Bus Priority with Highly Interruptible Traffic Signal Control

Simulation of San Juan's Avenida Ponce de Leon

Melanie Janos and Peter G. Furth

A traffic control strategy was designed explicitly to accommodate bus priority on Avenida Ponce de Leon between Old San Juan and the northern terminus of the metro, Tren Urbano, now under construction in Puerto Rico. Up to 36 buses operate per hour in mixed traffic northbound and in an exclusive contraflow lane southbound. The control strategy is characterized as "highly interruptible" because it responds quickly to, and recovers quickly from, priority interruptions, which are expected to occur frequently. It features full actuation and schedule-based conditional priority for buses. Buses may request to pass through the intersection or to advance to the bus stop that may be occupied by a standing queue. Priority tactics include green extension, early green, and early red. To preserve capacity and encourage dwell time to overlap with traffic delay, bus blockage at a stop forces an early and minimum length red. There is no background cycle; however, loose progression is provided by means of requests, of lower priority than bus requests, from upstream intersections to force or hold green until a platoon arrives. The control strategy was programmed and tested with VISSIM, whose useful features and weaknesses for this study are described. The results indicate significant transit travel time savings without slowing general traffic compared with the existing strategy of fixed time control with progression.

Construction is under way for the starter segment of the metro Tren Urbano in San Juan, Puerto Rico (*1*). However, the initial 17-km line does not extend all the way north to Old San Juan, the cultural center of the island and a major generator of transit trips. Until a contemplated extension is built, buses will be expected to provide high-quality transit service between Old San Juan and the northern terminus of Tren Urbano, a distance of 6.7 km. This paper reports on an effort to design and test by simulation a system of bus priority at traffic signals to improve service quality in this corridor.

SETTING

San Juan was originally one of only three walled cities in North America, located on the western tip of a small island (the Isletta). As the population grew and the walls were removed, the city grew to the eastern end of the Isletta and then across a narrow channel to the mainland. It continued to grow southward, first through an area called Santurce and then across a lagoon to an area called Hato Rey, which contains today's main business district; ultimately, it merged with Rio Piedras, a separate town that has since become part of the Municipality of San Juan. The growing city's spine was a trolley line from Old San Juan to Rio Piedras along Avenida Ponce de Leon, named after Puerto Rico's first colonial governor. From Rio Piedras northward, Tren Urbano follows that historic spine, but its initial segment ends at the Sagrado Corazon station at the Hato Rey–Santurce boundary.

The corridor of interest has two sections. The first section, 3.6 km long and presented in Figure 1, is through Santurce, a densely populated neighborhood with closely spaced, narrow streets, a busy commercial district, and several government offices and institutions. In this section, Avenida Ponce de Leon has three lanes. Two serve general northbound traffic; the other is an exclusive bus lane southbound. This section has 16 signalized and 29 unsignalized intersections. The signalized intersections operate under fixed time control with a 90-s cycle. Because of the one-way general traffic and low traffic volumes on most cross streets, all the intersections except one (Todd) operate with two phases. This section suffers from mild traffic congestion, which is exacerbated at times by double parking and poor traffic control in a one-block section used by cross traffic. There are 16 bus stops in this section, including Sagrado Corazon station.

The second section, 3.1 km long, is on the Isletta, ending at the Covadonga terminal on the edge of Old San Juan. Much of this section, with parks and institutions, is unpopulated. Buses mostly use a one-way pair of arterials with exclusive contraflow lanes. This section has eight bus stops and only three traffic signals that pose little delay to bus service. Although this study models the entire corridor, it pays primary attention to the Santurce section, leaving control on the Isletta unchanged. The morning peak hour, 7:15 to 8:14 a.m., was studied, when traffic on Avenida Ponce de Leon is greatest and the primary direction of bus passenger travel is northbound, the direction in which buses operate in mixed traffic.

Traffic counts taken at signalized intersections in 1996 and 1997 were extrapolated to 2003, assuming 3% annual growth. Discontinuities between neighboring intersections were treated as sources and sinks when the land use was appropriate (e.g., a sink at a large employment site, a small source at a side street) and smoothed otherwise. Peak-hour volumes range from about 1,750 vehicles per hour (vph) at the southern end and again 1.2 km north to 400 vph at the northern end, with volume-to-capacity ratio ranging from 0.77 to 0.32. Only two cross streets had volumes above 400 vph.

Bus routes, schedules, and volumes predicted for 2003, after Tren Urbano begins operation, were taken from a study done by Multisystems (*2*). Route 1, running with a 5-min headway the full length of the corridor, is the main subject of the study. Peak passenger volume northbound, 1,036 per hour, results in a rather high average peak load of 86 as buses leave Sagrado Corazon station. (An increase in Route 1's frequency of use is contemplated if demand

Department of Civil and Environmental Engineering, 400 SN. 360 Huntington Avenue, Northeastern University, Boston, MA 02115.

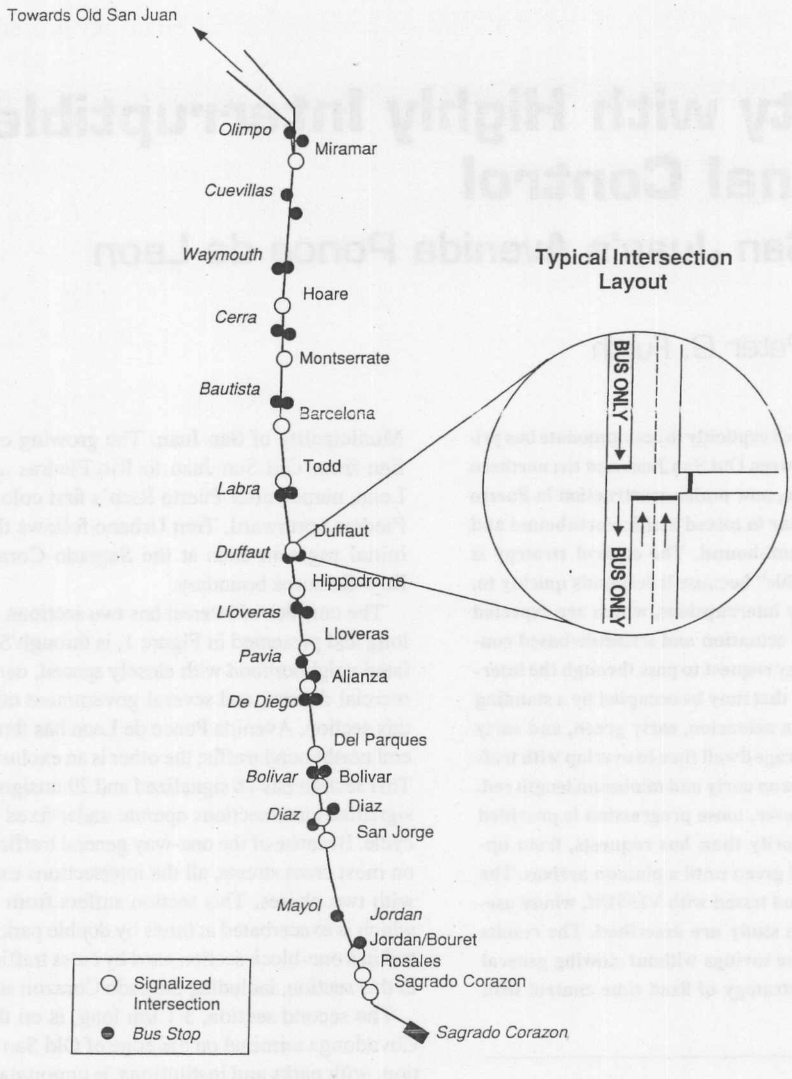

FIGURE 1 Corridor layout.

projections hold or are exceeded.) Peak load southbound, also at Sagrado Corazon station, is 20. Besides Route 1, several other bus routes operate along parts of Avenida Ponce de Leon. Including bus volume on Route 1. bus volumes in the Santurce section range from 36 to 21 per hour.

TRAFFIC SIGNAL CONTROL BASED ON TRANSIT PRIORITY

Priority at traffic signals is a powerful way to reduce transit delay and improve regularity. Although applied extensively in Northern Europe, such development in the United States is not yet mature. Applications and studies in the United States include Balke et al. (*3*), Hunter-Zaworski et al. (*4*), and Luh (*5*). U.S. applications typically accept an underlying signal control strategy with fixed cycle length and offsets selected to provide progression, allowing small deviations to favor detected buses.

Instead of trying to fit priority into an existing traffic control scheme, the approach was to develop a traffic control strategy from the ground up whose objective is first to provide high-quality bus service. From the start, any attempt to maintain a common cycle and progression offsets is abandoned. Unless active priority is limited to "borrowing" a few seconds from the side street, its essence is to interfere with progression. As Furth and Muller indicate (*6*), trying to return an intersection to the background cycle after a priority interruption can take a considerable length of time and can be even more disruptive than the original priority interruption. They found that, when the frequency of priority interruptions increased from four to eight per hour at an intersection programmed to freely grant priority but then returned to a background cycle, large overflow queues developed because the intersection could not recover from one priority interruption before the next occurred. For this reason, some signal systems in the United States will not accept a priority request for 15 min or so after a priority interruption. Imposing this kind of restriction on Avenida Ponce de Leon, where a Route 1 bus passes through each intersection every 2 min (counting both directions) would render priority meaningless.

To accommodate transit priority, a traffic signal control scheme must be highly interruptible—able to quickly grant priority and to quickly dissipate queues resulting from an interruption. This kind of background operation permits priority interruptions as needed, without restrictions on their frequency. To an extent, the control strategy developed in this study may be unique because of the unusual arrangement of Avenida Ponce de Leon, a one-way arterial with buses operating in mixed traffic in one direction and in an exclusive contraflow lane in the other direction. However, the authors believe the principle can be more widely applied.

In the next section, the traffic control strategy and its basis are described. Then, the simulation techniques used are described. Finally, the simulation results are presented and analyzed, which provides a strong proof of concept for the priority strategy developed.

TRAFFIC SIGNAL AND OPERATIONAL CONTROL STRATEGIES

As Muller emphasizes (*7*), traffic signal control is a way to accomplish societal objectives. The aim of this study was to develop traffic signal control logic appropriate to the physical setting, traffic and transit service characteristics, and road use priorities of the Avenida Ponce de Leon corridor. First, the overall objectives and the control principles that follow from those objectives are presented. Next, more detail about the priority and actuation tactics used is given. At the conclusion of this section, the control logic is summarized in a flowchart.

Societal Objectives and Control Principles

High-Quality Bus Service: Active, Conditional (Schedule-Based) Bus Priority

Renewing and reorganizing the San Juan metropolitan area around high-quality transit service is one of Puerto Rico's important societal objectives (*1*), which led to the investment of hundreds of millions of dollars in Tren Urbano. To accomplish that objective, Tren Urbano must be complemented by high-quality, attractive bus services. Nowhere is this need more critical than in the historic corridor linking Tren Urbano to Santurce and Old San Juan. Because bus service in this corridor needs to coordinate well with Tren Urbano, the most critical aspect of service quality is reliability, which becomes the first objective of the signal control system. Reliability can be interpreted as either maintaining the scheduled headways or maintaining the scheduled departure times at time points all along the route. Headway maintenance is important, as it is in any short headway service, to prevent overcrowding, long waits, and the resulting deterioration in service rates for boarding and alighting movements. Schedule adherence is important, especially in off-peak periods when headways are longer and making the bus–train transfer becomes more important. Even in peak periods, when headways are short, bus–train coordination is important because each Route 1 bus trip will be scheduled to carry the continuing load of one train. Because schedule maintenance implies headway maintenance, but not the reverse, this control program emphasizes schedule adherence.

Traffic signal control can contribute to schedule adherence in two ways. First, reducing intersection delays generally reduces the variance in intersection delays as well (*8*). The chief way traffic signal control can reduce intersection delays for transit is by giving active (i.e., detection-based) priority. Therefore, the first principle of the control system is to give active priority to transit. Specific priority tactics are discussed in the next section.

The second way traffic signal control can contribute to schedule reliability is by conditioning transit priority on schedule deviation, thereby providing a means of operational control (*6, 9*). By extending priority to late buses and withholding priority from early buses, a signal control program can provide the push–pull means of operational control necessary to compensate for inevitable disturbances. Therefore, the second principle of the control system is to give priority only to late buses, a strategy called conditional priority.

It should be noted that, for operational control, "late" should be measured in seconds, not minutes. In at least one recent U.S. application, transit priority is granted only if the bus is at least 5 min late. Although this definition of late is consistent with U.S. transit industry practice, it is useless for operational control. In such a system, either buses will rarely be late and therefore rarely get priority, or they will regularly fall 5 min behind schedule and then the priority system will keep them there. If the former is true, why bother to invest in a priority system? If the latter is true, why not shift all the departure times by 5 min so the buses run on time? For operational control, a bus should be given priority if its lateness is greater than the expected correction priority will provide. In this system, priority is extended if the bus is more than 60 s behind schedule.

