A Methodology for Obtaining Signal Coordination within A Distributed Real-Time Network Signal Control System with Transit Priority

François Dion

Research Scientist
Virginia Tech Center for Transportation Research, Blacksburg, VA
Phone: (540) 231-9619; Fax: (540) 231-5214; E-mail: fdion@ctr.vt.edu

Bruce Hellinga

Assistant Professor
Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada
Phone: 519-885-1211; Fax: 519-888-6197; Email: bhellinga@uwaterloo.ca

Accepted for Presentation at the 80th Annual Meeting of TRB and Publication in the Transportation Research Record

ABSTRACT

Real-time network signal control offer the potential to provide delay benefits over traditional fixed time and actuated control. The SPPORT (Signal Priority Procedure for Optimization in Real-Time) model is a fully distributed heuristic rule-based signal control method that explicitly considers transit priority. While a distributed architecture enables the network control problem to be decomposed in such way that local controllers can optimize individual intersections, it also prevents the explicit development of coordination along signalized corridors. In this case, coordination can only be achieved when local controllers are instructed to consider the timing plans of adjacent controllers, the vehicle departures from upstream adjacent intersections, and the projected vehicle arrivals at downstream adjacent intersections. This paper describes a coordination methodology that was developed for use within the SPPORT model to allow it consider traffic progression objectives. In this methodology, specific considerations are given to coordination with downstream signal timings, downstream queues, downstream transit activities, upstream signal timings, upstream queue spillback events, and upstream transit activities. An evaluation of the resulting model for a five-intersection arterial corridor with scenarios considering a range of traffic conditions finally shows the benefits that the application of the model can, particularly over fixed-time control.

INTRODUCTION

Over the past decades, numerous efforts have been devoted towards the development of efficient traffic responsive signal control methods. Examples of such methods are the OPAC (*1-3*), Rhodes (*4-6*), SCATS (*7*), SCOOT (*8*), PRODYN (*9*) and UTOPIA (*10*) traffic signal control systems. However, while past experiences show that real-time traffic signal control has a potential to improve traffic operations in urban areas through their ability to employ information from traffic detectors and automatically respond to detected changes in traffic demand, significant limitations currently exist when these methods are applied to networks in which passenger cars and transit vehicles share the right of way. First, none of the existing systems considers the effects traffic of transit vehicles stopping in the right of way to board and discharge passengers. While they are stopped, these vehicles can partially or completely block a traffic lane and create a bottleneck that may result in an inefficient use of green time. Furthermore, priority of passage is also often granted to transit vehicles without considering all the potential effects that sudden traffic signal changes might have on general traffic.

The *Signal Priority Procedure for Optimization in Real-Time* (SPPORT) model (*11*), was developed with the objective of addressing the above two limitations. A first effort resulted in the development of a model that could control individual intersections with a simple two-phase operation. The novelty of this model was the use of a heuristic rule-based signal operation procedure that generates timing plans on the basis of a certain number of proposed strategies. Later, Conrad et al. (*12*) expanded the model's applicability to any type of intersection configuration and signal control with more than two phases. They also enhanced the rule-based signal optimization procedure and redesigned the model's discrete-event microscopic traffic simulation module. Subsequently, an evaluation of the revised model for isolated intersections indicated its ability to provide benefits over optimal fixed time and actuated control for a range of traffic conditions (*13*).

Recent research efforts have extended the applicability of the SPPORT model to the control of signalized intersections within coordinated networks (*14-15*). As a result of these enhancements, modifications have been made to the traffic demand estimation process and the rule-based signal optimization module. In particular, a new method for determining the most appropriate time for implementing signal switching decisions that explicitly considers upstream and downstream coordination impacts has been developed and introduced in the signal optimization module.

This paper describes the methodology that is used to encourage progression within fully distributed network signal control systems. The scope of this paper is specifically limited to describing the methods developed for implementation within SPPORT. Consequently, limited performance evaluations of the model are presented in the paper. More exhaustive evaluations are the subject of a separate paper. In term of content, provides a brief description of the SPPORT model's system control framework. The following two sections then respectively provide descriptions of the signal optimization process and of the methods used to compute ideal phase implementation times. A fourth section then presents some evaluation results. Some conclusions and recommendations are finally made in the last section.

SYSTEM CONTROL FRAMEWORK

A general problem faced when designing traffic responsive signal control methods for coordinated networks is to develop methods capable of reacting to traffic variations while still maintaining a sufficient degree of coordination between adjacent intersections. Often, adding such flexibility may contradict the general objective of providing uninterrupted progression through sets of intersections. At the network level, it is usually desirable to control intersections so that vehicles released from one intersection will have a high probability of traversing downstream intersections without stopping. At the local level, the priority is usually to minimize stops and delays. Planning traffic movements on an network basis thus often involves constraining the operation of individual intersections, while providing optimum local control often implies destroying existing progression schemes to implement allow short-term tactical decisions. Therefore, the main problem is to design signal control methods that provide a reasonable balance between local and network control strategies in the presence of somewhat predictable variable demands.

System Architecture

Network real-time traffic signal control problems can be decomposed into a series of sub-problems of manageable complexity and size using either hierarchical or fully distributed approaches (*14*). In the hierarchical approach, the network optimization problem is transformed into a multi-level control problem with distinct optimization objectives at each level. Examples of this approach are found in UTOPIA and SCATS. In this case, however, developing an efficient hierarchical control structure necessitates finding ways to resolve the conflicts that often appears between the two control levels. Consequently, signal controllers in are either allowed to completely override the proposed network timings or subjected to network constraints reflecting average flow conditions that may not always correspond to prevailing demands. To date, no satisfactory solution to this problem has been found. In the distributed approach, which is exemplified by PRODYN, the network control problem is broken into a series of smaller problems that can be solved in parallel by independent processors. There are no upper or lower control levels, and no explicit global network optimization. Coordination between adjacent sub-systems then is achieved by strategically modeling the linkage that exists neighboring sub-systems.

The SPPORT model follows the distributed approach. In this case, the model attempts to provide signal control by entirely distributing the signal timing calculations among local controllers. This approach creates optimization problems that are independent from the size of the network. It also minimizes network optimization times by allowing each intersection to be optimized in parallel.

Real-Time Operation Principles

SPPORT provides real-time control using a discrete-time rolling horizon process similar to OPAC. Typically, the model generates a new phase plan for the next 40 to 120 seconds of signal operation every 5 to 10 seconds, depending on the model's parameter setup. Given the rapid rate at which signal timing plans are updated, only the decisions taken in the first few seconds of each new plan are then typically implemented. This allows for quick reactions to changes in detected traffic conditions.

Similar to PRODYN and UTOPIA, the model is also acyclic in nature, in that it does not consider pre-defined cycle times. New phase plans are built by evaluating at regular intervals whether the current phase should be terminated immediately or extended up to the next decision point. Signal timing changes thus occur only when certain traffic conditions, which will be described later, are met.

Network Control Principles

By dividing network control into a series of independent intersection optimization problems, the model loses the ability to coordinate large groups of intersections. In order to promote local coordination, SPPORT signal controllers are instructed to evaluate the effects of candidate signal timing plans on the approaches to the intersection they each control, as well as on links leading to neighboring intersections. This is achieved by providing each controller with a simulation modeling of dependent links and surrounding intersections similar to the one shown in **Figure 1**. In the SPPORT model for isolated intersections, only approaches to controlled intersection were considered. This often caused sub-optimal timing plans to be selected as candidate timings were evaluated only on the basis of their effect on approaching traffic. In the new model, since each intersection is in turn the one being optimized and an intersection with which coordination is attempted, any timing decision taken at one intersection will then gradually propagate its effect across the entire network. In this way, the entire network becomes interconnected and coordinated, even if all intersections are individually optimized. Global optimization cannot be guaranteed, but if all intersections are optimally operated, it can be hypothesized that the entire network would operate close to global optimal conditions.

