TRAFFIC SIGNAL OPTIMIZATION WITH CONDITIONAL TRANSIT SIGNAL PRIORITY FOR CONFLICTING TRANSIT ROUTES

Eleni M. Christofa
Institute of Transportation Studies, University of California, Berkeley
109 McLaughlin Hall, Berkeley, CA 94720-1720
Email: christofa@berkeley.edu

Alexander Skabardonis, Ph.D.
Institute of Transportation Studies, University of California, Berkeley
109 McLaughlin Hall, Berkeley, CA 94720-1720
Email: skabardonis@ce.berkeley.edu

ABSTRACT

Transit Signal Priority (TSP) is a strategy that has been extensively used to improve transit operations in urban networks. However, none of the existing signal control systems has been able to effectively address issues such as the impact of TSP strategies on auto traffic or the provision of priority to transit vehicles traveling in conflicting directions at traffic signals. There is a need to develop a TSP system that explicitly accounts for vehicle occupancy and the impact of the signal timings on the cross-street traffic while addressing the issue of conflicting transit routes in a systematic way.

A traffic responsive signal control system for signal priority on conflicting transit routes that incorporates vehicle occupancy is presented. The signal control system provides signal timings that minimize the total delay of a single intersection, while assigning weights to the vehicles based on their occupancy. The system presented is tested through simulation at a single intersection located in Athens, Greece. The intersection under consideration is characterized by heavy auto and transit traffic and nine bus lines traveling in conflicting directions. The results indicate that using the developed optimization process substantial reductions in the transit users' delay and the total person delay at the intersection can be achieved with only small increases in the delays of the auto users.

Keywords: Transit Signal Priority (TSP), signal control systems, delay, mathematical models

INTRODUCTION

There is an increasing need for designing efficient multimodal transportation systems to improve mobility in urban areas. In order to do so in a system optimal and equitable way high occupancy vehicles, mainly transit vehicles, should be treated differently. A way to achieve this is by granting priority to transit vehicles at bottlenecks such as signalized intersections, which are responsible for a big portion of their delay. Prioritizing transit vehicles through improvements in facility design (e.g., bus lanes) is not always feasible because of geometric and design restrictions. As a result, there is a clear need to optimize signal control systems such that they balance their treatment for transit and auto users by minimizing total person delays in the system.

Transit Signal Priority (TSP) is an emerging operational strategy that facilitates efficient transit operations by providing priority to transit vehicles at signalized intersections. TSP strategies have been implemented in several urban areas in the U.S. and Europe. Many studies report significant reductions in transit vehicles' control delay and an overall improvement of their operations. However, they are often disruptive for the auto traffic, leading to substantial increases in their delay. Most importantly, the majority of such implementations are site-specific and restricted to provide priority to transit vehicles traveling only on non-conflicting routes, leading to inequitable treatment among transit users.

Demand for balanced multimodal systems requires that signal control systems provide optimized signal settings for the auto traffic, such that the negative impacts of TSP on their delays are minimized. Combining traffic signal optimization with TSP strategies is the most cost-effective way to improve the level of service for transit operations and minimize the total person delays in signalized networks. This paper describes the development of a real-time, traffic responsive, signal control system with TSP for an isolated intersection that provides priority to transit vehicles even when traveling in conflicting directions while minimizing the negative impacts on the auto traffic. This is part of a major effort to develop TSP strategies for a range of operating conditions for arterials and grid networks. The system is expected to reduce delays for the transit vehicles and as a result improve the reliability of the transit system. Improvements in fuel consumption and pollutant emissions are also expected.

This paper is organized as follows: First, the background on TSP strategies and a summary of signal control systems with TSP are presented. A description of the methodology and the study site used for testing the traffic responsive signal control system follows. The results from the simulation tests performed and a discussion conclude the paper.

BACKGROUND

Transit Signal Priority Strategies

Existing transit signal priority strategies fall into two major categories: passive and active priority strategies. Passive priority strategies are developed off-line based on historical data, and they do not require any detection system. They mainly include changes in the signal settings (green times, offsets, and cycle lengths). Examples include adjustment of offsets to account for the slower bus speeds and the midblock dwell time (Skabardonis, 2000), and/or additional green time to the phases that serve transit vehicles so that the probability of a transit vehicle arriving at the intersection during the green interval is increased. Reduction in the cycle length is another passive priority strategy commonly used since it increases the turnover of the phases and as a result, decreases the delays for all vehicles.

Passive priority strategies are inexpensive to develop and easy to implement. However, their success depends on the validity of the assumption of non-variable traffic volumes. In addition, such strategies assume deterministic dwell times at the transit stops, which is not realistic for most transit operations.

Active priority strategies respond to traffic variations in real-time, and as a result they are usually more effective than passive priority strategies. Information in real-time about the transit vehicles' speed and location, obtained by sensing technologies, is required for the design of such strategies. Active priority strategies include holding the green until the transit vehicle clears the intersection (phase extension), or advance the start of the green for the phase(s) serving transit vehicles (phase advance). Other options include inserting a new phase that can serve the transit vehicle at the moment it is approaching the intersection (phase insertion) or rotating the phases so that transit vehicles are served as soon as possible (phase rotation).

While active priority strategies can be used in real-time and be more effective in improving transit operations than the passive priority strategies, they require detection and communication systems that increases their cost, without any guarantee of succeeding on a network-level basis. Active priority strategies often have detrimental impacts on the non-transit traffic (mainly the cross-street traffic), can cause confusion to the motorist, and in many cases are responsible for loss of signal coordination and interruption in the progression of the vehicle platoons which can result to excessive delays (Chang and Ziliaskopoulos, 2003; Skabardonis, 2000).

Signal Control Systems with TSP

A number of real-time signal control systems that incorporate active priority strategies exist in the literature. Such systems use detection of vehicular traffic at some point upstream and/or downstream of an intersection to predict the traffic conditions and adjust the signal settings in real-time. Using the available information the signal settings are optimized on a decision horizon in the order of one cycle to a few minutes (traffic responsive systems) or on a rolling horizon concept (adaptive systems). Contrary to traffic responsive systems, adaptive signal control systems do not maintain the concept of a cycle and as a result, their optimization process is characterized as acyclic.