A second aspect of service quality is speed, and the signal control system can help speed transit service by reducing signal delay. To an extent, this objective is parallel to the objective of improving service reliability. However, if all signal delay is eliminated by giving priority to every bus, traffic signal control loses its ability to exercise operational control and reliability will suffer. Absolute priority is not in transit's best interest. Buses that are early should be delayed, and it is easier, from the viewpoint of both the vehicle operator and the customer, for that delay to be imposed by a traffic signal than by other means. Furthermore, if a schedule is written that expects no delay at traffic signals, then priority cannot be used to push ahead a vehicle that has fallen behind schedule. Ideally, the bus schedule should be written with some traffic signal delay built in, and the control program should be able to reduce or increase that delay as needed to keep the buses on schedule.

Consequently, the objective of providing fast transit service yields to the objective of providing reliable transit service once the level of scheduled traffic delay has been reduced to a critical amount that should be retained for operational control. This strategy requires a detailed and carefully written transit schedule that includes scheduled running times for every section and scheduled departure times at every signalized intersection.

Preserve Capacity: Actuated Control, with Forced Red During Bus Blockage

Mobility, another important societal objective, leads to another objective of the traffic control system: to maintain traffic capacity and prevent queue overflows. On the main street, with buses running northbound in mixed traffic, queue overflows lead to large delays for transit as well as motorists. Queue overflows on side streets cause large delays for affected motorists and for the buses operating on routes that use some of those side streets. Priority interruptions can lead to unexpected queues as green times are extended, cut short, or skipped. A highly interruptible traffic control system must be able to recognize queues and give priority to dissipating them.

The concern here is not theoretical capacity as might be calculated by following a capacity manual: those capacity formulas assume an infinite supply of traffic flow and are maximized by long cycles that diminish the impact of lost time during phase changes. Instead, the concern here is throughput, recognizing real-time variations in traffic supply due to variations in demand, variations in upstream signal cycles, and frequent blockage by buses. Throughput is maximized by switching the phase from a traffic stream that has no flow or low flow to a traffic stream with a standing queue that is ready to go at saturation flow.

Therefore, the third principle of the control system is full actuation. On both the main street and the side streets, detectors should measure flow (or its inverse, gaps) while the signal is green and detect queues while the signal is red. And just as detectors actuate the start of a priority request, exit detectors should cancel priority requests to prevent unnecessary prolonging of a priority interruption.

On Avenida Ponce de Leon, there are no bus bays; buses stop in the travel lane, blocking the lane. One way to prevent a resulting capacity loss, with both near-side stops and far-side stops (although the effect is stronger with near-side stops), is to turn the signal red for the street with the bus (the "bus street") during the blockage. If the stop is near side and the cycle can be short enough so that green returns to the bus street by the time the bus is ready to advance, this strategy will be doubly beneficial. All too often, a bus stops while the signal is green, blocking the lane and inhibiting traffic, and when its dwell time is over and the bus is ready to advance, the signal turns red.

Therefore, the fourth principle of the control system is to force an early red for the bus street when a bus serving passengers is blocking a lane. We recognize that this strategy will impose some delay on buses, but the expected impact in the Avenida Ponce de Leon context is small. Because of two-phase control and narrow streets, a forced red on the main street can usually return to early green in 24 s (5 s each for two clearance intervals, plus 14 s minimum green for the side street). Allowing that 8 s will be used to open and close the doors and to begin advancing toward the stop line, no transit delay is expected if four or more passengers board or alight, which is usually the case during busy periods in Santurce. Even if only one passenger boards or alights, transit delay is limited to 12 s, a reasonable price for the improved traffic flow in which the buses operate.

Serve Pedestrian Crossings: Minimum Green and Maximum Cycle

Serving pedestrians well is another important societal objective in this corridor. Pedestrian safety requires adequate crossing time and precludes the idea of short, nonpedestrian phases, leading to the fifth principle of the control system: enforce minimum green times. Eleven seconds was used for the main street and 14 s was used for the side streets, based on roadway width. Good service for pedestrians also calls for short waits, meaning short red times, which follow from short cycles. The goal of having short cycles is already served by the principle of actuation, which nearly minimizes cycle length. However, to prevent occasional long waits, a sixth principle is needed: enforce maximum green times. Maximum green times were used that would keep the cycle time from exceeding 120 s.

Expediting Private Traffic: Requests from Upstream Intersections

Reducing motorist travel time is not a primary objective of the control system. For private traffic, Avenida Ponce de Leon's intended uses are local access and minor arterial. Most private cars on the road use only a short segment at the origin or destination end of their trip. For this intended use, speed is not important. A parallel freeway serves longer distance traffic. In contrast, for transit users, Avenida Ponce de Leon is a major arterial. Buses and many passengers travel the full length of the road. To serve its intended use, high travel speed for transit is important; high travel speed for private traffic is not. Of course, inducing a large increase in motorist travel time would certainly be politically problematic. However, because one of the control principles is to protect capacity, travel time should not exhibit the feared exponential increase that occurs when demand exceeds capacity.

To make reducing motorist delay a secondary objective that yields to a primary objective of serving transit well, a seventh control principle was adopted, which is to give an expected platoon priority over side street traffic. The expected arrival of a platoon, signaled from the neighboring upstream intersection when the side street red ends, is treated as a request to turn or hold the main street green. That request takes priority over a side street request but not over a bus priority request. After the expected arrival of the platoon, the request is canceled, and normal actuation will keep the signal green until the rear of the platoon has passed. In the absence of bus-induced interruptions, this strategy tends to make intersections cycle together with offsets that provide an imperfect green wave. One of the questions this study addressed is to what extent this strategy, in the face of interruptions for bus blockage and bus priority, would result in common cycles.

Summary of Control Principles

In conclusion, societal objectives lead to seven traffic control principles. Rearranged in order of precedence, they are as follows:

1. Minimum green time;
2. Active priority for buses;
3. Priority only for late buses, with a tight standard for lateness;
4. Forced red during bus blockage;
5. Maximum cycle time;
6. Priority for platoons arriving from an upstream intersection; and
7. Actuated requests and cancelation of all requests.

In addition, conditional (lateness-based) priority is also a transit operational control strategy, requiring a detailed and carefully calibrated schedule.

Priority Tactics

In this section, two types of priority requests and tactics used to respond to those requests are described. Buses can generate two kinds of priority requests. The first is a request to pass through the intersection, generated when there is no near-side stop or when a bus serving a near-side stop closes its doors. The second is a request to advance to the bus stop, generated when there is a near-side stop that may be blocked by a queue. The latter type of request does not apply in the southbound direction in which buses operate in an exclusive contraflow lane.

These two types of priority requests must be matched to the available priority tactics. Simple two-phase control (main street, side street) is appropriate for Avenida Ponce de Leon because of its one-way general traffic flow. Also, because of high pedestrian traffic, it would be unsafe to insert phases shorter than the pedestrian minimum. Therefore, the available control tactics are limited to green extension, early green (forcing the side street to red), and early red. The latter tactic involves forcing the main street green to red, then applying a minimum green to the side street, and then returning green to the main street.

Green extension is applied if the main street is green when the priority request is received and the expected total green time is less than 1.2 times the normal maximum green. Expected green is calculated based on the distance from the detector to either the stop line (for a request to pass) or the bus stop (for a request to advance to stop). The priority request is canceled when the bus passes a detector at the bus stop or stop line, as appropriate to the request type. Time-outs (after 10 s for a request to advance to the stop line; after 15 s for a request to pass the intersection) automatically cancel a request, which protects against detector malfunction and unexplained delay.

Early green is applied if the main street is red when the priority request is received. The side street green is forced off early, subject to a minimum green based on pedestrian crossing needs.

Early red followed by minimum side street green is applied in two cases. In one case, a priority request is received when the main street is green, but a green extension cannot be granted because the expected total green time would be too long. In the second case, a late bus arrives at a near-side stop. (If an early bus arrives at a near-side stop while the main street is green, there will be an early red as a capacity-protection measure but without forcing the side street to minimum green.) To protect against large queues forming on the side streets, the side street cannot be hurried by an early red request in successive cycles.

It was also necessary to consider conflicting priority requests when a bus operating in one direction requests continued green while a bus in the other direction requests early red. In such a conflict, a green extension request takes priority until it is canceled by the bus passing the exit detector at the stopline.

Volume–Density Control

Actuated end of green involves detecting a low flow rate of vehicle density. Instead of simple gap detection, which is problematic on multilane approaches, a running average of flow over the last 5 s is kept, based on a count of vehicles passing a detector 80 m upstream of the stopline. Dividing by speed (also measured by detectors) yields density, which is proportional to flow when speed reaches the practical speed limit, and favors queues when speeds are low. Following the strategy known as volume–density control, the threshold increases gradually to reflect the growing interest of waiting traffic and pedestrians on the cross street. By trial and error, an initial threshold of 80 vehicles per km was chosen, increasing by 1 each second after the minimum green had expired and any expected platoon from the upstream intersection had arrived.

To protect an approach with a queue from losing its green, even if there is little arrival flow 80 m upstream, there were presence detectors 30 and 55 m upstream of the stopline. If either of them is occupied by a vehicle traveling at 3 km/h or slower, that will hold the light green for the next second.

Control Logic Flowchart

A signal control flowchart summarizing the control logic is presented in Figure 2. The control program runs through the flowchart once at the end of each time step, beginning at "major green" or "minor green" depending on which street shows green. The logic leads back to the original state, indicating that the green will continue for another time step, or to the command to end green and start the appropriate interstage (change interval, consisting of yellow and all-red of specified duration). During change intervals the logic sequence is ignored.

The full control program for each intersection takes about 500 lines in the simulation software's controller language.

SIMULATION TECHNIQUES

One obstacle to the research and design of bus priority systems is that available traffic simulation packages do not offer all the features the analyst needs to model the roadway, the traffic, or the traffic signal control. Sometimes analysts have had to take cumbersome measures to model transit priority with available traffic simulators (*5*). This section describes key features that were found useful or lacking in the simulation package used and some of the techniques that were used to get around the model's weaknesses.

VISSIM, a microscopic, time step and behavior-based simulation model of urban traffic and public transit operations developed in Germany was chosen (*10*). It has been applied in several U.S. projects, and studies have shown that it is generally consistent with the U.S.-developed package CORSIM (*11*). Like CORSIM, it simulates, evaluates, and animates traffic operations under different scenarios of lane configuration, detector placement, traffic composition, traffic signals, and transit stops. However, it has several features that make it a useful tool for modeling transit priority.

One useful feature is modeling passenger movements. VISSIM tracks the load on each bus, generates passenger arrivals at stops following a Poisson distribution, and determines passenger alightings based on a fixed, user-supplied proportion of the load for each stop. With this feature, the effect of variable headway on dwell time and the resulting fundamental instability of close headway service is captured.

Another useful feature is bus detection and communication. In addition to standard traffic detectors, VISSIM features transit detectors that will work in both mixed traffic and in an exclusive lane or median and that will send a "telegram" to the traffic control system indicating the line, branch, direction, train length, load, priority, and delay. Thus, it was possible to give priority to Route 1 buses but not to buses operating on other routes.

A third important feature of VISSIM is that its traffic signal controller (called a "signal state generator") is user programmable and operates separately from the traffic simulator, with a limited interface. At the end of each time step, the simulator sends information to the controller detector. Controllers can also send codes to neighboring controllers. The controller then runs through its logic, changes signals if appropriate, and returns the signal state to the simulator for the next time step. Menus allow for simple input of standard control methods such as fixed time control, but any logic can also be programmed, provided the inputs can be obtained from the detectors. The controller language allows the user to define variables, do calculations, and control logical flow with comparisons. It supports subroutines and user-supplied data files.