Objective Function

To evaluate the performance of candidate signal control strategies, SPPORT requires the user to define the objective function that each signal controller should use. This done by assigning values to the coefficients k_d, k_s, k_{TT}, k_{TC} and o_v in **Equations 1** and **2**, which define a performance index *PI* combining stops, delay and travel time, and a terminal cost *TC* that estimates the delays incurred beyond the end of the decision horizon by vehicles left in a queue the end of the horizon as a result of the signal-switching decision taken during the horizon.

$$PI = \sum_{i=1}^{N_{sim}} o_v \left(k_d\, d_i + k_s\, s_i + k_{TT}\, TT_i \right) + \sum_{k=1}^{N_{int}} \left(k_{TC}\, TC_k \right) \quad [1]$$

$$TC_k = \sum_{j=1}^{N_{link\,k}} \left(\frac{{Q_{end\,j}}^2}{2\, q_{sat\,j}} \cdot \left[U_j + \alpha_R (1 - U_j) \right] + Q_{end\,j}\, R_{min\,j} \left[U_j \right] \right) \quad [2]$$

where:

where:	PI	=	Performance index.
	d_i	=	Delay incurred by vehicle i (seconds).
	k_d	=	Weight of delay ($k_d \geq 0$).
	k_S	=	Weight of stops ($k_S \geq 0$).
	k_{TC}	=	Weight of terminal cost ($k_{TC} \geq 0$).
	k_{TT}	=	Weight of travel time ($k_{TT} \geq 0$).
	α_R	=	Weight of stopped vehicles on an approach with a green signal with respect to stopped vehicles on an approach with a red signal ($\alpha_R \geq 0$).
	N_{int}	=	Number of intersections in network.
	N_{sim}	=	Number of vehicles entering the network during the control period.
	$N_{link\,k}$	=	Number of approach links to intersection k.
	$Q_{end\,j}$	=	Queue size on link j at end of decision horizon (pcu).
	$q_{sat\,j}$	=	Saturation flow on link j (pcu /second).
	U_j	=	Signal display on link j at end of decision horizon (0 if green, 1 if red).
	$R_{min\,j}$	=	Minimum remaining effective red interval on link j at end of decision horizon based on phase sequence and minimum green constraints (seconds).
	o_v	=	Relative importance of vehicles of type v ($o_v \geq 0$).

s_i = Number of stops incurred by vehicle i.

TC_k = Terminal cost for intersection k at end of performance evaluation period.

TT_i = Total travel time of vehicle i (seconds).

SIGNAL OPTIMISATION PROCESS

SPPORT makes signal-switching decisions following the heuristic rule-based optimization procedure of **Figure 2**, which was originally developed in response to concerns that exhaustive optimization procedures such a dynamic or linear programming may be too computationally demanding for real-time applications in networks with highly variable demands (*11*). The procedure is based on the recognition that signal switches usually occur after the realization of specific discrete traffic events. By ignoring all events that have no importance for the signal operation, the procedure specifically allows for a significant reduction in the number of potential switching combinations that need to be considered to find optimum solutions to traffic control problems.

Signal Switching Rules

Currently, the SPPORT model considers the following responses to key traffic events occurring on a given intersection approach:

Response to stop line queues:

1. If a queue of n vehicle exists and is not being served, start serving it as soon as possible.
2. If a queue is being served, continue serving it.
3. Switch the signal to green if a single stopped vehicle has been waiting for a green signal for more than n seconds.

Response to potential queue spillbacks:

4. Switch the signal to green if a stop line queue exceeds a user-defined length.
5. Maintain the current green signal if vehicles are still queued on the intersection approach passed a user-defined location.
6. Switch the signal to red if a queue on one of the approach's user-defined major exit links threatens to spill across the controlled intersection.

Response to incoming platoons:

7. If a platoon of *n* vehicles or more is approaching, switch the signal to green at a time that will allow the platoon to cross the intersection without being affected by vehicles stopped at the stop line.
8. If a platoon is being served, continue serving it.
9. If vehicles are crossing the stop line at a rate exceeding a user-defined value, continue serving them.

Response to incoming transit vehicles:

10. If a transit vehicle is approaching a transit stop, switch the signal to green at a time that will allow the vehicle to proceed uninterrupted up to its loading point.
11. If a transit vehicle is approaching the stop line, switch the signal to green at a time that will allow the vehicle to cross the intersection without having to stop.

Rule Prioritization

To account for the fact that different traffic events do not carry the same importance, SPPORT requires the user to prioritize the rules listed in the previous section. For each link, the user is required to assign a positive numerical value to each of the rules, with increasing values representing higher priority events. The values assigned to the different rules are directly converted into priority levels, except for Rules 1, 2, 7 and 8, for which priority levels are calculated using **Equation 3**.

$$P = Min(\alpha x^2, \beta) \qquad [3]$$

where:
- P = Priority level
- α = Incremental user-defined priority value.
- β = Maximum user-defined priority value.
- x = Current stop line queue size for Rule 1, queue size at green onset for Rule 2, Estimated platoon size for Rules 7 and 8.

Generation of Request Lists

At each intersection, the signal-switching decision process starts with the generation of signal requests reflecting the needs of current traffic demands on each intersection approach. This process is initiated by projecting traffic movements on each intersection approach and exit links using the discrete-event microscopic traffic simulation model embedded within SPPORT (*12*). The projection is done using traffic flow information provided by detectors installed at some distance upstream of the intersection, ideally just downstream of the upstream intersection, as is done with SCOOT. Similar to PRODYN, additional detectors located closer to the stop line or just downstream of a bus stop can also be used to provide updated traffic information along intersection approaches.

For approaches to the intersection being optimized, the one at the center of the each minim-network in **Figure 2**, projections are made assuming that a green signal is displayed on all approaches or can be displayed after the remaining minimum red has elapsed. This allows vehicle arrival rates at the intersection to be determined independent of previously generated signal timings. For the adjacent intersections, the current projected signal timing plans are assumed to remain in operation so that coordination impacts with these intersections could be fully considered.

Within the simulation model, traffic projections are made according to the traffic modeling provided by the user regarding link average speeds, saturation flows, bus dwell times, bus stop location along intersection approaches, and the degree to which link saturation flows are reduced near a transit stop when a transit vehicle is loading passengers. To improve the degree to which the simulation represents reality, real-time traffic information can also be used to update the simulation parameters at regular intervals; however, the description of how this could be achieved is beyond the scope of this paper.

After completing the traffic projection, traffic conditions at each decision point within the decision horizon are evaluated against the signal control rules and requests calling for either a green or red signal display at a specific time on a given link are then generated, together with their priority level, for each key traffic event identified. In this case, queue-related traffic events are easily determined by tracking the number and location of queued vehicles on intersection approaches over time. Platoon arrivals at signalized intersections are for their part determined not by tracking the movement of individual platoons, but by analyzing the projected sequence of vehicle arrivals at the intersection stop line. In this case, platoon dispersion between signalized intersections can be considered by randomly assigning varying speeds to the individual vehicles being simulated. Finally, transit needs are determined by keeping track of the position of each transit vehicle within the simulated network.

Signal-Switching Decision Process

The goal of the signal-switching decision process is to serve the highest number of high-priority conditions requiring immediate service while delaying service to the least number of high-priority conditions requiring service at some future time. To fulfill this goal, all candidate phases that can be implemented at a given time are first evaluated by compiling the benefits and future costs associated with their implementation at the decision point being considered. These benefits and costs are calculated by adding together the priority level of all requests that would be satisfied if the phase were implemented, and by subtracting the priority level of all requests that could not be satisfied in time. At the end, the phase having the highest rating is selected for implementation, or continued service if this phase is the one currently in operation.

Multi-Objective Optimization

While the use of prioritized events allows SPPORT to determine the relative importance of various traffic conditions, it is often difficult to determine beforehand which event should have the highest priority. To solve this problem, the user is allowed to provide the model with more than one prioritized list of events. In such case, the optimization algorithm generates a candidate signal timing plan for each list and then selects for implementation the one yielding the best performance index.

Signal Co-ordination Logic

Signal coordination between adjacent intersections is implemented through the three following channels:

- Rule for handling incoming platoons (Rules 7-9),
- Rules for handling queue spillback (Rules 4-6), and
- Request delaying logic.