Active priority strategies have been implemented with SCOOT (Split, Cycle, and Offset Optimization Technique), a cyclic model which optimizes phase splits, and cycle lengths in real-time based on saturation level constraints as well as offsets such that traffic progression improvements can be achieved (Hunt et al., 1982). Priority to transit vehicles is provided through phase extension or advance conditional on schedule-based and headway-based criteria, only when traffic conditions are below user defined levels of saturation (Bretherton et al., 2002). A similar TSP logic is followed by SCATS (Sydney Co-ordinated Adaptive Traffic System) (Cornwell et al., 1986). Another traffic responsive signal control system with TSP was recently developed by California PATH (Li, 2008). The system provides priority based on a trade-off between bus delay savings and the impact on the rest of the traffic.

PRODYN incorporates transit priority by including cost elements for the transit vehicles in the objective function that is optimized over a rolling horizon. The cost elements are weighted based on the priority level assigned *a priori* to each transit vehicle and its direction (Henry and Farges, 1994). The UTOPIA (Urban Traffic OPtimization by Integrated Automation) system in Turin, Italy, is an adaptive signal control system that provides unconditional priority to selected bus routes by continuously optimizing the signal settings on a rolling horizon concept and simultaneously improving mobility for private vehicles (Mauro and Di Taranto, 1989). SPPORT (Signal Priority Procedure for Optimization in Real-Time) is another adaptive signal control system that incorporates rule-based transit priority, while accounting for the impact of stopped transit vehicles on traffic operations (Conrad et al., 1998). Finally, the Centralized TSP system, an adaptive signal control system implemented in Los Angeles, provides priority to buses based on their schedule lateness (Li et al., 2008).

There are several issues that have not been successfully addressed by the systems described above. First of all, none of the existing systems addresses the issue of providing priority to transit vehicles traveling in conflicting directions in a systematic way. Existing work has dealt with this issue either by predetermining which route gets priority (Henry and Farges, 1994) or by constraining the implementation of the system on networks that include only transit vehicles traveling in non-conflicting directions (Cornwell et al., 1986). Moreover, transit priority is often provided unconditional on specific criteria, such as occupancy and schedule delay (Li, 2008), which would ensure improvement in the operations of transit vehicles while protecting cross-streets from reaching oversaturated conditions. In addition, the existing systems do not take into account the differential in the occupancy of autos and transit vehicles, instead optimizing their systems on a per vehicle basis (Bretherton et al., 2002; Cornwell et al., 1986). The provision of priority is often rule-based and as a result not explicitly included in the optimization process. Existing traffic signal control systems are based on site-specific implementations (Mauro and Di Taranto, 1989; Li et al., 2008), restricted mainly to isolated intersections and single arterials and do not result from systematic analyses, limiting even further their applicability in the real-world.

METHODOLOGY

We propose a traffic responsive signal control system to minimize the total person delay at the traffic signals. The goal of the formulation is to optimize the signal timings, such that conditional priority is granted for the transit vehicles based on their occupancy. Conditional priority is used as a way to assign priority when two or more transit vehicles that are candidates for priority are expected to arrive at the intersection at approximately the same time. In addition, the impact of TSP on the auto delays at the intersection is taken into account through the inclusion of the total person delay in the objective function for all the vehicles present at the intersection.

The formulation is based on the assumption of undersaturated conditions and constant cycle lengths with negligible lost times between phases. The auto arrivals as well as the service times for all vehicles at the signalized intersection are assumed to be deterministic. The arrivals of the transit vehicles at the intersection are assumed to be known in real-time. The phase design and sequence are predetermined and as a result, only phase extension and advance can be implemented. Pedestrian traffic that could add further constraints on the green interval durations is ignored for the moment. Another assumption is that the transit vehicles travel on mixed traffic lanes. However, the formulation holds even when dedicated rights-of-way exist.

The mathematical program minimizes the total person delay at the intersection, by changing the green times for each phase i, G_i within the cycle under consideration (indexed by T), constrained by the minimum green times for each lane group[1] j, $G_{j\,\min}$ and a fixed cycle length, C. The mathematical program that optimizes the signal settings for any design cycle T, is as follows:

$$\begin{aligned} \min \quad & \sum_{a=1}^{A} o_a d_a + \sum_{b=1}^{B} o_b d_b \\ \text{s.t.} \quad & \sum_{i \in I_j} G_i \geq G_{j\,\min} \\ & \sum_{i=1}^{N} G_i = C \end{aligned} \tag{1}$$

where:
o_a: passenger occupancy of auto a $[\frac{\text{pax}}{\text{veh}}]$
o_b: passenger occupancy of transit vehicle b $[\frac{\text{pax}}{\text{veh}}]$
d_a: control delay for auto a $[\frac{\text{sec}}{\text{veh}}]$
d_b: control delay for transit vehicle b $[\frac{\text{sec}}{\text{veh}}]$
A: total number of autos served during the design cycle T or the next one $T+1$
B: total number of transit vehicles served or arrived during the design cycle T
G_i: green time allocated to phase i [sec]
$G_{j\,\min}$: minimum green time allocated to lane group j [sec]
C: cycle length [sec]
I_j: set of phases that can serve lane group j
N: number of phases in a cycle

The objective function consists of the summation of the person delay for the auto and the transit vehicle passengers, which are explained in more detail in the following sections, along with the constraints of the mathematical program.

Auto Delay

The person delay for the auto passengers consists of the summation of the person delay that corresponds to the autos that will be served during the design cycle T, and the expected delay for those that will be served during cycle $T+1$. Delays for the passengers that are served during the next cycle $T+1$ need to be included to account for the impact that the design of the signal timings in the current cycle will have on the delays of the next. If no such expected measure of delay were to be included, the optimized scenario for the design cycle would be to the provision of the minimum green times to all the phases apart from the last one, which would increase the auto delay for the next cycle substantially.

