quality of traffic movements along an arterial and determine the locations, types, and extent of traffic delays. Common measures extracted from travel time runs include overall travel time, delay for through vehicles at individual intersections, number of stops, and the variability of delays and/or stops at intersections. The latter measure is often useful in demonstrating a change between uncoordinated and coordinated operation.

The floating car method is the most commonly employed technique used for travel time runs. The vehicle driver is instructed to "float" with the traffic stream, which is defined as traveling at a speed that is representative of the other vehicles on the roadway. This is accomplished by passing as many vehicles as pass the floating car (within reasonable safety limits). The floating car data is then reduced to determine the number of stops, delay time, and travel time over the route traversed by the vehicle. This process must be repeated several times to obtain good representative data.

A variety of techniques have been used to collect travel time runs, ranging from manual measurements with stop watches to more advanced techniques using portable Global Positioning System (GPS) devices. GPS output provides the vehicle's location (latitude and longitude) and velocity. The probe vehicle begins traveling along the corridor with the GPS set to record position and speed for each second of travel. The vehicle location and instantaneous velocity allows for an estimate of the performance measures of delay, percent runs stopped, and travel time. Special inputs can also be coded to denote an event such as the location of a stop bar or bus stop, or when a parking maneuver occurs.

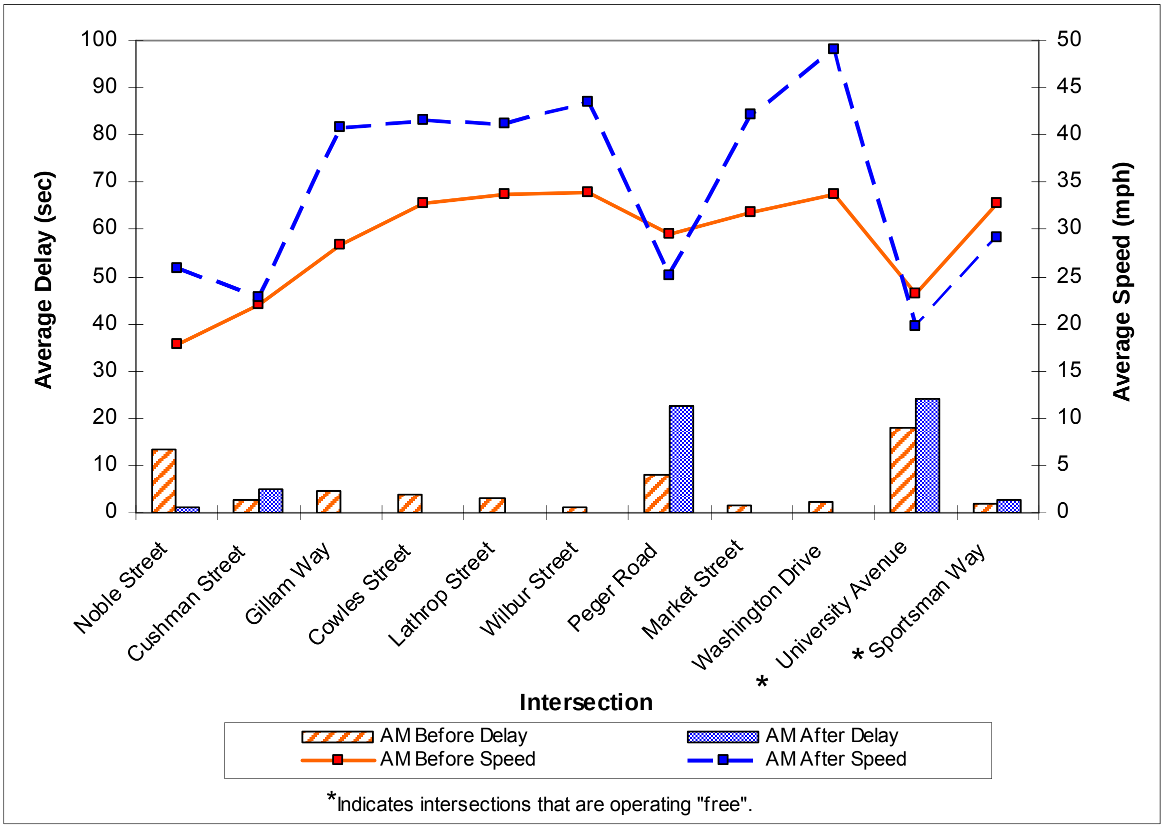
Of the variety of performance measures that can be used to evaluate arterial performance, four can be determined by travel time runs:

- *Travel Time* is the total elapsed time spent driving a specific distance. The average travel time represents an average of the runs for a particular link or for the entire corridor.
- *Delay* is the elapsed time spent driving in a stopped or near-stopped condition, often defined as speeds less than 5 mph. The average delay represents an average of the runs for a particular intersection or for the entire corridor.
- *Percent Runs Stopped* is the percentage of the total number of travel time runs conducted during which a vehicle stops.
- *Average Speed* is the average distance a vehicle travels within a measured amount of time (this is affected by the amount of time experienced in delay).

A sample travel time graph is shown in Figure 3-6 where distance along the corridor is represented along the *x*-axis and performance measures are plotted along the *y*-axis. This type of graphical display allows for ready comparison of before and after data along the corridor.

A critical factor in the success of a travel time study using vehicle probes is the skill level of the driver(s) involved. Inexperienced personnel may suddenly realize they are no longer traveling as part of the platoon and may adversely affect the platoon by inducing braking reactions in response to lane changes. This is more likely to occur in flows at or near capacity.

**Figure 3-6 Sample travel time run result graph**




### *Automated Collection of Measures of Effectiveness*

It is ideal to use a traffic signal system for automated collection of performance measures, but few agencies currently have that capability. In many cases, vehicle detectors are used to perform two different types of functions.

The first function is to provide inputs for actuated signal control. When detectors are used for this purpose, they are typically installed near the intersection on the approaches under actuated control, which makes them less effective for performance measurements.

The second function is to act as system sensors. These detectors are typically installed at a mid-block location, which provides accurate measures of vehicle volumes and speeds. The information provided by system sensors can be used to support the following system functions:

- acquiring traffic flow information to compute signal timing
- identifying critical intersection control (CIC)
- selecting timing plans

- developing on-line and off-line timing plans through optimization programs.

Thus, the local intersection detectors are used directly by actuated controllers, whereas the system sensors are used to support the functions of the central computer or arterial master (*5*).

To acquire data, the loop detector is best suited to computerized traffic signal control because it is reliable, accurate, and able to detect both presence and passage. Using the presence and passage outputs, a number of variables, such as volume, occupancy, speed, delay, stops, queue lengths and travel time, can be defined to a varying degree of accuracy. These variables are related to either traffic flow along the length of the detectorized roadway or in the immediate vicinity of the detectors (*5*). Some of these measures of effectiveness that can be collected through loop detectors are described below.

- *Volume* is the quantity that is most easy to obtain. This quantity is the number of pulses measured during a given period.
- *Occupancy* is the percent of time that a detector indicates a vehicle is present over a total time period. Occupancy can range from zero to 100 percent, depending on vehicle spacing. Volume and occupancy are the most important variables for use in selecting traffic responsive timing plans and for many operating thresholds. In some cases, volume can be used without occupancy. It generally follows that occupancies of over 25 percent are a reliable indicator of the onset of congestion. (*5*)
- *Speed* is another useful variable used for on-line or offline computation of signal timing plans using optimization programs. For a single inductive loop detector, speed is inversely proportional to occupancy at a fixed volume. Two loops (dual loop detectors—also known as speed traps) provide much more accurate speed measurements.

## 3.6 SAFETY ASSESSMENT

The safety performance of signal timing is often implicitly considered but seldom explicitly evaluated. The transportation profession is in the process of developing increasingly quantitative safety methodologies to allow analysis of safety measures of effectiveness comparable to their operational counterparts. The forthcoming *Highway Safety Manual* (*6*) represents a major step in this direction. In addition, further discussion on safety analysis as applied to signalized intersections may be found in FHWA's *Signalized Intersections: Informational Guide* (*1*); the content in this section is largely adapted from that document.

Statistical techniques for evaluating collision performance vary from basic to complex. They may compare the safety performance of a single signalized intersection to another group of similar intersections, or they may serve as a screening tool for sifting through a large group of sites and determining which site has the most promise for improvement. This section provides information about safety analyses focusing exclusively on whether or not there are safety issues at an intersection that will respond to signal timing; it also provides quantitative information about estimated safety impacts of signal timing changes.

There are many factors in addition to signal timing that contribute to the safety performance of an intersection. There are also many new tools being developed to support quantitative safety analyses. Therefore, in the situation where it has been determined an intersection may respond to safety improvements, the analyst should be sure to consider more than signal timing changes as a possibility for improving safety.

### 3.6.1 Crash Data Review

Crash data is of widely varying quality. In general, the less severe the collision, the less likely it is for a crash to be reported to the local jurisdiction. As a result, it is common to experience systematic underreporting of property damage only (PDO) crashes. Crashes are sometimes reported inaccurately or subsequently inaccurately entered into the database, making it difficult for the analyst to accurately identify factors contributing to crashes. It is also notable that the number of crashes that occur at a given location will vary from year to year partially due to changes in traffic volumes, weather conditions, travel patterns, and surrounding land uses, and partially due to regression to the mean. Regression to the mean is caused by the tendency for crashes at a location to fluctuate up and down over time as the average number of crashes per year converges to a long-term average. Therefore, the observed number of crashes in any given year may actually be higher or lower then the long-term average for the site.

The evaluation of crash frequency, crash patterns, and crash severity remains an important part of a safety analysis. The analyst should obtain three to five years of crash data; summarize the crashes by type, severity, and environmental collisions; and prepare a collision diagram of the crashes to help identify trends. A site visit can also be helpful in identifying causal factors. It is also possible to statistically determine if any of the crashes by type, severity, or other environmental conditions (e.g. weather, pavement, time of day) are over-represented as compared to other comparable facilities. A variety of tests and other evaluative tools are discussed further in FHWA's *Signalized Intersections: Informational Guide (1)*.

### 3.6.2 Quantitative Safety Assessment

Signal timing is one of many factors which may contribute to crashes. Other factors may be horizontal and vertical alignment conditions, roadside features, sight distance, driver compliance with traffic control, access management near intersections, driver expectations, and roadway maintenance and lighting. If the purpose of the project is to update intersection signal timing, the analyst should be aware of the potential impacts to safety. Alternatively, if the site has been identified as one needing safety improvements the analyst should be aware that signal timing is one of *many* possible countermeasures for improving safety at an intersection. Table 3-10 provides a summary of crash types and possible signal timing changes to benefit safety.

**Table 3-10 Summary of crash types and possible signal modifications to benefit safety**

| | Collision with Another Vehicle | | | | | Single-Vehicle Collision | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Signal Timing Change** | **Angle** | **Head-On** | **Rear-End** | **Sideswipe-Same Direction** | **Sideswipe-Opposite Direction** | **Collision with Bicycle** | **Collision with Parked Vehicle** | **Collision with Pedestrian** | **Overturned** | **Run Off Road** |
| Provide left-turn signal phasing | • | • | • | • | • | | | | | |
| Optimize clearance intervals | • | | • | | | | | | | |
| Restrict or eliminate turning maneuvers (including right turns on reds) | • | • | • | • | | • | | • | | |
| Employ signal coordination | • | | • | | | | | • | | |
| Implement emergency vehicle pre-emption | • | • | • | • | • | • | • | • | | |
| Improve traffic control of pedestrians and bicycles | | | • | | | • | | • | | |
| Remove unwarranted signal | | | • | | | | | | | |
| Provide/improve left-turn lane channelization* | • | • | • | • | • | | | | | |
| Provide/improve right turn lane channelization* | • | | • | • | | • | | • | | |

1 * *While not technically part of signal timing, these have a significant impact on safety and should be considered. See the FHWA Signalized Intersections: Informational Guide (1) for more information. Ref: 9, Appendix 10*

Accident modification factors (AMFs) are used as a way to quantify crash reductions associated with safety improvements. These factors are developed based on rigorous before-after statistical analysis techniques that account for, among many factors, regression to the mean, sample size, and the effects of other treatments from the estimation process. AMFs which have been calculated using reliable before-after analyses will be quantified and will often also include a confidence interval. AMFs below 1.0 will likely have a safety benefit, and AMFs greater than 1.0 may degrade safety conditions. For example, studies have found that converting a 4-legged urban intersection from stop control to signal control will have an AMF for all fatal and injury crashes of 0.77. This means that there will be an anticipated reduction in crashes of 23 percent if a traffic signal were installed.

Quantitative AMFs can be found in many sources (*7*, *8*, *9*). The most reliable sources are those that estimate AMFs using the more sophisticated before-after analytical techniques. In the future, the *Highway Safety Manual* (*6*) and the final report for NCHRP 17-25 “Crash Reduction Factors for Traffic Engineering and ITS Improvements” will also contain numerous AMFs. In addition, local jurisdictions can also develop their own AMFs based on studies conducted locally.

## 3.7 SIGNAL WARRANTS

Warrants for signalization are intended to create a minimum condition for which signalization may be the most appropriate treatment. Each of the warrants is based on simple volume, delay, or crash experience at the location before signalization is installed. None accounts for the specific design of the signal or the way it may be timed (e.g., pre-timed versus actuated). As a result, an engineering evaluation should be conducted in conjunction with the evaluation of signal warrants to determine that the proposed signalization plan actually represents an improvement over existing conditions.

The eight warrants presented in Chapter 4C of the 2003 MUTCD are as follows (*10*):

- **Warrant 1, Eight-Hour Vehicular Volume.** This warrant consists of two volume-based conditions, Minimum Vehicular Volume and Interruption of Continuous Traffic, of which one or both must be met over an eight-hour period.
- **Warrant 2, Four-Hour Vehicular Volume.** This warrant is similar to Warrant 1 but relies on volume conditions over a four-hour period.
- **Warrant 3, Peak Hour.** This warrant is primarily based on delay to minor movements during peak hour conditions.
- **Warrant 4, Pedestrian Volume.** This warrant is intended for conditions where delay to pedestrians attempting to cross a street is excessive due to heavy traffic on that street. It consists of minimum pedestrian volumes and available gaps in traffic.
- **Warrant 5, School Crossing.** This warrant is similar to Warrant 4 except that it is specifically intended for school crossing locations.
- **Warrant 6, Coordinated Signal System.** This warrant is intended to allow signals that may assist in progression of traffic. It is only intended for use where the resulting signal spacing is not less than 1,000 feet (300 m).
- **Warrant 7, Crash Experience.** This is the only specific safety-related warrant and is satisfied by a frequency of crashes over a specific period of time that can be corrected through the use of signalization.
- **Warrant 8, Roadway Network.** This warrant may be used to support the use of a signal to concentrate traffic at specific locations.

As noted in the MUTCD, "the satisfaction of a traffic signal warrant or warrants shall not in itself require the installation of a traffic control signal" (Section 4C.01, *10*).

**Signalization is not always the most appropriate form of traffic control for an intersection, and it is sometimes possible to create a larger benefit by removing a traffic signal than by retiming it.**

The MUTCD acknowledges this by stating that "since vehicular delay and the frequency of some types of crashes are sometimes greater under traffic signal control than under STOP sign control, consideration should be given to providing alternatives to traffic control signals even if one or more of the signal warrants has been satisfied." (*10*). Potential alternatives include the use of warning signs, flashing beacons, geometric modifications, and/or conversion of the intersection to a stop-controlled intersection or a roundabout.

## 3.8 REFERENCES


1. Rodegerdts, L., B. Nevers, B. Robinson, J. Ringert, P. Koonce, J. Bansen, T. Nguyen, J. McGill, D. Stewart, J. Suggett, T. Neuman, N. Antonucci, K. Hardy, and K. Courage. *Signalized Intersections: Informational Guide.* Report No. FHWA-HRT-04-091. Federal Highway Administration, Washington, D.C., August 2004.

2. Transportation Research Board. Highway Capacity Manual 2000. National Academy of Sciences, Transportation Research Board, Washington, D.C., 2000.

3. Institute of Transportation Engineers. *Manual of Transportation Engineering Studies.* H. D. Roberson, ed. Institute of Transportation Engineers, Washington, D.C., 2000.

4. Wallace, C. E., E.C.P. Chang, C. J. Messer, and K. G. Courage, "Methodology for Optimizing Signal Timing (M¦O¦S¦T), Volume 3: PASSER II-90 Users Guide," prepared for the Federal Highway Administration, COURAGE AND WALLACE, Gainesville, FL, December 1991.

5. *Traffic Detector Handbook, 2nd ed.* Report No. FHWA-IP-90-002. Federal Highway Administration, Washington, D.C., July 1990.

6. Transportation Research Board. Highway Safety Manual. Anticipated publication 2009.

7. National Cooperative Highway Research Program (NCHRP). *NCHRP Research Results Digest 299*. Washington, DC: Transportation Research Board of the National Academies, November 2005.

8. Elvik, R. and T. Vaa. *The Handbook of Road Safety Measures*. Amsterdam: Elsevier Ltd., 2004.

9. Antonucci, N.D., Hardy, K.K., Slack, K. L., Pfefer, R., and Neuman, T. R., "NCHRP Report 500 Volume 12: A Guide for Addressing Collisions at Signalized Intersections." Washington, D.C., Transportation Research Board, National Research Council, 2004.

10. Federal Highway Administration. Manual on Uniform Traffic Control Devices. Washington, DC: Federal Highway Administration, 2003.

# CHAPTER 4

# TRAFFIC SIGNAL DESIGN

## TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# 4.0 TRAFFIC SIGNAL DESIGN CONCEPTS

This chapter documents the concepts of traffic signal design as they apply to traffic signal timing. The principles related to geometric design and operation are addressed in the *Signalized Intersections: Informational Guide* (*1*) The elements addressed in this Chapter include: signal control type, signal phasing, detection layout, and how the decisions made during traffic signal design affect signal timing for isolated (described in Chapter 5) and coordinated (described in Chapter 6) operation.

This chapter consists of four sections. The first section provides an overview of the objectives of traffic signal design. The second section summarizes the basic concepts associated with key signal design elements. The third section describes a procedure for determining appropriate signal design elements for a given intersection. The last section provides guidelines for selecting signal design values and choosing from among design options.

## 4.1 OVERVIEW

The topic of traffic signal design broadly includes any of an intersection's traffic signal control elements that have a physical presence at the intersection. In this regard, the traffic control type (e.g., pre-timed, actuated) is implemented in the signal controller and cabinet. The signal phasing is implemented using signal heads, signal indications, and logic in the controller that governs their sequence. Detection is provided by various devices that identify users (such as vehicles and pedestrians) of the intersection. Other design elements exist (e.g., preemption); however, the aforementioned three elements are present at almost all intersections, and have a significant influence on intersection safety and efficiency. Hence, they are the subject of discussion in this chapter.

The objective of signal design is to produce a design that yields safe and efficient operation for the prevailing conditions. This objective is accomplished by making design choices that are tailored to the specific facility conditions. Signal timing settings (as described in Chapter 5) can be changed as needed to accommodate changes in traffic demand, pattern, but signal design elements are relatively static and are typically more difficult (or costly) to change and thus are discussed as an introduction. The efficiency of an intersection is directly impacted by its signal design, and the detection layout can have a significant effect on the safety associated with high-speed intersection approaches (*2*).

### 4.1.1 Relationship between Signal Timing and Traffic Control Design

The traffic signal design process should recognize and accommodate signal timing considerations to insure effective operation of the intersection. A robust detection system is needed for the traffic signal to be able to respond to changes in traffic conditions. Detection systems sense when pedestrians and vehicles are at a traffic signal and use that information to determine who will be served next and how long the phase is served. The quality of intersection operation is particularly dependent on the relationship between the detection layout and the signal controller settings. For optimum performance, the detector layout and signal settings should be "tuned" to the geometry of the intersection, its traffic volume, and the approach speed. The tuning process consists of finding a balance between detector location (relative to the stop line), detector length, passage time, and minimum green time for the prevailing conditions. However, there is no strong consensus in the industry with regard to what is the "best balance."

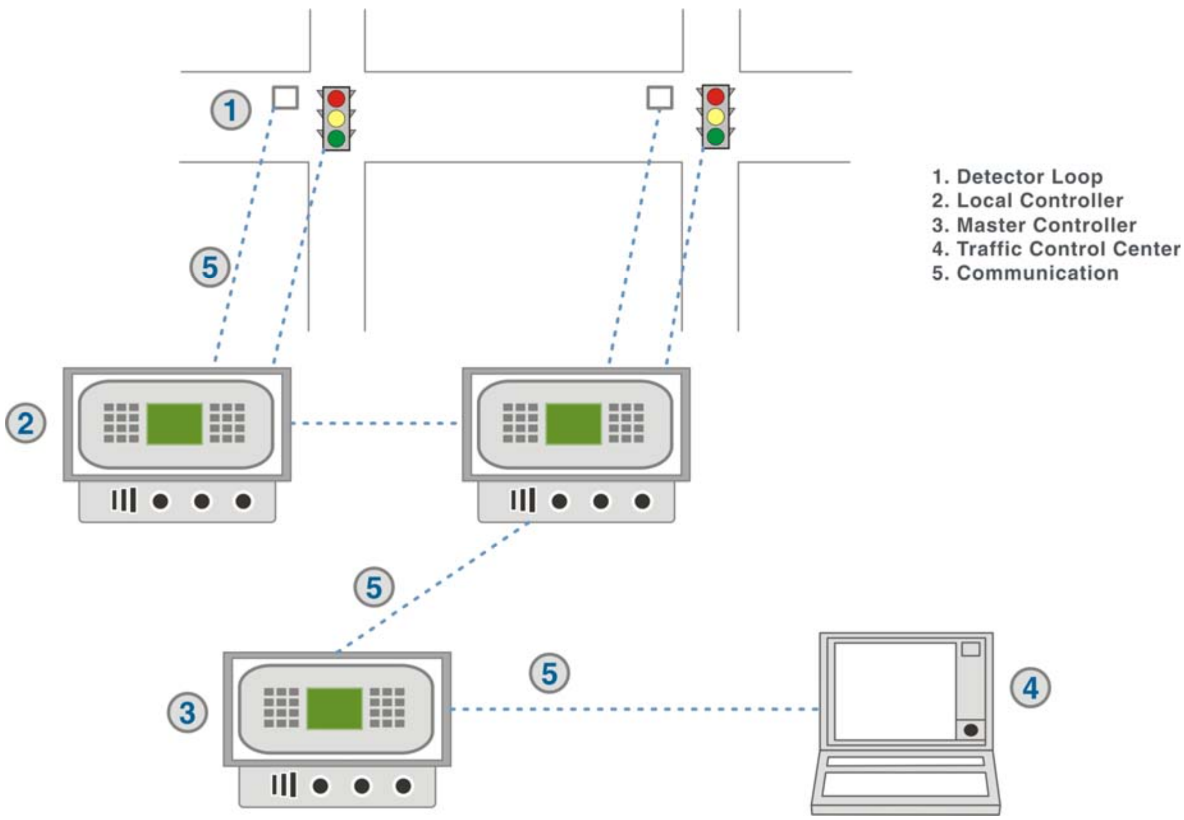
**Use of an iterative process in the design process results in an intersection that can take advantage of signal timing techniques to provide a high level of service to all users.**

The signal design can also be influenced by the traffic control device layout. Specifically, for a safe and effective signal design, it must be possible to properly position signal heads for maximum visibility for all movements. The MUTCD (*3*) describes the minimum standards for traffic signal displays. These standards address the number, size, mounting alternatives, physical arrangement, and placement of the signal heads. They also include special considerations for left-turn phasing that address the number and arrangement of lenses in the left-turn signal head, the location of this head, and the display sequence it presents during the signal cycle. The reader is referred to the MUTCD for further information on this topic. The intent of this chapter is to highlight specific issues that affect signal timing and not replicate the information contained within the MUTCD.

### 4.1.2 Traffic Signal System Design

Traffic signals may operate in a system of intersections. The application of timing plans depends on the infrastructure available in the signal system. The typical hardware components of a signal system are shown in Figure 4-1. These components are described below.

**Figure 4-1 Physical components of a signal system**




1. **Detection.** Detectors are used to gather information about the "conditions" used in the local traffic signal controller at each intersection and to allocate time at the intersection. Detection may also be used to collect data that can support monitoring, managing, and measuring performance. Few agencies achieve all three functions with their detection system.

2. **Local Controller.** (further described in Chapter 5 and 6) The local controller operates the displays through the load switches using the signal timing provided by the user. The local controller implements specific strategies from field inputs or directly from the central signal system operator.

3. **Master Controller.** The master controller is an optional component of a system that facilitates communication between the "central" signal system and the local controllers for control decisions. The primary functions of the on-street master controller are to select the timing plans for a group of intersections, to process and store detector count information, and to monitor equipment operation.

4. **Traffic Control Center (Signal System).** The traffic control center is used for a variety of functions. The central signal system is an alternative to using a master controller and contains the

operational database that stores controller data, monitors the system, and allows timing and other parameters to be modified. It may also facilitate the implementation of the advanced concepts described in Chapter 9.

5. **Communication.** Several components of the system communicate through many forms. Communication facilitates coordination between controllers (see below).

These system components works together to implement a control strategy, such as coordination. To understand the implementation of the strategy, one must have a rudimentary understanding of how a system is configured and how coordination is maintained at a system level. This is particularly important for troubleshooting.

## 4.2 PHASING OVERVIEW

Phasing represents the fundamental method by which a traffic signal accommodates the various users at an intersection in a safe and efficient manner. This section first presents definitions and terminology and then discusses the variety of phasing techniques used in common practice.

### 4.2.1 Definitions and Terminology

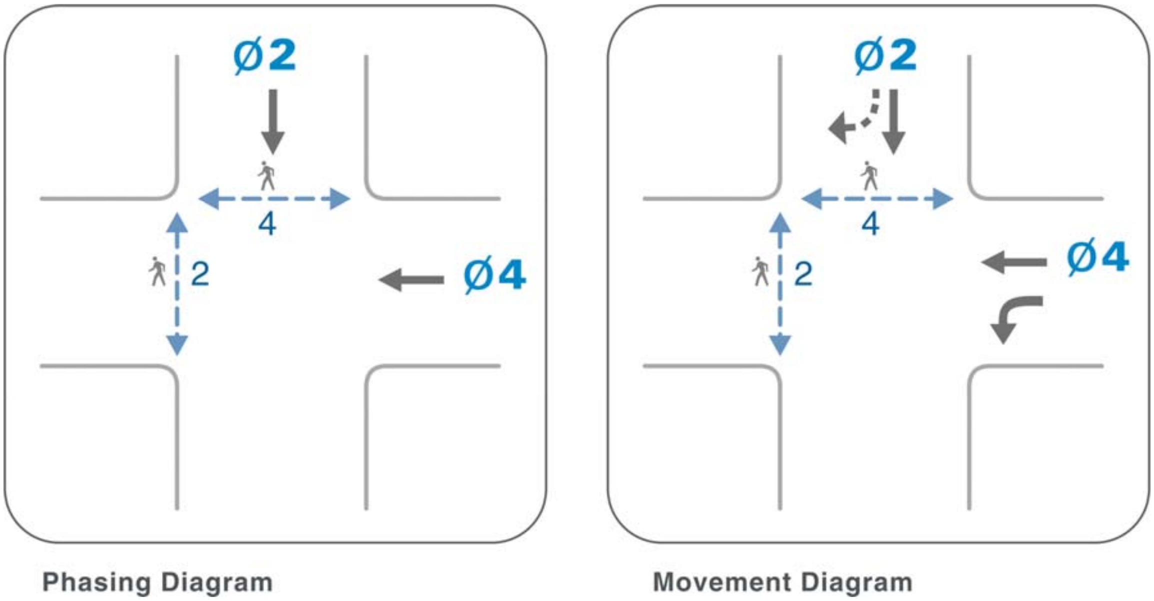
Definitions used in traffic signal timing have resulted in some confusion. In many cases, "the old way of doing things" confounds the ability of the industry to educate people new to the field. Over the years, the description of the "individual movements" of the dual-ring 8-movement controller as "phases" has blurred into common communicated terminology of "movement number" being synonymous as "phase number". Most signal designs and all controllers sold today provide eight standard phases within the signal controller; however, a four-phase intersection is commonly referred to in the literature to represent a standard four-legged intersection with protected left turns on all approaches. It is the intent of this document to reduce the ambiguous use of the term "phase" and for the purposes of this document, ***a phase is defined as a controller timing unit associated with the control of one or more movements.*** Each phase at an intersection has a set of timing, possibly containing vehicle and pedestrian timing. A phase may control both a through movement and a right turn movement on an approach.

> **The MUTCD defines a signal phase as the right-of-way, yellow change, and red clearance intervals in a cycle that are assigned to an independent traffic movement or combination of traffic movements..**

Two additional terms that are important for improving the use of terminology within the signal timing industry is to articulate is the use of movement and interval. A movement reflects the user perspective and is defined by the user type and the action that is taken (turning movement for a vehicle or pedestrian crossing). Two different types of movements include those that have the right of way and those that must yield consistent with the rules of the road or the Uniform Vehicle Code.

A simple example of the concept of movements is the intersection of one-way streets shown in Figure 4-1. In this example, the intersection is operated by two phases (2 and 4) and pedestrians are accommodated as concurrent movements to the traffic. Phase 2 will include a through and a right turn movement, while phase 4 will have a through and a left turn movement (appropriate turning movements are omitted from the diagram for simplicity). The right turn on phase 2 must yield to pedestrian traffic crossing the west leg of the intersection.

**Figure 4-2 Phasing and movement diagram for the intersection of two one-way streets**




An interval is a duration of time during which the signal indications do not change. For example, a pedestrian phase contains three intervals—Walk, Flashing Don't Walk, and solid Don't Walk—and within the Walk and Flashing Don't Walk intervals, the corresponding through movement will remain green.

The movements served at an intersection can be categorized by the various users: vehicles, pedestrians, bicyclists, and transit. Figure 4-2 illustrates the typical vehicle and pedestrian movements at a four-leg intersection. In the context of this figure, bicycle and transit movements track the same paths as vehicle movements. These movements are regulated by the signal controller through their allocation to one or more signal phases. The following list defines some of the terms used to describe vehicle and pedestrian phasing (*4*):

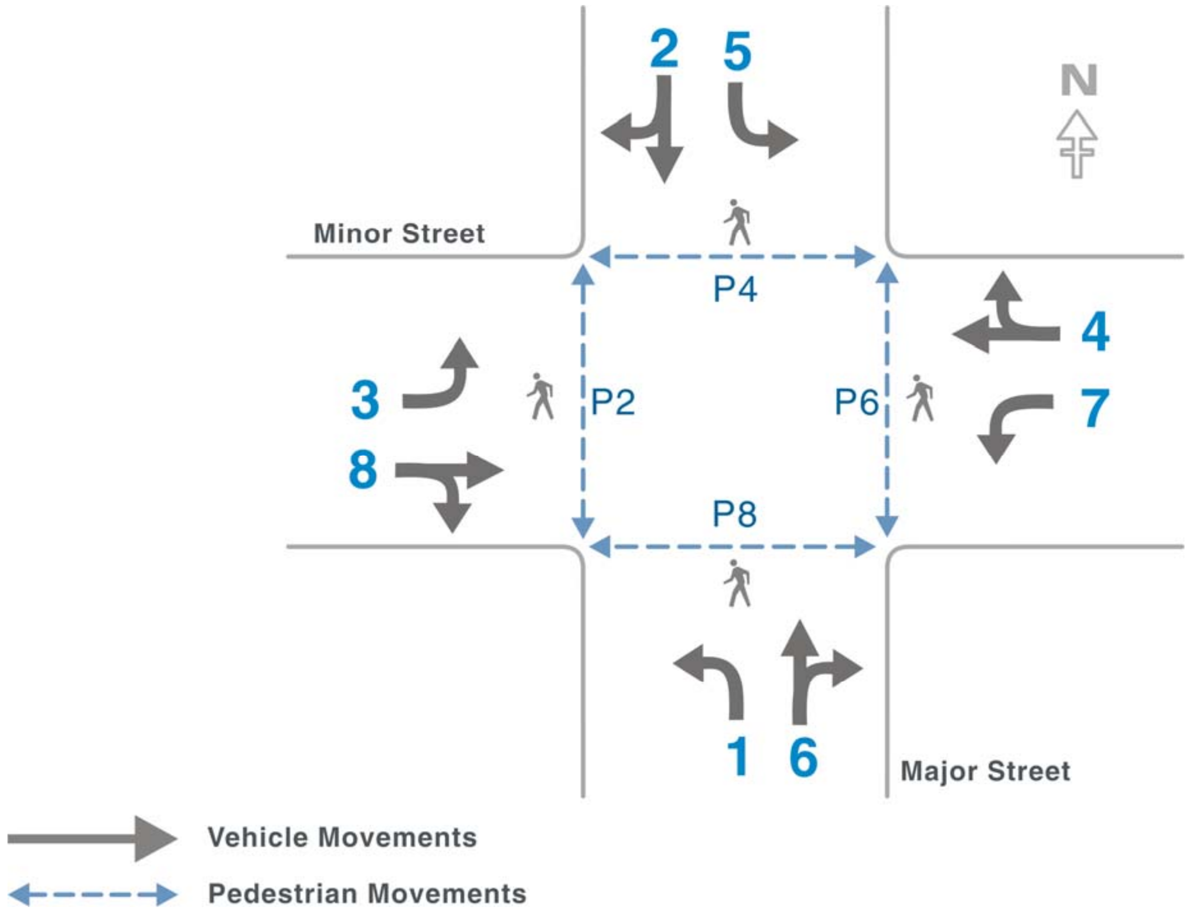
- A vehicular phase is defined as a phase that is allocated to one or more vehicular traffic movements, as timed by the controller unit.
- A pedestrian phase is defined as a traffic phase allocated to pedestrian traffic that may provide a pedestrian indication either concurrent with one or more vehicular phases, or separate from all vehicular phases.
- A traffic phase is defined as the green, change, and clearance intervals in a cycle assigned to specified movement(s) of traffic.
- A cycle is defined as the total time to complete one sequence of signalization for all movements at an intersection. In an actuated controller unit, the cycle is a complete sequence of all signal indications.

The assignments shown are typical for eight-phase controller operation although other assignments are possible. As can be seen in Figure 4-3, left-turn movements are assigned odd number phases, while through movements are assigned to even number phases. In this example, the southbound left-turn movement is protected and is associated with phase 5. The right-turn movements are not typically assigned to separate phases. For example, the westbound right-turn movement is compatible with the westbound through movement and thus shares phase 4.

Typically, a pedestrian movement is associated with the concurrent vehicular phase running parallel and adjacent to it. As indicated earlier, this phase may include the concurrent right-turn movement that is associated with the through movement. As illustrated in Figure 4-3, pedestrians crossing the northern leg of the intersection are assigned the concurrent westbound vehicular phase

(phase 4), which conflicts with the eastbound left turn (phase 3), this is further described in the next section. The operation of a concurrent phase is influenced by both the vehicular and pedestrian movements that it serves.

**Figure 4-3 Typical vehicular and pedestrian movements at a four-leg intersection**




### 4.2.2 Ring-and-Barrier Diagrams

Modern U.S. practice for signal control organizes phases by grouping them in a continuous loop (or ring) and separating the crossing or conflicting traffic streams with time between when they are allowed to operate, either by making the movements sequential or adding a barrier between the movements. The ring identifies phases that may operate one after another and are typically conflicting phases organized in a particular order. For instance, it may be desirable to separate the traffic traveling through the intersection in the northbound direction from the southbound left turn movement. A change interval and clearance time is used to separate that movement in time.

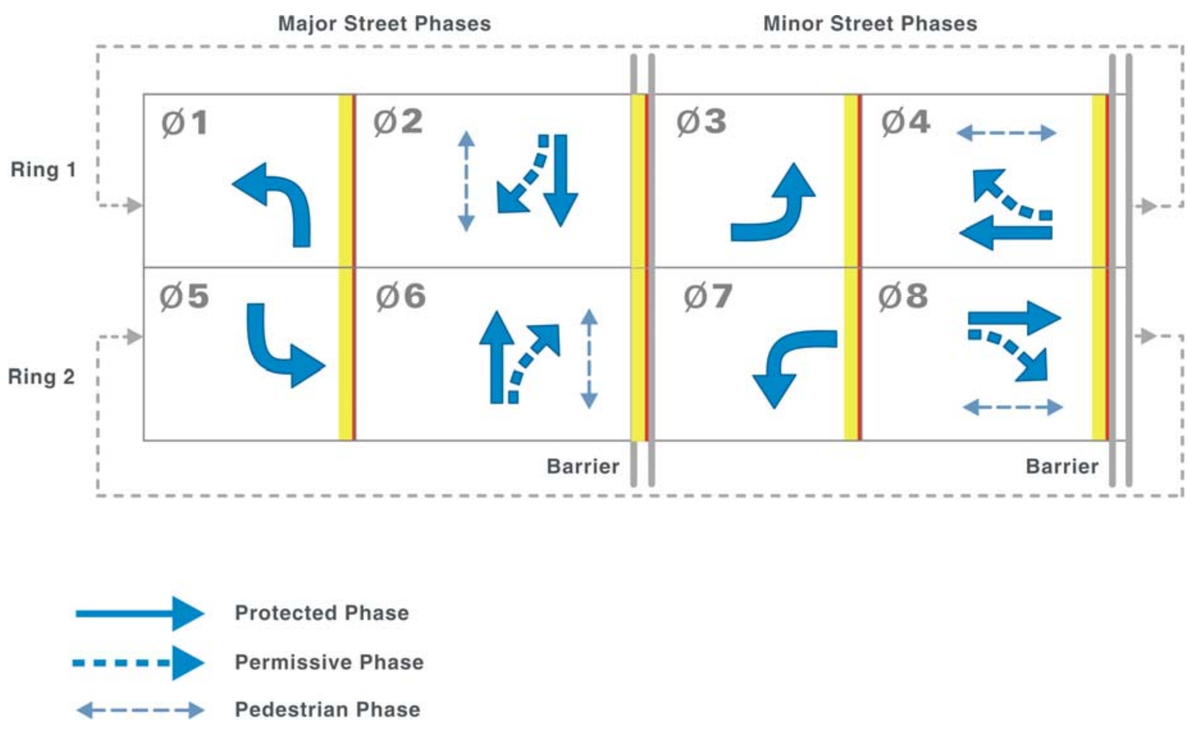
In Figure 4-4, a “barrier” would be used to separate the east-west movements from north-south movements to avoid operating conflicting movements at the same time. They are also used to define a relationship between the rings to assure compatible movements. The barrier represents a reference point in the cycle at which a phase in each ring has reached a point of termination; both rings must cross the barrier simultaneously.

The time sequence of phases can be described using a ring and barrier diagram. Some common rules for number phases at an intersection (which are applicable to the diagram) are provided in the following list, which assumes leading left turns and separate left-turn phases.

- Phases 1, 2, 3, and 4 are assigned to Ring 1. Phases 5, 6, 7, and 8 are assigned to Ring 2.
- Phases 1, 2, 5, and 6 are assigned to Barrier 1. Phases 3,4,7, and 8 are assigned to Barrier 2.
- A phase pair contains two phases within the same ring and barrier that cannot be displayed concurrently. Examples of phase pairs include 1+2, 5+6, 3+4, and 7+8. Phases within a phase pair can be reversed (e.g. 2+1, 6+5, etc.) Phase pairs within the same barrier must end simultaneously (i.e., end at the barrier). For example, phase pairs 1+2 and 5+6 must end simultaneously at the end of barrier 1 and phase pairs 3+4 and 7+8 must end simultaneously at the end of barrier 2.
- Phase pair 1+2 can operate concurrently with phase pair 5+6. Phase pair 3+4 can operate concurrently with phase pair 7+8. These phase pairs are also known as concurrency groups because they can time together.
- One common practice is to assign phases 2 and 6 to the major street movements and the phases on the other side of the barrier to the minor road movements. Another practice is to define this by direction (phase 4 may be the most northerly pointed phase).

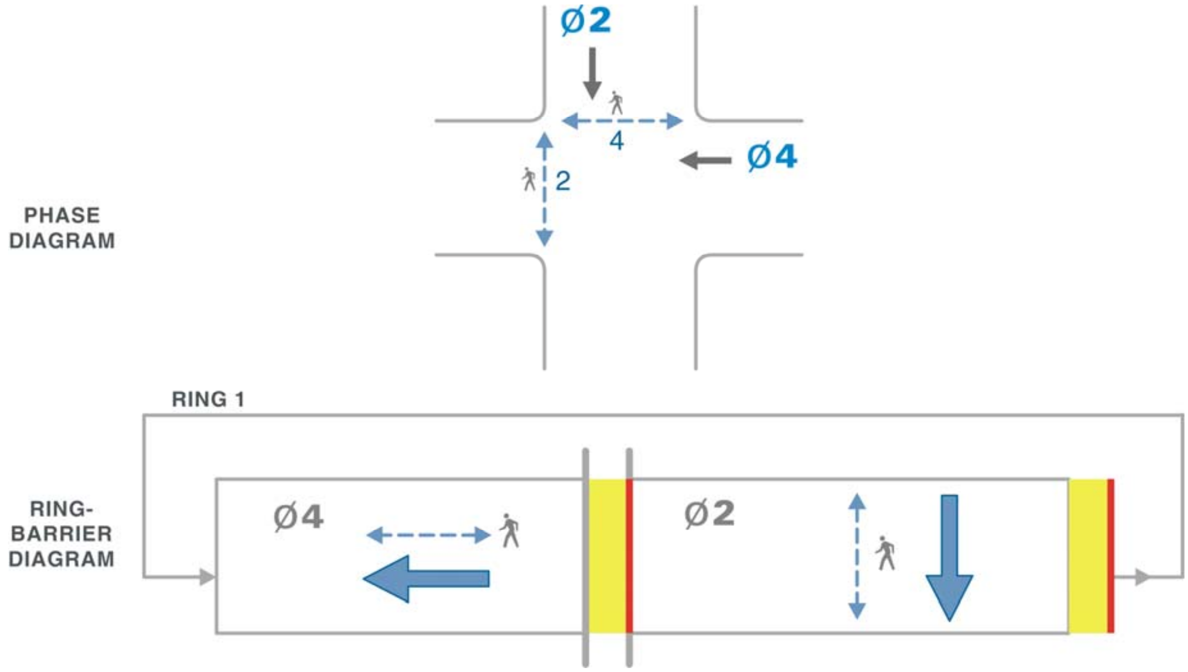
An example ring diagram is shown in Figure 4-4. The sequence of phases is shown as they occur in time, proceeding from left to right. The figure illustrates a phase sequence with left-turn movements leading the opposing through movements on both the major and minor streets. The diagram shows phase 1 and 5 ending at the same time, but they operate independently and can end at different times. The subsequent phase (phases 2 and 6 respectively) may begin once the previous phase has used its time. Once the barrier is crossed phases 3 and 7 operate followed by phases 4 and 8. The cycle ends with the completion of phases 4 and 8.

**Figure 4-4 Standard ring-and-barrier diagram**




When considering signal phasing it is helpful to start at the most basic, one where two one-way streets intersect. Figure 4-5 shows a ring and barrier structure that compliments the phase diagram described previously in Figure 4-3. All movements on the street are served (vehicles, bicycles, and pedestrians) in one direction during one phase, and all movements on the cross street during the other phase. It should be noted that it is common practice to assign pedestrians to a through traffic movement, with the underlying assumption that a motorist or bicyclist making a right turn must yield to pedestrians as per most Uniform Vehicle Code (*5*). However, care should be exercised with this concurrent pedestrian phasing approach when there are exclusive right-turn lanes with high right turn volumes (*6*].

**Figure 4-5 Ring-and-barrier diagram for intersection of two one-way streets**




In this example, the left-turning movements are either non-existent or prohibited (phase 2) or are protected (phase 4). Application of this concept at a more typical intersection of two-way streets uses the standard ring and barrier structure described in Figure 4-4.

**Assignment of phase numbers to signalized intersections is somewhat arbitrary based on historical design principles, but there are some rules that have been applied to standardize operation.**

Depending on the complexity of the intersection, 2 to 8 phases are typically used, although some controllers can provide up to 40 phases to serve complex intersections or sets of intersections. Developing an appropriate phasing plan begins with determining the left-turn phasing type at the intersection. Section 4.4 presents some of the guidance related to selection of appropriate left turn phasing based on traffic volumes and safety experience.

## 4.3 LEFT-TURN DISPLAY OPTIONS

There are five options for the left-turn phasing at an intersection: permissive only, protected only, protected-permissive, split phasing, and prohibited. Phasing can have a significant impact on signal system effectiveness for a number of reasons, including:

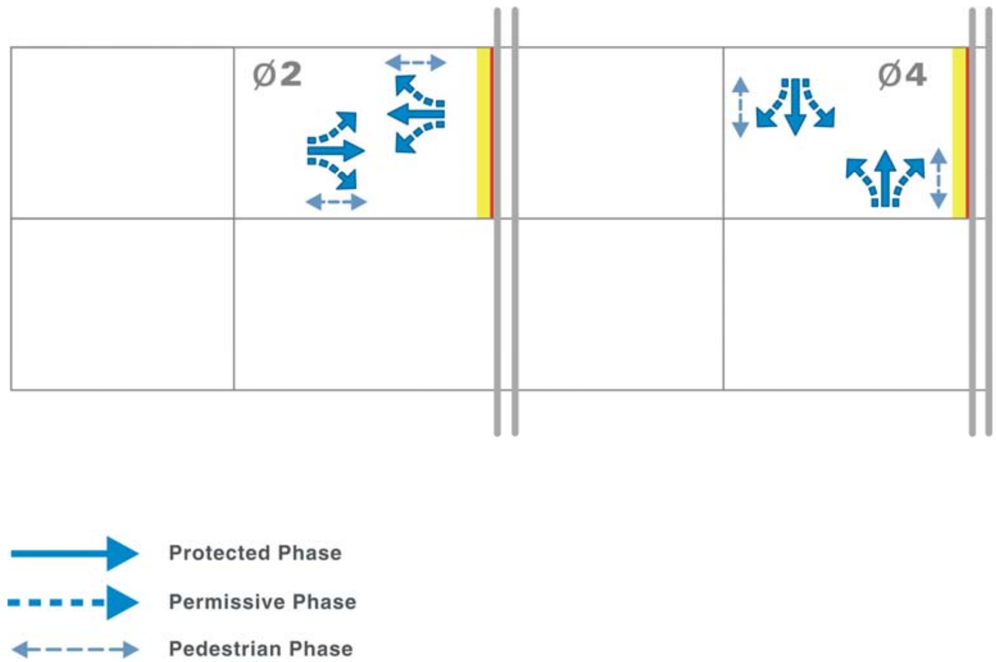
- Permissive only left turn operation may reduce delay for the intersection, but may adversely affect intersection safely, because it requires motorists to choose acceptable gaps.
- Protected only left-turn phases may reduce delay for turning vehicles but are likely to increase overall intersection delay.
- Protected-permissive left turn phases can offer a good compromise between safety and efficiency but could limit available options to maximize signal progression during coordination unless innovative displays are used.
- Split phasing may be applicable with shared lanes, but could increase coordinated cycle length if both split phases are provided a concurrent pedestrian phase.
- Prohibited left turns may be used selectively to reduce conflicts at the intersection.

### 4.3.1 Permissive Only Left-Turn Phasing

Permissive only operation requires left-turning drivers to yield to the conflicting vehicle and pedestrian traffic streams before completing the turn. In the permissive mode, the left-turn movement is served concurrently with the adjacent through movement. Both the left turn and the opposing through movements are presented with a circular green indication. Thus, in this left turn display option, a green arrow is never provided.

Permissive operation is primarily used when traffic is light to moderate and sight distance is adequate. This display option provides the most efficient operation for green allocation at the intersection. The efficiency of this mode is dependent on the availability of gaps in the conflicting streams through which the turn can be safely completed. This mode can have an adverse affect on safety in some situations, such as when the left-turn driver's view of conflicting traffic is restricted or when adequate gaps in traffic are not present. The yellow trap can occur for the “permissive only” left turn when the opposite direction has a lagging left turn movement. Figure 4-6 and the following figures are adapted from those presented in the Signalized Intersections: Informational Guide Report and thus, phase 2 is defined as the eastbound movement.

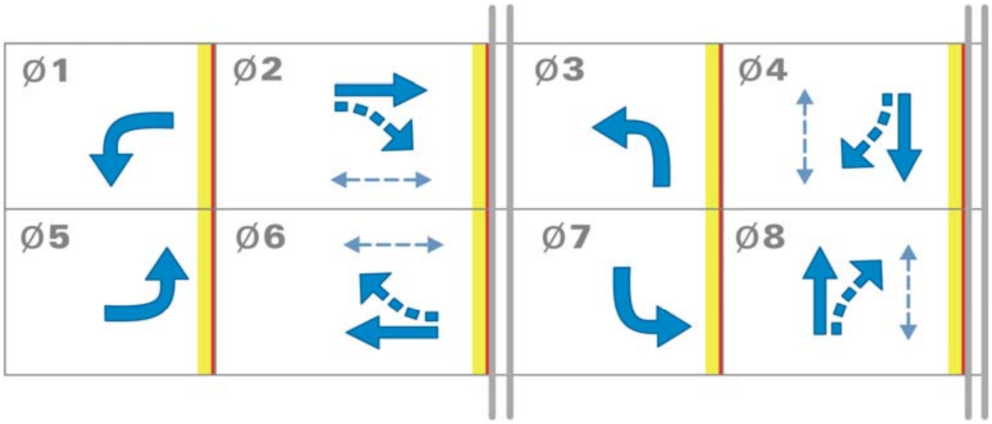
**Figure 4-6 Ring-and-barrier diagram showing permissive phasing**




### 4.3.2 Protected Only Left-Turn Phasing

Protected only operation assigns the right-of-way to drivers turning left at the intersection and allows turns to be made only on a green arrow display. This operation provides for efficient left-turn movement service; however, the added left-turn phase increases the lost time within the cycle length and may increase delay to the other movements. An exclusive left-turn lane is typically provided with this phasing as shown in Figure 4-7. The left-turn phase is indicated by a green arrow signal indication. This type operation is recognized to provide the safest left-turn operation.

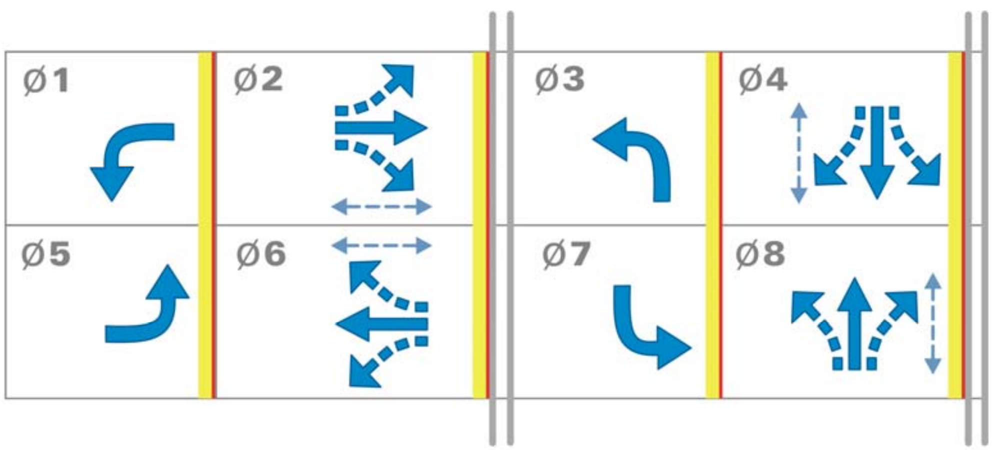
**Figure 4-7 Ring-and-barrier diagram showing protected phasing**




### 4.3.3 Protected-Permissive Left-Turn Phasing

Protected-permissive operation represents a combination of the permissive and protected modes. Left-turning drivers have the right-of-way during the protected left-turn phase. They can also complete the turn "permissively" when the adjacent through movement receives its circular green indication as illustrated in Figure 4-8. This mode provides for efficient left-turn movement service, often without causing a significant increase in delay to other movements. This mode also tends to provide a relatively safe left-turn operation, provided that adequate sight distance is available and turns during the permissive component can be safely completed.

**Figure 4-8 Ring-and-barrier diagram showing protected-permissive phasing**

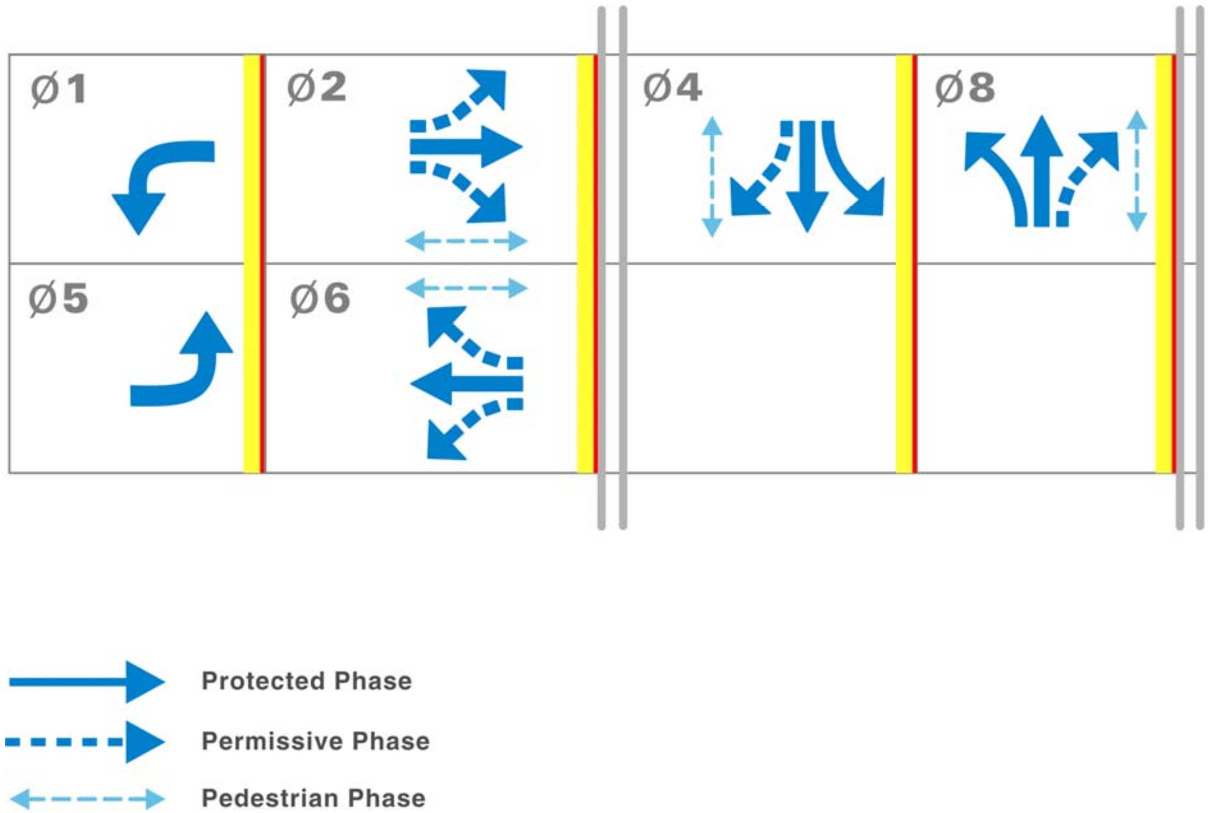



Protected-permissive phasing should be used with caution when a phasing sequence other than lead-lead left-turn phasing is being deployed. Section 4.4.2 provides additional detail on this matter.

### 4.3.4 Split Phasing

Split phasing represents an assignment of the right-of-way to all movements of a particular approach, followed by all of the movements of the opposing approach. This is depicted in Figure 4-9. Spilt phasing may be necessary when intersection geometry results in partially conflicting vehicle paths through the intersections or where the approaches are offset such that left turning vehicles would have to occupy the same space to complete their turns. Split phasing avoids the conflict of opposing left turn vehicle paths. Similarly, if the intersection has high left turn and through volume, the traffic engineer may have to use shared left turn and through lanes to make efficient use of the approach which would also result in split phasing for the approach.

**Figure 4-9 Ring-and-barrier diagram showing split phasing**




Split phasing may be helpful if any of the following conditions are present (*7*):

- There is a need to accommodate one or more left-turn lanes on each approach, but sufficient width is not available to ensure adequate separation in the middle of the intersection. This problem may also be caused by a large intersection skew angle.
- The larger left-turn lane volume is equal to its opposing through lane volume during most hours of the day ("lane volume" represents the movement volume divided by the number of lanes serving it.)
- The width of the road is constrained such that an approach lane is shared by the through and left-turn movements yet the left-turn volume is sufficient to justify a left-turn phase.
- One of the two approaches has heavy volume, the other approach has minimal volume, and actuated control is used. In this situation, the phase associated with the low-volume approach would rarely be called and the intersection would function more nearly as a "T" intersection.
- Crash history indicates an unusually large number of sideswipe or head-on crashes in the middle of the intersection and involving left-turning vehicles.

This phasing is typically less efficient than other types of left-turn phasing. It typically increases the cycle length, or if the cycle length is fixed, reduces the time available to the intersecting road.

### 4.3.5 Prohibition of Left-Turns as a Phasing Option

Prohibition of left turns on an approach is an option that has been implemented in some cases to maintain mobility at an intersection. In this case, a supplemental sign may be provided that indicates "no left turn". In some cases, these have been applied only during certain times of day, when gaps in traffic are unavailable and operation of permitted phasing may be unsafe. Figure 4-10 is an example from Toronto, Ontario, that prohibits left turns during the morning and evening periods.

**Figure 4-10 Prohibited left turns by time of day**


In general, the operational mode used for one left-turn movement on a road is also typically used for the other (opposing) left-turn movement. For example, if one left-turn movement is permissive, the opposing left turn is also permissive. However, this agreement is not required and the decision of mode should be movement-specific based on factors such as sight distance, volumes, number of turning lanes, number of opposing lanes, and leading vs lagging left turn operation.

### 4.3.6 Guidelines for Selecting Left-Turn Phasing

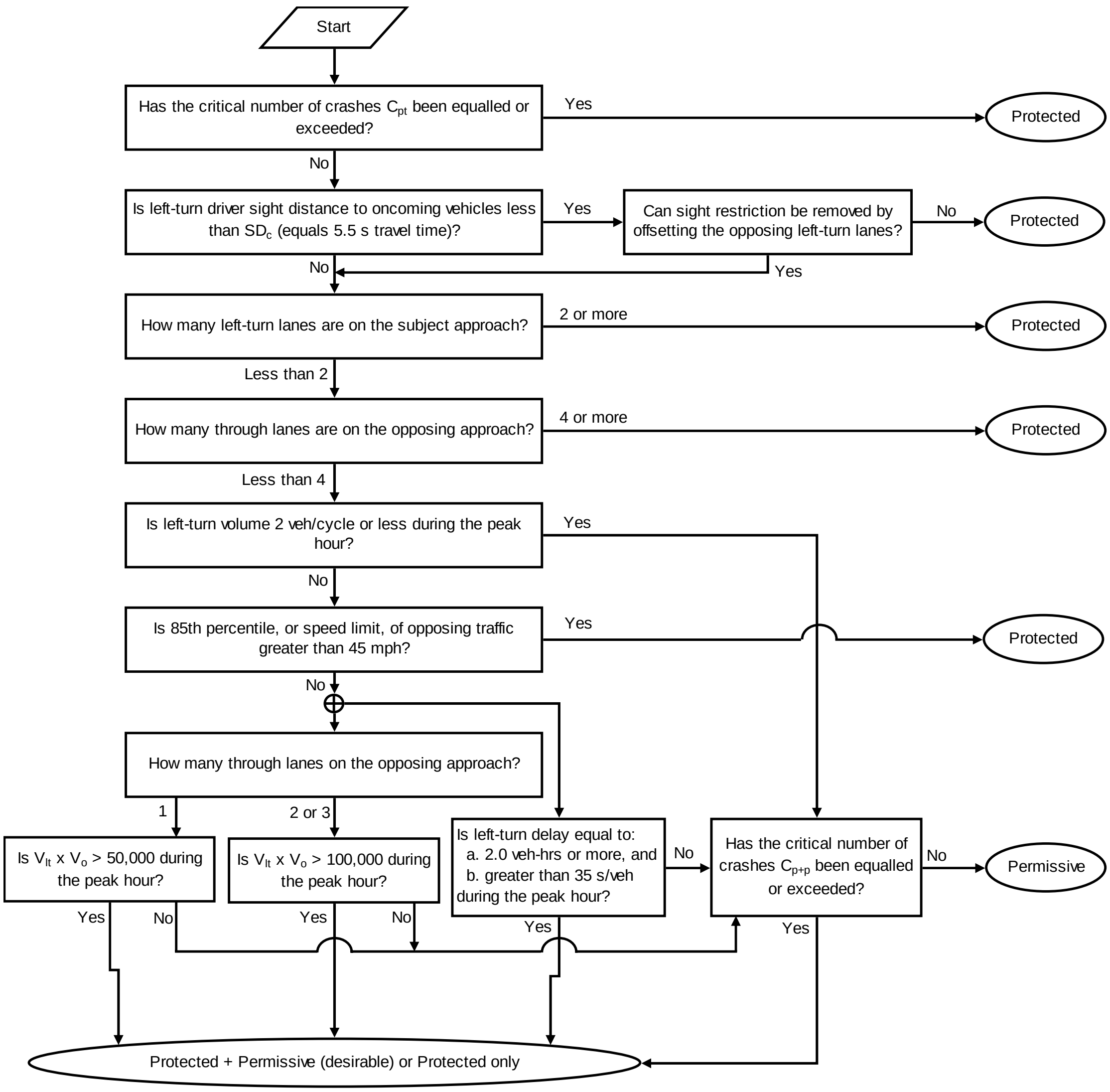
A variety of guidelines exist that have been developed to indicate conditions where the benefits of a left-turn phase typically outweigh its adverse impact to intersection operation. Many of these guidelines indicate that a left-turn phase can be justified based on consideration of several factors that ultimately tie back to the operational or safety benefits derived. These factors include:

- Left-turn and opposing through volumes
- Number of opposing through lanes
- Cycle length
- Speed of opposing traffic
- Sight distance
- Crash history

The flowchart shown in Figure 4-11 can be used to assist in the determination of whether a left-turn phase is needed and whether the operational mode should be protected or protected-permissive.

These guidelines were derived from a variety of sources (*8*;*9*). Application of the flowchart requires the separate evaluation of each left-turn movement on the subject road.

**Figure 4-11 Guidelines for determining the potential need for a left-turn phase**




| Number of Left-turn Movements on Subject Road | Period During Which Crashes are Considered (years) | Critical Left-Turn-Related Crash Count: When Considering Protected-only, $C_{pt}$ (crashes/period) | Critical Left-Turn-Related Crash Count: When Considering Prot.+Perm, $C_{p+p}$ (crashes/period) |
|---|---|---|---|
| One | 1 | 6 | 4 |
| One | 2 | 11 | 6 |
| One | 3 | 14 | 7 |
| Both | 1 | 11 | 6 |
| Both | 2 | 18 | 9 |
| Both | 3 | 26 | 13 |

| Oncoming Traffic Speed Limit (mph) | Minimum Sight Distance to Oncoming Vehicles, $SD_c$ (ft) |
|---|---|
| 25 | 200 |
| 30 | 240 |
| 35 | 280 |
| 40 | 320 |
| 45 | 360 |
| 50 | 400 |
| 55 | 440 |
| 60 | 480 |

Variables

$V_{lt}$ = left-turn volume on the subject approach, veh/h

$V_o$ = through plus right-turn volume on the approach opposing the subject left-turn movement, veh/h

Source: Adapted from (Kell and Fullerton, 1998; Orcutt, 1993; Traffic Engineering Manual, 1999).

The objective of the flow chart is to identify the least restrictive left-turn operational mode. A secondary objective is to provide a structured procedure for the evaluation of left-turn phasing for the purpose of promoting consistency in left-turn phase application.

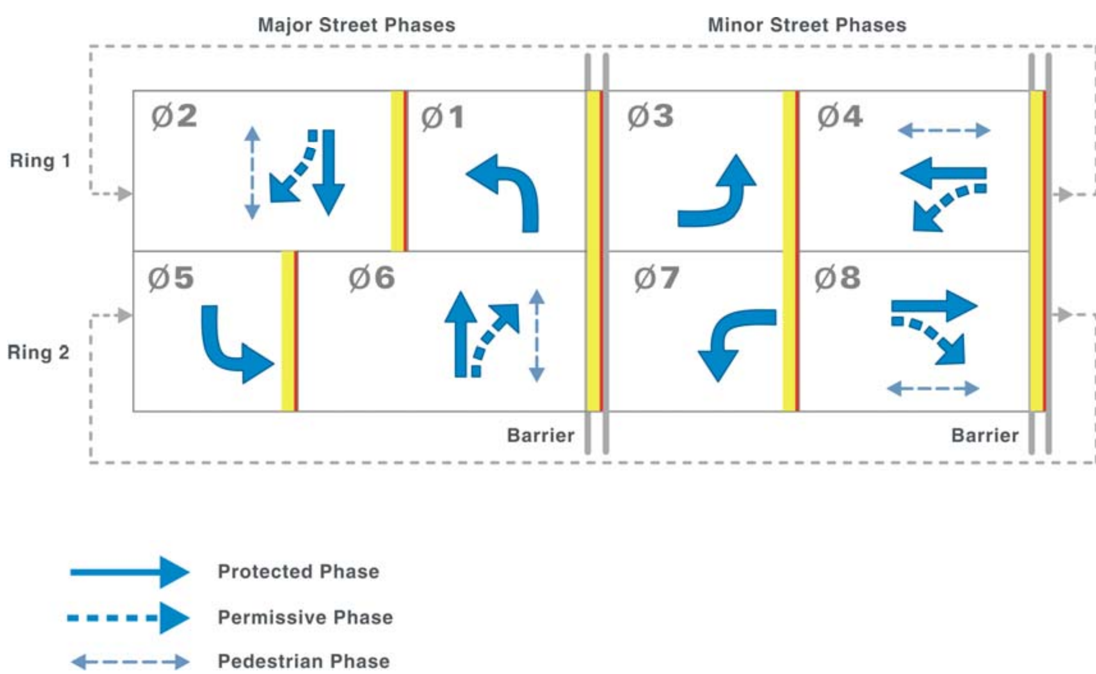
The critical left-turn crash counts identified in the figure are based on an underlying average critical crash frequency and recognize the inherent variability of crash data. The underlying averages are 1.3 crashes per year and 3.0 crashes per year when considering protected-permissive and protected only left-turn phasing, respectively. If the reported crash count for existing permissive operation exceeds the critical value, then it is likely that the subject intersection has an average left-turn crash frequency that exceeds the aforementioned average (5 percent chance of error) and a more restrictive operational mode would likely improve the safety of the left-turn maneuver.