To determine whether a bus is late, a file was supplied with the scheduled departure time for each Route 1 bus trip at each intersection, by direction. On receipt of a telegram from a bus crossing a transit detector, the controller compares the time with the scheduled departure time and determines whether the bus is late, using a user-

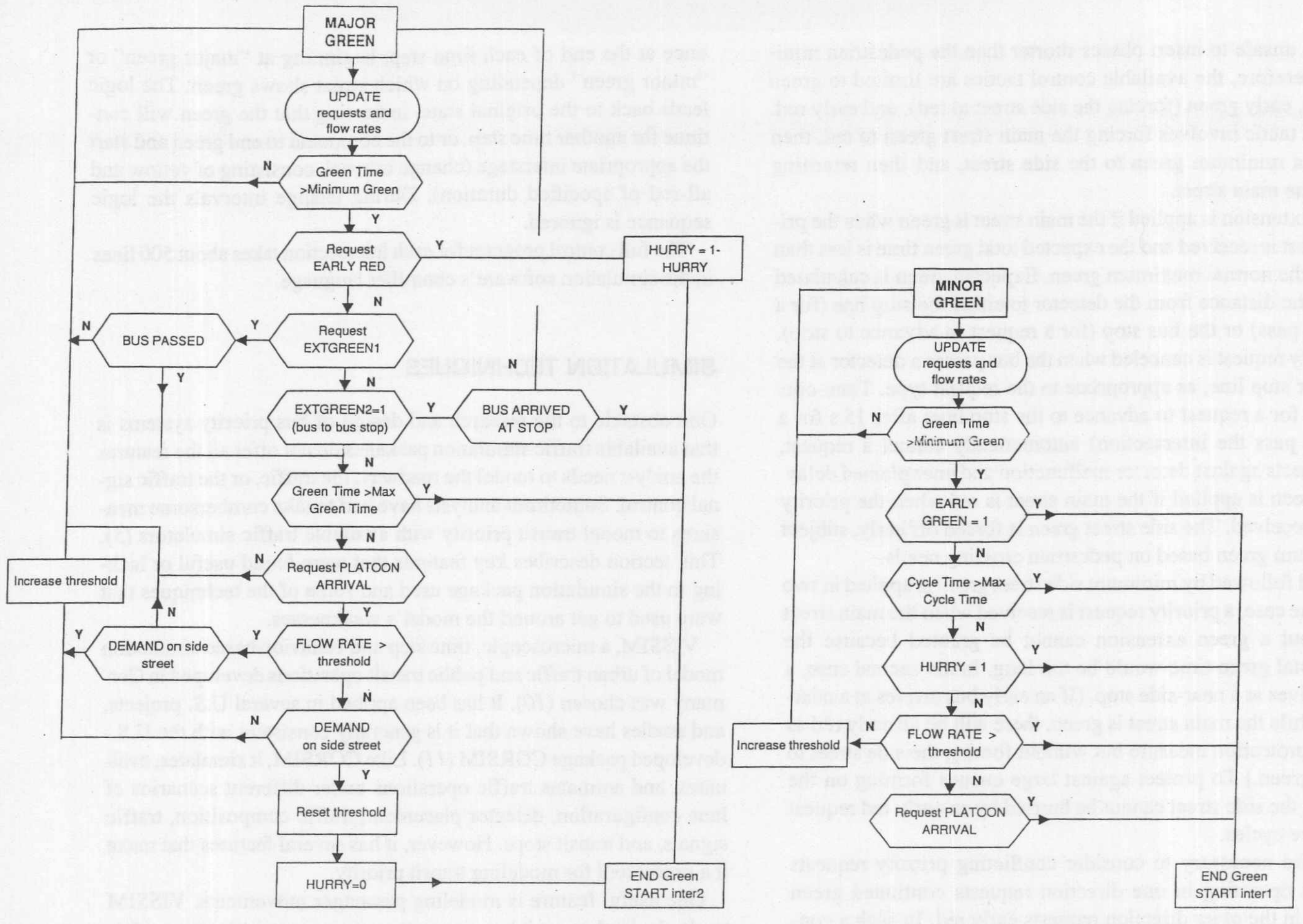

FIGURE 2 Traffic control logic.

supplied lateness tolerance, set to 60 s for conditional priority and to a large negative number for absolute priority.

VISSIM's transit detection capability does not include sending a message when the door closes, or, equivalently, when dwell time has expired. To model the end of dwell time so that buses serving a near-side stop can request priority after their stop has been served, a detector was placed just downstream of the bus stop. When the bus reaches that detector, a telegram is sent to the controller, whose logic interprets it as a transit request.

VISSIM's dwell time functions are limited and not user programmable. As mentioned, VISSIM does track passengers accumulating at stops (generated by a Poisson process) and load on buses; a stop-specific factor applied to load determines alightings (there is no randomness there). Dwell time can be a deterministic function of passenger boardings and alightings or a random variable drawn from a user-supplied distribution that is independent of passenger boardings and alightings. For bus service, the former is more appropriate. The deterministic functions available are limited to a choice of one-door or two-door operation, with fixed unit times per boarding and per alighting. In the case of two-door operations, which were used here, one door is for boarding and the other is for alighting; the door that needs the most time determines dwell time. In all the tested scenarios except one, unit boarding and alighting times of 4 s were used, based on a small sample collected on the route now operating in the corridor. In one test case, unit boarding and alighting time was made function of the crowding on the bus. Although VISSIM cannot model such a relation straightforwardly, a way was found to implement it, which is described in the results section as part of the discussion of that scenario.

RESULTS AND ANALYSIS

To test the new traffic signal control strategy and various other cases, three simulations of 2.5 h were run for each case and averaged. Evaluation recording began after the first half hour to allow the system to build up. Results are reported on a per-hour basis.

Base Case: Fixed Time Control with Progression

The base case is fixed time control with a 90-s cycle, matching the basic strategy now in place. To provide a fair basis of comparison against the other strategies, signals were retimed based on 2003 traffic volumes by using the *Highway Capacity Manual* (*12*). Green starts for Avenida Ponce de Leon northbound were initially offset to provide a 40 km/h green wave. Offsets were later manually adjusted to allow standing queues to clear, based on observation of the traffic animation.

Results for the base case and the other test cases are presented in Table 1. Travel times and delays are from the Santurce segment of Avenida Ponce de Leon only. Transit times and delays apply only to Route 1. The transit times, 15.6 min northbound and 14.9 min southbound, show only a small benefit of the exclusive contraflow lane southbound. It is worth noting that, whereas the contraflow bus lane frees buses from congestion with other traffic, it also traps them in congestion with other buses, since buses cannot overtake one another. With flows of 21 to 36 buses per hour, interference from other buses is not uncommon. Travel time for general traffic northbound is 8.7 min. The 7-min difference between bus and general traffic travel time northbound is not striking when one realizes that, because of the extreme passenger loads, northbound Route 1 buses spend on average 6.4 min of dwell time in Santurce. (Not counting boardings at the initial stop, the remaining stops generate on average 97 passenger movements through the critical door per bus trip, at 4 s per movement.) Naturally, some of that dwell time overlaps with traffic delay.

Delay reported in the table is the additional time a person or vehicle needs to get from one point to another compared with the free-flow travel and therefore includes traffic delay as well as (for transit) dwell time. Delay to all travelers includes transit travelers on Route 1 and on other routes. As a measure of regularity, the standard deviation of load measured approaching the ninth stop the buses reach in Santurce in their respective directions was used; these locations were chosen because they showed the greatest variance in passenger load. Northbound load varies little because it depends primarily on the load leaving Sagrado Corazon station, where dispatching there was assumed to be on time. The standard deviation, 5.3, is quite small in comparison with an average load of 57. In contrast, southbound load depends mostly on passengers picked up along the route. The standard deviation of 15.2 is large compared with the average load of only 18.

Actuated, Priority Control: Full and Incremental Effects

The highly interruptible, actuated control strategy with conditional priority is the fourth case reported in Table 1. Compared with the base case, it reduces transit travel time by just over a minute in each direction, while leaving travel time for general traffic unchanged. Transit delay falls by 13 person-hours per hour; that appears to be enough to justify considerable investment. Regularity, as measured by standard deviation in load, actually worsens but by too small an amount to be significant.

To better understand the incremental effects of the new control strategy, two intermediate cases are presented: actuated control (*a*) without active priority, and (*b*) with absolute priority. One can see that much of the benefit gained by the new control strategy is simply due to switching from coordinated fixed time control to actuated control (which, as described, includes loose coordination). Compared with conditional priority, absolute priority speeds southbound bus travel in the contraflow lane by 0.2 min, but it slows northbound travel by an equal amount, presumably because of its negative effect on general traffic flow. Also, regularity, as measured by the standard deviation of load, suffers noticeably with absolute priority.

To better understand the effect of the loose coordination strategy, average cycle times were tracked at the 15 signalized intersections in Santurce, indicated in Figure 3. All of them show an average cycle length below 90 s (grand mean = 52.8 s), which probably explains a good deal of the reduction in delay. The overall picture is one of diverse cycle lengths (between-intersection standard deviation = 8.2 s), yet with two strings of three or four neighboring intersections with equal or near-equal cycle lengths. This figure indicates that the system is working as envisioned—facilitating progression to a limited extent but also responding to transit requests in a manner that varies among intersections.

TABLE 1 Impacts

Case	TRAVEL TIME: Transit Travel Time [min], Northbound	TRAVEL TIME: Transit Travel Time [min], Southbound	TRAVEL TIME: Traffic Travel Time Northbound [min]	DELAY: Transit Delay [person-hours/hour]	DELAY: Transit Delay [vehicle-hours/hour]	DELAY: Delay to All Travelers [person-hours/hour]	DELAY: Delay to All Travelers [vehicle-hours/hour]	STD DEV[a] LOAD Taken from 9th Stop (Bautista/Pavia): Northbound	STD DEV[a] LOAD Taken from 9th Stop (Bautista/Pavia): Southbound
Fixed Time	15.6	14.9	8.7	76.7	1.7	219.6	97.9	5.3	15.2
Actuated, No Priority	14.7	13.9	9.0	65.9	1.4	199.7	96.9	5.9	15.7
Absolute Priority	14.6	13.6	8.8	61.2	1.3	191.6	90.9	6.2	16.9
Conditional Priority	14.4	13.8	8.7	63.6	1.3	183.0	85.0	5.9	15.5
Fixed + 10% More Traffic[b]	17.7	14.9	11.1	84.0	1.8	283.5	139.8	6.1	15.3
Conditional + 10% More Traffic[b]	14.5	13.8	8.9	65.6	1.4	218.2	113.5	5.7	16.5
Fixed + Crowding Impact	16.6	15.1	9.0	86.2	1.8	233.0	100.8	6.3	15.2

[a] STD DEV LOAD = standard deviation of load.
[b] Results for Traffic Delay have been scaled to match the number of vehicles in the base case.

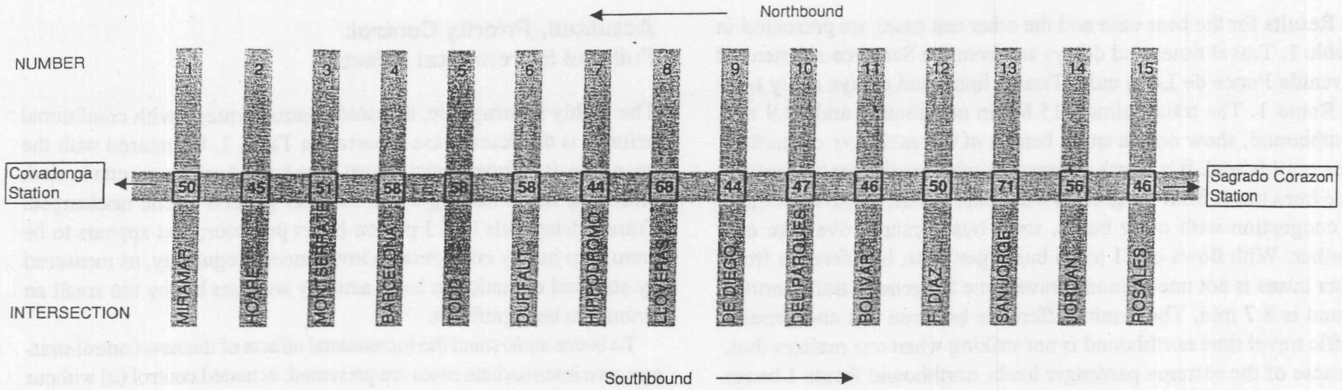

FIGURE 3 Average cycle times (seconds) with conditional priority.

As mentioned earlier, conditional priority is highly dependent on detailed and carefully constructed schedule running times by segment. Scheduled running times were generated following the approach suggested by Muller and Furth (*9, 13*), with 85th-percentile completion times predicted by simulations using absolute priority. This approach offers a balance by being generous (with 85th-percentile times, not mean times) under an operating environment that is unusually fast. However, as reported, transit travel times using absolute priority were scarcely faster than without priority at all. Because of rather generous resulting schedules and the assumption of on-time dispatching, buses rarely ran late and therefore made few priority requests. In the conditional priority case, the average number of priority requests per bus trip was only 0.9 northbound and 3.3 southbound. Research is needed to develop improved methods of generating scheduled run times for conditional priority.

Impact of Traffic Growth

What will happen if traffic increases or if the projected traffic volumes are low? To answer that, traffic volumes (except bus) were increased by 10% for two cases. With fixed time control, travel time increased substantially (by 2.1 and 2.4 min for bus and general traffic, respectively, northbound). With the interruptible strategy with conditional priority, increases in travel time are barely perceptible (below 0.2 min). This result further indicates the strength of the design with its traffic responsiveness and capacity-preserving features.