More specifically, the total delay D_j for one cycle for lane group j for vehicles that arrive deterministically with rate q_j, are served at saturation flow s_j, and experience a red time interval of $R_j = C - G_j^e$ and a green ratio of $\lambda_j = G_j^e/C$, where G_j^e is the the summation of the effective green

[1] A *lane group* is defined as one or more adjacent lanes (at each intersection approach) that can be served by the same phases (Highway Capacity Manual, 2000).

times for all the phases that can serve lane group j is given by:

$$D_j = \frac{1}{2}\frac{q_j C^2(1-\lambda_j)^2}{(1-\frac{q_j}{s_j})} \tag{2}$$

where $C(1-\lambda_j)$ is the red time lane group j experiences.

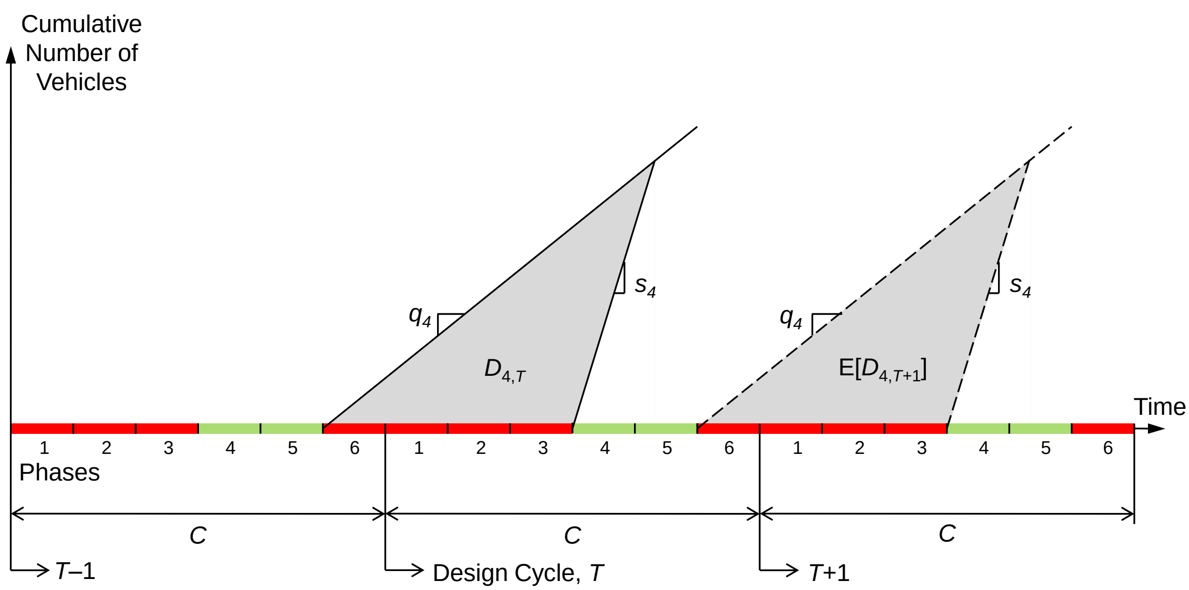

Figure 1 – Queueing Diagram for Lane Group 4 (Auto Delay)

Figure 1 illustrates the delay for the vehicles of a lane group j. It shows the cumulative number of vehicles present at an intersection for cycles $T-1$, T, and $T+1$ for lane group $j=4$. According to Figure 1 lane group 4 can be served by phases 4 and 5, so its effective green time will be: $G_4^e = G_4 + G_5$. The shaded area represents the total delay for the autos that belong to lane group j and are served by the design cycle T (solid lines) and the next cycle $T+1$ (dotted line). Such queueing diagrams can be drawn for all lane groups to allow for the estimation of the delay for autos and transit vehicles under the assumption of first-in-first-out (FIFO) queueing discipline. Next, the two cases for calculating auto vehicle delay are described. The examples illustrate how to calculate the red times for a lane group in each of the cases.

1. The total delay for the autos served by cycle T is derived from equation (2) as follows:

$$\begin{aligned} D_T &= \sum_{j=1}^{J} D_{j,T} \\ &= \frac{1}{2}\sum_{j=1}^{J}\frac{q_j}{1-\frac{q_j}{s_j}}\left(\sum_{i=l+1}^{N} G_{i,T-1} + \sum_{i=1}^{k-1} G_{i,T}\right)^2 \end{aligned} \tag{3}$$

where:
J: total number of lane groups
T: cycle index
q_j: arrival rate for lane group j $[\frac{\text{veh}}{\text{hr}}]$
s_j: saturation flow for lane group j $[\frac{\text{veh}}{\text{hr}}]$
l: the last phase in a cycle that can serve lane group j
k: the first phase in a cycle that can serve lane group j

Example: Vehicles in lane group 4, that can be served by phases 4 and 5, experience red time equal to the summation of the green time of phase 6 in the previous cycle $T-1$ and the green times of phases 1–3 in the design cycle T (Figure 1).

2. The total expected delay for the autos served by cycle $T+1$ is derived from equation (2) as follows:

$$E[D_{T+1}] = \frac{1}{2}\sum_{j=1}^{J}\frac{q_j}{1-\frac{q_j}{s_j}}\left(\sum_{i=l+1}^{N}G_{i,T}+\sum_{i=1}^{k-1}G_{i\min,T+1}\right)^2 \tag{4}$$

The expected delay that the autos will experience in the next cycle $T+1$, is estimated here assuming the most optimistic scenario that the next cycle $T+1$ will be designed to have the minimum green times, $G_{i\min}$ for all the phases except for the last one, which will be assigned the residual of the cycle length. The use of minimum green times results in the lowest possible delays for all vehicles that are served in cycle $T+1$. The choice of the expected signal timings is a user-specified factor which does not affect the structure of the formulas presented.

Example: Vehicles in lane group 4, that will be served by phases 4 and 5 in the next cycle $T+1$, experience red time equal to the summation of the green time of phase 6 in the design cycle T and the green times of phases 1–3 in cycle $T+1$ (Figure 1).

As a result, the first component of the objective function (auto person delay) becomes:

$$\sum_{a=1}^{A} o_a d_a = \bar{o}_a(D_T + E[D_{T+1}]) \tag{5}$$

Regarding the auto occupancy, an average value per auto $\bar{o}_a$ is used because total vehicle delay is calculated directly rather than accounting for each vehicle separately.