The flowchart has two alternative paths following the check of opposing traffic speed. One path requires knowledge of left-turn delay; the other requires knowledge of the left-turn and opposing through volumes. The left-turn delay referred to in the flowchart is the delay incurred when no left-turn phase is provided (i.e., the left-turn movement operates in the permissive mode).

## 4.4 LEFT-TURN PHASE SEQUENCE OPTIONS

It may be advantageous under certain circumstances to change the sequence in which left turns are served relative to their complementary through movements. This is done by reversing the sequence of a pair of complementary phases, as is shown for phases 1 and 2 in Figure 4-4. In this example, phase 1 is said to “lag” phase 2. Specifically, Figure 4-12 shows phases 2 and 6 starting and ending at different times in the cycle. This independence between the through phases can be desirable under coordinated operations because it can accommodate platoons of traffic arriving from each direction at different times.

**Figure 4-12 Ring-and-barrier diagram showing protected lead-lag left turns**




### 4.4.1 Lead-Lead Left-Turn Phase Sequence

The most commonly used left-turn phase sequence is the "lead-lead" sequence which has both opposing left-turn phases starting at the same time. If a single ring structure is used, then the two phases also end at the same time. If an actuated dual ring structure is used, then each left-turn phase is assigned to a different ring such that each can end when the left-turn demand is served (i.e., they can end at different times). The advantages of this phasing option are: (1) that drivers react quickly to the leading green arrow indication and (2) it minimizes conflicts between left-turn and through movements on the same approach, especially when the left-turn volume exceeds its available storage length (or no left-turn lane is provided). A more detailed discussion of the advantages of leading left-turn phases is provided in Chapter 13 of the *Traffic Engineering Handbook* (*10*).

### 4.4.2 Lag-Lag Left-Turn Phase Sequence

This left-turn phase sequence is most commonly used in coordinated systems with closely spaced signals, such as diamond interchanges. It has both opposing left-turn phases ending at the same time. If it is implemented in a single ring structure, then the two phases also start at the same time. If a dual-ring structure is used, then each left-turn phase is assigned to a different ring such that each can start when the left-turn demand is served (i.e., they can start at different times).

Lagging left-turn phasing is also recognized to offer operational benefits for the following special situations:

- At "T" intersections when the one left-turn phase that exists is combined with a protected-permissive mode.
- At the intersection of a two-way street and a one-way street where the one left-turn phase that exists is combined with a protected-permissive mode.
- At an interchange, or a pair of closely-spaced interconnected intersections, where both intersections have a left-turn phase and each are combined with the protected-permissive mode.

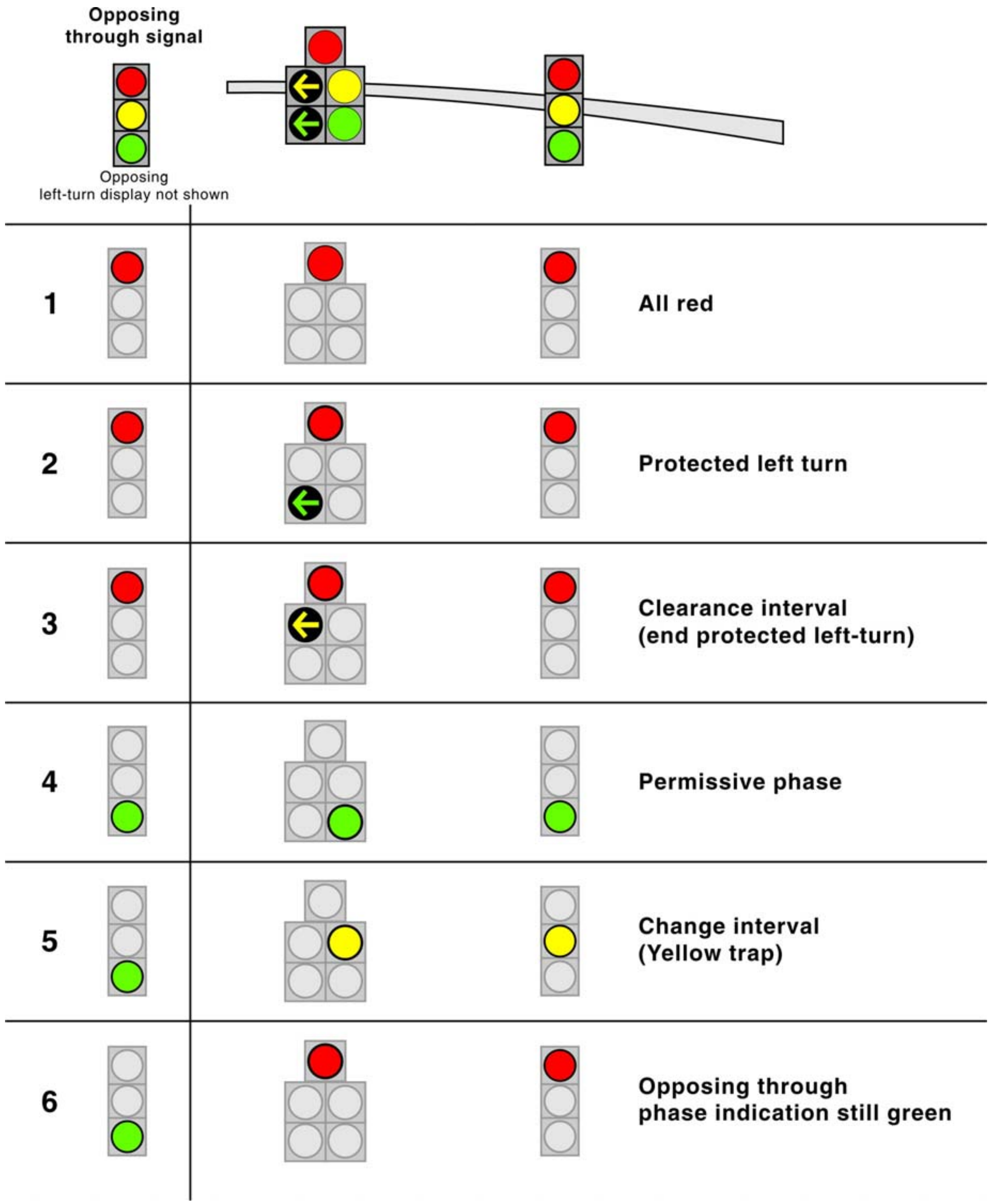
When used with protected phasing, this phase sequence provides a similar operational efficiency as a lead-lead or lead-lag phase sequences. However, differences emerge when they are used with protected-permissive mode. One disadvantage of lagging left-turn phases is that drivers tend not to react as quickly to the green arrow indication. Another disadvantage is that, if a left-turn bay does not exist or is relatively short, then queued left-turn vehicles may block the inside through lane during the initial through movement phase.

When lag-lag phasing is used at a four-leg intersection where both phases are used with the protected-permissive mode, then both left-turn phases must start at the same time to avoid the "yellow trap" (or left-turn trap) problem, illustrated in Figure 4-13. This problem stems from the potential conflict between left-turning vehicles and oncoming vehicles at the end of the adjacent through phase. Of the two through movement phases serving the subject street, the trap is associated with the first through movement phase to terminate and occurs during this phase's change period. The left-turn driver seeking a gap in oncoming traffic during the through phase, first sees the yellow ball indication; then incorrectly assumes that the oncoming traffic also sees a yellow indication; and then turns across the oncoming traffic stream without regard to the availability of a safe gap.

The "yellow trap" problem can be alleviated by using one of the following techniques:

- Use the protected-only mode for both left-turn movements.
- Use a single-ring structure to ensure that both through movement phases end at the same time (use with the protected-permissive mode).
- Use a flashing yellow arrow or "Dallas Display" for both left-turn signal heads and use the protected-permissive mode for both left-turn movements.

**Figure 4-13 Illustration of the yellow trap**




Source: FHWA Signalized Intersection Guide

In fact, under at least one condition, the second technique can operate more efficiently than dual-ring lead-lead phasing. This condition occurs when the left-turn volume is moderate to heavy and relatively equal on both approaches. Regardless, a detailed operational evaluation should always be used to confirm that lag-lag phasing operates more efficiently than other phasing options.

The third technique avoids the yellow trap by using an overlap in the controller and a five-section left-turn signal head. An overlap is a controller output (to the signal head load switch) that is associated with two or more phases. In this application, the left-turn green and yellow arrow indications are associated with the subject left-turn phase; and the left-turn green, yellow, and red ball indications are associated with the opposing through movement phase (as opposed to those of the adjacent through phase). Two types of protected-permissive displays have been developed to provide more operational flexibility and avoid the “yellow trap” problem: the flashing yellow arrow and the “Dallas Display”.

The flashing yellow arrow is contained within a three-, four-, or five-section head and provides a permissive indication to the driver that operates concurrent with the opposing through movement rather than the adjacent through movement. Researchers for National Cooperative Highway Research Program (NCHRP) Project 3-54 studied alternatives to the green ball indication for permissive left turn movements. This study was conducted over a 7-year period and comprised a very comprehensive research process, including engineering analyses, static and video-based driver comprehension studies, field implementation, video conflict studies, and crash analyses. This study (*11*) recommended that a flashing yellow arrow be allowed as an alternative to the circular green for permissive left-turn intervals.

- The flashing yellow arrow was found to be the best overall alternative to the circular green as the permissive signal display for a left-turn movement.
- A flashing yellow arrow was found to have a high level of understanding and correct response by left-turn drivers, and a lower fail-critical rate than the circular green.
- The flashing yellow arrow display in a separate signal face for the left-turn movement offers more versatility in field application. It is capable of being operated in any of the various modes of left-turn operation by time of day, and is easily programmed to avoid the "yellow trap" associated with some permissive turns at the end of the circular green display.

Subsequent to the publication of the NCHRP research, FHWA has issued an Interim Approval and no longer requires a Request to Experiment for this display. As of this writing, it has been proposed for inclusion in the next revision to the MUTCD.

In the “Dallas Display”, the signal head uses louvers on the yellow and green ball indications to restrict visibility of the left-turn display to adjacent lanes. The louvered signal head is referred to as the "Dallas Display." With this display, both left-turn phases can operate in the protected-permissive mode and the trap is avoided.

#### 4.4.3 Lead-Lag Left-Turn Phase Sequence

This left-turn phase sequence is generally used to accommodate through movement progression in a coordinated signal system. The aforementioned "yellow trap" may occur if the leading left-turn movement operates in the protected-permissive mode and the two through movement phases time concurrently during a portion of the cycle. The "yellow trap" problem can be alleviated by using one of the following techniques:

- Use the protected only mode for the leading left-turn movement.
- Use a single ring structure that ensures that the two through movement phases do not time concurrently (use with the protected-permissive mode).
- Use a flashing yellow arrow or "Dallas Display" for the left-turn signal head associated with the leading left-turn phase and use the protected-permissive mode for both left-turn movements.

The first two techniques will likely have an adverse effect on operations, relative to a dual ring implementation of lead-lag phasing with protected-permissive operation. However, they avoid the potential adverse effect a yellow trap would have on safety.

The third technique avoids the yellow trap by using an overlap in the controller and a five section left-turn signal head. This use of overlap was described in the previous discussion of lag-lag left-turn phasing. However, in practice, the Dallas Display is used for *both* the leading and the lagging left-turn signal heads because it improves operational performance (*12*).

Lead-lag phasing is also recognized to offer operational benefits for the following special situations:

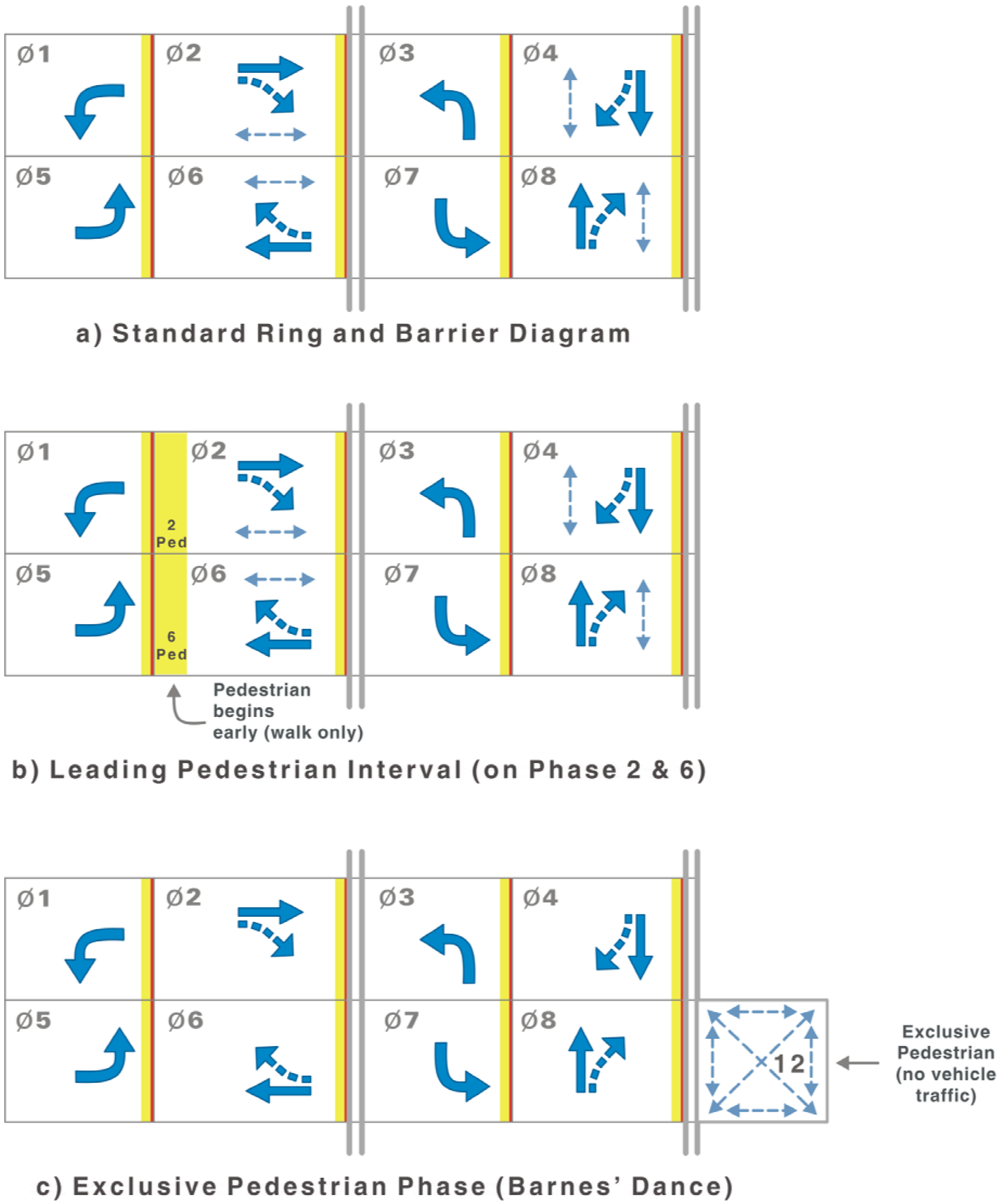
- Where there is inadequate space in the intersection to safely accommodate the simultaneous service of the opposing left-turn movements. However, in this application, a single ring structure (or equivalent functionality in a dual ring structure) should be used to ensure that the two left-turn phases never time concurrently.
- Intersections where the leading left-turn movement is not provided an exclusive lane (or the available left-turn storage is relatively small).

## 4.5 PEDESTRIAN PHASING

Pedestrian movements are typically served concurrently with the adjacent through movement phase at an intersection. This is done to simplify the operation of the intersection primarily and is largely a legacy issue in our application of signal logic and control. Typical application of pedestrian operation puts pedestrians in conflict with right-turning vehicles and left-turning vehicles that operate in a permissive mode, by inviting their movement at the same time. There are specific measures that can be used to mitigate this potential conflict, three common options include:

- *Leading pedestrian interval.* As shown in Figure 4-14b, a leading pedestrian interval starts a few seconds before the adjacent through movement phase. This allows pedestrians to establish a presence in the crosswalk and thereby reduce conflicts with turning vehicles. This option supports improved safety for pedestrians by allowing them increased visibility within the intersection and is applicable to intersections where there are significant pedestrian-vehicle conflicts (*13*).
- *Lagging pedestrian interval.* A lagging pedestrian interval option operates similarly to a leading pedestrian interval, except that the pedestrian walk interval starts several seconds after the adjacent through movement phase. This option allows a waiting right-turn queue to clear before the pedestrian walk indication is presented and reduces conflicts with right-turning vehicles. It is applicable to intersections where there is: (1) a high right-turn volume and (2) either an exclusive right-turn lane (or lanes) or the two intersecting roads have one-way traffic (*14*).
- *Exclusive pedestrian phase (also “pedestrian scramble” or “Barnes’ Dance”)* As shown in Figure 4-14c, an exclusive pedestrian phase dedicates an additional phase for the exclusive use of all pedestrians. This additional phase is configured such that no vehicular movements are served concurrently with pedestrian traffic. During this phase, pedestrians can cross any of the intersection legs and may even be allowed to cross the intersection in a diagonal path. This type of phasing has an advantage of reducing conflicts between right-turning vehicles and pedestrians, but it comes at a penalty of reduced vehicular capacity and longer cycle lengths (which increases delay to all users). The exclusive pedestrian phase is not frequently used but can be found in the central business districts of several cities, including Denver and San Francisco.

**Figure 4-14 Ring-barrier diagrams showing a leading pedestrian interval and an exclusive pedestrian phase**



## 4.6 RIGHT-TURN PHASING

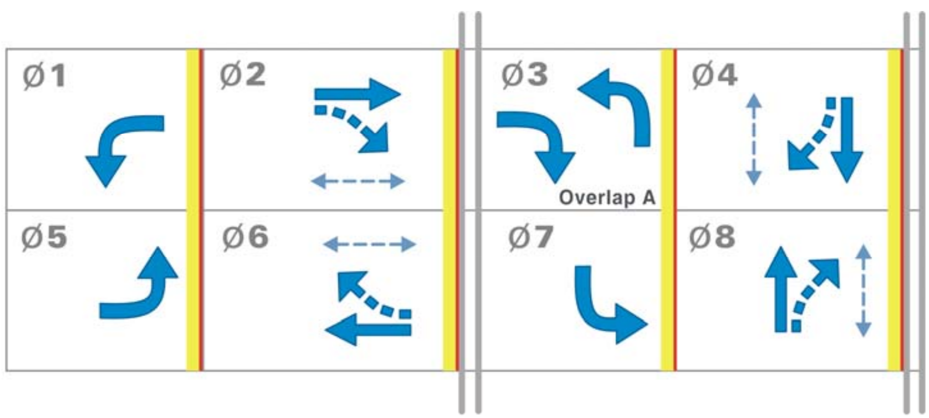
Two types of right-turn phasing are addressed in this section. The first type is based on the addition of a phase to the signal cycle that exclusively serves one or more right-turn movements. This type of right-turn phasing is rarely used. If it is being considered, then its operational or safety benefits should be evaluated and shown to outweigh its adverse impact on the efficiency of the other intersection movements.

The second type of right-turn phasing is based on the assignment of the right-turn movement to the phase serving the complementary left-turn movement on the crossroad. The following conditions should be satisfied before using this type of right-turn phasing:

- The subject right-turn movement is served by one or more exclusive right-turn lanes.
- The right-turn volume is high (300 vehicles per hour or more) and is a critical movement at the intersection (see Chapter 3 for more details).
- A protected left-turn phase is provided for the complementary left-turn movement on the intersecting road.
- U-turns from the complementary left-turn are prohibited.

If the aforementioned conditions are satisfied, then the appropriate operational mode can be determined. If the through movement phase for the subject intersection approach serves a pedestrian movement, then the right-turn phasing should operate in the protected-permissive mode. As shown in Figure 4-15, the permissive right-turn operation would occur during the adjacent through movement phase, and the protected right-turn operation would occur during the complementary left-turn phase.

**Figure 4-15 Phase diagram illustrating a right-turn overlap**




If the through movement phase for the subject intersection approach does not serve a pedestrian movement, then the right-turn phasing should operate in the protected only mode during both the adjacent through movement phase and the complementary left-turn phase. A controller overlap may be used to provide this sequence.

## 4.7 DETECTION FUNDAMENTALS

Detection at an intersection informs the signal controller that a user desires service (often called "demand" for service). Detectors place calls into the traffic signal controller. The controller uses this information and the signal timing to determine the display provided to the users. Detection for pedestrians is limited in most cases to push buttons as shown in Figure 4-16, although accessible pedestrian signal detectors are increasing in their use. The FHWA's *Pedestrian Facilities User Guide-Providing Safety and Mobility (15)* is a resource describing more detail related to other equipment.

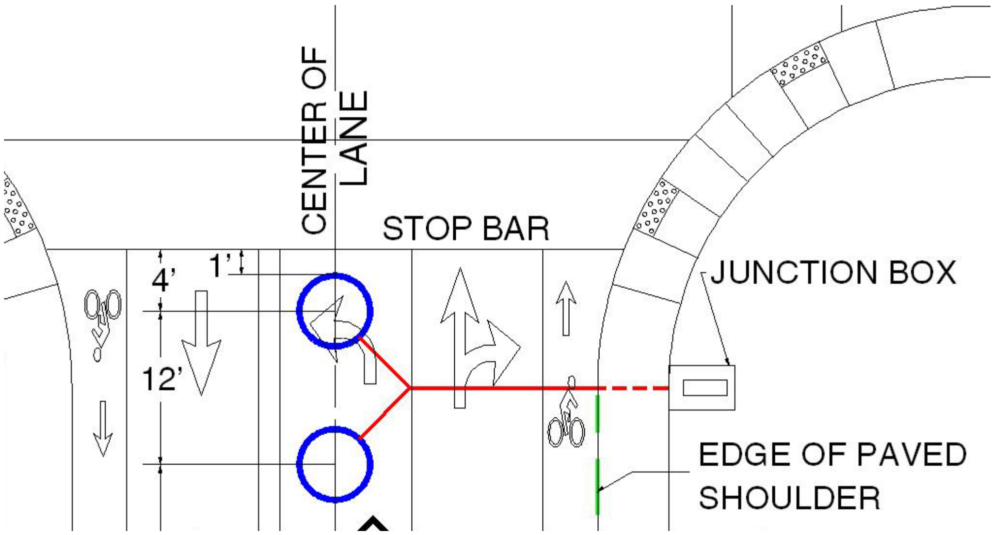
**Figure 4-16 Examples of pedestrian push buttons**


Source: Fred Ranck,FHWA, Illinois, Naperville, intersection of Washington Street at Shuman Blvd

There are various forms of vehicle detection technologies, and strengths and weaknesses of each are described in the *Traffic Detector Handbook, 3rd* edition (*2006).* The detection design for an intersection describes the size, number, location, and functionality of each detector. Most engineering drawings include the wiring diagram for how detectors are associated to phases. Signal timing settings such as the passage time, delay, extend, and other related parameters are described in more detail in Chapter 5.

The size and location of detectors is an important element in traffic signal design. Detectors can consist of one 6-foot-by-6-foot inductive loop detector, a series of closely spaced 6-foot-by-6-foot loop detectors (may be circular in shape as shown in Figure 4-17), one long (6-foot-by-40-foot) loop detector, or alternative detection technology (e.g., video, microwave, etc.) that can monitor an equivalent length of roadway. This detection zone can be used to meet the objectives described below.

**Figure 4-17 Example of vehicle detector design (through lane on side street)**




(a) Schematic design of loop detector design

(b) Physical representation of loop detectors

A "call" represents the controller's registration of a request for service on a specified phase. A call can be triggered by an actuation from any detection, vehicular, pedestrian, or other or through a controller function. When triggered by a vehicle detection unit, the call's start and end time can be equal to that of the actuation from the detection unit, or it can be modified using the controller's detector parameters (e.g., delay, extend, call, queue, etc.). These parameters are described in Chapter 5.

### 4.7.1 Detection Design Objectives

The objective of detection is to detect vehicle presence and identify gaps in vehicle presence that are sufficiently long to warrant terminating the phase. There are many objectives of detection design that can be characterized with the following statements:

1. Identify vehicle presence on a phase.
2. Extend the phase to serve queued traffic and that which is progressed from upstream traffic signals.
3. Identify gaps in traffic where the phase may be ended and extend the green.
4. Provide a safe phase termination for high-speed movements by minimizing the chance of a driver being in the indecision zone at the onset of the yellow indication.

The first and fourth objectives are safety related. The first objective addresses expectancy, while the fourth specifically addresses the potential crashes as a result of phase termination. The fourth objective is achieved by using advance detectors on the approach. The location of these detectors can vary and depends on the detection technology used as well as intersection approach speed. The safety benefit of this design tends to be more significant on high-speed approaches. The other objectives focus on intersection efficiency.

The second and third objectives are designed to address efficiency. During low volume (late night or off-peak) conditions, detection should seek to serve all traffic identified without stopping. In peak conditions, green allocation should seek to measure flows and maintain flows that are near saturation flow as described in Chapter 3. Detection timing to achieve this objective will be discussed in more detail in Chapter 5.

### 4.7.2 Detector Operating Modes

The detection operating mode refers to the way the detection unit measures activity and is set in the detection unit. It also affects the duration of the actuation submitted to the controller by the detection unit. One of two modes can be used: presence or pulse. The presence mode is typically the default mode.

#### *Pulse Mode*

Pulse mode is used to describe a detector which detects the passage of a vehicle by motion only (point detection). This is characterized by a short "on" pulse sent to the controller of 0.10 to 0.15 seconds duration. The actuation starts with the arrival of the vehicle to the detection zone and ends after the pulse duration. This mode is typically used when the detectors are located upstream of the stop line and the associated detector channel operates in the locking mode.

#### *Presence Mode*

Presence mode is used to measure occupancy and the actuation starts with the arrival of the vehicle to the detection zone and ends when the vehicle leaves the detection zone. Thus, the time duration of the actuation depends on the vehicle length, detection zone length, and vehicle speed.

> **Presence mode measures the time that a vehicle is within the detection zone and will require shorter extension (or gap) timing with its use.**

Presence mode is typically used with long-loop detection located at the stop line. In this application, the associated detector channel in the controller is set to operate in the non-locking mode (see next section). In this mode, the delay or extend parameters in the controller (described in detail in Chapter 5) can be used to modify the call start and end times. Alternatively, the delay or extend functions in the detection unit could also be used to adjust the start and end time of the actuation. The combination of presence mode operation and long-loop detection typically require a small passage time value to maintain efficiency. This characteristic tends to result in an operational benefit through efficient queue service.

> **Modifiers to the detector settings are commonly handled in the controller. This has increased in popularity because it provides a single location for information on all phases at the intersection.**

### 4.7.3 Controller Memory Modes

Controller memory modes refer to the controller's ability to "remember" (i.e., retain) a detector actuation. One of two modes can be used: non-locking or locking. This mode is set in the controller for each of its detector channel inputs. It dictates whether an actuation received during the red interval (and optionally, the yellow interval) is retained until the assigned phase is served by the controller. All actuations received during the green interval are treated as non-locking by the controller. The non-locking mode is typically the default mode.

#### *Non-locking Mode*

In the non-locking mode, an actuation received from a detector is not retained by the controller after the actuation is dropped by the detection unit. The controller recognizes the actuation only during the time that it is held present by the detection unit. In this manner, the actuation indicates to the controller that a vehicle is present in the detection zone and the controller converts this actuation into a call for service. This mode is typically used for phases that are served by stop line detection. It allows permissive movements (such as right-turn-on-red) to be completed without invoking a phase change. In doing so, it improves efficiency by minimizing the cycle time needed to serve minor movement phases. Non-locking mode is not typically used with pulse detection due to an inability to detect vehicle presence after the pulse duration elapses.

#### *Locking Mode*

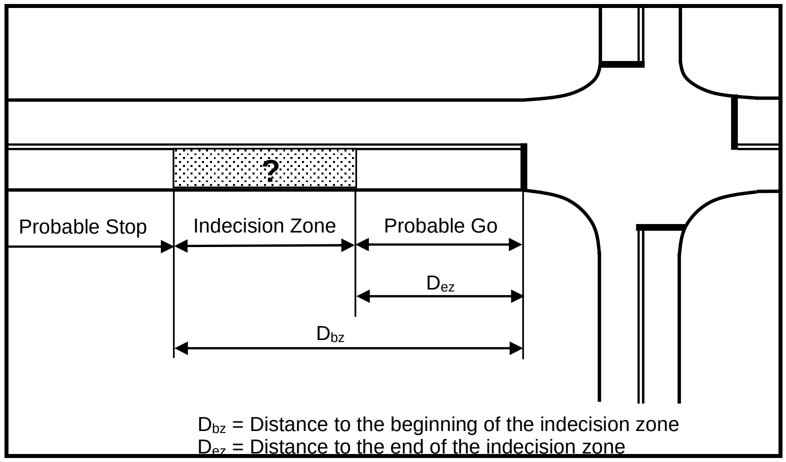
In the locking mode, the first actuation received by the controller on a specified channel during the red interval is used by the controller to trigger a continuous call for service. This call is retained until the assigned phase is serviced, regardless of whether any vehicles are waiting to be served. This mode is typically used for the major-road through movement phases associated with a low percentage of turning vehicles (as may be found in rural areas). One advantage of using this mode is that it can eliminate the need for stop line detection, provided that advance detection is provided and that it is designed to ensure efficient queue service.

### 4.7.4 Detection Design for High-Speed Approaches

Detection designs for high speed approaches (speeds greater than 35 mph) have the objective to not only service the queue at the beginning of green but also to safely terminate the phase in the presence of a conflicting call. Stop bar detection is usually used to clear the queues and the multiple upstream detectors are used to safely terminate the phase. For efficient operation, the stop bar detector should be programmed as a queue detector so that the stop bar detector is disconnected after the queue clears and only the upstream detectors are used to safely terminate the phase. When stop bar detectors are not used, volume density functions should be used to provide appropriate minimum green time to clear the queues.

The design of advance detection on high-speed approaches requires special attention. Drivers within a few seconds travel time of the intersection tend to be indecisive about their ability to stop at the onset of the yellow indication. This behavior yields an "indecision zone" (also known as a "dilemma zone") in advance of the stop line wherein some drivers may proceed and others may stop. The location of this zone is shown in Figure 4-18.

**Figure 4-18 Indecision zone boundaries on a typical intersection approach**

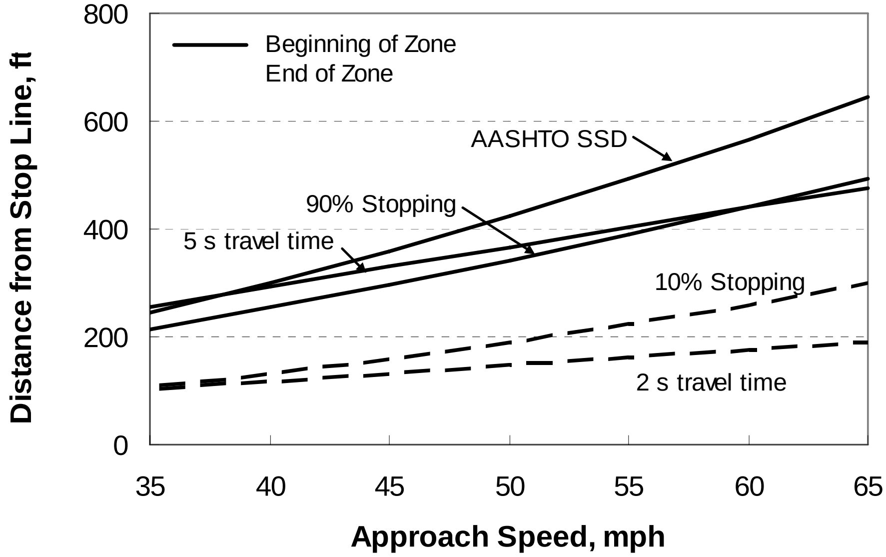



The indecision zone location has been defined in several ways.

- *Distance from stop line.* Some researchers have defined it in terms of distance from the stop line (*16*;*17*). They define the beginning of the zone as the distance beyond which 90 percent of all drivers would stop if presented a yellow indication. They define the end of the zone as the distance within which only 10 percent of all drivers would stop. The distance to the beginning of the zone recommended by Zegeer and Deen corresponds to about 5 seconds of travel time. That recommended by ITE increases exponentially with speed, ranging from 4.2 to 5.2 s travel time, with the larger values corresponding to higher speeds.
- *Travel time.* Another definition of the indecision zone boundary is based on observed travel time to the stop line. Chang et al. (*18*) found that 85 percent of drivers stopped if they were more than 3 s from the stop line, regardless of their speed. Similarly, they found that drivers less than 2 s from the stop line would almost always continue through the intersection.
- *Stopping sight distance.* A third definition of the beginning of the indecision zone is based on safe stopping sight distance (SSD). A method for computing this distance is described in Chapter 3 of the AASHTO document, *A Policy on the Geometric Design of Highways and Streets* (*19*).

The zone boundaries obtained by these three definitions are compared in Figure 4-19. The boundaries based on distance typically have an exponential relationship. Those based on travel time have a linear relationship. Based on the trends shown in the figure, the beginning and end of the indecision zone tend to be about 5.5 and 2.5 seconds, respectively, travel time from the stop line. These times equate to about the 90$^{th}$-percentile and 10$^{th}$-percentile drivers, respectively.

**Figure 4-19 Distance to the beginning and end of the indecision zone**




For these types of designs, the furthest detector upstream of the stop bar is usually located at the beginning of the indecision zone of the approach design speed ($85^{th}$-percentile approach speed). This is usually at a distance of 5 to 5.5 seconds of travel time. Subsequent detectors have a design speed of 10 mph lower than the upstream detector. Typically 3 to 4 detectors are used to enable safe termination of the high speed approach phase. The detectors are allowed to extend the phase by the passage time programmed in the controller or by the extension time on the detector itself (see Chapter 5).

Although the concept of the indecision zone has been known for many years, comprehensive research has not been completed to conclude what the appropriate way to address the human factors associated with intersection design and signal timing display. The indecision should not be confused with the dilemma faced by drivers determining whether there is distance to stop and, if not, to travel through the intersection before a conflicting movement receives a green indication.

### 4.7.5 Detection Design for Low-Speed Traffic Movements

Detection design for low speed traffic movements (speeds of 35 mph or less) has a different objective compared to the detection design for high speed traffic movements. The primary objective of detection on low speed approaches is to call a phase and clear the queue while minimizing delay. Due to lower speeds, there is less emphasis on protection from dilemma zone or indecision zone on the approach. Hence, some agencies use only stop bar detection for low speed approaches. Stop bar detectors are usually operating in the presence mode. This facilitates the primary objective of detector calling the phase and clearing the queues.

## 4.8 DETECTION APPLICATIONS

This section presents a series of detection applications.

### 4.8.1 Basic Fully-Actuated Design

This design is based on the following assumptions:

- A detection zone located at the stop line.
- The detectors are operating in the presence mode.
- Non-locking memory is used for the associated detector channel in the controller.

- No recall is used for the phase to which the detector is assigned.

The key element of this design is the determination of detection zone length. The optimal length represents a trade off in the desire to avoid both premature gap out and excessive extension of green. According to Lin (*20*), the ideal length of the stop line detection zone is about 80 feet. This length allows the passage time setting to be small such that the design is very efficient in detecting the end of queue while minimizing the chance of a premature gap-out. However, the installation and maintenance of such a long detector is often cost prohibitive and multiple detectors of a shorter length are often used. The following guidelines should be used to determine the appropriate length of the stop line detection zone:

- The detection zone should not be smaller than 20 feet.
- The zone should be positioned such that a queued vehicle cannot stop between its trailing edge and the crossroad.
- The zone should consist of one long detection zone or a series of loops. Other sensors that can provide the equivalent length of detection can also be used.

### 4.8.2 Volume-Density Design

This design is based on the following assumptions:

- One 6-foot detector is located upstream of the stop line.
- The detector unit is operating in the pulse mode.
- Locking memory is used for the associated detector channel in the controller.
- No recall is used for the phase to which the detector is assigned.

This design is not well-suited to approaches with a significant percentage of turning vehicles because these vehicles may unnecessarily submit a call for the through movement phase.

A key element of this design is the location of the detector. This location should be based on the desired maximum allowable headway for the design. Research has shown that maximum allowable headways in the range of 1.8 to 2.5 seconds yield the snappiest operation, values of 2.6 to 4.5 seconds will allow detected vehicles to use the green, but may result in extension of green during low flow periods (this is discussed further in Chapter 5). Lower values are more appropriate for higher volume conditions.

The advance detector should be located such that the travel time from the detector to the stop line for a vehicle traveling at the 85$^{th}$ percentile speed is equal to the maximum allowable headway. In this manner, the green interval is not unnecessarily extended by a vehicle that has crossed the stop line (as it is with stop line detection). Based on this principle, the recommended location for the detector is listed in Table 4-1 for a range of approach speeds.

With this design, the passage time setting is equal to the maximum allowable headway (which may vary if gap-reduction is used). Guidelines related to the detection design as it relates to signal timing are described in Chapter 5 for determining the appropriate minimum green interval. Gap-reduction is not essential with a maximum allowable headway of 3.0 s but should be considered for a maximum allowable headway of 4.0 s. If gap-reduction is used, then the minimum gap should be equal to 2.0 s. Guidelines for setting the time before reduction parameter and the time to reduce parameter are described in Chapter 5.

**Table 4-1 Recommended distance between stop line and detector**

| 85th Percentile Approach Speed, mph | Distance between Detector and Stop Line, ft | |
|---|---|---|
| | 3.0 s Max. Allowable Headway | 4.0 s Max. Allowable Headway [1] |
| 20 | 90 | 115 |
| 25 | 110 | 145 |
| 30 | 130 | 175 [2] |
| 35 | 155 | 205 [2] |
| 40 | 180 | 235 [2] |
| 45 | 200 | 265 [2] |

Notes:
1 - Use the controller's gap-reduction feature with a minimum gap 2.0 s.
2 - Use the controller's variable initial feature to ensure the minimum green duration is not unnecessarily long during low-volume conditions (see Chapter 5, Table 5-3).

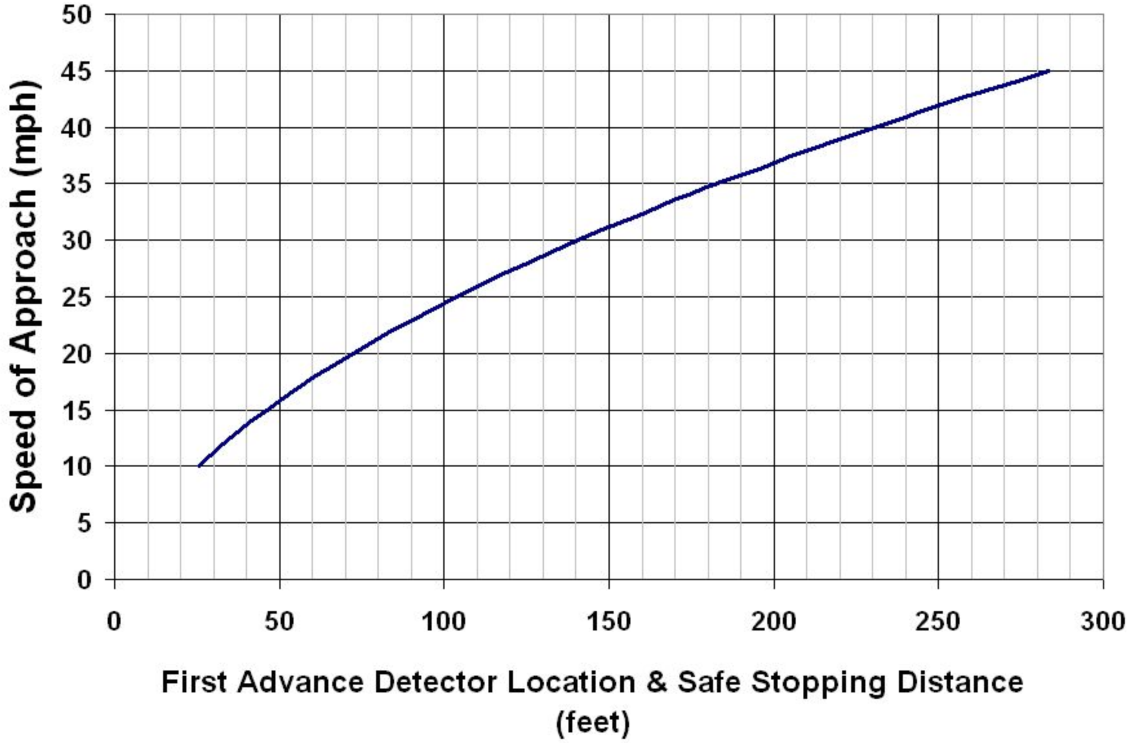
### 4.8.3 Multiple-Detector Design

This design technique, used by the city of Portland, Oregon, is based on the following assumptions:

- One 6-foot detector is placed at 60 feet (because safe stopping distance ends at 15 mph), this is the last detector the vehicle will cross.
- One 6-foot detector is located at the safe stopping distance upstream of the stop line, based on the approach speed; this is the first detector the vehicle will cross.
- At higher speeds, one or two additional detector(s) is(are) placed between the first advance detector and the last detector (at 60 feet).
- The detectors operate in presence mode.
- Non-locking memory is used for the associated detector channel in the controller.
- No recall is used for the phase to which the detector is assigned.

The first detector registers demand for the vehicles as they approach the stop bar. The second detector is used like a speed sieve to determine whether the vehicle can stop safely upon its approach to the intersection (based on the safe stopping distance from the second detector. This design is well-suited to approaches with a significant percentage of turning vehicles because it is likely these vehicles will gap out as the vehicle slows to make the turn at the intersection or upstream driveway. This is further illustrated in Figure 4-20.

**Figure 4-20 Detector location for multiple detector technique (level approach)**




Source: City of Portland, Oregon

A key element of this design is the use of detectors and extend values within the controller timing to reduce the time between successive vehicles. A sample table of extend values from the City of Portland is shown in Table 4-2.

**Table 4-2 Recommended detector locations and timing settings for multiple detector technique**

| Approach Speed, mph | Location of Advance Detector (feet) | Location of Second Detector (feet) | Location of Detector Nearest Stop Bar (feet) | Extend (Carryover) Value from Upstream Detectors (secs)[1] |
|---|---|---|---|---|
| 25 | 105 | 60 | | 0.8 |
| 30 | 140 | 60 | | 2.1 |
| 35 | 183 | 115 | 60 | 1.0 |
| 40 | 230 | 130 | 60 | 1.6 |
| 45 | 283 | 190 | 115 & 60 | 1.0 |

*Notes*:

1 - A minimum gap of 0.5 seconds is used to allow the vehicle to leave the last detector.

Source: City of Portland

### 4.8.4 Left-Turn Movements

The guidelines provided in this section can be used to design the left-turn movement detection when this movement has an exclusive lane (or lanes). In general, the detection design for a left-turn movement should follow the guidelines offered for through movements, as described previously for basic fully-actuated design in Section 4.8.1.

Some agencies use an additional detector approximately 100 feet upstream of the stop bar to provide some advance notice of a vehicle approaching the intersection. In this case, an extension (carryover) setting is provided to allow a vehicle leaving the first advance detector to reach the detection zone at the stop bar and a short gap setting is used for the phase (assuming stop bar detection in presence mode).

In cases where the left-turn movement operates in the protected-permissive mode and the conflicting through movement phase is on recall, then the stop line detection zone could be set back from the stop line one or two vehicle lengths. This technique assumes that one or two vehicles can be served at the end of the conflicting through phase. A call for the left-turn phase would only be registered if the left-turn queue extends back over the detection zone, in which case there would be more vehicles that could clear at the end of the through phase.

If the left-turn movement operates in the protected-permissive mode and the conflicting through movement phase is likely to rest in recall, then the controller's delay parameter can be used with the channel assigned to the left-turn lane detector to minimize unnecessary phase changes. The delay value used should range from 3 to 15 seconds, with the larger values used when higher speeds and volumes exist on the conflicting approach. Most controllers have a detector switching setting that can be used to send calls from left-turn detectors to the adjacent through phases to extend the permissive green phase for both movements.

If the left-turn movement operates in the permissive or protected-permissive mode and the adjacent through movement phase is not on recall, then it may be desirable to extend the stop line detection zone beyond the stop line. This extension would be intended to minimize the potential for stranding a turning vehicle in the intersection at the end of the permissive period. It may also ensure that the left-turn movement is detected under low-volume conditions.

### 4.8.5 Right-Turn Movements

The guidelines provided in this section can be used to design the right-turn movement detection when this movement has an exclusive lane (or lanes). In general, the detection design for a right-turn movement should follow the guidelines offered for through movements, as described previously for basic fully-actuated design in Section 4.8.1.

If the right-turn volume is moderate-to-high and the volume on the intersecting road is relatively low (such that many gaps for right-turn-on-red exist), then the controller's delay parameter can be used with the channel assigned to the right-turn lane detector to minimize unnecessary phase changes. The delay parameter should also be considered when a conflicting phase is on recall. The delay value used should range from 7 to 15 s, with the larger values used when higher speeds and volumes exist on the intersecting road.

If the right-turn volume is high but there are few gaps for right-turn-on-red, then the use of the delay parameter may not be appropriate because it may only increase the delay to the right-turning vehicles.

## 4.9 REFERENCES


1 Rodegerdts, L.A., B. Nevers, B. Robinson, J. Ringert, P. Koonce, J. Bansen, T. Nguyen, J. McGill, D. Stewart, J. Suggett, T. Neuman, N. Antonucci, K. Hardy, and K. Courage. *Signalized Intersections: Informational Guide*. Report FHWA-HRT-04-091. Federal Highway Administration, Washington, D.C., 2004.

2 Wu, C-S., C.E. Lee, R.B. Machemehl, and J. Williams. "*Effects of Multiple-Point Detectors on Delay and Accidents*." Transportation Research Record 881. National Research Council, Washington, D.C., 1982, pp. 1-9.

3 *Manual on Uniform Traffic Control Devices*. U.S. Department of Transportation, Washington, D.C., December 2003.

4 *National Transportation Communications for ITS Protocol: Object Definitions for Actuated Traffic Signal Controller (ASC) Units – 1202 v01.07*. National Electrical Manufacturers Association, Rosslyn, Virginia, January 2005

5. National Committee on Uniform Traffic Laws and Ordinances, http://www.ncutlo.org/, visited February 26, 2007

6 Hubbard, S.M.L., R. Awwad, and D.M. Bullock, "*A New Perspective on Assessing Impact of Turning Vehicles on Pedestrian Level of Service at Signalized Intersections*," Transportation Research Record, TRB, National Research Council, Washington, DC, TRB Paper ID# 07-0473, in press.

7 Rodegerdts et al. 2004

8 Kell, J. H., and Fullerton, I. J., *Manual of Traffic Signal Design.* 2nd Edition. Institute of Transportation Engineers, Washington, D.C., 1998.

9 Orcutt, F.L. *The Traffic Signal Book*. Prentice Hall, Englewood Cliffs, New Jersey, 1993.

10 *Traffic Engineering Manual.* Section 3 - Signals. Florida Department of Transportation, Tallahassee, Florida, March 1999

11 NCHRP Report 493

12 De Camp, G., and R.W. Denny. "Improved Protected-Permitted Left-Turn Signal Displays -- The Texas Approach." *ITE Journal*. Institute of Transportation Engineers, Washington, D.C., October, 1992, pp. 21-24

13 Lalani, N. *Alternative Treatments for At-Grade Pedestrian Crossings.* Institute of Transportation Engineers, Washington, D.C., 2001

14 Lalani, N., 2004

15 Zegeer, C.V., C. Seiderman, P. Lagerwey, M. Cynecki, M. Ronkin, and R. Schenider. "*Pedestrian Facilities Users Guide – Providing Safety and Mobility*." Federal Highway Administration, Washington, D.C., March 2002.

16 ITE Technical Committee 18 (P.S. Parsonson, Chair). "Small-Area Detection at Intersection Approaches." *Traffic Engineering.* Institute of Transportation Engineers, Washington, D.C., February 1974, pp. 8-17

17 Zegeer, C.V., and R.C. Deen. "Green-Extension Systems at High-Speed Intersections." *ITE Journal*. Institute of Transportation Engineers, Washington, D.C., November, 1978, pp. 19-24

18 Chang, M.S., C.J. Messer, and A.J. Santiago. "Timing Traffic Signal Change Intervals Based on Driver Behavior." *Transportation Research Record No. 1027.* Transportation Research Board, National Research Council, Washington, D.C., 1985, pp. 20-30.

19 *A Policy on Geometric Design of Highways and Streets.* American Association of State and Highway Transportation Officials, Washington, D.C., 2004

20 Lin, F.B. “Optimal Timing Settings and Detector Lengths of Presence Mode Full-Actuated Control.” *Transportation Research Record No. 1010.* Transportation Research Board, National Research Council, Washington, D.C., 1985, pp. 37-45

# CHAPTER 5

# BASIC SIGNAL TIMING CONTROLLER PARAMETERS

## TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# 5.0 BASIC SIGNAL TIMING PROCEDURE AND CONTROLLER PARAMETERS

This chapter documents the principles of basic traffic signal timing at an intersection. Signal timing is a collection of parameters and logic designed to allocate the right-of-way at a signalized intersection. A major focus of this chapter is to describe basic signal timing parameters necessary to operate an intersection and guidelines for selecting values for those parameters. The principles described in this chapter are generally applicable to all signalized intersections. To maximize the usefulness and transferability of the information provided, the chapter uses the terminology defined in current traffic signal control standards, such as National Transportation Communications for ITS Protocol (NTCIP) Document 1202 (1) and National Electrical Manufacturers Association (NEMA) Standards Publication TS 2-2003 (2), with alternative definitions in some cases.

This chapter contains the following sections:

5.1 – Terminology and Key Definitions

5.2 – Modes of Traffic Signal Operation and Their Use

5.3 – Phase Intervals and Basic Parameters

5.4 – Actuated Timing Parameters

5.5 – Volume-Density Features

5.6 – Detection Configuration and Parameters

5.7 – Guidelines for Time-Base Controls

## 5.1 TERMINOLOGY AND KEY DEFINITIONS

This section identifies and describes basic terminology used within this chapter. Additional terms can be found in the Glossary section of the Manual.

*Actuated Signal Control*
A type of signal control where time for each phase is at least partially controlled by detector actuations.

*Call*
An indication within a controller that a vehicle or pedestrian is awaiting service from a particular phase or that a recall has been placed on the phase.

*Extend*
A detector parameter that increases the duration of a detector actuation by a defined fixed amount.

*Gap Out*
A type of actuated operation for a given phase where the phase terminates due to a lack of vehicle calls within a specific period of time (passage time).

*Interval*
The duration of time during which the indications do not change their state (active or off). Typically, one or more timing parameters control the duration of an interval. The pedestrian clearance interval is determined by the pedestrian clearance time. The green interval duration is controlled by a number of parameters including minimum time, maximum time, gap time, etc.

*Isolated intersection*
An intersection located outside the influence of and not coordinated with other signalized intersections, commonly one mile or more from other signalized intersections.

*Minimum Gap*

A volume density parameter that specifies the minimum green extension when gap reduction is used.

*Minimum Green*

A parameter that defines the shortest allowable duration of the green interval.

*Minimum Recall*

A parameter which results in a phase being called and timed for at least its minimum green time whether or not a vehicle is present.

*Movement –*

Movements reflect the user perspective. Movements can also be broken down into classes (car, pedestrians, buses, LRT, etc.). Typical movements are left, through and right. Movement is an activity in response to a "go" (green ball, green arrow, walk, white vertical transit bar) indication.

*Max Out*

A type of actuated operation for a given phase where the phase terminates due to reaching the designated maximum green time for the phase.

*Passage Time (Vehicle Interval, Gap, Passage Gap, Unit Extension)*

A parameter that specifies the maximum allowable duration of time between vehicle calls on a phase before the phase is terminated.

*Pedestrian Clearance Interval*

Also generally known as "Flashing Don't Walk" (FDW). An indication warning pedestrians that the walk indication has ended and the don't walk indication will begin at the end of the pedestrian clearance interval. Some agencies consider the pedestrian clearance interval to consist of both the FDW time and the yellow change interval.

*Phase*

A timing unit associated with the control of one or more indications. A phase may be timed considering complex criteria for determination of sequence and the duration of intervals.

*Pre-timed control*

A signal control in which the cycle length, phase plan, and phase times are predetermined and fixed.

*Queue*

A line of vehicles, bicycles, or persons waiting to be served by a phase in which the flow rate from the front of the queue determines the average speed within the queue. Slowly moving vehicles or people joining the rear of the queue are usually considered part of the queue. The internal queue dynamics can involve starts and stops. A faster-moving line of vehicles is often referred to as a moving queue or a platoon.

*Recall*

A call is placed for a specified phase each time the controller is servicing a conflicting phase. This will ensure that the specified phase will be serviced again. Types of recall include soft, minimum, maximum, and pedestrian.

*Semi-Actuated Control*

A type of signal control where detection is provided for the minor movements only.

*Volume-Density*

A phase timing technique that uses a series of parameters (variable initial, minimum gap, time before reduction, time to reduce) to provide alternative, variable settings for the otherwise fixed parameters of minimum green and passage time.

## 5.2 MODES OF TRAFFIC SIGNAL OPERATION AND THEIR USE

Traffic signals operate in either pre-timed or actuated mode or some combination of the two. Pre-timed control consists of a series of intervals that are fixed in duration. Collectively, the preset green, yellow, and red intervals result in a deterministic sequence and fixed cycle length for the intersection.

In contrast to pre-timed control, actuated control consists of intervals that are called and extended in response to vehicle detectors. Detection is used to provide information about traffic demand to the controller. The duration of each phase is determined by detector input and corresponding controller parameters. Actuated control can be characterized as fully-actuated or semi-actuated, depending on the number of traffic movements that are detected.

Table 5-1 summarizes the general attributes of each mode of operation to aid in the determination of the most appropriate type of traffic signal control for an intersection. The attributes of the various modes of operation are discussed in additional detail in the following subsections.

**Table 5-1. Relationship between intersection operation and control type.**

| | Pre-timed | | Actuated | | |
|---|---|---|---|---|---|
| **Type of Operation** | **Isolated** | **Coordinated** | **Semi-Actuated** | **Fully Actuated** | **Coordinated** |
| **Fixed Cycle Length** | Yes | Yes | No | No | Yes |
| **Conditions Where Applicable** | Where detection is not available | Where traffic is consistent, closely spaced intersections, and where cross street is consistent | Where defaulting to one movement is desirable, major road is posted <40 mph and cross road carries light traffic demand | Where detection is provided on all approaches, isolated locations where posted speed is >40 mph | Arterial where traffic is heavy and adjacent intersections are nearby |
| **Example Application** | Work zones | Central business districts, interchanges | Highway operations | Locations without nearby signals; rural, high speed locations; intersection of two arterials | Suburban arterial |
| **Key Benefit** | Temporary application keeps signals operational | Predictable operations, lowest cost of equipment and maintenance | Lower cost for highway maintenance | Responsive to changing traffic patterns, efficient allocation of green time, reduced delay and improved safety | Lower arterial delay, potential reduction in delay for the system, depending on the settings |

### 5.2.1 Pre-timed Control

Pre-timed control is ideally suited to closely spaced intersections where traffic volumes and patterns are consistent on a daily or day-of-week basis. Such conditions are often found in downtown areas. They are also better suited to intersections where three or fewer phases are needed (3).

Pre-timed control has several advantages. For example, it can be used to provide efficient coordination with adjacent pre-timed signals, since both the start and end of green are predictable. Also, it does not require detectors, thus making its operation immune to problems associated with detector failure. Finally, it requires a minimum amount of training to set up and maintain. On the other hand, pre-timed control cannot compensate for unplanned fluctuations in traffic flows, and it tends to be inefficient at isolated intersections were traffic arrivals are random.

**Modern traffic signal controllers do not explicitly support signal timing for pre-timed operation, because they are designed for actuated operation. Nevertheless, pre-timed operations can be achieved by specifying a maximum green setting that is equal to the desired pre-timed green interval and invoking the maximum vehicle recall parameter described below.**

### 5.2.2 Semi-Actuated Control

Semi-actuated control uses detection only for the minor movements at an intersection. The phases associated with the major-road through movements are operated as "non-actuated." That is, these phases are not provided detection information. In this type of operation, the controller is programmed to dwell in the non-actuated phase and, thereby, sustain a green indication for the highest flow movements (normally the major street through movement). Minor movement phases are serviced after a call for their service is received.

Semi-actuated control is most suitable for application at intersections that are part of a coordinated arterial street system. Coordinated-actuated operation is discussed in more detail in Chapter 6. Semi-actuated control may also be suitable for isolated intersections with a low-speed major road and lighter crossroad volume.

Semi-actuated control has several advantages. Its primary advantage is that it can be used effectively in a coordinated signal system. Also, relative to pre-timed control, it reduces the delay incurred by the major-road through movements (i.e., the movements associated with the non-actuated phases) during periods of light traffic. Finally, it does not require detectors for the major-road through movement phases and hence, its operation is not compromised by the failure of these detectors.

The major disadvantage of semi-actuated operation is that continuous demand on the phases associated with one or more minor movements can cause excessive delay to the major road through movements if the maximum green and passage time parameters are not appropriately set. Another drawback is that detectors must be used on the minor approaches, thus requiring installation and ongoing maintenance. Semi-actuated operation also requires more training than that needed for pre-timed control.

### 5.2.3 Fully-Actuated Control

Fully-actuated control refers to intersections for which all phases are actuated and hence, it requires detection for all traffic movements. Fully-actuated control is ideally suited to isolated intersections where the traffic demands and patterns vary widely during the course of the day. Most modern controllers in coordinated signal systems can be programmed to operate in a fully-actuated mode during low-volume periods where the system is operating in a "free" (or non-coordinated) mode. Fully-actuated control can also improve performance at intersections with lower volumes that are located at the boundary of a coordinated system and do not impact progression of the system (4). Fully-actuated control has also been used at the intersection of two arterials to optimize green time allocation in a critical intersection control method.

There are several advantages of fully-actuated control. First, it reduces delay relative to pre-timed control by being highly responsive to traffic demand and to changes in traffic pattern. In addition, detection information allows the cycle time to be efficiently allocated on a cycle-by-cycle basis. Finally, it allows phases to be skipped if there is no call for service, thereby allowing the controller to reallocate the unused time to a subsequent phase.

The major disadvantage of fully-actuated control is that its cost (initial and maintenance) is higher than that of other control types due to the amount of detection required. It may also result in higher percentage of vehicles stopping because green time is not held for upstream platoons.

## 5.3 PHASE INTERVALS AND BASIC PARAMETERS

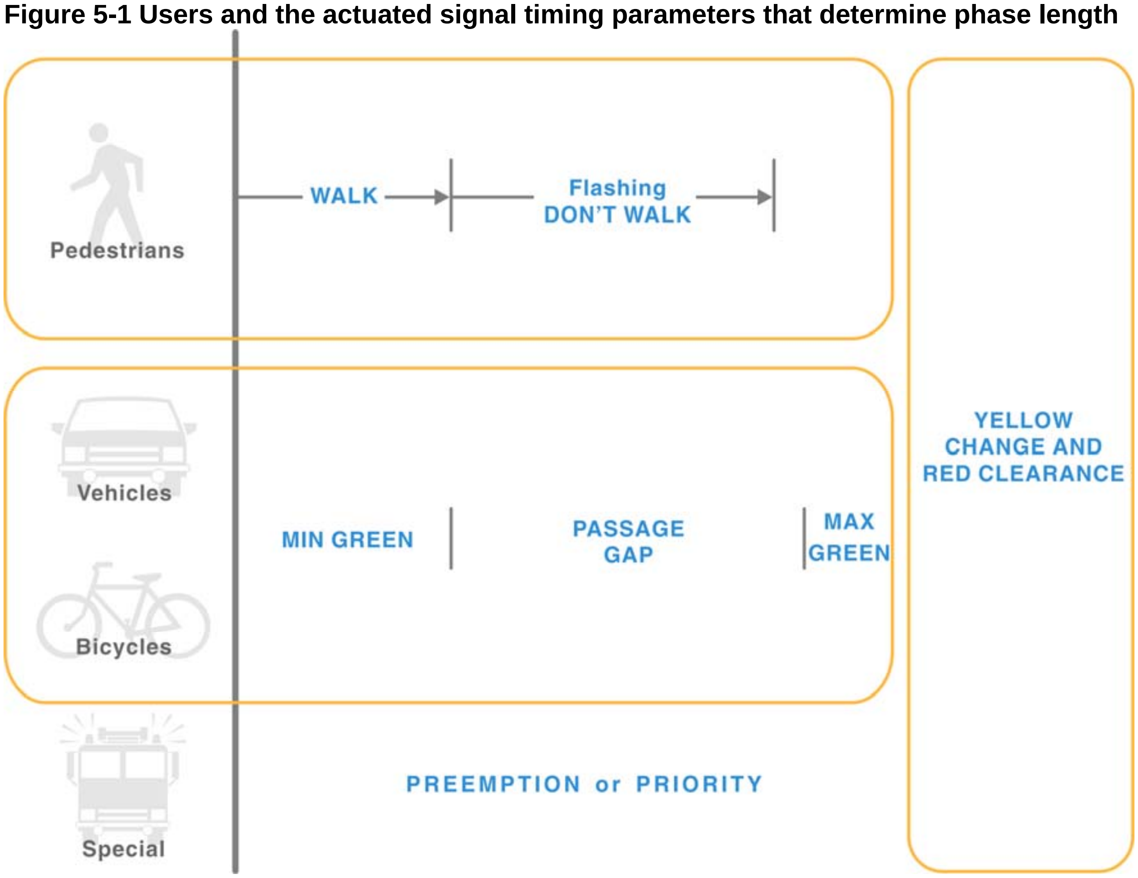
An interval is defined in the NTCIP 1202 standard as “a period of time during which signal indications do not change.” Various parameters control the length of an interval depending on the interval type. For example, a pedestrian walk interval (the time period during which the Walking Person signal indication is displayed) is generally controlled by the single user-defined setting for the walk parameter. The vehicular green interval, on the other hand, is generally controlled by multiple parameters, including minimum green, maximum green, and passage time.

This section describes guidelines for setting basic parameters that determine the duration of each interval associated with a signal phase. These intervals include:

- Vehicular Green Interval
- Vehicle Change and Clearance Intervals
- Pedestrian Intervals

Parameters related to these intervals and discussed in this section include minimum green, maximum green, yellow-change, red clearance, pedestrian walk, and pedestrian flashing don’t walk (FDW). Figure 5-1 depicts the relationship between these parameters and the user group associated with each interval that may time during a phase. These intervals time concurrently during a phase. Although shown here, signal preemption and priority are addressed in Chapter 9. Additional timing parameters related to actuated control (e.g., passage time) may also influence the duration of an interval and are discussed in Section 5.4.

**Figure 5-1 Users and the actuated signal timing parameters that determine phase length**



### 5.3.1 Vehicular Green Interval

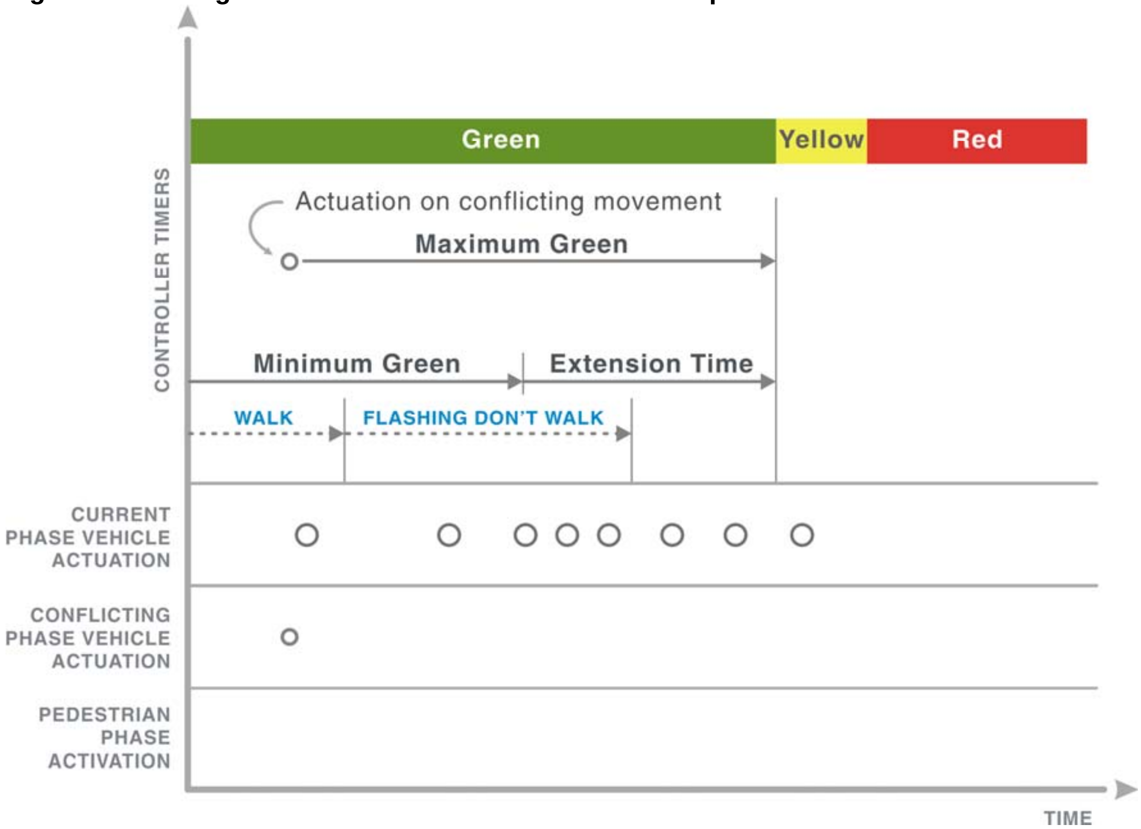
The vehicular green interval is the time dedicated to serving vehicular traffic with a green indication. This interval is defined primarily by the minimum and maximum green parameters in the case of an isolated intersection. At an actuated controller, other parameters (e.g., passage time) also determine the length of this interval. Those parameters are discussed in Section 5.4. It is also possible that the duration of the vehicle green interval may be defined by the length of the associated pedestrian intervals.

#### *Minimum Green*

The minimum green parameter represents the least amount of time that a green signal indication will be displayed for a movement. Minimum green is used to allow drivers to react to the start of the green interval and meet driver expectancy. Its duration may also be based on considerations of queue length or pedestrian timing in the absence of pedestrian call buttons and/or indications. A minimum green that is too long may result in wasted time at the intersection; one that is too short may violate driver expectation or (in some cases) pedestrian safety. The minimum green interval is shown in Figure 5-2, as it relates to other intervals and signal control parameters. Calls placed on the active phase during the minimum green have no bearing on the duration of the green interval as the interval will time at least as long as the minimum green timer.