Impact of Crowding on Unit Boarding and Alighting Times

Finally, a scenario was run to model the impact of crowding on unit boarding and alighting times. With this model, which is based on speculation because of the lack of data, when vehicle load exceeds 40, the unit boarding and alighting time increases linearly by the factor $f = 0.025 \times (\text{load} - 40)$. Because the only part of VISSIM that is user programmable is the traffic signal controller, traffic signals were used to model this effect. By using a helpful VISSIM feature, a transit-only traffic signal was installed a few meters beyond the bus stop, with transit detectors just before and just after the bus stop. From the difference in time registered at the transit detectors, and allowing a few seconds for the bus movement to the transit signal, it was possible to estimate VISSIM's nominal dwell time. If the load on arrival was greater than 40, the controller logic calculated f and held the transit traffic signal red for a time equal to $f \times$ (nominal dwell time). For example, if load is 60, $f = 0.5$, so unit boarding time should increase from 4.0 to 6.0 s. If five people board, nominal dwell time is $5 \times 4.0 = 20$ s, while the dwell time accounting for the crowding effect should be $5 \times 6.0 = 30$ s. VISSIM will first hold the bus at the bus stop for the nominal dwell time and then hold it again for $f \times 20 = 10$ s at the transit traffic signal.

The tested scenario with the crowding impact used the traffic control of the base case (fixed time with progression). In the southbound direction, where there is little crowding, there was little impact. In the northbound direction, transit travel time increased by a full minute compared with the base case, the standard deviation of load increased by 1, and travel time for traffic also increased slightly. This result indicates the importance of better studying the effect of crowding on unit boarding and alighting times, an effect rarely accounted for in transit simulations. If the magnitude of the effect is at all close to what was speculated, then this effect makes a measurable difference in transit and traffic performance and deserves to be accounted for in models. The result also underscores the importance of taking measures to speed boardings and alightings (e.g., low floors, wide doors, off-vehicle payment) and of operational control to prevent high levels of crowding.

CONCLUSION

This study demonstrates the feasibility and societal benefits of a highly interruptible traffic signal control program with transit priority. It shows that, at least on a road with one-way general traffic and two-way bus travel, adequate progression can be maintained without a fixed or common cycle, allowing more flexibility for priority interruptions and therefore a greater measure of transit priority.

More research is needed in applying the principle of highly interruptible signal control to settings with general traffic in both directions and more complex signal phasing that offers a greater variety of priority tactics. Research is also needed in determining optimal schedules for schedule-based priority.

ACKNOWLEDGMENT

This research was supported by the Puerto Rico Highway and Transportation Authority through the Tren Urbano Technology Transfer program.

REFERENCES

1. Bienvenidos al Tren Urbano. www.dtop.gov.pr/TU/frtuo.htm. Accessed Aug. 1, 2001.
2. Multisystems. *Tren Urbano Feeder Service Plan—Final Report.* Tren Urbano Office, San Juan, Puerto Rico, Dec. 2000.
3. Balke, K., C. Dudek, and T. Urbanik. *Development and Evaluation of an Intelligent Bus Priority Concept.* Institute of Transportation Studies, University of California, Berkeley, 1999.
4. Hunter-Zaworski, K. M., W. C. Kloos, and A. R. Danaher. Bus Priority at Traffic Signals in Portland: The Powell Boulevard Pilot Project. In *Transportation Research Record 1503,* TRB, National Research Council, Washington, D.C., 1995, pp. 29–33.
5. Luh, J. Z. Case Studies for Simulation of Roadway and Traffic Operations Using CORSIM. *ITE Journal,* Vol. 71, No. 7, 2001, pp. 34–40.
6. Furth, P. G., and T. H. J. Muller. Conditional Bus Priority at Signalized Intersections: Better Service with Less Traffic Disruption. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1731,* TRB, National Research Council, Washington, D.C., 2000, pp. 23–30.
7. Muller, T. H. J. *Traffic Signal Control.* Delft University of Technology, Netherlands, 1997.
8. Turnquist, M. A. Strategies for Improving Reliability of Bus Transit Service. In *Transportation Research Record 818,* TRB, National Research Council, Washington, D.C., 1981, pp. 7–13.
9. Muller, T. H. J., and P. G. Furth. Integrating Bus Service Planning with Analysis, Operational Control, and Performance Monitoring. *Proc., Intelligent Transportation Society of America Annual Meeting,* Boston, May 2000, Intelligent Transportation Society of America, Alexandria, Va., 2000.
10. *VISSIM User Manual, Version 3.50.* PTV Vision, Karlsruhe, Germany, 2001.
11. Bloomberg, L., and J. Dale. A Comparison of VISSIM and CORSIM Traffic Simulation Models on a Congested Network. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1727,* TRB, National Research Council, Washington, D.C., 2000, pp. 52–60.
12. *Special Report 209: Highway Capacity Manual,* 3rd ed. (1997 update). TRB, National Research Council, Washington, D.C., 1998.
13. Muller, T. H. J., and P. G. Furth. Trip Time Analyzers: Key to Transit Service Quality. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1760,* TRB, National Research Council, Washington, D.C., 2001, pp. 10–19.

Publication of this paper sponsored by Committee on Traffic Signal Systems.

Paper No. 02-4030

Program for Optimizing Diamond Interchanges in Oversaturated Conditions

Vijay G. Kovvali, Carroll J. Messer, Nadeem A. Chaudhary, and Chi-Leung Chu

Traditionally, optimization of diamond interchange timings has been done with PASSER III for standard and special diamond phasing sequences. PASSER III is limited because it is designed for undersaturated conditions. It applies vertical stacking of queues and is not capable of modeling queue spillback conditions in its current form. This deficiency is addressed by the arterial signal coordination software (ASCS), whose capabilities in timing diamond interchanges in undersaturated and oversaturated conditions are presented here. ASCS consists of three modules: (*a*) input module, a user interface through which volume and geometry inputs can be provided to the program; (*b*) optimization module, a genetic algorithm-based optimization routine that can optimize signal timings; and (*c*) analysis module, which consists of a bandwidth analysis routine and a delay analysis routine (DAR). The DAR is a mesoscopic simulator that applies a second-by-second analysis of flows for modeling flows accurately. DAR applies horizontal stacking of queues and shock wave analysis to estimate the performance of traffic operations. Validation of ASCS for oversaturated arterial links against PASSER III and CORSIM was conducted. The results indicate that delay and throughput estimation in ASCS are realistic. The genetic algorithm-based optimization routine in ASCS was applied to estimate diamond interchange timings for three scenarios. Where queue spillback occurred, ASCS clearly outperformed PASSER III. ASCS produced near-optimal results for all scenarios studied.

Diamond interchanges serve as critical links between roadway facilities of two different classifications: freeways and surface street systems (*1*). During peak traffic conditions, inefficient operation of a diamond interchange and adjacent traffic signals may cause either system to become a bottleneck, downgrading not only the capacity of the interchange but also that of the arterial and, in some cases, even the capacity of the freeway ramps.

Operational issues in the analysis of signalized diamond interchanges are as follows:

- Traffic patterns at signalized interchanges have a high percentage of turning volumes. This requires that special attention be paid to the origin-destination patterns at the diamond.
- Queue spillback considerations are extremely important at compressed and tight urban diamonds.

Signal control at diamond interchanges traditionally has been provided by either a three-phase pretimed signal sequence or the four-phase, two-overlap signal phase sequence (TTI 4-phase). Current signal controls at diamond interchanges are typically variations and/or combinations of these two basic phasing sequences with pretimed or actuated controllers (*2*). For optimizing the timings at a diamond interchange, PASSER III is most commonly used (*1*).

Although PASSER III (*3*) can analyze and optimize signal timings for minimizing delay within each interchange, the delay analysis routine (DAR) used by PASSER III is applicable only to undersaturated conditions. Also, because the delay routine of PASSER III applies vertical stacking of queues, the optimization could be inaccurate for short links even during undersaturated conditions. Currently, no other program exists that has been specifically designed for optimizing signal timings at diamond interchanges.

A genetic algorithm (GA)-based program that is capable of optimizing signal timings at a diamond interchange is described here. The GA model uses a delay analysis model as its fitness function, which can analyze measures of effectiveness during undersaturated and oversaturated conditions.

BACKGROUND

The operating conditions at signalized interchanges are unique. Because of the immediate upstream metering of traffic flows, downstream flows are not subject to random arrivals. At the same time, these flows cannot be characterized by any of the simple progression types described in the *Highway Capacity Manual* (HCM) (*4*). Also, in urban areas, diamond interchanges are usually formed by two closely spaced signalized intersections operating at the same cycle length.

A review of the software evaluation techniques for interchange operations was conducted by Chen et al. (*5*). They found that TRANSYT-7F, HCS, PASSER II, and PASSER III were most commonly used. Another survey was recently conducted by Chaudhary et al. (*1*). The survey findings were that most engineers in Texas applied PASSER III for optimization of diamond interchanges. The survey also found that special phasing sequences were quite often applied for timing diamond interchanges.

The available software for optimizing diamond interchanges, including special phasing sequences, is limited to PASSER III. Although Synchro 4.0 (*6*) is capable of analyzing TTI 4-phase timings for a given cycle length, it is not capable of optimizing within the special phasing sequences while varying the cycle lengths.

DEVELOPMENT OF ARTERIAL SIGNAL COORDINATION SOFTWARE

Many agencies in Texas apply PASSER III for optimizing diamond interchange timings (*1*). As already stated, the PASSER III optimization procedure is limited to undersaturated conditions. A new

V. G. Kovvali, Innovative Transportation Concepts, 1128 Northeast 2nd Street, Suite 204, Corvallis, OR 97330. C. J. Messer, Suite 304C, N. A. Chaudhary, Suite 304A, and C.-L. Chu, Suite 309C, CE/TTI Building, Texas Transportation Institute, Texas A&M University, College Station, TX 77843-3135.

arterial signal coordination software (ASCS) was developed that is capable of timing oversaturated diamond interchanges. ASCS consists of three main modules: the input module, the optimization module, and the analysis module. The input module consists of a user interface through which input volumes and geometric conditions can be entered. The optimization module consists of a GA-based signal timing optimizer. The analysis module consists of two routines: a bandwidth analysis routine and a DAR. The model was developed for a generic arterial. In this paper, only the work related to timing diamond interchanges is presented. ASCS has the following features, which allow it to be applied to diamond interchanges:

- The saturation flow module in ASCS is the same as the one used in PASSER III-99. A saturation flow rate prorating model based on volumes was used in both programs.
- ASCS has the capability of including an origin-destination trip matrix similar to PASSER III, which allows traffic patterns to be coded accurately.
- The program includes capabilities for calculating special diamond phase timings for different cycle lengths. Basic three-phase, extended three-phase, and TTI 4-phase timings can be calculated. Timings for standard phasing sequences can also be obtained.
- An improved DAR is included in the ASCS. DAR is capable of simulating traffic flows and estimating the delay trends realistically in both undersaturated and oversaturated conditions. The DAR model makes the diamond signal optimization included in ASCS more realistic and accurate than does PASSER III-99. A description of the DAR model is presented in the next section.

DAR

Estimation of delay in the DAR model is conducted in two steps: initialization and analysis.

Initialization

All simulation models require an initialization step. This initialization step is the warm-up period that allows the bias due to initial conditions to be removed to achieve at least near-steady-state conditions. Park (*7*) summarized the various methods of initialization applied for traffic simulation:

- Perform plenty of computer runs so that the data from the initialization period are not significant compared with the steady-state data.
- Run the simulation for a particular warm-up period, when the simulation results are not recorded. After the end of the warm-up period, the simulation results are recorded. Estimation of the warm-up period becomes an important issue.
- Choose initial starting conditions that are more typical of steady-state conditions.

The first two options are computationally more intensive than the third option. Hence, the third option was selected. The DAR model applies initial starting conditions in two steps:

- In the first step, for one cycle, an extension of the delay-difference-of-offset model (*8*) is applied. Starting from the first upstream link in the A direction, all the computations on each link are conducted and the flows are projected to the next link. This process is performed for all links in the A direction first. The process is repeated for the B direction starting with the first upstream link in the B direction. Available link storage is not considered in this step. Hence, queue lengths greater than the link length can occur in the calculations.
- In the second step, the process is applied for one cycle. The methodology applied in this step is the same as the one applied in the analysis step explained in the next section. This step allows the artery to reach an equilibrium condition. At the end of this step, flow storage and internal link storages are updated for all movements.

Analysis

In the analysis step, the process is applied for n number of cycles. For each time step, starting from the downstream link in each direction, the flows are updated through all the links. The data are recorded for the last m cycles, where $m \leq n$. Both m and n are variable numbers that can be selected by the user.