Transit Delay

The person delay for the transit vehicles consists of the summation of the person delay that corresponds to the transit vehicles that are served during the design cycle T and the expected delay for those that arrive in T but are served in $T+1$. Under the assumption that information on the transit vehicles' location and arrival times is available only for the design cycle, the transit vehicles that arrive during cycle $T+1$ are not taken into account. The exclusion of such vehicles is not expected to affect significantly the results. The signal timings for one cycle earlier are necessary in order to determine the delays of the transit vehicles that arrive during cycle $T-1$ but will be served during the design cycle T. In addition, real-time information about the transit vehicles' location is required in order to determine their actual arrival times.

The arrivals of the transit vehicles at the intersection are assumed to be known in real-time. As a result, the delay that each transit vehicle experiences is equal to the delay an automobile arriving at the same time at the lane group's queue would experience, and thus it can be calculated using queueing diagrams (Figure 2).

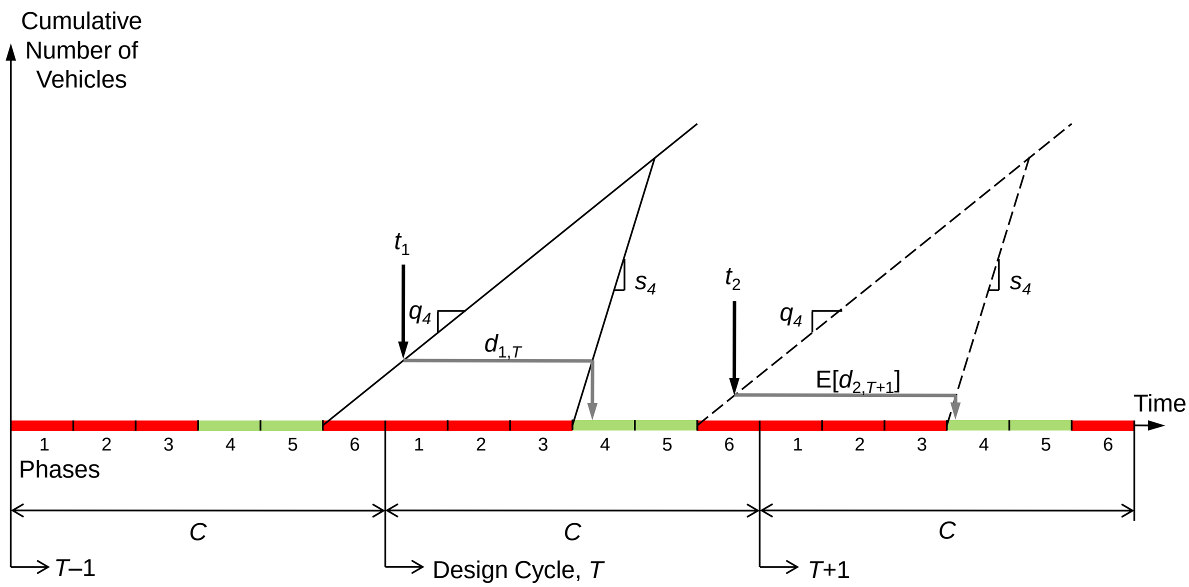

Figure 2 – Queueing Diagram for Lane Group 4 (Bus Delay)

The estimation of the transit delay used in the optimization of each cycle T depends on the actual arrival time of each of the transit vehicles t_b which can be summarized in the following two cases:

1. If a transit vehicle that belongs to lane group j arrives either after the last phase that could serve its lane group in the previous cycle $T-1$ or before the beginning of those phases in the current cycle T, its delay is the same as an auto that arrives at the same time at the lane group's queue. The delay for a transit vehicle that belongs to a lane group j can be calculated by queueing diagrams (Figure 2) and it is:

$$d'_{b,T} = \frac{q_j}{s_j}\left(t_b - (T-1)C + \sum_{i=l+1}^{N} G_{i,T-1}\right) + (T-1)C + \sum_{i=1}^{k-1} G_{i,T} - t_b \quad (6)$$

$$\text{for} \quad \sum_{i=1}^{l} G_{i,T-1} < t_b \leq \sum_{i=1}^{l} G_{i,T}$$

   If the transit vehicle arrives after the clearance of its lane group's queue and at a time within the phases that can serve it, the equation will give a negative delay, which implies that the delay for such a transit vehicle will actually be 0. So,

$$d_{b,T} = \max\{d'_{b,T}, 0\} \quad \text{for} \quad \sum_{i=1}^{l} G_{i,T-1} < t_b \leq \sum_{i=1}^{l} G_{i,T} \tag{7}$$

Example: If a bus in lane group 4 arrives during phases 6 of cycle $T-1$ or phases 1, 2, or 3 of cycle T (e.g., $t_b = t_1$ in Figure 2), it will be served during cycle T, and its delay is indicated on the queueing diagram as $d_{1,T}$ (Figure 2).

2. If the transit vehicle arrives during cycle T after the last phase that can serve its respective lane group, two things could happen: 1) either the phase that can serve it will be extended so that it can serve the transit vehicle during the current cycle T, or 2) the transit vehicle will be served during the next cycle $T+1$. For optimization purposes, the expected delay that such a transit vehicle would experience if the green time of the phase that can serve it is not extended is included in the calculation of the delay for cycle T. In order to estimate the expected delay of that transit vehicle, the green times for the phases of the next cycle $T+1$, $G_{i,T+1}$ are assumed to be the same as the minimum green times for each phase $G_{i\min}$, as explained before. The expected delay of such a transit vehicle is thus given by:

$$E[d_{b,T}] = \frac{q_j}{s_j}\left(t_b - (T-1)C - \sum_{i=1}^{l} G_{i,T}\right) + TC + \sum_{i=1}^{k-1} G_{i\min,T+1} - t_b \tag{8}$$

$$\text{for} \quad t_b > \sum_{i=1}^{l} G_{i,T}$$

If the transit vehicle ends up being served during the current cycle T, its delay will be zero.