Lin (5) conducted extensive simulation analysis of fully-actuated controlled intersections to determine the effect of minimum green intervals on delay. Through these simulations, he found that delay was minimal when the minimum green interval was less than 4 seconds. Delay for the intersection under the scenarios studied tended to increase slightly as the minimum green interval increased from 4 to 8 seconds.

**Figure 5-2 Settings that define the duration of a vehicle phase**




The intent of the minimum green interval is to ensure that each green interval is displayed for a length of time that will satisfy driver expectancy. When stop-line detection is not provided, variable initial, as described in Section 5.4, should be used to allow vehicles queued between the stop line and the nearest detector at the start of green to clear the intersection. In cases where separate pedestrian signal displays are not provided, the minimum green interval will also need to be long enough to accommodate pedestrians who desire to cross in a direction parallel to the traffic movement receiving the green indication. These considerations and the conditions in which each applies are shown in Table 5-2.

**Table 5-2 Factors considered when setting the minimum green interval**

| Phase | Stop Line Detection? | Pedestrian Button? | Considered in Establishing Minimum Green? | | |
|---|---|---|---|---|---|
| | | | Driver Expectancy | Pedestrian Crossing Time | Queue Clearance |
| Through | Yes | Yes | Yes | No | No |
| | | No | Yes | Yes | No |
| | No | Yes | Yes | No | Yes, if actuated |
| | | No | Yes | Yes | Yes, if actuated |
| Left-Turn | Yes | Not applicable | Yes | Not applicable | No |

To illustrate the use of Table 5-2, consider a through movement with stop-line detection and a pedestrian push button. As indicated in the table, the minimum green interval should be based solely on driver expectancy. However, if a pedestrian call button is not provided (and pedestrians are expected to cross the road at this intersection), the minimum green interval should be based on driver expectancy and pedestrian crossing time.

#### *Minimum Green to Satisfy Driver Expectancy*

The duration of minimum green needed to satisfy driver expectancy varies among practitioners. Some practitioners rationalize the need for 15 seconds or more of minimum green at some intersections; other practitioners use as little as 2 seconds minimum green. If a minimum green parameter is set too low and violates driver expectancy, there is a risk of increased rear-end crashes. The values listed in Table 5-3 are typical for the specified combination of phase and facility type.

**Table 5-3 Typical minimum green interval duration needed to satisfy driver expectancy**

| Phase Type | Facility Type | Minimum Green Needed to Satisfy Driver Expectancy ($G_e$), s |
|---|---|---|
| Through | Major Arterial (speed limit exceeds 40 mph) | 10 to 15 |
| | Major Arterial (speed limit is 40 mph or less) | 7 to 15 |
| | Minor Arterial | 4 to 10 |
| | Collector, Local, Driveway | 2 to 10 |
| Left Turn | Any | 2 to 5 |

#### *Minimum Green for Pedestrian Crossing Time*

The minimum green duration must satisfy pedestrian crossing needs for through phases that are not associated with a pedestrian push button *but have a pedestrian demand*. Under these conditions, the minimum green needed to satisfy pedestrian considerations can be computed using Equation 5-1. Methodology for computing walk and pedestrian clearance interval durations are provided in Section 5.3.3.

$$G_P = PW + PC \tag{5-1}$$

Where: $G_p$ is the minimum green interval duration needed to satisfy pedestrian crossing time,

$PW$ is the walk interval duration, and

$PC$ is the pedestrian clearance interval duration, s (all values in seconds).

#### *Minimum Green for Queue Clearance*

The duration of minimum green can also be influenced by detector location and controller operation. This subsection addresses the situation where a phase has one or more advance detectors and no stop-line detection. If this detection design is present, and the added initial parameter (as discussed later) is not used, then a minimum green interval is needed to clear the vehicles queued between the stop line and the advance detector. The duration of this interval is specified in Table 5-4.

**Table 5-4 Typical minimum green interval duration needed to satisfy queue clearance**

| Distance Between Stop Line and Nearest Upstream Detector, ft | Minimum Green Needed to Satisfy Queue Clearance [1, 2] ($G_q$), s |
|---|---|
| 0 to 25 | 5 |
| 26 to 50 | 7 |
| 51 to 75 | 9 |
| 76 to 100 | 11 |
| 101 to 125 | 13 |
| 126 to 150 | 15 |

Notes:

1 - Minimum green values listed apply only to phases that have one or more advance detectors, no stop line detection, and the added initial parameter is not used.

2 - Minimum green needed to satisfy queue clearance, $Gq = 3 + 2n$ (in seconds), where $n$ = number of vehicles between stop line and nearest upstream detector in one lane. And, $n = Dd / 25$, where $Dd$ = distance between the stop line and the downstream edge of the nearest upstream detector (in feet) and 25 is the average vehicle length (in feet), which could vary by area.

If a phase has one or more advance detectors, no stop-line detection, and the added initial parameter *is* used, then the minimum initial interval should equal the minimum green needed to satisfy driver expectancy. Timing of minimum greens using the added initial parameter is discussed in Section 5.4.

### *Maximum Green*

The maximum green parameter represents the maximum amount of time that a green signal indication can be displayed in the presence of conflicting demand. Maximum green is used to limit the delay to any other movement at the intersection and to keep the cycle length to a maximum amount. It also guards against long green times due to continuous demand or broken detectors. Ideally, the maximum green will not be reached because the detection system will find a gap to end the phase, but if there are continuous calls for service and a call on one or more conflicting phases, the maximum green parameter will eventually terminate the phase. A maximum green that is too long may result in wasted time at the intersection. If its value is too short, then the phase capacity may be inadequate for the traffic demand, and some vehicles will remain unserved at the end of the green interval.

Most modern controllers provide two or more maximum green parameters that can be invoked by a time-of-day plan or external input (i.e., Maximum Green 2). As shown in Figure 5-2, the maximum green extension timer begins timing upon the presence of a conflicting call. If there is demand on the phase that is currently timing and no conflicting calls, the maximum green timer will be reset until an opposing call occurs.

It should be noted that the normal failure mode of a detector is to place a continuous call for service. In this case, a failed detector on a phase will cause that phase's maximum green to time every cycle.

Many modern controllers also provide a feature that allows the maximum green time to be increased to a defined threshold after maxing out a phase a certain number of consecutive times (or alternatively to select among two or three maximum green values). The maximum green time may then be automatically decreased back to the original value after the phase has gapped out a certain number of times. The exact methods and user settable parameters for this feature vary by manufacturer.

The maximum green value should exceed the green duration needed to serve the average queue and, thereby, allow the phase to accommodate cycle-to-cycle peaks in demand. Frequent phase termination by gap out (as opposed to max out) during low-to-moderate volumes and by occasional max out during peak periods is commonly used as an indication of a properly timed maximum green duration.  Example values are listed in Table 5-5.

**Table 5-5 Example values for maximum green duration**

| Phase | Facility Type | Maximum Green, s |
|---|---|---|
| Through | Major Arterial (speed limit exceeds 40 mph) | 50 to 70 [1] |
| | Major Arterial (speed limit is 40 mph or less) | 40 to 60 |
| | Minor Arterial | 30 to 50 |
| | Collector, Local, Driveway | 20 to 40 |
| Left Turn | Any | 15 to 30 |

Note:
[1]Range is based on the assumption that advance detection is provided for indecision zone protection. If this type of detection is not provided, then the typical maximum green range is 40 to 60 s.

Two methods are commonly used to establish the maximum green setting. Both estimate the green duration needed for average volume conditions and inflate this value to accommodate cycle-to-cycle peaks. Both of these methods assume that advance detection for indecision zone protection is not provided. If advance detection is provided for indecision zone protection, the maximum green setting obtained from either method may need to be increased slightly to allow the controller to find a “safe” time to terminate the phase by gap out.

One method used by some agencies is to establish the maximum green setting based on an 85$^{th}$ to 95$^{th}$ percentile probability of queue clearance (6). The procedure requires knowledge of the cycle length, or an estimate of its average value for actuated operation. If the cycle length is known, then the maximum green setting for a signal phase can be obtained from Table 5-6.

**Table 5-6 Maximum green duration as a function of cycle length and volume**

| Phase Volume per Lane, veh/hr/ln | Cycle Length, s | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 |
| | Maximum Green ($G_{max}$)[1], s | | | | | | | |
| 100 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| 200 | 15 | 15 | 15 | 15 | 16 | 18 | 19 | 21 |
| 300 | 15 | 16 | 19 | 21 | 24 | 26 | 29 | 31 |
| 400 | 18 | 21 | 24 | 28 | 31 | 34 | 38 | 41 |
| 500 | 22 | 26 | 30 | 34 | 39 | 43 | 47 | 51 |
| 600 | 26 | 31 | 36 | 41 | 46 | 51 | 56 | 61 |
| 700 | 30 | 36 | 42 | 48 | 54 | 59 | 65 | 71 |
| 800 | 34 | 41 | 48 | 54 | 61 | 68 | 74 | 81 |

Note:
[1]Values listed are computed as: $G_{max} = (V\,C)/(1200\,n) + 1$ where, $V$ = design hourly volume served by subject phase (in vehicles per hour); $n$ = number of lanes served by subject phase, and $C$ = cycle length (in seconds). A 15-s minimum duration is imposed on the computed values.

The values listed are based on the equation shown in the table footnote. Due to the approximate nature of this equation, the actual percentage probability of queue clearance varies between the 85$^{th}$ and 95$^{th}$ percentiles for the values listed.

A second method for establishing the maximum green setting is based on the equivalent optimal pre-timed timing plan (7). This method requires the development of a pre-timed signal timing plan based on delay minimization. The minimum-delay green interval durations are multiplied by a factor ranging from 1.25 to 1.50 to obtain an estimate of the maximum green setting (8).

The maximum green time used for a particular phase is calculated differently for low and high levels of saturation. During periods of low volume, when the green phase times rarely reach their maximum values, the maximum green time can be set fairly high (up to 1.7 times the calculated average time for the phase). This accommodates most fluctuations in vehicle arrival rates. During conditions at or near saturation, it is important to set the maximum green times as if they were fixed time, equitably allocating the green based on the critical lane volumes as described in Chapter 3.

To this end, application of the maximum green times show significant disparity in the techniques reported for determination of maximum phase time, which ultimately may result in a wide variation of cycle lengths at intersections. In many cases, maximum green times are set at one value throughout the day and don't reflect the needs of the intersection during various times of day. In some cases, these maximum green time values result in cycle lengths that are too long for efficient operations.

### 5.3.2 Vehicular Change and Clearance Intervals

The intent of the vehicle phase change and clearance intervals is to provide a safe transition between two conflicting phases. It consists of a yellow change interval and, optionally, a red clearance interval. The intent of the yellow change interval is to warn drivers of the impending change in right-of-way assignment. The red clearance interval is used when there is some benefit to providing additional time before conflicting movements receive a green indication.

***Yellow Change***

The duration of the yellow change interval is typically based upon driver perception-reaction time, plus the distance needed to safely stop or to travel safely through the intersection.

A state's Uniform Vehicle Code directly affects yellow change interval timing, as it determines whether a *permissive* or *restrictive* yellow law is in place.

- Permissive Yellow Law: A driver can enter the intersection during the entire yellow interval and be in the intersection during the red indication as long as the vehicle entered the intersection during the yellow interval. Under permissive yellow law, an all-red clearance interval must exist as a timing parameter to ensure safe right-of-way transfer at an intersection. This rule is consistent with paragraph 11-202 of the Uniform Vehicle Code (9).
- Restrictive Yellow Law: There are two variations of this law (10). In one variation, a vehicle may not enter an intersection when the indication is yellow unless the vehicle can clear the intersection by the end of yellow. This implies that the yellow duration should be sufficiently long as to allow drivers the time needed to *clear* the intersection if they determine that it is not possible to safely stop. In the other variation, a vehicle may not enter an intersection unless it is impossible or unsafe to stop. With restrictive yellow law, the presence of an all-red interval is optional and good engineering judgment should be applied.

Due to the varying interpretations of the yellow change use, it is encouraged that traffic engineers refer to the local and regional statutes for guidance in determining the purpose of the yellow change time.

### ***Red Clearance***

The red clearance interval, referred to in some publications as an all-red interval, is an interval at the end of the yellow change interval during which the phase has a red-signal display before the display of green for the following phase. The purpose of this interval is to allow time for vehicles that entered the intersection during the yellow-change interval to clear the intersection prior to the next phase. Note that the use of the "all-red" nomenclature is generally incorrect, as the red clearance interval only applies to a single phase, not to all phases.

The use of a red clearance interval is optional, and there is no consensus on its application or duration. Recent research has indicated that the use of a red clearance interval showed some benefit to the reduction of red-light-running violations. In these studies, there was a significant reduction in right-angle crashes after implementing a red clearance interval. Other research suggests that this reduction may only be temporary. A comprehensive study of long-term effects for the Minnesota Department of Transportation (11), indicated short-term reductions in crash rates were achieved (approximately one year after the implementation), but long-term reductions were not observed, which implies that there may not be safety benefits associated with increased red clearance intervals.

A disadvantage of using the red clearance interval is that there is a reduction in available green time for other phases. At intersections where the timing for minor movements is restricted (e.g., to split times under coordinated operation (see Chapter 6)), the extra time for a red clearance interval comes from the remaining phases at the intersection. In cases where major movements are already at or near saturation, the reduction in capacity associated with providing red clearance intervals for safety reasons should be accounted for in an operational analysis.

The MUTCD provides guidance on the application and duration of the yellow change and red clearance intervals. It recommends that the interval durations shall be predetermined based on individual intersection conditions, such as approach speed and intersection width. The MUTCD advises that the yellow change interval should last approximately 3 to 6 seconds, with the longer intervals being used on higher-speed approaches. It also advises that the red clearance interval should not exceed 6 seconds. A recent survey conducted by The Urban Transportation Monitor indicated that practitioners who used a standard red clearance interval used a range from 0.5 to 2.0 seconds.

Kell and Fullerton (12) offer the following equation for computing the phase change period (yellow change plus red clearance intervals):

$$CP = \left[ t + \frac{1.47v}{2(a + 32.2g)} \right] + \left[ \frac{W + L_V}{1.47v} \right] \quad (5\text{-}2)$$

where:

$CP$ = change period (yellow change plus red clearance intervals), s;
$t$ = perception-reaction time to the onset of a yellow indication, s;
$v$ = approach speed, mph;
$a$ = deceleration rate in response to the onset of a yellow indication;
$g$ = grade, with uphill positive and downhill negative (percent grade / 100), ft/ft;
$W$ = width of intersection, ft; and
$L_v$ = length of vehicle.

Equation 5-2 is based on driver reaction time, approach speed, approach grade, and intersection width and consists of two terms. The first term (yellow change) represents the time required for a vehicle to travel one safe stopping distance, including driver perception-reaction time. This permits a driver to either stop at the intersection if the distance to the intersection is greater than one safe stopping distance or safely enter the intersection (and clear the intersection under the restrictive yellow law) if the distance to the intersection is less than one safe stopping distance. The second term (red clearance) represents the time needed for a vehicle to traverse the intersection ($[W + L_v]/v$). Although values will vary by driver population and local conditions, the values of $t$ = 1.0 s, $a$ = 10 ft/s$^2$, and $L_v$ = 20 ft are often cited for use in Equation 5-3 (13,14,15). These values of perception-reaction time and deceleration rate are different from those cited in highway geometric design policy documents because they are based on driver response to the yellow indication, which is an expected condition. They are not based on the longer reaction time necessary for an unexpected (or surprise) condition.

When applying Equation 5-2 to through movement phases, the speed used is generally either the 85$^{th}$-percentile speed or the posted regulatory speed limit, depending on agency policy (16). When applying Equation 5-2 to left-turn movement phases, the speed used should reflect that of the drivers that intend to turn. This speed can equal that of the adjacent through movement but it can also be slower as left-turn drivers inherently slow to a comfortable turning speed. Regardless, if the left-turn phase terminates concurrently with the adjacent through phase, it will have the same total change and clearance interval durations as the through phase because the phases are interlocked by the ring-barrier operation.

The width of the intersection is often defined by local policy or state law. For instance, in Arizona intersection width is defined by state law as the distance between prolongations of the curb lines. Where intersection width is not defined by local policies, engineering judgment should be used when measuring the width of the intersection, $W$. One approach is to measure from the near-side stop line to the far edge of the last conflicting traffic lane along the subject movement travel path. If crosswalks are present at the intersection, some agencies have policies to measure from the near-side stop line to the far side of the pedestrian crosswalk on the far side of the intersection (for through-movement phases) or to the far side of the pedestrian crosswalk across the leg of the intersection which the left-turn is entering. This is a jurisdiction-wide issue that must be carefully applied.

**Table 5-7 Duration of change period intervals**

| Approach Speed, mph | "$t + v/2a$" Terms, s (YELLOW) | Width of Intersection, ft | | | | |
|---|---|---|---|---|---|---|
| | | 30 | 50 | 70 | 90 | 110 |
| | | "$(W+L_v)/v$" Term, s (ALL-RED) | | | | |
| 25 | 3.0[a] | 1.4 | 1.9 | 2.5 | 3.0 | 3.5 |
| 30 | 3.2 | 1.1 | 1.6 | 2.0 | 2.5 | 3.0 |
| 35 | 3.6 | 1.0 | 1.4 | 1.8 | 2.1 | 2.5 |
| 40 | 3.9 | 0.9 | 1.2 | 1.5 | 1.9 | 2.2 |
| 45 | 4.3 | 0.8 | 1.1 | 1.4 | 1.7 | 2.0 |
| 50 | 4.7 | 0.7 | 1.0 | 1.2 | 1.5 | 1.8 |
| 55 | 5.0 | 0.6 | 0.9 | 1.1 | 1.4 | 1.6 |
| 60 | 5.4 | 0.6 | 0.8 | 1.0 | 1.2 | 1.5 |

[a] The 2003 MUTCD recommends a minimum duration of 3 seconds for the yellow change interval.

The values for the yellow change interval in Table 5-7 are based on negligible approach grade. They should be increased by 0.1 second for every 1 percent of downgrade. Similarly, they should be decreased by 0.1 second for every 1 percent of upgrade. To illustrate, consider an approach with a 30 mph approach speed, 70-foot intersection width, and 4-percent downgrade. The estimated change period is 5.6 seconds (= 3.2 + (0.1 x 4 + 2.0)).

States that follow the "restrictive yellow" rule may equate the yellow change interval to the value obtained from Equation 5-2 (i.e., the sum of both terms). If a red-clearance interval is needed, its value may be set at 0.5 to 2 seconds, as determined by engineering judgment.

States that follow the "permissive yellow" rule will typically set the yellow change interval equal to the value obtained from the first term of Equation 5-2 (i.e., column 2 of Table 5-7), but not less than 3.0 seconds. This duration will allow drivers that do not have the necessary distance to stop the time needed to reach the intersection before the red indication is presented. If a red clearance interval is needed, its value is typically based on the second term of Equation 5-2 (i.e., columns 3 through 7 of Table 5-7). Some agencies reduce the value of the second term by 1.0 second in recognition of the perception-reaction time of drivers in the next conflicting phase to be served (17).

### 5.3.3 Pedestrian Timing Intervals

The pedestrian phase consists of three intervals: walk; pedestrian clearance, commonly referred to as flashing don't walk (FDW); and solid don't walk. The walk interval typically begins at the start of the green interval and is used to allow pedestrians to react to the change to walk at the start of the phase and move into the crosswalk. This interval corresponds to the WALKING PERSON indication on the pedestrian signal (18). The pedestrian clearance interval follows the walk interval and informs pedestrians the phase is ending. During this interval, the UPRAISED HAND indication flashes on the pedestrian signal. The solid don't walk interval follows the pedestrian clearance interval and is indicated by a solid UPRAISED HAND indication. This interval is an indication to the pedestrian that they should have cleared the crosswalk and opposing vehicle movements could begin. The solid don't walk time is not a programmable parameter in the controller. The duration of the solid don't walk interval is simply the length of the cycle minus the walk and pedestrian clearance intervals.

Although the illustration in Figure 5-2 does not include a pedestrian phase activation, it does show that the pedestrian timers (walk and FDW) would time concurrently with the vehicle intervals if there was a pedestrian activation. In the case of Figure 5-2, the pedestrian intervals are shown as requiring less time than allowed by the maximum green timer. In this case, if there was continuing vehicle demand, the pedestrian indication would show a solid don't walk until the vehicle phase terminated due to lack of demand or the maximum green timer expired. However, if the pedestrian intervals required more time than permitted by the maximum green timer, the vehicle phase would continue to time until the pedestrian flashing don't walk interval finished timing.

***Walk***

The walk interval should provide pedestrians adequate time to perceive the WALK indication and depart the curb before the pedestrian clearance interval begins. It should be long enough to allow a pedestrian that has pushed the pedestrian push button to enter the crosswalk. In many cases, the pedestrian phase will be set to rest in the walk interval to maximize the walk display during a vehicle green. Some controllers have a mechanism to specify that the walk interval begins before, or even after, the onset of the green interval. The walk interval may be extended in some controllers during coordination. A pedestrian recall mode, as discussed in a later section, can be used to eliminate the need for a pedestrian to push buttons and ensures that the pedestrian phase is presented each cycle.

The length of the walk interval is usually established in local agency policy. The MUTCD (19) indicates that the minimum walk duration should be at least 7 seconds, but indicates that a duration as low as 4 seconds may be used if pedestrian volumes are low or pedestrian behavior does not justify the need for 7 seconds. Consideration should be given to walk durations longer than 7 seconds in school zones and areas with large numbers of elderly pedestrians. In cases where the pedestrian push button is a considerable distance from the curb, additional WALK time is desirable. Table 5-8 summarizes the recommended walk interval durations based on the guidance provided in the MUTCD and the Traffic Control Devices Handbook (20). At intersections where older pedestrians are present, the MUTCD recommends that the WALK time allows for a pedestrian to reach the middle of the street at a 3.0 feet per second walking speed.

**Table 5-8 Pedestrian walk interval duration**

| Conditions | Walk Interval Duration (*PW*), s |
|---|---|
| High pedestrian volume areas (e.g., school, central business district, sports venues, etc.) | 10 to 15 |
| Typical pedestrian volume and longer cycle length | 7 to 10 |
| Typical pedestrian volume and shorter cycle length | 7 |
| Negligible pedestrian volume | 4 |
| Conditions where older pedestrians are present | Distance to center of road divided by 3.0 feet per second |

**WALK times that allow pedestrians to cross to a wide median before the FLASHING DON'T WALK may reduce the potential for pedestrians stopping within the median.**

***Pedestrian Clearance***

The pedestrian clearance interval follows the walk interval. When the pedestrian clearance interval begins, pedestrians should either complete their crossing if already in the intersection or refrain from entering the intersection until the next pedestrian walk interval is displayed. The MUTCD currently stipulates that the pedestrian clearance interval must be calculated assuming the distance from the curb to the far side of the opposing travel way, or to a median of sufficient width for pedestrians to wait. Note that previous editions of the MUTCD only required the clearance time to be as long as needed for the pedestrian to reach the *center of the farthest traveled lane*.

Pedestrian clearance time is computed as the crossing distance divided by the walking speed. The speed of pedestrians is a critical assumption in determining this parameter. The MUTCD recommends a walking speed value of 4.0 feet per second (ft/s). The Americans with Disabilities Act (ADA) Accessibility Guidelines for Buildings and Facilities recommended use of 3.0 ft/s. Recent work completed by LaPlante and Kaeser has suggested that a speed of 3.5 ft/s be used to calculate the pedestrian clearance (FDW and Yellow) duration for curb to curb clearance and 3.0 ft/s be used for the total pedestrian time (WALK, FDW, and Yellow) duration for top of ramp to far curb clearance. The *Pedestrian Facilities User Guide* (21) recommends a maximum walking speed of 3.5 ft/s. This guide also suggests that a slower walking speed should be used in areas where there is a heavy concentration of elderly persons or children. A survey by Tarnoff and Ordonez (22) suggests a range of 3.0 to 3.5 ft/s is typically used by agencies to compute crossing time for these special-needs pedestrians. Pedestrian clearance time for typical pedestrian crossing distances can be obtained from Table 5-9.

**Table 5-9 Pedestrian clearance time**

| **Pedestrian Crossing Distance, ft** | **Walking Speed, ft/s** | | |
|---|---|---|---|
| | **3.0** | **3.5** | **4.0** |
| | **Pedestrian Clearance Time (*PCT*), s** | | |
| 40 | 13 | 11 | 10 |
| 60 | 20 | 17 | 15 |
| 80 | 27 | 23 | 20 |
| 100 | 33 | 29 | 25 |

Note:
1 - Clearance times computed as $PCT = D_c / v_p$, where $D_c$ = pedestrian crossing distance (in feet) and $v_p$ = pedestrian walking speed (in feet per second).

In general, agencies use one of two methods to determine the setting for the pedestrian clearance parameter. Some agencies require that the pedestrian clearance time conclude with the onset of the yellow change interval. This approach provides additional time (equal to the change period) for pedestrian clearance—time that is sometimes of benefit to pedestrians who walk slower than average. The pedestrian clearance interval duration for this practice is computed using Equation 5-4.

$$PC = PCT \qquad (5\text{-}3)$$

Other agencies allow a portion of the pedestrian clearance time to occur during the change period (i.e., yellow change or yellow change plus red clearance intervals). This practice minimizes the impact of pedestrian service on phase duration and allows it to be more responsive to vehicular demand. This pedestrian clearance interval duration is computed using Equation 5-5.

$$PC = PCT - (Y + R) \qquad (5\text{-}4)$$

where:
$PC$ = pedestrian clearance interval duration, s;
$PCT$ = pedestrian clearance time, s;
$Y$ = yellow change interval, s; and
$R$ = red clearance interval, s (optional).

**The practice of exclusing the change and clearance intervals may place pedestrians at risk if a concurrent permissive left turn movement is receiving a yellow and the vehicles from that movement are expected to clear the intersection during the yellow interval. Some agencies using flashing yellow applications choose to omit the permissive left turn portion of a protected-permissive left-turn movement during a pedestrian call.**

The pedestrian clearance time that transpires during the green interval coincides with a flashing "DON'T WALK" indication. At the onset of the yellow interval, a steady "DON'T WALK" indication is presented. It is noted that some agencies display the flashing "DON'T WALK" until the end of the change period. However, the MUTCD (Sections 4E.07 and 4E.10) states that if countdown pedestrian signals are used, the pedestrian clearance interval must finish timing before the onset of the yellow clearance interval.

## 5.4 ACTUATED TIMING PARAMETERS

Research has shown that the best form of isolated operation occurs when fully-actuated controllers are used. Actuated controllers operate most effectively when timed in a manner that permits them to respond rapidly to fluctuations in vehicle demand (*23*). This section describes several of the more commonly used settings and parameters that influence phase function or duration in an actuated controller, including phase recall, passage time, simultaneous gap, and dual entry. In addition, this section discusses the volume-density technique.

### 5.4.1 Phase Recalls

Recall causes the controller to place a call for a specified phase each time the controller is servicing a conflicting phase, regardless of the presence of any detector-actuated calls for the phase. There are four types of recalls: minimum recall (also known as vehicle recall), maximum recall, pedestrian recall, and soft recall. These are specified as phase option parameters in NTCIP Document 1202 (24).

***Minimum Recall (Vehicle Recall)***

The minimum recall parameter causes the controller to place a call for vehicle service on the phase. The phase is timed at least for its minimum green regardless of whether there is demand on the movement. The call is cleared upon start of green for the affected phase and placed upon start of the yellow change interval. This may be used where detection has failed.

Minimum recall is the most frequently used recall mode. It is frequently used for the major-road through-movement phases (commonly designated as phases 2 and 6) at semi-actuated non-coordinated intersections. This use ensures that the controller will always return to the major-road through phases regardless of demand on the major-road through phases, thus providing a green indication as early as possible in the cycle.

***Maximum Recall***

The maximum recall parameter causes the controller to place a continuous call for vehicle service on the phase. It results in the presentation of the green indication for its maximum duration every cycle as defined by the maximum green parameter for the phase. When the maximum recall parameter is selected for a phase, the maximum green timer begins timing at the beginning of the phase's green interval, regardless of the presence of a conflicting call or lack thereof.

There are at least three common applications of maximum recall:

- *Fixed-time operation is desired:* Each phase is set for maximum recall. The maximum green setting used for this application should be equal to the green interval durations associated with an optimal fixed time plan.
- *Vehicle detection is not present or is out of service:* Maximum recall for a phase without detection ensures that the phase serves the associated movement. However, maximum recall can result in inefficient operation during light volume conditions (e.g., during night times and weekends) and should be used only when necessary. In some of these situations, a lower maximum green or MAX 2 (50 to 75% of the typical MAX GREEN value) may be desirable.
- *Gapping out is not desired:* Maximum recall can be used to prevent a phase from gapping out. An example application of this is under coordinated operations where a left turn phase is lagging. By setting the lagging left turn phase to maximum recall, the phase will time for its maximum duration, allowing the adjacent coordinated phase to also time

for its intended maximum duration. This type of operation is typically only used on a time-of-day basis in conjunction with a particular coordinated plan (see Chapter 6).

***Pedestrian Recall***

The pedestrian recall parameter causes the controller to place a continuous call for pedestrian service on the phase, resulting in the controller timing its walk and flashing don't walk operation. Coordination plans may invoke pedestrian calls using a rest in walk command, which dwells in the pedestrian walk interval, while awaiting the yield point.

There are at least two common applications of pedestrian recall:

- *Pedestrian detection is not present or is out of service:* Pedestrian recall for a phase without pedestrian detection ensures that the phase times pedestrian walk and clearance intervals each cycle.
- *High pedestrian demand:* Pedestrian recall is sometimes used to activate the Walk and Pedestrian clearance intervals for phases and time periods that are likely to have high pedestrian demand. This is a common application during periods of high pedestrian activity in downtown environments or at intersections near schools as students are arriving or leaving school for the day.

***Soft Recall***

The soft recall parameter causes the controller to place a call for vehicle service on the phase in the absence of a serviceable conflicting call. When the phase is displaying its green indication, the controller serves the phase only until the minimum green interval times out. The phase can be extended if actuations are received. This may be used during periods of low traffic when there is a desire to default to the major street.

The most typical application for soft recall is for the major-road through movement phases (usually phases 2 and 6) at non-coordinated intersections. The use of soft recall ensures that the major-road through phases will dwell in green when demand for the conflicting phases is absent.

### 5.4.2 Passage Time

Passage time, sometimes called passage gap, vehicle extension, or unit extension, is used to extend the green interval based on the detector status once the phase is green. This parameter extends the Green Interval for each vehicle actuation up to the Maximum Green. It begins timing when the vehicle actuation is removed. This extension period is subject to termination by the Maximum Green timer or a Force Off.

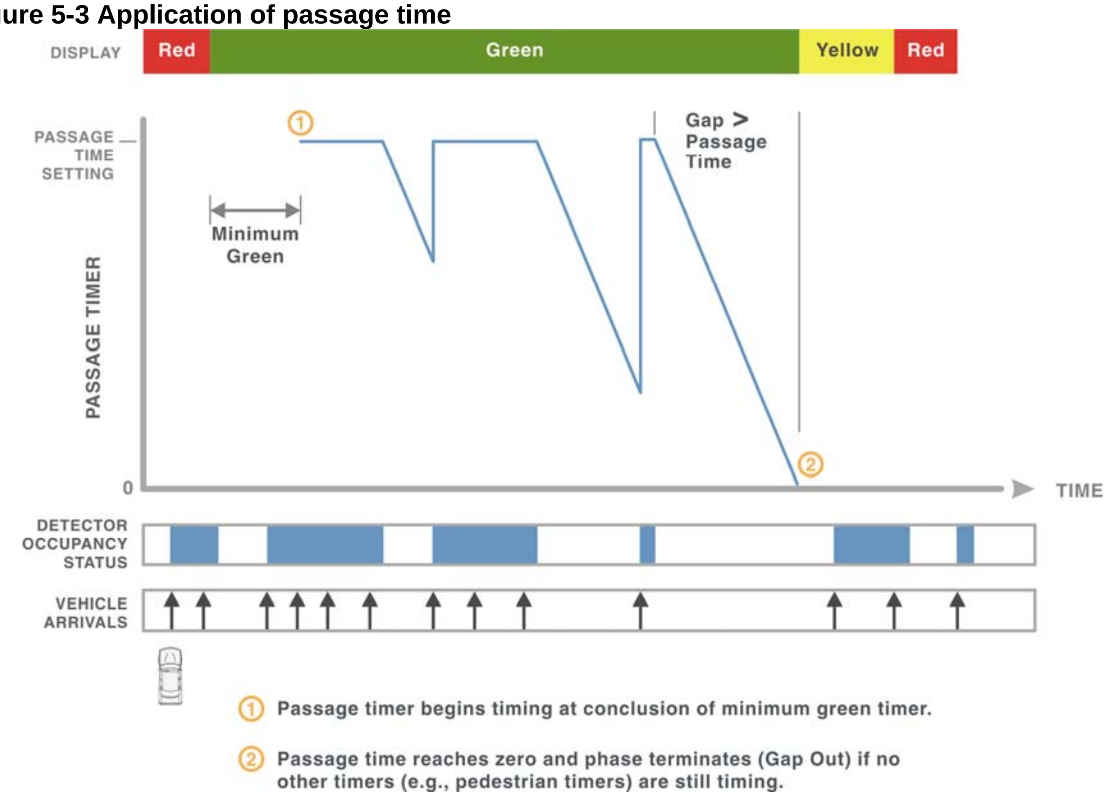
Passage time is used to find a gap in traffic for which to terminate the phase, essentially it is the setting that results in a phase ending prior to its maximum green time during isolated operation. If the passage time is too short, the green may end prematurely, before the vehicular movement has been adequately served. If the passage interval is set too long, there will be delays to other movements caused by unnecessary extension of a phase (*25*,) resulting in delay to the other movements at the intersection. The appropriate passage time used for a particular signal phase depends on many considerations, including: type and number of detection zones per lane, location of each detection zone, detection zone length, detection call memory (i.e., locking or nonlocking), detection mode (i.e., pulse or presence), approach speed, and whether lane-by-lane or approach detection is used. Ideally, the detection design is established and the passage time determined to ensure that the "system" provides efficient queue service and safe phase termination for higher speed approaches. Detection design procedures that reflect these considerations are described in Chapter 4.

The passage timer starts to time from the instant the detector actuation is removed. A subsequent actuation will reset the passage timer. Thus, the mode of the detector, pulse or presence, is extremely important in setting the passage time. The pulse mode essentially measures headways between vehicles and the passage time would be set accordingly. The speed of the vehicles crossing the detectors and the size of the detectors is an important consideration in determining passage time when using presence mode. Longer passage times are often used with shorter detectors, greater distance between the detector and stop line, fewer lanes, and slower speeds.

When the passage timer reaches the passage time limit, and a call is waiting for service on a conflicting phase, the phase will terminate, as shown in Figure 5-3. When this occurs, it is commonly termed as a “gap out”. In the figure, vehicle calls extend the green time until the gap in detector occupancy is greater than the passage time. In this example, presence detection is assumed.

Research by Tarnoff suggests that the vehicle extension interval is one of the most important actuated controller settings, but the variety of techniques for determining proper settings suggest that there is either a lack of knowledge on the availability of this information or disagreement with the conclusions presented (26).

**Figure 5-3 Application of passage time**




The objective when determining the passage time value is to make it large enough to ensure that all vehicles in a moving queue are served but to not make it so large that it extends the green for randomly arriving traffic. This objective is broadened on high-speed approaches to ensure the passage time is not so large that the phase cannot be safely terminated.

Many professionals believe that keeping one lane of traffic (in a left turn or a minor street) moving in deference to a major street with multiple lanes results in inefficient operation. Research has shown that measuring flow rates across lane groups and comparing them with the potential demand at an approach may provide improved decision making within the signal control logic.

The guidelines provided in this section are based on the assumption that non-locking memory is used and that one source of detection is provided (per lane) for the subject signal phase. This source of detection could consist of one long detector loop at the stop line, a series of 6-foot loops that are closely spaced and operate together as one long zone of detection near the stop line, or a single 6-foot loop located at a known distance upstream of the stop line (and no detection at the stop line). As discussed in Chapter 4, passage time is a design parameter for detection designs that include multiple detectors for the purpose of providing safe phase termination (i.e., indecision zone protection). The passage-time value for this application is inherently linked to the detection design and should not be changed from its design value.

Passage time defines the maximum time separation that can occur between vehicle calls without gapping out the phase. When only one traffic lane is served during the phase, this maximum time separation equals the maximum allowable headway (MAH) between vehicles. Although the maximum time separation does not equal the maximum allowable headway when several lanes are being served, the term "MAH" is still used and it is understood that the "headway" represents the time interval between calls (and not necessarily the time between vehicles in the same lane).

Figure 5-4 illustrates the relationship between passage time, gap, and maximum allowable headway for a single-lane approach with one detector. This relationship can be used to derive the following equation for computing passage time for presence mode detection. Gap as shown in this figure is the amount of time that the detection zone is unoccupied.

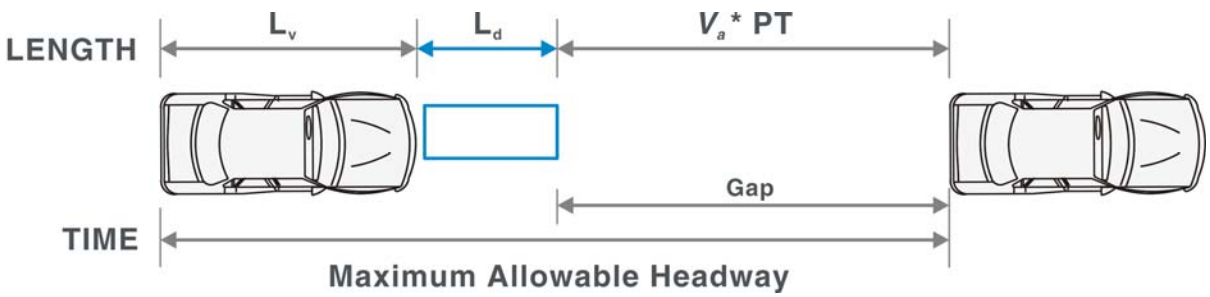
(5-5)

$$PT = MAH - \frac{L_v + L_d}{1.47\, v_a}$$

where,

$PT$ = passage time, s;
$MAH$ = maximum allowable headway, s;
$v_a$ = average approach speed, mph;
$L_v$ = length of vehicle (use 20 ft); and
$L_d$ = length of detection zone, ft.

**Figure 5-4 Relationship between passage time, gap, and maximum allowable headway**




If Equation 5-3 is used with pulse-mode detection, then the length of vehicle $L_v$ and the length of detector $L_d$ equal 0.0 ft, and the passage time is equal to the *MAH*.

The duration of the passage time setting should be based on three goals (27):

1. *Ensure queue clearance.* The passage time should not be so *small* that the resulting MAH causes the phase to have frequent premature gap-outs (i.e., a gap-out that occurs before the queue is fully served). A premature gap-out will leave a portion of the stopped queue unserved and, thereby, lead to increased delays and possible queue spillback. If the queue is extraordinarily long and cannot be accommodated without creating a cycle length that is longer than desirable, this goal may not apply.
2. *Satisfy driver expectancy.* The passage time should not be so *large* that the green is extended unnecessarily after the queue has cleared. Waiting drivers in conflicting phases will become anxious and may come to disrespect the signal indication.
3. *Reduce max-out frequency.* The passage time should not be so *large* that the resulting MAH causes the phase to have frequent max-outs. A long MAH would allow even light traffic volumes to extend the green to max-out. Waiting drivers in higher-volume conflicting phases may be unfairly delayed.

Research by Tarnoff and Parsonson (28) indicates that there is a range of passage times within efficient intersection operations. This range extends from about 1 to 4 seconds for presence mode detection, with lower values being more appropriate under higher volume conditions. Values outside this range tend to increase delay. These passage times correspond to *MAH* values in the range of 2.0 to 4.5 seconds, depending on detection zone length and location.

Based on the previous discussion, the following MAH values are recommended for use with Equation 5-3 to determine passage time:

- Gap reduction not used: *MAH* = 3.0 s
- Gap reduction used: *MAH* = 4.0 s

The recommended MAH values may be increased by 0.1 s if the approach is on a steep upgrade and by 1.0 seconds if there is a large percentage of heavy vehicles.

The passage time computed from the recommended MAH values for a range of speeds and detection zone lengths is provided in Table 5-10 for presence mode detection. It is critical that the relationship of passage time to vehicle speed, detector length, and detector location be considered.

**Table 5-10 Passage time duration for presence mode detection**

| Maximum Allowable Headway, s | Detection Zone Length, ft | 85th Percentile Approach Speed, mph [1] | | | | |
|---|---|---|---|---|---|---|
| | | 25 | 30 | 35 | 40 | 45 |
| | | Passage Time (*PT*), s | | | | |
| 3.0 | 6 | 2.2 | 2.3 | 2.4 | 2.5 | 2.6 |
| | 15 | 1.9 | 2.1 | 2.2 | 2.3 | 2.4 |
| | 25 | 1.6 | 1.8 | 2.0 | 2.1 | 2.2 |
| | 35 | 1.3 | 1.6 | 1.8 | 1.9 | 2.1 |
| | 45 | 1.0 | 1.3 | 1.6 | 1.7 | 1.9 |
| | 55 | 0.7 | 1.1 | 1.3 | 1.6 | 1.7 |
| | 65 | 0.4 | 0.8 | 1.1 | 1.4 | 1.5 |
| | 75 | 0.1 | 0.6 | 0.9 | 1.2 | 1.4 |
| 4.0 | 6 | 3.2 | 3.3 | 3.4 | 3.5 | 3.6 |
| | 15 | 2.9 | 3.1 | 3.2 | 3.3 | 3.4 |
| | 25 | 2.6 | 2.8 | 3.0 | 3.1 | 3.2 |
| | 35 | 2.3 | 2.6 | 2.8 | 2.9 | 3.1 |
| | 45 | 2.0 | 2.3 | 2.6 | 2.7 | 2.9 |
| | 55 | 1.7 | 2.1 | 2.3 | 2.6 | 2.7 |
| | 65 | 1.4 | 1.8 | 2.1 | 2.4 | 2.5 |
| | 75 | 1.1 | 1.6 | 1.9 | 2.2 | 2.4 |

Note:
1 - Average approach speed is computed as 88 percent of the 85th percentile approach speed.

### 5.4.3 Simultaneous Gap

Simultaneous gap defines how a barrier is crossed when a conflicting call is present. If enabled, it requires all phases that are timing concurrently to simultaneously reach a point of being committed to terminate (by gap-out, max-out, or force-off) before they can be allowed to jointly terminate. If disabled, each of the concurrent phases can reach a point of being committed to terminate separately and remain in that state while waiting for all concurrent phases to achieve this status. Simultaneous gap out should be enabled when advance detection is used to provide safe phase termination.

### 5.4.4 Dual Entry

The dual (double) entry parameter is used to call vehicle phases that can time concurrently even if only one of the phases is receiving an active call. For example, if dual entry is active for Phases 2 and 6 and Phase 1 receives a call but no call is placed on Phase 6, Phase 6 would still be displayed along with Phase 1. The most common use of dual entry is to activate the parameter for compatible through movements. If the dual entry parameter is not selected, a vehicle call on a phase will only result in the timing of that phase in the absence of a call on a compatible phase.

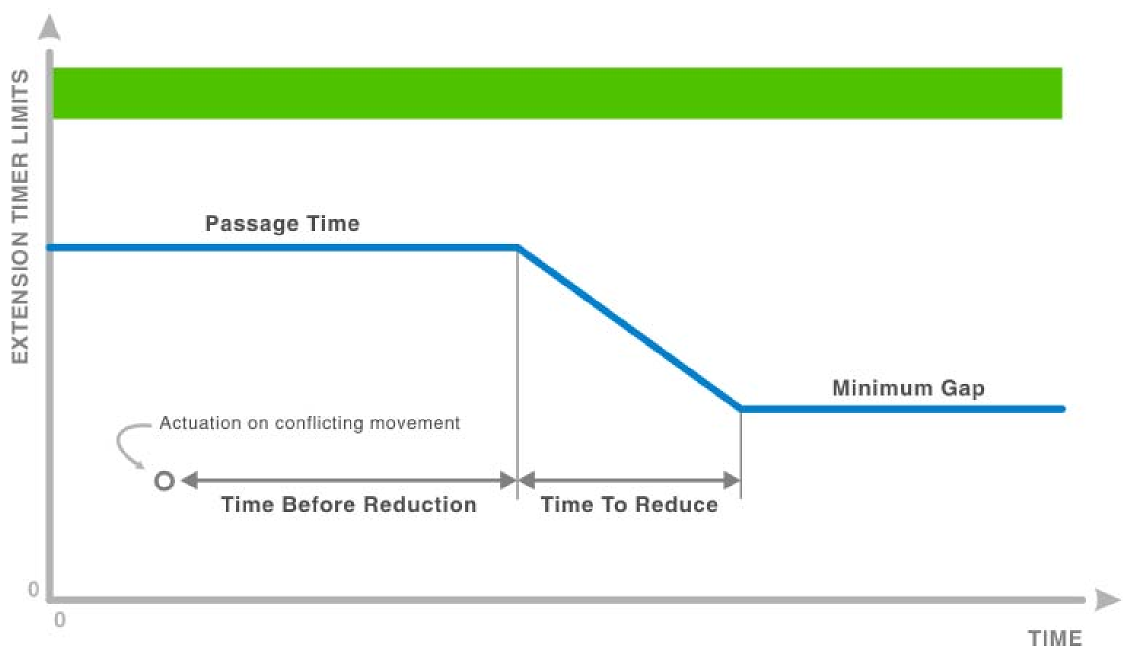
## 5.5 VOLUME-DENSITY FEATURES

Volume-density features can be categorized by two main features: gap reduction and variable initial. These features permit the user to provide variable alternatives to the otherwise fixed parameters of passage time (gap reduction) and minimum green (variable initial). Gap reduction provides a way to reduce the allowable gap over time, essentially becoming more aggressive in looking for an opportunity to end the phase. Variable initial provides an opportunity to utilize cycle by cycle traffic demand to vary the minimum time provided for a phase. These features increase the efficiency of the cycle with the fluctuations in demand, which can result in lower delay for users at the intersection.

### 5.5.1 Gap Reduction

The gap reduction feature reduces the passage time to a smaller value while the phase is green. Initially, the gap sought between actuations is the passage time value. Then, after a specified time (Time Before Reduction), the passage timer is reduced to a minimum gap using a gradual reduction over a specified time (Time To Reduce). This functionality is achieved by programming the following controller parameters: time before reduction, time to reduce, and minimum gap. Their relationship is shown in Figure 5-5.

**Figure 5-5 Use of volume-density to change the extension time**




The time-before-reduction parameter establishes the time that is allowed to elapse after the arrival of a conflicting call and before the extension timer limit is reduced. This period begins when the phase is green and there is a serviceable call on a conflicting phase. Once the time-before-reduction period expires, the extension timer limit is reduced in a linear manner until the time-to-reduce period expires. Thereafter, the extension timer limit is set equal to the minimum-gap parameter. Like the Passage Time, this parameter extends the green interval by up to the Minimum Gap time for each vehicle actuation up to the Maximum Green. It begins timing when the vehicle actuation is removed. This extension period is subject to termination by the Maximum Green or a Force Off.

The gap-reduction feature may be desirable when the phase volume is high and it is difficult to differentiate between the end of the initial queue and of the subsequent arrival of randomly formed platoons. This feature allows the user to specify a higher passage time at the beginning of a phase and then incrementally reduce the passage time as a phase gets longer and the delay to conflicting movements increases.

With gap reduction, a MAH of 2.0 seconds is recommended for use with Equation 5-3 to determine the minimum gap. This MAH may be increased by 0.1 second if the approach is on a steep upgrade, and by 1.0 second if there is a large percentage of heavy vehicles.

If Equation 5-3 is used with pulse-mode detection, then the length of vehicle $L_v$ and the length of detector $L_d$ equal 0.0 ft, and the minimum gap is equal to the MAH.

The minimum gap is computed and shown in Table 5-11 using the recommended MAH values for a range of speeds and detection zone lengths provided for presence mode detection.

**Table 5-11 Minimum gap duration for presence mode detection**

| Maximum Allowable Headway, s | Detection Zone Length, ft | 85th Percentile Approach Speed, mph [1] | | | | |
|---|---|---|---|---|---|---|
| | | 25 | 30 | 35 | 40 | 45 |
| | | Minimum Gap, s | | | | |
| 2.0 | 6 | 1.2 | 1.3 | 1.4 | 1.5 | 1.6 |
| | 15 | 0.9 | 1.1 | 1.2 | 1.3 | 1.4 |
| | 25 | 0.6 | 0.8 | 1.0 | 1.1 | 1.2 |
| | 35 | 0.3 | 0.6 | 0.8 | 0.9 | 1.1 |
| | 45 | 0.0 | 0.3 | 0.6 | 0.7 | 0.9 |
| | 55 | 0.0 | 0.1 | 0.3 | 0.6 | 0.7 |
| | 65 | 0.0 | 0.0 | 0.1 | 0.4 | 0.5 |
| | 75 | 0.0 | 0.0 | 0.0 | 0.2 | 0.4 |

Note:
1 - Average approach speed is computed as 88 percent of the 85th percentile approach speed.

A number of different policies may be employed in determining the value of time-before-reduction. An example policy is to make the time-before-reduction setting equal to the minimum green interval and the time-to-reduce setting equal to half the difference between the maximum and minimum green intervals (29). This guidance is illustrated in Table 5-12.

**Table 5-12 Gap reduction parameter values**

| Minimum Green Interval, s | Time Before Reduction [1], s | Maximum Green, s | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 70 |
| | | Time To Reduce, s | | | | | | | | | | |
| 5 | 10 | 8 | 10 | 13 | 15 | 18 | 20 | 23 | 25 | 28 | 30 | 33 |
| 10 | 10 | 5 | 8 | 10 | 13 | 15 | 18 | 20 | 23 | 25 | 28 | 30 |
| 15 | 15 | N/A | 5 | 8 | 10 | 13 | 15 | 18 | 20 | 23 | 25 | 28 |
| 20 | 20 | N/A | N/A | 5 | 8 | 10 | 13 | 15 | 18 | 20 | 23 | 25 |

Notes:
1 - Time before reduction should always be 10 s or more in length.
N/A - Gap reduction is not applicable to this combination of minimum and maximum green settings.

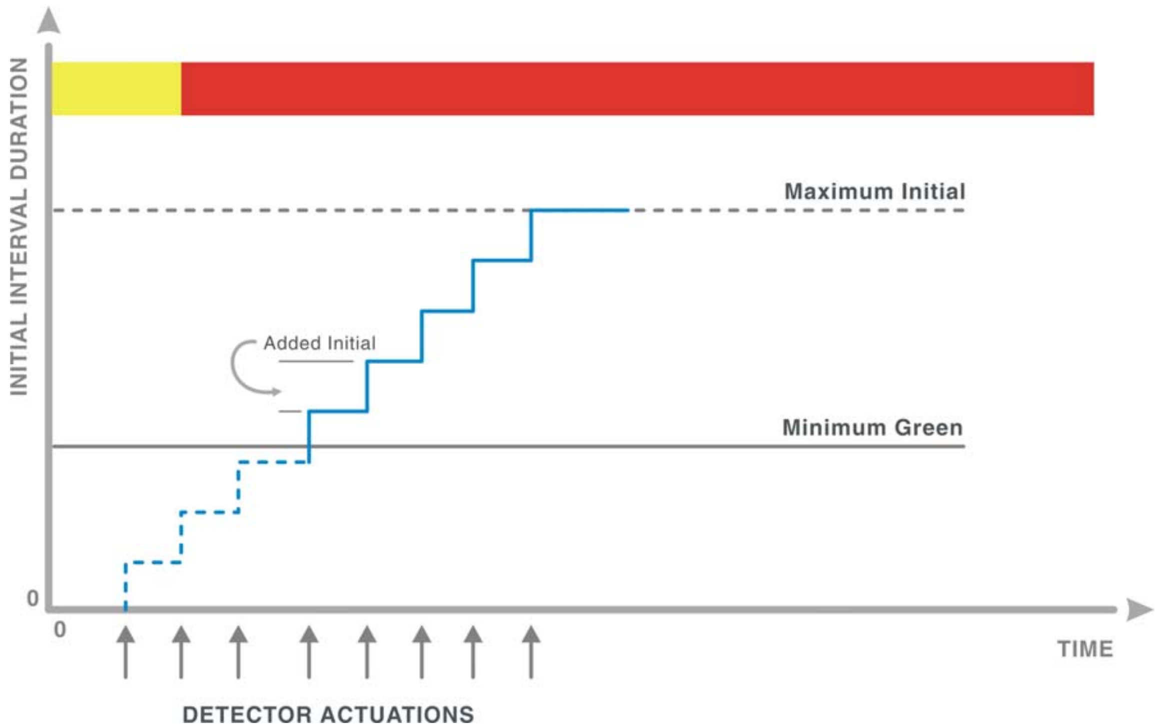
### 5.5.2 Variable Initial

Variable initial is used in some cases to ensure that all vehicles queued between the stop line and the nearest upstream detector are served. Variable initial uses detector activity to determine a minimum green. Vehicles arriving on red that are not able to reach the upstream detector due to a standing queue will be detected and will extend the green by an amount sufficient to allow them to be served using the passage time. This feature is applicable when there are one or more advance detectors, no stop-line detection, and wide fluctuations in traffic volumes between peak and off-peak hours. Variable initial timing is achieved by programming the following controller parameters: minimum green, added initial, and maximum initial. Their relationship is shown in Figure 5-6.

*Added Initial* – This interval times concurrently with the minimum green interval, and is increased by each vehicle actuation received during the associated phase yellow and red intervals. The initial green time portion is the greater of the minimum green or added initial intervals. The Added Initial cannot exceed the Maximum Initial (Figure 5-6).

*Maximum Initial* – This is the maximum period of time for which the Added Initial can extend the initial green period. The Maximum Initial can not be less than the Minimum Green (Figure 5-6).

**Figure 5-6 Use of Added Initial to modify minimum green**




As with other volume-density parameters, there are varying policies that may be employed in determining the values of maximum-initial and added-initial. Generally, the maximum initial setting should be determined using the calculation for minimum green for queue clearance shown in Section 5.3.1.

A common policy for selecting the added-initial parameter is to set this value at approximately 2.0 seconds per actuation if the phase serves only one traffic lane, 1.5 seconds per actuation if it serves two traffic lanes, and 1.2 seconds per actuation if it serves three or more lanes. Slightly larger values can be used if the approach has a significant upgrade or a significant number of heavy trucks. Bicycle traffic may also warrant higher values depending on the intersection width. Some agencies have developed more specific calculations for determining the added initial parameter. For example, the Los Angeles Department of Transportation uses Equation 5-4 to calculate the added initial setting.

$$AI = 2 + \frac{3}{\left\lfloor \frac{MI - 3}{2} \right\rceil} \qquad (5\text{-}6)$$

where,
AI = added initial, s;
MI = maximum initial, s.

## 5.6 DETECTION CONFIGURATION AND PARAMETERS

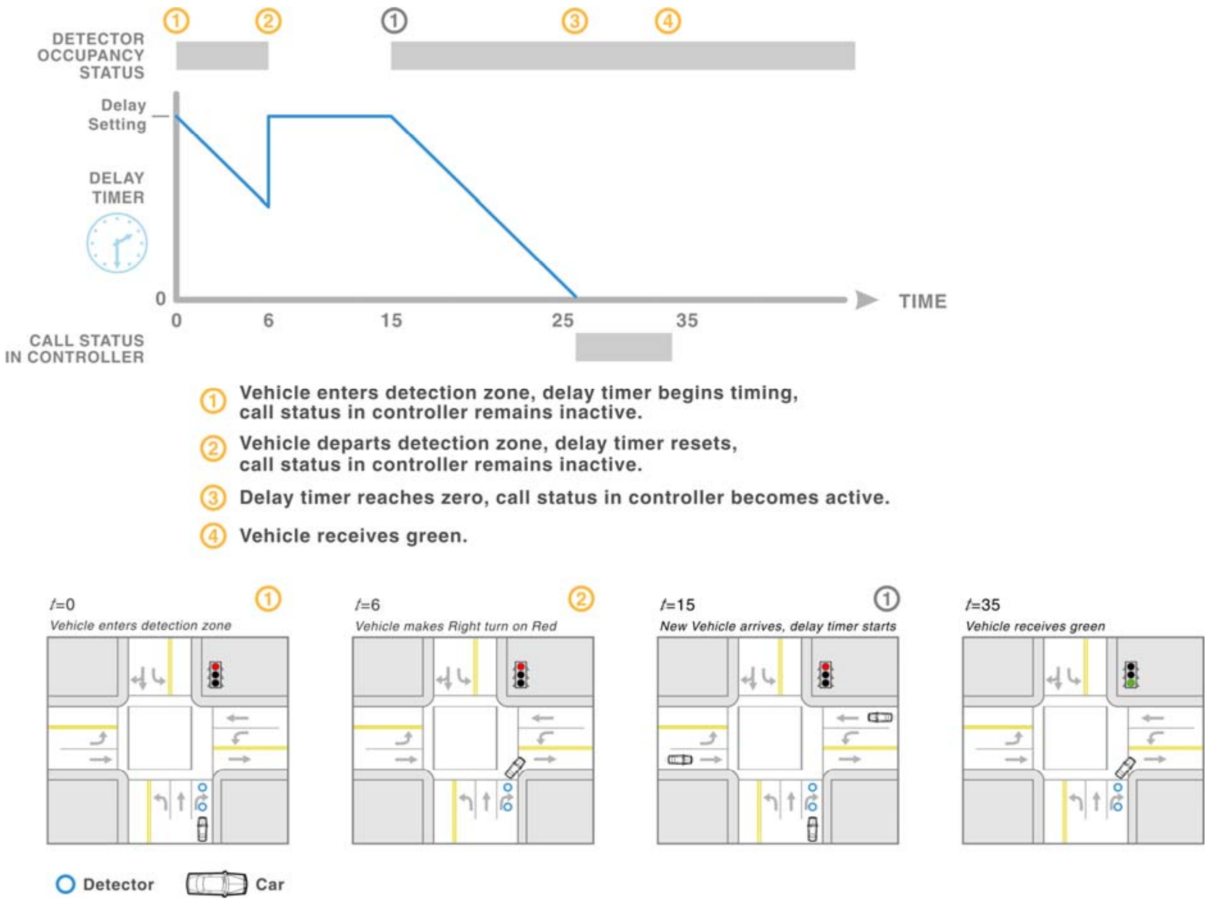
Traffic signal controllers have several settings that can be used to modify the vehicle actuations. Traditionally, this functionality was available only in the detection unit that served as an interface between the vehicle detector and the signal controller. Its implementation in the controller unit has streamlined the signal timing process and can duplicate functionality that may be in the detection unit.

The parameters discussed in this section include: delay, extend, call, and queue. The latter two parameters are actually elements of the detector options in NTCIP Document 1202 (30).

### 5.6.1 Delay

A delay parameter can be used to postpone a vehicle actuation for a detector input on a phase. By using a delay timer, an actuation is not made available until the delay timer expires and the actuation channel input is still active (i.e., the detection zone is still occupied). Once an actuation is made available to the controller, it is continued for as long as the channel input is active. Application of the delay timer is illustrated in Figure 5-7**.**

**Figure 5-7 Application of delay timer**




Common applications of a delay parameter on detection include the following:

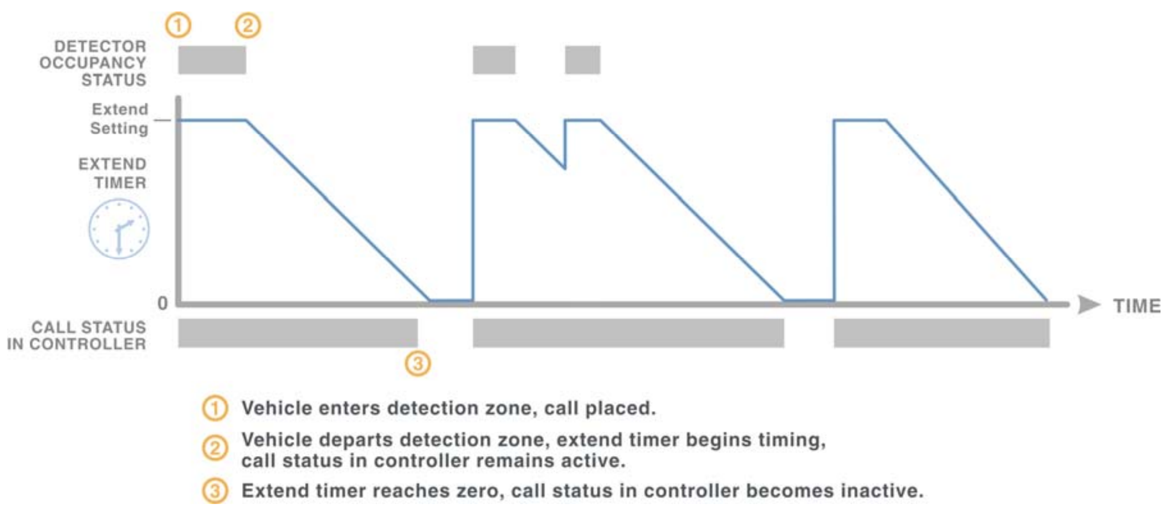
- Delay is sometimes used with stop-line, presence mode detection for turn movements from exclusive lanes. For right-turn-lane detection, delay should be considered when the capacity for right-turn-on-red (RTOR) exceeds the right-turn volume or a conflicting movement is on recall. If RTOR capacity is limited, then delay may only serve to degrade intersection efficiency by further delaying right-turn vehicles. The delay setting should range from 8 to 12 seconds, with the larger values used for higher crossroad volumes (31).

- If the left-turn movement is protected-permissive and the opposing through phase is on minimum (or soft) recall, then delay should be considered for the detection in the left turn lane. The delay setting should range from 3 to 7 seconds, with the larger values used for higher opposing volumes (32). In this case, a minimum recall should also be placed on the adjacent through phase to ensure that a lack of demand on the adjacent through phase does not result in the left-turn movement receiving neither a permissive nor a protected left-turn indication.
- Delay may also be used to prevent an erroneous call from being registered in the controller if vehicles tend to traverse over another phase's detector zone. For example, left-turning vehicles often cut across the perpendicular left-turn lane at the end of their turning movement. A detector delay coupled with non-locking memory would prevent a call from being placed for the unoccupied detector.

### 5.6.2 Extend

The extend parameter is used to increase the duration of the actuation for a detector or phase. The extend timer begins the instant the actuation channel input is inactive. Thus, an actuation that is one second in duration at the channel input can be extended to three seconds, if the extend parameter is set to two seconds. This process is illustrated in Figure 5-8.

**Figure 5-8 Application of extend timer**




Extend is typically used with detection designs that combine multiple advance detectors and stop-line detection for safe phase termination of high-speed intersection approaches. Extend is used with specific upstream detectors to supplement the passage-time parameter, to ensure that these detectors can extend the green interval by an amount of time equal to the sum of the passage time and call extension. The magnitude of the extension interval is dependent on the passage time, approach speed, and the distance between the subject detector and the next downstream detector. Typical values range from 0.1 to 2.0 seconds.

The objective when used at high-speed approaches is to extend the green interval to ensure that a vehicle approaching the intersection has just enough time to reach the next downstream detector and place a new call for green extension. The procedure for identifying when call extension is needed and computing the amount of the extension time is specific to the detection design. Refer to the *Manual of Traffic Detector Design* by Bonneson and McCoy (33) for additional guidance on this application.

### 5.6.3 Carryover

Carryover is a term commonly used for the Extend setting in controller manuals. It is another way to describe the time provided for a vehicle to traverse from one detector to the next.

### 5.6.4 Call

The call parameter is used to allow actuations to be passed to the controller for the assigned phase when it is not timing a green interval. Actuations received during the green interval are ignored. The call parameter is sometimes used with detection designs that include one or more advance detectors and stop-line detection. With this design, the call-only parameter is used with the stop-line detectors to ignore the actuations these detectors receive during the green interval. The advance detectors are used to ensure safe and efficient service during the green interval. When an appropriate detection design is combined with this parameter, intersection efficiency can be improved by eliminating unnecessary green extension by the stop-line detection.

### 5.6.5 Queue

A detector can be configured as a queue service detector to effectively extend the green interval until the queue is served, at which time it is deactivated until the start of the next conflicting phase. This functionality is offered as a parameter in most modern controllers. However, if it is not available as a parameter, equivalent functionality can be acquired by using the features of many modern detector amplifiers.

This functionality is sometimes used with detection designs that include one or more advance detectors and stop-line detection. With this design, the queue service functionality is used to deactivate the stop line detection during the green interval, but after the queue has cleared. The advance detectors are then used to ensure safe phase termination. When combined with an appropriate detection design, this functionality can improve intersection efficiency by eliminating unnecessary green extension by the stop-line detection.

## 5.7 GUIDELINES FOR TIME-BASE CONTROLS

Most controllers provide a means to externally apply signal timing parameters by time of day; typically these include maximum green, phase omit, and minimum recall on a time-of-day basis. Depending on the manufacturer, time-of-day selection of pedestrian omit, maximum recall, pedestrian recall, detector switching, overlap omit, additional maximums, alternate walk intervals, and other parameters may also be available. The approach specified by NTCIP 1202 for activating phase and ring controls invokes a timing pattern that can be selected on a time-of-day basis (34). In NTCIP protocol, a timing pattern consists of a cycle length, offset, set of minimum green and maximum green values, force off (determined by splits in some cases), and phase sequence. It also includes specification of phase parameters for minimum or maximum vehicle recall, pedestrian recall, or phase omit. This will be further described in Chapter 6.

There are a number of controls that can be used to modify controller operation on a time-of-day basis. A remote entry to one of these controls will invoke the corresponding parameter. The most common time-based controls are maximum green 2 (Max 2), phase omit, and minimum recall. The method of activating these controls varies from manufacturer to manufacturer.

There are two typical uses of phase omit. One use is when a left-turn phase is only needed during the peak traffic period. A second use is where a left-turn movement is prohibited during the peak period. In this situation, the associated left-turn phase is omitted during the turn prohibition period.

Minimum recall is used primarily on the major-road phase(s) of a fully-actuated, non-coordinated intersection. If the volume on the minor road is low only during certain times of the day, minimum recall for the major-road phases could be activated during these time periods.