A brief description of the procedures applied in the analysis step is as follows:

- For each second, all the links are analyzed from the most downstream link to the upstream link in each direction. This allows both precedence and dependence relationships to be incorporated.
- The flows are categorized as external-to-external, external-to-internal, internal-to-external, and internal-to-internal flows. The delays for external-to-external flows are computed by applying the HCM delay equation.
- For a given second, for a particular link, the downstream flows are first updated. The arrival flows at the downstream flows are obtained by applying the TRANSYT-7F platoon dispersion algorithm to upstream flows. The input flows are added to flow storage for each of the downstream movements. The internal-to-external flows occurring at the downstream link are updated from the flow storage available at that particular second and the output profile at that particular second.
- The internal-to-internal flows and the external-to-internal flows are updated by calculating available storage at the back of the queue. This available flow storage is calculated by applying shock wave theory (*9*) to the back of the queue.
- Delays and throughputs are calculated by summing the second-by-second flow storage values and output flows, respectively. Stops, queue spillback, and green starvation time were also calculated.
- The procedure is repeated for the time period specified and the system delays, throughputs, and stops are calculated.

Validation of DAR

Validation of DAR required that the different features available in the model be validated separately. Two different scenarios were selected that required close inspection to determine whether the model performs accurately in congested situations. These scenarios are as follows:

- Arterial system with spillback conditions due to short downstream link, and
- Arterial system with spillback conditions due to oversaturated downstream link.

Arterial System with Spillback Conditions Due to Short Downstream Link

Hadi and Wallace (*10*) studied two closely spaced intersections for validating TRANSYT-7F version 8.1. The artery (Figure 1) presented in their research was used for validating the congested conditions modeling in DAR. The stop-line to stop-line length of the link is 100 m (328 ft). Considering that the upstream node has two northbound lanes of width 3.9 m (12 ft) each, the internal storage length is 92.7 m (304 ft) per lane. The jam distance headway applied in this research is 7 m (23 ft) per vehicle.

This model allows the conditions where the queue spillback results from a short undersaturated downstream link length to be investigated. Both the upstream and downstream links are undersaturated with a volume/capacity (v/c) ratio of about 0.9. Although the system is undersaturated, the queue in the link can extend to the upstream link because of the limited storage space available. Depending on the offset, two boundary conditions can be studied for this arterial system: demand starvation and upstream traffic movement blocking. Demand starvation is the condition in which a movement has a green, but no flow output occurs because of zero inflow. This condition occurs usually because of a bad relative offset. Demand starvation is represented in DAR as the time during which the movement has a green and no output flow occurs because of zero inflow and queues. Upstream movement blocking decreases the throughput. Blocking can occur because of a bad relative offset or because of queue buildup. For the arterial studied, it is due to a bad offset relationship.

The maximum downstream throughput can be achieved when the relative offset is such that the downstream queue does not block upstream movements. Because the link is undersaturated, the maximum possible downstream throughput is equal to the traffic flow feeding this movement, which is 1,400 vehicles per hour (vph). Similarly, the minimum arterial throughput occurs when the offset is such that only the vehicles stored on the link are output during the green. For the jam distance headway considered, a total of 13.22 vehicles can be stored in each lane per cycle. Hence, the throughput for a 120-s cycle length is 26.43 vehicles, which equals 793 vph. Figure 2 presents the relationship between offset and throughput for the arterial downstream movement as estimated by CORSIM and DAR. In Figure 3, demand starvation for the arterial for the different offsets is presented. A study of the throughput in Figure 2 and demand starvation in Figure 3 indicates that starvation leads to a decrease in throughput, as expected.

As indicated in Figure 2, the DAR model throughput values are very similar to CORSIM results. For the offset values of 40 to 80, the downstream throughput is equal to the vehicles stored in the link. CORSIM allows the storage of vehicles within the intersection area. Because of this, two additional vehicles are stored in link during each cycle, which resulted in the increased throughput of about 60 vph compared with the DAR throughput for these offset values. A comparison of the arterial delay is presented in Figure 4. The trends indicate that the decrease in throughput due to queue spillback is modeled accurately.

Arterial System with Spillback Conditions Due to Oversaturated Downstream Link

The arterial system used for investigating queue spillback due to an oversaturated downstream link is presented in Figure 5. This scenario is also obtained from Hadi and Wallace (*10*). For this scenario, the downstream arterial through-movement capacity is less than the feeding volumes to this movement (v/c about 1.2). The distribution of the traffic to this movement is 75% arterial traffic and 25% ramp traffic.

The boundary conditions studied by investigating this artery are (*a*) queue buildup and queue blocking, and (*b*) shock wave application for studying the movement of the back of the queue.

The length of the link is 305 m (1,000 ft). Because the downstream v/c is only about 1.2, the queue buildup per cycle is low enough that the link does not get blocked until the simulation has run for a few cycles. The queue buildup and blocking was modeled by simulating the arterial for eight cycles (about 15 min) beyond the initialization period.

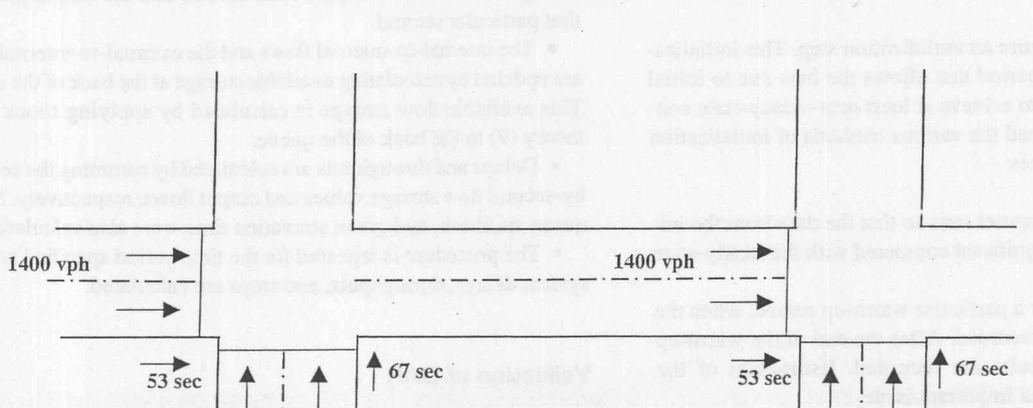

FIGURE 1 Arterial system with spillback conditions due to short downstream link (vph = vehicles per hour).

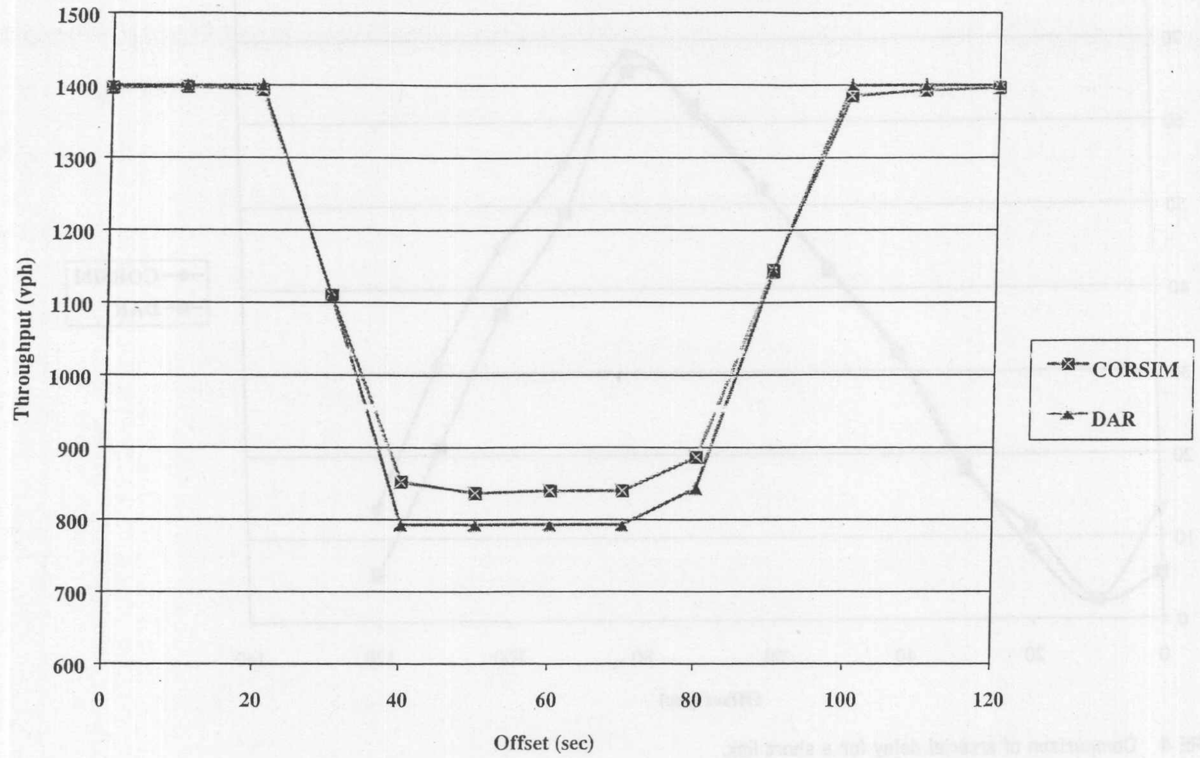

FIGURE 2 Comparison of downstream arterial throughput for a short link.

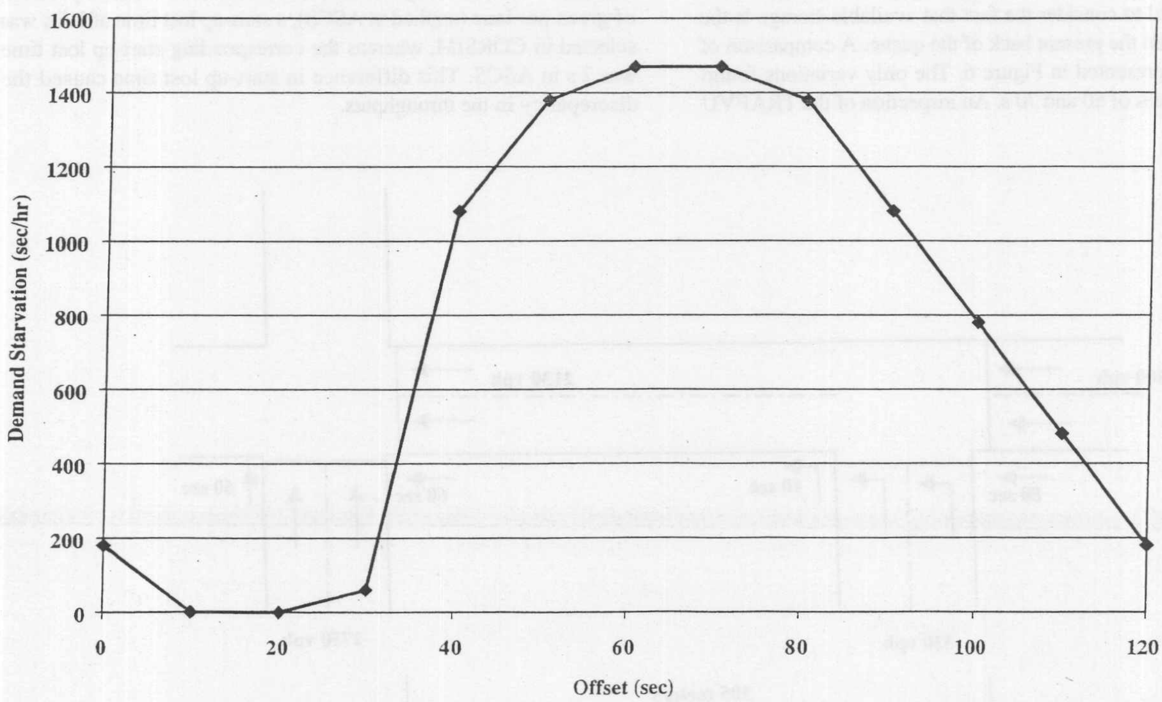

FIGURE 3 Demand starvation modeled at the downstream arterial movement for a short link in DAR.