Example: If a bus in lane group 4 arrives during phase 6 in cycle T (e.g., $t_b = t_2$ in Figure 2), it can either be served by phase 5 of the current cycle T, if it is possible to extend the green by a sufficient amount to serve the bus, or it will be served in the next cycle $T+1$ during phase 4. For the optimization process of cycle T the expected delay the bus would experience if it was to be served during the next cycle $T+1$ is taken into account.

Constraints

The first constraint refers to the minimum green times for each lane group. Minimum green times $G_{j\min}$ are necessary to ensure undersaturated conditions for each lane group, i.e., $G_{j\min} \geq \frac{q_j}{s_j}C$, where q_j is the flow for lane group j, and s_j is the saturation flow for lane group j. Since each phase does not always coincide with serving one lane group and a lane group can be served by more than one phase, there are many different combinations of minimum green times for the phases that can satisfy the minimum green time requirement for each lane group. As a result, constraining the green times for each lane group is the strictest possible way to ensure that all the queues will clear, allowing more flexibility in the allocation of green time to the different phases. This makes it possible to reduce delays more than if minimum green times for the phases were imposed. The second constraint ensures that the green time for each phase, which will be the outcome of the optimization, add up to the cycle length, C, which is kept constant for every cycle within a specific time period.

TEST SITE

The traffic responsive signal control system is being tested through simulation at the intersection of Katechaki and Mesogeion Avenues, which is located in Athens, Greece. The intersection of Katechaki and Mesogeion Avenues was selected for two reasons:

1. the high auto demand on all approaches,
2. the existence of conflicting bus routes.

The intersection's layout is presented in Figure 3. As the Figure shows, the main through movement for Mesogeion Avenue, passes underneath the intersection for both directions. Auto volumes are available from loop detectors placed 40 m upstream for each approach at a rate of once per second.

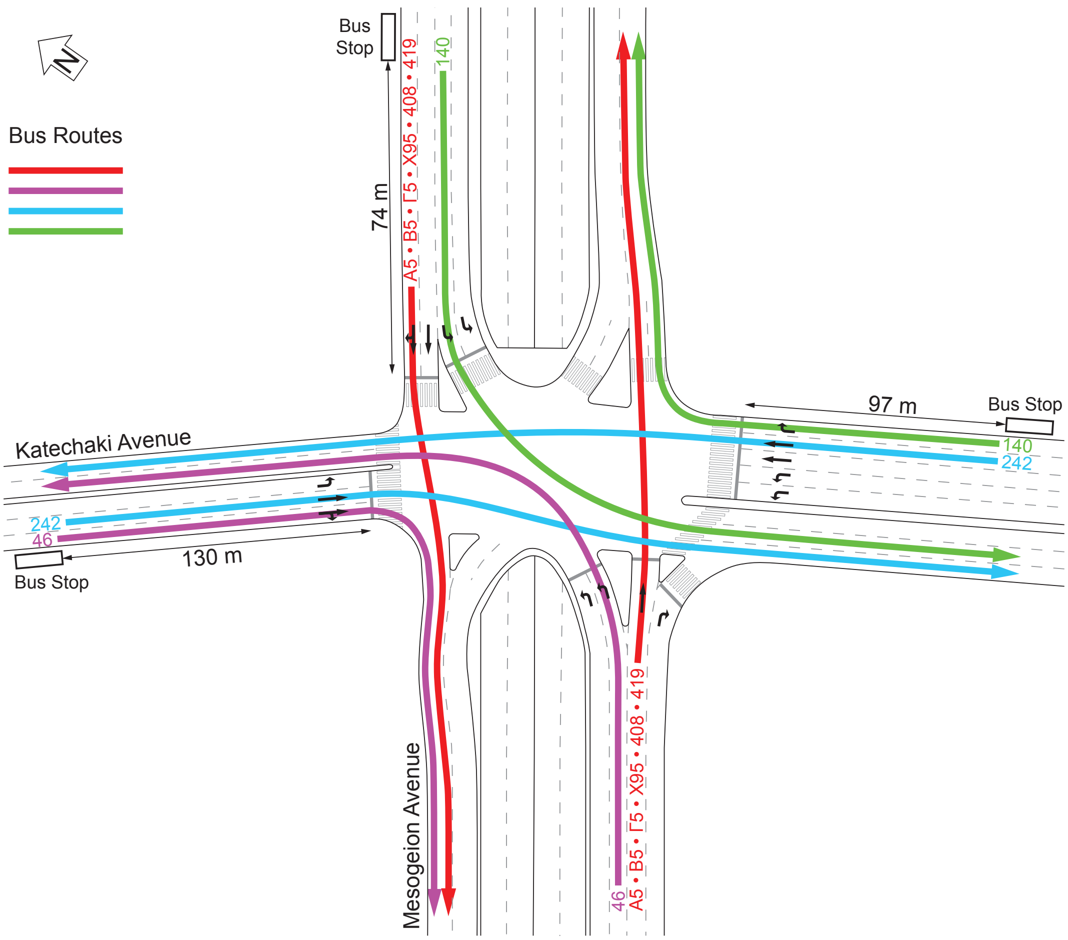

Figure 3 – Intersection Configuration and Bus Routes

Nine bus routes travel through the intersection in mixed traffic lanes with headways that vary from 8 to 20 minutes during peak hours (Figure 3). The numbers next to the directional arrows in Figure 3 correspond to the different bus routes. The bus routes travel in four conflicting directions and their bus stops are located nearside (i.e., upstream of the intersection). The southwest approach bus stop is not shown here because of its longer distance from the stop line, which diminishes its impact on the traffic operations of the intersection. However, the impact of the bus stops on the operations of the intersection are ignored for the moment. Information about the bus schedule is available at the Athens Urban Transport Organization website (http://www.oasa.gr).