On each approach, the combined effect of the incoming platoon rules is to promote the implementation of green signals when platoons are projected to reach an intersection. Since incoming arrival patterns are functions of the timings implemented at upstream intersections, any attempt to avoid stopping incoming platoons thus results in implicit signal coordination with the upstream intersections. This type of coordination is similar to the maximization of progression opportunities in TRANSYT-7F (*16*), where a progression opportunity is seen as the simple ability for a vehicle leaving an intersection at a given time to travel uninterrupted across the next intersection.

The main objective of the queue spillback rules is to prevent queues of vehicles from spreading across adjacent intersections. The first two rules in the group promote coordination with upstream intersections by attempting to dissipate queues of vehicles before they reach these intersections. The third rule, on the other hand, promotes coordination with downstream intersections by commanding a red signal switch to prevent the blockage of the intersection being optimized by a growing queue on one of its exit links.

The contributions of the request delaying logic to the signal coordination are for their part discussed in detail in the next section.

CALCULATION OF IDEAL REQUEST TIMES

The most important signal coordination component is the logic used to adjust the times at which green signals are requested on each intersection approach following the identification of key traffic events. When generated, traffic signal requests have an associated signal switch time that is initially calculated by the corresponding rule. For instance, a request to serve a queue usually calls for an immediate signal change to green. On the other hand, a request to serve an incoming platoon has a time corresponding to the expected stop line platoon arrival time minus the time required to serve all vehicles traveling in front of it. This section describes the adjustments that are made to these initial request times to better coordinate traffic signal operations with upstream and downstream intersections.

Coordination with Downstream Signal Timings

The first task performed by the request delaying logic is to adjust the times at which green signals are requested at an intersection to reflect current and projected signal timings at adjacent downstream intersections. The objective of this change is to delay the green initiation at the controlled intersection so as to allow vehicles to reach the downstream intersections during a scheduled local green signal.

For each green signal request, the calculation of the time adjustment begins by assuming that all vehicles will leave the controlled intersection at the initial request time and will travel uninterrupted up to the next intersection. Following these assumptions, the time of arrival of the first vehicle to reach the downstream intersection is determined and used to calculate the amount of time by which each request should be delayed at the controlled intersection to allow uninterrupted progression across the downstream intersection:

- If the first vehicle is projected to reach the intersection during a scheduled green, no adjustment is implemented.
- If the vehicle is projected to reach the intersection during a red interval and a future local green interval has already been scheduled, then the requests are delayed by an amount of time that would allow the vehicle to reach the downstream intersection just at the beginning of the scheduled effective green.
- If the vehicle is projected to reach the downstream intersection during a red interval and no future green interval has been scheduled, or after the period for which projected signal timings are known, then the request is delayed as a function of the soonest time at which a green indication can be displayed at the downstream intersection when following local phase sequence and minimum green constraints.

This delay calculation is only performed for the requests associated with approach links for which the user has identified at least one exit link with which coordination is to be attempted. If more than one coordinated exit link is specified, the ideal request time is calculated using **Equation 4** as the weighted average of all the individual ideal request times that were calculated for each exit link.

$$RT_{del} = RT_o + \frac{\sum_{j'=1}^{N_{exit\,j}} Turn_{j\,j'} \left(RT_{del\,j} - RT_o \right)}{\sum_{j'=1}^{N_{exit\,j}} Turn_{j\,j'}} \qquad [4]$$

where: RT_{del} = Delayed request time (seconds).

$RT_{del\,j}$ = Delayed request time based on traffic conditions on exit link *j* (seconds).

RT_o = Initial request time (seconds).

$Turn_{j\,j'}$ = Turning proportion from approach link *j* to exit link *j'*.

$N_{exit\,j}$ = Number of exit links from approach link *j*.

Coordination with Downstream Queues

To avoid unnecessary stops and delays, traffic should also arrive at the adjacent intersections after the complete dissipation of any stop line queue that has been generated during the last red interval. To promote these conditions, SPPORT further adjusts green signal request times on each intersection approach to consider queue dissipation times on the approach's main exit links. In this case, the objective is to delay the green initiation so that vehicles leaving the intersection could travel uninterrupted along the entire length of the links leading to downstream coordinated intersections.

As shown in **Figure 3**, the offset calculation begins at the stop line of the downstream intersection and then works its way back, segment-by-segment, up to the stop line of the intersection being optimized. In this case, a segment-by-segment approach is use since the simulation model used by SPPORT typically models traffic links using 25- to 50-meter segments.

For each segment, the soonest time at which vehicles can enter at its upstream end and travel uninterrupted along its full length is computed using **Equations 5** and **6**:

$$\Delta RT_{queue\ m} = \frac{q_{veh\ m}}{q_{max\ m}} - \frac{L_m}{V_m} \qquad [5]$$

$$q_{\max m} = \min[q_{sat\,m}, q_{max\,m-1}] \qquad [6]$$

where: $\Delta RT_{queue\ m}$ = Change in ideal request time due to queuing on segment *m* (seconds).

$q_{max\ m}$ = Maximum exit flow rate from segment *m* (pcu/second).

$q_{sat\ m}$ = Saturation flow of segment *m* (pcu/second).

$q_{veh\ m}$ = Number of vehicles on segment *m* (pcu).

L_m = Length of segment *m* (meters).

V_m = Average free-flow speed on segment *m* (meters/second).

At the beginning of the calculation, the initial ideal request time is assumed to correspond to the soonest time at which vehicles can start to cross the downstream intersection when considering minimum vehicle headway, the existing phase display, phase sequence and minimum green requirements, as well as green interval start-up lost times. The final ideal request time at the controlled intersection is then computed by applying **Equation 7** for each segment on the link being considered up to the stop line of the intersection approach for which the request has been generated

$$RT_{queue\ m} = max\left\lfloor t_{\Delta t}\ ,RT_{queue\ m\text{-}1} + \Delta RT_{queue\ m}\right\rfloor \qquad [7]$$

where: $RT_{queue\ m}$ = Delayed request time after consideration of segment *m* (seconds).

$t_{\Delta t}$ = Time of current decision point (seconds).

In these calculations, the maximum rate at which vehicles are assumed to be able to exit from a given segment corresponds to the lowest saturation flow rate of all the downstream segments that were previously considered. This reflects the constraining effect that physical bottlenecks such as those caused by stopped transit vehicles have on upstream traffic flow conditions.

Coordination with Downstream Transit Activities

Similar to the previously described adjustments, the primary goal of the adjustments presented in this section is to delay, by a minimum amount of time, the initiation of the green at an intersection so that vehicles released from that intersection will avoid the queuing or blockage caused by a transit vehicle dwelling in the right-of-way on a downstream link. In this case, the extent to which a request is delayed is a function of the maximum rate at which vehicles can pass a dwelling transit vehicle. If the reduced rate is below a user-defined threshold, the request delaying calculations are carried out so that vehicles released from the upstream intersection will arrive at the transit stop just after the transit vehicle has left the stop. In all other cases, the green initiation is delayed so that vehicles will reach the transit stop just after the queue that is expected to form behind the dwelling transit vehicle will have completely dissipated.

Equation 8 calculates the additional amount of time that green signal requests should be delayed at an intersection as a function of the temporary capacity reductions caused by transit vehicles dwelling in the right-of-way on the intersection's exit links.

$$\Delta RT_{transit} = t_{interf\ dwell} \left(1 - f_{transit} \frac{q_{sat\ m}}{q_{sat\ m-1}} \right) \qquad [8]$$

where: $\Delta RT_{transit}$ = Transit interference request delay (seconds)

$t_{interf\ dwell}$ = Interfering dwell time (seconds)

$f_{transit}$ = Proportion of saturation flow available on segment *m* during dwell time

$q_{sat\ m}$ = Saturation flow on segment *m* when no transit vehicle is dwelling (pcu/second)

$q_{sat\ m-1}$ = Saturation flow on segment *m-1* immediately upstream of transit stop segment *m* (pcu/second)

The request delay that is calculated in this case corresponds to the additional time required to serve the vehicles upstream of a transit stop as a result of the temporary bottleneck created by a transit vehicles that is partially or completely blocking traffic while dwelling. To illustrate, consider a queue of ten vehicles behind a loading transit vehicle. When there is no transit interference, only ten seconds are required to serve the ten vehicles at a rate of one vehicle per second. However, 20 seconds, 10 more than under normal conditions, are required to dissipate the same queue if the transit vehicle temporarily reduces by half the rate at which vehicles can pass the transit stop while it is dwelling.