## 5.8 REFERENCES


1 *National Transportation Communications for ITS Protocol: Object Definitions for Actuated Traffic Signal Controller (ASC) Units – 1202 v01.07*. National Electrical Manufacturers Association, Rosslyn, Virginia, January 2005

2 *NEMA Standards Publications TS 2-2003 v2.06: Traffic Controller Assemblies with Controller Requirements.* National Electrical Manufacturers Association, Rosslyn, Virginia, 2003

3 Orcutt, F.L. *The Traffic Signal Book*. Prentice Hall, Englewood Cliffs, New Jersey, 1993

4 Skabardonis, A., R. Bertini, B.Gallagher, "Development and Application of Control Strategies for Signalized Intersections in Coordinated Systems" *Transportation Research Record No. 1634.* Transportation Research Board, National Research Council, Washington, D.C., 1998, pp. 110-117

5 Lin, F.B. "Optimal Timing Settings and Detector Lengths of Presence Mode Full-Actuated Control." *Transportation Research Record No. 1010.* Transportation Research Board, National Research Council, Washington, D.C., 1985, pp. 37-45

6 Parsonson, P.S. *NCHRP Synthesis 172: Signal Timing Improvement Practices*. Transportation Research Board, National Research Council, Washington, D.C., 1992

7 Kell, J. H., and Fullerton, I. J., *Manual of Traffic Signal Design. 2nd Edition*. Institute of Transportation Engineers, Washington, D.C., 1998

8 Skabardonis, Alexander, *Progression Through a Series of Intersections with Traffic Actuated Controllers, Volume 2 User's Guide*, FHWA RD-89-133, Washington, D.C., October 1998

9 *Uniform Vehicle Code - Millennium Edition.* National Committee on Uniform Traffic Laws and Ordinances. Alexandria, Virginia, 2000

10 Parsonson, P.S. 1992

11 Souleyrette, R. R., O'Brien, M. M., McDonald, T., Preston, H., and Storm, R., "Effectiveness of All-Red Clearance Interval on Intersection Crashes", Center for Transportation Research and Education, Iowa State University, Minnesota Department of Transportation, MN/RC-2004-26, http://www.lrrb.org/PDF/200426.pdf, May 2004

12 Kell, J. H., and Fullerton, I. J. 1998

13 *Ibid*

14 *Traffic Engineering Handbook*. 5th Edition. J.L. Pline editor. Institute of Transportation Engineers, Washington, D.C., 1999

15 *Traffic Control Devices Handbook*. J.L. Pline editor. Institute of Transportation Engineers, Washington, D.C., 2001

16 Tarnoff, P.J. "Traffic Signal Clearance Intervals." *ITE Journal*. Institute of Transportation Engineers, Washington, D.C. April 2004, pp. 20-24

17 *Traffic Control Devices Handbook*. 2001

18 *Manual on Uniform Traffic Control Devices*. U.S. Department of Transportation, Washington, D.C., December 2003

19 *Ibid*

20 *Traffic Control Devices Handbook*. J.L. Pline editor. Institute of Transportation Engineers, Washington, D.C., 2001

21 Zegeer, C.V., C. Seiderman, P. Lagerway, M. Cynecki, M. Ronkin, and R. Schneider. *Pedestrian Facilities User's Guide - Providing Safety and Mobility.* Report No. FHWA-RD-01-102. U.S. Department of Transportation, Federal Highway Administration, Washington, D.C., 2002

22 Tarnoff, P.J., and J. Ordonez. *Signal Timing Practices: State of the Practice*. Institute of Transportation Engineers, Washington, D.C., March 2004

23 I*bid*

24 *National Transportation Communications for ITS Protocol:* January 2005

25 Henry, RD, "Signal Timing on a Shoestring", Federal Highway Administration, FHWA-HOP-07-006, March 2005.

26 Tarnoff, P. and Ordonez, 2004.

27 Bonneson, J. A., and P.T. McCoy. *Manual of Traffic Detector Design, 2nd Edition*. Institute of Transportation Engineers, Washington DC, May 2005

28 Tarnoff, P.J., and P.S. Parsonson. *NCHRP Report 233: Selecting Traffic Signal Control at Individual Intersections*. Transportation Research Board, National Research Council, Washington, D.C., 1981

29 *UDOT Procedural Update - Timing of Traffic Signals*. Traffic Operations Center, Utah Department of Transportation, Ogden, Utah, April 15, 2004

30 *National Transportation Communications for ITS Protocol:* January 2005

31 Bonneson, J.A., and M. Abbas. *Intersection Video Detection Manual*. Report FHWA/TX-03/4285-2. Texas Department of Transportation, Austin, Texas, September 2002

32 Bonneson, J. A., and P.T. McCoy. 2005

33 *Ibid*

34 *National Transportation Communications for ITS Protocol:* January 2005

# CHAPTER 6

# COORDINATION

## TABLE OF CONTENTS

## LIST OF TABLES

# LIST OF FIGURES

# 6. COORDINATION

This chapter presents the concept of coordination of traffic signals. Coordination is a tool to provide the ability to synchronize multiple intersections to enhance the operation of one or more directional movements in a system. Examples include arterial streets, downtown networks, and closely spaced intersections such as diamond interchanges. This chapter identifies coordination concepts using examples from research and practice. It contains four sections. The first section provides an overview of coordination including a summary of objectives, the fundamental concepts, and expectations of coordination timing. The second section describes the concepts for coordination, its effect on time allocation, implementation issues, and time-space diagrams. The third section provides guidelines for developing coordination timing plans, and the fourth section describes complexities associated with coordinated operations. The intent of this chapter is to provide necessary background for the development of timing strategies.

## 6.1 TERMINOLOGY

This section identifies and describes basic terminology used within this chapter. Additional terms can be found in the Glossary section of the Manual.

*Coordination*

The ability to synchronize multiple intersections to enhance the operation of one or more directional movements in a system.

*Double Cycle*

A cycle length that allows phases to be serviced twice as often as the other intersections in the coordinated system. This is also referred as a “Half Cycle”.

*Early Return to Green*

A term used to describe the servicing of a coordinated phase in advance of its programmed begin time as a result of unused time from non-coordinated phases.

*Force-off*

A point within a cycle where a phase must end regardless of continued demand. These points in a coordinated cycle ensure that the coordinated phases are provided a minimum amount of green time.

*Fixed Force-off*

A force-off mode where force-off points cannot move. Under this mode, non-coordinated phases can use unused time of previous phases.

*Floating Force-off*

A force-off mode where force-off points can move depending on the demand of previous phases. Under this mode, non-coordinated phase times are limited to their defined split amount of time and all unused time is dedicated to the coordinated phase. Essentially, the split time is treated as a maximum amount for the non-coordinated phases.

*Master Clock*

The background timing mechanism within the controller logic to which each controller is .referenced during coordinated operations.

*Offset*

The time relationship between coordinated phases defined reference point and a defined master reference (master clock or sync pulse).

*Offset Reference Point (Coordination Point)*

The defined point that creates an association between the local clock at a given signalized intersection and the master clock.

*Permissive Period*

A period of time after the yield point where a call on a non-coordinated phase can be serviced without delaying the start of the coordinated phase.

*Time-Space Diagram*

A chart that plots the location of signalized intersections along the vertical axis and the signal timing along the horizontal axis. This is a visual tool that illustrates coordination relationships between intersections.

*Yield Point*

A point in a coordinated signal operation that defines where the controller decides to terminate the coordinated phase.

## 6.2 PRINCIPLES OF COORDINATED OPERATION

The decision to use coordination can be considered in a myriad of ways. There are numerous factors used to determine whether coordination would be beneficial. Establishing coordination is easiest to justify when the intersections are in close proximity and there is a large amount of traffic on the coordinated street. The MUTCD provides the guidance that traffic signals within 800 meters (0.5 miles) of each other along a corridor should be coordinated unless operating on different cycle lengths.

### 6.2.1 Coordination Objectives

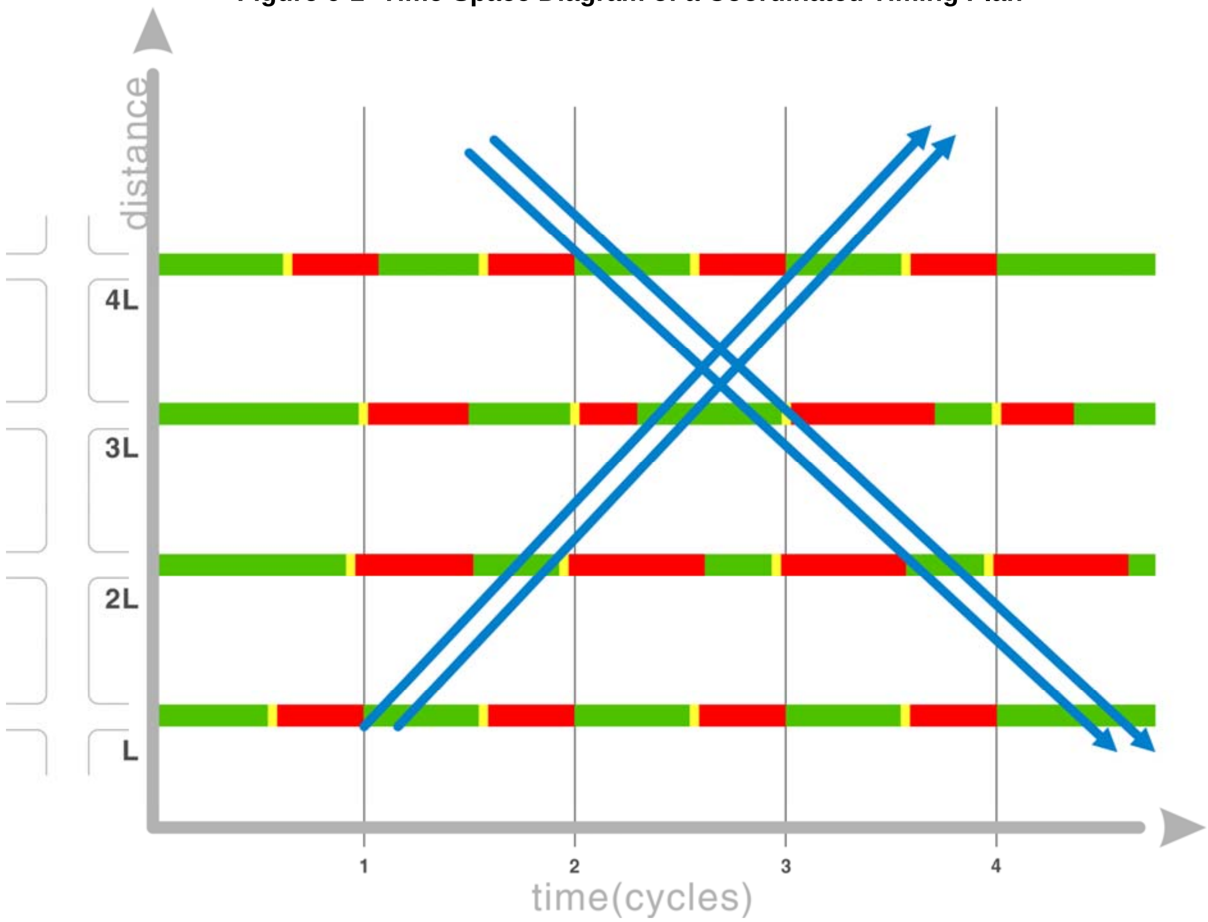
Coordination is largely a strategic approach to synchronize signals together to meet specific objectives. While there are numerous objectives for the coordination of traffic signals, a common objective is stated succinctly in the National Report Card:

**The intent of coordinating traffic signals is to provide smooth flow of traffic along streets and highways in order to reduce travel times, stops and delay(1).**

A well-timed, coordinated system permits continuous movement along an arterial or throughout a network of major streets with minimum stops and delays, which, reduces fuel consumption and improves air quality (2). Figure 6-1 illustrates the concept of moving vehicles through a system of traffic signals using a graphical representation known as a *time-space diagram*. The time-space diagram will be described in additional detail in later sections of this chapter.

The time-space diagram is a chart that plots ideal vehicle platoon trajectories through a series of signalized intersections. The locations of intersections are shown on the distance axis, and vehicles travel in both directions (in a two-way street). Signal timing sequence and splits for each signalized intersection are plotted along the time axis. It is very important these plots are to scale so that a consistency between units can be maintained. The time axis illustrates what motorists on the arterial will experience as they travel down the street. Left turns are shown as angled lines that either are operated with a concurrent green for the same direction or not.

**Figure 6-1  Time-Space Diagram of a Coordinated Timing Plan**




The result of signal coordination is illustrated on the time-space diagram above. The start and end of green time show the potential trajectories for vehicles on the street. It is these trajectories that determine the performance of the coordination plan. Performance measures include stops, vehicle delay, and arterial travel time; they can also include other measures such as changes in delay to transit vehicles or existence of spill back queuing between closely spaced intersections. In general, effective signal coordination should enable the engineer to meet the objective defined as a part of the study or relevant policies associated with the community. While the effectiveness of the coordination timing plan is directly related to the performance measures defined from the policy, it is also determined by the user experience and their perception of signal displays. Successful coordinated signal timing plans are usually characterized by both the audience and the measure:

- Downtown merchants may favor pedestrian traffic over vehicular traffic;
- A neighborhood may seek reduced traffic and lower speeds;
- The state may be interested in traffic volume throughput on state highways; and
- A transit agency may be concerned with signal delay to buses and/or light rail vehicles.

Ultimately, the perceived effectiveness of a coordination plan will depend upon local transportation policies, the elected leadership, relevant stakeholders, and the selected performance measures specific to the community.

### 6.2.2 When to Use Coordination

Numerous factors can be used to determine whether coordination would be beneficial. Establishing coordination is easiest to justify when the intersections are in close proximity to one another and when traffic volumes between the adjacent intersections are large. The need for coordination can be identified through observation of traffic flow arriving from upstream intersections. If arriving traffic includes platoons that have been formed by the release of vehicles from the upstream intersection, coordination should be implemented. If vehicle arrivals tend to be random and are unrelated to the upstream intersection operation, then coordination may provide little benefit to the system operation.

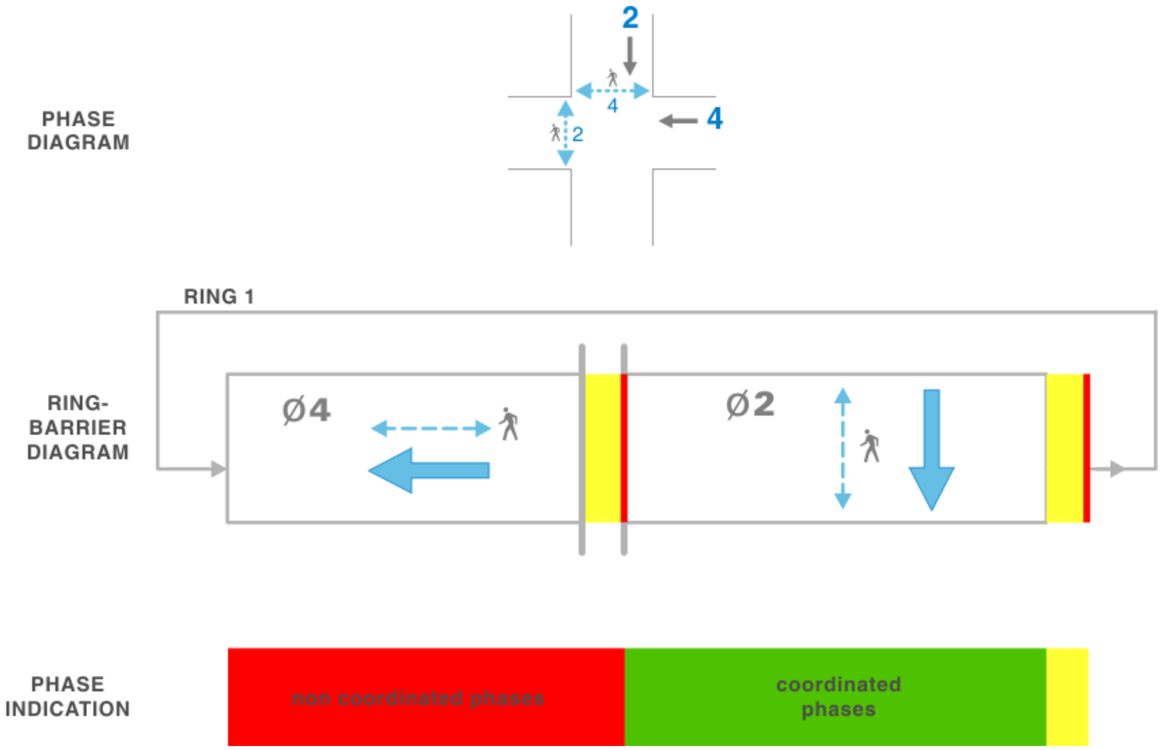
Information presented in the FHWA report, *Signal Timing on a Shoestring*, revealed that using both simple and complex procedures worked for identifying intersections for coordination. In short, when intersections are close together (i.e., within ¾ mile of each other) it is advantageous to coordinate them. At greater distances (i.e., ¾ mile or greater), the traffic volumes and potential for platoons should be reviewed to determine if coordination would be beneficial to the system operations. In both cases, the traffic conditions and the community policies should be considered as a part of the decision.

### 6.2.3 Fundamentals of Coordination

With the modern signal controller, coordination is accomplished by adding a layer of logic (that is, coordination logic compliments some basic features such as when a phase can begin or end) to the basic actuated logic used for isolated signal timing operations (discussed in Chapter 5). In previous chapters, the details of the controller settings were limited to those applied at isolated or independent intersections (maximum green, pedestrian timing, etc). Signal coordination establishes an additional set of time constraints among a series of signalized intersections by establishing a background cycle length (on each ring of the phase diagram). This cycle length includes a series of timers for each phase and requires the designation of one phase (in some cases one phase on each ring) as the coordination phase. This designation identifies the phase that will be the last (or first depending on your outlook) one to receive its allocation of green time. This coordinated phase is distinguished from other actuated phases because it always receives a minimum amount of assigned green time. While it is possible to have a portion of the coordinated phase be actuated, the important point is that there is a non-actuated interval for the designated coordinated phase(s) that is “guaranteed” every cycle for the purposes of coordination.

Figure 6-2 shows the intersection of two one-way streets, which results in a simple intersection with just two phases. It also shows the relationship between a phase diagram with movements assigned to the indications at the intersection, a ring-and-barrier diagram that illustrates the logic used in the controller for phase sequence and to establish the relationship between conflicting and complimentary movements, and the time-space diagram which would be used to display the relationship between this intersection and others along an arterial or within a street network (previously shown in Figure 6-1). This intersection only has one ring within the ring-and-barrier diagram and a simple time-space diagram. Phase 2 is identified as the coordinated phase. This intersection represents the most basic configuration for a signalized intersection.

**Figure 6-2  Phase and ring-and-barrier diagrams of intersection of two one-way streets**




In an actuated-coordinated system, the first event in the cycle is the non-coordinated phase, which would serve the demand on that phase (in this case, just phase 4). If demand is less than the time allocated to that phase, it would gap out, and the remaining time for that movement would be reallocated to the coordinated phase (phase 2).

In Chapter 4, the ring-and-barrier diagram was introduced for a full-movement four-legged intersection, where Ring 1 consists of a sequence of phases and Ring 2 consists of the concurrent phases for the intersection, while the barrier separates intersecting movements (east-west and north-south). Figure 6-3 illustrates Ring 1 of the diagram and provides a graphical example of the phase indications for left-turn phasing. In this example, phase 2 is the coordinated movement, which means that phases 3, 4, and 1 all have an opportunity to use portions of their allotted time before phase 2 begins. If any time allotted for phases 3, 4, or 1 is unused, phase 2 will start before the normal start time (commonly referred to as early return to green) and then rest in green until its end point. The following sections of this chapter explain in more detail the nuances between different settings that affect the coordination relationship. Essentially, this is an additional layer of constraint (coordination logic) that can be applied to signalized intersections to improve the operation of a system of traffic signals.

**Figure 6-3 Example of coordination logic within one ring**

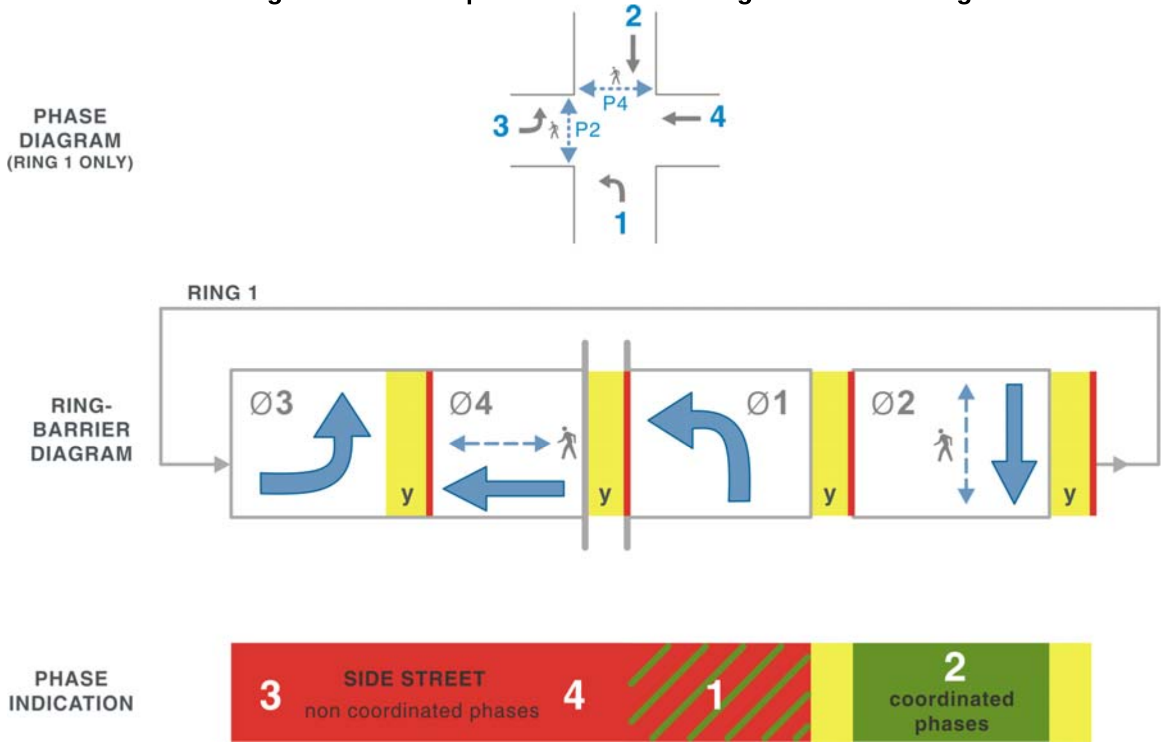
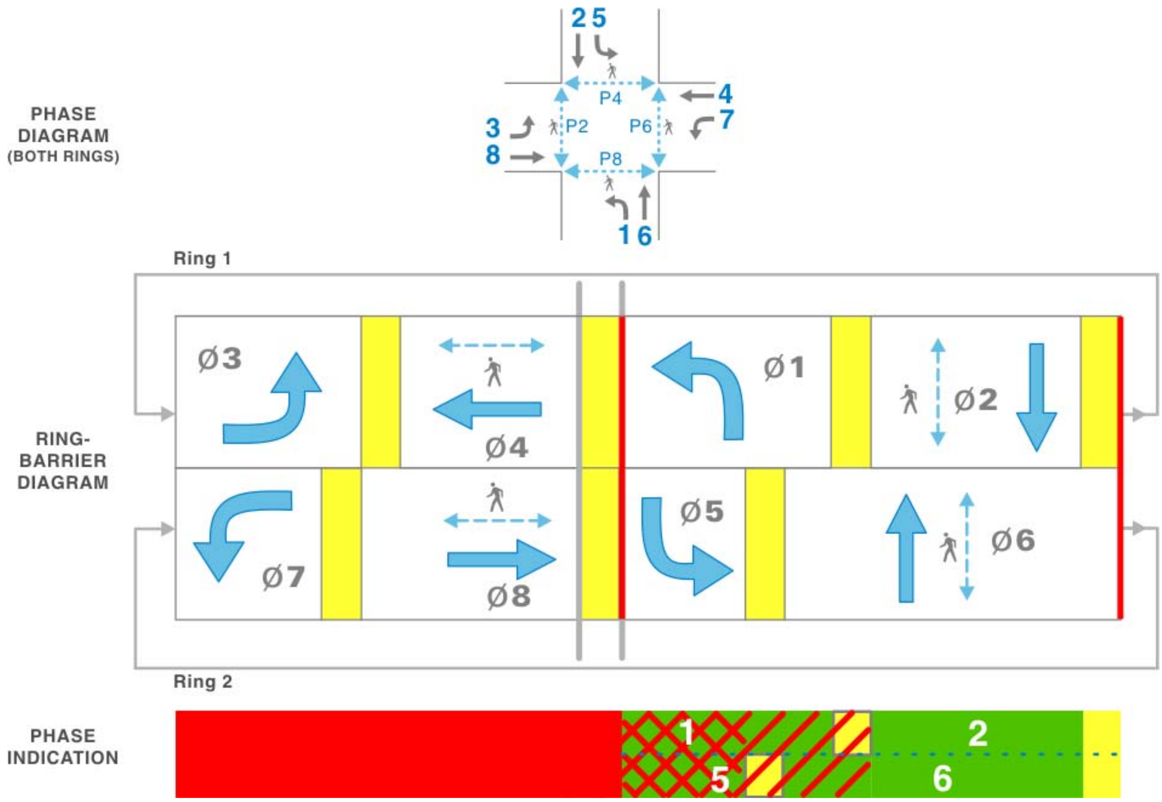



Figure 6-4 shows an example relationship between two rings (rings 1 and 2) at a typical 8-phase intersection. As seen in the figure, phase 5 uses less time than 1. As a result, phases 1 and 6 will time concurrently before phase 1 ends and phase 2 begins. In the time-space diagram, this concurrent operation of phases 1 and 6 is indicated by downward sloping left-to-right red bars, which indicate that the southbound movement is able to move through the intersection, but not the northbound movement. Additional examples of this are shown in section 6.2.4.

**Figure 6-4  Coordination using two rings**




The effects of coordination at an individual intersection depend on the timing plan and the operations of adjacent intersections. The coordination may be beneficial to a vehicle traveling between the two intersections, however it may negatively impact a pedestrian or vehicle crossing the street, waiting for the signal to provide the right of way. Many of the complaints from citizens related to the use of coordination address the restrictions placed that inhibits responsiveness to demand. The last section of this chapter describes the complexity of coordination in greater detail. This will be further exacerbated if the cycle length selected is unnecessarily long or the coordination plan is operating when traffic volumes are lower than typical (holidays that fall on a weekday).

### 6.2.4 Summary

Coordination applications can range from two signals controlling a diamond interchange, to dozens of signals controlling a combination of actuated and fixed time controllers controlling an arterial system that bisects a grid network. When applied, coordination strategies should follow the policies and resulting objectives for signal timing established by the agency. Once these are defined, performance measures can be established to determine whether the application is beneficial. Various performance measures can be used to evaluate signal coordination. Isolated intersection performance measures include delay, queue length, and safety; however, performance measures for coordinated systems may be slightly different. Evaluation methods include the number of stops reduced for the main street through movements or queue management.

**Designating the traffic movement with the greatest peak hour demand as the coordinated phase is the most common practice. The coordination logic provides unused green time for the coordinated phase especially when demand for the other movements is low, which can result in fewer stops for the traffic movements with greatest demand. This is most typically through major street through movement.**

As previously mentioned, other measures of effectiveness include reduced through movement and intersection delay, reduced travel times, lower emissions, lower fuel consumption, and maximized bandwidth. The *Highway Capacity Manual* quantifies some of these measures of effectiveness (through movement delay and resulting through-through travel speed), this manual does not. Instead, this chapter presents the concepts needed to operate a signal system. It should also be noted that from a system perspective the *Highway Capacity Manual* procedures are also not sensitive to the policy issues discussed here. The effectiveness of the signal system is (and should be) based on the policies and expectations for the various agencies. Measures of effectiveness can vary from travel speed, travel times, number of stops, pedestrian safety, pedestrian delay, transit efficiency, and overall intersection delay. In cases of public approval, the number of complaints or phone calls can be used as a measure of effectiveness, but may be biased toward a narrow perspective.

As previously described, well-defined objectives should be the starting point of the system evaluation. If an agency is focused on efficiency of automobiles, then the objectives will correspond to reduced travel time and delay for a given movement or intersection. The signal timing will reflect the priority given to the coordinated movement and, as a result, the through movement will have a higher percentage of the cycle time. However, if the agency wants to provide a system that is focused on moving people, then transit efficiency measures, such as percent on-time, ridership, and travel time, should be evaluated.

Many performance measures are difficult to quantify. Computer software programs can estimate many performance measures, including network delay, emissions, and fuel consumption, that are not easily measured in the field. However, one must be careful in selecting software tools, making sure they reflect the capabilities of the control software being used including timing settings and the detailed design of the detection system. Some performance measures are more difficult to quantify, which makes it more difficult to evaluate them objectively and to use them explicitly in an optimization exercise. These include public perceptions measured by phone calls (positive and negative), differences in perception of wasted time when conflicting traffic is present compared to when the intersection appears empty, and differences in perception of stops at minor intersections compared to major intersections. These measures require the traffic engineer to employ judgment—some might call it art—in balancing these less quantifiable measures with the other more scientific measures.

## 6.3 COORDINATION MECHANICS

Three fundamental parameters distinguish a coordinated signal system: cycle length, offset and split. These settings are necessary inputs for coordination. Figure 6-5 shows the cycle length and set of splits for an intersection, along with the offset between two intersections and the relationship to the master clock. There are several ways these inputs can be interpreted by the controller and thus a description of how the inputs are used to develop the relationship between the various intersections is provided here.

### 6.3.1 Cycle Length

Cycle length defines the time required for a complete sequence of indications. Cycle lengths must be the same for all intersections in the coordination plan to maintain a consistent time based relationship. (One exception would be an intersection that "double cycles," serving the phases twice as often as the other intersections in the system.) The cycle length is measured from the deterministic point defined by the user. Coordination occurs most commonly along an arterial at an interchange or between at least two signals, but network coordination in downtown or other grid networks is also common. Professionals have determined cycle length through a variety of ways. The guidelines section of this chapter further discusses how to establish the cycle length for a coordinated timing plan.

### 6.3.2 Yield Point

The first component of coordination is often referred to as the yield point, but may be better defined as the deterministic point. This point is necessary for coordination to operate because it is a point where the controller makes a decision to terminate the coordinated phase. The controller will not leave the coordinated phase immediately because it has to confirm there is a conflicting call and time some coordinated phase clearance intervals (pedestrian clearance if resting in walk or just the yellow and red clearance).

Most controllers confirm what time-of-day plan should be running at this point in the cycle and transition to try to get to the appropriate plan at this point. It is also at this point that the controller will seek to serve the next phase in the sequence that has a call. In the instance where there is a call on any other phase, the coordinated phase would begin to terminate based on coordinated phase dwell state (walk or don't walk) and clearance (yellow and red clearance time) and the next phase would receive a green indication based on the demand. If there is no demand on any other phase, the cycle would continue in the coordinated phase until the end of one or more yield periods, or the next occurrence of the yield point, where the controller would serve the phases in sequence. For each cycle, the controller decides at the yield point (or later with permissive periods) what phase(s) will serve.

### 6.3.3 Splits

Within a cycle, splits are the portion of time allocated to each phase at an intersection. These are calculated based on the intersection phasing and expected demand. Splits can be expressed either in percentages of the cycle or in seconds. Split percentages typically include the yellow and all-red associated with the phase; as a result, the green percentage is less than total split for a phase. For implementation in a signal controller, the sum of the phase splits must be equal to (or less than) the cycle length, if measured in seconds, or 100 percent, if measured as a percent (some traffic signal controllers are more finicky about this than others). In traditional coordination logic, the splits for the non-coordinated phases define the minimum amount of green for the coordinated phases.

**Figure 6-5 Cycle Length and Split**

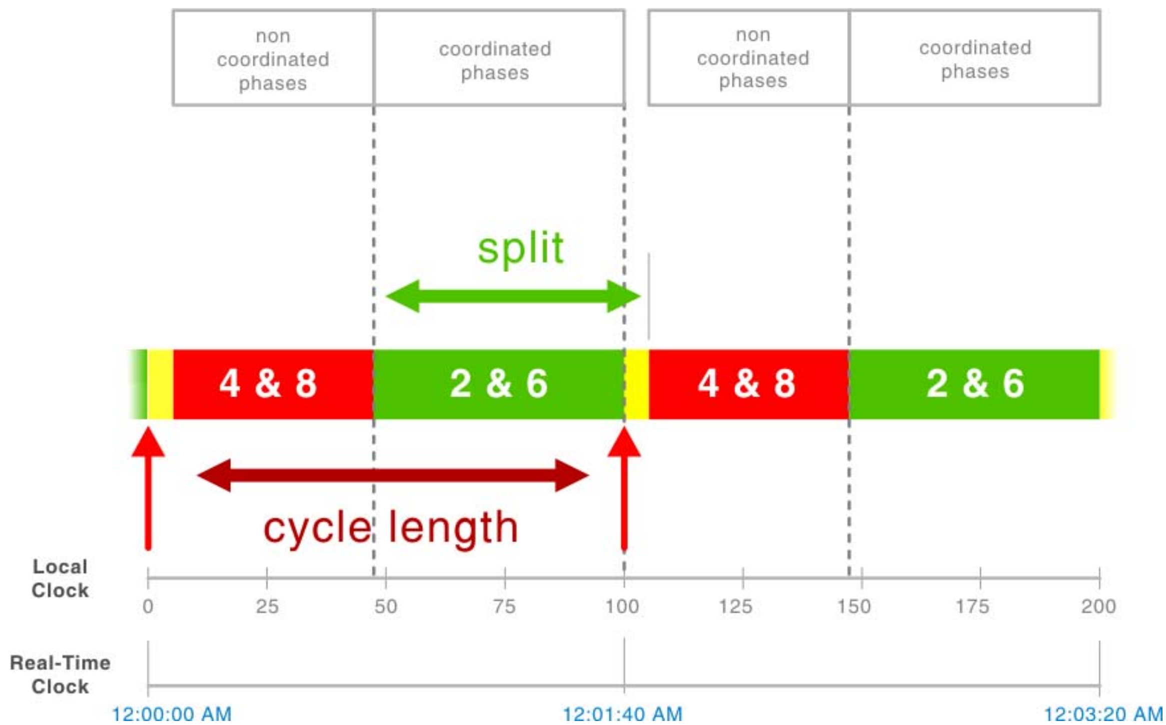



**Figure 6-5 is a time-space diagram that shows a simplification of the signal indications for the coordinated and non-coordinated phases. The measured split for a phase consists of its green time, yellow change, and red clearance times. The cycle length is the sum of time for the complete sequence of indications. The measured split may be longer than what is input into the controller because of the early return to green. In an actuated-coordinated system, the cycle length must be measured from a defined observable point, typically the end of the coordinated phase green or beginning of coordinated phase yellow. Measuring the cycle length from the observed start of green at an actuated coordinated intersection will result in erroneous results because of the early return to green that can occur.**

***Force-offs***

Force-offs are used in some controllers as an alternate way to control the phase splits. The force-offs are points where non-coordinated phases must end even if there is continued demand. The use of force-offs overlays a constraint on all non-coordinated phases to ensure that the coordinated phase will receive a minimum amount of time for each cycle.

In some controllers, this might be less than the pedestrian timing requirements, which offers the engineer some flexibility in timing. However, this flexibility comes at the price of potentially losing coordination if the controller does not return to the coordinated phase at its assigned time. Losing coordination under light traffic or only very occasionally due to pedestrian calls may be an acceptable option.

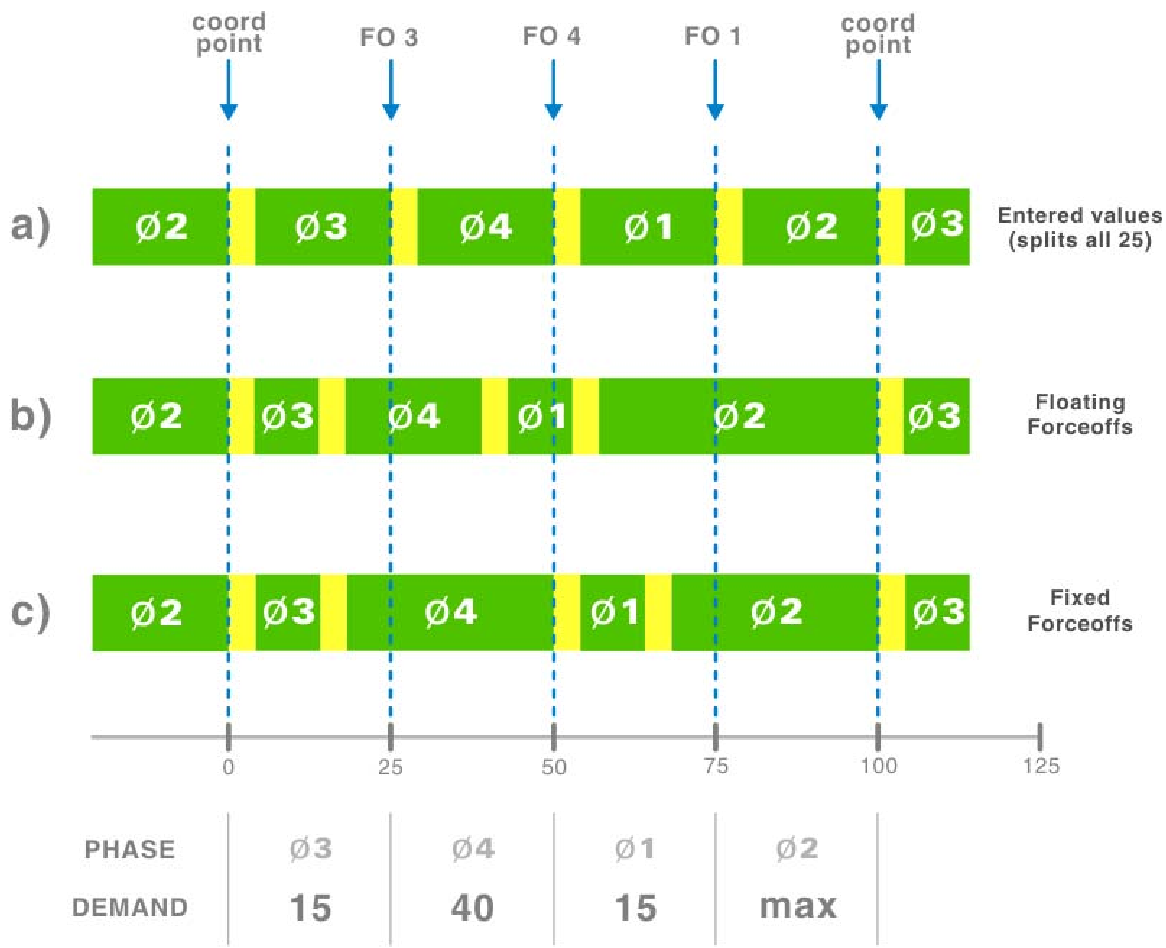
There are two options for programming force-offs in controllers, fixed or floating. The fixed force-off maintains the phase's force-off point within the cycle. If a previous non-coordinated cycle ends its phase early, any following phase may use the extra time up to that phase's force-off. This is beneficial if there are fluctuations in traffic demand and a phase needs more green time. One of the outcomes of this is that a phase later in the sequence (before the coordinated phase) may receive more than its split time (provided the maximum green is not reached). It should be noted that the phase directly after the coordinated phase will never have an opportunity to receive time from a preceding phase, regardless of the method of force-offs.

Floating force-offs are limited to the duration of the splits that were programmed into the controller. The force-off maintains the non-coordinated maximum times for each non-coordinated phase in isolation of one another. Floating force-offs are more restrictive for the non-coordinated phases. If a phase does not use all of the allocated time, then all extra time is always given to the coordinated phase. This is illustrated in Figure 6-6.

The maximum green timer, if allowed, may also result in the phase not reaching this force-off value. In addition to the maximum green timer, a definable controller parameter, known as Inhibit Max, may be invoked to prevent the controller from using the maximum green to limit the time provided to a phase during coordinated operation.

Figure 6-6 shows the difference between fixed and floating force-offs. The first row (“row a”) illustrates a scenario where demand exceeds the allotted green time and each phase is terminated at the respective force-off points. The second and third rows (“row b” and “row c”) illustrate the concepts of the floating and fixed force-off concepts. To better illustrate the differences in the two concepts, the demand for the phases are different. In this example, phases 1 and 3 experience a demand of 15 seconds (10 seconds shorter than the split time), and phase 4 experiences a demand of 40 seconds (15 seconds longer than the split time).

**Figure 6-6  Fixed and Floating Force-offs**




In this example phase 2 is the coordinated phase, phases 1 and 3 gap out, and phase 4 maxes out due to high demand. With fixed force-offs, the green time for phase 4 is extended to serve an increased demand up to the force-off point; in this case, it receives additional time from phase 3. The coordinated phase is given additional green time due to the previous phase (phase 1) gapping out. The green time is not taken from the other phases. For the same scenario under floating force-offs, phase 4 would be forced off even with the higher demand at its split value, 25 seconds. A Texas Transportation Institute report summarizes advantages and disadvantages of fixed force-offs (3):

- Fixed force-offs are useful to allow use of the time available from phases operating under capacity by phases having excess demand, which varies in a cyclic manner. This is the case when the phase(s) earlier in the phasing sequence is under capacity more often than the other phases.
- Fixed force-offs may reduce the early return to coordinated phases, which can be helpful in a network with closely spaced intersections. An early return to the coordinated phase at a signal can cause the platoon to start early and reach the downstream signal before the onset of the coordinated phase, which results in poor progression.
- Fixed force-offs reduce the early return to the coordinated phase which can also be a disadvantage. Under congested conditions on the arterial, an early return can result in the queue clearance for coordinated phases. Minimizing early return to coordinated

phases can cause significant disruption to coordinated operations. This disadvantage can be overcome by adjusting the splits and/or offsets at the intersection to minimize disruption.

***Permissive Periods***

A permissive period represents a period of time during the cycle in which calls on conflicting phases will be accepted. If a vehicle arrives after this period, it will have to wait until the next cycle to be served. In older actuated controllers, using coordinated operation, it was necessary to specify the permissive periods and force-offs for each phase. These values may be needed for entering signal timing plans in traffic controllers as well as traffic simulation models. Newer controllers generally automatically calculate (if allowed) the maximum permissive periods for the actuated phases. Care should be given to the selection of the coordination mode so as to not limit the benefit of larger permissive periods.

### 6.3.4 Offsets

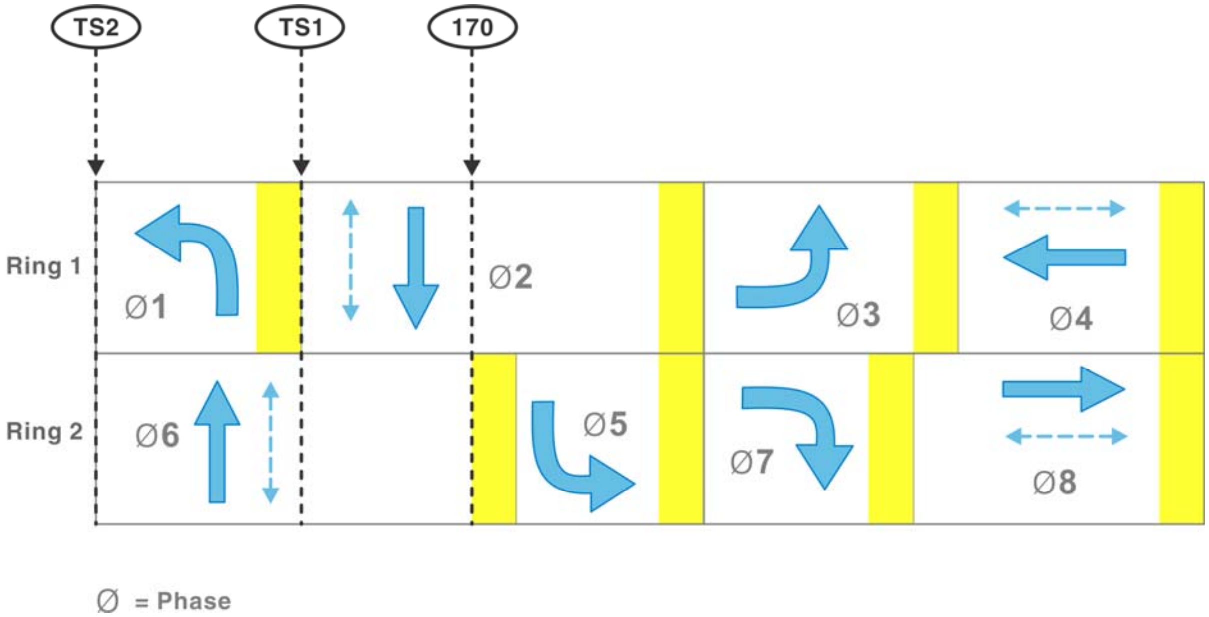
The term offset defines the time relationship, expressed in either seconds or as a percent of the cycle length, between coordinated phases at subsequent traffic signals. The offset is dependent on the offset reference point, which is defined as that point within a cycle in which the local controller's offset is measured relative to the master clock. It is not necessarily the same as the deterministic point (or yield point) within the cycle. The master clock is the background timing mechanism within the controller logic to which each controller is referenced during coordinated operations. This point in time (midnight in some controllers, user defined in others) is used to establish common reference points between every intersection. Each signalized intersection will therefore have an offset point referenced to the master clock and thus each will have a relative offset to each other. It is through this association that the coordinated phase is aligned between intersections to create a relationship for synchronized movements.

**The location of the yield point and the offset reference point describes the relationship between the coordination plan at the individual intersection and the master clock.**

Different offset reference points are associated with each of the three major controller types: NEMA TS1, NEMA TS2, and the Type 170. The NEMA TS1 references the offset point at the start of the coordinated phases (e.g., phases 2 *and* 6). The NEMA TS2 references the offset point from the start of the green indication of the first coordinated phase (e.g., phases 2 *or* 6). The 170 typically references the offset point from the start of the coordinated phase yellow. For each of these controller types, software may allow variations of these designations. Figure 6-7 illustrates the differences of each offset reference point on a two-ring diagram.

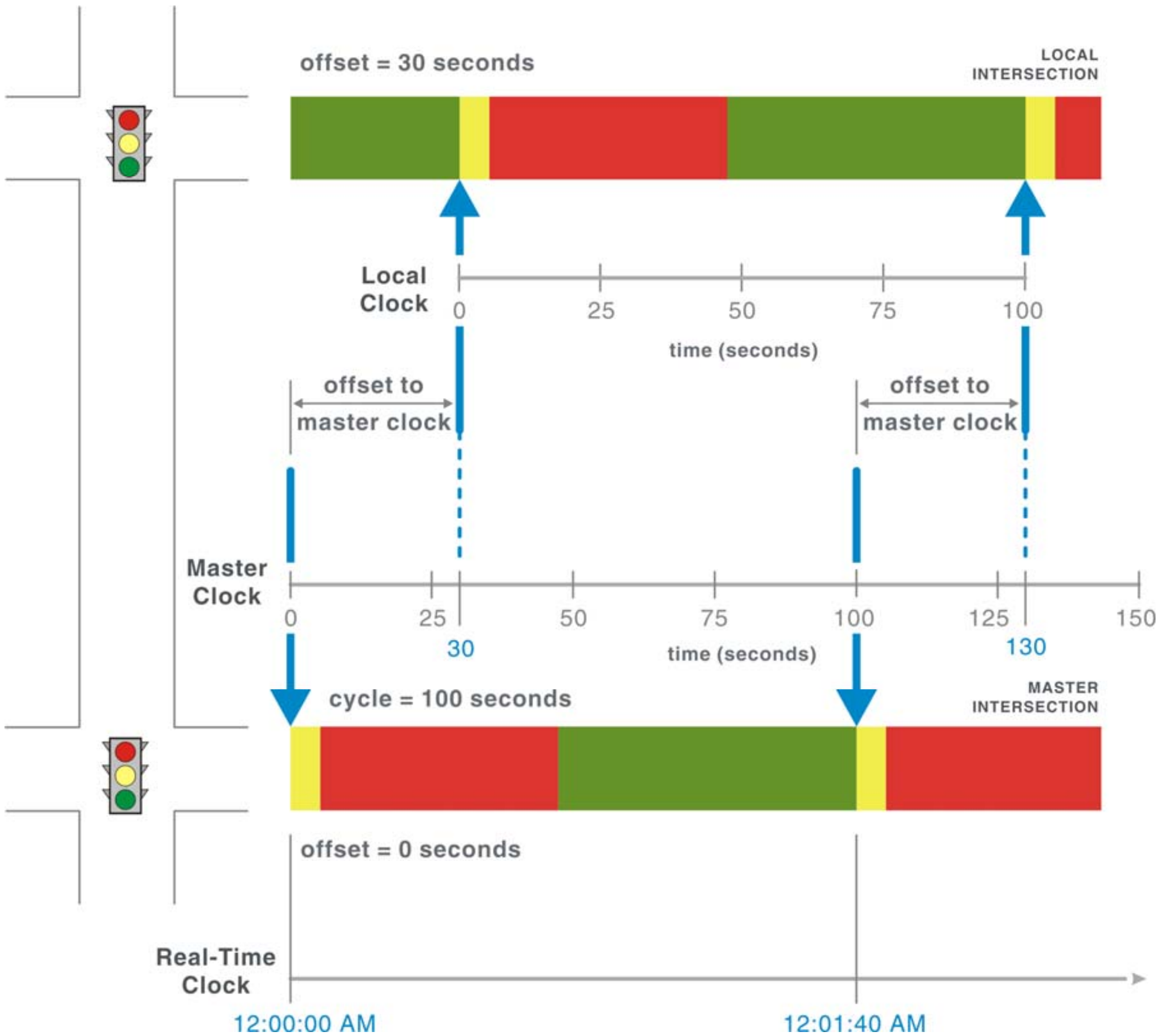
Of the three reference points, only the use of the start of coordinated phase yellow is readily observable in the field. Under this type of designation, if Intersection B has an offset of 20 seconds after Intersection A, one should see Intersection B's yellow twenty seconds after Intersection A's yellow. For both NEMA designations, the use of start of coordinated phase green as an offset reference point is not a fixed point due to the variability in the start of green (early return to green) under typical actuated-coordinated operations. However, knowledge of the assigned split for the coordinated phase can allow one to calculate the observable fixed point in the cycle. In some cases, the start of coordinated phase Flashing Don't Walk is used as an offset reference point. The majority of figures in this manual use the beginning of the coordinated phase yellow as the offset reference point, as it is easily observed in the field, although other references (including the HCM) often use start of coordinated phase green as the offset reference point.

**Figure 6-7  Standard Offset Reference Points for Type 170, NEMA TS1, and NEMA TS2 Controllers**




Once the reference point is identified, the offset is defined as the time that elapses between when this reference point occurs at the master clock and when it occurs at the subject intersection. Figure 6-8 illustrates this concept.  In this example, the offset reference point is at the start of coordinated phase yellow, and a cycle length of 100 seconds is used for both intersections. The offset of the intersection on the bottom of the figure is zero and thus matches the master clock, which is referenced to midnight. The top intersection is set to an offset of 30 seconds. The coordinated phase begins its yellow at 30 seconds and 130 seconds (12:00:30 AM and 12:02:10 AM), always 30 seconds after the bottom intersection. The relative offset is observed by the user from intersection to intersection, but this can be different from the offset to the master clock.

**Figure 6-8 Relationship between the Master Clock, Local Clock, and Offset**




It is important that each intersection have consistent master clocks to enable time-of-day plans and preemption to use this as a base point. It is also important to understand that when the cycle length is changed, most controllers calculate a new "sync" point based on the master clock reference point. The start of the master clock, also known as Pattern Sync Reference, may occur at midnight or other times during the early morning, i.e. 1:00 AM or 3:00 AM. The selected time of day should avoid a transition during significant traffic volumes. This time-based reference requires the controller be configured to keep track of subtle issues, such as if the area follows daylight savings time and/or when daylight savings time begins and ends. For example, recent changes in Indiana in 2006 required all controllers in the state to be reconfigured to acknowledge daylight savings time. In 2007, every traffic controller in the country in states using daylight savings time had to be adjusted to account for a different start of daylight savings time.

### 6.3.5 Other Coordination Settings

There are a number of controller settings that may also be known as coordination modes in some controllers. Each controller type uses different coordination modes to give flexibility to the user. Such coordination modes include "Rest in Walk", "One or more Permissive Periods", and "Actuated-Coordinated Mode". The modes operate in different manners, but each is designed to provide flexibility towards serving the users' needs. Applications of these modes may vary with respect to high pedestrian volumes, transit priority, or leading and lagging left turns.

### 6.3.6 Pre-timed and Actuated Comparison

In pre-timed systems, typical of downtown closely spaced intersections, the time relationships are "fixed." Today, most pre-timed systems use actuated controllers with phases recalled to their maximum time. In these types of systems, care should be given to the selection of walk times in order to provide a pedestrian friendly environment. Some controllers have a "fixed-time" mode which maximizes the walk time.

In systems with actuated phases, these relationships are less rigid and more complex. The "coordinator" uses many parameters to define where those time relationships vary and by how much. While the basic timing parameters of coordination are cycle length, splits, force-offs, coordinated phases, and offsets; it is very important to understand the complexities of these settings and their effect on coordinated actuated operation. The following section presents the theoretical construct of coordination in an attempt to break down the complexities into the basic fundamentals which are necessary to implement coordination consistent with the signal timing design undertaken.

## 6.4 TIME-SPACE DIAGRAM

The time-space diagram is a visual tool for engineers to analyze a coordination strategy and modify timing plans. The main components in a time-space diagram that are inputs include individual intersection locations, cycle length, splits, offset, left turn phasing (on the arterial in the direction of the diagram), and speed limit. The phase lengths may be approximations of their duration; in an actuated system this changes on a cycle-by-cycle basis. The outputs of a time-space diagram include bandwidth (or vehicle progression opportunities), estimates of vehicle delay, stops, queuing and queue spillback. The following sections describe the components of a time-space diagram and how the diagram can be used to evaluate signal coordination.

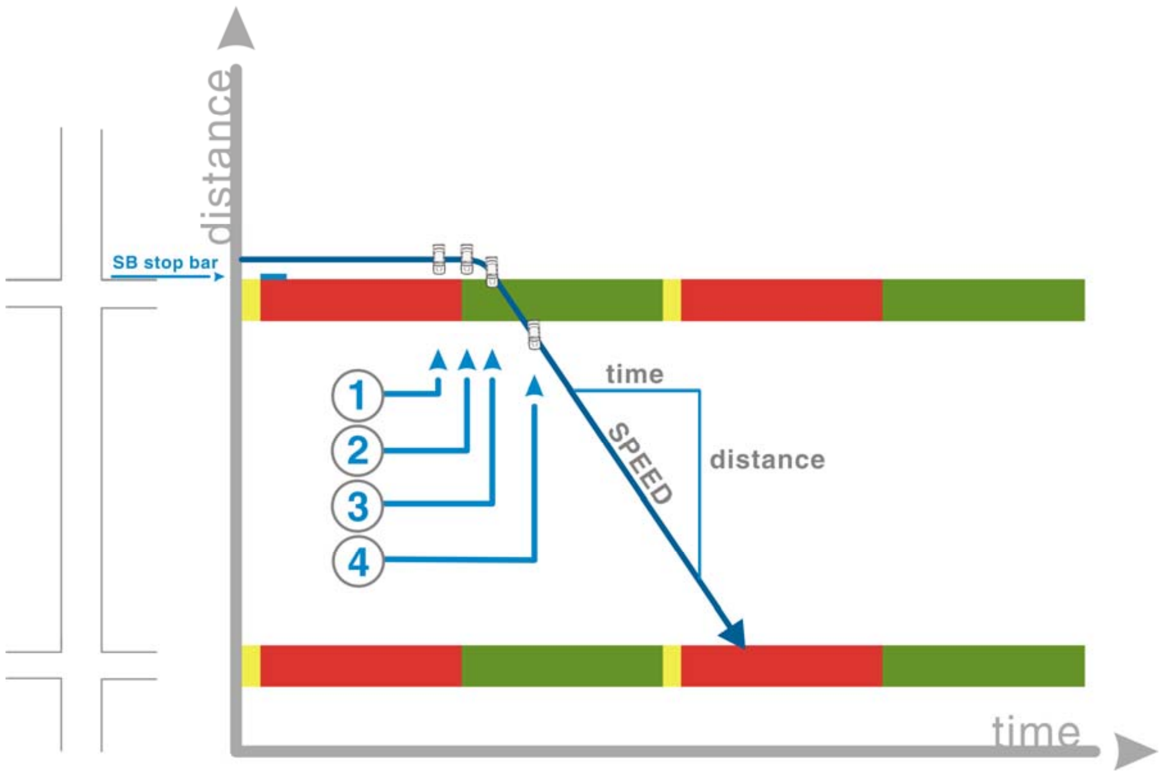
### 6.4.1 Basic Concepts (Time, Distance, Speed, and Delay)

A time-space diagram is drawn with time on the horizontal axis and distance (from a reference point) on the vertical axis. The time is relative from the master clock described earlier. Vehicle trajectories are plotted on the time-space diagram and the difference in distance over time (distance divided by time, or change in y divided by x) represents the speed or a sloped line on the diagram. The trajectories always move left to right along with time, and as shown the distance traversed can be either northbound (bottom to top of the diagram) or southbound (top to bottom). Vehicles can have a positive or negative slope that indicates the movements on a street network. Stopped vehicles (no change in distance) are shown as horizontal lines. The assumed speed for coordination on the corridor may be the speed limit, the 85$^{th}$-percentile speed, or a desired speed. The resulting speeds on the corridor are affected by the presence of other traffic, the signal timing settings, the acceleration and deceleration rates of the vehicles, and other elements within the streetscape. The acceleration rates are especially important considering the departures of standing queues at intersections.

Figure 6-9 shows the one-way street described in Figure 6-2 adjacent to another signalized intersection. In this diagram, the motorist experiences four different conditions in moving from a stop to the progression speed, these are:

1. Vehicles delayed (no change in distance as time moves forward);
2. Driver perception--reaction time at the onset of green;
3. Vehicle acceleration; and
4. Running speed of the vehicle (often assumed to be the speed limit or an estimated progression speed)

**Figure 6-9  Time-Space Diagram – Basic Concepts**

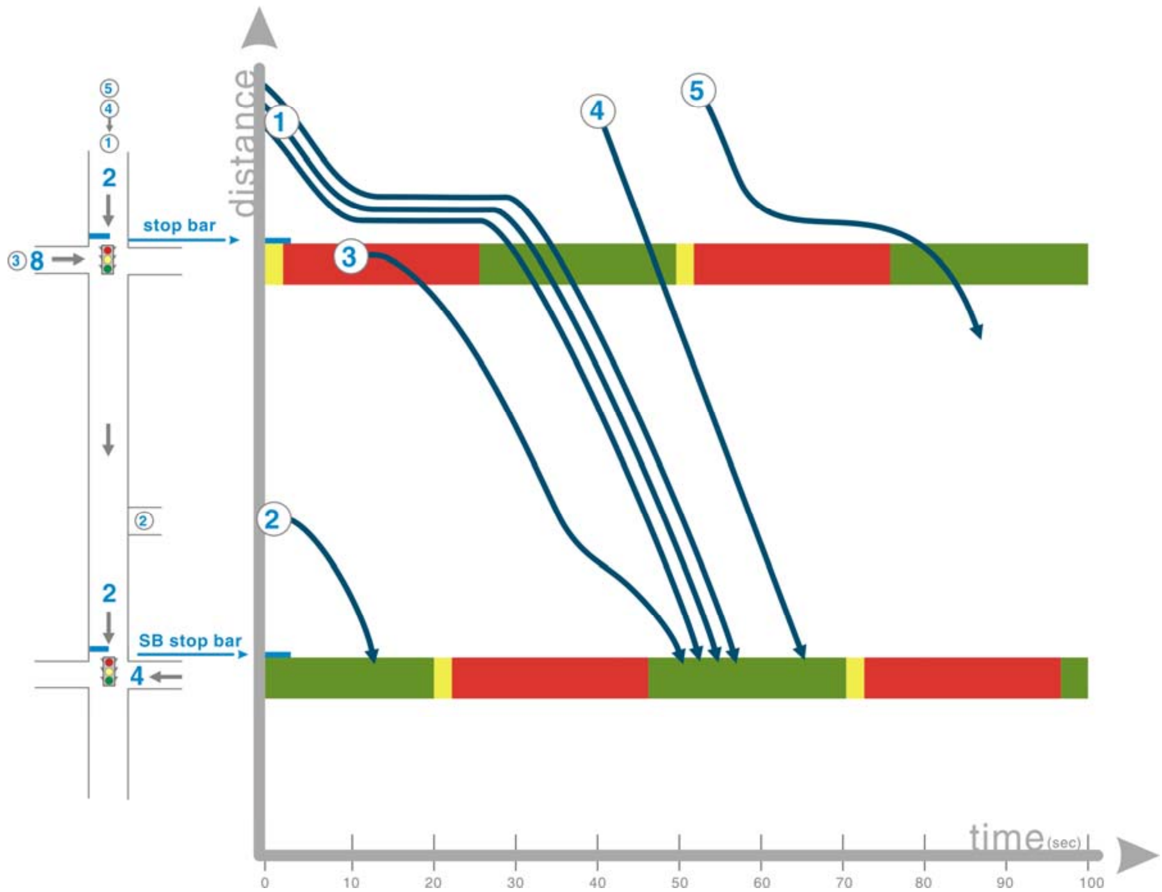



Vehicles on the time-space diagram are shown as trajectories between intersections. Vehicles that turn off of the arterial are considered separately from the through-through vehicles depending on the analysis. The vertical distance between two trajectories is the headway between vehicles.

Figure 6-10 shows more detail related to the range of possible trajectories for vehicles at the two intersection network described. The representation of the one-way street is continued for simplicity. The vehicle trajectories in this figure labeled 1 through 5 are described below:

1. Vehicle travel (as in Figure 6-10);
2. Vehicle from mid-block traveling through the downstream intersection;
3. Vehicle from side street traveling to the downstream intersection, note the vehicle enters the arterial link during the red of the upstream intersection because it is on the side street turning left;
4. Vehicle traveling at the progression speed through the intersections; and
5. Vehicle delayed at the upstream intersection.

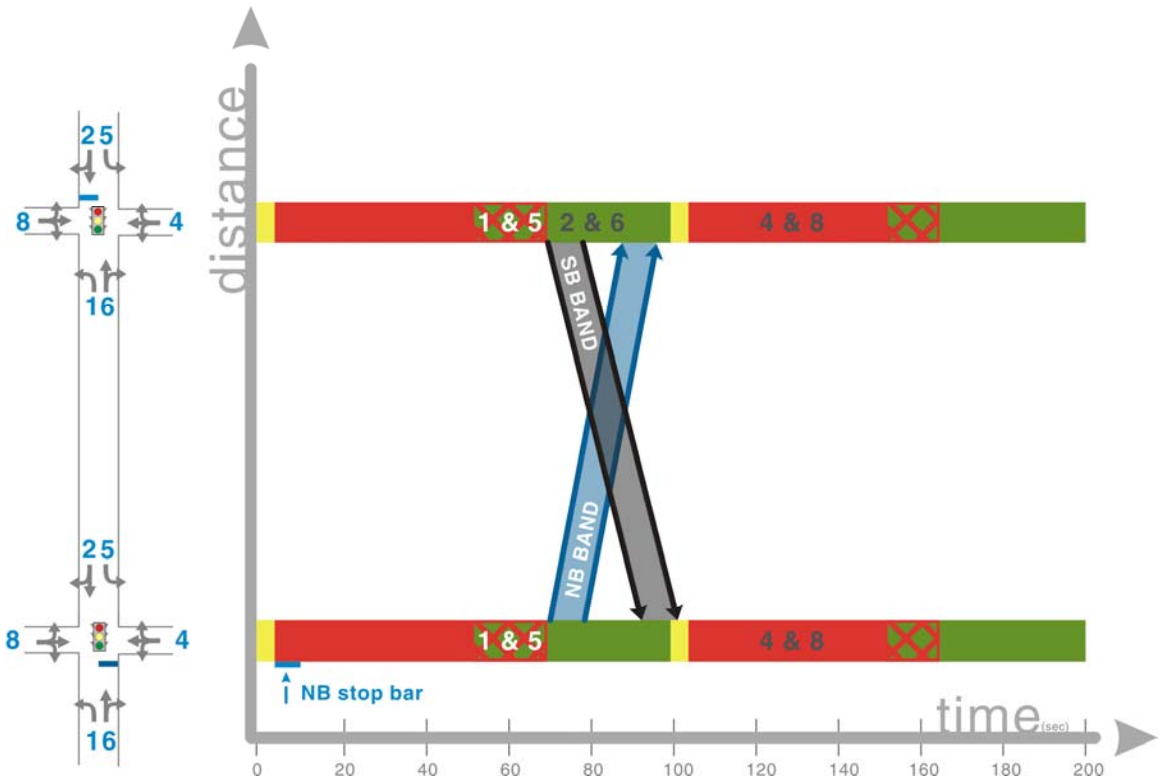
**Figure 6-10Time-Space Diagram – One-Way Street Operation**




The time-space diagram illustrates the signal phasing for the coordinated phases for each of the signalized intersections. Each signal operates with two phases, with phase 2 as the coordinated phase.

Protected left-turn phases and two-way operation on the arterial street complicates the time-space diagram slightly. Figure 6-11 shows the additional phases and two-way operation on the arterial. A 100-second cycle length is assumed. Bandwidth is shown as the shaded area between the intersections.

**Figure 6-11Time-Space Diagram – Two-way Street Operation**




As illustrated in Figure 6-11, the left-turn phases result in less time for the arterial phase. Different types of phasing impacts the arterial street in various ways. Accommodating left-turning vehicles at signalized intersections is a balance between intersection safety, capacity, and signal delay.

Several traffic flow assumptions are used with time-space diagrams. As stated earlier, time-space diagrams typically consider the through movements on the street in deference to the turning traffic and other modes on the roadway. The time-space diagram and estimation of traffic flow are complicated by the interactions between pedestrians and turning traffic, vehicular interactions at midblock driveways, impedance from shared traffic lanes, and other users of the facility. Careful consideration of these conditions must be taken into account when using the time space diagram.