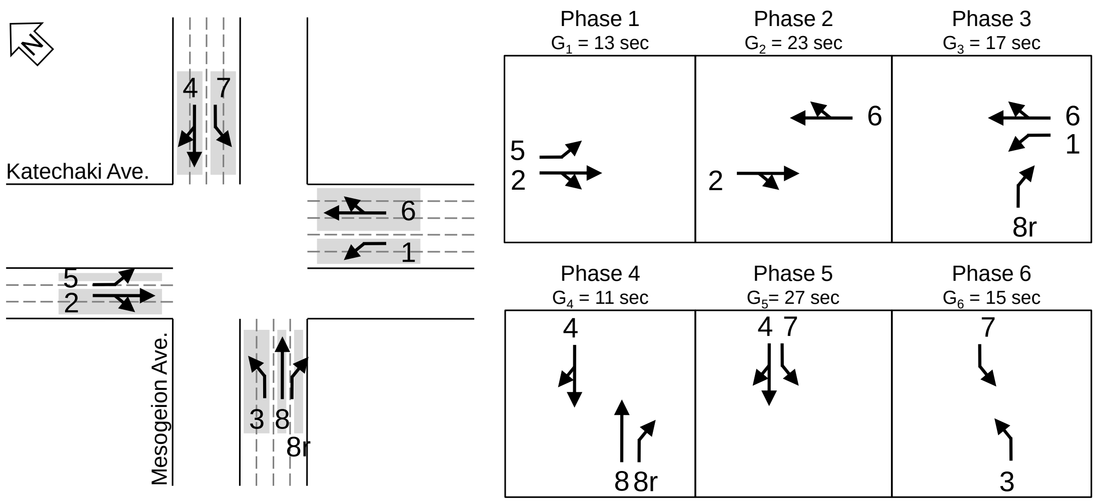

Figure 4 – Lane Groups and Phases for Katechaki-Mesogeion Intersection

Traffic volumes during the morning peak hour (7–8 am) are used as a representative peak volume. The intersection signal is operated on a fixed 6-phase cycle with a cycle length of 120 seconds during that time period. Figure 4 presents the phase design for the intersection during peak hours (on the right labeled 1–6) and the lane groups (on the left labeled 1–8r).

Since no real-time information was available, the actual bus arrival times at the intersection were estimated based on a shifted normal distribution around their scheduled arrival time. The average auto occupancy $\bar{o}_a$ is assumed to be 1.25 which is a reasonable value for autos at this specific intersection during the morning peak hour (7–8 am) used for this analysis. Regarding the transit vehicles, the passenger arrivals at the bus stops are assumed to be deterministic. As a result, the bus occupancy is a function of the time between the actual arrivals of two consecutive buses of the same route (i.e., the buses are assumed to operate as if they arrive empty at the bus stop just upstream of the intersection under consideration). The occupancy of each bus that arrives at the intersection is given by:

$$o_b = \phi_r(t_{b,r} - t_{b-1,r}) \tag{9}$$

where:
ϕ_r: demand for bus route r $\left[\frac{\text{pax}}{\text{hr}}\right]$
$t_{b,r}$: actual arrival time of bus b of route r
$t_{b-1,r}$: actual arrival time of bus $b-1$ of route r

In that way, despite the fact that the schedule delay of the buses was not considered initially it is implicitly included through the higher occupancy that late buses are expected to have. For the initial testing of the signal control optimization, an average bus occupancy of 40 passengers per vehicle was assumed.

RESULTS

Several scenarios with different auto vehicle demands and bus occupancies were evaluated through a one hour simulation. For each scenario a warm up period equal to one cycle length was used. In addition, each scenario was evaluated ten times in order to account for the stochastic variation in bus arrivals at the intersection. A total of 300 test cases were performed.

Real data from the study site intersection described in the Test Site section were used to test the outcomes of two optimization scenarios: 1) when only vehicle delay is minimized and 2) when total person delay for both bus and auto passengers is minimized (i.e., signal control optimization with TSP). The corresponding person delays for the optimized signal timings for the total number of passengers traveling through the intersection as well as the respective ones for the auto and bus passengers for the morning peak hour (7–8 am) are shown in Table I. The Table also contains a comparison of the person delays for the two optimized scenarios.

Table I – Person Delays for the Morning Peak Hour Auto Demand and Average Bus Occupancy of 40 pax/veh

	Auto Delay (pax-hrs)	Bus Delay (pax-hrs)	Total Delay (pax-hrs)
Base Case Scenario (Existing Signal Settings)	63.78	33.27	97.06
Scenario 1: Vehicle-based Optimization	59.04	41.53	100.56
Scenario 2: Person-based Optimization	62.56	25.63	88.19
% Improvement over Scenario 1	–5.97%	38.28%	12.30%

A comparison of the person delays obtained from our optimization process (scenario 2), with the corresponding ones obtained from the minimization of vehicle delays (scenario 1) indicate the significance of the improvement achieved. Total person delay for the intersection was reduced by 12%, while bus passenger delay was reduced by 38%, indicating the improvement for transit operations achieved by the use of conditional priority. The 6% increase in the auto passenger delay is expected as a result of their lower occupancy which makes them being weighted much less during the optimization process. The percentages translate to an increase in the auto delay in the order of 2 seconds per vehicle on average and a decrease in the bus delay in the order of 16 seconds per bus on average. In general, the impacts depend on the auto and transit demand (i.e., vehicle flows) as well as on the transit characteristics such as transit occupancies, headways and the number of routes traveling through the intersection.

Effect of Auto Demand

In order to capture the effect of auto vehicle demand on the results of the optimization process, five scenarios were tested. These scenarios were defined by scaling up and down the initial auto volumes for the morning peak hour (7–8 am), while keeping the average auto occupancy to

1.25 passengers per vehicle. Also, for each auto demand scenario, we tested different average bus occupancies (20, 30, and 40 passenger per vehicle), a total of 15 cases. While changing the auto volumes, adjustments in the minimum green times and cycle lengths were required for undersaturated conditions to be maintained. The adjustment in the cycle length was such that the volume to capacity ratio remains constant for all the lane groups. So, the impact that auto demand has on the results could be isolated from the impact changes in the cycle length could have.

Figure 5 illustrates the changes in the person delay of auto and bus passengers, as well as the total person delay, achieved by the person-based optimization with respect to the vehicle-based optimization results for the different auto demand scenarios and for an average bus occupancy of 40 passengers per vehicle. The results of Figure 5 indicate consistent patterns in the person delay changes for all the scenarios. More specifically, the higher the auto demand the lower the benefits we could achieve in both the total and the bus passenger delay, as expected. Higher auto demands increase the delays for all the vehicles traveling through the intersection and constrain the provision of priority to the buses. As a result, the reductions in the delays achieved compared to the vehicle-optimized signal settings are diminishing with higher auto volumes. For very high auto volumes, i.e., 1.25 and 1.5 times the initial auto volumes, the two ways of optimizing the signal settings result in the same person delays, the high auto flow outweighs the higher occupancies of the buses.