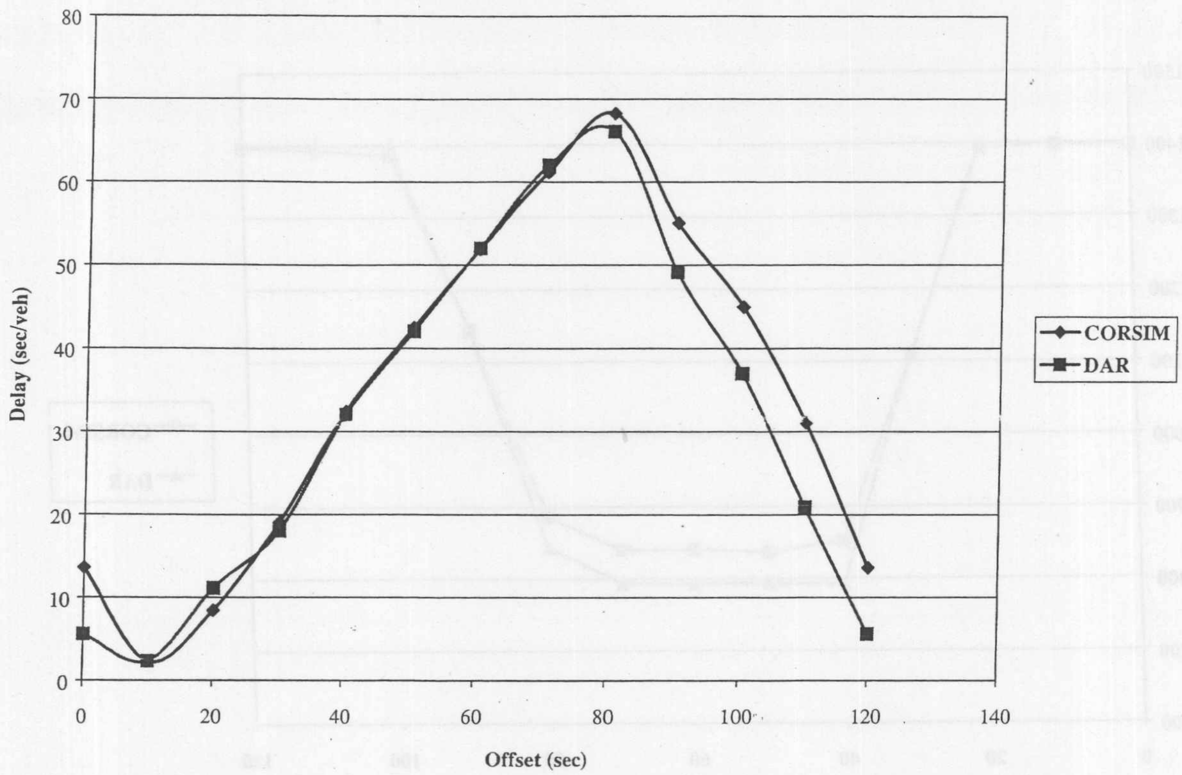

FIGURE 4 Comparison of arterial delay for a short link.

Traditional macroscopic and mesoscopic models apply simple input and output modeling for considering available storage. This usually leads to inaccurate estimation of throughput and delays. Shock waves were applied to consider the fact that available storage is the space available from the present back of the queue. A comparison of the throughput is presented in Figure 6. The only variations found were for offset values of 60 and 70 s. An inspection of the TRAFVU simulations of these offset conditions indicated that these differences occurred mainly because of the calibration process applied to CORSIM. For obtaining saturation flows of 1,900 vehicles per hour of green per lane (applied in ASCS), a start-up lost time of 3.5 s was selected in CORSIM, whereas the corresponding start-up lost time was 2 s in ASCS. This difference in start-up lost time caused the discrepancy in the throughputs.

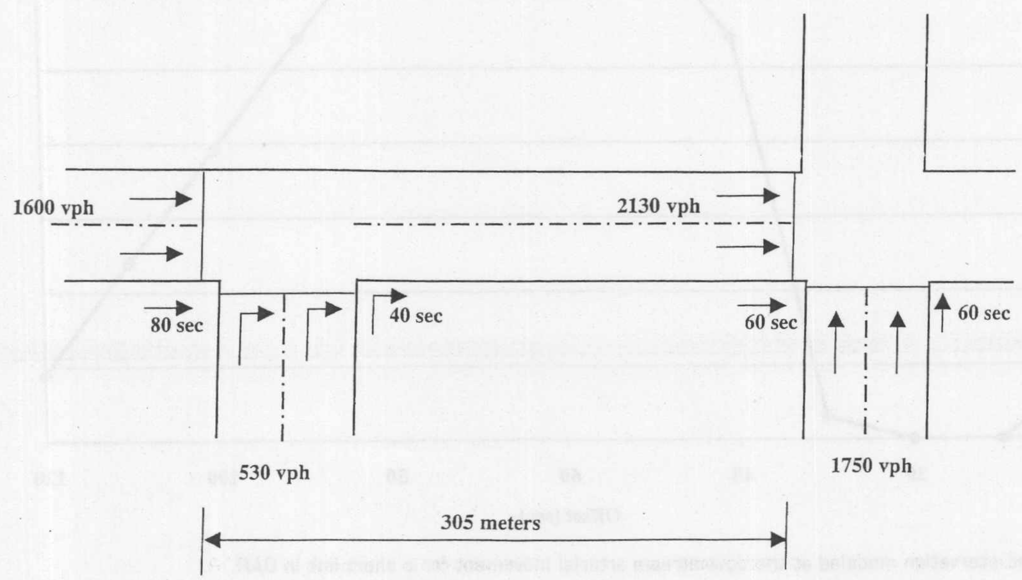

FIGURE 5 Arterial system with spillback conditions due to oversaturated downstream link.

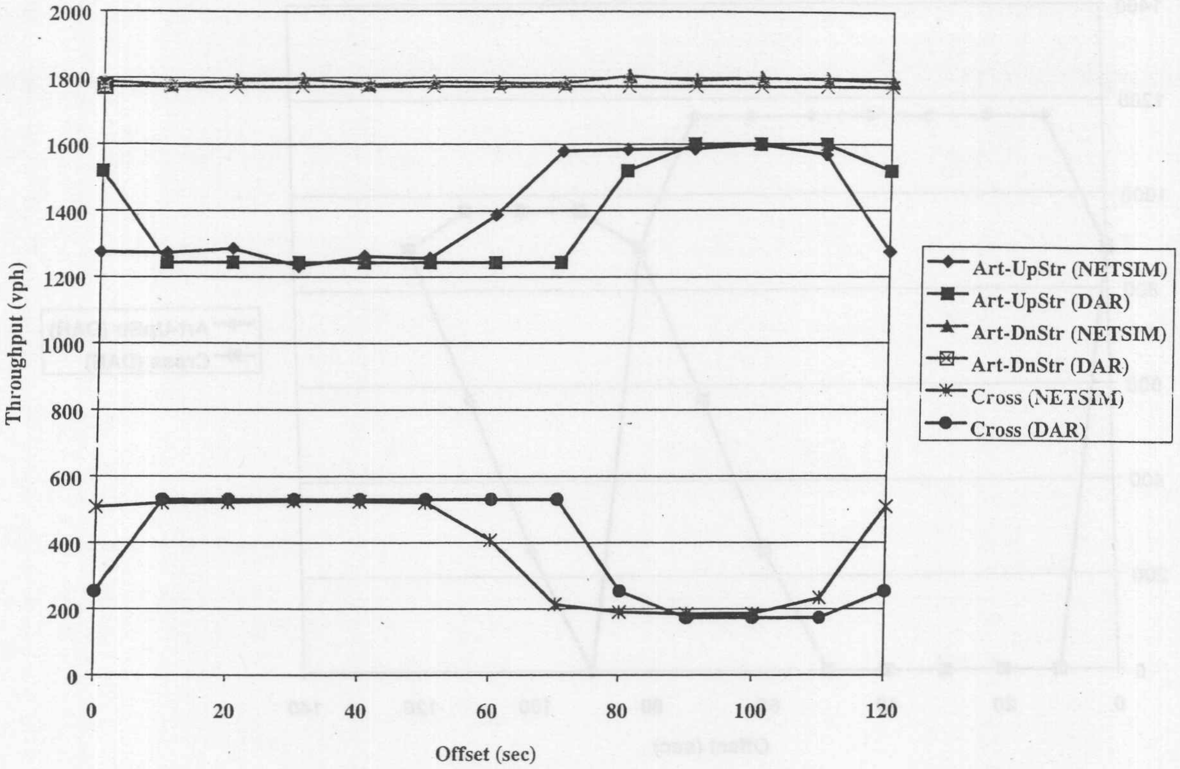

FIGURE 6 Throughput comparison for oversaturated link.

Regardless of the offset chosen, the downstream through movement is always saturated. The effect of offset is in determining the portion of the downstream link capacity assigned to the upstream arterial and cross-street traffic flows. A study of the spillback for the two movements clearly illustrates this phenomenon, as indicated in Figure 7.

APPLICATION OF ASCS TO DIAMOND INTERCHANGES

Comparison with PASSER III-99 and CORSIM

The program was compared with PASSER III and CORSIM for three different geometric conditions (interchange spacing of 66, 132, and 198 m, which correspond to 200, 400, and 600 ft, respectively) and two different volume conditions. The outputs for basic three-phase and TTI 4-phase are presented in Figures 8 and 9 for 66-m (200-ft) spacing and 300 vph per lane (vphpl) of volumes. DAR results conform rather well to CORSIM and PASSER III results for both basic three-phase and TTI 4-phase timing plans.

In Figures 10 and 11, the analysis is repeated for 66-m (200-ft) spacing but for higher volumes of 600 vphpl. These higher volumes correspond to a v/c of about 0.9 but because of the short spacing lead to queue spillback. The basic three-phase results for this condition indicate that PASSER III considerably overestimates interior delays. This is because of PASSER III's vertical queue stacking, which causes more vehicles to be queued in the link than possible. The results obtained from DAR are more consistent with the CORSIM results. The variability in CORSIM and DAR is partly due to lane blocking applied by CORSIM, which is not modeled in DAR, and the lack of a random delay component in DAR. Results for the other volume and geometric conditions were also found to be consistent with CORSIM output.

GA Routine for Optimization of Diamond Interchange Timings

In PASSER III, all the scenarios selected are analyzed and the optimal result is presented. If a lead-lead phasing is selected with a cycle range of 60 to 150 s and a 5-s increment, for each cycle length the delay is analyzed for each offset value (0 to Cycle 1). This method is computationally very expensive, especially for a more involved delay routine like the one used in ASCS.

The ASCS model can optimize a diamond for a selected range of cycle lengths for all the phasing sequences available in PASSER III. Because traffic patterns and queue spillback are explicitly considered, the results were more realistic than PASSER III in oversaturated conditions. The optimization can be conducted by trying all combinations, or a GA-based optimization routine (GA model) can be applied. The GA model provides a smarter and quicker optimization.

GA Representation

Selection of GA parameters is an important process for obtaining near-optimal results. A study of sensitivity analysis of GA parameters by Kovvali and Messer (*11*) concluded that, for each problem, a specific preliminary investigation should be conducted for obtaining the parameters. The solution space for diamond interchanges consists of a combination of at most three variables: cycle length, phasing sequence, and ring lag/internal offset:

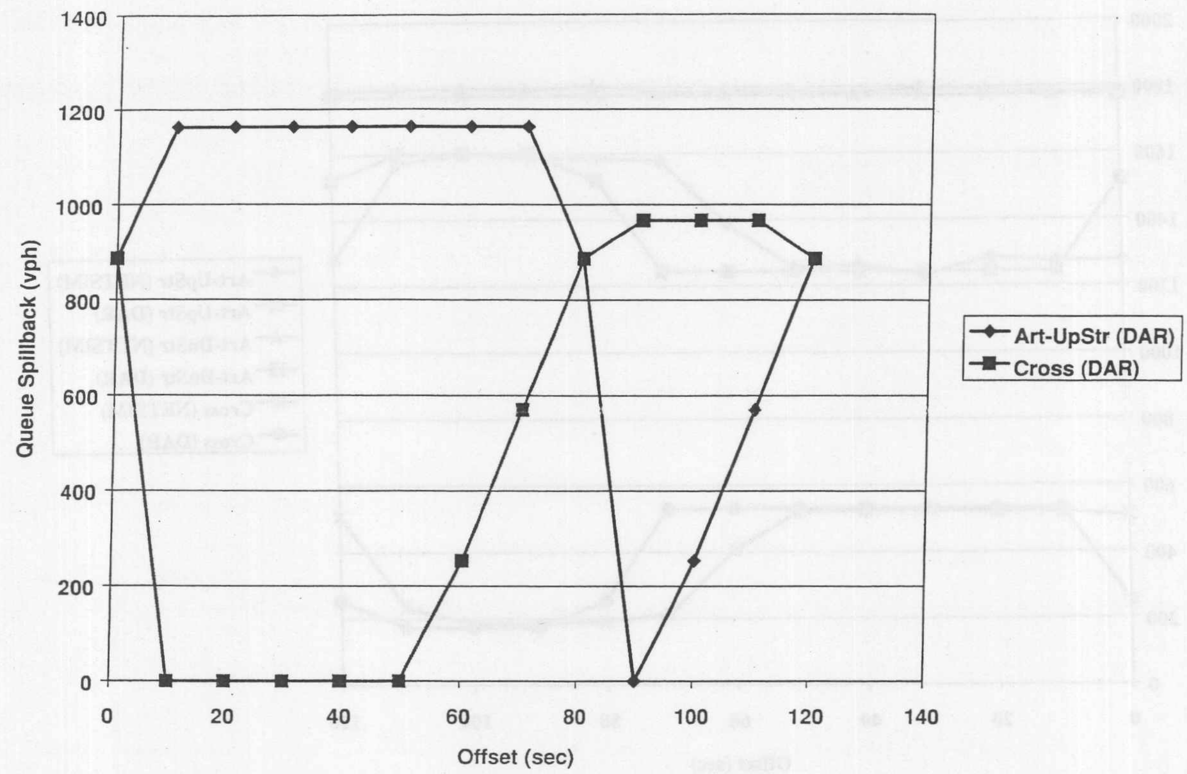

FIGURE 7 Modeling of spillback due to blocking of downstream link for an oversaturated link.