In **Equation 7**, the interfering dwell time is the portion of dwell time during which the transit vehicle truly interferes with the progression of other traffic. Often, the interfering dwell time is less than the total dwell. For example, transit vehicles stopped in the right-of-way do not interfere with the progression of other traffic if they are loading passengers while stopped within a queue. Little interference also exists if the vehicles that are disrupted by a dwelling transit vehicle join the tail of an existing queue before reaching the downstream end of the link.

Figure 4 illustrates how the interfering dwell time is determined. The interfering dwell time starts at the soonest time at which vehicles leaving the transit stop can expect to travel without interruption along the remaining portion of the intersection approach. For a transit vehicle that has not yet reached its loading point, the interfering dwell time start at the vehicle's projected arrival time at the transit stop. The end of the interfering dwell time depends on whether the vehicles behind the transit vehicle can all pass the transit stop before the end of the loading period. To determine which case prevails, the time that would be required to serve all the vehicles currently upstream of the transit stop under the reduced flow rate is determined. If the service time is less than the transit dwell time, the interfering dwell corresponds to the service time. Otherwise, it can be expected that the transit vehicle will affect traffic conditions until it departs, thus causing the interfering dwell time to corresponds to the remaining dwell time.

When performing the calculations, the transit dwell times defined in the simulation model are used. These dwell times can be either fixed or varying according to a given probability distribution. In all cases, average observed dwell time should be used. Exact time cannot be used, as such times are usually not known before passenger loading has terminated. To automatically account for the variability of dwell times, traffic detectors capable of selectively detecting transit vehicles can be installed just downstream of transit stops. Since traffic signals are re-optimized every few seconds, the information that a transit vehicle has left its stop, or is remaining stopped longer than expected, could then be used to make appropriate adjustments in the traffic projections and signal timings.

Figure 5 presents two examples of request delaying showing application of the downstream queue and transit activities offset. Specifically, **Figure 5a** illustrates the case in which all the vehicles upstream of the transit vehicles can pass it before it finishes loading passengers, while **Figure 5b** illustrates the opposite case. In both cases, the contribution of the queuing and transit offsets in the total delay imposed to the requests considered is very apparent. **Figure 5b** also illustrates the final request offset if the reduced saturation flow rate during the transit dwell time is less than the user-defined transit interference critical flow rate. In this case, the request time is delayed so that vehicles would reach the transit stop immediately after the transit vehicle has finished dwelling.

Coordination with Upstream Timings

Another important problem considered by the request delaying logic is the possibility that the logic itself may cause undesirable control loops to occur. By delaying the green initiation at an intersection, vehicles would then arrive later at the downstream intersection. In turn, the optimizer at the downstream intersection might decide to change its projected signal timings and to delay the time at which a green signal would be displayed at its intersection to account for the shift in vehicle arrivals. This could then result in further delaying of the green signal requests and green onset at the upstream intersection and in the initiation of an undesirable control loop.

To avoid such situations, a limit is imposed on the maximum delay that can be assigned to a request on links for which coordination with upstream intersection is attempted. **Figure 6** ndicates how the maximum request delay is determined. **Figure 6a** illustrates the case in which an alignment of green intervals already exists and in which no request delay is therefore allowed. **Figure 6b** illustrates the case in which no progressive pattern currently exists. In this case, requests can be delayed up to the furthest point in time at which a signal switch from red to green can be implemented without interfering with the progression of the vehicles coming from the upstream intersection. Finally, no request delaying is considered for request that initially call for a signal switch beyond the maximum allowed delayed request time.

Coordination with Upstream Queue Spillback Events

Under normal conditions, the objective of signal optimization is to minimize driver's real and perceive delays. However, this control objective is generally inappropriate when one or more movements become over-saturated. When queues threaten to spill across upstream intersections, it becomes more important to serve the existing queues to avoid intersection blockage, thus shifting the control strategy from stop and delay minimization to throughput maximization.

To provide such a control shift, special coordination rules were developed to reduce the scope of the coordination with downstream traffic conditions on approaches on which a queue threatens to spill across the upstream intersection. As an example, **Figure 7** illustrates how potential spillback conditions can affect the application of the request delaying logic.

Figure 7a first illustrates a scenario in which there is currently no spillback potential from link A-B or link B-C. In this case, the main control objective remains the minimization of stops and delay and the request delaying logic is applied without restriction.

Figure 7b illustrates a second scenario in which the queue on link A-B exceeds the critical queue reach, but only a consequence of the disruptions caused by a transit vehicle blocking one traffic lane while loading. In this case, no special action is taken to reduce the scope of the coordination with link B-C on the basis that the number of vehicle on link A-B remains below the link's user-defined critical content, which typically correspond to the minimum between half the link's total queuing capacity and its queuing capacity downstream of the critical queue reach. The rationale for this operating criterion is that no immediate signal changes is needed to avoid potential spillback problems across intersection A if there is still enough available queuing capacity on link A-B to store incoming traffic from intersection A. If a queue threaten to spills across the intersection but link's A-B occupancy is less than 50 percent of its capacity, it is usually an indication that the problem is either not created by the downstream traffic signal, as in the case illustrated, or that the downstream signal had already been turned green in response to the growing queue, and so, that the queue is currently being dissipated and that no additional action is currently necessary.

Figure 7c finally illustrates a scenario in which the queue extends the critical queue reach and in which the number of vehicles on the links exceeds its critical content. This condition normally results in a request calling for an immediate green signal switch on link A-B to serve the queue. Normal application of the request delaying logic would then typically delay the request to allow the queue on link B-C to start moving before releasing traffic onto the link. If the delay is long enough, it could then cause the queue on link A-B to spill across intersection A. Since the queue on link A-B is obviously the product of the signal timing operation, the request delaying logic only considers in this case traffic conditions on the portion of link B-C that can accommodate the minimum number of vehicles that each request is expected to send across the intersection so as to delay the green initiation on link A-B as little as possible and reduce the potential spillback problem with intersection A. If there is not enough storage capacity on the receiving link B-c, then the request delaying logic is applied normally across the entire length of the downstream link B-C to avoid creating potential spillback conditions on this link.

MODEL EVALUATION

To evaluate the SPPORT model for network control, a series of test scenarios considering a north-south five-intersection arterial were developed. These scenarios consider three levels of traffic demand (low medium and high) as well as two types of vehicle arrival patterns (constant and peaking). The main difference between each demand level is a 25-percent increase in link traffic flows for all simulated links, with no change in turning percentages. The three levels of demand considered allow SPPORT to be evaluated in situations in which minimum green interval constraints the signal operation, in which the network operates near capacity, and in intermediate traffic conditions. For each demand level, the two types of arrival patterns that are considered featured the same average hourly flows. While the constant arrival pattern assumes that the instantaneous vehicle arrival rates on each link always correspond to the given hourly average rate, the peaking pattern assumes that the instantaneous flow rates gradually increase from 90 to 115 percent of the specified average hourly flow, before decreasing again to 90 percent.

Figure 8 first demonstrates the benefits of allowing SPPORT to consider alternative control strategies when determining optimum signal timing plans. As illustrated, a significant reduction in the value of the performance index is obtained when more than one prioritized list are used by SPPORT. In this case, the best control performance is obtained when the six defined lists are used. This result illustrates the difficulty faced by engineers in selecting a priori traffic signal control strategies that can results in optimal control decisions at each decision point. In the figure, specifically five of the six overall worst control strategies are associated with the use of a single list. However, the same lists produced much improved results when SPPORT is allowed to choose the best one at each decision point.