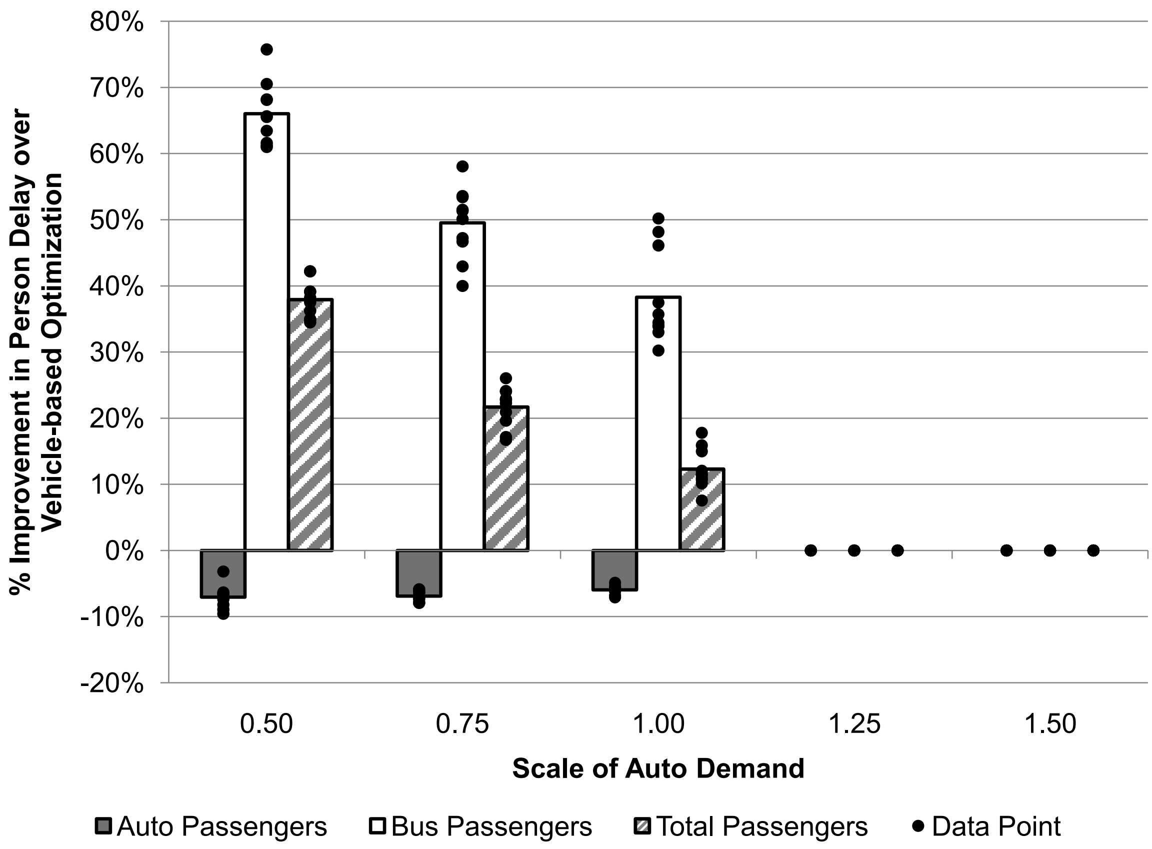

Figure 5 – Improvement in Person Delays for Different Auto Demand Scenarios

Figure 5 also presents the results for each of the ten simulation runs, used to account for the stochasticity in bus arrivals for each auto demand level (shown as Data Points in the Figure). The results indicate that the higher the auto demand, the higher the variation in the improvement of the delays for the bus passengers users and the intersection travelers as a whole.

Effect of Average Bus Occupancy

The effect of buses' average occupancy was captured by simulating scenarios with average bus occupancies that varied between 20, 30 and 40 passengers per vehicle. For each average bus occupancy level, we tested different auto demand scenarios, by scaling up and down the initial auto volumes for the morning peak hour (7–8 am). The average auto occupancy was kept constant for all the scenarios and equal to 1.25 passengers per vehicle. Figure 6 illustrates the changes in the person delay of auto and bus passengers, as well as the total person delay, achieved by the person-based optimization with respect to the vehicle-based optimization results during a one hour simulation period for the different average bus occupancy scenarios and for the initial auto volumes of the morning peak hour. The results indicate that the higher the occupancy of the buses, the higher the savings for their passengers and the higher the delays for the auto users, always compared with the vehicle-based optimization person delays.