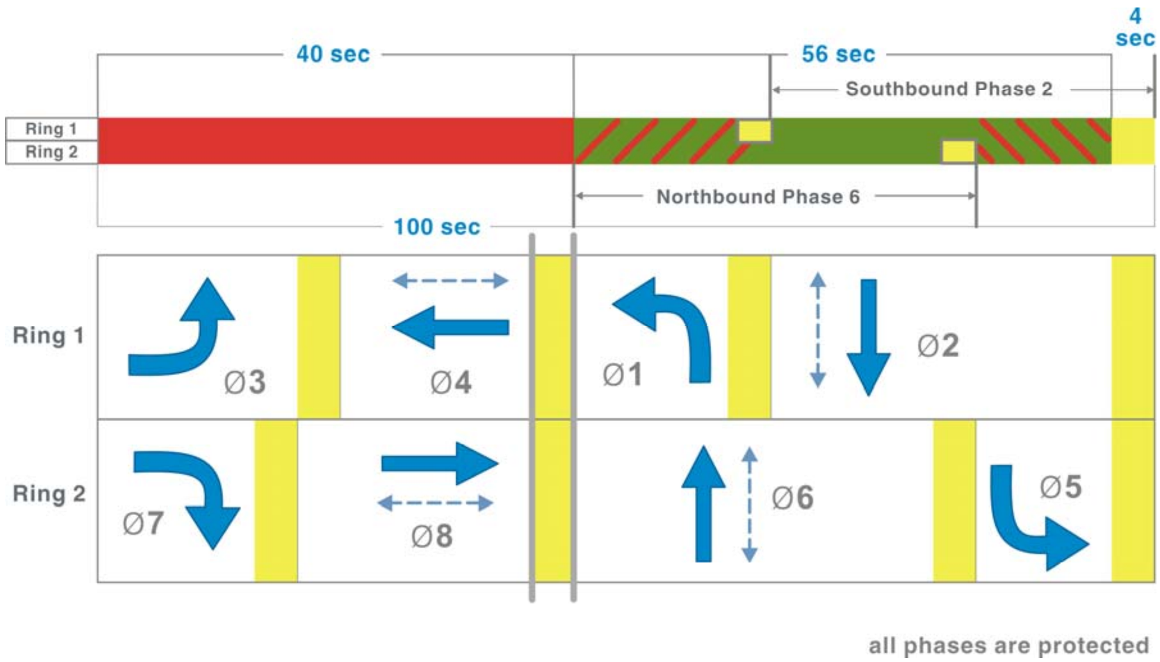
### 6.4.2 Left-Turn Phasing

As a phase times, the utilization of that green depends on the demand in close proximity or approaching the stop bar. In cases where traffic demand has not arrived, delaying the through movement may be beneficial. A corollary to this is that the traffic platoon from the upstream intersection may reach the downstream intersection too early. In either case, lagging the left turn may be beneficial to improve progression or make more efficient use of the green time.

The time-space diagram focuses on the arterial through movements (typically the coordinated phases). There are times when there are other concurrent movements occurring with the coordinated phase, such as the left turn movements. Within a time-space diagram, the left-turn movements are represented by hatching that is in the same direction as the coordinated phase. (In previous figures, phase 2 is a northbound movement and phase 6 is a southbound movement.)

There are situations when other movements are assigned to the coordinated phase, such as the left turn movements. This may be done within an interchange when a left-turn movement is particularly heavy or needs additional time.

**Figure 6-12 Sequence of Left Turn Phasing as Shown in a Time-Space Diagram**

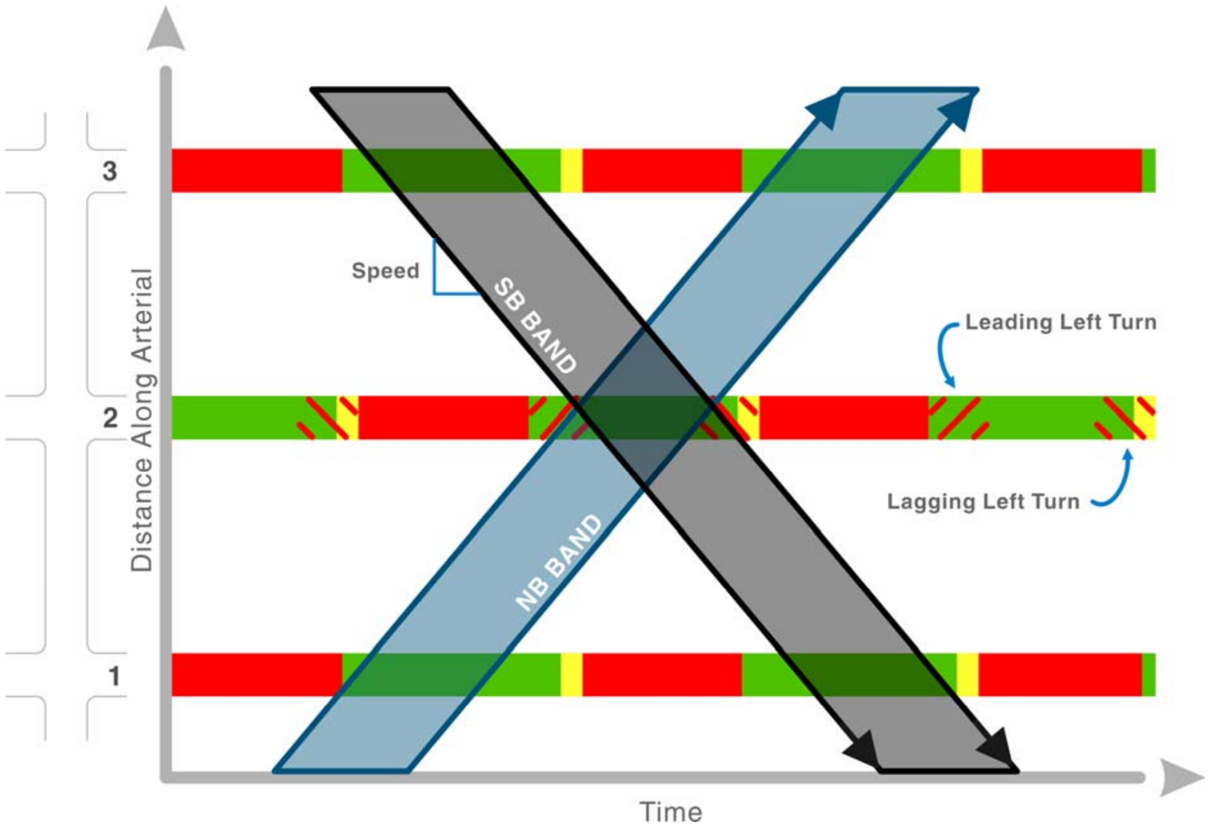



As shown in Figure 6-12, the hatching for phase 5 is in the direction of phase 2 vehicle trajectories and the hatching for phase 1 is in the direction of phase 6 vehicle trajectories. When the left-turn movements occur at the same time (leading or lagging left turns), the hatching crisscrosses to show a period of time where through movements are not possible in either direction.

**Lagging one or both of the left turns along an arterial to promote progression is common. Altering the order of the phases in the sequence may improve the use of the green provided, i.e. vehicles may arrive on green for their phase at the right time.**

Lagging one of the left turns separates the start of the through phase from the start of the left-turn phase, which is particularly useful when upstream intersections from either direction are not equally spaced or have different offsets. This is demonstrated in Figure 6-13 below. Further discussion and examples are provided in Section 6.5.

**Figure 6-13Time-Space Diagram Example of Benefits of Lagging Left Turns**




As shown in Figure 6-13, the bandwidth is increased (as compared to Figure 6-11) by lagging the left turns for subsequent intersections. The benefits of the lead-lag left-turn phasing are further enhanced with protected/permissive lead-lag phasing. By allowing vehicles to turn left during the permissive interval, required left-turn green phase time is reduced, which allows more green time for the coordinated movements. This technique is especially effective for coordinated arterial signals where the progressed platoons in each direction do not pass through the signal at exactly the same time. Al Grover recently completed a comprehensive evaluation of the impacts of protected/permissive phasing associated with coordinated signal timing in western San Bernardino County, California (4). Grover documented a 30- to 50-percent reduction in vehicle delay when comparing protected-only to protected/permissive left-turn phasing.

The Arizona Insurance information association studied lagging left-turn operation in 2002.Tucson, AZ, uses a lagging left-turn phase operation throughout the city. A particular benefit of lagging the left turn is that with PPLT control, the drivers have an opportunity to find a gap, but are providing an opportunity to clear the lane at the conclusion of the permitted phase. The Arizona Insurance information association studied this operation in 2002[5]. The study found that Tucson, AZ had lower crash rates than the leading left-turn operations in the Phoenix, AZ. Area, and this benefit was attributed in part to the use of lagging-left phases. The City of Tucson uses lagging left-turn phase operation throughout the city.

Application of lagging left turns in conjunction with protected-permitted phasing presents the yellow trap to motorists. Understanding of yellow-trap issues is necessary before implementing lead-lag phasing (6). The “Yellow Trap” is a condition that leads the left-turning driver into the intersection when it is possibly unsafe to do so even though the signal displays are correct. During a signal indication change from permissive movements in both directions to a lagging protected movement in one direction, a yellow trap is presented to the left turning driver whose permissive left-turn phase is terminating. Use of the flashing yellow arrow as described in Chapter 4 alleviates the concerns associated with the traditional use of the protected-permitted in the doghouse display.

### 6.4.3 Bandwidth

Bandwidth is described as the amount of time available for vehicles to travel through a system at a determined progression speed. This is an outcome of the signal timing that is determined by the offsets between intersections and the allotted green time for the coordinated phase at each intersection. The bandwidth is calculated by the difference between the first and last vehicle trajectory that can travel at the progression speed without impedance. Bandwidth is a parameter that is commonly used to describe capacity or maximized vehicle throughput, but in reality it is only a measure of progression opportunities. Bandwidth is independent of traffic flows and travel paths and for that reason it may not necessarily be used by travelers. In other words, on an arterial with 10 signalized intersections, a bandwidth solution would be established to allow vehicles to travel through the entire system. **In reality, one must consider how many vehicles desire to travel through all intersections without stopping.**

A few important points to understand related to bandwidth:

- Bandwidth is different for each direction of travel on the arterial and dependent on the assumed speed on the time-space diagram:
- As additional intersections are added to the system, it is increasingly difficult to achieve and measure the impact of an additional signal:
- During periods of oversaturated conditions, bandwidth solutions may result in poor performance, often simultaneous offsets are more effective: and
- Timing plans that seek the greatest bandwidth increase network delay and fuel consumption.

## 6.5 TRANSITION LOGIC

Transition is the process of either entering into a coordinated timing plan or changing between two plans. Transition may also be necessary after an event such as preemption or loss of coordination due to a pedestrian crossing. In general, traffic signals do not operate within the same pattern parameters and cycle lengths at all times. The pattern may change during the day due to a number of reasons:

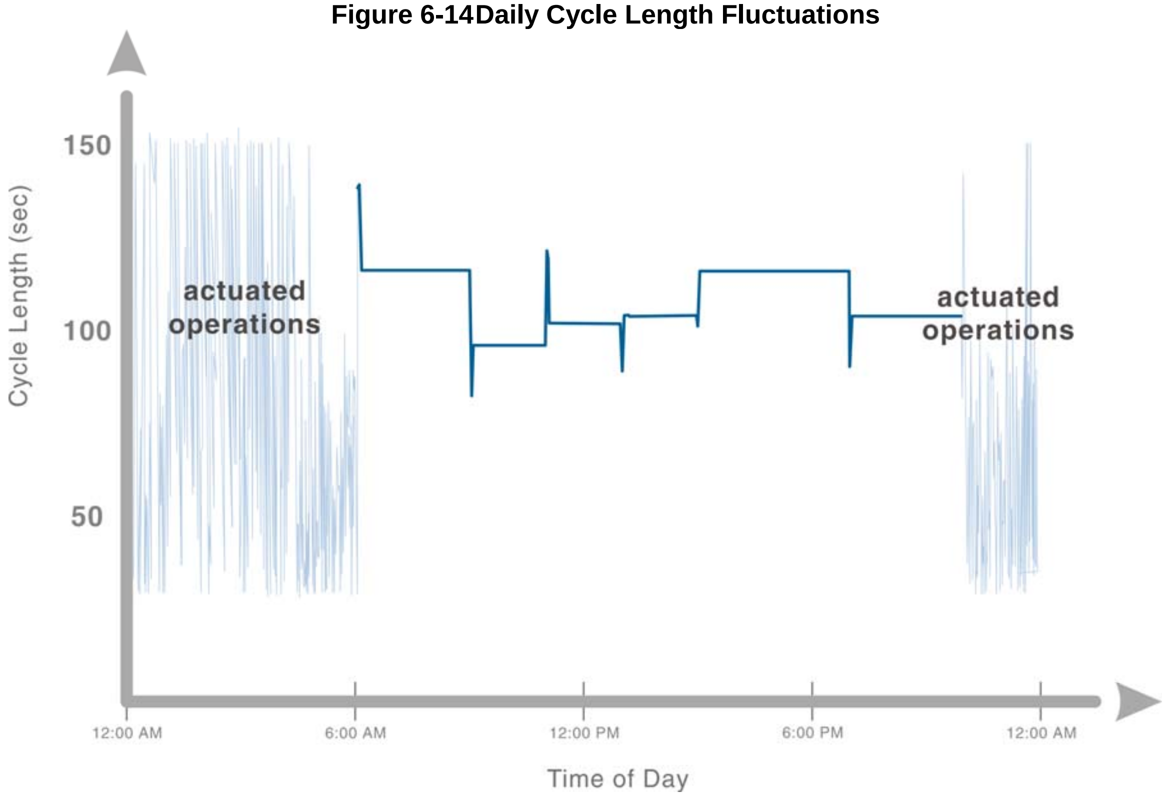
- Time-of-day scheduled changes
- Manual operator selection
- Traffic-responsive pattern selection
- Emergency vehicle, rail-road, or other preemption
- Adaptive control system pattern selection
- Corrections to controller clock
- Pedestrian demand
- Power loss and restoration

The time-of-day schedule determines what time a plan will be active. The simplest schedules typically define an a.m., off-peak, and p.m. peak for weekdays and a different set of plans for weekends. However, all controllers have extensive scheduling options that allow users to define several dozen plans that can be activated by individual day of week, month, pre-defined holidays, or major events.

When the controller clock reaches a point where it is necessary to change the coordination plan, the cycle, split, and offset are changed. If just the splits are changed, transition is trivial because the controller simply starts using the new splits. However, if either the offset or cycle change, the controller must shift the local offset reference point. This requires the use of an algorithm that may either shorten or lengthen the cycle to make that shift. That transition algorithm typically operates for one to five cycles, depending upon the transition mode selected and how much the cycle needs to be shifted. Consequently, the split durations during the period of transition may be different from either the previously defined splits or the new ones.

For example in the example cycle length plot shown in Figure 6-14, one can see the system runs with a fixed background cycle from 6:00 AM to 10:00 PM, with cycle changes at 9:00 AM, 11:00 AM, 1:00 PM, 3:00 PM, and 7:00 PM. During each of these plan changes, the controller goes into transition, resulting in variable cycle lengths for a couple of cycles to adjust to the new plan.

**Figure 6-14Daily Cycle Length Fluctuations**




### 6.5.1 Example Application of Time Based Coordination Transition

Traffic signal coordination requires adjacent signals to operate at the same cycle length or at a multiple of the cycle length with pre-determined offset and coordination points. The most common method for achieving that specified offset is called time-base coordination, where coordinated signals are configured to use the same sync reference time, such as midnight or 3:00 a.m. When the signal is operating in a coordinated mode, it can calculate when the current cycle should begin by effectively counting forward from that sync reference time. The definition of the "start of cycle" depends on the offset reference point used by that signal, such as the start of green for a coordinated phase. The transition is initiated to re-align the local zero point (when the cycle begins) with the system sync reference time (when a timing plan is initiated).

As an example of time based coordination, consider a signal that is operating a coordination pattern that calls for a 90-second cycle and a 20-second offset, with 12:00 a.m. as the sync reference time. From this information, we know that this signal should begin a 90-second cycle at 20 seconds after 3:00 a.m. (3:00:20 a.m.) and every 90 seconds thereafter. An adjacent signal operating a pattern with the same cycle length but an offset of 45 seconds should begin a 90 second cycle at 45 seconds after 3:00 AM (3:00:45) and every 90 seconds thereafter. Hence these two signals will always have the same relative offset of 25 seconds (the difference between the two absolute offsets of 20 seconds and 45 seconds).

When a signal controller begins to operate a new coordination pattern, it must establish the cycle length and offset of that pattern. The same applies when re-establishing an offset after a cycle is disrupted by preemption or a pedestrian time that exceeds the split time. By calculating back to the sync reference time (12:00 a.m. in the above example), the controller can determine when the offset reference point is scheduled to occur within the current cycle. In general, the next start-of-cycle time will be several seconds earlier or a few seconds later than the current cycle start time. The larger the difference between the two cycles and/or the offset values of the two patterns, the longer the transition will take. While central or master-based systems typically communicate the selection of a new timing plan to all signals in a group at the same time, the actual transition logic is executed independently at each signal, without explicit regard for the state of adjacent signals. Most controllers allow three or four transition modes, which govern the precise details of how the signal resynchronizes to the new cycle and offset. The transition modes differ significantly from one controller manufacturer to the next. Some vendors may also refer to transition as offset seeking, offset correction, or coordination correction (7). The next section gives a brief overview of signal timing during transition for the most commonly available transition modes. However, no matter which mode is selected, traffic control can be significantly less efficient during the transition between timing plans than it was during coordination.

### 6.5.2 Transition Modes

There are three basic techniques for achieving an offset transition: dwell, lengthen, and shorten. The first technique is to stay (dwell) in the coordinated phase(s) until the new offset is achieved. That is, the current cycle is increased in length as needed, with all the additional time being assigned to "dwelling" or coordinated phases. The second technique also expands or adds to the cycle length as needed, but distributes that additional time between all phases. The third technique shortens the cycle length, taking time from all phases to the extent allowed by their minimum green settings and any pedestrian activity during the phase.

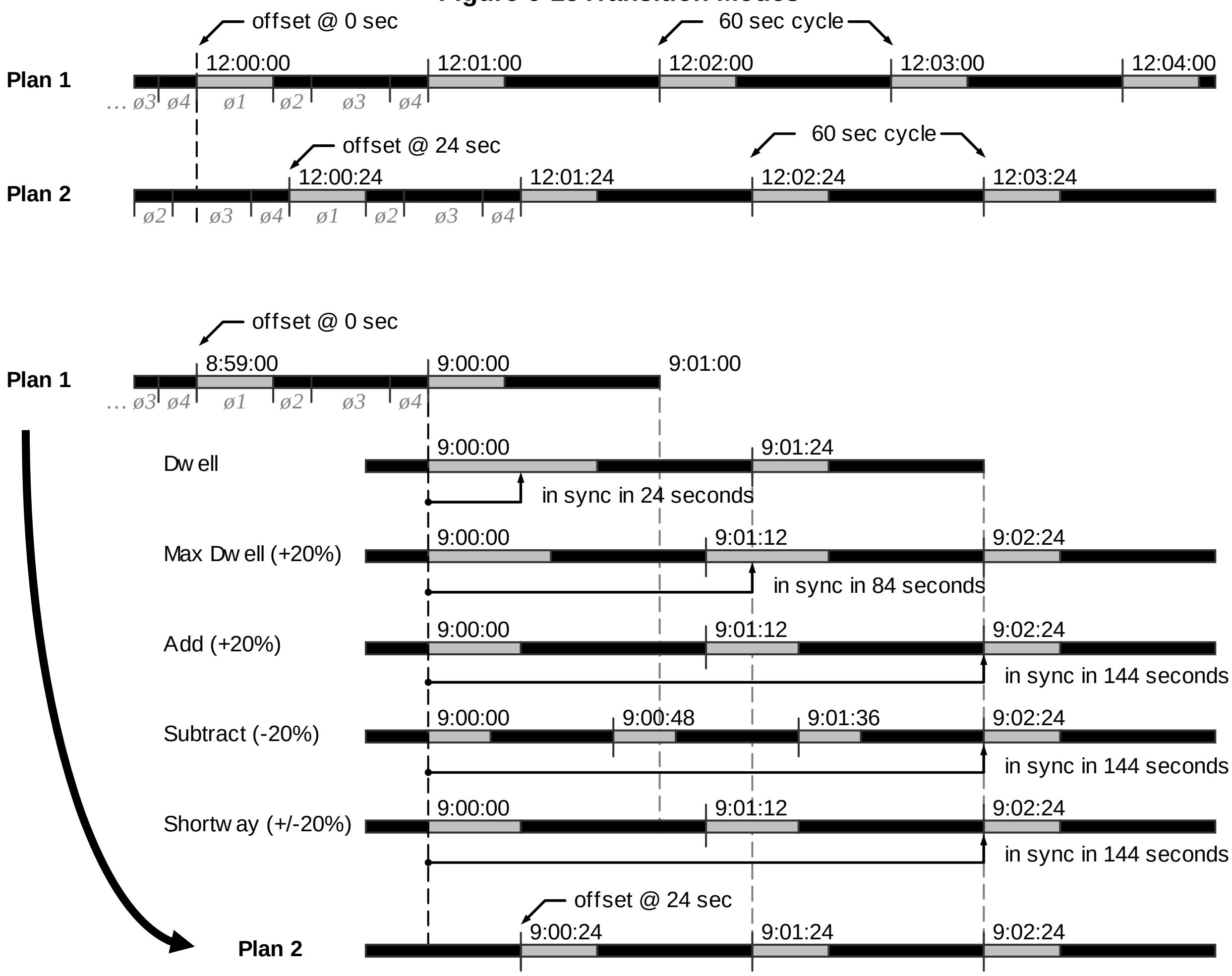
To avoid an excessive cycle length or phase greens that are too short during transitioning, it is common for controllers to limit the maximum amount of adjustment that can be made in one cycle. If such a limit is imposed, and it usually is, the signal may not be able to complete a given offset transition within one cycle. Signal controllers also commonly compute their adjustments so that transition is completed in a set number of cycles for the worst case scenario, typically a maximum of three to five cycles.

The most common transition modes in signal controllers in the United States include Dwell, Max Dwell, Add, Subtract, and Shortway, of which at least two or more modes are offered. (*8*) Figure 6-15 portrays two timing plans where the gray sections indicate when the coordinated phase is green and black sections indicate when the coordinated phase is red. Switching from Plan 1 to Plan 2 entails (for simplicity's sake) no change to the cycle length, but a 24-second shift in the offset. Figure 6-15 also shows a separate timeline for each of five types of transition between Plan 1 and Plan 2. The descriptions of these modes discuss the example in Figure 6-15:

- **Dwell** – At the next display of green in the coordinated phase, the controller begins transition by holding (or dwelling) in this state until the new local zero point is achieved, at which time the signal is considered in sync and begins the new timing plan. As shown in Figure 6-15, this adjustment appears as one prolonged cycle. In general, this transition mode may result in an uneven allocation of split times during transition. It can be a very helpful mode to use when troubleshooting a coordinated system because it is very deterministic.
- **Max Dwell** – This modified version of Dwell also adjusts the start of the cycle by extending the green time of the coordinated phase. However, only a limited amount of extra green time may be added to each cycle. The example in Figure 6-15 is constrained to add no more than a certain percentage (20% is used in the example) to the cycle length, thus two cycles are required to achieve coordination.

- **Add** – This mode synchronizes by shifting the start of the cycle progressively later, by timing slightly longer than programmed cycle lengths. The Add mode increases the green time of all phases in the sequence, whereas the Dwell modes add time only to the coordinated phase(s). This is illustrated in Figure 6-15, where Max Dwell and Add both increase cycle lengths by a certain percentage (20% again is used), but the allocation of extra time to the coordinated phase is less in Add mode because extra green time is distributed proportionally amongst all phases. If a signal is subject to preemption, selecting the "Add-only" transition prevents splits on the phases omitted during preemption from being reduced during the transition period. The only real disadvantage of this mode is that if a cycle needs to shift one second backwards, it must shift the entire cycle forward one second less than the cycle length. This results in longer cycles during the transition period that could potentially cause unexpected storage problems in left turn lanes or between closely spaced intersections.
- **Subtract** – This mode shifts the start of the cycle progressively earlier, subtracting time from one or more phases in the sequence (subject to their minimum green time requirements). As shown in Figure 6-15, Subtract mode decreases all phases in the cycle by a certain percentage (20% is assumed) during transition. Though Subtract mode takes three cycles in Figure 6-15, the total transition time of three short cycles is equivalent to the time taken by two long cycles during Add transition.
- **Shortway** – This mode (also sometimes called Smooth) resynchronizes by applying either Add or Subtract transition logic, choosing the mode that presents the “shortest path” to achieving sync. The specific details of this mode, such as its name and the exact logic for determining the “shortest path”, can vary significantly from one controller vendor to the next. In general this is the default mode on most controllers and is typically appropriate to use, unless the signal is subject to preemption.

**Figure 6-15Transition Modes**




.

### 6.5.3 Operational Guidelines

The example shows the effect transition has at a signal. Adjacent signals that change patterns simultaneously may take different amounts of time to complete their offset transitions, depending on which method is used and the size of the change required. In addition to disrupting the progression of vehicles between signals, offset transitioning artificially lengthens or shortens phases, leading to inefficient use of green time or unusual queuing. Therefore, it is best to avoid changing patterns during congested conditions when the signals need to operate at maximum efficiency. A peak-period pattern is best implemented early to ensure all offset transitioning is completed before the onset of peak traffic flows. Similarly, it is desirable to avoid frequent pattern changes.

When using the subtract transition technique, the new cycle length may be close to the sum of the minimum phase green times (or pedestrian times), meaning only a small adjustment in cycle length can be made by shortening. In this case, it may take many cycles to complete an offset transition. Hence, it is not practical to require the controller to use the subtract technique exclusively. To avoid this problem and allow use of the subtract technique when it works well, most controllers offer some version of the Shortway method. If the user selects this option, the controller will investigate both add and subtract techniques and automatically choose the one that will complete the offset transition (get the signal "in step" or "in sync") most quickly. Field experience and laboratory experiments have shown that Shortway typically provides the least-disruptive effects on traffic than any other methods, except in very rare cases. Many studies have also shown that excessive changes to timing plans, in an attempt to match traffic patterns closely and improve performance, can be a detriment because the system never achieves coordination for more than a few minutes at a time. Because of this, it is generally recommended to remain in a coordinated pattern for at least 30 minutes.

**In locations with low pedestrian demand, it may be desirable to allow coordination to be lost when a pedestrian requests service. Care should be taken in using this mode of operation if vehicle progression performance is valued; a few pedestrians may cause one or more signals in a coordinated system to be in transition almost all of the time.**

## 6.6 COORDINATION TIMING PLAN GUIDELINES

This section provides guidelines for selecting coordination settings. The information provided is based on established practices and techniques and some acknowledgement that additional research is necessary on this subject. The guidelines address the following topics:

- Coordinated Phase Assignment
- Cycle Length Selection
- Split Distribution
- Offset Optimization

Each of these topics is addressed separately in the remainder of this section.

### 6.6.1 Coordinated Phase Assignment

At each intersection, a coordinated phase is designated to maintain the relationship between intersections. Common practice is to designate the main street through phase as the coordinated phase because the coordination logic in most controllers dwells in this phase, and additional time provided to what is normally the busiest movement results in better performance. In many cases this is phases 2 and 6 for the main street through phases.

With some applications, assigning a phase other than the major street through movement has proven effective. A diamond interchange where queue management strategies are desirable is one example where a traffic signal may operate more effectively with such a designation.

### 6.6.2 Cycle Length Selection

Cycle length selection should reflect local objectives and users of the system. Theoretically, shorter cycle lengths result in lower delays to potential users. As we consider more users, the cycle length may increase and tradeoffs are made between competing objectives. Cycle length selection may be completed independently or by considering a system of intersections, but in most cases a first step is to assess each intersection for its minimum cycle length.

Each intersection is assessed for its cycle length requirement as a part of the selection process. The result is typically different “optimal” cycle lengths at the intersections. As these intersections are aggregated into systems, decisions have to be made regarding whether to include the intersection in a systems and which system to include the intersection in. There are cases where intersections are included in a coordination strategy which results in longer cycle lengths for the other intersections in the system. One should always consider whether intersections with long cycle length requirements would better operate independently of the system.

**Cycle length selection is typically based on traffic data that is collected during representative periods. In reality, there is a wide variety of volumes that occur throughout the operation of the timing strategy. Automated data collection would improve the data and ultimately the cycle length selection, particularly at sites such as shopping districts that experience heavy weekend traffic that often is problematic for agencies to collect.**

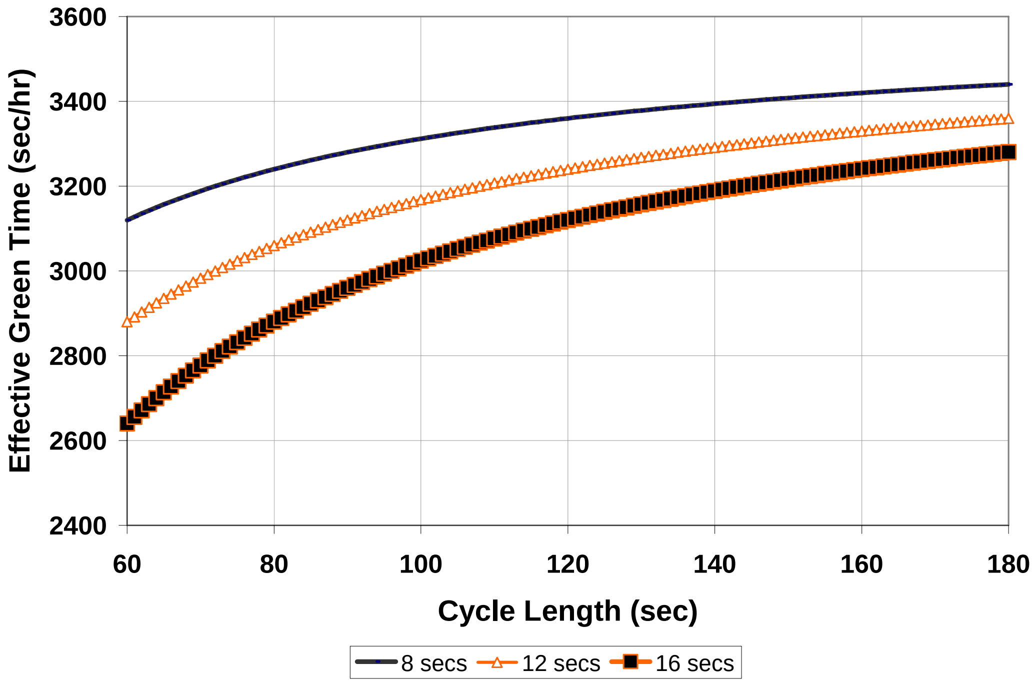
Similarly, because pedestrian timing may influence cycle length, various timing strategies are used to ensure effective coordination (8). The effect of pedestrians is further described in the next section. A key decision is whether pedestrians will be accommodated within the coordinated cycle length. In cases where pedestrian volumes (and the resulting actuations) are low, a pedestrian call may not be accommodated within the cycle length which will result in temporary disruption to the coordination timing.

The selection of a cycle length affects intersection capacity and delays. Longer cycle lengths can increase capacity, but only marginally. Shorter cycle lengths usually result in reduced delays. Thus, the objective of choosing a cycle length is to determine the smallest value of cycle length that provides the desired level of vehicular capacity at the intersection while being appropriate for the needs of other users such as pedestrians.

The selection of cycle length should also be influenced by the desired progression speed for a roadway. This is a complicated geometric relationship (to be discussed later) that must be recognized when selecting cycle length. Cycle lengths that are too long may increase congestion rather than reducing it due to the impacts of long waiting queues on side streets and the arterial alike. Cycle lengths also result in establishing relationships between the intersections and in some cases some values work better than others due to the time space relationship between the intersections.

In general, it is preferred that the cycle lengths for conventional, four-legged intersections not exceed 120 seconds, although larger intersections may require longer cycle lengths(9). Theoretically, intersection capacity increases as the cycle length becomes longer because a smaller portion of the time is associated with lost time. However, it is important to recognize that the improvements are modest and this assumes that all lanes are operating with saturated flows. This requires that turn bays are long enough to provide sufficient demand to a particular movement; auxiliary lanes are more likely to be blocked with longer cycle lengths. As shown in the figure, the change of cycle length from two minutes (120 sec.) to three minutes (180 sec.) results in a modest 2 percent increase in capacity. This calculation was made by estimating the start-up delay resulting from the signal changes, as it relates to the number of times that the signal intervals change during the course of an hour. The message conveyed by the information presented in Figure 6-16 is that one should avoid placing too much emphasis on longer cycle lengths as a panacea for congested conditions.

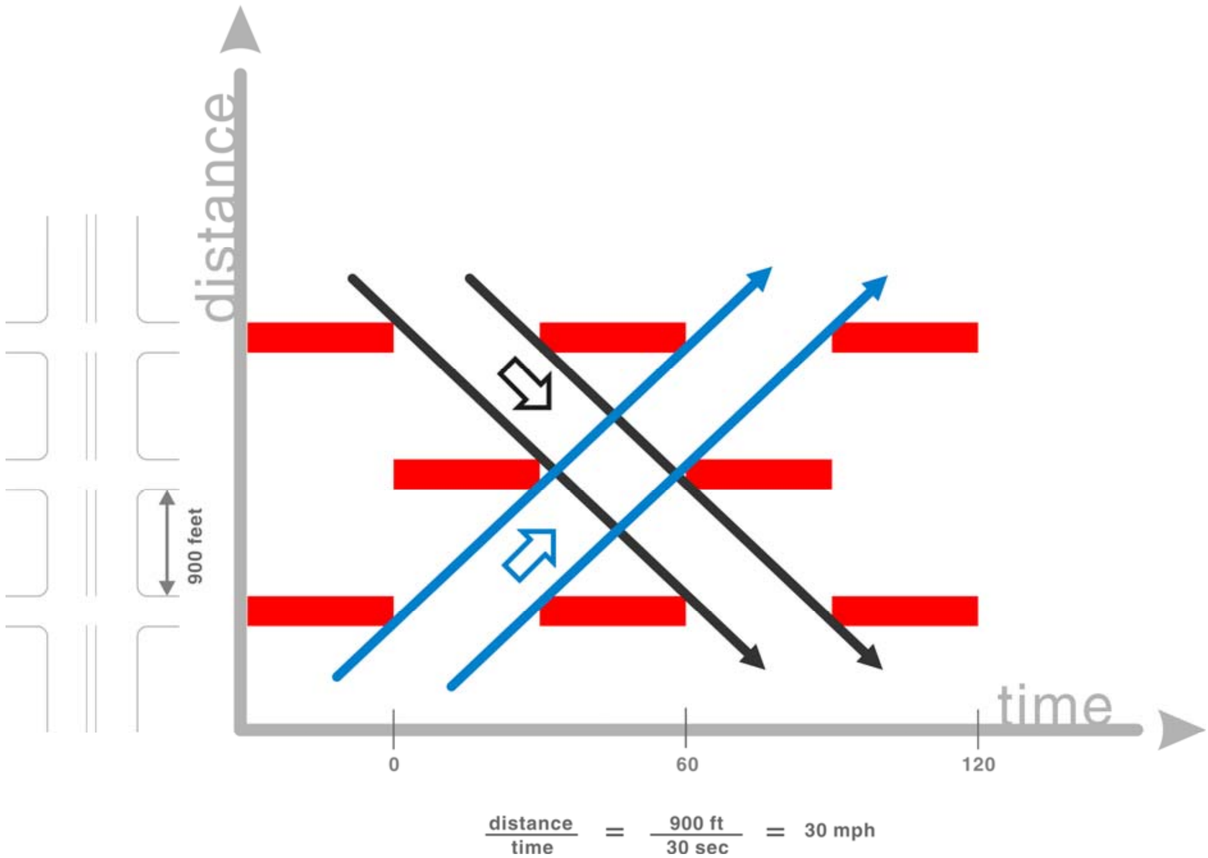
**Figure 6-16Cycle Length & Theoretical Capacity**




***Manual Methods***

A manual methodology for determining cycle lengths in a traffic signal network was first developed in Los Angeles. The City used a traffic signal timing strategy that was elegantly simple for its robust grid system. Most of the signals in the City had permitted left turns with signals at 1/4-mile spacing. A 60-second cycle length was selected and the green time was equally distributed to the cross street traffic and the arterial. Given these timings and spacing, a vehicle traveling at 30 miles per hour would reach the next 1/4-mile signal in 30 seconds, or half of a cycle length. The result is in an “alternating” system of offsets between subsequent intersections as shown in Figure 6-17. This system permitted two-way coordination along corridors. Under this configuration, the cycle length must be carefully selected with regards to the signal spacing and the anticipated or desired speed in the system. Although the optimization models used today can provide additional refinement to the coordination plans developed using this manual method, this method of alternating offsets serves as a reliable default timing strategy.

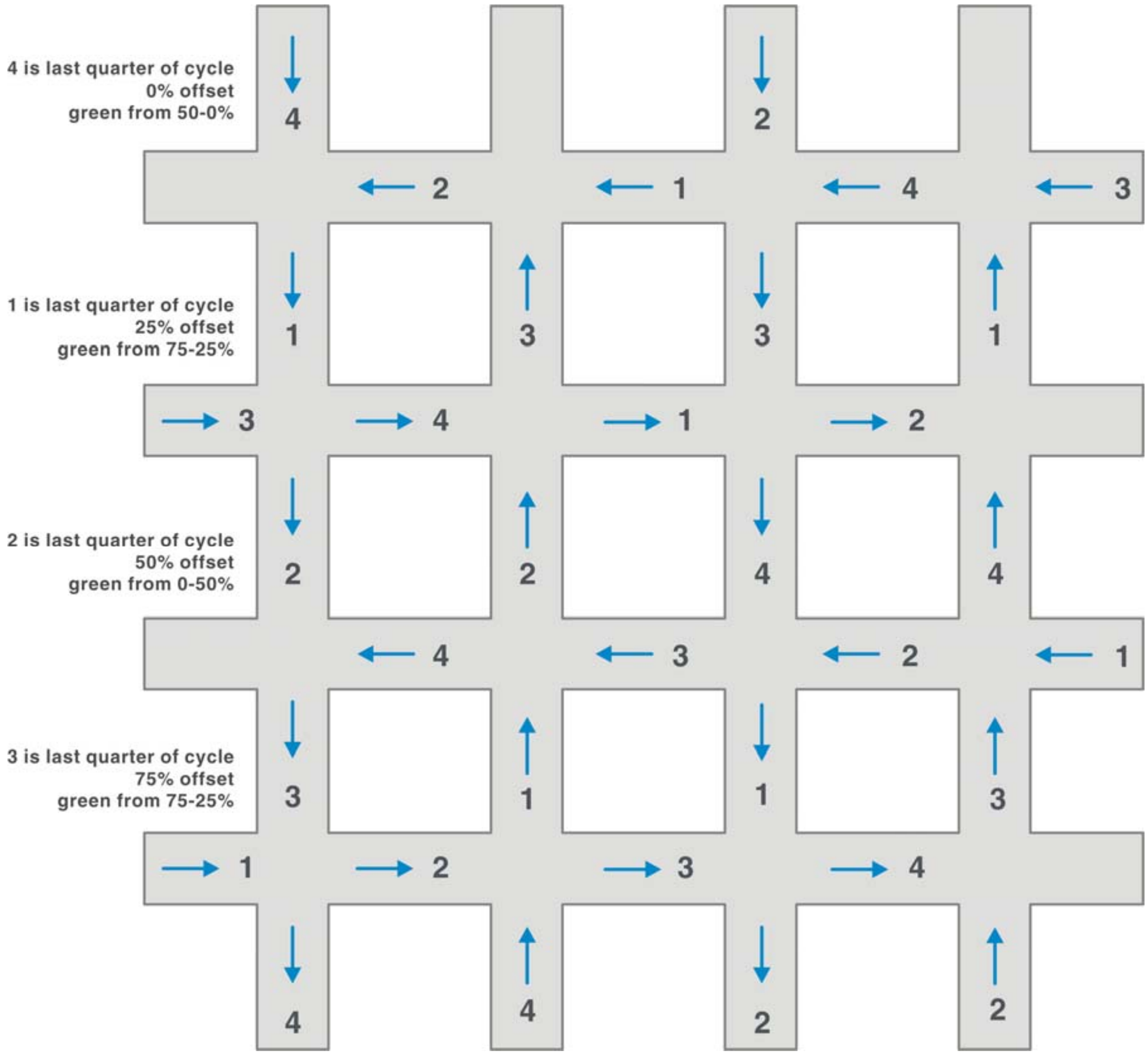
**Figure 6-17 Alternating Offsets System of Intersections**




Although the simulation and optimization models used today can refine the ways in which signals are timed, the alternating offset methods of yesteryear still work well as reliable default timing strategies10.

A similar method of manual coordination timing can be applied to downtown grid networks. This method has been deployed in downtown Portland, Oregon by separating intersections into a quarter cycle offset pattern. The block spacing in downtown Portland is fairly uniform and relatively short (280 feet) and the grid is a one-way network. Each subsequent intersection is offset by a quarter of the cycle length, which is selected to progress traffic in both directions. The result is a progression speed that is dependent upon the cycle length. This approach establishes a relationship in both directions of the grid and permits progression between each intersection in each direction based on the speed that is a result of the selected cycle length and the block spacing. As shown in Figure 6-18 cross-coordination throughout the grid is achieved using the quarter cycle offset method. This approach can be adjusted to account for turning movements within the grid and subtle modifications to the distribution of green time.

**Figure 6-18 Quarter Cycle Offset Example Model**




Some agencies have established fixed cycle lengths for various types of streets based on intersection spacing, signal phasing, travel speeds, and pedestrian crossing requirements. As an example, Harris County, Texas, uses cycle lengths shown in Table 6-1 below (11).

**Table 6-1 Harmonic cycle lengths based on street classification in Harris County, Texas.**

| Minor Arterial (seconds) | Major Arterial (seconds) |
|---|---|
| 60 | 90 |
| 70 | 105 |
| 80 | 120 |
| 90 | 135 |

This ensures a consistent approach is taken throughout the system for working between individual signals and intersecting corridors. This approach also suggests that the absolute value of the cycle length is less important than maintaining a relationship between adjacent systems of traffic signals.

### *Critical Intersection Methods (Webster, HCM)*

Much of early research regarding cycle length selection recommended evaluating the intersections identifying the critical intersection which was typically the intersections with the highest demand. A cycle length is established for this location and is selected so that it will be sufficient to maintain undersaturated conditions. This fundamental assumption is currently being investigated to determine if there are further strategies for dealing with oversaturated conditions.

The critical intersection approach considers a signalized intersection in isolation to other intersections and for this reason may not always yield the optimum cycle length. Most of the analytical tools developed for cycle length selection focus on undersaturated flow (12). The tools also do not consider the constraints of the intersections beyond the lost time and saturation flow rate. These critical intersection approaches to cycle length selection are primarily for isolated intersections and are all based on the assumption that vehicular delay is most important. This approach analyzes the intersection with the heaviest traffic to determine a minimum cycle length and used that to set the remaining intersections. The first step is to consider each intersection as though it is isolated to determine the minimum (optimum) cycle length needed at each intersection, as though it were isolated(13). The traditional models use Webster's model to determine optimal cycle length. Webster used computer simulation and field observation to develop a cycle-optimization equation intended to minimize delays when arrivals are random. (14) The formula is as follows:

**(6-1)**

$$C = \frac{1.5L + 5}{1.0 - Y}$$

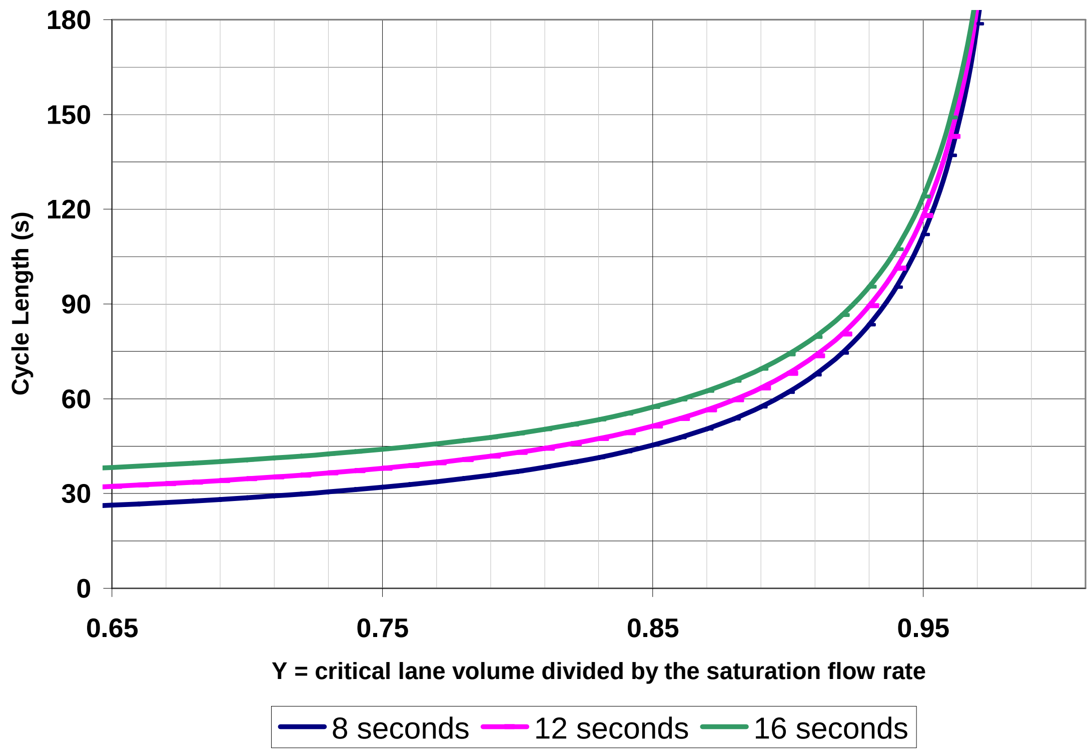
where:

$C$ = optimum cycle length, (s)

$Y$ = critical lane volume divided by the saturation flow, summed over the phases

$L$ = lost time per cycle, (s).

**Figure 6-19Webster's Optimum Cycle Length**




In practice, much of the assessment of signalized intersections is completed using the Highway Capacity Manual (HCM) procedure on "Signalized Intersections." The HCM provides few pieces of guidance on cycle lengths, but also notes limitations to the methodology. The current methodology does not take into account the potential impact of downstream congestion can have on intersection operation. Nor does the methodology detect or adjust for turn-pocket overflows and the impacts they have on through traffic and intersection operation.

The HCM offers a quick estimation method for the selection of a cycle length. The formula for cycle length estimation is as follows:

**(6-2)**

$$C = \frac{L}{1 - \frac{\min(CS, RS)}{RS}}$$

where:

$C$ = cycle length (s),

$L$ = total lost time (s),

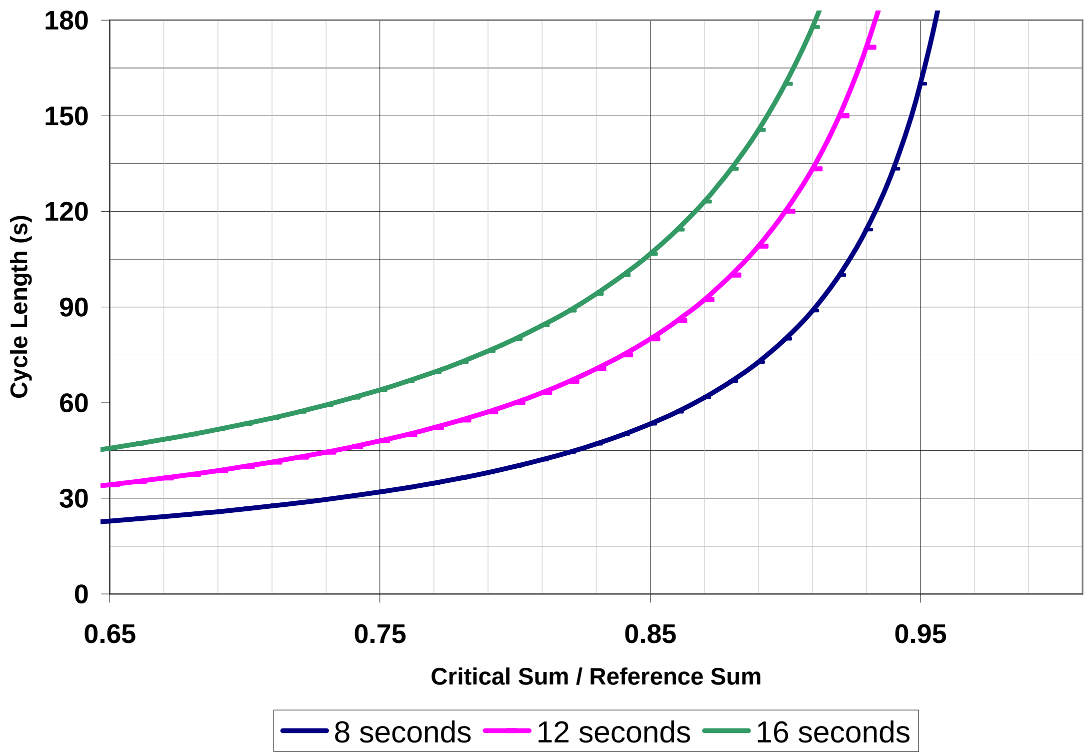
$CS$ = critical sum of traffic volumes from the critical movement analysis (veh/h),

$RS$ = reference sum flow rate = 1,710 · $PHF \cdot f_a$ (veh/h),

$f_a$ = area type adjustment factor (0.90 if CBD, 1.00 otherwise).

Primarily, this calculation is intended for planning-level analyses. This equation suggests that as the intersection approaches capacity, the cycle length should increase up to a maximum value, which the HCM suggests is set by the local jurisdiction (such as 150 seconds). The minimum cycle length suggested for use is 60 seconds. The equation does not explicitly address the pedestrian crossing requirements, left turn type and minimum green times necessary to meet driver expectancies.

**Figure 6-20HCM Cycle Length Estimation**




In both Webster's and the HCM's estimation, the sum of the critical lane flows is a representation of the demand at the intersection. The critical lane is defined as the intersection approach with the greatest demand of all the approaches that are serviced during a given signal phase. For example, during the main street phase, on a street with two-way traffic, the critical lane would be the one lane in either direction that has the greatest demand. $Y$ is the sum of the critical lane flows divided by 1900, which is the percent of available intersection capacity that is in demand. If $Y$ = 1, the intersection is saturated and the equation is no longer applicable.

The situations where longer cycle lengths degrade intersection performance are a result of specific elements that lead to poor performance. Longer cycle lengths will increase congestion in cases such as when:

- Upstream throughput exceeds downstream link capacity. Long cycles may move more vehicles through an intersection than can be handled downstream
- Turning bay storage is exceeded. Long cycle lengths may cause vehicles in left-turn bays to back up into through lanes. In a similar manner, long cycles may cause through traffic to back up beyond turn bays, restricting their access
- Increased variability in actuated green times. Long cycles result in high variability in the side street green time used, which may result in poor arrival types at the downstream intersection. This is particularly noteworthy when split times exceed 50 seconds (15).

The delay experienced by a motorist depends on cycle length and volume. Higher volumes always lead to longer delays. Shorter cycle lengths reduce delay, provided they do not result in inadequate intersection capacity. **Oversaturated conditions require special considerations, and these models are not valid during that range of conditions.**

#### *Network Approaches*

The network approach to cycle length selection considers multiple intersections to determine an optimal cycle length. Most applications of network approaches use signal timing optimization models. There are a number of computer programs that can be used to assist in selection of a cycle length. The Federal Highway Administration's (FHWA's) Traffic Analysis Toolbox describes additional resources (http://www.ops.fhwa.dot.gov/trafficanalysistools/toolbox.htm). Three of the more popular programs of this type are Synchro, PASSER™ II, and TRANSYT-7F (16).

These signal timing optimization models consider the network being analyzed and determine an optimal solution based on a given set of inputs. The range of cycle lengths is based on the users input. The optimization models use the individual intersection characteristics, the volume to capacity ratios of each intersection, the link speed, and the distance between the intersections to estimate the performance for each individual cycle length and resulting plan. The models make assumptions based on the inputs related to the splits and offsets to determine performance measures that can be compared to timing policies. Because the models are imperfect a significant amount of effort is necessary to take an initially screened plan to a point that can be field implemented. The timing policies described previously, the optimization policies, and the criteria for determining which criteria to use to select a signal timing plan must be considered prior to and as a part of the optimization process.

A recent FHWA publication devotes a significant number of pages describing various optimization software packages available (17), so only the pertinent elements will be described here. The optimization models change with new versions of the software and for that reason, the documentation is best handled by the individual software producer. Their guidance related to the development of timing plans is most important.

The PASSER program uses the concepts described in Webster's equation to determine the appropriate cycle length. The program uses a hill climbing algorithm to estimate delay at each intersection with the cycle length input to further quantify the performance of the system and maximize bandwidth using the pre-calculated splits as input to that model. At the optimization stage, it can find the cycle length, offsets, and phase sequences that produce maximum two-way progression. Essentially, PASSER uses the concepts described in Webster's equation for selection of cycle length and, "after calculating the minimum delay cycle length for all intersections in the arterial street, the largest minimum delay cycle length is selected as the shortest cycle length for the system and that the longest allowable cycle length should be no more than 10 to 15 seconds longer than the shortest allowable cycle length to minimize the excess delay at the non-critical intersections." (18)

TRANSYT-7F allows the user to define the performance function used for optimization. TRANSYT-7F was initially designed to select signal timings that produce minimum network delay and stops. Subsequent modifications added the capability to select several other objectives, including minimization of fuel consumption and maximization of progression opportunities. During its optimization process, TRANSYT-7F generates second-by-second flow profiles of vehicles on all links in the network and analyzes these profiles to determine performance measures. This model considers the formation and dissipation of queues in space. In addition, it accounts for flow interactions on adjacent links through a step-by-step analysis of all links in the system. TRANSYT-7F assesses cycle length by calculating equal saturation splits and applies a hill-climbing method to optimize signal offsets and splits.

In similar fashion, Synchro uses its algorithm to estimate arrivals at each intersection in the network and to calculate percentile signal delay, stops, and a queue penalty, which addresses the impact of queuing on arterial performance(19). The performance index is calculated for each cycle length based on the splits and offsets assumed within the model as constrained by the user. There are various steps to developing an "optimal" timing plan, but one of the limitations of this model is the inability to define the parameters within the Performance Index calculation.

$$PI = \frac{D \cdot 1 + St \cdot 10 + QP \cdot 100}{3600} \tag{6-3}$$

where

$PI$ = Performance Index

$D$ = Percentile Signal Delay (s)

$St$ = Vehicle Stops (vph)

$QP$ = Queue Penalty (vehicles affected)

A detailed network analysis using an optimization tool assesses the cycle length for a coordinated system. Computer models facilitate multiple iterations of varying cycle combinations to determine the optimum signal timing parameters. (20)

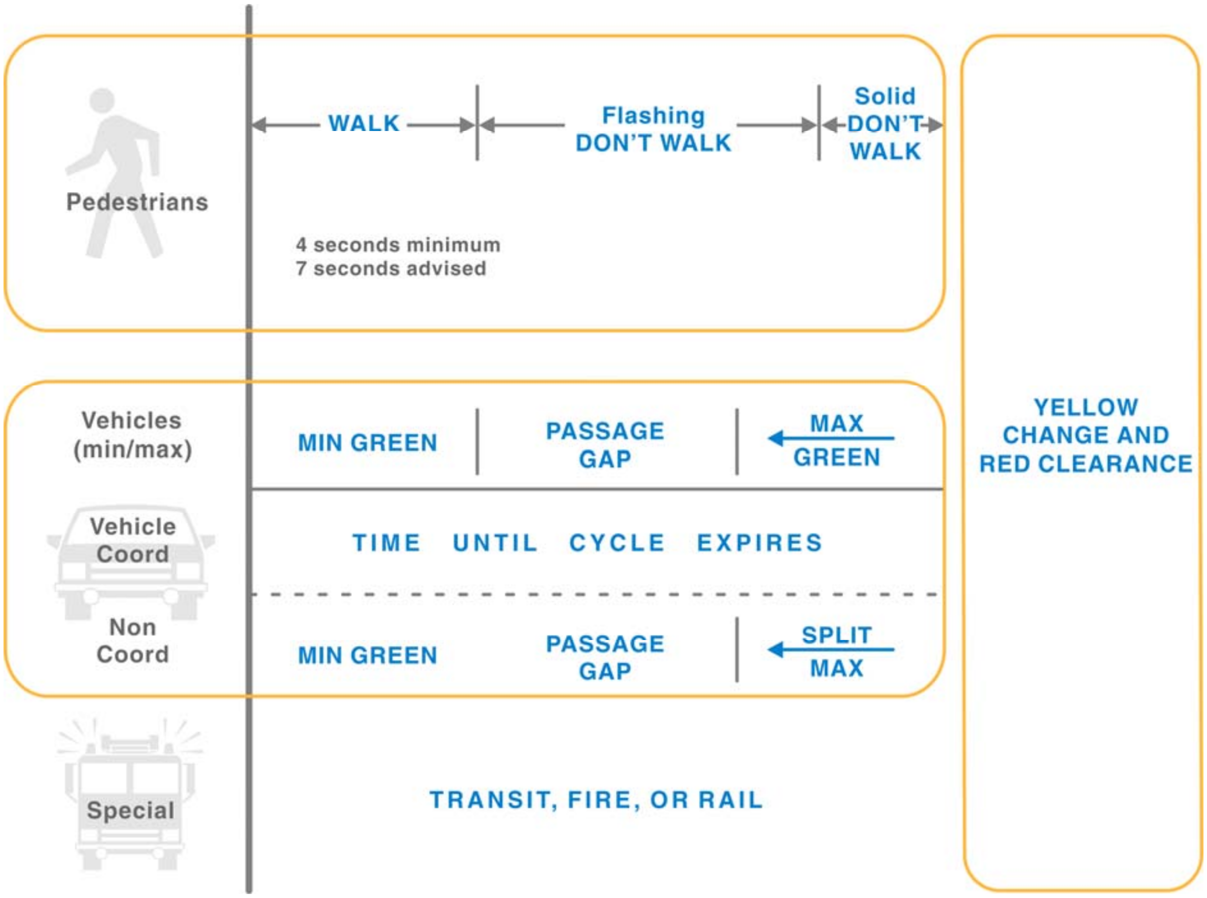
### 6.6.3 Split Distribution

The splits operate as a part of the coordinated timing plan, essentially acting as another set of maximum green times for the non-coordinated phases. Once a cycle length is determined, split distribution is the process of determining how much of the cycle should be provided to each of the phases. These are maximum durations a phase may be served before it must terminate and yield to the next phase. Splits are typically allocated to provide a design level of capacity to all of the minor movements, with the remaining residual time allocated to the coordinated movement (21).

The coordination split for a phase $i$, expressed in seconds, is calculated by the sum of the green, yellow and red times, $g_i + y_i + r_i$. A split for phase $i$, expressed as a percentage, is calculated as $100(g_i + y_i + r_i)/C$. This $g_i$ value is independent of the maximum green time for phase $i$ and the walk and flashing don't walk (if applicable). The actuated logic described in Chapter 5 applies to the phase and thus the phase may not use the entire split percentage allocated within the cycle. Figure 6-21 shows these three timers, the basic timing parameters for vehicles and pedestrians and the corresponding split time associated with the coordination plan. The coordinated phase is slightly different in this regard, in that it receives the remaining time available in the cycle length.

Determining adequate split times can be challenging. If a split time is too long, other approaches may experience increased delays, while if a split time is too short, the demand may not be served. There are often opportunities to vary controller parameters to allow for the fluctuations in daily traffic flow. As shown in Figure 6-21, there are many factors that should be considered in developing signal timing for both a single and a series of intersections.

**Figure 6-21How a phase times**




**Maximum green values may be ignored during coordination using the INHIBIT MAX feature in many controllers. This allows the phase to extend beyond its normal maximum green value.**

There are various policies for determining the necessary split time for a movement. The intent with split times is to provide sufficient time to avoid oversaturated conditions for consecutive cycles, but over the course of an analysis period (15 minutes or one hour) split distributions seek to provide a volume to capacity ratio that is consistent with the operating agency's design standards. In many cases, this will provide an opportunity for fluctuations to be met with the slack time or variable green time, and the actuated operation will reduce phases as necessary to maintain efficient operations. Slack time is defined as the additional time in a cycle that is more than the minimum split times for the phases at the intersection. One common policy allocates the green time such that the volume to capacity ratios for the intersection critical movements are equal in the coordinated cycle length. Another policy is to allocate a minimum amount of time to the minor streets and the remainder to the major or coordinated phases to enhance progression opportunities and maximize bandwidth. The latter methodology is used in traditional coordination, assuming the non-coordinated phases gap out. With many controller parameters and features, this allocation of green time can depend on pedestrians, transit phases, and gap settings.

### *Coordinated Phase*

The length of the coordinated phase split is defined by the demand on the other movements. The coordinated phase receives the time within the cycle that is unused by the other phases of each ring.

There is a close relationship between the rings with the barrier on the intersecting street and therefore the various movements must be considered carefully. In periods of low demand on the non-coordinated phases, the coordinated phase may receive the entire cycle in the absence of an opposing call. The opposing call must be received before the permissive window expires.

***Non-Coordinated Phase(s)***

For a non-coordinated phase to time during a cycle, a call must be active, or the phase could be activated if a corresponding phase in the other ring is active and dual entry is enabled. For instance, if phase 4 has an active call, the last movement (phase 8 is most likely) will be green as long as phase 4 is active and dual entry is enabled for phase 8. Once the phase is active the basic signal timing settings (described in Chapter 5) and the split defined in the coordination plan determine the length of the phase.

### 6.6.4 Offset Optimization

Offsets should consider the actual or desired travel speed between intersections, distance between signalized intersections, and traffic volumes. In an ideal coordinated system, platoons leaving an upstream intersection at the start of green should arrive at the downstream intersection near the start of the green indication. For the users, this is a relative offset, where the time-distance relationship is observable and promotes progression. The actual offset is not always observable because of the actuated logic within the controller that can provide an early return to green.

The HCM suggests that an analyst should review the time-space diagrams to analyze arterial progression and the effectiveness of offsets for a set of signal timing plans. The actuated coordination logic of each signal controller causes the green time allocated to the side street to vary on a cycle by cycle basis. Thus, the time-space diagram is dynamic because of the phenomenon of "early return to green" that results from variable demand on the non-coordinated phases. The HCM translates this assessment of offsets into an arrival type that is used to modify the second delay term of the delay equation. Determining the quality of progression factor (PF) term of the HCM average intersection delay equation is a difficult task, even if observed in the field. In fact, a study where traffic engineers were asked to observe several identical video clips of vehicles arriving at a traffic signal indicating those subjective assessments had wide ranges in estimated arrival types.

Traditional methods for field optimization have included an engineer or technician observing in the field to determine whether the timing plan is operating and whether the offsets are effectively progressing traffic between intersections. These observations provide the engineer or technician with a limited ability to review conditions. With a 100-second cycle, there are 36 cycles during an hour, and the effect of an early return to green on one cycle may not be indicative of the next cycle's performance. Assuming an observation period of at least three cycles leads the field performance assessment with a limited review of the conditions. Engineers are supplementing the field review with improved data.

## 6.7 COORDINATION COMPLEXITIES

This section discusses the various complexities of signal coordination. There are many variables that that must be considered to achieve an acceptable coordination plan. The guidelines address the following topics:

- Hardware limitations
- Pedestrians
- Phase sequence
- Early return to green
- Heavy side street volumes
- Turn bay interactions
- Oversaturated conditions

Each of these topics is addressed separately in the remainder of this section.

### 6.7.1 Hardware Limitations

The microprocessor-based traffic signal controllers used today allow vendors to add new features by changing firmware or software. Advancements in processing power have led to many developments (e.g., alternative phase sequences, number of available timing plans, etc) as well as a variety of signal timing practices. The incompatibility of the equipment and various functionalities has led to some frustration for maintenance personnel. Various government agencies and the National Electrical Manufacturers Association have been actively developing standards to address some of the inconsistencies in hardware.

Although the actual traffic signal system technology and standards have evolved significantly in recent years, many issues such as funding, political support, management, training, inter-jurisdictional coordination, and common regional visions for system operation have significant opportunities for improvement (22). Limitations of funding to upgrade older traffic signal controllers have led to systems that meet today's needs, but "do not provide the building blocks for cost-effectively implementing integrated and interoperable systems." (23).

A particular example of this, as it relates to coordination, is the use of various cycle lengths throughout a day and during different times of the year. It is conceivable that a different set of timing plans is needed for summer and winter months (associated with tourist or other traffic trends), and five different timing plans are desired during a typical weekday and five additional plans warranted on the weekend. This may require the traffic signal controller to operate 20 different plans, which may not be possible due to memory storage in older versions of the hardware.

### 6.7.2 Pedestrians

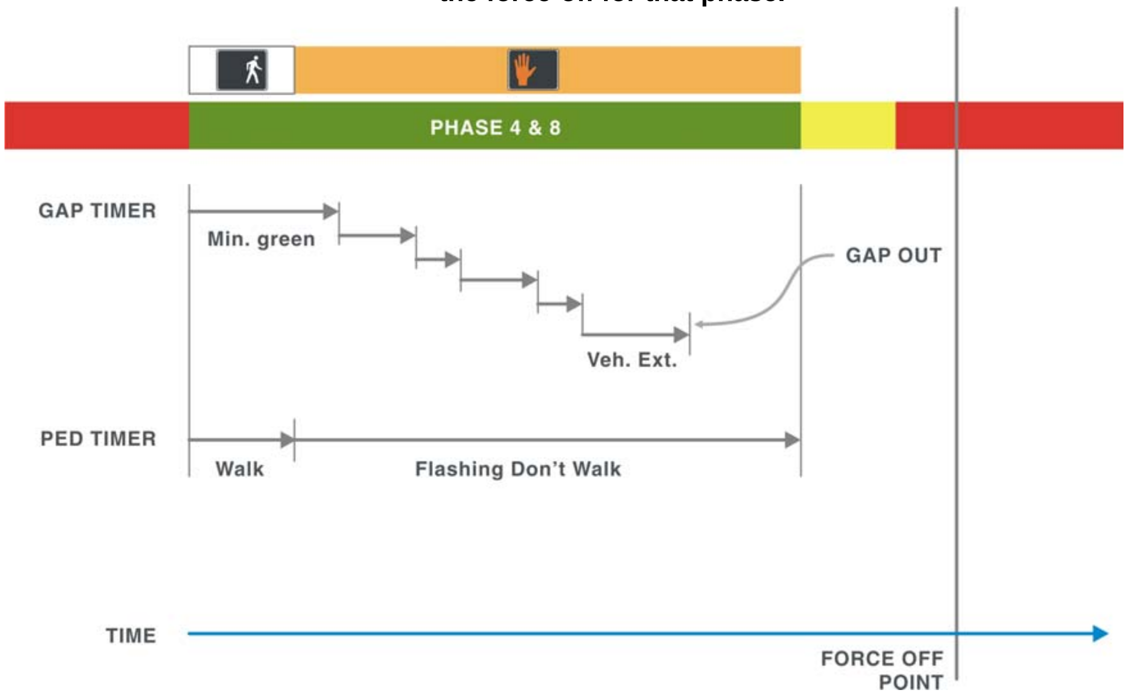
Pedestrian operations can have a direct impact on the ability to maintain coordination along an arterial. For some agencies, pedestrian crossing time is provided for all coordination plans within the split time for the phase, while other agencies (and controllers) allow traffic signals to suspend coordination when there is a pedestrian call, requiring the controller to resynchronize after a pedestrian call. Providing for pedestrian crossing time every cycle may result in a larger cycle and a reduction in green time available for main street movements. This may result in a less than optimal timing plan.