Figure 9 illustrates the change in traffic performance along the test arterial when signal control is switched from optimal TRANSYT-7F fixed timing plans to SPPORT. For this comparison, SPPORT control was performed using the six six control strategies described in **Figure 8**. As it can be observed, the application of the SPPORT model on the test arterial resulted in traffic improvements in all the scenarios considered. For both the constant and peaking demands, greater improvements were obtained with increasing levels of traffic demand. However, the more significant improvements occurred in scenarios where transit vehicles dwelling in the right of way create major disruptions to traffic progression, thus demonstrating SPPORT's ability to efficiently react to observed traffic conditions.

To evaluate the real-time applicability of the model, optimization times were finally compiled for scenarios considering various demand levels and number of prioritized lists. For scenarios considering a single list, optimization times varied between 0.2 to 3 seconds on a 233 MHZ-Pentium II computer, depending on the demand level considered. When six lists were used, optimization times remained between 5 seconds for most scenarios. The highest optimization time, 14 seconds, was obtained for a scenario considering high demands and six prioritized lists. While this last optimization time is a bit long, it nevertheless demonstrates the applicability of the methods in a real-time environment, especially if it considered that improvements in computer technology will most likely keep shrinking optimization times.

CONCLUSIONS AND RECOMMENDATIONS

This paper described a coordination methodology that was developed to allow the SPPORT model to effectively consider traffic progression objectives. Since the model features a fully distributed real-time control architecture, it cannot directly calculate and implement coordinated signal timing plans covering a large set of intersections. Within such a control environment, traffic progression between successive intersections can only be achieved by enabling local signal controllers to develop local signal timing plans that are sensitive to traffic conditions and projected signal timings at adjacent upstream and downstream intersections. Within the SPPORT model, such timing plans are generated using control rules handling incoming platoons and potential queue spillbacks. Rules attempting to adjust the times at witch green signal switches are requested on intersection approaches in response to the identification of key traffic events are also used. These rules, which the request delaying logic, specifically make adjustments to the green signal request times by considering the following effects:

- initial request time
- implemented and projected signal timings at downstream intersections,
- traffic conditions on downstream links,
- transit interference on downstream links,
- implemented and projected signal timings at upstream intersections, and
- potential queue spillback at upstream intersections.

This method of encouraging progression within a fully distributed signal control system that has been implemented in SPPORT satisfies several objectives. First, the effect of transit vehicles loading and discharging passengers within the shared right-of-way is explicitly considered. Second, the desired change in signal control objective, from minimizing stops and delays in under-saturated conditions, to maximizing discharge flow in over-saturated conditions, is also explicitly recognized. Finally, the signal control method is sensitive to the presence of queues on the downstream link that limit the storage and progression of new arrivals, as well as to the projected times when these queues are to be served.

An evaluation of the model with a five-intersection test arterial indicated the ability of the model to effectively generate efficient signal timing plans for a range of traffic conditions. In particular, the evaluations that were conducted demonstrated the benefits of considering alternative control strategies when generating optimum signal timings, the potential benefits that the SPPORT approach may provide over optimal fixed-time control, and the real-time applicability of the model.

Despite these results, it is recommended that this control structure employed by the SPPORT be evaluated more extensively in order to determine its potential benefits and its control characteristics under a range of traffic conditions. Of particular interest would be the determination of typical sets of prioritized rules for application in various traffic conditions. Such an evaluation should also provide insights into the relative importance of the different elements defined within the model, notably by examining the impacts that each element has on the minimization of stops and delays. Ultimately, a field should also be conducted to test in a real-world environment the effectiveness of the SPPORT optimization method.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada and the Ministry of Transportation of Quebec.

REFERENCES

1. Gartner, N.H. OPAC: A Demand Responsive Strategy for Traffic Signal Control. In *Transportation Research Record 906*, TRB, National Research Council, Washington, D.C., 1983, pp. 75-81.
2. Gartner, N.H., P.J. Tarnoff, and C.M. Andrews. Evaluation of Optimized Policies for Adaptive Control Strategy. In *Transportation Research Record 1324*, TRB, National Research Council, Washington D.C., 1990, pp. 105-114.
3. Pooran et al, 1999
4. Head, K.L., P.B. Mirchandani, and D. Sheppard. Hierarchical Framework for Real-Time Traffic Control. In *Transportation Research Record 1360*, TRB, National Research Council, Washington, D.C., 1992, pp. 82-88.
5. Head, K.L. An Event-Based Short-Term Traffic Flow Prediction Model. In *Transportation Research Record 1510*, TRB, National Research Council, Washington, D.C., 1995, pp. 45-52.
6. Dell'Olmo, P., and P. Marchandani. An Approach for Real-Time Coordination of Traffic Flows on Network. In *Transportation Research Record 1494*, TRB, National Research Council, Washington, D.C., 1995, pp. 106-116.
7. Sims, A.G., and K.W. Dobbinson. SCAT - The Sydney Co-ordinated Adaptive Traffic System – Philosophy and Benefits. *Proc., International Symposium on Traffic Control Systems*, Vol. 2B, University of California, Berkeley, pp. 19-41.
8. Hunt, P.B., D.I. Robertson, R.D. Bretherton, and R.I. Winton. SCOOT - A Traffic Responsive Method of Coordinating Signals. Report LR1014, Transportation and Road Research Laboratory, Crowthorne, England, 1981.
9. Henry, J.J, J.L. Farges, and J. Tuffal (1983). "The PRODYN Real Time Traffic Algorithm." In *IFAC Proceedings Series 1984,* No. 3, Pergamon Press, New York, pp. 305-310.
10. Donati, F., V. Mauro, G. Roncolini, and M. Vallauri (1984). "A Hierarchical-Decentralized Traffic Light Control System - The First Realization: Progetto Torino". Proceedings, 9th IFAC Triennial World Congress, Budapest, Hungary, pp. 2853-2858.
11. Han, B. and S. Yagar. Real-Time Control of Traffic with Bus and Streetcar Interactions. Proceedings, 6th IEE International Conference on Road Traffic Monitoring and Control, Conference Publication No. 355, IEE, London, 1992, pp. 108-112.

12. Conrad, M., F. Dion, and S. Yagar. Real-Time Traffic Signal Optimization with Transit Priority: Recent Advances in the SPPORT Model. In *Transportation Research Record 1634*, TRB, National Research Council, Washington D.C., 1998, pp. 100-109.
13. Dion, F. and B. Hellinga. A Rule-Based Real-Time Traffic Responsive Signal Control System: Application to an Isolated Intersection. Accepted for publication in Transportation Research B.
14. Dion, F. and S. Yagar. Distributed Approach to Real-Time Control of Complex Signalized Networks. *Transportation Research Record 1554*, TRB, National Research Council, Washington D.C., 1996, pp. 1-8.
15. Dion, F. Development and Evaluation of a Distributed System for the Real-Time Control of Signalized Netowks. Ph.D. Thesis, University of Waterloo, Waterloo, Ontario, 1998.
16. Hadi, A.H. and C.E. Wallace. Progression-Based Optimization Model in TRANSYT-7F. In *Transportation Research Record 1360*, TRB, National Research Council, Washington D.C., 1992, pp. 74-81.