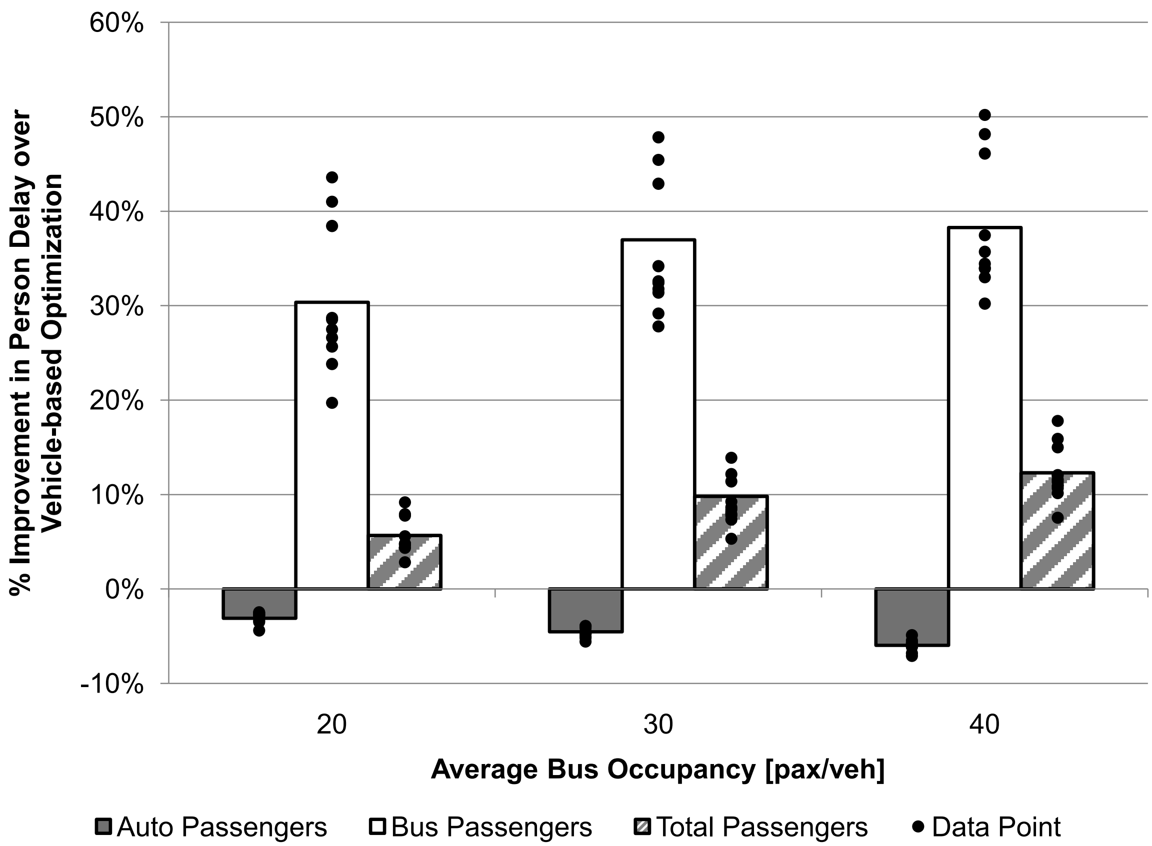

Figure 6 – Improvement in Person Delays for Different Average Bus Occupancy Scenarios

Figure 6 also illustrates the variation in the results of the ten different simulation runs, used to account for the stochasticity in bus arrivals, for each average bus occupancy scenario (shown as Data Points in the Figure). The results indicate that the bus passengers experience higher variation in their delay savings compared to the variations in the delay changes for the auto users and the total number of passengers traveling through the intersection. Testing the different average bus occupancy scenarios for the five auto demand scenarios, similar patterns were observed.

CONCLUSIONS

A real-time, traffic responsive signal control system with TSP has been developed and tested at an isolated intersection. The optimization method used explicitly accounts for the occupancy of autos and transit vehicles to assign priority in an equitable way, even for transit vehicles that travel in conflicting directions.

The results from the application of the optimization method on a real-world intersection shows the effectiveness of the proposed traffic responsive TSP system in reducing the overall person delay as well as the bus passengers' delay by providing priority to buses traveling in conflicting directions. The total person delay of all passengers was reduced by 12% and the delay of bus passengers was reduced by 38% compared to the vehicle-based optimization results for the morning peak hour conditions. At the same time, the increase in the auto passengers' delay was on the order of only 6%.

The optimization was shown to be effective in reducing total passengers' delays at the isolated intersection for a wide range of auto demand and average bus occupancies. The tests showed that an increase in the auto demand lowers the benefits for both the total and the bus passengers. For very high auto volumes the developed optimization process is expected to reveal the same results as the vehicle-based optimization. Similar patterns were also observed for different average bus occupancy values. The tests also showed that the higher the bus occupancy, the higher the reduction in the person delay for the intersection and the bus passengers and the higher the increase for the auto users.

Next steps in the study include testing additional TSP strategies, e.g., phase rotation, for a wide range of traffic and design characteristics. Note that the test intersection is characterized by more complicated geometry and phasing (Figure 4) than at most intersections in the U.S. We therefore expect that the testing of the proposed method will yield similar or better results when implemented at simpler intersections. We are also planning to evaluate the proposed system based on several performance measures, e.g., transit schedule adherence, energy consumption, and emissions. Next, using the presented formulation for the isolated intersection as the stepping stone, we will extend the system to arterials and grid networks.

REFERENCES

Bretherton, D., Bowen, G., and Wood, K. (2002). Effective urban traffic management and control: SCOOT Version 4.4. In European Transport Conference.

Chang, E. and Ziliaskopoulos, A. (2003). Data challenges in development of a regional assignment: Simulation model to evaluate transit signal priority in chicago. Transportation Research Record: Journal of the Transportation Research Board, 1841:12–22.

Conrad, M., Dion, F., and Yagar, S. (1998). Real-time traffic signal optimization with transit priority: Recent advances in the signal priority procedure for optimization in real-time model. Transportation Research Record: Journal of the Transportation Research Board, 1634:100–109.

Cornwell, P., Luk, J., and Negus, B. (1986). Tram priority in SCATS Melbourne. Traffic Engineering and Control, 27:561.

Henry, J. and Farges, J. (1994). P.T. priority and Prodyn. In Proceedings of the 1st World Congress on Application of Transport Telematics and Intelligent Vehicle-Highway Systems, volume 6, pages 3086–3093.

Hunt, P., Bretherton, R., Robertson, D., and Royal, M. (1982). SCOOT on-line traffic signal optimisation technique. Traffic engineering and control, 23:190–2.

Li, M. (2008). Toward deployment of adaptive transit signal priority systems. Technical report, California Partners for Advanced Transit and Highways.

Li, Y., Koonce, P., Li, M., Zhou, K., Li, Y., Beaird, S., Zhang, W., Hegen, L., Hu, K., Skabardonis, A., et al. (2008). Transit signal priority research tools. Technical report, California Partners for Advanced transit and Highways.

Mauro, V. and Di Taranto, C. (1989). UTOPIA. In Proceedings of the 2nd IFAC-IFIP-IFORS Symposium on Traffic Control and Transportation Systems, pages 575–597.

Skabardonis, A. (2000). Control strategies for transit priority. Transportation Research Record: Journal of the Transportation Research Board, 1727:20–26.

Transportation Research Board (2000). Highway Capacity Manual, Special Report, National Research Council, Washington, DC.