The provision of pedestrian timing and the effects of that pedestrian timing on coordination are two distinct concepts. Pedestrian timing is required for all phases that serve pedestrians. However, when pedestrian activity is relatively low, it may be desirable to allow a pedestrian call to have an impact on coordination because the network system is more efficient without accommodating pedestrians within the coordinated cycle length.

***Pedestrian timing for non-coordinated phases***

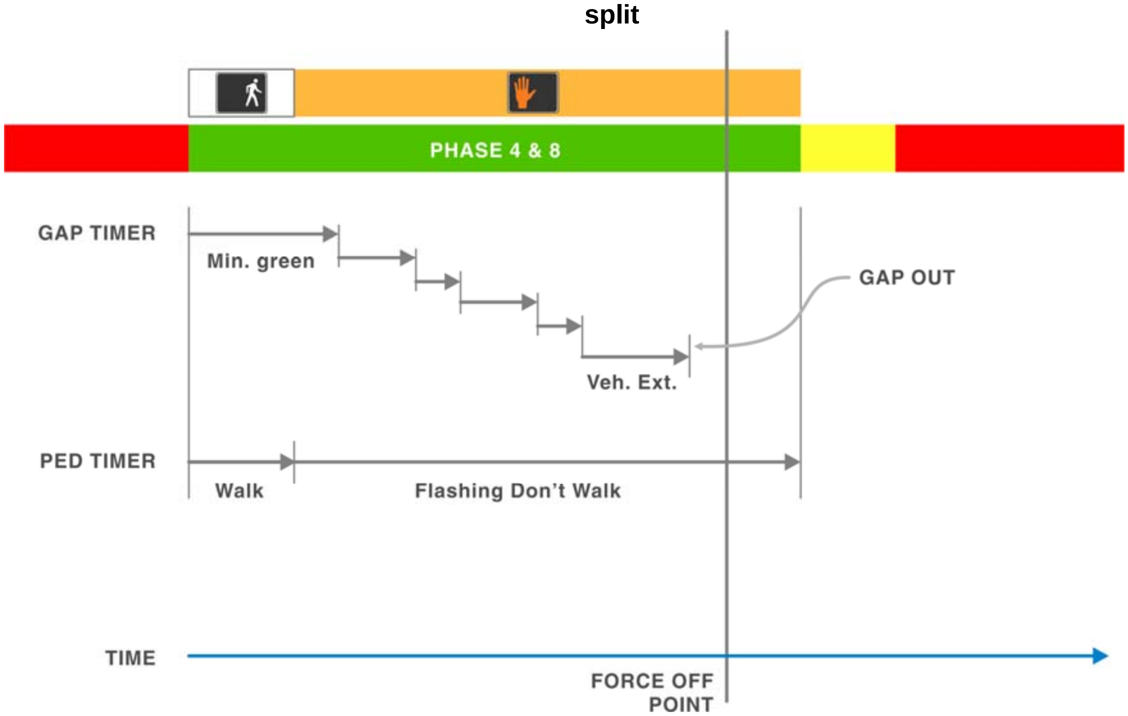
The effect of pedestrian timing on coordination is most commonly seen as it affects minor street timing. Figure 6-22 illustrates the basic principle of pedestrian timing for the minor street where the vehicle split is sufficient to accommodate the required pedestrian time.

**Figure 6-22 Non-coordinated phase operation with pedestrian timing completed before the force-off for that phase.**



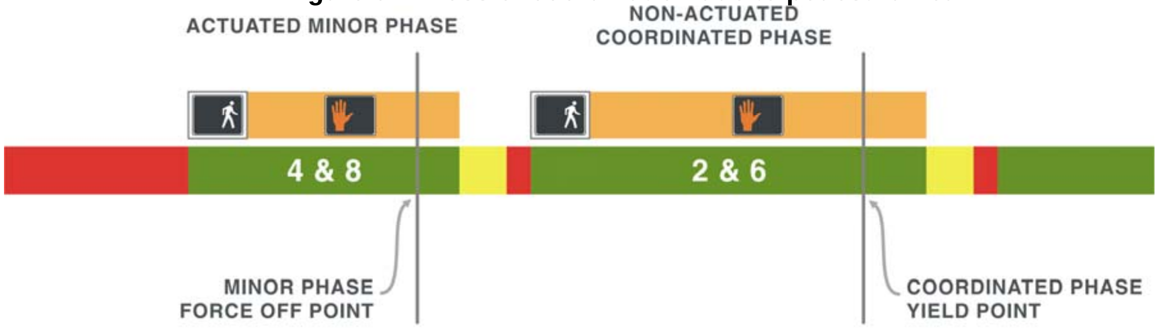

When the split for the phase in question is not sufficient to cover the pedestrian timing, the controller times the phase beyond its force-off point, as illustrated in Figure 6-23. The response of the controller depends on a two factors: (1) demand for subsequent non-coordinated phases and (2) non-actuated versus actuated operation for the coordinated phases.

**Figure 6-23 Non-coordinated phase operation with pedestrian timing exceeding phase split**



When the coordinated phases are non-actuated, the coordinated phases must begin timing sufficiently in advance of the controller's yield point to enable full vehicle timing (minimum green) and pedestrian timing (walk plus flashing don't walk). Should the amount of time be insufficient to cover these timing requirements, the controller will time the coordinated phase past the yield point and fall out of coordination, as shown in Figure 6-24. It is at the yield point that the controller logic determines the method by which the controller will transition back into coordination.

**Figure 6-24Loss of coordination due to pedestrian call**




As a general rule, it is desirable to accommodate pedestrian timing entirely within the split for a given phase. By doing so, any pedestrian calls that may occur can be accommodated without causing the controller to time the phase beyond its force-off point. In these circumstances, the controller loses coordination and must transition back into coordination.

In practice, it is possible to use smaller splits than are needed to cover pedestrian timing without adversely affecting coordination. The ability to do this depends on the capability of the controller. For example, using one particular brand of controller, coding vehicle split times of 85 to 90 percent of the pedestrian timing (walk plus flashing don't walk plus vehicle clearance interval timing) results in an immediate loss of coordination.  In other cases, when these force-offs are combined with cycle lengths that are long enough to allow a controller to temporarily shorten its cycle length during transition without violating the controller minimum (typically at least 10 percent greater than the controller minimum cycle length), the controller will typically resynchronize within a cycle or two, thus having minimal adverse effect on coordination.

A questionnaire survey from the NCHRP 172 project (24) indicated that as a general rule, pedestrian minimum time should be used for the side street when a pedestrian call occurs more than 20 percent of the cycles.

***Pedestrian timing for coordinated phases***

The amount of time needed to serve vehicle volume or provide bandwidth along the major street usually results in coordinated phase splits that are sufficient to accommodate pedestrian timing. Many controllers require that pedestrian timing be accommodated within the coordinated split timing to allow any type of coordinated operation.

For controllers operating with non-actuated coordinated phases, the major street splits must be large enough to accommodate all vehicle and pedestrian minimum timing requirements. For actuated coordinated phases, however, it is sometimes possible to provide a split for the coordinated phases that is less than that required to serve pedestrians. In practice, this works acceptably only if (1) pedestrian demand along the major street results in relatively few pedestrian calls, and (2) demand for the non-coordinated phases is frequently less than the split. In these cases, the controller can take advantage of the unused time from the non-coordinated phases to serve the coordinated pedestrian timing without passing the yield point and falling out of coordination.

### 6.7.3 Phase Sequence

The sequence of phases, particularly those of left turns, can provide measurable benefit to arterial operation. The most common phase sequencing decision, whether to lead or lag left turns, can have a particularly strong impact on the ability to provide bandwidth in both directions of an arterial. Other phase sequence decisions, such as the sequence of left turns on the minor street or the sequence of split phasing on the minor street, do not directly impact arterial bandwidth but can affect arterial delay. These two concepts are discussed in the following sections.

***Major street left-turn phase sequence***

Modern controllers allow left turn phase sequences to be varied by time of day. This has traditionally been done only for protected left-turn operations, but the use of specific display techniques allows this to be extended to protected-permissive operations (see Chapter 4).

The basic concept of lagging a major street left turn is to time the left turn after the opposing through movement (assumed to be one of the coordinated phases) terminates. Figure 6-25 illustrates a typical time-space diagram showing an arterial with only leading lefts and the same arterial with both leading and lagging lefts. The arterial demonstrated in the figure has a major intersection on each end and a minor intersection in the middle. As can be seen in the figure, a lagging left turn at the middle intersection facilitates better progression in both directions because it allows the two platoons to arrive at different times in the cycle. In addition, the two major intersections benefit to some degree from selective lagging left turns.

**Figure 6-25 Vehicle trajectory diagrams showing the effect of changes in phase sequence**

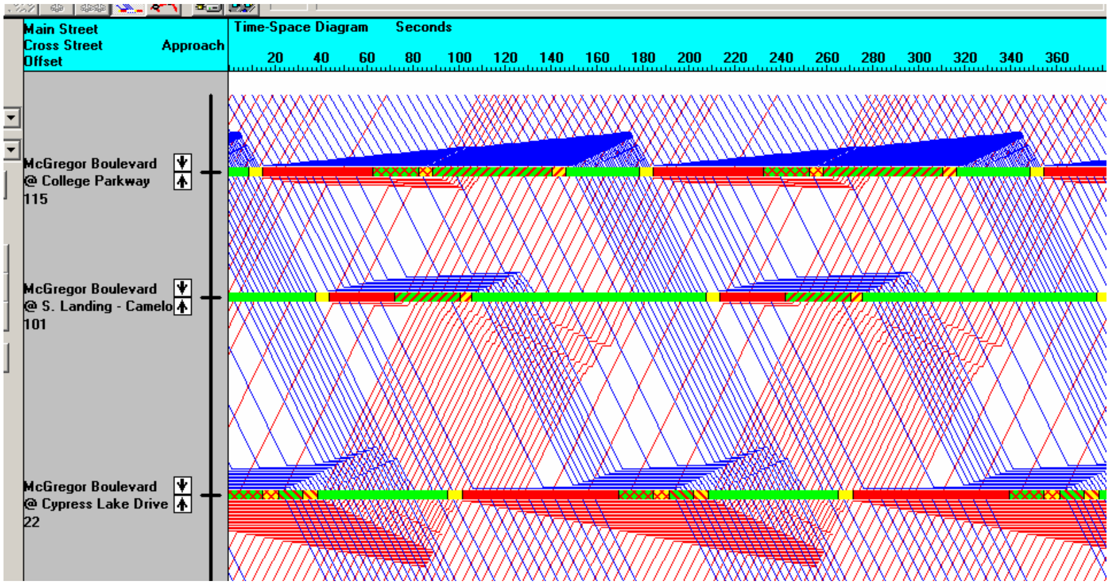



Figure 6-25a is a vehicle trajectory diagram for an arterial with only leading left turns on the major street.

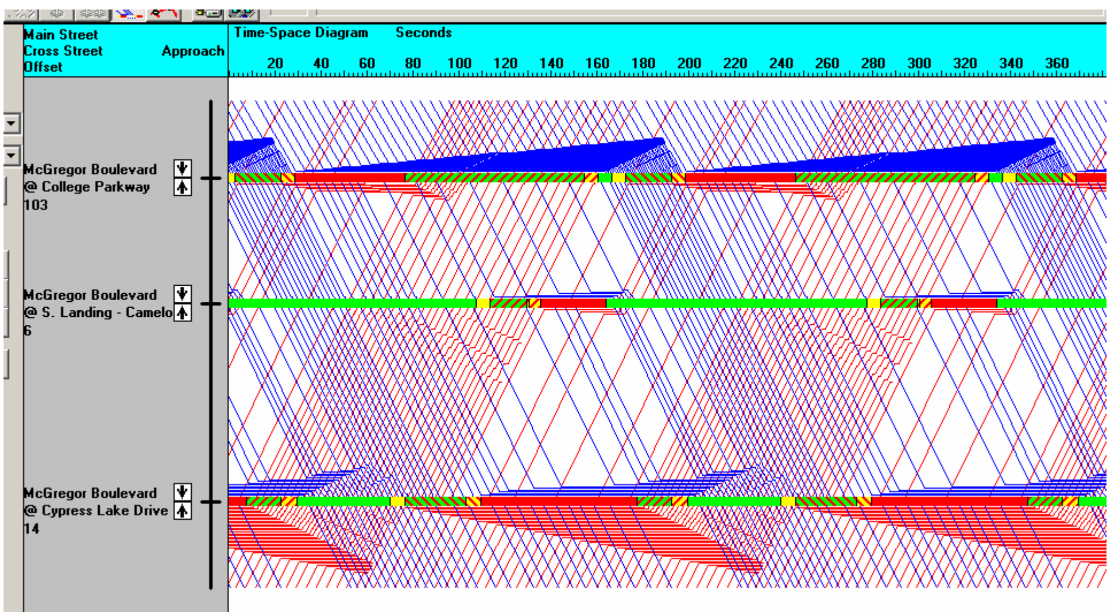



Figure 6-25b is a vehicle trajectory diagram for the same arterial but using selective lagging left turns on the major street.

One of the potential consequences of lagging left turns that are actuated is that the end of the adjacent coordinated phase becomes less predictable. In terms of dual-ring operation, the lagging left turn is typically served after the deterministic (yield) point is reached. The lagging left turn extends the concurrent (adjacent) through movement time indirectly, not as a result of any particular timing within the coordinated phase itself. As a result, only the detection for the lagging left turn is used to determine when to gap out the lagging left turn phase and the adjacent coordinated phase. Therefore, it is possible that the adjacent coordinated phase may gap out earlier than expected from cycle to cycle.

One technique that has been used to eliminate this variability is to use a maximum recall on the lagging left-turn phase. In most controllers, this can be set by time of day and can often be paired with the specific timing plan containing the lagging left turn. The use of a maximum recall on the lagging left turn makes the end of the adjacent coordinated phase more predictable. On the other hand, if the demand for the lagging left turn is highly variable or is less than the split coded, the use of the recall on the lagging left turn may give the appearance of sluggish operation or defective detection.

In addition to the operational differences associated with lagging left turns, some have expressed concern over potential safety differences with having left turn sequencing changing by time of day. Even if lagging operation is used during coordinated operation during the majority of the day, the intersection is often configured to revert to leading-left operation when operating free (uncoordinated) during nighttime operations. There is no definitive research offering consensus on whether changing the left-turn sequence throughout the day has any negative safety consequences.

***Minor-street phase sequence***

It may be advantageous in some circumstances to adjust the left-turn phase sequence for the minor street. In doing this, it may be possible to reduce the delay and queuing for minor-street left turns as they enter the major street and arrive at downstream intersections. Although such adjustments may affect system-wide delays and stops, they will have no effect on the theoretical bandwidth for the coordinated phases.

### 6.7.4 Early Return to Green

One of the consequences of coordinated, actuated control is the potential for the coordinated phase to begin earlier than expected. This “early return to green” occurs when the sum total of the time required by the non-coordinated phases is less than the sum total of the vehicle splits coded for the phases. While this may reduce delay at the first intersection, it may increase system delay because of inefficient flow at downstream intersections or, most important, the critical intersection of the network. Figure 6-26 illustrates this within a time-space diagram.

**Figure 6-26 Time-Space Diagram Example of Early Return to Green**

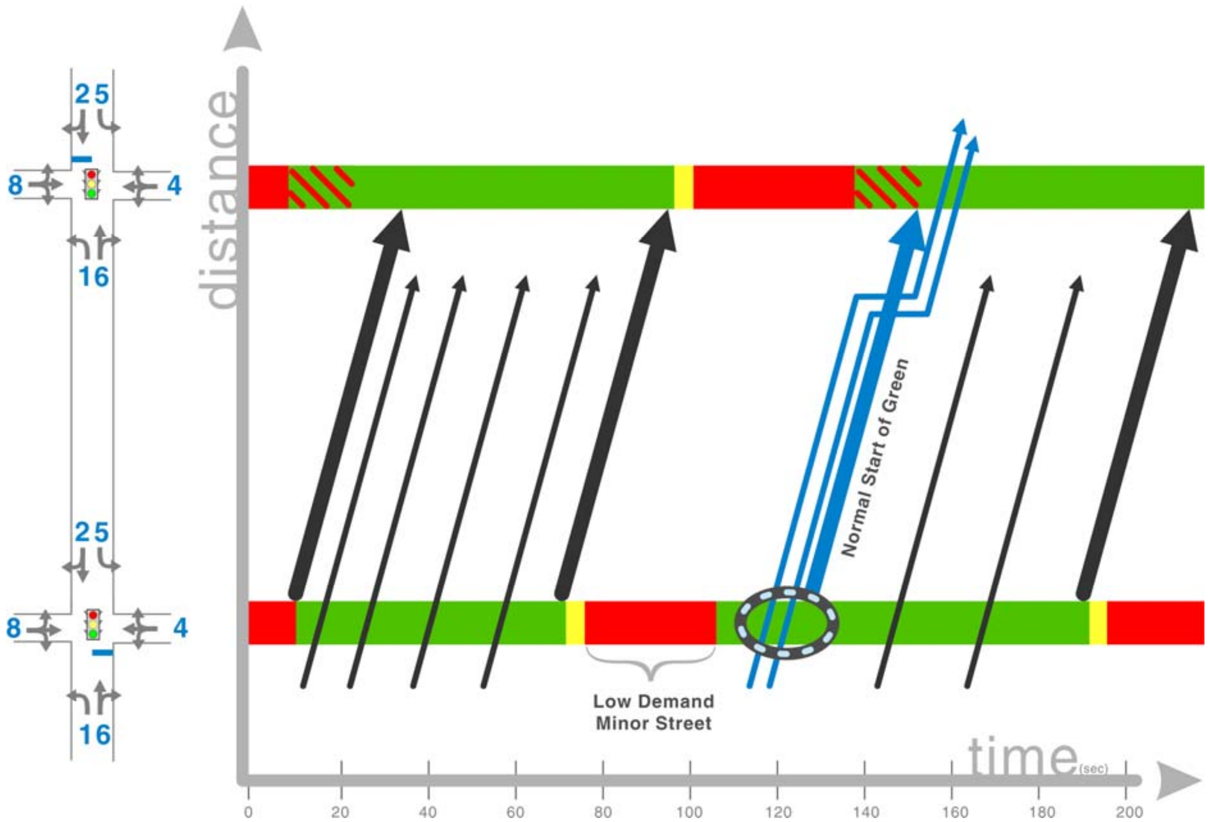



Figure 6-26 shows that vehicles in coordinated phases that begin early may be forced to stop at one or more downstream intersections until they fall within the "band" for that direction of travel. This can result in multiple stops for vehicles and a perception of poor signal timing.

Early return to green can have a substantial negative effect on the performance of the coordinated phases. Research for offset transitioning has been completed to "*smooth progression of a platoon through an intersection using the volume and occupancy profile of advance detectors*" (25). Early return to green can be difficult to manage along a corridor, and it rarely can be completely prevented without eliminating most of the benefits of actuation. One technique that is sometimes used is to delay the start (offset shifted to the right) of the coordinated phase at a critical upstream intersection with sufficient non-coordinated demand (thus making its operation more predictable). Similarly, minor intersections downstream of this critical offset can be started earlier (offset is shifted left in a time-space diagram) to minimize the likelihood of a stop due to an early return to green. In either case, the engineer should use caution when shifting offsets to address early return to green in one direction may adversely affect operation in the opposite direction.

### 6.7.5 Heavy Side Street Volumes

Heavy side street volumes can affect the ability to progress through movements along an arterial. These volumes can come from either signalized intersections within the coordinated signal system or from unsignalized intersections, or from driveways between coordinated signals. Interchanges are a common source of heavy side street volumes.

In many cases, this additional demand proceeds along the remainder of the arterial and becomes part of the major street through demand at downstream intersections. However, this demand often enters the system outside the band established for through movements traveling end-to-end along the arterial. It is usually desirable to adjust downstream intersection timing to allow these heavy side street movements to proceed with a minimum of stops. In these cases, solutions that seek to optimize arterial bandwidth may be counterproductive to effective signal timing.

### 6.7.6 Turn Bay Interactions

Turn bay (or turn pocket) interactions can significantly reduce the effective capacity of an intersection. This is experienced when either demand for the turning movement exceeds the available storage space or when vehicle queues block the entrance of a turn bay. If left-turn demand cannot enter the left-turn bay due to impeding through vehicles, the left-turn phase will gap out early due to a perceived lack of demand. This results in some left turning traffic requiring more than one cycle to be served. In addition, the through movements lose capacity due to the impedance of left turns. In some cases, this may effectively remove an entire through lane from being able to effectively serve through demand.

Turn bay overflows also adversely impact progression. The impedance created by left-turn vehicles stored in the through lane prevents through traffic from proceeding to downstream intersections. As a result, any platoon of through vehicles passing through the affected intersection tends to be dispersed. This reduces the ability for downstream intersections to efficiently provide green time for the platoon.

Turn bay overflows can occur under both under- and oversaturated conditions. Even if enough green time is provided to serve a given turning movement, turn bay overflow can occur if the available storage is insufficient to store the queue for a given cycle. For cases where the turn bay is fed from a two-way left-turn lane, turn bay overflow rarely has significant adverse consequences for the adjacent through movements. For locations with raised medians, on the other hand, turn bay overflow can result in left-turning vehicles extending into the adjacent through lane.

Turn bay overflow can be managed in a number of ways:

- Turn bays can be extended to accommodate the necessary storage. This is typically the best solution, but it may be infeasible due to physical constraints, access needs, turn bay requirements for adjacent intersections, or other factors.
- Shorter cycle lengths can be used to keep overall queue lengths shorter, thus reducing the likelihood of overflow or turn bay blocking.
- If the left turn is protected, protected-permissive left-turn phasing may be considered to allow some of the left turn demand to be processed during the permissive portion of the phase. This reduces the overall queue length.
- If the left turn is permissive, protected-permissive left turn phasing may be considered to provide a period of higher saturation flow rates (the protected portion of the phase). This technique, however, may result in longer cycle lengths that partially offset the gain in capacity. Some agencies use a lagging protected period after the permissive period that is called only if there remains any unserved left-turn demand at the end of the permissive portion of the phase.
- Conditional service for the phase may be invoked, bringing up the movement twice during the cycle.
- If two receiving lanes are available, the adjacent through lane can be designated as a shared left-through lane and the phasing changed to split phasing. While this is rarely desirable for major street movements, it may be an appropriate solution for minor street movements.
- At an intersection with one heavy left-turn movement on the major street, it may be preferable to designate the left-turn phase as one of the coordinated movements paired with the adjacent through movement (e.g., phases 1 and 6 or phases 2 and 5). This allows any unused time to roll over to the left-turn phase of interest, thus reducing its effective red time and associated queue formation.

### 6.7.7 Critical Intersection Control

A challenging aspect of timing an arterial street or a network of streets is the need to provide enough capacity for major intersections without creating excessive delay for minor intersections. Ideally, all of the intersections to be coordinated operate optimally with similar cycle lengths. However, most arterial streets do not have this optimal arrangement due to a mixture of simple signals (e.g., two phases) with more complex signals (e.g., eight phases), wide ranges in cross street volume (e.g., major arterials versus collectors), and variations in left-turning volumes.

The techniques to determine the ideal cycle length for each intersection in isolation were covered earlier. The critical system cycle length is the maximum optimal cycle length of any intersection in the system.

Several techniques can be used where there is a significant disparity in the ideal cycle length for each intersection:

- Each intersection is timed using the critical system cycle length. This ensures the ability to coordinate all of the intersections in the system. However, it may result in excessive delay at minor intersections.
- Each intersection is timed to either the critical system cycle length or to half that value. This technique is commonly referred to as “double cycling” (a minor intersection cycles twice as frequently as a major intersection) or “half cycling” (a minor intersection has half the cycle length of the major intersection). This method can often produce substantially lower delays at the minor intersections where double cycling is employed. However, it may become more difficult to achieve progression in both directions along the major arterial, which may result in more arterial stops than desired.
- The major intersections are operated free, and the minor intersections are coordinated using a shorter cycle length. Because the major intersections are operating free, it is impossible to provide coordination through the major intersections. Therefore, major street vehicles are likely to stop at both the major intersection and at a downstream intersection due to randomness in arrival at and departure from the major intersection. This technique can often result in lower overall system delay at the expense of additional stops along the major street.

### 6.7.8 Oversaturated Conditions

Timing for oversaturated conditions requires different strategies than those used for undersaturated conditions. An intersection that is operating at or over capacity requires all movements to operate at a saturation flow rate to serve demand. Beyond this, the timing plan may favor the movements with the most lanes to maximize the throughput of the intersection. Obviously, the timing plan must consider whether undesirable effects such as turn bay overflow or other conditions exacerbate the problem.

Under these conditions, arriving vehicles must join the back of a queue to ensure that they enter the intersection with a minimum amount of headway (maximum saturation flow rate); however, it is impossible for vehicles on the arterial to maintain a travel speed traditionally desired for coordination.

In addition, an oversaturated approach cannot serve all arriving demand, thus creating a residual queue at the end of a cycle that carries over to subsequent cycles. This residual queue depends on demand conditions and can grow from cycle to cycle. Even when demand drops, the residual queues create saturated conditions beyond the time period when the arriving demand would create saturated conditions by itself. These residual queues can extend to adjacent intersections and prevent traffic from exiting upstream intersections if the intersections are closely spaced.

The general technique for accommodating oversaturated conditions involves managing queues. The following sections present options available for accommodating oversaturated conditions, including benefits and trade-offs.

### *Queue storage on minor movements to favor major movements*

It is sometimes possible to accommodate oversaturated conditions by favoring the coordinated movements at the expense of minor street movements. Under this strategy, the coordinated phases are timed for a volume-to-capacity ratio typically no higher than 0.95 to 0.98. The minor movements receive less green time, which results in demand-to-capacity ratios exceeding 1.0. This method has a few advantages. The major street receives priority, which helps maintain traffic flow along the major street. This typically benefits the heaviest movements through the intersection, as well as the transit and emergency vehicles who frequently use the major street. In addition, demand held on side streets cannot enter downstream intersections, thus improving the actual downstream flow rate.

However, this type of timing strategy can have significant disadvantages. While this method can theoretically keep the major street moving, it creates extensive delay and queuing on side streets. This can have negative safety repercussions and highly negative public feedback. More importantly, it may be impossible to reduce splits for side-street through movements due to pedestrian timing requirements. If pedestrian calls are frequent enough to cause the side street through movement to time to its full split, the only movements with time available for use are the major street left turns and the minor street left turns (if present). Queue spillback for the major street left turns can exceed available storage and spill into the adjacent through movement, creating operational and safety problems. As a result of these disadvantages, this technique is often not desirable.

### *Use of short cycle lengths*

One of the most effective techniques in managing oversaturated conditions is the use of shorter cycle lengths. Shorter cycle lengths allow more frequent servicing of all movements at only a minor expense of additional loss time during the peak time period. This frequent cycling provides more equitable servicing of all movements and allows drivers to visibly experience progress, even if it takes multiple cycles to be served at a given intersection.

### *Queue management on major street movements*

An alternative technique is to selectively store queues on major street movements. Candidates for this treatment frequently include major intersections that are spaced far enough from other signalized intersections to allow the queues to grow without creating upstream intersection impacts. In addition, drivers may be more accepting of congestion at major intersections than at minor intersections.

For network applications (e.g., downtown grids), it is often best to store queues outside the network using key signals to meter traffic into the grid network. While this creates congestion at some intersections, it allows a network of intersections to operate undersaturated, thus enabling traffic to progress through the network.

### *Use of actuated uncoordinated operation*

Removing the cycle length constraint during the oversaturated period can result in efficient allocation of green time, provided gap timers are set appropriately. There are emerging strategies such as lane-by-lane detection and measurement of flow used as opposed to presence for control logic decision-making and operation.

## 6.8 REFERENCES


1. National Traffic Signal Report Card, National Transportation Operations Coalition, 2005.

2 Fundamentals of Traffic Engineering (15th ed.), UCB-ITS-CN-01-1, January 2001

3 Srinivasa R. Sunkari, Roelof J. Engelbrecht, and Kevin N. Balke, Evaluation Of Advance Coordination Features In Traffic Signal Controllers

4 Albert Grover & Associates, “Minimize Delay, Maximize Progression with Protected Permissive Lead/Lag Phasing”, Technical Workshop. Inland Empire Section of ITE, December 14, 1995.

5 Frederickson, J. Letter to the Arizona Insurance information association. October 2000.

6 *Kittelson & Associates, Evaluation of Traffic Signal Displays for Protected/Permissive Left-Turn Control. NCHRP 3-54(2), Final Report, FHWA, U.S. Department of Transportation, Washington, DC, 2003*.

7 *Shelby, S. G., Bullock, D. M., and Gettman, D. Transition Methods in Traffic Signal Control. Transportation Research Record: Journal of the Transportation Research Board, No. 1978, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 130-140.*

8 Zong Z. Tian, Tom Urbanik, Kent Kacir, Mark Vandehey, and Howard Long, “Pedestrian Timing Treatment for Coordinated Signal Systems”, viewed at wolfweb.unr.edu/homepage/zongt/Publications_files/PedTimingICTTS.pdf on January 17, 2007.

9 Signalized Intersections Guide

10 Los Angeles Department of Transportation, “Early Methods for Traffic Signal Network Timing” http://www.lacity.org/ladot/TopicsAndTales/TrafficSigTiming6.pdf, visited March 5, 2007.

11 Nelson, Eric “Timing Strategies for Managing Oversaturation in Harris County”, Presentation at the TRB Summer Meeting, Woods Hole, MA, July 11, 2006

12 Lieberman, Ed, NCHRP 3-38 (4), Transportation Research Board, 1992.

13 Parsonson, “Signal Timing Improvement Practices”NCHRP Synthesis 172 National Research Council, Transportation Research Board, Washington, D.C., February 1992

14 Webster, F.V., “Traffic Signal Settings,” Road Research Technical Paper No. 30, HMSO, London (1958)

15 Teply, Stan “Saturation Flow at Signalized Intersections through a Magnifying Glass,”Proceedings of the Eight International Symposium on Transportation and Traffic Theory University of Toronto Press, 1983.

16 Alexiadis, V. K. Jeannotte, A. Chandra, Traffic Analysis Toolbox Volume I: Traffic Analysis Tools Primer, FHWA -HRT-04-038, June 2004.

17 Signal Timing Process, Federal Highway Administration, Sabra Wang, December 30, 2003

18 Fambro, Sunkari, Sangineni, “Guidelines for Retiming Arterial Networks”, Texas Transportation Institute, February 1993

19 Trafficware. Synchro 6 Users Manual, 2004.

20 MnDOT Traffic Signal Timing and Coordination Manual, March 2005

21 Bullock, D. and A. Sharma, MOST PROJECT Laboratory 6. Actuated Traffic Signal Control, DRAFT, January 2007.

22 Transportation Infrastructure, Benefits of Traffic Control Signal Systems Are Not Being Fully Realized. GAO/RCED-94-105, Report to the Chairman, Committee on Energy and Commerce, House of Representatives, United States General Accounting Office, Washington, D.C., March 1994

23 Bullock, Darcy and Tom Urbanik, “Addressing Diverse Technologies and Complex User Needs”, http://onlinepubs.trb.org/onlinepubs/millennium/00116.pdf, visited January 17, 2007

24 Ibid

25 Abbas, Montasir, Darcy Bullock, and Larry Head, “Real-time offset transitioning algorithm for coordinating traffic signals”, Transportation Research Record 1748, Washington D.C. 2001

# CHAPTER 7

# DEVELOPING SIGNAL TIMING PLANS

## TABLE OF CONTENTS

## LIST OF FIGURES

# 7.0 DEVELOPING SIGNAL TIMING PLANS

This chapter documents the process for developing signal timing plans. The intent of this chapter is to describe the steps necessary to develop a traffic signal timing plan for a series of intersections. These steps, along with calculations, must be made to produce a timing plan for field implementation. The information presented in this chapter builds on the concepts, guidelines, and information presented in the previous chapters.

## 7.1 OVERVIEW

Timing plans can be developed for an isolated intersection or several intersections on a coordinated arterial. At a single intersection an adjustment may be in response to a citizen complaint or agency staff member's observation regarding a specific problem. Field observation by agency staff is often critical to identifying localized problems and solutions. This can include minor adjustments to the detector settings or fine tuning to adjust the split and/or offset at the problem intersection for the time of day during which the problem was observed. This may also include adjustments to pedestrian and clearance intervals in response to a perceived safety problem. These types of adjustments are essential for effective signal system operation. They are restricted to that particular intersection and do not usually consider the impact the timing changes may have on nearby intersections. These adjustments are solely driven by a localized change in traffic conditions (and in some cases geometric conditions) or the enhancement of existing timing at a single intersection. Care should always be taken to make sure the problem at a single intersection is not transferred to another location resulting in overall diminished system performance.

Area-wide timing is quite different in scope from localized adjustments. The area-wide timing is a systematic update to a group of coordinated intersections. It provides the appropriate and necessary timing plans for each intersection in terms of its individual needs as well as the combined needs of a series of intersections (1). This includes a system-wide cycle length that provides the best overall performance. There is a trade-off between the efficiency of the entire system and that of individual intersections, and good area-wide timing should strike a balance between the two. Equally important, the area-wide timing process considers the number of different timing plans required for varying levels of traffic demand and considers the best times and/or traffic conditions for the changes between plans. It includes a rigorous process of engineering and testing both in the office and the field to ensure that the timing plans being implemented are well suited for the conditions. Minor adjustments are an important aspect of traffic signal operations but are not a substitute for carefully updated area-wide timing.

### 7.1.1 Signal Timing Process

The signal timing process has eight distinct steps that can be tailored depending on whether the effort is a slight adjustment of the timing or is an area-wide or corridor retiming effort. The signal timing process steps were introduced in Chapter 2. Refer to Chapter 2 to ensure that the relevant policies and objectives are defined and incorporated in the scope of the project (the first step in the process). The steps include:

1. Project Scoping;
    - Review Policies, Determine Objectives, and Identify Problems;
    - Confirm Standards and Procedures;
    - Identify the study system, corridor, or intersections;
    - Divide the System into Sections;
    - Select Performance Measures;
    - Identify the Number of Timing Plans;
2. Data Collection;

3. Model Development;
    - Data Input;
    - Analysis;
    - Draft Timing Plans;
4. Field Implementation and Fine Tuning;
5. Evaluation of Timing;
    - Performance Measurement;
    - Policy Confirmation; and
    - Reporting.

The decisions and calculations made in these steps should be based on field data or first-hand observations of traffic operations at the subject intersection. This process is discussed in detail in Section 7.3 as it relates to developing timing plans for an area-wide project.

### 7.1.2 Frequency of Timing Updates

Traffic signal timing should be reviewed every three to five years and more often if there are significant changes in traffic volumes or roadways conditions. Retiming traffic signals every three to five years is generally considered to be good engineering practice. This frequency of signal retiming is particularly important for jurisdictions or localized areas within a jurisdiction in which land use changes lead to rapid changes in traffic patterns.

**Key to the ability to identify the right frequency for signal timing updates is a proactive monitoring program. Monitoring should include information from sources external to the traffic signal operators, including planning or development review staff or other agency department staff that are in the field. It may also include citizen complaints collected from phone calls or through emails.**

Signal timing should be revisited when any of the following occur:

- Increased demand or changing turning movements (5 to 10 percent; lower percentage change critical when intersections operate at or near capacity);
- Reduced traffic flow (typically 10 to 15 percent reduction);
- Changed vehicle mix (increased truck percentages);
- Construction activities on the study route or on a parallel route;
- Changed land uses associated with development or travel patterns due to shifts in employment centers;
- Observed queue spillback caused by oversaturation of adjacent traffic signal; or
- Other factors (such as the addition of a new traffic signal to a corridor associated with new development or changes in the roadway network).

These changes can be produced by a number of physical alterations to the roadway network, including new roadways, geometric changes to existing roadways, or added or relocated bus stops. Changes may also result from modifications to surrounding land uses, such as new or expanded residential, business, retail or educational facilities. Most communities experience many of these transformations over a period of three to five years. Changes in traffic flow and roadway network characteristics often lead to the need for new signal timing. Equally important, they may also require changes in the times of day during which the timing is implemented. In the case of reduced demand, they may also lead to situations in which signal removal or flash operation during certain periods of the day is appropriate.

These alternative approaches are typically used to determine when new timing plans should be developed:

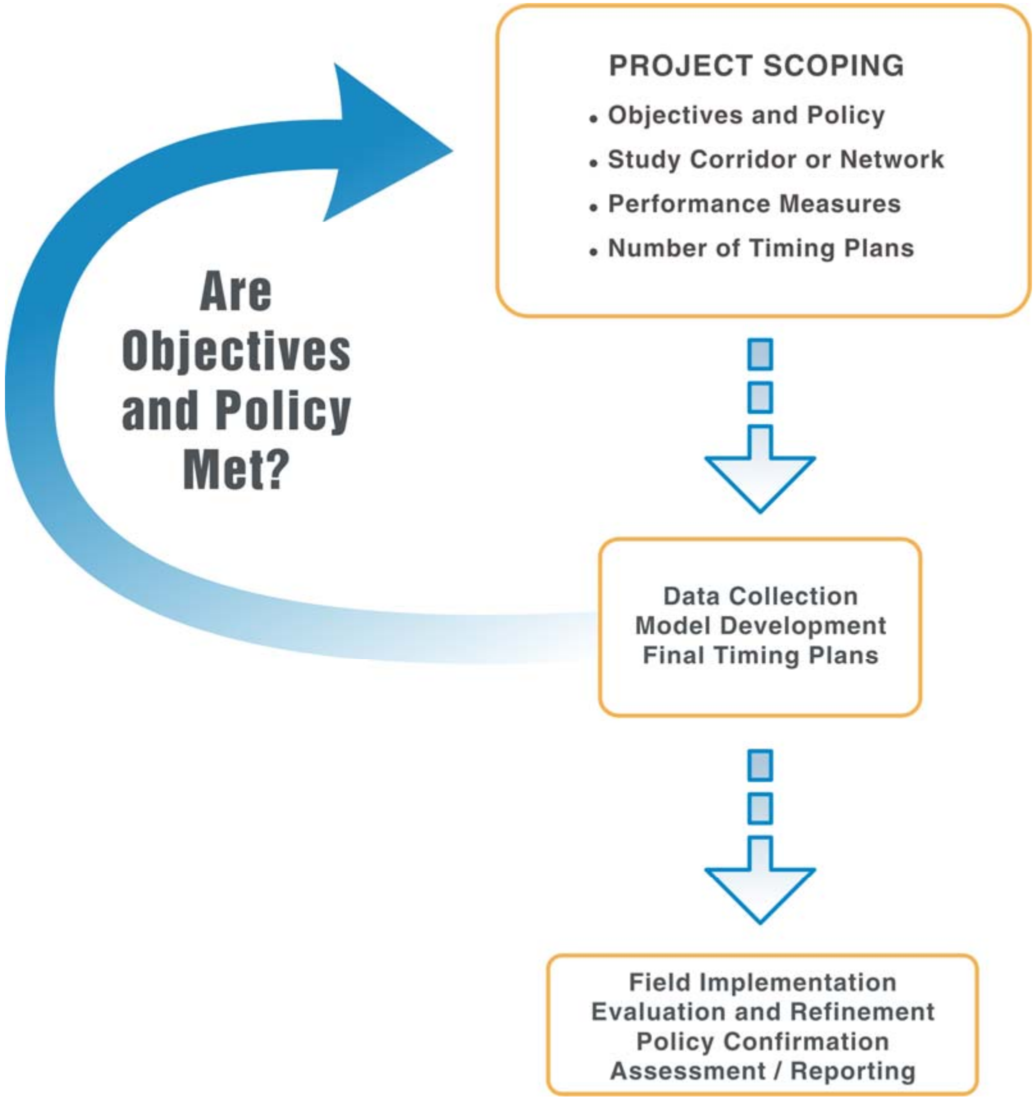
- The agency may develop a regular schedule (e.g., three years) and retime 1/3 of their system annually.
- The agency may identify the need for qualitative new timing and selecting areas for retiming based on the observation of degraded performance or areas undergoing development or redevelopment. The latter is due to increased intersection delays or symptoms such as cycle failure (queues do not completely discharge during each signal cycle), spillback from left turn bays or between intersections, and unused green time on side streets.
- The agency may perform an engineering analysis of areas where signal timing may be warranted. This can be done through the use of a simulation or by methods that compare the performance of the existing timing system against its optimized timing performance developed with signal timing software (such as Synchro, TRANSYT 7F, TEAPAC or PASSER II).

In some cases, agencies have found that signal systems that initially require timing every three years may require less frequent timing (every five years) as neighborhoods and their attendant traffic flows mature and stabilize. However, the need for retiming should be reexamined every three to five years using a formalized process, such as described above.

### 7.1.3 Steps

This section describes a procedure to develop a signal timing plan for a series of coordinated intersections. The procedure comprises a series of steps that describe the decisions and calculations that need to be made to produce a timing plan that will yield an effective relationship between a series of signalized intersections. Each step is described in detail in the following sections. Figure 7-1 highlights the overall steps involved in a signal retiming project.

**Figure 7-1 Project Scoping**




As highlighted in Figure 7-1, a key component of the signal retiming process is ensuring that the agency objectives and policies are being met with the signal timing plan development. Once these are confirmed, field implementation, evaluation and refinement, and reporting can be performed for the project.

## 7.2 PROJECT SCOPING

The project scoping is a key component of signal timing development. Project scoping establishes objectives, standards and procedures, study area, performance measures, and the number of timing plans.

### 7.2.1 Determine Objectives based on Signal Timing Policies

The signal timing policies and the associated objectives developed using the process described in Chapter 2 are the basis of the development of the area wide timing plan. Common objectives include reducing stops, delay, and travel time for a corridor. It is important in scoping a project to recognize that there is a point when the objectives of smooth traffic flow are irrelevant. Recent research conducted by the FHWA describes this in detail.

During the project scoping period, consideration should also be given to identifying known problems. The identification of these problems may be the result of public comments, staff observations, or known discrepancies with established policies. The identification of problems during project scoping may be an iterative exercise done in conjunction with determining objectives. In other words, previous problems may help shape the project objectives and/or, by clearly defining project objectives, deficiencies may be more apparent and easily addressed.

### 7.2.2 Confirm Standards and Procedures

Beyond the big picture policies determined above, there are also specific procedures and standards that should be confirmed as a part of the scoping process. The standards identify parameters used for the timing of change and clearance intervals, actuated timing settings, and pedestrian timing. The MUTCD defines several of these values and provides standard policies for agency use.

### 7.2.3 Dividing the System into Sections

The third step in project scoping is to identify logical groups (sections) of signals to be included in the timing process. The concept of sections is an important one. Typically, every intersection in a section changes to a new timing plan at the same time of day. Each intersection in a section likely operates in the same control mode—manual, time of day, or traffic responsive. For example, if an incident occurs and an operator would like to manually select an incident timing plan, the incident plan is implemented for every intersection in the affected section. In some cases during oversaturated conditions, intersections may be operated free to insure efficient green allocation.

Operational reasons are not the only justification for sectionalization. When traffic signal optimization software is used, the algorithm that is used to determine the best (optimized) timing is most sensitive to changes in smaller numbers of intersections than changes in an entire system equaling a large number of intersections (20 or more). Smaller groups of signals are easier to deal with. This includes management of data, collection of traffic data, implementation of new timing in the field, and fine-tuning adjustments to installed timing.

From the preceding discussion of sections, it should be clear that sections are autonomous groups of signals that can be controlled independent from other groups of signals. The groups of signals don’t have to operate the same cycle length; double cycling is evaluated when a common cycle length is not desirable. This understanding of sectionalization leads to the definition of the following section characteristics:

- The best size for a section is in the range of 2 to 30 intersections, understanding that as the number of intersections increase, the likelihood of a solution with a very small and/or discontinuous progression band also increases.
- It is possible for a single intersection to serve as a section, such as the case of an isolated intersection, or an intersection that is at the junction of two crossing arterials, each of which is a separate section.
- Section sizes can often exceed 30 intersections in cases where there are no logical breaks in intersection operation, such as in a large central business district where there is equal block spacing.
- If possible, section sizes should not exceed 100 intersections because the data management and installation requirements become unmanageable.
- Sections should include groups of intersections that are logical to control as a “unit.” This can be determined by common changes in traffic flow among all intersections (i.e., the peak period occurs at nearly the same time at all intersections in the section), by common changes in the time at which saturated conditions occur, by which intersections

are impacted by the presence of incidents, or by the geographic location of the intersection.

- Adjacent intersections are considered when the distance between intersections is less than one mile and the traffic flow along the corridor is greater than 500 vehicles per hour per direction.
- To the greatest degree possible, sections should be created among groups of intersections whose operation does not have to be coordinated with intersections outside the section boundary. In this manner, one section can change its timing plan or control mode without concern for the impact on intersections outside the section. A commonly used natural boundary for a section is the intersection of an arterial with a freeway interchange.

However, there are also a number of misconceptions about sections, including:

- **It is necessary for all intersections in the section to operate using the same cycle length—False.** A common cycle length is not required. The definition indicates that all the intersections in the section must operate using the same timing plan, in which the cycle, split, and offset is defined for every intersection. However, the specific values do not need to be common to all the intersections in the section.
- **It is not possible to coordinate sections with each other—False.** Systems using common time bases for all intersections and operating on the same cycle length or a multiple of one another can be coordinated across all section boundaries.
- **Intersections cannot be reassigned to different sections on a time-of-day basis—False.** Some systems permit reassignment, which means it may be possible to vary intersection-to-section assignments by time-of-day.

As previously indicated, section boundaries are selected in a manner that will minimize the need for interaction with the signals in adjacent sections. For example, the signals on roadways connecting the two sections should be far enough apart that there is little need for signals in one section to be coordinated with the signals in the adjacent section. The traffic engineer should look for groups of signals that are surrounded by a buffer zone, which may be in the form of long intersection spacing, natural boundaries such as railroad tracks or bridges, changes in land use, or other factors that lead to differing traffic characteristics.

### 7.2.4 Select Performance Measures

As the project scope is defined, performance measures must be selected to evaluate the success of the retiming effort. Most operators use stops and delays as measures because they are the most sensitive to changes in signal timing and are the most obvious improvements that motorists will observe. However, several things should be considered prior to arriving at this decision' including funding sources, the need to report results to the public, environment, etc. The FHWA and NTOC have projects that highlight a standard set of performance measures that emphasize the importance of travel time in addition to stops, delays, and average speeds.

During under-saturated conditions, stops and delay are the performance measures that are typically used. During oversaturated conditions, the objective is to minimize the time period during which these conditions exist and manage queue spillback between intersections. The performance measures in use include queue lengths, numbers of cycle failures, and the percent time that intersections are congested. This is accomplished by prioritizing movements and controlling the directions of queue build-up to avoid spillback and to minimize cycle failures.

Once the measures have been selected, the locations and routes that are candidates for performance measurement should be selected. Measurements should be made before new timing is installed and again at the completion of the signal timing process. In this way, the effectiveness of the retiming effort can be evaluated and communicated with agency management and the public. The before-and-after evaluation can also be used to identify locations where additional adjustments are needed because their performance falls short of expectations.

### 7.2.5 Identify the Number of Timing Plans

Since timing plans are developed for specific sets of traffic conditions, it is important to define the times of day when those traffic conditions exist, and therefore, the times of day when each plan should be used. It is also important to determine the number of unique sets of traffic conditions that will require signal timing plans. For example, tourist areas that experience a large fluctuation in traffic flow throughout the year may find it beneficial to develop timing plans for the tourist and non-tourist seasons of the year.

A number of factors constrain the amount of plans that can practicably be developed, including:

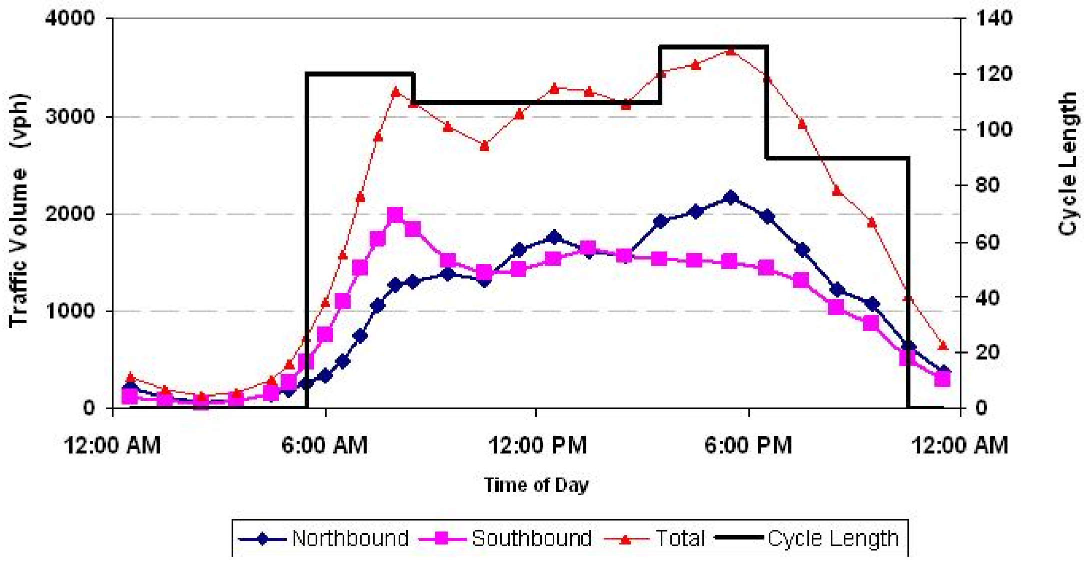
- **Resources.** Collecting data, developing plans, installing the plans and fine-tuning plans is a labor-intensive process; this work is constrained by available personnel and fiscal resources.
- **Control Equipment.** An important step in developing signal timing plans is making sure that the project engineer/technician has a full understanding of the controller type/equipment being used by the agency and its features, functions, and possible limitations of that equipment. Many systems have limits on the number of plans, phase sequences, and other signal timing operations that can be accommodated within the controller. This limitation should be carefully considered, and if resources are available for modernization they should be sought. If possible, a review of the equipment should be completed during the scoping process or, at a minimum, at the start of the project.
- **Data Availability.** Some plans, such as for snow storms, hurricanes, incidents, and store sales are based on conditions that occur on rare occasions. Collection of traffic data during these transitory conditions can be difficult if not impossible. Many agencies respond to these “unusual” circumstances through adjustment of existing plans. For example, agencies might increase the cycle length and slow down the speed of progression for plans that are used during snowstorms.

The process of selecting the number of timing plans and the times of day when they operate may be determined through a combination of reviewing traffic data along the corridor, such as 24-hour directional traffic counts, intersection turning movement counts (TMCs), and traffic engineering judgment. Ideally, the following steps are used to determine the number of different plans and appropriateness for time of day changes:

- For each section of the system (this is another reason why sectionalization is important), select a sample of representative intersections. These might be the most important intersections in the section (if some parts of the section are more important than others), or they may be a group of intersections whose traffic conditions are typical of those existing throughout the section. In either case, the set of representative intersections must include only locations that are adjacent to each other. Note that this process will require the availability of hourly counts for the entire time period being analyzed.
- Prepare a graph that plots the traffic volume as a function of time of day for the two or three most important intersections in the sample. Using the graph, identify the AM, PM and off-peak time periods for the sample. This step is executed using judgment in the same manner that a traffic engineer would manually determine the time periods for these conditions.
- Further analysis could include using the signal timing optimization model to determine whether the performance of the system would be improved with different timing plans. This requires substantially more data and may be completed later in the process.

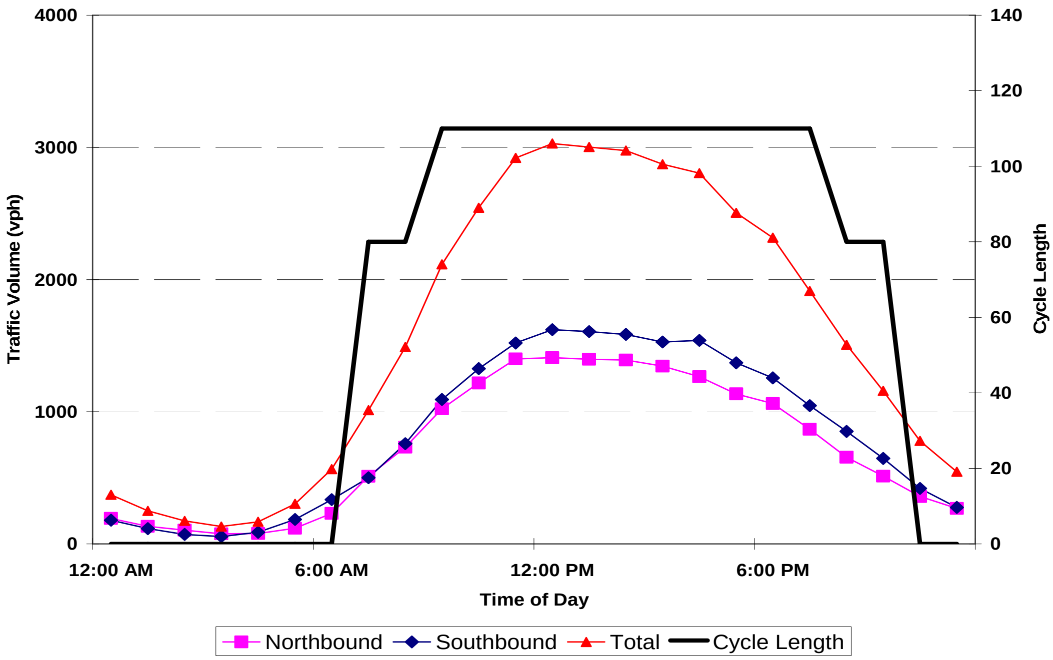
Figure 7-2 has been prepared to represent the process described above. This figure shows the 15-minute periods that are included in each of the three major time periods (AM, PM and off-peak). The bounds or number of major periods can be varied as described above. Frequently, the greatest benefits can be found in breaking the off-peak period from the peak periods. For example, it may be desirable to add a noontime plan to accommodate increased traffic during the lunch hour. However, there must be a balance in the number of plans provided as many traffic signal controllers may need several cycles before achieving a coordinated offset and cycle following plan transition.

**Figure 7-2 Traffic Volumes Summary Used to Determine Weekday Time of Day Plans**



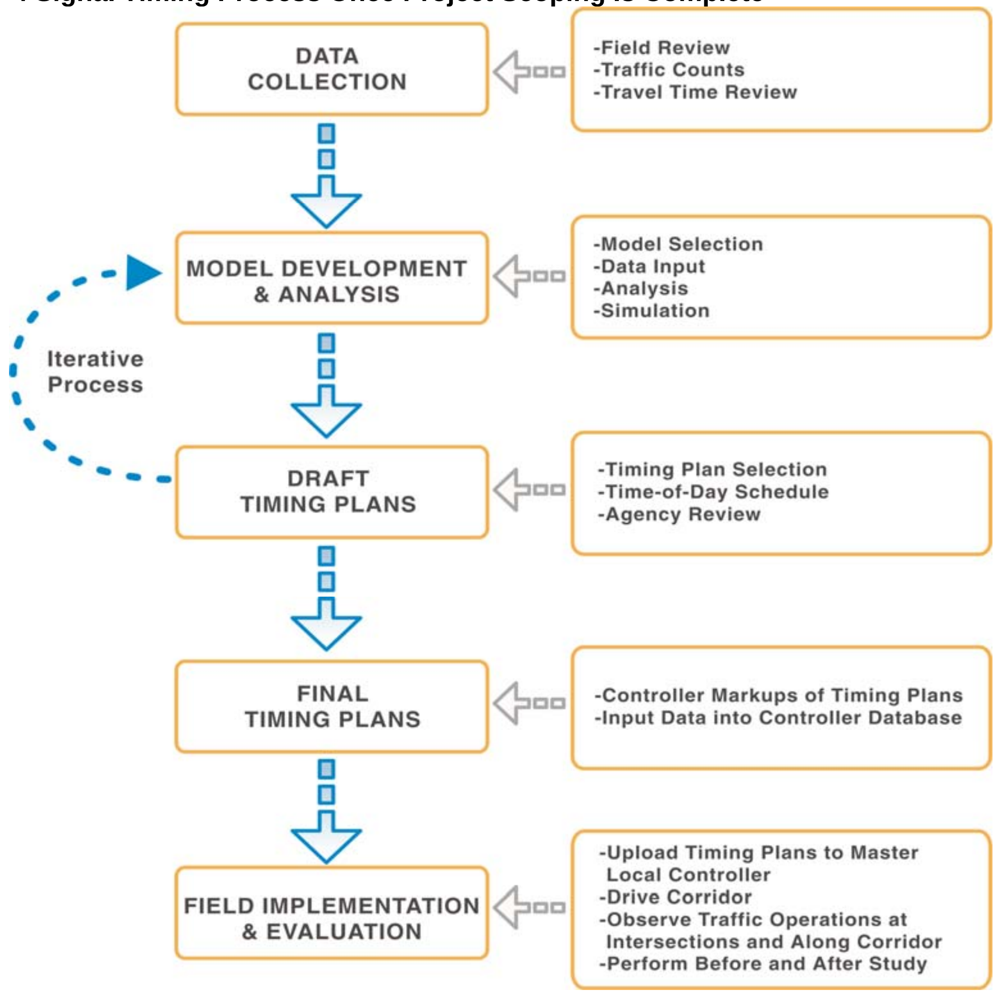

In addition to time-of-day plans, separate plans for weekdays, weekends or even specific days (such as Friday in a high weekend travel area) may be warranted based on variation in travel patterns and volumes. Figure 7-3 illustrates this process for a weekend time-of-day volume profile. For weekend time-of-day plans, it is common to use one timing plan for most of the day due to the balanced traffic volumes. A second timing plan might be used between uncoordinated operations and the peak traffic volumes.

**Figure 7-3 Traffic Volumes Summary Used to Determine Weekend Time of Day Plans**




Through the above-mentioned process, the number of timing plans should be determined and included in the scope of work for the project. Figure 7-4 outlines the next steps used to develop signal timing plans that can be implemented in the field.

**Figure 7-4 Signal Timing Process Once Project Scoping Is Complete**




## 7.3 DATA COLLECTION

The data collection effort for a signal timing project can be quite costly and time intensive. This section summarizes several types of data used, as well as techniques that can be used to reduce the data collection effort and expense.

The data typically can be categorized as: traffic characteristics, traffic control devices, intersection geometry, and crash history. Data associated with each category are described in subsequent sections. The description focuses on the collection of data at an existing intersection. If the intersection does not exist, then estimates will need to be provided. In many instances, the data described herein will have already been collected as part of the engineering study associated with the signal warrant analysis.

### 7.3.1 Traffic Volumes

***24-Hour Weekly Volume Profiles***

Weekly traffic volumes should be collected at critical locations along the corridor. These locations are determined by the peak hour time periods and traffic patterns of the corridor. The weekly traffic volume profiles are an important element in the data collection effort and are used to identify:

- The number of timing plans used during the weekdays and weekend;
- When to transition from one timing plan to the next;
- Volume adjustment factors for developing turning movement counts; and
- Directional distribution of traffic along the corridor.

Weekly traffic volumes can be collected using tubes, system detectors, or other ITS devices that are in place along the corridor. Using system detectors or other ITS devices will reduce the time and cost of the data collection effort.

#### *Turning Movement Counts*

Turning movement counts are typically collected at each of the subject intersections under consideration for retiming. Depending on the traffic volumes and traffic patterns along the corridor, turning movement counts often only need to be conducted during peak periods, commonly the weekday morning, midday, and evening peak time periods. The daily traffic volume profiles should be used to identify the specific time periods for conducting the counts, as well as for use in developing volume adjustment factors for the weekend and/or off-peak traffic volumes. At some locations, the traffic volumes may be higher or more critical during the weekend time periods, which could lead to performing turning movement counts for the weekend and the weekday time periods. Additionally, seasonal traffic patterns may need to be considered and incorporated in the scope of work. Procedures for conducting, recording, and summarizing traffic counts are described in the *Manual of Transportation Engineering Studies (2)*.

Intersection turning movement counts are collected for representative traffic periods. The traffic count should include a count of all vehicular traffic at the intersection, as categorized by intersection approach (i.e., northbound, southbound, etc.) and movement (i.e., left-turn, through, or right-turn movement), pedestrians, and vehicle type (including transit). If heavy vehicles are significant, the vehicular count should be further categorized by vehicle classification. The presence of special users at the intersection (e.g., elderly pedestrians, school children, bicycles, emergency vehicles, etc.) should also be documented.

#### *Vehicular Speed*

Data on vehicular traffic speed should be gathered to identify the approach speeds to the intersection. This is especially important in determining the signal phasing. In cases where the traffic signal is a new installation, detection layouts (number of detectors and location) are based on the approach speed. Traffic speeds are especially important when considering the detector timing described in Chapter 5.

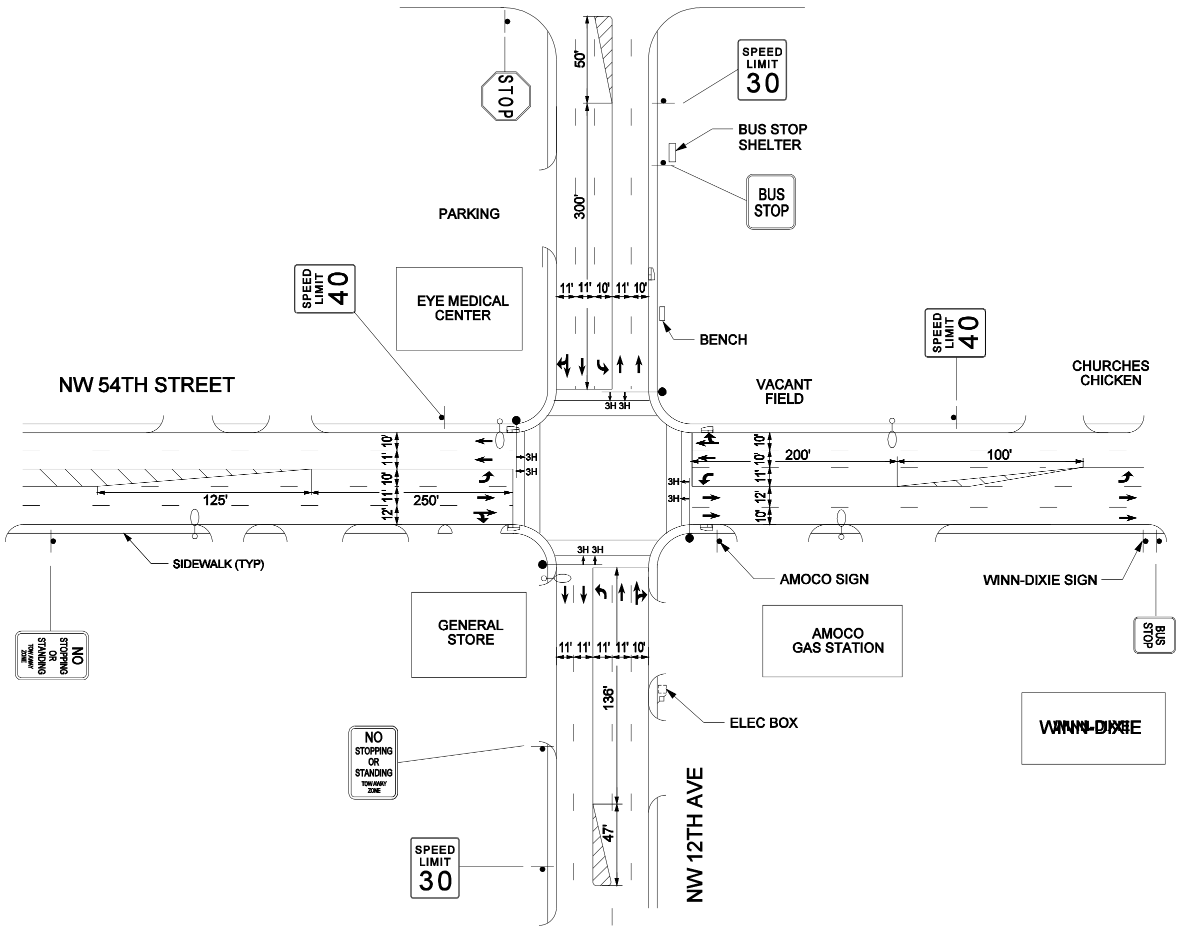
#### *Travel Time Runs*

Travel time runs can be used to calibrate the existing analysis model and to compare the corridor operations before and after the new timings are implemented. Travel time runs are collected by driving the subject corridor and recording the delay, stops, and running time using electronic equipment. The following information can be collected from the study and later analyzed and used to develop the timing plans: average travel times, intersection delays, intersection stops, and speeds. Travel time and speed are common measures to report to elected officials and the public when doing signal timing improvements. It is also a good practice to monitor travel time as a means of determining the quality of existing timing and the need to retime.

### 7.3.2 Intersection Geometry and Control

A site survey should be conducted to record relevant geometric and traffic control data. These data include: number of lanes, lane width, lane assignment (i.e., exclusive left turn, through only, shared through and right turn, etc.), presence of turn bays, length of turn bays, length of pedestrian crosswalks, and intersection width for all approach legs. An effective method for recording this information is with a condition diagram. An example condition diagram is shown in Figure 7-5.

**Figure 7-5 Example intersection condition diagram.**




Other information that can have an impact on signal operations include: approach grades, sight line restrictions, 85$^{th}$ percentile approach speed (if different from the speed limit), presence of on-street parking, presence of loading zones, presence of transit stops, dips in approach profile near the intersection, and intersection skew angle. Many of these factors will have an impact on the capacity of one or more movements at the intersection; they may also influence intersection safety.

If the intersection is being re-designed and existing conditions are the basis for documenting the benefits associated with the proposed design, then information about the existing detection design is also needed. This information includes: the number and location of existing detectors, their length and type, the number of detector channels associated with the detection for each traffic movement, and the type of controller memory associated with each detector channel and the detection mode.

All traffic control devices at the intersection should be inventoried during the site survey. This inventory should be categorized by intersection approach and include the speed limit signs, parking prohibition signs, turn movement prohibition signs, and one-way street signs. This information includes: the number and location of existing vehicular and pedestrian signal heads, phase sequence, use of overlaps, and signal timing settings.