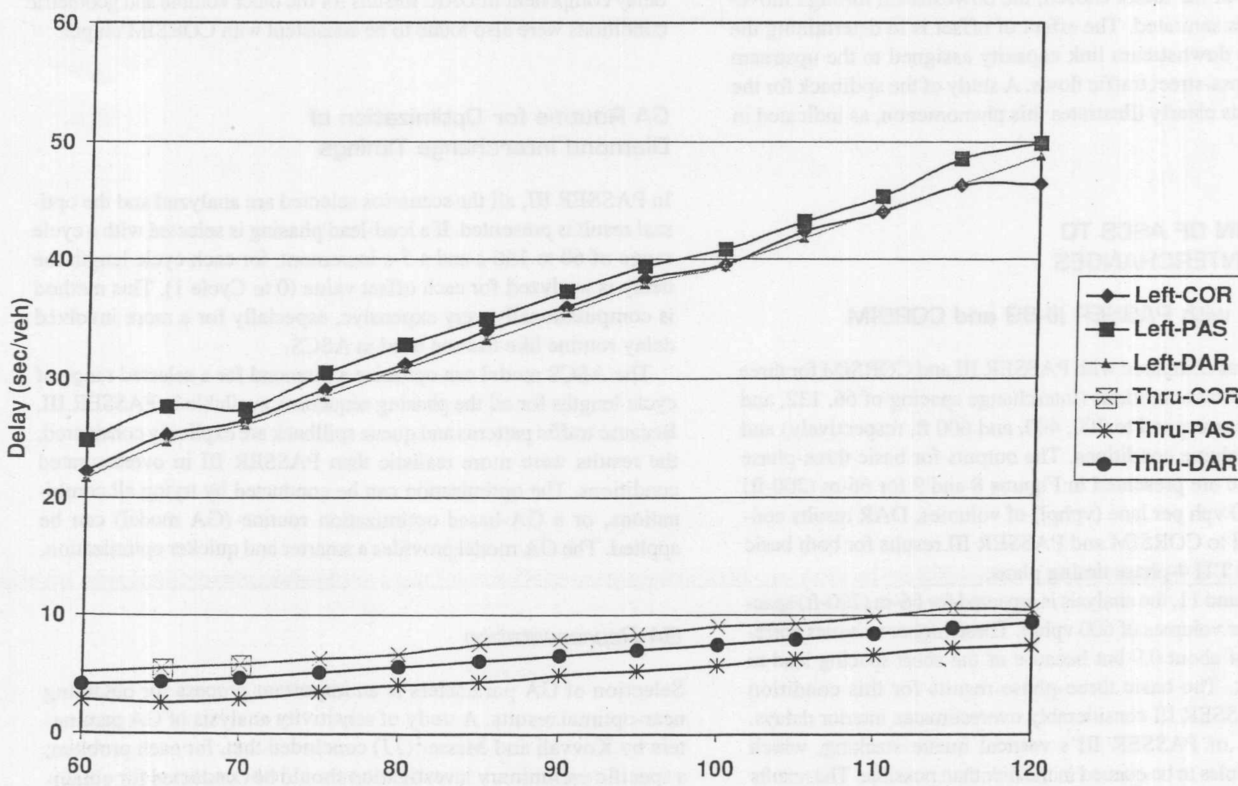

FIGURE 8 Comparison of DAR with PASSER III and CORSIM for basic three-phase for 300 vphpl for 66-m (200-ft) link length.

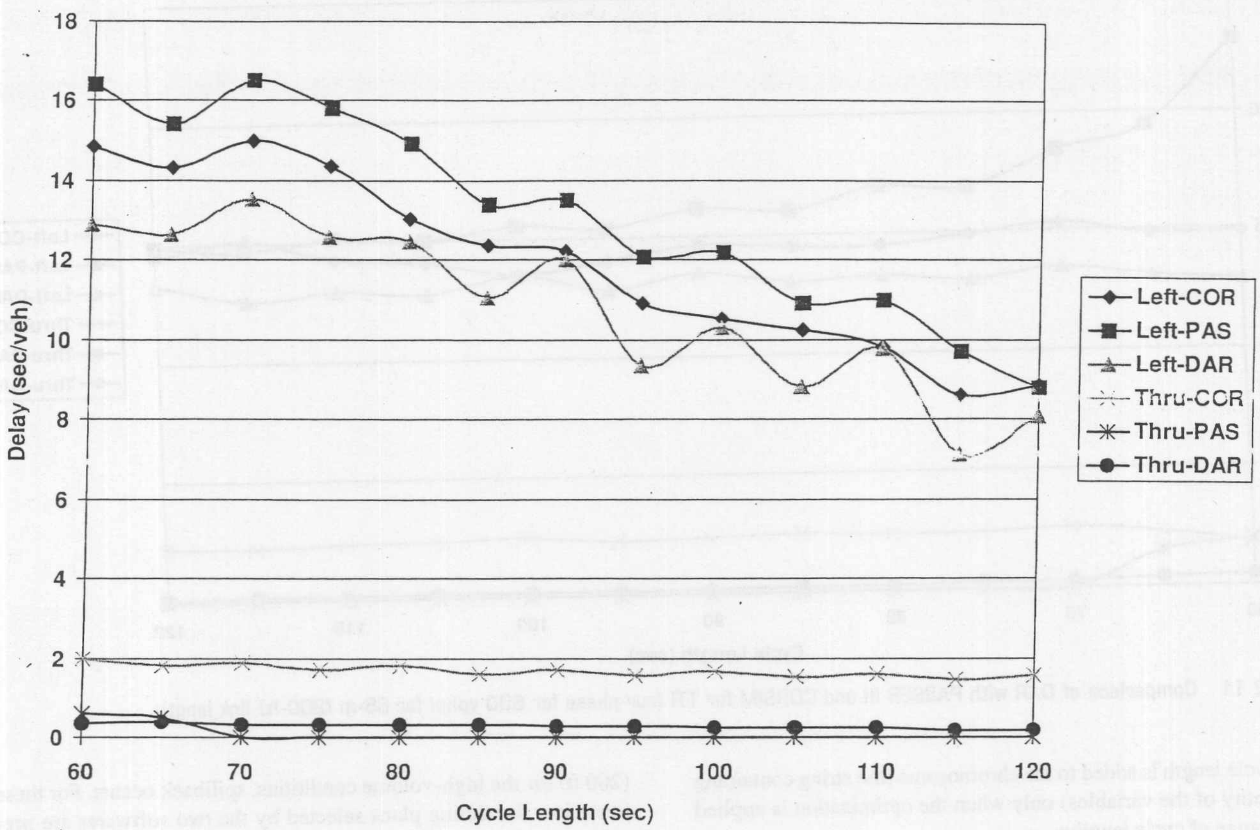

FIGURE 9 Comparison of DAR with PASSER III and CORSIM for TTI four-phase for 300 vphpl for 66-m (200-ft) link length.

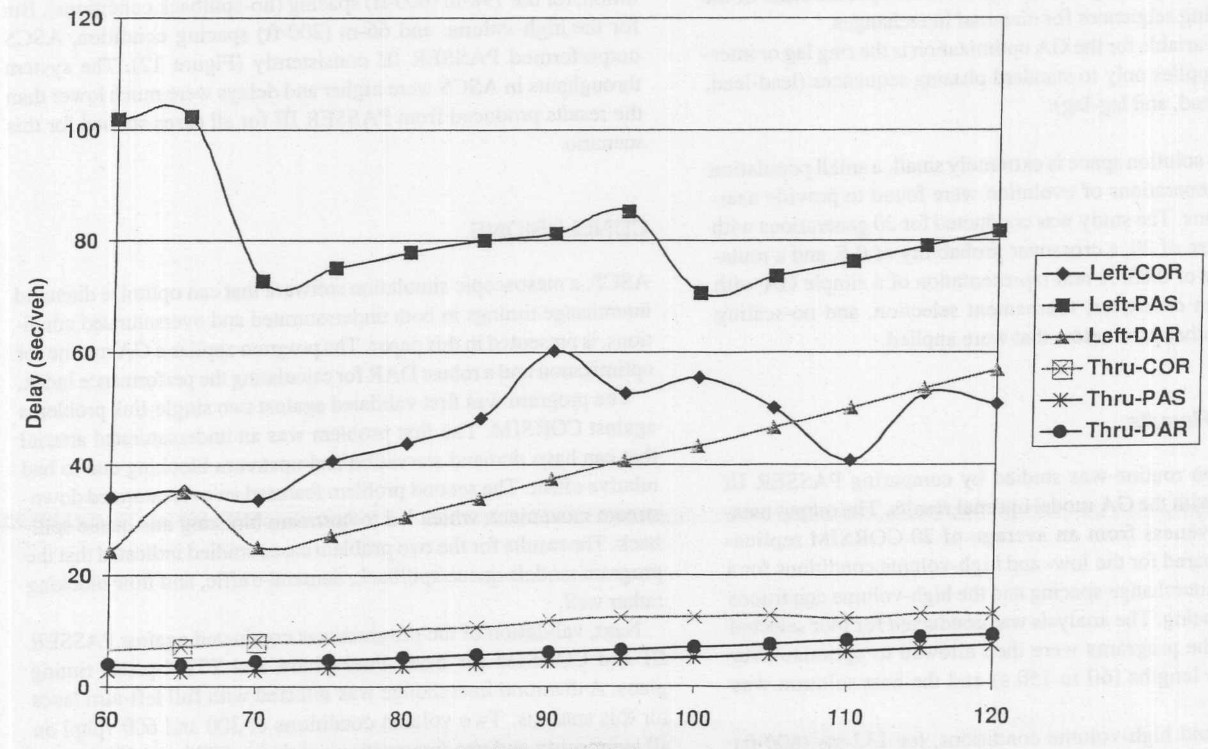

FIGURE 10 Comparison of DAR with PASSER III and CORSIM for basic three-phase for 600 vphpl for 66-m (200-ft) link length.

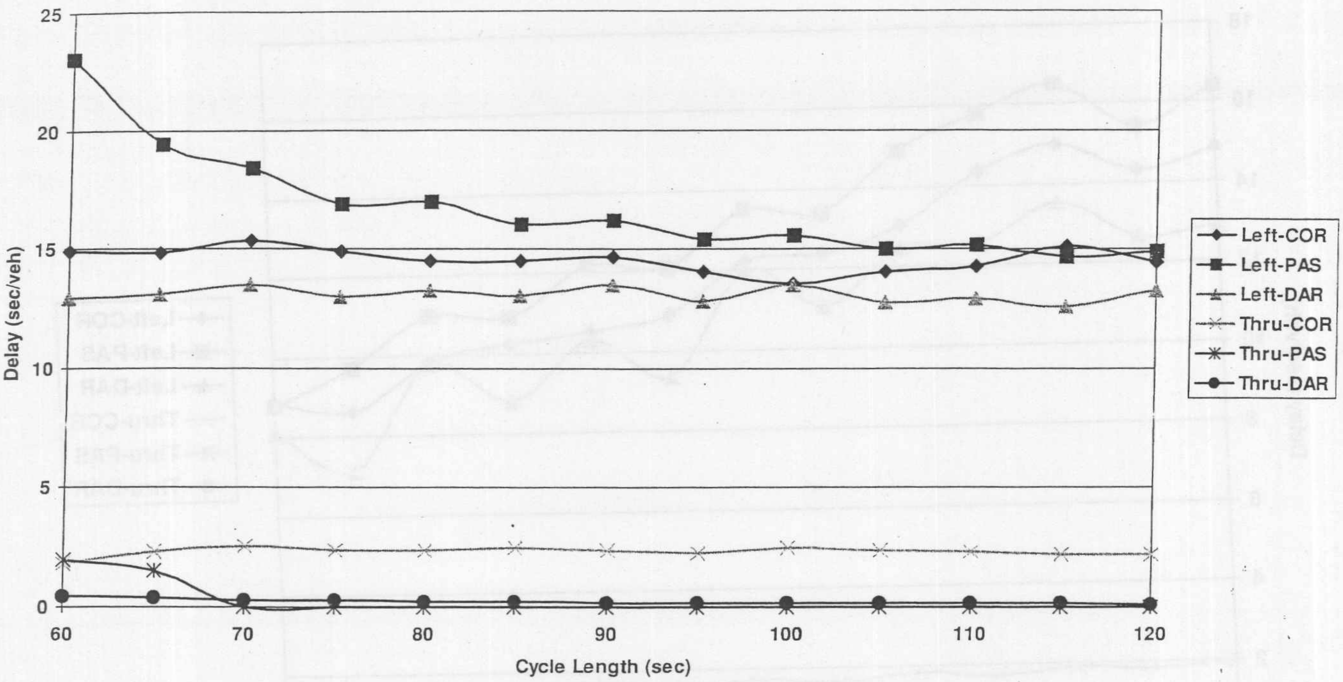

FIGURE 11 Comparison of DAR with PASSER III and CORSIM for TTI four-phase for 600 vphpl for 66-m (200-ft) link length.