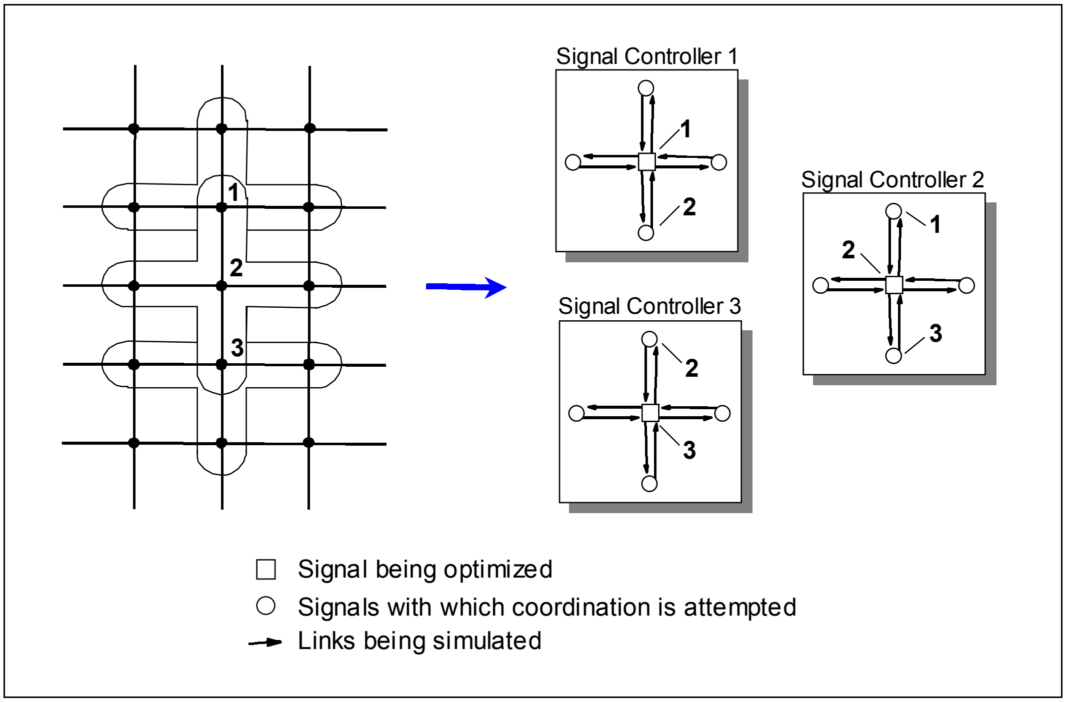


FIGURE 1. Individual Intersection Optimization Network Modeling

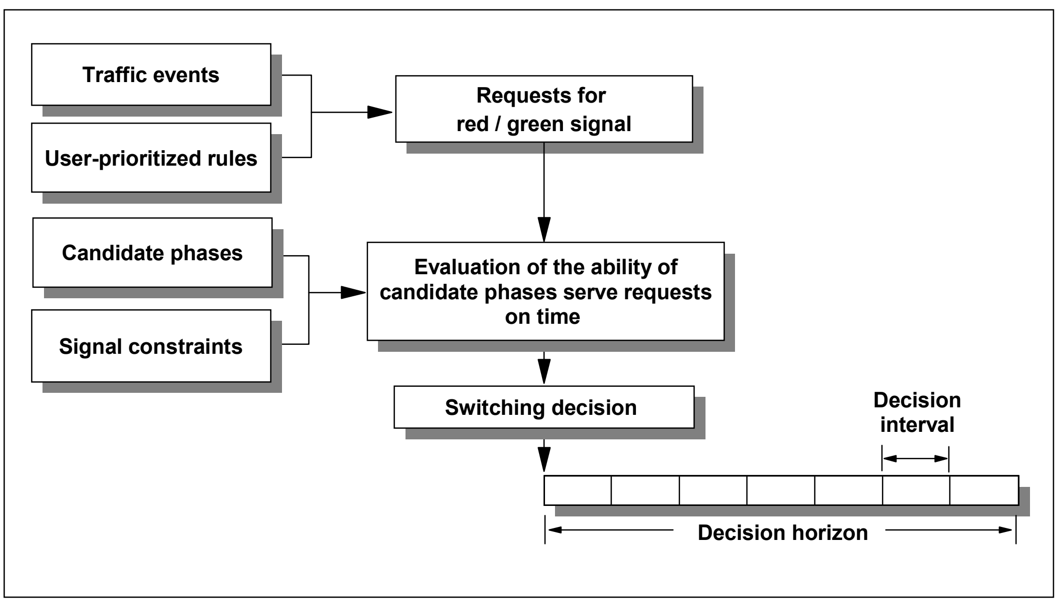


FIGURE 2. Rule-Based Decision Process

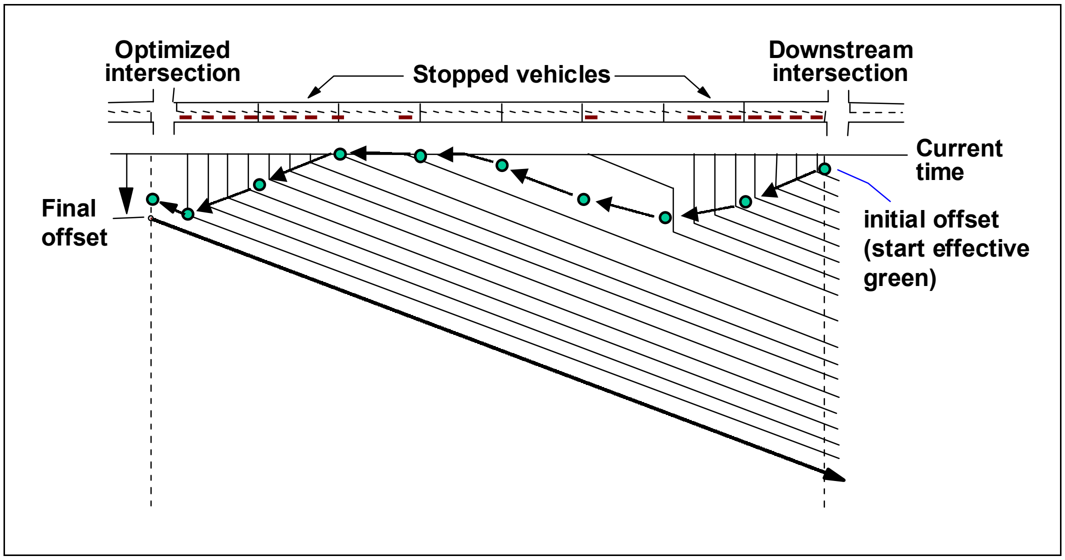