### 7.3.3 Field Review

An important component of the signal timing development process is performing a field review at the signalized intersections and of the subject corridor. The key elements to consider are location and operation of signal equipment, intersection geometry, signal phasing, intersection operations, vehicle queuing, adjacent traffic generators, posted speed and/or free flow speed. The free flow speed will assist with calibrating the traffic model and establishing offsets that reflect the traffic conditions. However, the use of the free flow speed or posted speed should be discussed during the scoping process.. The field review is an excellent opportunity for the engineer or technician to become familiar with the operations, geometry, and potential constraints at the intersection, as well as to identify signal equipment (i.e., pedestrian indicators, detectors, or controller) that might not be operating properly. It is also an excellent opportunity to observe operational issues that would not be obvious by the hard data alone, such as queue spillback and approaches that are not serving the full demand. The data collection sheet shown in Figure 7-6 provides some of the information that might be useful during the field review.

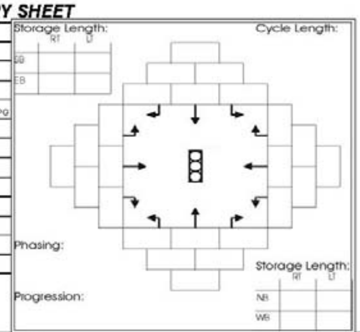
**One of the challenges of field data collection is to visit the corridor in a way that allows observation of the most critical time periods and important intersections. Identifying congested locations will be important to isolate where traffic analysis may need additional attention.**

**During the field review, if observations are made that reveal the signal timings are working adequately, then this should be noted. Many locations have traffic patterns that may be stable and have not undergone any geometric or signal phasing improvements, resulting in the existing operations being acceptable. This initial review may result in a smaller scope of project and could lead to focusing the effort on other deficiencies.**

**Figure 7-6 Signal System Data Collection Sheet**

***SIGNAL SYSTEM INVENTORY SHEET***

| | |
|---|---|
| PLACE AGENCY OR COMPANY LOGO HERE | Project Name: |
| | Project Number: |
| | Analyst: |
| | Date: |
| | Filename H:\profile\6125\photos\Data Collection\[Inventory Sheet.xls]DATA INVENTORY (PG |




| | | | |
|---|---|---|---|
| **Intersection Name** | | | |
| **Section #** | | **Zone #** | |

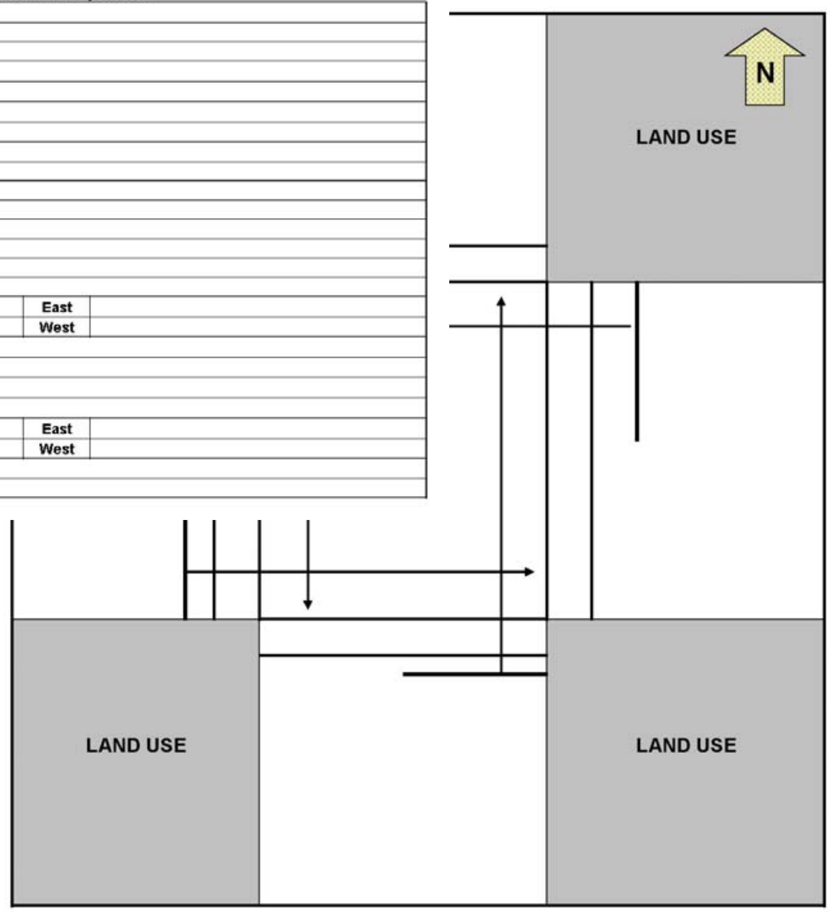
| | |
|---|---|
| **Picture #'s** | |
| | |
| | |
| | |

**Communication/System Inventory**

| | | | |
|---|---|---|---|
| **Communication Type** | | **Cabinet Type** | |
| **Emergency Pre-Emption** | | **Cabinet Size** | |
| **Detection** | | **Controller Type** | |
| **Special Settings** | | | |
| **Line Of Sight (Photos)** | | | |
| **Other** | | | |

**Qualitative Assessment - Intersection Operations**

| | | | | |
|---|---|---|---|---|
| **Level of Service** | | | | |
| **Queue Lengths by Approach** | **North** | | | |
| | **South** | | | |
| | **East** | | | |
| | **West** | | | |
| **Pedestrians** | | | | |
| **Cycle Length** | | | | |
| **Coordination** | | | | |
| **Speed Limit by Approach** | **North** | | **East** | |
| | **South** | | **West** | |
| **Land Use by Quadrant** | **NW** | | | |
| | **NE** | | | |
| | **SW** | | | |
| | **SE** | | | |
| **Crosswalk Distance by Approach** | **North** | | **East** | |
| | **South** | | **West** | |
| **Other** | | | | |



Turning movement counts are typically one of the more costly items in the scope of a traffic signal retiming study. FHWA published “Signal Timing on a Shoestring,” which provides some guidance on ways to minimize the data collection effort through a short count method or by estimating turning movement counts using link traffic volume data. These guidelines may be useful on corridors or networks where the traffic conditions are predictable. However, in developing areas where traffic patterns are rapidly changing, it may be more appropriate to collect two- to three-hour turning movement counts for the various study-time periods.

**Use of existing detection for collection of turning movement counts is a practice recommended by the Institute of Transportation Engineers in a recent Informational Report.**

### 7.3.4 Existing Signal Timing

This information is important to obtain at the subject intersections as it is used to develop an existing analysis model. Key information from the existing signal timing includes phase sequence, yellow and all-red intervals, pedestrian walk and flashing don’t walk intervals, minimum green, and detector settings. Additionally, if the intersection is operating in coordination, then the cycle length, splits and/or force-offs, offsets, and reference phases should be obtained from the existing signal timing. Some unique settings may be used in the field that can be identified by reviewing the existing signal timing, such as conditional use, overlap, or limitations in the phase sequence that can be incorporated in the existing model and timing development. The existing signal timing helps the user understand what currently exists in the field and provides a baseline for improvement.

### 7.3.5 Intersection Analysis

The building blocks of traffic signal timing rely on the analysis procedures highlighted in Chapter 3 and the intersection design elements described previously in Chapter 4. Combining the analysis with the design elements leads to the development of potential changes as a part of the signal timing process.

#### *Assess Signal Phasing*

During this step, data collected is used to determine a reasonable signal phasing for the conditions. Specifically, the key determinations in this step are whether a left-turn, right-turn, or pedestrian phase is needed for the various approaches in combination with the opposing approach. Guidelines for making these determinations are provided in Section 4.3. The phase sequence for leading or lagging left turns is described in Section 4.4. Pedestrian phasing is covered in Section 4.5.

#### *Identify Traffic Signal Control Mode of Operation*

During this step, the need for coordination should be evaluated. Additionally, whether the intersection should operate as pre-timed, semi-actuated, or fully actuated and whether it should be coordinated with adjacent signals is determined. It is important to consider the daily volume variation, prevailing approach speed, and objectives for the operation of the traffic signal. Guidelines for making this determination are provided in Section 5.2. If the intersection is part of a coordinated signal system, then the control mode is likely to be coordinated-actuated, where the minor movements are actuated and the major-road through movements are coordinated.

#### *Verify Detection Design*

An intersection's detection design depends on the mode of operation and the prevailing speed on each intersection approach. An intersection operating with pre-timed control does not require detection. In contrast, an intersection operating with semi-actuated or fully-actuated control will need detection for each actuated traffic movement. Guidelines for designing various types of detection are given in Chapter 4.

### *Refine (or Develop) Basic Signal Timing Parameters*

During this step, the basic signal timing parameters are evaluated to determine the settings that will operate as a part of the timing plan. In the absence of coordination, these parameters control the green allocation and cycle length of the intersection. Intersections operated as part of a coordinated system will rely on the parameters to provide boundaries for the traffic signal timing and modifiers to the coordinated plans. It is very important that the basic signal timing parameters, as outlined in Chapter 5, be considered during any retiming effort as they may have a significant impact on both coordinated and uncoordinated operations. An evaluation of the basic timing parameters may be most effectively performed through field observation (e.g., performance of vehicle detection due to passage time settings or adequacy of clearance times).

## 7.4 MODEL DEVELOPMENT

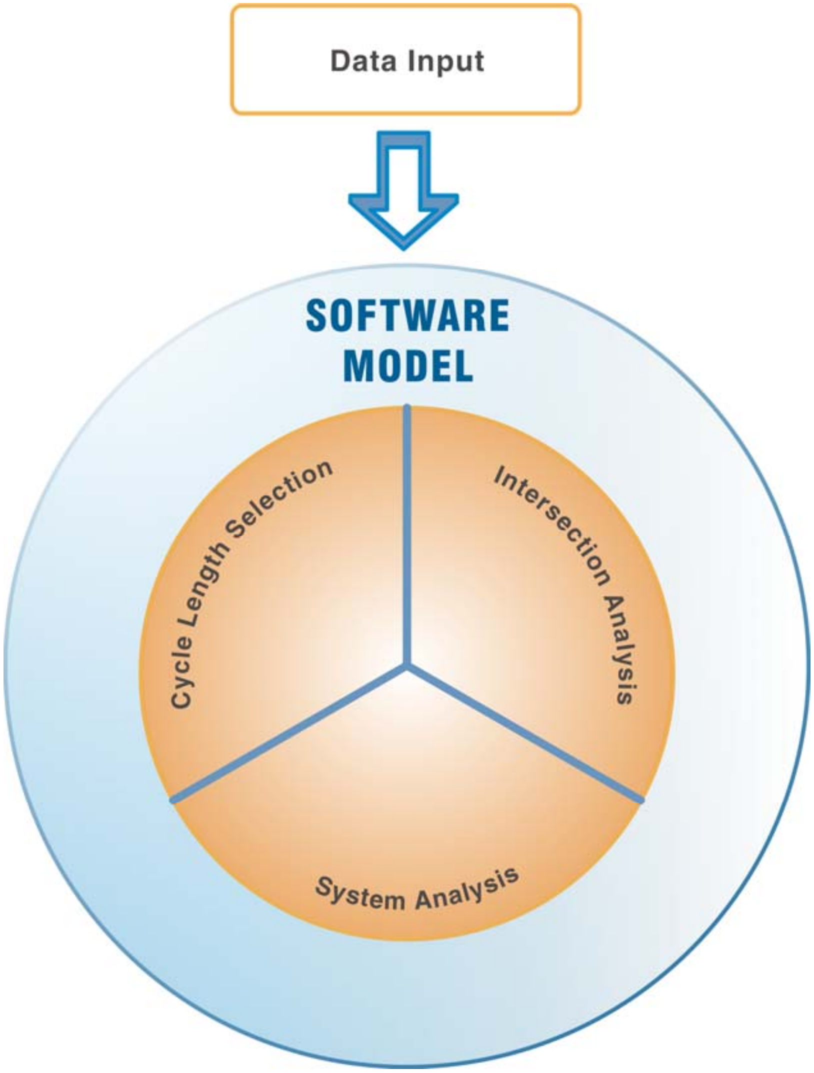
This section addresses the steps needed to develop an analytical model for the signal retiming study. The first step is to select the model. Chapter 6 discusses the various models used for signal retiming and the use of manual methods.

Many computer-based tools are available to calculate and evaluate signal timing. Since many of these tools assume the presence of under-saturated conditions, it is important to recognize their capabilities and limitations. It is recommended that users develop a thorough understanding of the selected computer programs, their use, and how they relate to the field conditions.

The process of evaluating existing timing plans and developing new timing plans is iterative and involves detailed reviews of the information to successfully meet the objectives of the study. Several elements should be considered in the study:

- Establish a “standards and conventions” document (i.e., file naming, map settings, base data parameters, analysis settings) that provides the user with consistency through the retiming process;
- Review the plan development in levels or stages to ensure efficiency;
- Coordinate with the respective agencies; and
- Include quality assurance and quality control measures.

**Figure 7-7 Use of a Signal Timing Model in Initial Assessment**




As shown in the figure, the data (traffic volumes, link network, lane geometry, existing signal timing, etc.) is coded in the software model and used to perform a cycle length analysis. From this analysis, a preferred cycle length is selected for the section based on the objectives of the study. This cycle length is refined through an iterative process of evaluating individual intersection operations with the system operations. Through this process, a final timing plan is developed and used during field implementation.

The requirements for developing timings for saturated and under-saturated conditions should be considered as the model is developed. Saturated conditions require careful distribution of green time that should include timing settings that seek to balance queue buildup and meter incoming demand. Because under-saturated flow tends to be stable (similar from one cycle to the next) traffic signal timing software can be used to calculate the best signal timing. Congestion exists if traffic conditions in the section are unstable with growing queues, spillback, cycle failure, or left turn bay overflows. Under these conditions, a separate process of calculating timing for saturated conditions should be used.

In saturated conditions, congestion is difficult to avoid. Therefore, the control policy should be aimed at postponing the onset and/or the severity of secondary congestion, which is caused by the blockage of upstream intersections.

**Turning movement counts may not capture the actual demand at intersections when conditions are oversaturated. Typical data collection measures for turning movement counts result in the traffic flow through the intersection rather than the actual demand that may still be part of a queue. It is important to validate the model to identify periods and locations where demand exceeds capacity.**

### 7.4.1 Data Input

Depending on the size of the network, the data input work effort varies in intensity. With a few intersections the work effort is quick. However, if you have a corridor with 15 intersections or a downtown network with 100 intersections, the work effort becomes time consuming. To be successful, data input must be organized. The typical data used in the model include lane geometry, link speeds and distances, phase numbering, left- and right-turn phasing, existing signal timing (i.e., yellow and all-red intervals, pedestrian walk and flashing don't walk intervals, minimum green, and detector settings), controller type, and coordinated reference phases. The data can be categorized into two sets: Network Parameters and Signal Timing Parameters.

- **Network Parameters -** These are parameters that are typically fixed through the analysis process. Examples of these include lane geometry, link speeds and distances, phase numbering, and left- and right-turn phasing. The link distances can be identified using aerial photography or a GIS map as base mapping in the software.

- **Signal Timing Parameters –** These are isolated signal timing parameters that are typically reviewed and modified during a retiming study. Examples of these parameters include yellow and all-red intervals, pedestrian walk and flashing don't walk intervals, minimum green, and detector settings. These parameters are reviewed, and changes may be recommended at the beginning of the project. Once these parameters are reviewed and approved, they are used in the network and to develop the new coordinated signal timing plans. Chapter 5 describes methods to calculate signal timing parameters.

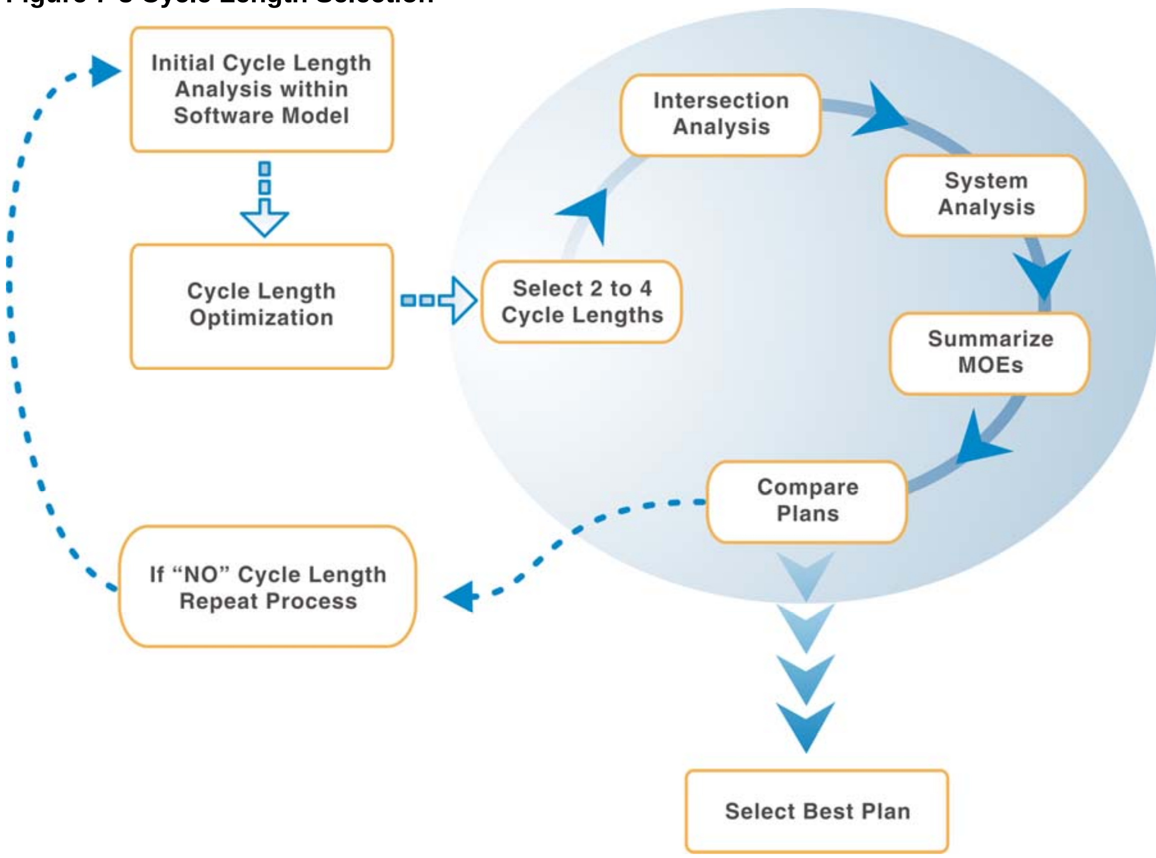
The network parameters and signal timing parameters are inputted in the selected model (i.e., Synchro, PASSER, TRANSYT-7F, or TEAPAC) to establish the base network. Certain parameters, such as traffic volumes and signal timing, may be inputted in the model through an automated process using the Uniform Traffic Data Format (UTDF). A base network is established for each of the study time periods (i.e., morning, midday, p.m., etc.)

Field observations should be compared with the traffic operation results for each time period in the model. If necessary, the base networks for each of the time periods should be calibrated using travel time, delay, and queue data collected from the field. Parameters within the models that can be adjusted to calibrate the existing base networks with actual field conditions include saturation flow rates, right-turning-vehicles-on-reds, lane utilization, and other parameters. A review of the calibrated model should be performed prior to moving forward with the timing plan development analysis.

### 7.4.2 Analysis

Several steps are typically used to analyze the different time periods when trying to establish a set of timing plans for the corridor. Figure 7-8 illustrates the process used to analyze the data and move towards plan selection.

**Figure 7-8 Cycle Length Selection**




As shown in the Figure, the base network for each time period is used to select a cycle length, evaluate intersection and system operations, and identify the best plan based on the objectives of the study. These tasks are described below.

### *Cycle Length Selection*

Cycle length selection should reflect local policies and users of the system. Theoretically, shorter cycle lengths result in lower delays to all potential users. However, as tradeoffs are made among the various users, cycle length selection becomes more complex. As we consider more users, the cycle length may increase; each intersection is assessed for its cycle length requirement independently to determine a minimum cycle length.

The cycle length for the subject time period is determined based on the performance measures that have been identified for the system. The analysis of cycle length is performed using either simulation or signal timing software, during which the performance of the section is evaluated for a range of different cycle lengths. The range of cycle lengths used begins with the minimum cycle length (which is determined using Webster’s equation or a similar process) up to a maximum value determined by signal timing policy and controller capabilities. Typically, the range is evaluated at two second intervals.

After the initial analysis, two to four cycle lengths should be identified based on the objectives of the study and results of the initial analysis. The selected cycle lengths should be carried forward and analyzed for both intersection and system operations.

It may be helpful to summarize some of the MOEs from the different cycle length analyses to assist in selecting cycle lengths for further evaluation. As mentioned in previous chapters, a MOE that is often used to identify acceptable operations is volume-to-capacity ratio. A volume-to-capacity ratio of 0.90 or less is typically an acceptable measure; a volume-to-capacity ratio of greater than 1.0 would be a measure of oversaturated conditions (provided the assumed saturation flow rate is accurate). Figure 7-9 and Figure 7-10 illustrate intersection and system MOEs for various cycle lengths.

**Figure 7-9 Cycle Length Analysis – Volume-to-Capacity Ratio Summary**

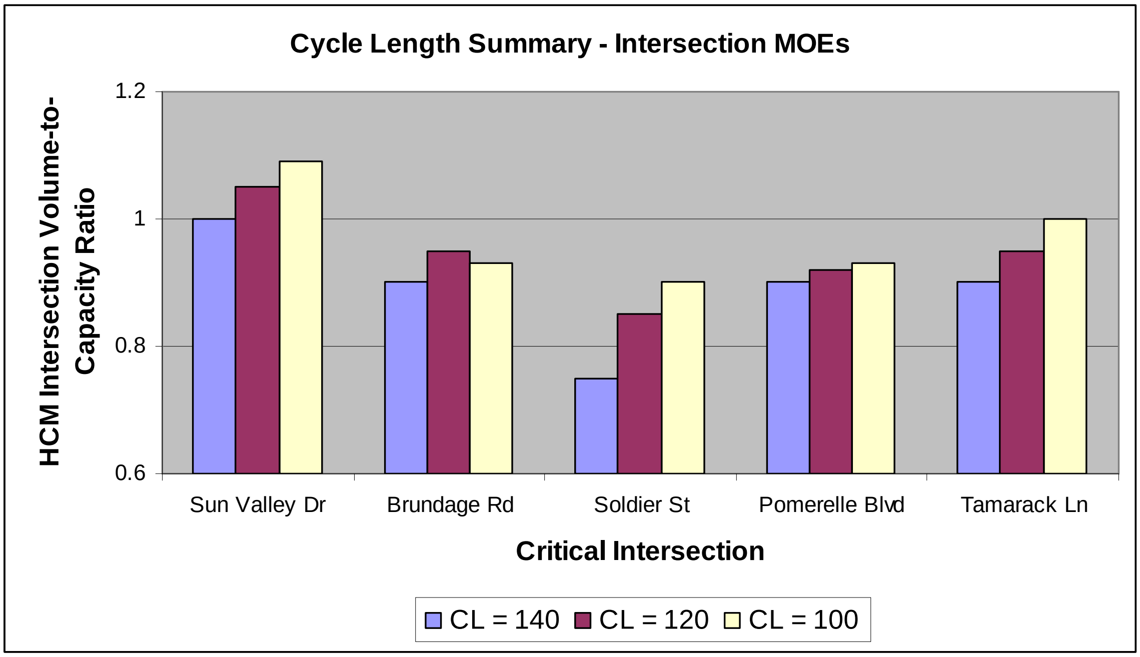
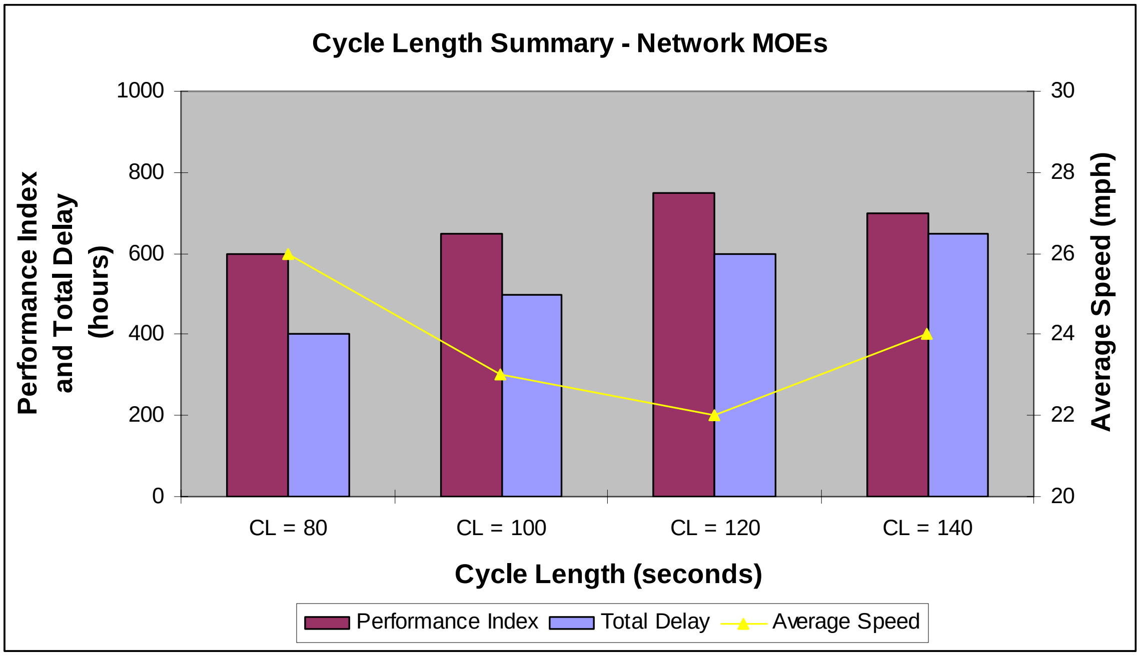



**Figure 7-10 Cycle Length Analysis – Performance Index, Delay, and Speed**



### *Cycle Length Refinement – Intersection Analysis*

Once a cycle length is selected, the volume-to-capacity ratios, movement splits, minimum splits, and vehicle queues should be evaluated at each of the subject intersections for the two to four cycle lengths identified for further refinement. This analysis allows the user to assess if the cycle length will meet specific objectives of the study. For instance, if an objective is to ensure that pedestrian timings are accommodated within each phase split, then this assessment would make sure that the phase splits accommodate the pedestrian timings. The key elements are further shown in Figure 7-11.

**Figure 7-11 Key Elements to Consider During the Intersection Analysis**

1. **Review movement volume-to-capacity ratios.**

   V/C should be less than 0.90 (if possible)

2. **Review movements splits.**

   Splits should accommodate traffic demand

3. **Review minimum splits.**

   Are pedestrian times met?

   Is there an agency policy for violating pedestrian minimums?

4. **Review vehicle queues.**

   Are the queues excessive?

Signal retiming is an iterative process that requires time to arrive at the most appropriate solution for the subject corridor. At the conclusion of the intersection analysis, the user should be able to identify one or two timing plans that would work for intersection operations.

### *Corridor Refinement - System Analysis*

After the analysis is complete, a system analysis of the subject corridor or network should be completed to provide an analysis of:

- Vehicle progression along the corridor (i.e., Is there an acceptable bandwidth?);
- Intersection-to-intersection interaction (i.e., Do vehicle queues spillback?); and
- Left-turn operations (i.e., Is there benefit to lagging a left-turn movement?).

The user should use the following steps to ensure a detailed analysis is performed on the subject corridor or network:

- Review time-space diagram and/or traffic flow profile.
- Identify locations where the through band is being cut off (refer back to the objectives of the study) or where the front of the platoon encounters a yellow indication or only a few seconds of remaining green.
- Identify locations where movements, such as platoons on the major street or left-turns from a minor street, are being stopped.
- Identify locations where excessive queuing occurs.
- Evaluate locations for lagging left-turns.
- Evaluate locations where an “early green” may impact progression either beneficially or negatively.

In this system analysis, several signal timing parameters might be changed to improve the traffic flow, reduce vehicle queues, and meet the objectives the study. Some of these parameter adjustments could include offset changes, an increase in phase splits, and/or changes to the phase sequence of left turns on the major or minor streets. Care should be taken that if lagging left turn phasing is used, the requirements of the MUTCD are met so as not to create a yellow trap condition.

After this review is complete, the information associated with the intersection and system analyses should be summarized for the different cycle lengths. The summary comparison should be referenced with the objectives of the study to ensure the proper selection of a cycle length and resulting timing plan.

***Simulation***

After the various timing plans are reviewed or the best timing plan is selected, it might be desirable to evaluate the plans through a simulation program. Depending on the budget of the study, complexity of the corridor, and the degree of congestion, a simulation tool can be used to provide a more detailed analysis of the timing plans. The simulation results can be used to determine the best timing plan for meeting the objectives of the study and to increase confidence that the timing plans will operate as designed. Simulation does not need to be used on every retiming study, but it may be useful for evaluating timing plans with over-saturated traffic conditions.

The simulation program can be used to review the signal timing by incorporating the new timing plans in the simulation program and entering the traffic conditions for which they were developed. The simulation is then used to compare the selected performance measures resulting from the new timing with the measures produced by the existing timing or other cycle lengths. In this way, the potential effectiveness of the new timing can be evaluated. The simulation outputs can be used to identify locations where manual adjustments to the timing are required due to the presence of unusually long queues or other symptoms of poor performance.

Additionally, simulation programs have animated graphical outputs that approximate aerial views of the simulated network. These animated outputs permit visual assessment of system performance, a capability that cannot be duplicated by any other form of output, including time-space diagrams. Care must be used when reviewing animation from simulation models, as these should be reflective of an average run which can be selected only after review of the results. The Traffic Analysis Toolbox (3) describes simulation as  useful for these and other applications, such as:

- To evaluate signals incorporating actuated-coordinated operations. Signal optimization programs treat these controllers as if they were pre-timed and then make suitable adjustments to approximate the operation of actuated controllers. The simulation program (if coded correctly) may provide a more realistic assessment of the effectiveness of the actuated controller within the section being evaluated.
- To confirm the likelihood of presence of queue spillback between intersections;
- To evaluate special features such as transit priority;
- To evaluate system performance in the presence of saturated conditions because these conditions are not adequately considered by signal timing optimization programs;
- To evaluate the effectiveness of manual adjustments to signal timing;
- To evaluate fuel consumption and emissions or system travel time resulting from a given set of signal timing; and,
- To demonstrate the improvements to public officials as it paints a clearer picture than numbers alone.

### 7.4.3 Draft Timing Plans

After the cycle lengths are evaluated, the intersections and system analyzed, and the timing plans compared, a preferred timing plan should be selected and carried forward to final review. This preferred timing plan is often called a draft timing plan.

In addition to the draft timing plans, a time-of-day (TOD) schedule is developed that identifies the time periods for when the draft timing plans will be in operation. Figure 7-12 and Figure 7-13 illustrate a TOD schedule superimposed on a 24-hour traffic volume plot for a typical weekday and Saturday, respectively. On the weekday schedule, combined traffic peaks in the morning and evening during rushhour. On the Saturday schedule, traffic peaks in the middle of the day.

**Figure 7-12 Weekday Time-of-Day Schedule**

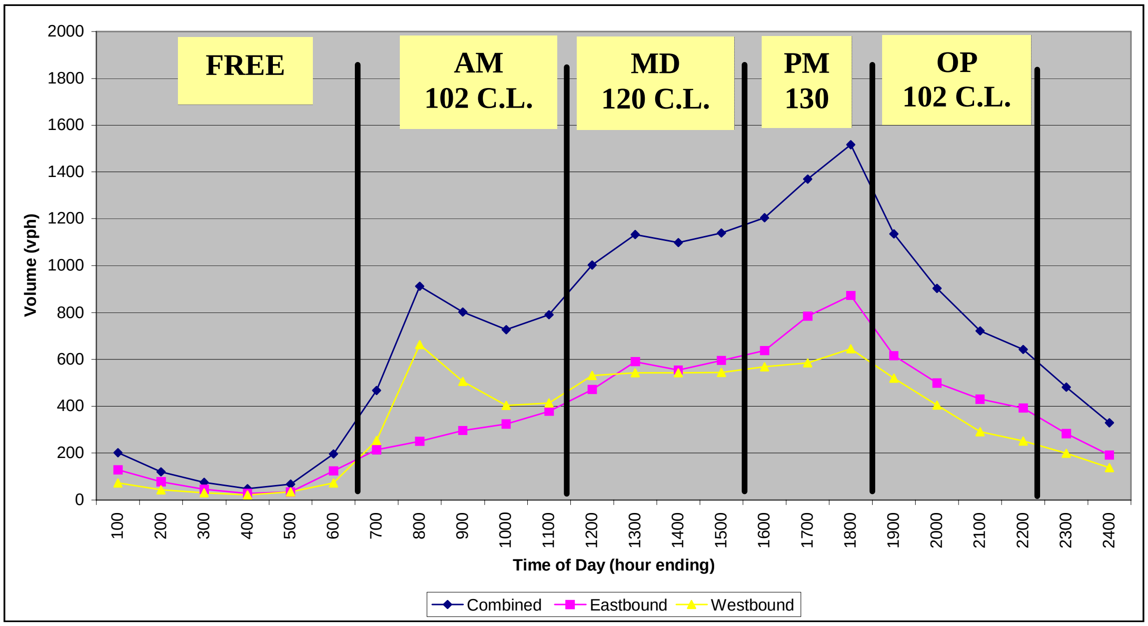
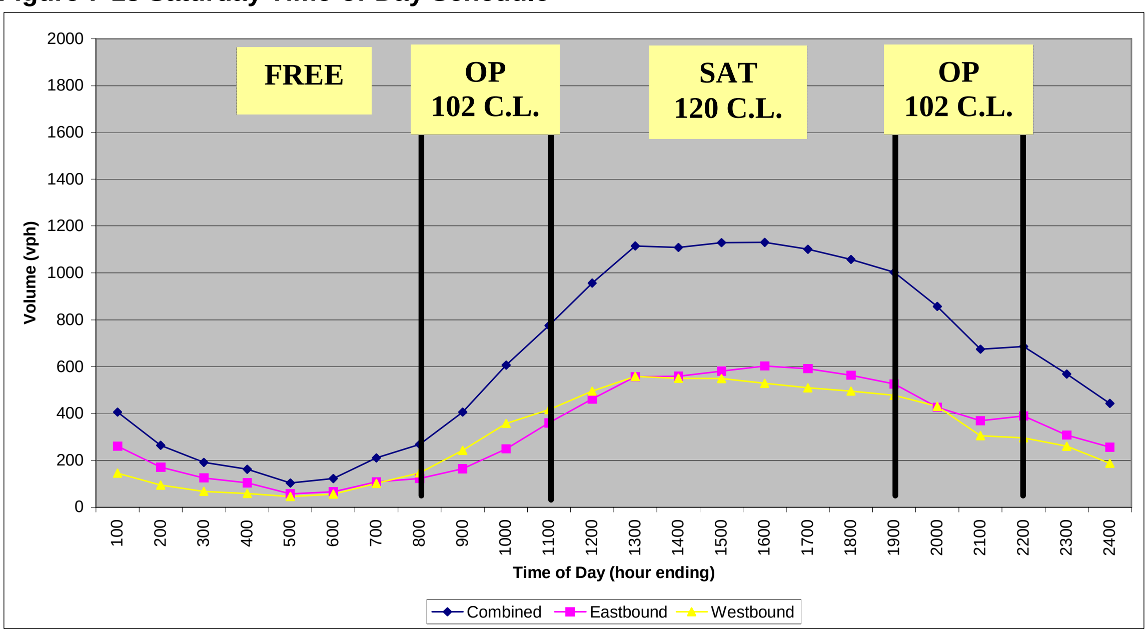



**Figure 7-13 Saturday Time-of-Day Schedule**



The TOD schedule will vary on corridors depending on the types of surrounding land uses. For instance, commercial corridors may have a low volume during the morning hours, begin to peak in the late morning, and continue to increase in demand through the evening time periods. Locations that consist of theme parks may have two peaks, one in the morning when users are arriving to the park and one in the late evening near the park closing. It is important to understand the surrounding land uses and how they impact the traffic operations throughout the day.

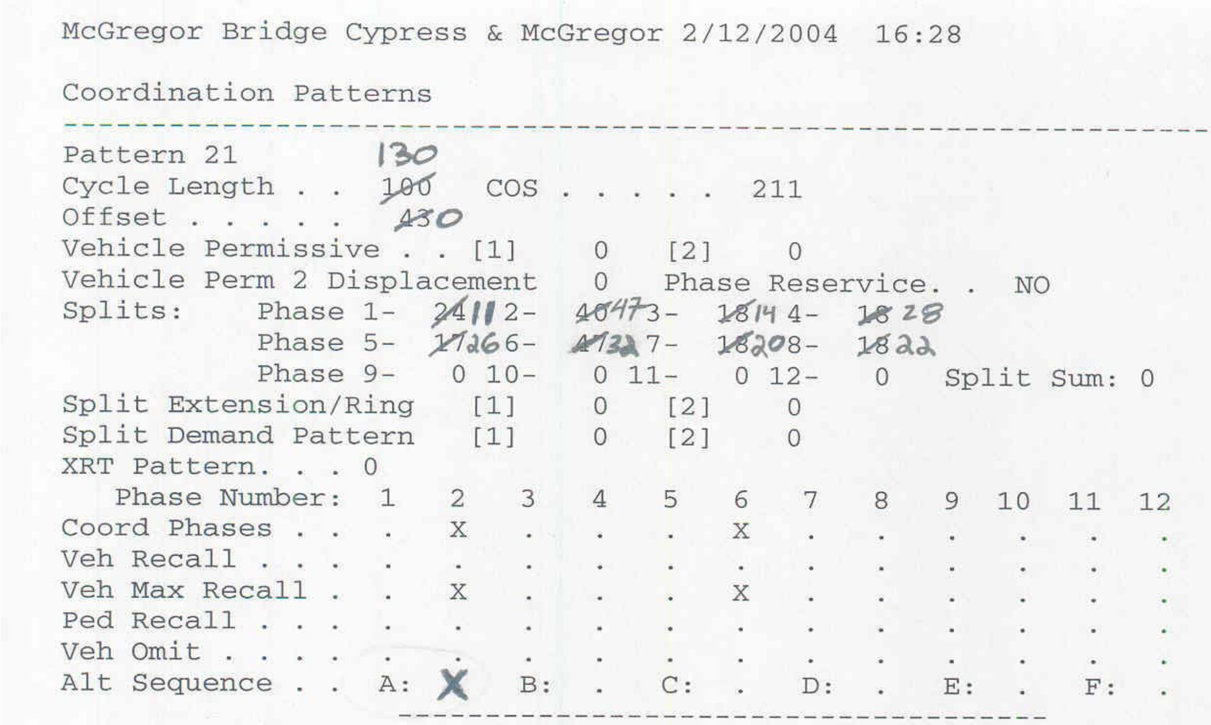
Typically, a summary of the MOEs and objectives of the study for each draft timing plan is prepared along with the draft TOD schedule and is submitted to the review agency. It is important for the parties to meet and discuss the draft timing plan to ensure that the objectives of the study are being met. After this review, the draft timing plans should be updated to reflect the final timing plans that will be used for field implementation.

### 7.4.4 Final Timing Plans

Once the review agency has completed its review and the comments have been incorporated into the draft timing plans, the timing plans are ready to be deemed final. A few steps need to be completed to prepare for field implementation:

- Identify the isolated and coordinated signal timing parameters at each intersection for each timing plan. Most software models provide a summary of this information by intersection. Typical parameters that might be modified as part of a retiming project are outlined below.
  - **Coordinated Signal Timing Parameters**
    - Cycle length
    - Splits or Force offs (depends on type of controller)
    - Offsets
    - Coordinated Phase
    - Phase Sequence (Leading or lagging left turns)
    - Vehicle Recalls
  - **Basic Signal Timing Parameters**
    - Minimum Green Time
    - Yellow and All-Red Intervals
    - Walk and Flashing Don’t Walk Intervals
  - Other Signal Parameters
    - Rest in Walk
    - Inhibit Max
    - Transition modes
- Identify the TOD schedule for the signal timing plans.
- Develop controller markups of the final timing plans and TOD schedule. Figure 7-14 illustrates an example markup for a controller.

**Figure 7-14 Controller Markup**



McGregor Bridge Cypress & McGregor 2/12/2004 16:28

Coordination Patterns

Pattern 21 130
Cycle Length . . ~~100~~ COS . . . . . . 211
Offset . . . . . ~~4~~30
Vehicle Permissive . . [1] 0 [2] 0
Vehicle Perm 2 Displacement 0 Phase Reservice. . NO
Splits: Phase 1- ~~24~~11 2- ~~40~~47 3- ~~18~~14 4- ~~18~~28
Phase 5- ~~17~~26 6- ~~47~~32 7- ~~18~~20 8- ~~18~~22
Phase 9- 0 10- 0 11- 0 12- 0 Split Sum: 0
Split Extension/Ring [1] 0 [2] 0
Split Demand Pattern [1] 0 [2] 0
XRT Pattern. . . 0

| Phase Number: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coord Phases . . | . | X | . | . | . | X | . | . | . | . | . | . |
| Veh Recall . . . | . | . | . | . | . | . | . | . | . | . | . | . |
| Veh Max Recall . | . | X | . | . | . | X | . | . | . | . | . | . |
| Ped Recall . . . | . | . | . | . | . | . | . | . | . | . | . | . |
| Veh Omit . . . . | . | . | . | . | . | . | . | . | . | . | . | . |

Alt Sequence . . A: X B: . C: . D: . E: . F: .

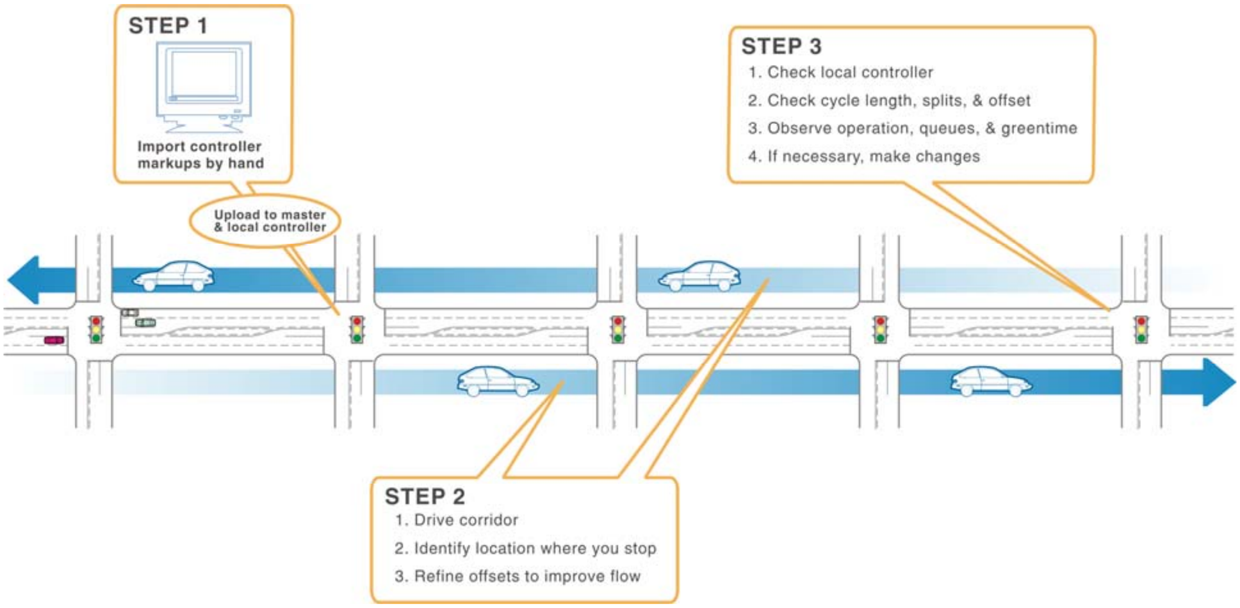
Once the draft controller markups are completed, a detailed review of this information should be performed to ensure that the correct timings have been marked up on the controller sheet. Additionally, this review may identify other existing timing parameters that are unique for a certain intersection that might need to be modified to ensure operation of the new timings. Depending on the agency, the final controller markups should be used to update the electronic files for each controller. Again, a detailed review of each electronic file should be performed within the agency or prior to sending the final files to the agency and being implemented into the controller. This process assists with minimizing the number of data entry errors during the process and allows for adequate preparation for field implementation. Finally, if yellow and/or red settings are being altered or if phasing is being revised, a formal approval by the agency's traffic signal engineer may be required before the changes can be implemented.

## 7.5 FIELD IMPLEMENTATION AND FINE TUNING

Once the final timing plans have been completed and the controller markups have been prepared, the plans are ready to be implemented and observed in the field. Field implementation is the most critical part of the signal retiming process. Both science and finesse are needed to fully realize a good timing plan in the field. However, care should be taken to not draw an erroneous conclusion for a snapshot of time while on the corridor. Cycle-by-cycle variations occur, and one must be patient during observations. It is relatively difficult to see an offset error in the field when upstream intersections are releasing traffic early. Figure 7-15 illustrates the typical steps used during field implementation.

**To assist with field implementation, a field notebook should be generated that includes hard-copies of the existing and proposed signal timing plans, time-space diagrams, traffic volumes, and a method for noting changes in the field.**

**Figure 7-15 Field Implementation Process**




As highlighted in Figure 7-145, the final signal timing plans are uploaded from the traffic control center to the master controller (if applicable), where this information is transferred to each individual controller. Once this process has occurred, the engineers/technicians should drive the subject corridor or network to verify that the new timings are operational at each intersection. Observations of the intersection operations (i.e., Is the intersection operating the correct cycle length, phasing sequence, splits, etc.?) include side-street delay, major street left-turn delay, and review of vehicle queuing to ensure that the cycle length and movement green time are adequately dispersed for the traffic demand. If new signal timings are developed and implemented for the pedestrian Walk and Flashing Don't Walk times, the engineer/technician performing the field implementation should check all of these times at each intersection to ensure that the crossing times are adequate and entered correctly in the controller database.

After visiting each intersection and confirming adequate operations, the team should drive the corridor to review vehicle progression and determine if any changes to the offsets or left-turn phase sequence is necessary. Depending on the history of the corridor and the number of proposed timing plans, field implementation and observations of the timings should consist of a minimum of three days (two days for weekday operations). If necessary, more days in the field should be added to ensure that the signal timings are operating acceptably.

Any changes that are made in the field should be documented in a field notebook and in the cabinet notebook at each intersection. These changes will be incorporated in the model network for developing as-builts of the implemented signal timing plans.

Once the signal timing is installed in a section, field observers or closed circuit television cameras should be positioned so that the entire system can be monitored. This activity will ensure that major problems, such as data entry errors, are quickly identified and corrected. Once the system appears to be operating correctly, it is necessary to identify the need for additional fine-tuning. This can be done through measurement of stops and delays at important intersections using the process of counting standing queues at intersections at regular intervals . As a final step, travel times should be evaluated using floating cars driven over the same routes that were followed during the evaluation of the "before" signal timing so that a comparison can be made.

**If feasible, observing the traffic operations simultaneously both in the field and via a traffic management center can provide an efficient way to identify problem areas along the corridor due to short split times, excessive queuing, cycle failure, and/or hardware issues.**

## 7.6 EVALUATION OF TIMING

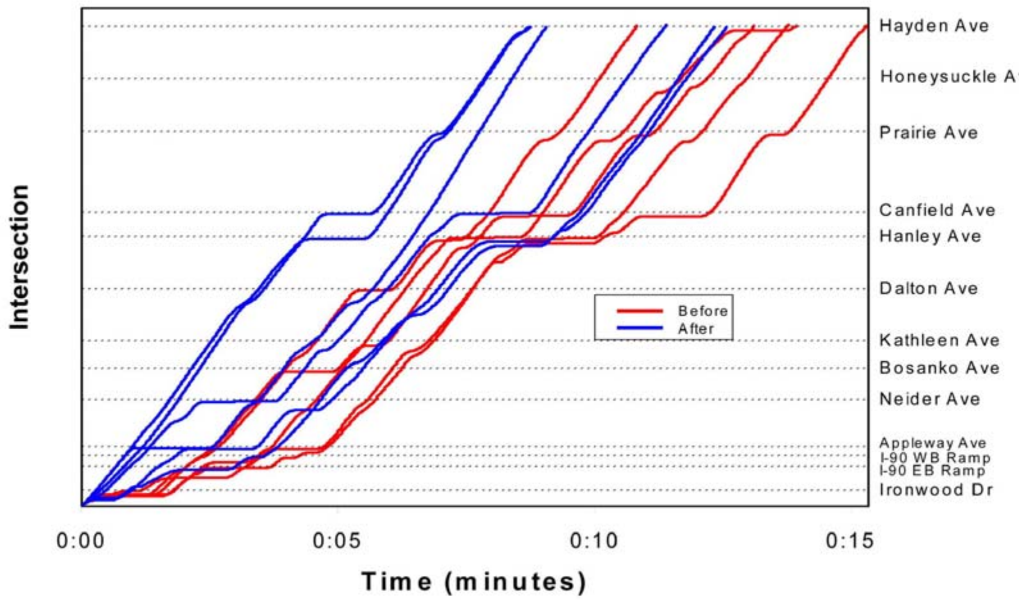
With any type of project that is implemented in the field, the engineer should evaluate the performance of that project once it is in operation. Beyond the evaluation tools, a signal retiming project is no different. Some evaluation techniques include:

- Observing the traffic flow along the corridor and vehicle queues at individual intersections;
- Conducting "after" travel time and/or delay studies for the study network;
- Fielding phone calls and listening to the public;
- Verifying the new traffic operations are consistent with the agency policy and objectives of the study; and
- Comparing "before" and "after" measures of effectiveness from the analysis model and data collection.

A travel time and delay study is a common tool applied to most signal retiming projects. A disadvantage is that the travel time and delay information collected during the study is commonly limited to the through movements for the major route of the corridor and may not identify possible complaints from pedestrians or motorists crossing the arterial. Generally, if the objective of the study is to reduce stops and travel time along the corridor, a travel time and delay study is appropriate with observations of queue lengths on the side street approaches.

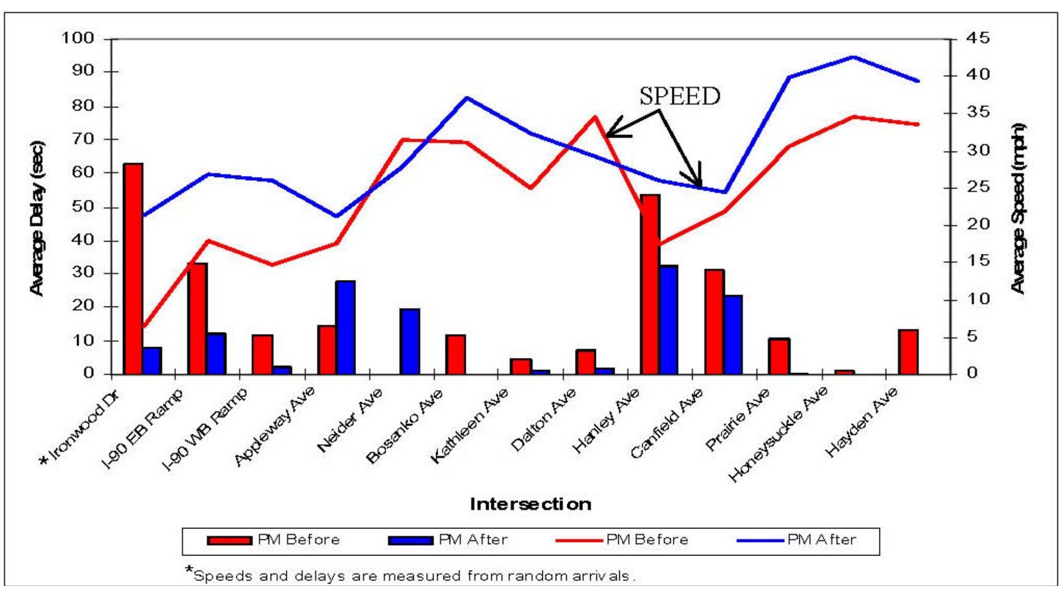
One of the summaries that can be generated from the study is a before-and-after projectile of vehicles along the corridor (i.e., a time-space diagram). The time-space diagram highlights the locations on the corridor that experience stops and long delays under the before and after conditions. Figure 7-16 illustrates a typical time-space diagram from a before-and-after study.

**Figure 7-16 Time Space Diagram (Before-and-After Study)**




Another set of measures of effectiveness include delay and speed, which are typically used to evaluate the success of the project. Figure 7-17 illustrates a before-and-after summary of delay and speed for the corridor.

**Figure 7-17 Delay and Speed Summary (Before-and-After Study)**




### 7.6.1 Performance Measurement

It is critical to collect performance measures at an intersection to determine how well that intersection operates and serves the public. Reviewing and updating the intersection-specific timing and operational aspects of individual signalized intersections on a regular basis is extremely important, especially where changes in traffic volumes and/or adjacent land uses have occurred.

The most common performance measures used to describe signalized intersection operation are outlined in the *Highway Capacity Manual* (HCM). The methods used are vehicle oriented and focus on collecting manual data that is costly to gather. In most cases, it is not feasible for agencies to schedule manual data collection activities for unusual periods, such as evening school events or weekend peaks at shopping malls that do not allow timing plans to be developed to address all situations. Consequently, it is desirable to have signal systems collect their own data to assess performance through an integrated general purpose data collection module within the controller (4). However, obtaining meaningful real-time performance data from signalized intersections has historically been difficult because detection systems are not designed to provide good count information. Data is summarized in averages that may not be calibrated, and access to the information is limited to personnel managing the signal system.

### 7.6.2 Policy Confirmation and Reporting

After the data collection, signal timing analysis, field implementation, and evaluation are complete, there are two important final tasks that are needed: to confirm that the established policies were met and to prepare a final report.

The final report is an opportunity to confirm that the objectives of the retiming effort were met. Some agencies use this step in the process to inform elected leaders about the success of the project and to document future opportunities. A final report documenting the before-and-after study is typical and may be posted on the agency websites and/or passed along to news agencies for the public to view. However, more time spent developing a final report results in higher project costs and can reduce the time for the signal timing analysis, field implementation, and fine tuning efforts. During the scope of work process, it is important to determine the type of report necessary to assist with future retiming efforts, but this report should not take away time spent for the signal timing analysis and field implementation.

A report that is focused on the before-and-after study might include the following items:

- Introduction
- Travel Time and Delay Study Methodology and Analysis
- Travel Time and Delay Study Summary
- Savings Methodology and Analysis
- Savings Summary
- Summary

A report that is focused on the entire study might include the following items:

- Executive Summary
- Introduction
- Data Collection
- Signal Timing Development
- Timing Plan Summary
- Field Implementation
- Before and After Study
- Conclusions

The signal retiming process includes many steps to develop a successful set of signal timing plans. To promote a successful retiming project, it is critical to develop a comprehensive scope of work that is consistent with the agency objectives and local policy.

## 7.7 REFERENCES

1 National Traffic Signal Report Card, National Transportation Operations Coalition, 2005.
2 Manual of Transportation Engineering Studies, Institute of Transportation Engineers, 2000.
3Traffic Analysis Toolbox, Volume I: Traffic Analysis Tools Primer, Federal Highway Administration, June 2004.
4 Smaglik, Edward, Sharma, A., Bullock, B. Sturdevant, J., and Duncan, G., “Event Based Data Collection for Generating Actuated Controller Performance Measures”, TRB Paper 07-1094, November 6, 2006.

# CHAPTER 8

# SIGNAL TIMING MAINTENANCE: OPERATIONS AND MONITORING

## TABLE OF CONTENTS

# LIST OF TABLES

# 8. SIGNAL TIMING MAINTENANCE: OPERATIONS AND MONITORING

The purpose of this chapter is to summarize the various steps necessary in maintaining effective traffic signal timing plans. It will consider the various field settings that are important for response to citizen inquiries and will identify ways to address day-to-day operations of the signal system.

This chapter contains four sections. The first section presents a short overview of the many activities that correspond to the maintenance of signal timing. The second section presents a discussion of the types of activities that are completed by agencies. It will describe methods used to identify changes in the street network and use of traffic signal systems of various sizes. The third section will present a checklist of typical events (public complaints, weather, etc.) and the range of possible responses, with cross-references to the appropriate chapters in the manual. This section will also present a series of common questions and answers that can assist personnel involved in the direct line of communication with the public. The final section will highlight issues raised during the ITE traffic signal self assessment and communicate recommended staffing levels for public agency jurisdictions.

## 8.1 OVERVIEW

As discussed in this manual, traffic signal timing is one component of a traffic signal. As with any component of a traffic signal, a maintenance element is important to ensuring that the traffic signal will continue to operate at the level expected by the agency policies and the general public. For a traffic signal, maintenance activities range from system-oriented, such as managing a Traffic Management Center (TMC) or providing training opportunities for staff, to local-oriented activities, such as inspection and replacement of traffic signals, controllers, and detectors and reviewing traffic signal timings on a periodic basis, to other activities, such as public relations or collecting traffic data at the intersection. All of these activities are important ones to ensure acceptable operations of the signal system.

Table 8-1 highlights some of the many maintenance activities of a traffic signal system. Cells that are highlighted represent the activities specific to the maintenance of signal timing.

**Table 8-1 Traffic Signal System Maintenance Activities**

| System-Oriented Activities | Local Activities | Other Activities |
|---|---|---|
| Personnel Management | Maintaining Traffic Signal Operations and Inspection | Public Relations, Education, and Outreach |
| Traffic Management Center | Utilizing and Maintaining ITS Devices Operations and Inspection | Traffic Data Collection and Inventory |
| Equipment Inventory and Training | Updating Traffic Signal Timing | Utilizing Website and Information Dissemination Outlets |
| Updating Software Inventory and Training | Modifying Traffic Signal Design for Construction Projects | |
| Maintenance Inventory and Management | Maintaining Detector Operations and Inspection | |
| Traffic Signal System Design and Operations Training | Traffic Signal Turn-On's | |
| Updating and Maintaining Communication Design and Inventory | | |
| Developing and Maintaining Inter-Agency Coordination | | |

## 8.2 TRAFFIC SIGNAL OPERATIONS

This section presents a discussion of the types of activities that are completed by agencies. It will describe methods used to identify changes in the street network and use of traffic signal systems of various sizes.

### 8.2.1 Signal Timing Maintenance Activities

As presented above, signal operators have several maintenance activities to ensure that traffic signals are functioning properly. Some of these activities related specifically to traffic signal timing are outlined below:

- Retiming of traffic signals due to the following:
    - Land use changes
    - Population growth
    - Change in flow profiles (volume and classification)
    - Incident management
    - Special Events
    - Traffic signal turn-on
    - Construction work zone or temporary traffic signal
    - Traffic signal equipment change
    - Scheduled or periodic traffic signal retiming
    - High frequency or rate of crashes
- Inventory of signal timing at each traffic signal
- Maintaining a database of the traffic signals, including signal design, signal timing, and history of updates
- Staff training for use of software and equipment
- Responding to public comments
- Observing traffic conditions via a Traffic Management Center or on-site field visits
- Coordination between designers, operators, and technicians

### 8.2.2 Reasons for Signal Timing Maintenance

It is not possible to retime an entire group of signals every time adjustments are being considered for a single intersection. This is true, even though the intersection requiring adjustment may be part of a coordinated system. When making adjustments to a single intersection, the challenge is to do so without making things worse at neighboring intersections. In other words, the adjustments should have a positive impact on the intersection at which they are made, leading to the improved performance of the overall system.

The adjustments being made are sometimes known as retiming. Retiming may involve modifying the phasing or mode of operation to accommodate unusual congestion, adjusting the split to reflect a change in demands at the intersection, or changing an offset or green time to accommodate an incident in the field. In some cases, the offset may be modified to reflect changes in travel behavior from adjacent intersections.

The need for retiming may result from the professional judgment of the jurisdiction’s engineering staff, an incident, a new traffic signal, or the result of citizen’s calls. In all cases, the request for retiming should be taken seriously, since it reflects the observation of either trained observers or frustrated motorists, field incidents, or construction.

**Signal Retiming:** The need for review and adjustment of signal timing at a single intersection could be the result of a variety of different factors including:

- Changes in traffic demand since the intersection was last timed. This could include changes in side street demand, turning movement volume or spill back, main street demand, or vehicle mix (for example a higher percentage of trucks). Changes in vehicle demand could also be reflected in general increases in demand that cause the need for longer periods with peak period timing, and the modification of night time flash operations.
- Changes in intersection operations (for example addition of an approach lane or the moving a bus stop from near side to far side) that influence the need for timing.
- Changes in pedestrian traffic due to land use changes (for example the opening of a residence for the elderly which required longer pedestrian clearance times) or the need for handicapped features.
- Changes to agency policies or national standards, such as the Manual on Uniform Traffic Control Devices.
- Temporary changes in roadway operations due to construction
- Observations of previously unnoticed conditions by an alert motorist or staff member or through use of a Traffic Management Center
- Agreements with other jurisdictions to coordinate with their signal systems, or to provide coordinated response to incidents on parallel facilities

From the length of this list, it is clear that, in most systems, signal retiming will be required frequently at various intersections throughout the system. It is important to recognize that there is a point at which so many localized adjustments have been made, that it becomes essential to initiate system-wide signal retiming as described in Chapter 7.

**Traffic Signal Inventory:** One of the items related to signal timing maintenance is the importance of maintaining an inventory of the traffic signals. The database should include information related to the traffic signal, such as location, signal layout, signal timing, coordinated or uncoordinated signal operation, communication, operating agency, history of updates, etc. A database provides the agency with a knowledge base of what changes have occurred and what might need to be updated in the near and long term at the signalized intersection.

**Staff Training:** The equipment and software utilized by many agencies on its traffic signal system is only as good as the availability of skilled and trained staff. Therefore, a valuable component of the traffic signal maintenance is ensuring that the staff managing and maintaining the traffic signals has been trained to operate the system. Critical training elements highlighted in the FHWA Guidelines for Transportation Management Systems: Maintenance Concept and Plans (1) included:

- Training by Vendors: Procurement contracts should include a requirement for on-site training of Agency staff in maintenance and operations of the equipment, preferably conducted by the vendor.
- Training by Contractors: Procurement contracts should also include a requirement for on-site training of Agency staff in the maintenance and operation of the assembled systems, including software, hardware, and devices.
- Training Library: The operating Agency should maintain a library of system documentation and, if available, a videotape or DVD library of training.
- Staff Retention: This can be difficult in a high-tech environment, but there are ways to improve retention, such as providing for additional training, allowing travel to technical conferences, and workshops and other non-salary related perquisites for Agency staff.

As noted in these guidelines, the key to success in this ever-changing environment is flexibility and a good understanding of priorities for both operational and maintenance concepts, requirements, and training.

**Incident Management:** Traffic patterns for special events, roadway construction, inclement weather, crashes, etc. are different from those that exist during normal traffic conditions. It is important for agencies to develop and utilize an incident management plan for its signal system. Goals of these plans related to signal timing might include sustaining or increasing corridor capacity during an event through longer green times at a traffic signal, enhancing public safety though modified time settings could reduce delays for emergency responders traveling to an incident on the roadway, and guiding motorists to a certain destination. Chapter 9 provides further discussion on this topic.

**Public Comments:** Calls from the public are one of the most common reasons for reviewing intersection operations. It is important that each agency has a process in place to field public phone calls, emails, etc. and addresses their responses in a timely, professional manner. More discussion on this item is presented in the next section.

## 8.3 DAY-TO-DAY OPERATIONS

This section will present a checklist of typical events (public complaints, weather, etc.) and the range of possible responses, with cross-references to the appropriate chapters in the manual. This section will also present a series of common questions and answers that can assist personnel involved in the direct line of communication with the public.

### 8.3.1 Signal Retiming

Since the need for retiming could have been identified by a number of different sources for a variety of different reasons, it is difficult to define a single procedure that might be applicable to all possible sets of conditions. However, in most cases, the following procedure should be followed, with all actions recorded:

- Schedule the field visit for the time-of-day and type-of-day (weekend, weekday, etc.) for which the problem was identified.
- Assemble the timing and configuration information for the intersection being visited. If the intersection is included in a system, timing information should include master clock, offsets, and time-of-day schedules for plan changes. Information should also include controller settings. If available, traffic count data should also be included.
- If the intersection is included in a system, coordinate with system operators to ensure that operations personnel will be available by radio contact to support the field activities.
- When arriving at the intersection, observe the physical condition of the street hardware including poles, mast arms or span wires, signal head positioning, signal lamp operation, pedestrian indication operation, and cabinet condition.
- Open the cabinet and perform a physical inspection of the cabinet interior including cabling, physical condition and operation of cabinet components, air filter and fan.
- Check operability of all cabinet components either through observation or suitable maintenance diagnostics.
- Review controller timing by comparing settings with timing documentation
- Qualitatively compare traffic conditions at the intersection with the traffic count data. Determine whether major changes in demand have occurred since the traffic counts were taken to support the development of the timing plans currently in use. If major changes have occurred, determine whether they are temporary (due to nearby construction) or permanent. If they are temporary, it is still desirable to fine-tune the

intersection timing.  However the log completed as the final step should include a notation that a second set of fine-tuning may be required when the construction has been completed.

- If the deficient intersection operation occurs during heavy traffic conditions, the next step would be to determine whether the adjustments are intended for congested or under-saturated conditions.  The procedures should be followed from the appropriate section below.
- If the deficient intersection operation occurs during light traffic conditions, the procedures of the section for “Other Types of Traffic Conditions” should be used.
- If the deficient intersection operation is related to the use of an incorrect timing plan, (assuming a correct plan is available), the procedures associated with the selection and scheduling of plans under “Other Types of Traffic Conditions” should be used.
- In all cases, after timing and/or scheduling changes have been made, the impact of the changes should be evaluated through observation of the intersection operation.
- The final step of this process is to log the actions taken.  This is essential for response to the individual initiating the fine-tuning, as well as records that must be maintained by the agency for a variety of engineering, operational and legal reasons.

Refer to Chapters 2, 5, 6, and 7 for more details on retiming at an intersection or a series of intersections.

**Retiming for Under-Saturated Conditions:** It is likely that the most frequent requests for intersection fine-tuning will occur during normal flow conditions.  These are the conditions that impact a large number of motorists, all of which expect high signal timing quality.  Motorists will be annoyed at instances of wasted green time during which they have to wait at a red signal indication when there are no vehicles on other phases.  Motorists will also become annoyed at locations where they have to stop at successive signals due to poor offsets.  In some cases their complaints are unjustified, since they are not aware of other constraints on signal timing, such as pedestrian clearance times and the need to make compromises in progression in order to accommodate flow in the reverse direction.  However, in other cases, their complaints are justified, and should be addressed whenever possible by the agency’s engineering and technician staff.

When retiming an intersection for free-flow (under-saturated) conditions, the following steps are recommended:

- Perform a qualitative evaluation of the intersection performance to determine whether any obvious improvements are possible.
- Adjust the split to reflect demand on competing approaches.
- Adjust the offset to reflect platoon arrival times.
- Review the cycle length for possible improvements.