- Cycle length is added to the chromosome (the string containing the identity of the variables) only when the optimization is applied over a range of cycle lengths.
- The seven possible phasing sequences are special phasing sequences (TTI 4-phase, basic three-phase, and extended three-phase) and standard phasing sequences (lead-lead, lead-lag, lag-lead, and lag-lag). The special phasing sequences are specific cases of the standard phasing sequences for diamond interchanges.
- The last variable for the GA optimization is the ring lag or internal offset; it applies only to standard phasing sequences (lead-lead, lead-lag, lag-lead, and lag-lag).

Because the solution space is extremely small, a small population size and few generations of evolution were found to provide near-optimal solutions. The study was conducted for 20 generations with a population size of 20, a crossover probability of 0.5, and a mutation probability of 0.05. A real representation of a simple GA with elitism, uniform crossover, tournament selection, and no-scaling comprises the other parameters that were applied.

Optimization Results

The optimization routine was studied by comparing PASSER III optimal results with the GA model optimal results. The output measures of effectiveness from an average of 20 CORSIM replications were compared for the low- and high-volume conditions for a 198-m (600-ft) interchange spacing and the high-volume conditions for the 200-ft spacing. The analysis was conducted for four selected cycle lengths. The programs were then allowed to optimize over a range of cycle lengths (60 to 150 s) and the best solution was selected.

For the low- and high-volume conditions, for 132-ms (600-ft) spacing, no spillback occurs. But when the spacing is only 66 m (200 ft) for the high-volume conditions, spillback occurs. For these scenarios, the timing plans selected by the two softwares are presented in Table 1. The results indicate that the two programs selected similar timing plans when queue spillback was not an issue.

No significant differences were found between PASSER III and ASCS for the low-volume condition and for the high-volume condition, for the 198-m (600-ft) spacing (no-spillback conditions). But for the high-volume and 66-m (200-ft) spacing condition, ASCS outperformed PASSER III consistently (Figure 12). The system throughputs in ASCS were higher and delays were much lower than the results produced from PASSER III for all cases studied for this scenario.

CONCLUSIONS

ASCS, a mesoscopic simulation software that can optimize diamond interchange timings in both undersaturated and oversaturated conditions, is presented in this paper. The program applies a GA routine for optimization and a robust DAR for calculating the performance index.

The program was first validated against two single-link problems against CORSIM. The first problem was an undersaturated arterial that can have demand starvation and upstream blocking due to bad relative offset. The second problem featured an oversaturated downstream movement, which led to upstream blocking and queue spillback. The results for the two problem cases studied indicated that the program models queue spillback, demand traffic, and flow blocking rather well.

Next, validation of the program was conducted against PASSER III and CORSIM for basic three-phase and TTI 4-phase timing plans. A diamond interchange was selected with full left-turn lanes for this analysis. Two volume conditions of 300 and 600 vphpl on all approaches and two intersection spacings of 200 and 600 ft were compared. For the low-volume condition, the results obtained from

TABLE 1 Optimal Timing Plans Selected by PASSER III and ASCS

	66 meters – High Volume		198 meters – Low Volume		198 meters – High Volume	
Cycle	PASSER III	ASCS	PASSER III	ASCS	PASSER III	ASCS
90	Basic 3-Phase	Lag-Lead 87 sec Offset*	Basic 3-Phase	Basic 3-Phase	Basic 3-Phase	Basic 3-Phase
100	Basic 3-Phase	Lead-Lag 0 sec Offset	Basic 3-Phase	Basic 3-Phase	Basic 3-Phase	Basic 3-Phase
110	Basic 3-Phase	Lag-Lead 107 sec Offset	Basic 3-Phase	Basic 3-Phase	Basic 3-Phase	Lag-Lead 89 sec Offset
120	Basic 3-Phase	Lag-Lead 117 sec Offset	Basic 3-Phase	Lag-Lead 5 sec Offset	Basic 3-Phase	Lag-Lead 97 sec Offset
Best Solution	Basic 3-Phase 90 sec Cycle	Lag-Lag 75 sec Cycle 74 sec Offset	Lead-Lead 60 sec Cycle 15 sec Offset	Basic 3-Phase 60 sec Cycle	Basic 3-Phase 90 sec Cycle	Basic 3-Phase 75 sec Cycle

* Offset in ASCS is defined as the time lag between arterial through movements in the A-direction

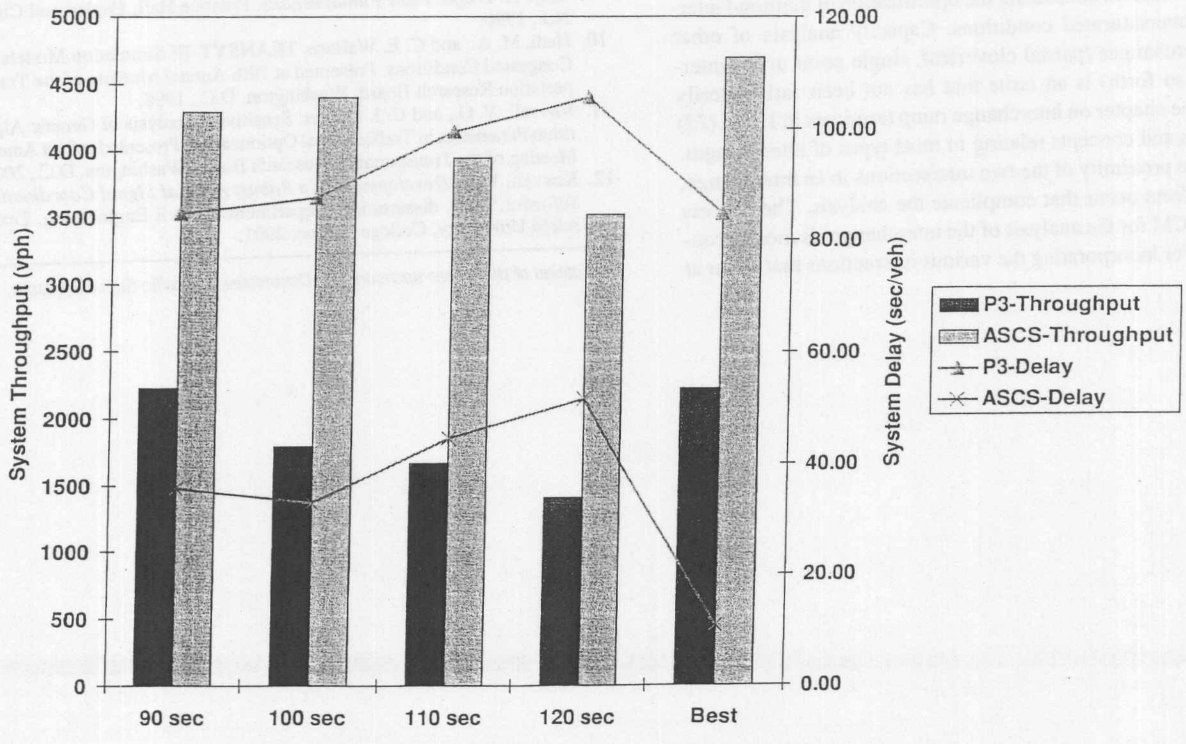

FIGURE 12 Comparison of optimal solution results of PASSER III and ASCS for 66-m (200 ft) spacing and 600 vphpl volumes.

both PASSER III and ASCS conformed well with CORSIM results. ASCS provided a better estimation of the delay trends than did PASSER III for the higher volume conditions compared with CORSIM. PASSER III overestimated delays for these conditions, especially for the basic three-phase timings. This is to be expected because queue spillback occurring within the interchange is not modeled in PASSER III, which occurs for basic three-phase timings and not in TTI 4-phase timings. For all the comparisons conducted, ASCS provided good estimation of delays and throughputs.

Finally, a comparison of the optimal solution results from PASSER III and ASCS was conducted. This was conducted for interchange spacing of 198 m (600 ft) and volume conditions of 300 and 600 vphpl, and for interchange spacing of 66 m (200 ft) and volume conditions of 600 vphpl. First, optimization capabilities of the software were studied by keeping the cycle lengths constant between the two programs. Then, for obtaining the best solution, the two programs were allowed to optimize over all phasing sequences possible at a diamond (standard and special) for a cycle length range of 60 to 150 s. Results indicate that ASCS output was similar to that of PASSER III when queue spillback did not occur. But when queue spillback occurred, ASCS clearly outperformed PASSER III.

The software described here, ASCS, is capable of providing signal coordination for both diamond interchanges and arterial signals. Kovvali's dissertation (*12*) provides more information about the programs features, capabilities, and performance.

FUTURE RESEARCH

This paper focuses on addressing the optimization of diamond interchanges in oversaturated conditions. Capacity analysis of other forms of interchanges (partial cloverleaf, single point urban interchange, and so forth) is an issue that has not been satisfactorily addressed. The chapter on interchange ramp terminals in HCM (*12*) presents ideas and concepts relating to most types of interchanges. Because of the proximity of the two intersections in an interchange, interactive effects occur that complicate the analysis. The process provided in HCM for the analysis of the interchanges is mostly conceptual (*12*). For incorporating the various interactions that occur at an interchange and analyzing the performance, simulation processes are useful. ASCS provides a user-friendly interface and capability for analyzing diamond interchanges. The software also provides an extensible framework that can readily incorporate methodologies for analyzing various diamond interchange forms. The software should be extended in the future to allow estimation of the capacity of other forms of interchanges.

REFERENCES

1. Chaudhary, N. A., C. L. Chu, K. N. Balke, and V. G. Kovvali. *Coordination of Diamond Interchanges with Adjacent Traffic Signals.* Research Report 4913-1. Texas Transportation Institute, Texas A&M University System, College Station, Oct. 2000.
2. Fambro, D. B., and T. Urbanik II. In *Transportation Research Circular 430: Overview of Signalized Interchange Operations.* TRB, National Research Council, Washington, D.C., July 1994, pp. 16–23.
3. Venglar, S, P. Koonce, and T. Urbanik II. *PASSER III-98 Application and User's Guide.* Texas Transportation Institute, Texas A&M University System, College Station, 1998.
4. *Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 2000.
5. Chen H., H. Lieu, and A. Santiago. Interchange Operations: Software Evaluation Techniques in Use [with Discussion]. In *Transportation Research Circular 430: Interchange Operations on the Local Street Side: State of the Art.* TRB, National Research Council, Washington, D.C., July 1994, pp. 23–31.
6. Husch, D., and J. Albeck. *Synchro 4.0 User Guide.* Albany, Calif., 1999.
7. Park, B. *Development of Genetic Algorithm-Based Signal Optimization Program for Oversaturated Intersections.* Ph.D. dissertation. Texas A&M University, College Station, 1998.
8. Jain, S. *An Enhanced Model for Calculating Delay as a Function of Offset.* M.S. thesis. Texas A&M University, College Station, 1996.
9. May, A. *Traffic Flow Fundamentals.* Prentice Hall, Englewood Cliffs, N.J., 1990.
10. Hadi, M. A., and C. E. Wallace. TRANSYT-7F Simulation Models for Congested Conditions. Presented at 79th Annual Meeting of the Transportation Research Board, Washington, D.C., 1998.
11. Kovvali, V. G., and C. J. Messer. Sensitivity Analysis of Genetic Algorithm Parameters in Traffic Signal Optimization. Presented at 81st Annual Meeting of the Transportation Research Board, Washington, D.C., 2002.
12. Kovvali, V. G. *Development of a Robust Arterial Signal Coordination Software.* Ph.D. dissertation. Department of Civil Engineering, Texas A&M University, College Station, 2001.

Publication of this paper sponsored by Committee on Traffic Signal Systems.