FIGURE 3. Signal Offset Calculation Example

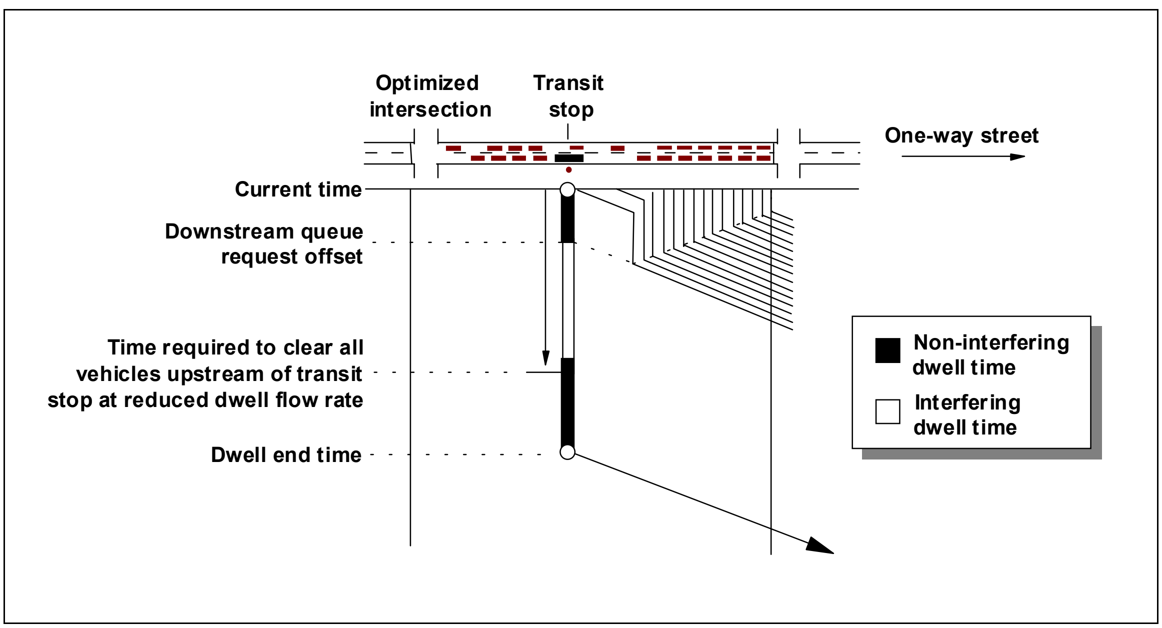


Figure 4 - Interfering Dwell Time Example

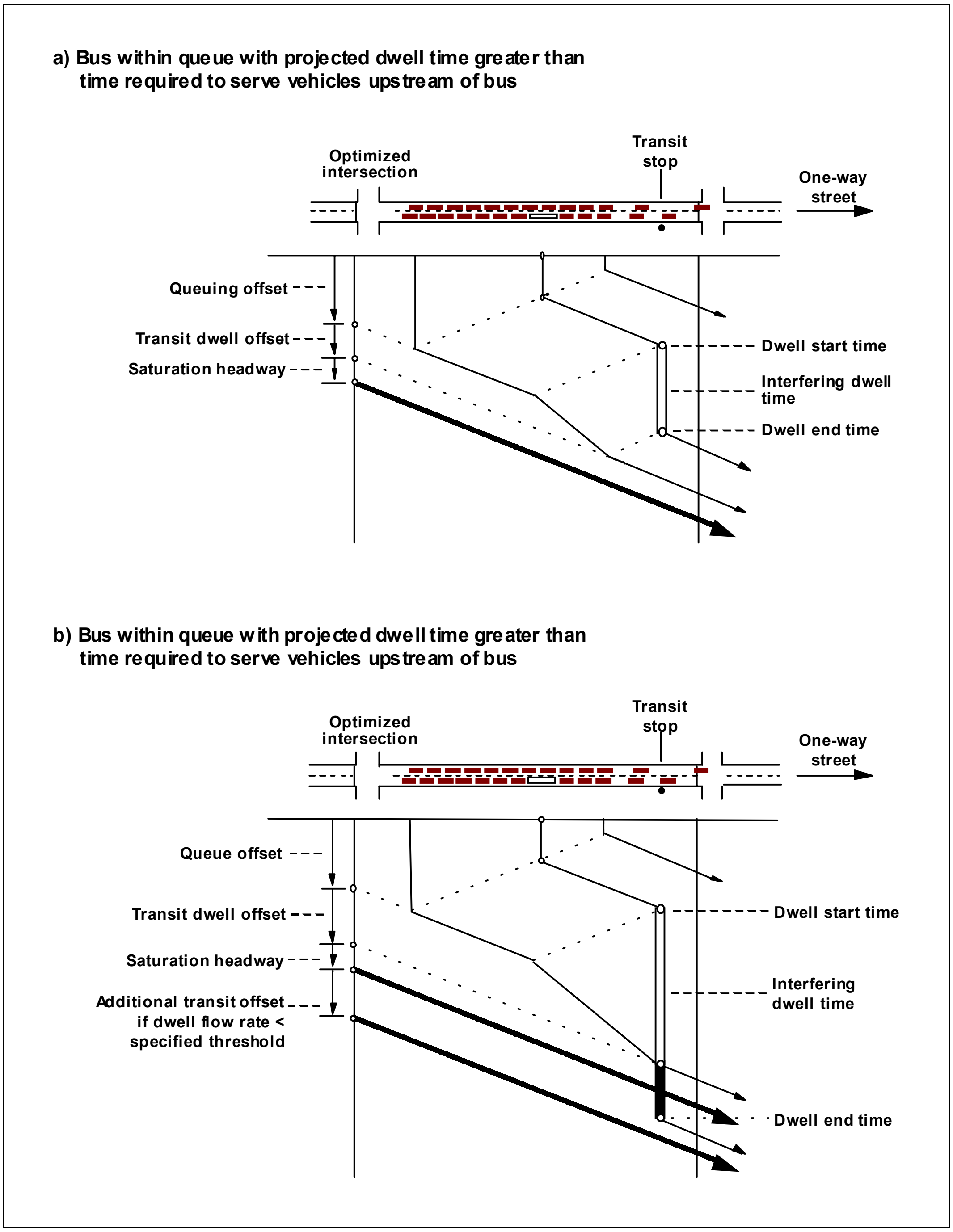


FIGURE 5. Examples of Request Delaying in Response to Queue and Transit Activities