**Retiming for Congested Conditions:** It is important to recognize different traffic signal timing strategies for networks that experience congested traffic conditions or more normal traffic flow conditions.  Strategies begin to change from mobility and progression to queue management.  Congestion can be recognized by the presence of queues at signalized intersections that are not completely discharged during the green period.  This is known as cycle failure. If the deficient intersection operation occurs and consider measures such as:

- during heavy traffic conditions, the next step would be to determine whether the adjustments are intended for congested or under-saturated conditions.  The procedures should be followed from the appropriate section below.

- during light traffic conditions, the procedures of the section for “Other Types of Traffic Conditions” should be used.
- related to the use of an incorrect timing plan, (assuming a correct plan is available), the procedures associated with the selection and scheduling of plans under “Other Types of Traffic Conditions” should be used.
- In all cases, after timing and/or scheduling changes have been made, the impact of the changes should be evaluated through observation of the intersection operation.

When re-timing a congested intersection, the split and cycle length should be reviewed and adjusted as necessary. When an intersection experiences congested conditions, there may not be a need to consider offset adjustments, except for the qualitative evaluation of intersection operations, since continued traffic flow between intersections may not possible. The following process is should be used:

- Perform a qualitative evaluation of intersection operation.
    - Adjust the downstream intersection signal offset to permit earlier discharge of left turning vehicles, to minimize the possibility that these vehicles might block through movements.
    - Change the signal phasing as needed to avoid spillback from left turn bays, or to make more effective use of concurrent phases.
    - Modify pedestrian phasing to minimize pedestrian/vehicle conflicts
- Review and adjust the split as required.
- Review and adjust the cycle length if necessary.
- If these approaches prove ineffective, consider using some of the “other measures” described below.
    - Driveways near the intersection which potentially block the through flows, and which may include weaving movements might be eliminated
    - Parking can be banned
    - Double parking regulations might be rigorously enforced
    - Bus stops might be moved from near-side to far side. This is particularly useful in the presence of heavy right turning traffic
    - The number of phases might be reduced. For example, some jurisdictions ban left turns at critical intersections to increase available capacity for through movements
    - Phasing might be changed to avoid the problem of left turning vehicle blocking through movements. This might be accomplished by substituting leading left turn phases for lagging phases, or in some cases both leading and lagging phases might be provided.
    - Increasing the length of turning bays can reduce the problem of left turn blockages
    - Other geometric improvements are possible including addition of lanes and overpasses

If all else fails, metering traffic entering the control area should be considered. Metering is performed by reducing the number of vehicles traveling on the congested arterial, through significant reduction of green times available to entering side street traffic. The use of this technique avoids spillback into upstream intersections, intersection blockage and overflow of turning bays.

### 8.3.2 Signal Timing Inventory

A short discussion was provided in Chapter 7 regarding this topic. Area-wide or corridor-focused signal retiming projects include a lot of information, such as traffic data, intersection geometry, signal phasing, signal timing, controller type, and posted speeds. Thus, it is important that this information be managed in a database to assist with a record as well as for future retiming updates. Documentation of the goals and outcome of the retiming process is also important. The following is a list of items that are typically collected in a signal retiming project:

- Types of signal control at each intersection
- Intersection signal phasing
- Signal timing plans
- Measures of effectiveness (both estimated and evaluated)
- Traffic volumes and speeds for the network
- Geometric description of the network
- Input data files for simulation and traffic signal timing software
- Software files that include network and traffic signal timing for the system

The signal timing process is greatly simplified, and its cost reduced, when this data is stored in a database that can be accessed each time area-wide signal timing is developed. However, as with any database, its value will be reduced if it is not kept up to date. It is essential that the database be updated each time controller operation is modified, including all changes in phasing, type of control, timing, or intersection geometrics. In this way the value of the database can be preserved for access during the area-wide signal timing process.

Furthermore, the above information relates to maintaining a database from a signal retiming project. Additional information can be added based on the following items to create a comprehensive database of the traffic signal system.

- Signal equipment failures and changes
- Public comments
- History of signal timing changes
- Crash data
- Other incidents reported near the intersection
- Changes in land use

By recording the above information in a comprehensive database, the agency will improve its effectiveness on managing a traffic signal system.

### 8.3.3 Staff Training

A valuable component of the traffic signal maintenance is ensuring that the staff managing and maintaining the traffic signals has been trained to operate the system. Depending on the size of the traffic signal system and number of staff, a regional or local training program may be established to provide training opportunities for staff.

Training activities could be as simple as peer exchanges within or between agencies or offering a series of technical sessions. These technical sessions could consist of bringing in outside experts to discuss important signal timing, software, and maintenance topics and offered on a regular monthly basis. Other topics might be directed towards TMCs and specific signal timing projects. Other training opportunities might be provided through attendance at conferences, education seminars, or universities.

Lastly, a public relation component that incorporates training might be offering short work sessions or seminars open to the public or other agency officials that would provide insight into signal timing and traffic operations.

### 8.3.4 Responding to Citizen Calls

Citizen's calls and emails are one of the most common reasons for reviewing intersection operations. The public may obtain the contact information for emailing or making a phone via a newspaper, website, or television/radio show. Some agencies include a sticker with the logo, phone number, and catch phrase on the outside of the traffic signal controller cabinet to assist the public with obtaining the correct contact information. The call could have been initiated for a number of reasons including:

- Lack of understanding of intersection and controller operations
- A signal that was in transition between two different timing plans
- An equipment failure
- A legitimate observation regarding a shortcoming in the existing timing
- An incident near or at the intersection that impacts the traffic operations

Motorists often have a surprisingly sophisticated understanding of intersection operations resulting from their familiarity with a given roadway. For this reason, as well as for possible reasons of safety, their calls should be taken very seriously. A well-managed signal operations organization will employ the following procedures in response to citizen's calls which could arrive by telephone, email or letter:

1. Identify the name and contact information of the caller.
2. Identify the location that is the subject of the contact.
3. Define the time-of-day for which the problem is being described.
4. Ask for a description of the problem in terms of traffic conditions and traffic signal operations.
5. Assure the caller that the problem will be investigated within a predefined number of days that has been established by agency policy.
6. Enter all information provided along with the time and date of the contact into a database.

Investigation of the problem should be scheduled as part of the agency's maintenance operations. If adjustments to the signal operation are required, the procedures described later in this chapter should be followed. In all cases, the results of the investigation should be recorded in a database and described in a response to the caller using the same media (telephone, email or mail) that was used to make the original contact. If no change was made, the reason for maintaining the status quo should also be explained.

**Some jurisdictions have websites that permit citizens to report problems and concerns using an automated questionnaire that guides them through the process of providing the information described above.**

A website may be used to record input and provide an estimate of response time, and enters the information provided into a database. This is a particularly effective technique in regions where multiple jurisdictions are involved with signal operations and maintenance. In this case, callers are uncertain of the agency to be contacted. A website could be established for the region, which directs problems and complaints to the appropriate agency without requiring citizens to determine the responsible agency. Examples of such websites include the Cities of Scottsdale, Arizona, Durham, North Carolina, and Tampa Bay, Florida.

While some agencies utilize sophisticated call-processing software which handles the database functions described here, this capability is not necessarily required. Smaller agencies can use simple spreadsheets to keep track of the disposition of citizen's calls. In either case, it is critical to ensure that all calls are investigated, and that a response is provided to the caller in a timely manner. Ideally, a response should be received by the caller within one week of the date that the initial contact has been made. Responsive service is the key to good customer relations.

### 8.3.5 Incident Management and Planned Special Events

Signal timing can play a role in managing and even mitigating certain types of non-recurring congestion. In particular, the high volumes of traffic generated by planned special events, the reduction of corridor capacity from roadway incidents, or the increased travel demand triggered by region wide evacuations can necessitate signal timing changes. A primary goal of readjusting signal timing in these circumstances would be to give priority to specified movements and to minimize the overall delay experienced by users from the non-recurring congestion. One way to achieve this objective would be to sustain and/or increase the throughput of traffic at certain intersections by increasing the green time for those movements. Traffic signals with modified timings settings perform this function by essentially "flushing" the preferred movement.

The techniques to modify signal timing during planned special events, roadway incidents, or evacuations involve processes not just in traffic operations and planning, but also require some management and coordination at the policy and institutional levels. On the operations/planning side, an initial step would be to determine the specific route and intersections where traffic signals would be retimed. This route could be a particular arterial that is parallel to a certain freeway. In the event of an incident on the freeway, traffic could then be diverted from the freeway to this arterial, which could provide additional capacity with the modified timing settings in place at its various traffic signals.

Adjustments to the cycle length and green time for a particular movement are typical components of an incident management plan. Additionally, utilization of a traffic management center would be required for monitoring of the intersections during these changes to ensure that the operations are working, as well as, if necessary, permits the use of manual control by an operator.

These operational procedures may not be effective or possible without a sufficient level of coordination among the jurisdictions impacted by the traffic from planned special events or emergency situations. This inter-jurisdictional coordination is needed among such institutions as law enforcement, public safety organizations, and various transportation/transit agencies to share resources, seamlessly exchange the required information, and to implement the required traffic control/signal timing plans. With such effort to coordinate between many stakeholders, it would be necessary to develop various response and contingency plans ahead of time and updated on a regular basis. Overall, from achieving both the non-operational and operational sets of goals, adjusting signal timing can offer significant benefits to eventually reduce delays for motorists during special events, roadway incidents, or evacuations. More discussion is provided on this topic in Chapter 9.

## 8.4 STAFFING NEEDS

This section will highlight issues raised during the ITE traffic signal self assessment and communicate recommended staffing levels for public agency jurisdictions.

### 8.4.1 Background Information

This section provides some background information regarding staffing needs that are included in several literature documents.

**ITE Traffic Engineering Handbook and Traffic Control System Operations: Installation, Management and Maintenance Manual:** These documents suggest labor requirements of 20 to 25 hours per intersection for traffic signal retiming and estimates as a rule of thumb that one traffic engineer is needed to properly operate and maintain every 75 to 100 signals and one technician to operate and maintain every 40 to 50 signals. As a rule of thumb these estimates are adequate; however the current transportation environment requires much more detailed estimates.

**Traffic Signal Operations and Maintenance Staffing and Resource Requirements Guidelines:** These guidelines are part of a new project put forth by the FHWA. The objective of this effort is to develop a guideline to assist agencies in developing a staffing and resource plan to effectively operate and maintain traffic signal systems. The background information provided as part of this project suggests that the current guidance available to agencies is very general and not achievable for most jurisdictions. The lack of a credible guideline for traffic signal operations and maintenance staffing and resource needs is one of the factors that has resulted in the inefficient operation and maintenance of traffic signals on a national scale.

### 8.4.2 Staff Positions and Roles

An agency may need a variety of staff positions and roles to adequately operate and maintain its traffic signal system. Depending on the size of the signal system, some of these positions may be combined due to a combination of small signal system and limited funding available. The roles of each position described below are based on information from agencies in addition to relevant ITE and FHWA literature. Some of the positions and respective roles are.

**Traffic Signal Engineer** - This staff person is responsibly for the day-to-day operations of the signal system. Tasks include the following: Responding to public comments, approving new signal turn-on's, assisting in the TMC, evaluating signal timing on existing arterials, managing signal operations staff and coordinating with the signal design and maintenance supervisors.

**Traffic Signal Technician/Analyst** - Staff assist the Traffic Signal Engineer with their day-to-day operations. Focus areas include signal timing, new signals, and the TMC.

**ITS Engineer** - This staff person is responsibly for the implementation of ITS projects. Tasks include the following: Responding to public comments, evaluating new products, assisting in the TMC, managing ITS contractors and vendors and coordinating with the signal design and maintenance supervisors.

**Traffic Signal Maintenance Technician** – Staff are generally responsible for troubleshooting and maintenance of the physical traffic signal equipment.

**Electronic Specialist** – Staff are responsible for the complex electronic equipment at the heart of the signal system. Some tasks include:

- Closed circuit television system repair, field and central system
- Fiber optic cable system testing, repair, termination
- Telecommunications systems maintenance and repair
- Traffic management center systems maintenance and repair
- Traffic signal controller electronics testing repair and inventory
- Other ITS devices repair

**TMC Operators** – Staff are responsible for observing the traffic conditions, responding to incidents that occur in the field, and providing support to homeland security efforts. Their role is critical to the rapid response and resolution of the situation.

**Public Relations Coordinator** – Staff are responsible for field phone calls from the public, coordinating with the Traffic Signal Engineer and Technician on responses, and marketing the TMC, incident management plan, and traffic signal operations to the public. Depending on the size of the agency, this position could be a full-time position or these tasks might be passed on to the Traffic Signal Engineer and Technician.

### 8.4.3 Staff Needs

The ITE "Traffic Control System Operations" manual suggests that a traffic signal system should have one traffic engineer per 75 to 100 traffic signals and one signal technician per 40-50 traffic signals or other field devices. An NCHRP report (Synthesis 245) also suggests 38 to 43 signals per technician. The manual also provides staffing guidelines for a continuously operated TMC which includes one center manager, two supervisors, and five system operators.

Overall, the current literature provides limited guidance on staffing for complex traffic signal systems that include a multitude of components ranging from traffic signals to video detection to ITS devices to incident management plans and a TMC. With the above limitations in mind, Table 8-2 provides general guidelines on staffing needs for a traffic signal system as it relates to signal retiming.

**Table 8-2 Summary of Staffing Needs**

| Position | 1 to 50 Traffic Signals | 51 to 100 Traffic Signals | 101 to 200 Traffic Signals | 201 to 500 Traffic Signals | 501 to 1000 Traffic Signals |
|---|---|---|---|---|---|
| Traffic Signal Engineer | 0 to 1 | 1 | 1 to 2 | 2 to 5 | 5 to 10 |
| Traffic Signal Analyst/Technician | 0 to 1 | 0 to 1 | 1 | 1 to 3 | 3 to 5 |
| ITS Engineer | - | - | 0 to 1 | 1 | 1 to 3 |
| Traffic Signal Maintenance Technician | 1 to 2 | 2 to 4 | 4 to 7 | 7 to 17 | 17 to 33 |
| Electronic Specialists | 1 | 1 | 1 to 2 | 2 to 4 | 4 to 9 |
| TMC Operators | - | - | 2 | 2 to 4 | 4 to 9 |
| Public Relations Coordinator | 0 to 1 | 0 to 1 | 1 | 1 | 2 |

Additional research in this area is necessary because all agencies vary in terms of the infrastructure, their staff's skill levels, and the environment the equipment operates. For example, a signal technician that has ten years of experience with a particular traffic controller is likely able to troubleshoot a problem in less time then a technician with less experience. In many cases, the difference in staff time needed is significantly higher, while the cost to the agency may be insignificant. This particularly effects agencies with high levels of staff turnover and is a notable problem for many agencies throughout the country.

Another example of this is use of older infrastructure may be less user friendly or it may fail or malfunction more often because of the environmental conditions which the traffic signal equipment operate. As with staff experience there is not specific research on the performance of traffic signal equipment over time and new equipment can result in other problems for agencies, but by utilizing a systems engineering approach during the selection and procurement of new traffic signal equipment, an agency may reduce the maintenance costs (staffing and consultant time) needed for the signal system.

**Several agencies maintain signal systems that operate in the DOS environment (an operating system that was before Windows). In most cases, these systems are not supported by Information Technologies (IT) staff and access to data leads to inefficient operations or requires special maintenance by agency personnel.**

Proximity to traffic signals is an important consideration, especially in staffing needs for technicians that are tasked with responding to field complaints or maintenance problems. A city with a large downtown may have 100 signals within a square mile (Manhattan, Portland, OR for example) that are fixed time without detection as opposed to a rural district with 100 signals over a 100 square mile area. Obviously, in these cases, the staffing needs are likely different and may require different skill sets. This may be exacerbated by an agency that utilizes different controller types supplied by various vendors. Many agencies are transitioning from an older standard, thus requiring careful scheduling for staff training and equipment management.

The Federal Highway Administration has recently conducted several Regional Traffic Operations Program Assessments to determine the sufficiency of staff within a region to perform basic signal operations activities for individual jurisdictions. In many cases, these reviews have shown that there are opportunities for improved operations through the use of regional approaches to problem solving, equipment procurement and testing, staff training, and performance measurement.

Additional research in this area is needed to further identify specific needs by agency, as each jurisdiction has special circumstances that may either overstate or understate the need described in Table 8-2. Overall, findings from the National Traffic Signal Report Card indicate that traffic signal operations could be improved with increased investment.

## 8.5 REFERENCES

---

1 Vick, C. and R. Sumner, "*Guidelines for Transportation Management Systems Maintenance Concept and Plans"*, Federal Highway Administration, 2002.

## CHAPTER 9

# ADVANCED SIGNAL TIMING TOPICS

## TABLE OF CONTENTS

# LIST OF FIGURES

# 9.0 ADVANCED SIGNAL TIMING CONCEPTS

This chapter covers some advanced concepts and applications within signal timing. The intent of this chapter is to introduce the concepts and to point the reader to references and information available to provide additional details. Each concept section provides an overview, discusses the effects on signal timing, and offers examples where applicable.

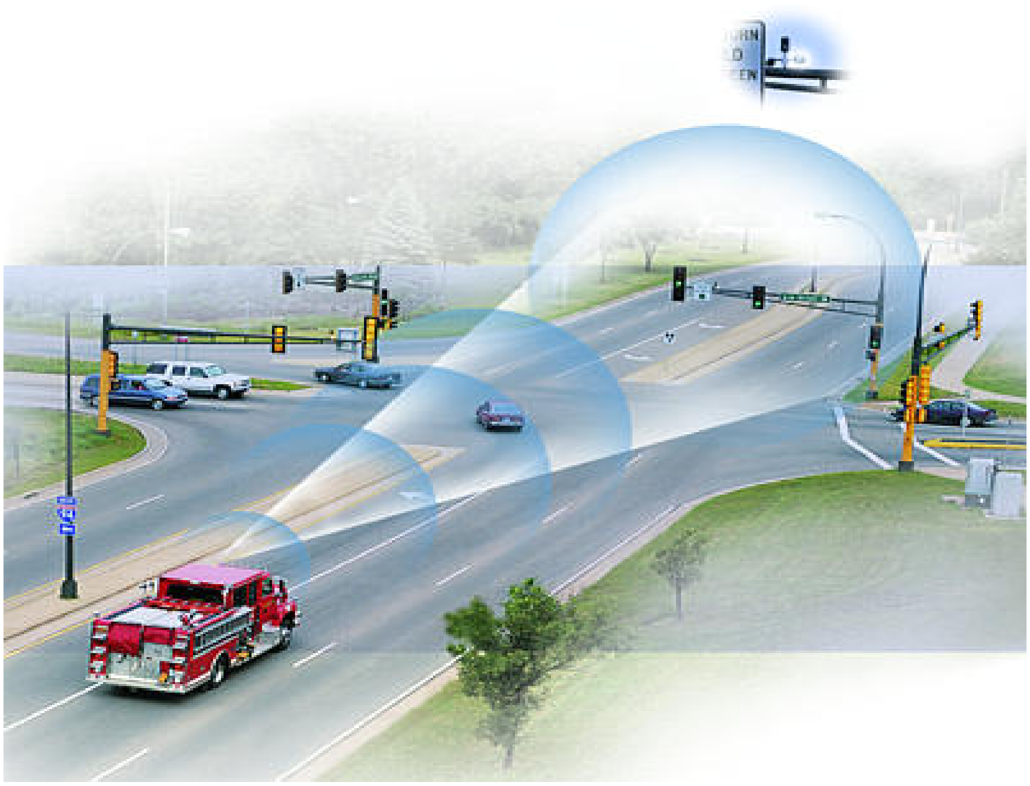
## 9.1 TRAFFIC SIGNAL PREEMPTION

### 9.1.1 Preemption Overview

The 2003 Manual on Uniform Traffic Control Devices (MUTCD) defines traffic signal preemption as "the transfer of normal operation of a traffic control signal to a special control mode of operation" (*1*). Preemptive control is designed and operated to give the most important classes of vehicles the right of way at and through a signal. This right of way is usually achived with a green indication on the approach of the vehicle requesting preemption. Preemptive control may be given to trains, boats, emergency vehicles, and light rail transit. It is commonly used for fire engines because the size of their vehicles makes them less able to move through traffic without the aid of preemption. Signal preemption controls the movement of traffic that is of greater importance than general vehicle and pedestrian traffic. Preemptive control is necessary to avoid collisions (e.g., trains versus automobiles) and/or give right of way to vehicles in an emergency situation (e.g., fire engines responding to an emergency).

Several types of technologies are available to detect vehicles requesting preemption , and include the use of light (strobe), sound (siren), pavement loops, radio transmission, and push buttons approaching an intersection, to request immediate service (green indication from the signal). Figure 9-1 shows an example of an emergency vehicle signal preemption using optical detection.

**Figure 9-1 Emergency Vehicle Signal Preemption Example (*2*)**



Preemption interrupts normal signal operations to transfer right of way to the direction of an approaching emergency vehicle, but a green indication is not always guaranteed immediately after preemption is requested. The MUTCD states that the shortening or omission of any pedestrian walk interval and/or pedestrian change interval shall be permitted. This is often necessary with intersections adjacent to rail as it is often infeasible to provide clearance given the limitations of the locations of railroad track circuits and the speed of the approaching vehicles, and sometimes due to very long crosswalks versus close detection distances.

For marine transport, the preemption would not be delayed as well for the servicing of pedestrians at such locations as drawbridge crossings. The signal preemption typically occurs not at the drawbridge itself but at a signalized intersection immediately adjacent to the drawbridge. At the intersection there may well be pedestrian signals. Whether or not the ped intervals are shortened or omitted when the drawbridge preemption occurs is a matter of engineering judgment, but the MUTCD allows the ped intervals to be shortened or omitted for any type of preemption but not for "priority control".The immediate servicing of pre-emption requests for these transportation modes is due to the need to maintain continued flow for rail and marine mobility.

Preemption is different from signal priority, which alters the existing signal operations to shorten or extend phase time settings to allow a priority vehicle to pass through an intersection. Traffic signal priority is discussed in greater detail in Section 9.2 of this chapter.

### 9.1.2 Effect on Signal Timing

Preemptive control has a profound affect on signal timing because, in the controller, it totally replaces normal timing and logic with preemptive timing and logic to serve a specific vehicle type. The preemptive systems can extend the green time on an approach up to a preemptive maximum, that is irrespective of the maximum green or coordination settings. Preemptive service is followed by a recovery or transition period where the controller transitions to normal signal operations and coordination timing plans (if applicable).

On a signal preemptive system in the Washington, DC, metropolitan area, once a signal was preempted, the coordinated systems took anywhere between 30 seconds to 7 minutes to recover to base time coordination(*3*). Signal phase sequencing and methodology for recovery or transition should be developed to minimize the impact preemption has on traffic operations and safety.

Part 4 of the MUTCD calls out the standards for how to transition into and out of preemption. The key signal timing aspects of preemption are listed here.

- *Transition Into Preemption*
  - The yellow and all-red vehicle clearance interval shall not be shortened or omitted.
  - Pedestrian walk or clearance intervals may be shortened or omitted.
  - A return to the previous steady green signal indication shall be permitted following a steady yellow signal indication in the same signal face, omitting the red clearance interval.
- *Transition out of Preemption*
  - The yellow and all-red vehicle clearance interval of the preempted approach shall not be shortened or omitted.
  - A signal indication sequence from a steady yellow signal indication to a steady green signal indication shall not be permitted.

In addition, traffic signals that can receive multiple requests for signal preemption should prioritize the requests by importance of vehicle right of way and/or by difficulty in stopping the type or class of vehicle. The amount of time a signal has to transition into preemption is predicated on the distance upstream where the preempting vehicle can be detected.

Some of the benefits associated with traffic signal preemption are:

- Improved response time/travel times for emergency, rail, waterway, and other preempting vehicles.
- Improved safety and reliability for vehicles receiving preemption right of way (e.g. emergency vehicles, trains, and boats).
- Improved safety and clarity of right of way for other roadway users (i.e. avoids drivers having to yield right-of-way on their own without prompting from traffic control for an emergency vehicle or etc.).

The nature of signal preemption varies greatly in its application, e.g. heavy rail crossings near a signalized intersection must be approached differently than providing preemption for emergency vehicles. Further references should be consulted beyond the general overview presented here to fully understand the various complexities associated with signal preemption. Two such resources are the National Cooperative Highway Research Program (NCHRP) Report 3-66, *Traffic Signal State Transition Logic Using Enhanced Sensor Information*, which describes preemption and advanced preemption due to heavy rail and light rail vehicles (*4*), and *Traffic Signal Preemption for Emergency Vehicles, A Cross-Cutting Study* by the Federal Highway Administration (FHWA) and the National Highway Traffic Safety Administration (*5*).

### 9.1.3 Example Applications

Examples of preemptive control vary widely, but could include the following:

- The prompt display of green signal indications at signalized locations ahead of fire vehicles and other official emergency vehicles (many cities have determined that law enforcement and ambulances are nimble enough to use their siren and can navigate efficiently without the aid of signal preemption, which reduces disruption to the signals and reserves the preemption for first repsonders from the fire department);
- A special sequence of signal phases and timing to provide additional clearance time for vehicles to clear the tracks prior to the arrival of a train;
- A special sequence of signal phases to display a red indication to prohibit movements turning toward the tracks during the approach or passage of a train or transit vehicle; and
- The prompt display of green signal indications at a freeway ramp meters to progress a standing queue through the meter to avoid queue spillback into upstream traffic signals.

Various challenges are identified in these applications originating from lessons learned. These include the following:

- Typically, a signal cycles a number of times before it returns to normal timing plan operations after a preemption call is carried out. This causes less-than-optimum timing splits, offsets, and corridor progression, which results in additional delays and queues, particularly during peak traffic volume periods.
- Delays to roadway traffic can be exacerbated if preemption is used for a transit service on light rail lines especially during peak hours when frequent calls for preemption would be issued at rail crossings (6). Furthermore, rail crossing gates that provide clearance for rail vehicles requesting preemption may be held down for too long, causing further delays to roadway traffic (6). This can occur if sensor equipment, detecting the rail vehicle, is not placed in the correct locations or is not operating reliably (6).

- Preemption can significantly shorten pedestrian walk and flash don't walk intervals. Care should be given to allow as much time as possible to ensure safe pedestrian crossing or return to curb.
- Preempt trap at signalized intersections near railroad crossings can occur if insufficient track clearance green time is allotted in association with an advanced preemption. The track clearance green time allows vehicles in a queue to be served on the near-side approach and to clear the railroad track area. The potential problem occurs when the clearance green time is not long enough to clear the queue. Under this scenario, a vehicle could be trapped on the tracks when the railroad crossing lights come on and the railroad gates come down.
- Pre-emption calls for rail transit systems can delay emergency vehicles, thus creating a significant "public safety concern" (6). Also, if "preempt confirmatory lights" do not provide clear indications on whether an approaching vehicle has control of the signal, emergency vehicles may wrongly assume they have received preemption. This can create a potential hazard if the preemption call was issued to the rail system that is in conflict with the emergency vehicle's approach.
- If there are multiple vendors supporting a preemption system, then interoperability will likely be required among the various proprietary technologies (e.g., between the signal controller hardware of vendor A and the traffic control software of vendor B). Such integration can be especially difficult to achieve if the vendors are market competitors (6).
- Implementing a successful signal preemption for rail is likely to require a multifaceted, coordinated systems level approach:
    - Requires coordination among multiple stakeholders such as transit authorities, emergency responders, roadway agencies to minimize any adverse impacts from preemption system on each stakeholder's operations.
    - The preemption system at a particular intersection must not only take into account standard roadway traffic characteristics, but also various attributes of the transit system that it's servicing such as headways and frequency of service. Overall, potential impacts to all relevant modes should be accounted for when designing and implementing a preemption system. It may require further analysis to comprehensively assess the impacts of signal preemption and to better understand the interaction between "rail and traffic signalizing systems" (6)

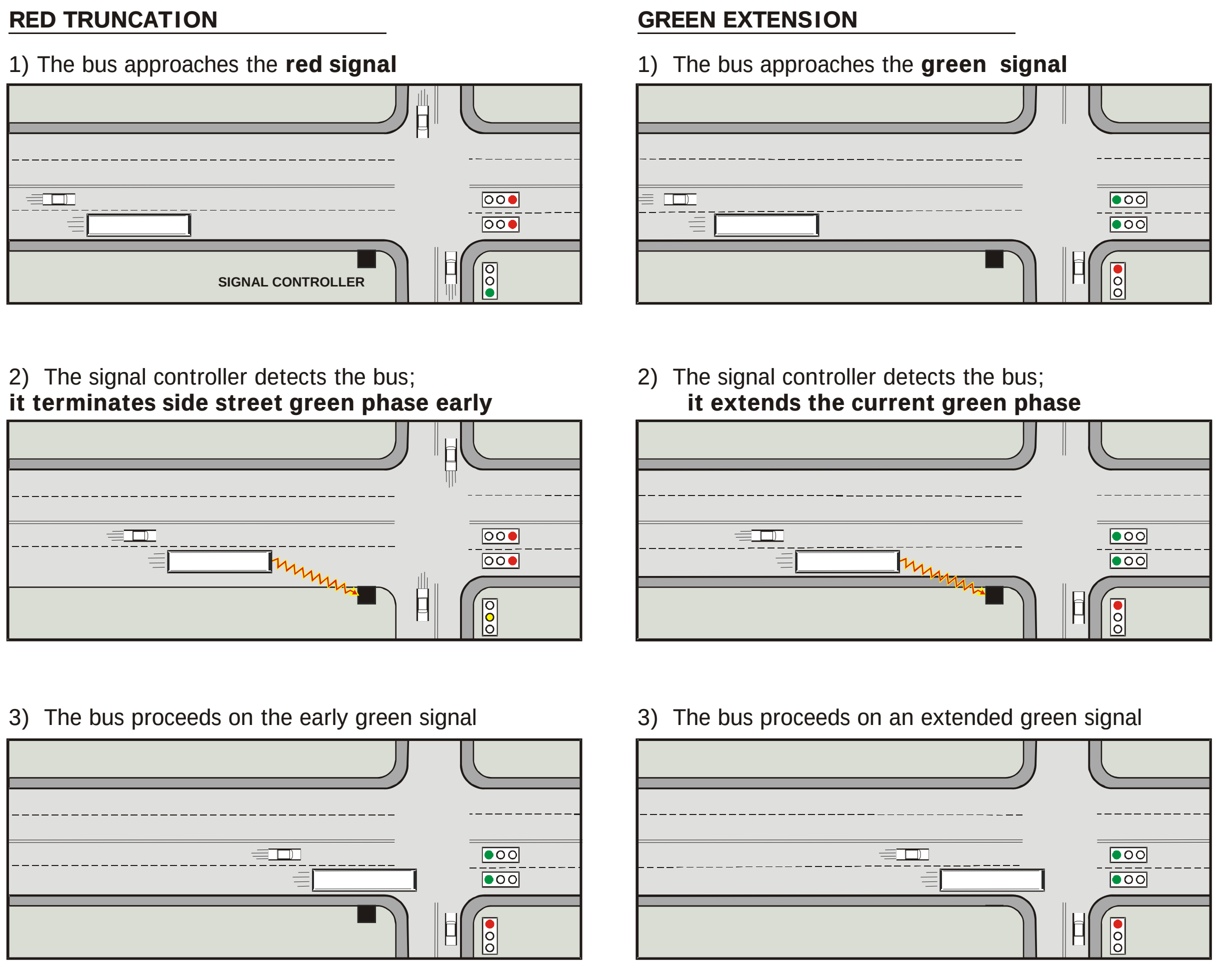
## 9.2 TRAFFIC SIGNAL PRIORITY

### 9.1.4 Traffic Signal Priority Overview

Transit Signal Priority (TSP) is an operational strategy that is applied to reduce the delay transit vehicles experience at traffic signals. TSP involves communication between buses and traffic signals so that a signal can alter its timing to give priority to transit operations. Priority may be accomplished through a number of methods, such as extending greens on identified phases, altering phase sequences, and including special phases without interrupting the coordination of green lights between adjacent intersections. Ultimately, TSP has the potential to improve transit reliability, efficiency, and mobility.

With TSP, there are two basic methods to adjust the signal timing at an intersection for an approaching bus: reducing the red time (red truncation) or extending the green time (green extension). Figure 9-2 illustrates these methods.

**Figure 9-2 Effect of TSP to Adjust Signal Timing**




### *Differences from signal preemption*

TSP is different from signal preemption, which interrupts the normal signal cycle to accommodate special events (e.g., a train approaching a railroad grade crossing adjacent to a signal or an emergency vehicle responding to a call). For example, a fire engine may send a preemption request that instantly alters the traffic signal timing and or phasing to provide a green indication. In this case, the normal signal operations process would be disrupted. More specifically, with pre-emption certain phases may be skipped or replaced for approaches to the intersection that are not receiving the signal pre-emption treatment (*7*). Note: preemption is discussed in Section 9.1.

With TSP, however, the transit detection system communicates a priority request to the traffic signal that may or may not be granted. If such a request is granted, the traffic signal timing is altered to serve the priority request without disrupting coordination. In this situation, the normal signal operations process and overall signal cycle are maintained (*6*). With TSP, side-street phases would not be skipped, although the timing of these phases is likely to be altered.

To achieve greater uniformity in the deployment and implementation of ITS and applications such as TSP, a “family” of communication standards was developed known as The National Transportation Communications for ITS Protocol (NTCIP) (*6*). NTCIP 1211 was established by the Signal Control and

Prioritization (SCP) working group, and it provides a set of communication standards for exchanging information among SCP systems such as TSP.

NTCIP 1211 represents the physical elements needed to provide signal priority, typically four components associated with the bus detection system and the traffic signal controller. The detection system lets the TSP system know where the vehicle requesting priority is located. The detection system communicates that message with the priority request generator (PRG) and the priority request server (PRS) manages those requests. The fourth and final component of the system may be the transit AVL or traffic management center which monitors the system and logs data.

### 9.1.5 Effect on Signal Timing

Transit signal priority has a limited affect on signal timing because it adjusts to normal timing and logic to serve a specific vehicle type. The priority algorithm modifies the green allocation and may work within the constraints of coordination settings or maximum green. The NTCIP 1211 standard requires that priority allow the coordination logic to be maintained without a recovery or transition period after the priority request.

The most commonly reported benefits of using signal priority include reduced signal delay for transit vehicles and improved transit travel time. In some cases, improved reliability has been provided through the integration of a transit system's Automatic Vehicle Location (AVL) system to request priority only when the vehicle is behind schedule.

### *9.1.6* Examples of Transit Signal Priority

TSP has been employed in various urban transit networks throughout the United States. King County, Washington, which includes the Seattle metropolitan area, is one region where TSP is actively operating. Portland, Oregon, is another. While both of these systems ultimately alter the timing of traffic signals to provide some benefit to transit vehicles, they differ in the way signal priority requests are generated.

#### *King County, WA*

With the TSP system for King County Metro, the transit vehicle is not responsible for generating a NTCIP 1211 standard priority request for signal priority. Instead, the vehicle only communicates its presence to the traffic signal system, which in turn has the ability to generate a signal priority request for the approaching bus.

As the vehicle approaches an intersection, an Automated Vehicle Identification (AVI) system within the bus transmits information, such as a vehicle's ID number, to a roadside reader. The reader detects the presence of a transit vehicle approaching the intersection and sends a message to a transit interface unit located in the individual signal controller cabinet.

TSP has resulted in transit performance improvements in King County. For example, the combination of TSP and signal optimization is responsible for a 40% reduction in transit signal delay along two transit corridors. Another one of the observed benefits from TSP in King County was a 35–40% reduction in travel time variability (*7*).

#### *Portland, OR*

In contrast with the TSP system in King County, transit vehicles in Portland, Oregon have the ability to decide whether or not to request priority through the AVL integration described above. . The vehicle is equipped with detection equipment (optical-based) similar to that used to provide an emergency vehicle with a preemption request. Unlike preemption, however, the transit vehicle using ITS only transmits input to the signal control system when it needs priority, such as when it is behind schedule. This process generally results in "first in, first out" operation, where the first vehicle to transmit a need for priority is served first, regardless of relative need.

Like King County, transit vehicles in Portland have experienced some measurable benefits with the TSP implementation. Some of those benefits are a 10% improvement in travel time and a 19% reduction in travel time variability. With the increased reliability, less schedule recovery time was needed to keep the buses on schedule. Moreover, the benefits of TSP enabled the Portland transit agency, TRIMET, to avoid having to purchase an additional bus (*7*).

## 9.3 TRAFFIC RESPONSIVE OPERATION

### 9.1.7 Traffic Responsive Overview

It is common for a coordinated traffic signal to operate different timing plans at different times of the day and days of the week. This is done by utlizing a predetermined timing plan that best suits the current traffic conditions. For example, at different times, a signal may operate an a.m. Peak, a p.m. Peak, or an Off Peak plan. It may also operate in free (uncoordinated) mode at other various times, such as overnight. The most common means of determining when to change timing plans is to use a time-of-day and day-of-week schedule. This is referred to as time-of-day (TOD) plan selection and is described in more detail in Chapter 7. The time-of-day plan selection schedule can be implemented in either the controller, field master, or a central computer system. However, during incidents or other unusual conditions, plans may also be changed manually through the same means. In any case, signal coordination usually requires that the timing plan be changed simultaneously at all signals within a coordinated group. Therefore, if time-of-day plan selection is done locally at the controller, all controllers within a coordinated group need to be configured with the same time-of-day schedule.

Time-of-day plan selection works well when traffic conditions are consistent and predictable – that is, similar traffic patterns generally occur during the same times each day. When an incident, a planned event (e.g., construction, county fair, football game, etc.), extreme weather, or any other unusual occurrence causes a significant change in the normal traffic conditions, the timing plan selected by the time-of-day method may not be the plan best suited to current conditions. To address this situation, the traffic responsive plan selection method uses data from traffic detectors, rather than time of day, to automatically select the timing plan best suited to current conditions.

Plan selection for responsive operations may also be invoked manually. There are several instances where agencies operating traffic signals from their traffic management centers will use predetermined plans for planned special events or recurring congestion on an as needed basis.

**To implement traffic responsive operations, it may be necessary to update TOD/coordination plans. Along with fine tuned plans, it is critical to confirm that the local controller clocks are in sync to maintain the coordination plans.**

### 9.1.8 Techniques - Operational

Traffic responsive plan selection (TRPS) normally takes place in a field master or a central computer system. When the master or computer selects a new timing plan, it sends a command to all signals in a coordinated group to instruct them to change to the new plan simultaneously, thus ensuring coordination is maintained.

The master or central computer monitors volume and/or occupancy data from multiple vehicle detectors. The data from the detectors are weighted, merged, and otherwise processed to calculate values for a few key parameters that are compared to thresholds. When a threshold is crossed, one of the predetermined plans is implemented for the conditions represented by the threshold categories selected. Different signal system suppliers use different parameters and algorithms for traffic responsive plan selection. A user needs to refer to documentation specific to their signal system and must prepare these in advance. Most algorithms involve separate calculations and threshold comparisons for each cycle length (total volume of traffic), offset (direction of traffic with largest volume), and split (relative occupancy and/or volume on different streets).

Regardless of the algorithm used, traffic responsive plan selection requires the user to enter potentially complex configuration parameters. When setting up TRPS, considerable effort may be needed to identify the vehicle detectors that will provide an adequate representation of traffic conditions, to establish appropriate parameter values associated with those detectors, to establish appropriate thresholds and associated plans, and to fine tune the configuration based on its performance once TRPS is implemented. Historical traffic count (and preferably occupancy) data should be available for candidate detectors before the detector selection and setup process begins. It is often necessary to repeatedly adjust parameters, especially thresholds, based on observation of calculated values relative to actual traffic conditions until effective settings are established.8

The detectors used in TRPS generally need to be located away from the stop line. (Advance detectors or departure-side detectors are commonly used.) The detectors need to be configured so that the controller can generate reasonably accurate count and occupancy data separately for each direction of travel even if served by the same signal phase, and preferably for each lane. Such detectors are often referred to as system detectors. Detectors used in TRPS must be actively monitored, reliable and faults must be repaired quickly.

If the volume of traffic trying to make a movement at an intersection exceeds the capacity of the traffic signal's phase serving that movement, traffic queues can grow over multiple cycles. In this case, measuring the volume of traffic served by the phase will not detect the overload condition. Occupancy (temporal density, or the portion of time that a detector is occupied, in the range 0.00 to 1.00) on an advance detector within the queuing area can be used to detect the onset of additional queuing, even though the served volume doesn't change. Occupancy can therefore be used to select a plan with a cycle length and split that will accommodate the excess traffic demand. On the other hand occupancy alone is relatively insensitive to traffic volume changes under free flow conditions, when queues don't extend onto detectors.

Algorithms usually allow the use of a combination of volume and occupancy in plan selection. One technique is to combine these into a single dimensionless value by summing the volume (count) and a multiple of the occupancy. This process is often referred to by the expression Volume (V) + Factor (k) x Occupancy (O), or V+kO. Increasing the value of the factor k increases the sensitivity to occupancy changes, relative to volume.

The algorithms commonly used in TRPS include hysteresis to prevent oscillations across a threshold. Even so, it is common for timing plans to change too frequently, resulting in inefficiencies due to offset transitions. On the other hand, the time period over which selection values are calculated can result in a coordinated system being too slow to change plans at the onset of a rapid change in traffic volumes such as may result from a major incident. Nonetheless, when well configured and fine tuned, TRPS can result in improved traffic signal operation compared to time-of-day plan selection, especially for coordinated signal groups subject to significant unpredictable changes in traffic flows. A related problem is that the plan change may occur during peak traffic conditions when an offset transition has a greater negative impact.

Despite the problems sometimes encountered, a well configured and fine tuned traffic responsive plan selection process can result in improved traffic signal operation compared to time-of-day plan selection for coordinated signal groups subject to significant unpredictable changes in traffic flows. Such conditions are often found, for example, adjacent to a major event venue and on a route that serves as a by-pass for a blocked freeway segment.

If the actual traffic conditions are quite different from that for which any of the available plans was designed, as often happens during both planned and unplanned events, it may be necessary to develop special timing plans for unusual conditions. For example, a special plan may be designed to serve extra heavy traffic (in one direction) leaving a sporting event or bypassing a freeway blockage. Traffic responsive plan selection can automatically implement such a plan when those conditions occur.

TRPS merely selects a timing plan to operate, but does not make changes to the timings specified in the timing plan. That is the role of adaptive traffic signal control, described in Section 9.5.

## 9.4 ADAPTIVE TRAFFIC SIGNAL CONTROL

### 9.1.9 Adaptive Traffic Signal Control Overview

Adaptive traffic signal control is a concept where vehicular traffic in a network is detected at an upstream and/or downstream point and an algorithm is used to predict when and where the traffic will be and to make signal adjustments at the downstream intersections based on those predictions. The signal controller utilizes these algorithms to compute optimal signal timings based on detected traffic volume and simultaneously implement the timings in real-time. This real-time optimization allows a signal network to react to volume variations, which results in reduced vehicle delay, shorter queues, and decreased travel times

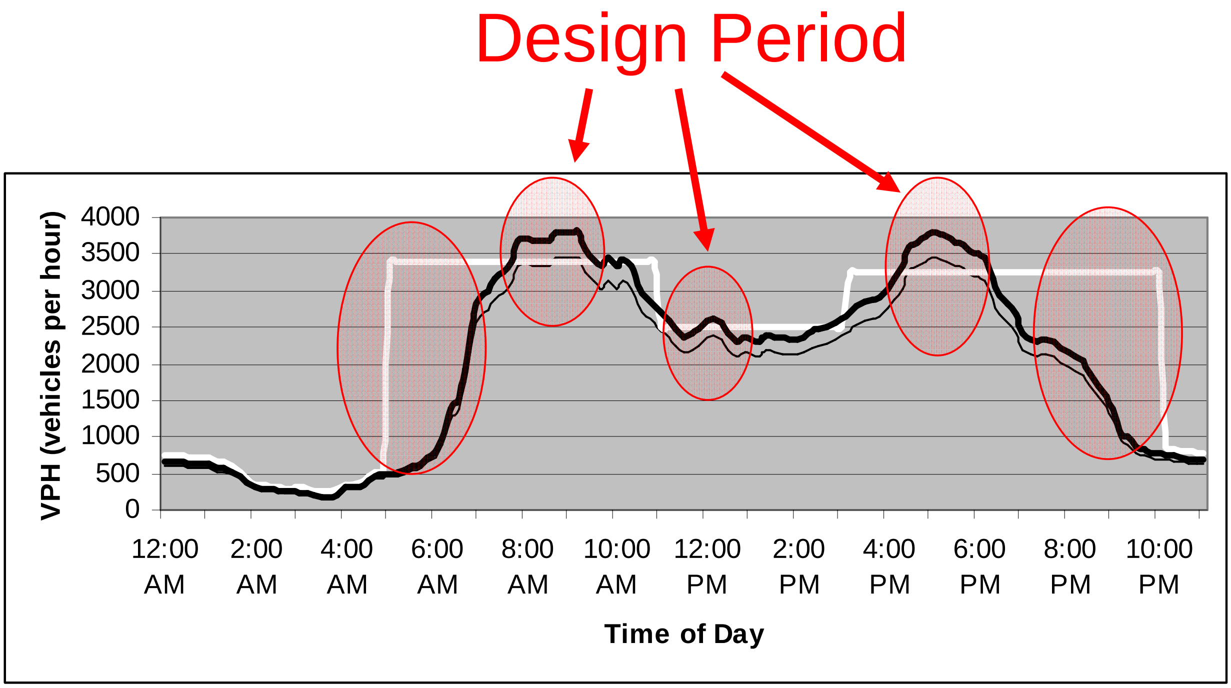
**All adaptive systems are critically linked to good detection systems. While some adaptive systems will have better tolerance of detector faults than others, the reliability and accuracy of the decisions made by the adaptive algorithms cannot be achieved without well-maintained detection. While it is important to consider maintenance of detector systems for all types of traffic signal control, maintenance of detection for adaptive systems is particularly important.**

Adaptive signal control autonomously adjusts signal timing parameters in real-time, to respond to actual, real-time traffic conditions. By adjusting the traffic control parameters to more closely align with traffic conditions, adaptive systems can reduce traffic delay, increase average speeds, improve travel times, and decrease travel time variability. Many studies have shown that adaptive signal control improves average performance metrics by 10%, with some systems improving particularly poor conditions by 50% or more. In some cases, comparisons suggest that adaptive systems do not improve traffic conditions significantly. These studies have typically compared adaptive control with "well-tuned" actuated and pre-timed systems. Thus, good and sound signal timing practices, such as those promoted in this manual, can go a long way to improve traffic conditions. However, adaptive systems can still have a profound impact even starting from well-designed baseline signal timings where:

- Traffic conditions fluctuate randomly on a day-to-day basis
- Traffic conditions change rapidly due to new or changing developments in land use
- Incidents, crashes, or other events result in unexpected changes to traffic demand
- Other disruptive events, such as preemption, require a response

The underlying reason for adaptive system performance is quite simple; when system conditions are stable, there is little need to modify the control parameters if they are set appropriately. In the situation where fixed parameters have not been chosen by thorough analysis, adaptive systems can find better traffic control parameters. When system conditions change, system performance can only be improved by modifying the control parameters. As shown in Figure 9-3, a typical signal timing time-of-day plan schedule can be adequate for a majority of daily volume fluctuations, but may not be optimal under variable flow.

**Figure 9-3 Sample Time of Day Volume Profile Compared to Time of Day Timing Schedule**




In addition, where congestion is high, adaptive systems typically provide benefits above fixed-time and fixed-parameter systems by:

- Delaying the onset of oversaturation
- Recovering more quickly from saturation after traffic demand has dissipated

On the other hand, there are cases where fixed-time and fixed-parameter systems may outperform adaptive control approaches when traffic volumes are very low. In this case, sound traffic engineering practice is adequate for effective traffic management.

### 9.1.10 Adaptive Control Concepts

Many adaptive systems have been developed over the last 30 years. New systems continue to evolve and new techniques have emerged as technology and traffic control system architectures have taken shape. In the United States, there are several adaptive systems available from a variety of vendors.

All adaptive control systems are driven by a similar conceptual process:

1. Collect data in real-time from sensor systems to identify traffic conditions
2. Evaluate alternative signal timing strategies on a model of traffic behavior
3. Implement the "best" strategy according to some performance metric
4. Repeat steps 1,2,3 again and again

Each adaptive system is distinguished by how it uses different components or approaches to these four key steps in the control of the traffic system.

All adaptive systems need accurate and comprehensive traffic detection systems, but some need a particular size or length or placement of detectors, while others are more flexible. Most adaptive systems require more detection than would be deployed for semi-actuated traffic control and traffic responsive plan selection methods.

Most adaptive traffic systems use some form of a traffic model to evaluate alternative traffic control strategies. This model may be based on virtually moving individual vehicles down the street and predicting their movements, estimating platoons of vehicles, measuring statistics of occupancy or volumes over time, and other approaches. All adaptive systems must have some way of internally evaluating the question: "Is traffic control scheme A better or worse than control scheme B? (and C and D and E …)". To some extent, all adaptive systems:

- Represent the current state of the traffic system
- Predict how the conditions of the traffic system will change if the control strategy changes (as well as what will happen if it doesn't) through various algorithms

Based on the complexity of the traffic model, the adaptive control system ranges from thousands of parameters (based on the number of intersections of the deployment) to very few. In most cases, parameters require calibration and adjustment. This can be a time-consuming and expert-driven process during deployment. The better tuned the traffic model becomes (i.e. the more it is matched to what is happening in the real-world), the better the system will perform. Some systems can automatically adjust their internal model parameters and others require human judgment and interaction. In any event, the traffic model and the prediction methodology (and how accurate its predictions are with respect to what really happened) used by the adaptive system determine how effectively it operates.

Finally, all adaptive systems evaluate alternative traffic control strategies and pick the best alternative according to a performance metric. There are three components to this element of adaptive control:

- The process for searching through as many possible alternatives as quickly as possible
- Evaluating each alternative using a performance metric
- Trading off improvements at individual intersections with system-wide performance considerations

The search methodology itself is not typically important for signal timing, but how the traffic engineer is able to constrain or influence the search of alternative timing strategies is critical to the success of adaptive system deployment. For example:

- The minimum and maximum phase lengths
- Which phase sequences are allowed or disallowed
- How rapidly or slowly the system allows timing parameter changes

Other parameters, specific to individual adaptive systems, are key elements to guide the adaptive system in searching appropriate and effective signal timing strategies. Training and staffing are important to keep the system operating at its best.

Similar to the process to identify constraints for signal timing optimization, the influence of the traffic engineer in determining the performance metric is also critical to the success of the adaptive system. Typical performance metrics may be to minimize total intersection delay, or some weighted combination of intersection delay and stops. Other systems look at different metrics such as progression efficiency, or maximizing a green-band along an arterial. Whatever the metric is in the adaptive system, the traffic engineer must ensure that the metric matches what he or she envisions as the proper optimization objective for the traffic where the adaptive system is deployed.

#### 9.1.11 Examples of Adaptive Traffic Control Systems

In the United States, many adaptive control systems have been deployed, but adaptive systems still control less than 1% of all the traffic signals in the nation. As detection and communications technology improves, adaptive systems are certainly becoming more popular and effective. This section provides a brief description of some systems that have been deployed in the U.S. over the last 20 years.

### *Split Cycle Offset Optimisation Technique (SCOOT)*

Developed in the United Kingdom, SCOOT is the most widel deployed adaptive system in existence. SCOOT uses both stop-line and advance detectors, typically 150-1,000 feet (50-300 meters) upstream of the stop line (or exit loops (loop detectors located downstream of the intersection), measuring vehicles leaving the upstream detector). The advance detectors provide a count of the vehicles approaching at each junction. This gives the system a high-resolution picture of traffic flows and a count of the number of vehicles in each queue, several seconds before they touch the stop line (allowing time for communication between the traffic signal controller and the central SCOOT computer). SCOOT also provides queue length detection and estimation. Under the SCOOT system, green waves can be dynamically delayed on a "just in time" basis based on the arrival of vehicles at the upstream detector, which allows extra time to be allocated to the previous green phase, where warranted by heavy traffic conditions. SCOOT controls the exact green time of every phase on a traffic controller by sending “hold” and “force-off” commands to the controller.

The SCOOT model utilizes three optimizers: splits, offsets, and cycle. At every junction and for every phase, the split optimizer will make a decision as to whether to make the change earlier, later, or as due, prior to the phase change. The split optimizer implements the decision, which affects the phase change time by only a few seconds to minimize the degree of saturation for the approaches to the intersection.

During a predetermined phase in each cycle, and for every junction in the system, the offset optimizer makes a decision to alter, all the offsets by a fixed amount. The offset optimizer uses information stored in cyclic flow profiles and compares the sum of the performance measures on all the adjacent links for the scheduled offset and the possible changed offsets.

A SCOOT system is split into cycle time “regions” that have pre-determined minimum and maximum cycle times. The cycle optimizer can vary the cycle time of each REGION in small intervals in an attempt to ensure that the most heavily loaded NODE in the system is operating at 90% saturation. If all stop bars are operating at less than 90% saturation, then the cycle optimizer will make incremental reductions in cycle time.

### *Sydney Co-ordinated Adaptive Traffic System (SCATS)*

Developed in Australia, SCATS uses a split plan selection technique to match traffic patterns to a library of signal timing plans and scales those split plans over a range of cycle times. SCATS gathers data on traffic flows in real-time at each intersection. This data is fed to a central computer via the traffic control signal. The computer makes incremental adjustments to signal timing based on second by second changes in traffic flow at each intersection. SCATS performs a vehicle count at each stop line and measures the gap between vehicles as they pass through each junction. As the gap between vehicles increases, green time efficiency for the approach decreases, and SCATS seeks to reallocate green time to the greatest demand. SCATS selects a timing plan on the controller, and thus the local actuated controller uses its own inherent gap-out and force-off logic to control the intersection second-by-second.

### *Real Time Hierarchical Optimized Distributed Effective System (RHODES)*

RHODES uses a peer-to-peer communications approach to communicate traffic volumes from one intersection to another in real-time. By passing the data back and forth over a high-speed communication network, RHODES is able to predict the impacts of traffic arriving 45-60 seconds upstream and plan for traffic phase sequence and phase durations accordingly. RHODES continually re-solves its planned phase timings, every 5 seconds, to adapt to the most recent information. RHODES requires upstream and stop-bar detectors for each approach to the intersections in the network and has a wide variety of parameters that are used to calibrate the traffic model to real-world conditions. RHODES over-rides the local controller by sending “hold” and “force-off” commands to the controller to set the exact duration of each phase.

### *Optimized Policies for Adaptive Control (OPAC) “Virtual Fixed Cycle”*

The OPAC adaptive control system uses a predictive optimization with a rolling horizon. This congestion control strategy, which attempts to maximize throughput, adjusts splits, offsets, and cycle length, but maintains the specified phase order. For un-congested networks, OPAC uses a local level of control (at the intersection) to determine the phase durations, and a network level of control for synchronization which is provided either by fixed-time plans (obtained offline), or by a virtual cycle (determined online). The levels of local and global influence are flexible and can be adjusted by the traffic engineer. The state of the system is predicated using detectors located approximately 10-15 seconds upstream on the approaches to the intersection. OPAC sends “hold” and “force off” commands to the local controller to set the exact duration of every phase on the signal.

***Adaptive Control Software Lite (ACS-Lite)***

ACS-Lite was developed to reduce the costs to deploy adaptive control systems, by consolidating the adaptive processing into a master control unit that supervises local field controllers. ACS-Lite downloads new split, offset, and cycle parameters to the local controllers every 5-15 minutes in response to changing traffic conditions. ACS-Lite is based on a very simple traffic model that has very few tunable parameters and requires modest calibration. Of all actuated systems, ACS-Lite may be the slowest to respond to rapid changes in traffic flows. ACS-Lite sends cycle, offset, and split values to the local controller. The gap-out and force-off logic of the controller works normally with the updated parameters.

## 9.5 PLANNED SPECIAL EVENTS, INCIDENT, & EMERGENCY MANAGEMENT

### 9.1.12 Overview

Signal timing can play a role in managing and even mitigating certain types of non-recurring congestion. In particular, high volumes of traffic generated by planned special events, reduced corridor capacity from roadway incidents, or increased travel demand triggered by region-wide evacuations can necessitate signal timing changes. A primary goal of readjusting signal timing in these circumstances would be to give priority to specified movements and to minimize the overall delay experienced by users from the non-recurring congestion. One way to achieve this objective would be to sustain and/or increase the throughput of traffic at certain intersections by increasing the green time for those movements. Traffic signals with modified timing settings perform this function by essentially “flushing” the preferred movement.

While increasing the roadway throughput, signal timing can also help to sustain or increase corridor capacity during an event. For example, during an evacuation from a hurricane, longer green times at signal lights for the preferred movement could enable an arterial to serve greater traffic volumes per hour.

Signal timing is modified during these events to meet other objectives besides those related to traffic operations and arterial performance, such as enhancing public safety. Traffic lights, for example, operating with these modified settings could reduce delays for emergency responders traveling to an incident on the roadway (9).

Another goal of changing signal timing would be to guide motorists to a certain destination, which could be the venue for a planned special event or the destination for evacuated traffic. More specifically, with certain roadways providing greater capacity or increasing traffic flow from the modified signal timing settings, motorists are “attracted” to these less-delayed corridors. In turn, these corridors become a pre-designated route for the traffic of a special event or roadway incident.

### 9.1.13 Techniques – Operational

The techniques to modify signal timing during special events, roadway incidents, or evacuations involve processes in traffic operations and planning, and some management and coordination at the policy and institutional levels. On the operations/planning side, an initial step would be to determine the specific route and intersections where traffic signals would be retimed. This route could be a particular arterial that is parallel to a certain freeway. In the event of an incident on the freeway, traffic could be diverted from the freeway to this arterial, which could provide additional capacity for a preferred movement with the modified timing settings in place at its various traffic signals.

Before signal timing is changed in response to certain non-recurring traffic, field observations or collected data would need to be collected before and during the event to determine whether such adjustments are warranted. This data may include traffic volumes/speeds and other traffic flow characteristics to detect the occurrence of a roadway incident or the presence of special event traffic. In one example, video cameras placed along certain arterials could help to verify whether an accident is causing non-recurring congestion. In another example, sensors placed at the ingress/egress points of various parking garages could determine when motorists are leaving a special event (such as a basketball game). These devices could indicate when to activate special signal timing plans.

Once the non-recurring traffic is detected, a particular change to the signal timing settings may be to increase green time for preferred or diverted movements. Meanwhile, the signal timing for the other movements remains the same. This results in an increased cycle length.

After making such adjustments, “real-time” monitoring would be required to ensure a more rapid implementation of appropriate timing plans and to permit operator manual control as needed (9).

### 9.1.14 Policy/Institutional Strategies

These operational procedures may not be effective or possible without a sufficient level of coordination among the jurisdictions impacted by the traffic from special events or emergency situations. This inter-jurisdictional coordination is needed among such institutions as law enforcement, public safety organizations, and various transportation/transit agencies to share resources, seamlessly exchange the required information, and to implement the required traffic control/signal timing plans (10). Overall, adjusting signal timing can reduce delays for motorists during special events, roadway incidents, or evacuations.

### 9.1.15 Example Implementations

***Coordinated Highways Action Response Team (CHART) – State of Maryland***

CHART is a cooperative group between the Maryland Department of Transportation, Maryland Transportation Authority, and Maryland State Police with several other federal and local agencies. This group started in the mid-1980s and focused on improving travel to and from the eastern shore, but has developed into a multi-jursidictional and multi-discplinary program ranging from Baltimore to the greater Washington D.C. area.

As part of the CHART system, many of the area highways are monitored via system detectors, video cameras, a cellular phone system, and weather sensors. Once an incident is detected, information is provided to the stakeholders and users. The stakeholders determine what is needed to mitigate the incident, perhaps a tow truck or lane closure. The user is informed of alternate routes or notified to expect delays. It is estimated that the program results in a total delay time reduction of **almost 30 million vehicle-hours,** and a total fuel consumption reduction of approximately **5 million gallons** (11)

The incident detection is integrated with the signal system software and allows pre-set signal timing plans to be implemented when needed. The system is developed so that when a change occurs in the traffic conditions on the highway, a message is given to various stakeholders. Then the stakeholders will decide if a diversion and a change of signal timing is needed for adjacent arterials.

***City of Portland***

The City of Portland, in conjuction with the Oregon Department of Transportation and Multnomah County, has developed various incident management plans along Interstates 5 and 205. A key to the plan is detection of incidents on the highway system. Upon detection—either by video, system loop detectors, or visual—dynamic message boards are programmed to inform motorists of the incident and possible diversion. If there is a diversion, the signal timing along adjacents arterial is modified to accommodate the influx of volume.

## 9.6 WEATHER-RELATED FACTORS THAT INFLUENCE SIGNAL TIMING

### 9.1.16 Weather-Related Factors Overview

Traffic signals are exposed to a range of weather conditions that may influence how users move through the intersection. Depending on the region, these conditions can include heavy rain, thunderstorms, slush, ice, and even snow. Fog is another weather-related factor that can reduce visibility and make it more difficult for drivers to see an approaching intersection. There are also cases of extreme weather disturbances such as hurricanes, tornadoes, and blizzards that may necessitate changes in signal timing at the effected intersections.

### 9.1.17 Techniques – Operational

Any of these above conditions or extreme incidents could warrant a change in signal timing settings. For example, fog could reduce visibility to the point of lowering intersection approach speeds. Inclement weather can also lead to increasing headways between vehicles and reductions in saturation flow rates through an intersection (*12*). In particular, roadways can become significantly congested with traffic evacuating areas impacted by such extreme weather conditions as hurricanes. At the same time, during snowy conditions roadways may be less congested, but there is likely to be an increase in start-up lost time and a reduction in tire pavement friction at an intersection.

Due to these effects on roadways and traffic patterns, modifying signal timing plans could mitigate or prevent increased delays at intersections. In particular, changing signal timing settings during inclement weather may improve the use of the traffic signal display. For example, with the increasing headways between vehicles in bad weather, phase times can be readjusted to maintain signal coordination for the traffic stream.

Another benefit can occur along roadways congested with traffic leaving a bad weather region. For example, in Clearwater, Florida, system operators have increased green times along roadways for traffic leaving the nearby beaches in the event of thunderstorms. One study has determined that this signal timing modification resulted in reduced delay at intersections and thus increased “roadway mobility” along major artierials (15).

In spite of these performance improvements, it is still uncertain whether modifying existing signal timing plans would be necessary for maintaining or improving the level of service at an intersection. Fewer vehicles on the roadway during inclement weather result in lower delays and thus the overall performance of the signalized intersection would not significantly decrease even with revised signal timing settings. Results in a few studies show that corridor operations would not be “radically affected” by bad weather, thereby making a modified signal timing plan less necessary (*13*).

The retiming of traffic signals in bad weather could be beneficial from a safety perspective. In particular, increasing the “amber-all-red interval” could improve drivers’ ability to either pass through or stop at the intersection safely. At the same time, changing signal timing plans during bad weather could help to reduce traffic speeds that are unsafe for the weather conditions. In Charlotte, NC, traffic signals at intersections were modified during inclement weather to operate with increased cycle lengths or with peak period timing plans. The result of these actions led to a reduction in travel speeds, and thus have “[minimized] the probability and severity of crashes” according to one study (16).

## 9.7 REFERENCES


*1* Manual on Uniform Traffic Devices for Streets and Highways (MUTCD), 2003 Edition. United States Department of Transportation, Federal Highway Administration, Washington D.C., 2003.

*2* 3M™ Opticom™ Priority Control System, http://solutions.3m.com/wps/portal/3M/en_US/Traffic_Safety/TSS/Offerings/Product_Catalog/?PC_7_0_56OU_root=GST1T4S9TCgv&PC_7_0_56OU_output=html&PC_7_0_56OU_gvel=PFJZV1PKJ1gl&PC_7_0_56OU_vroot=1PGXVVLN9Xge&PC_7_0_56OU_node=W7MPSPH3XVbe&PC_7_0_56OU_theme=tss_portal&PC_7_0_56OU_command=AbcPageHandler, visited March 2, 2007.

*3* Rodegerdts, Lee, Brandon Nevers, Bruce Robinson, John Ringert, Peter Koonce, Justin Bansen, Tina Nguyen, John McGill, Del Stewart, Jeff Suggett, Tim Neuman, Nick Antonucci, Kelly Hardy and Ken Courage, "Signalized Intersections: Informational Guide." FHWA-HRT-04-091, Federal Highway Administration, Washington D.C., August 2004.

*4* Urbanik, Thomas, Scott Beaird, Doug Gettman, Larry Head, Darcy Bullock, Ed Smaglik, Rick Campbell, and Matt Albett, "National Cooperative Highway Research Program Project 3-66: Traffic Signal State Transition Logic Using Enhanced Sensor Information."

*5* "Traffic Signal Preemption for Emergency Vehicles: A Cross-Cutting Study, Putting the "First" in First Response." FHWA, NHTSA, Washington D.C., January 2006.

*6* "Review of Signal Operations in the Minneapolis Hiawatha Avenue Light Rail Corridor", Federal Highway Administration Resource Center, November 2004.

*7* ITS America, Transit Signal Priority (TSP): A *Planning and Implementation Handbook* May 2005.

*8* Montasir M. Abbas, Nadeem A. Chaudhary, Anuj Sharma, Steven P. Venglar, and Roelof J. Engelbrecht. Guidelines for Determination of Optimal Traffic Responsive Plan Selection Control Parameters. Report 0-4421-1. Texas Transportation Institute, 2004.

*9* Managing Traffic for Planned Special Events, United States Department of Transportation, Federal Highway Administration, Washington D.C. September 2003.

*10* Managing Travel for Planned Special Events Fact Sheet, Report No. FHWA-OP-04-034, United States Department of Transportation, Federal Highway Administration, Washington D.C. November 2003.

*11* U.S. Department of Energy, Energy Information Administration, Short Term Energy Outlook – January 2002, Washington, DC: 2002, p. 23

*12* Al-Kaisy, Ahmed & Freedman, Zachary. November 2005. Weather Responsive Signal Timing: Practical Guidelines. Paper submitted to the Transportation Research Board 85th Annual Meeting January 22-26, 2006.

*13* Maki, Pamela J. 1999. *Adverse Weather Traffic Signal Timing*. Prepared by Short Elliott Hendrickson, Inc., for the Minnesota Department of Transportation.

15 Federal Highway Adminstration, Mitretrek Systems. *Best Practices for Road Weather Management. Version 2.0. City of Clearwater, Florida Weather-Related Signal Timing*.

16 Federal Highway Adminstration, Mitretrek Systems. *Best Practices for Road Weather Management. Version 2.0. City of Charlotte, North Carolina Weather-Related Signal Timing*.