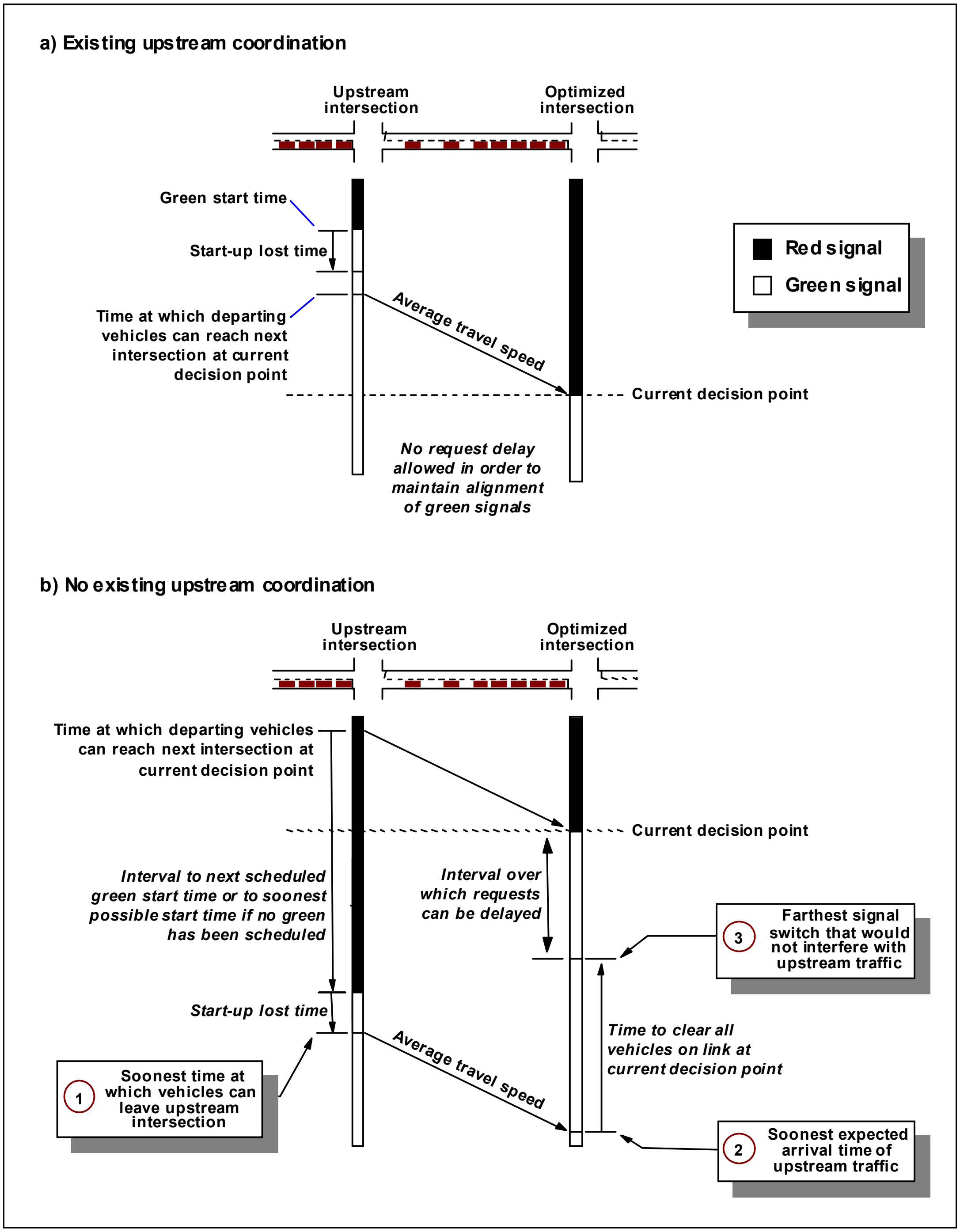


Figure 6 – Determination of Maximum Permitted Request Delay

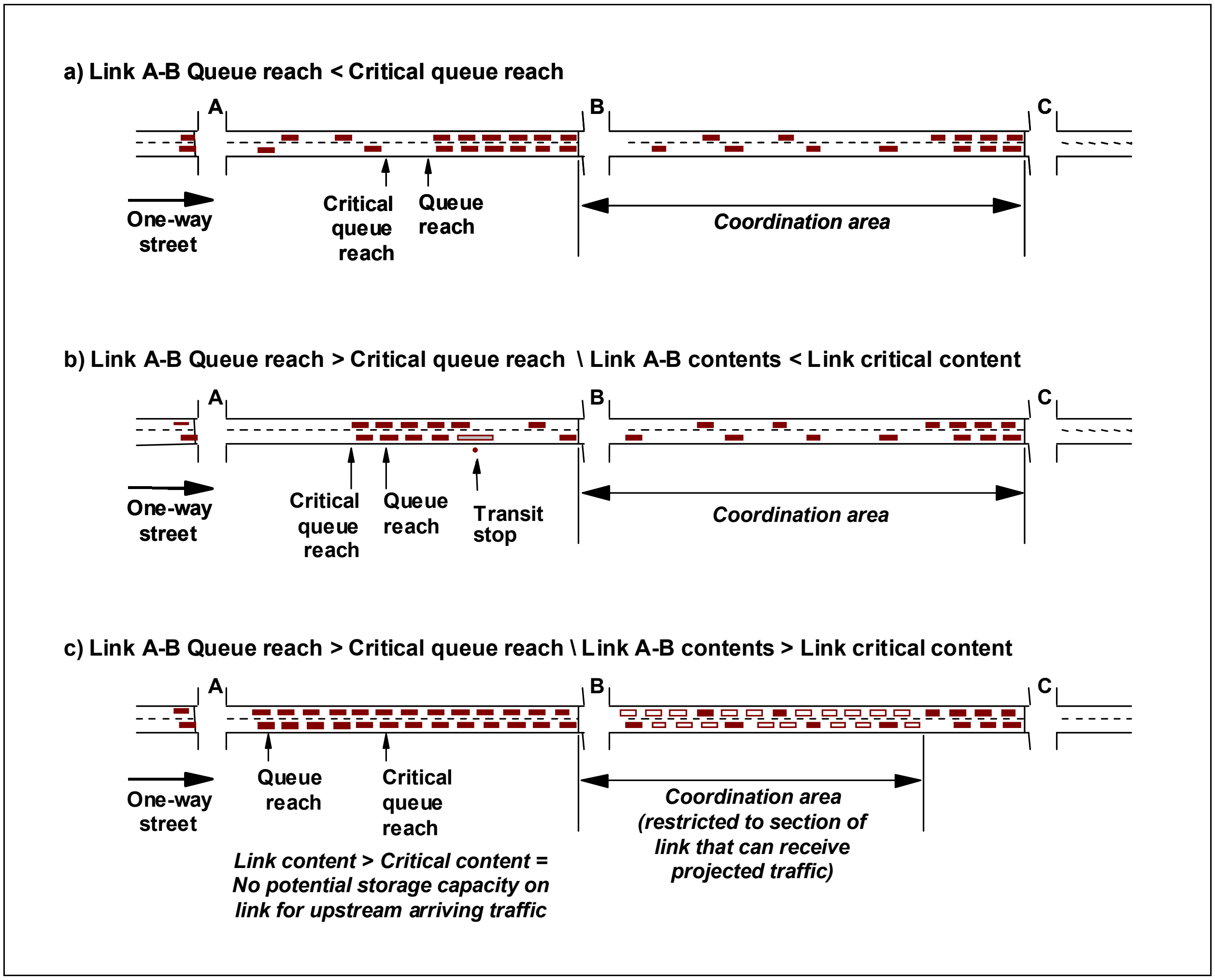
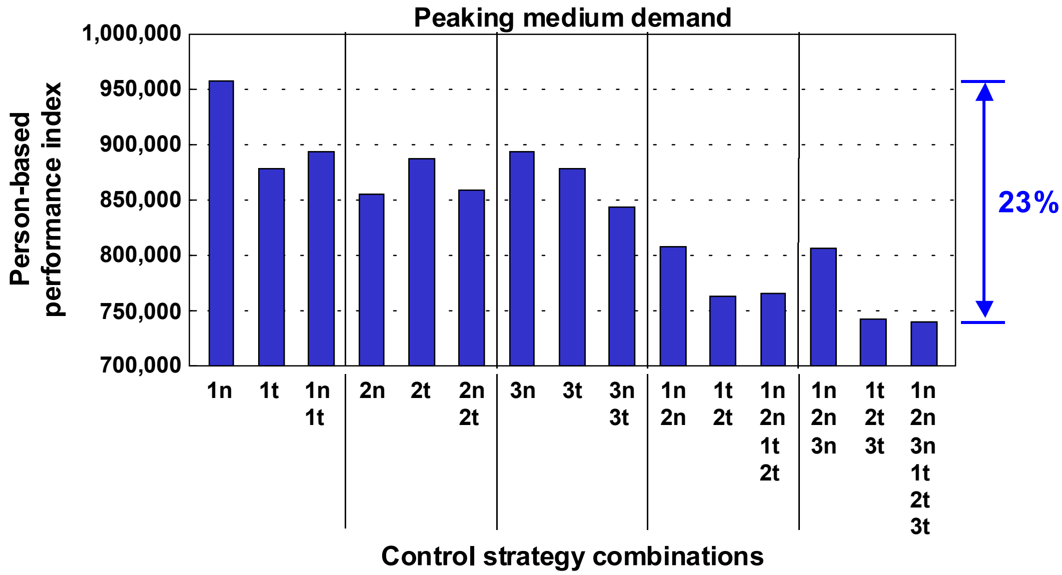


Figure 7 - Coordination Area on Intersection Exit Links within Request Delaying Logic



Prioritized control rule lists

1 = NB platoon progression
2 = SB platoon progression
3 = NB + SB platoon progression
n = Without transit priority
t = With transit priority

Performance index

$$PI = \Sigma_{cars}\ 1.5\ (d_i + 20\ s_i) + \Sigma_{buses}\ 60\ (d_i + 20\ s_i) + \Sigma TC_k$$

FIGURE 8. Effectiveness of Mutli-Objective SPPORT Operations

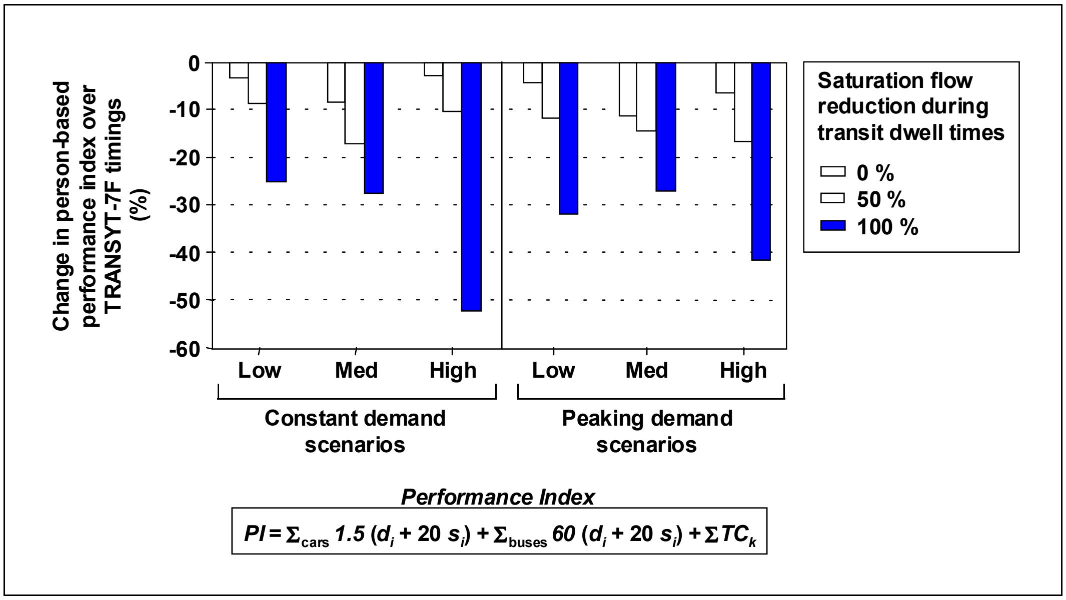


FIGURE 9. SPPORT performance against TRANSYT-7F optimal fixed timings on a five-intersection test arterial with mixed-traffic