Scheduling Buses to Take Advantage of Transit Signal Priority

Selman Z. Altun and Peter G. Furth

Transit signal priority can improve bus operations when it is applied to a route without making any changes to its route design or management; however, benefits can be greater if service design and management policies are purposely altered to take advantage of transit signal priority. Service design issues include generating carefully constructed schedules for use with conditional priority and deciding whether to locate stops on the near side or far side of intersections. Management issues include using conditional priority as a means to give priority only to late buses and deciding whether to hold buses at bus stops until a scheduled departure time, in case conditional priority itself does not offer the desired level of operational control. An optimal level of aggressiveness in the running time schedule is sought as a balance between mean running time with and without priority, assuming that priority will be conditional. With a more aggressive schedule, buses will be late and therefore will request (and get) priority more often; however, a more aggressive schedule also offers less slack to compensate for random delays. Analysis uses both a spreadsheet-based simulation model and a traffic microsimulation model that account for random delays at dispatch and at traffic signals, the effect of crowding on dwell time, and choice of priority and operational control tactics. It is found that optimal performance uses conditional priority and holding, but with an aggressive schedule, leading to substantial reductions in mean running time, standard deviation of running time, headway irregularity, and crowding.

Transit signal priority allows preferential treatment of buses at signalized intersections. It gives transit vehicles a little extra green time or a little less red time at traffic signals to reduce the time they are slowed by traffic signals (*1*). It also reduces the variability in running times by reducing signal delays. Variability in running times is a major challenge for transit authorities, as it not only causes overcrowding on buses but also increases the 90th percentile running time that is often used to determine allowed time (scheduled running plus recovery time), which in turn determines fleet size and thereby operating cost. A disturbance such as dispatching delay (starting late) or an excessive signal delay perceived to be small compared to the overall running time may cause buses to bunch after several stops. Signal priority helps reduce bunching by preventing excessive signal delays.

Conditional priority means giving priority to late buses while withholding priority from the early ones. Late vehicles that are granted priority get through intersections with little or no delay. Thus, they are sort of pushed ahead compared with the early buses that are likely to experience regular delays at the intersections. This creates a push–pull effect as Muller and Furth describe (*2*), making signal priority a powerful tool for operational control.

Conditional priority, if designed effectively, has advantages over absolute priority in several ways. First, there is less traffic disruption if conditional priority is used, since buses ask for priority only when they are late. Previous research showed that compared to absolute priority, conditional priority resulted in only minor impacts on cross traffic (*3, 4*).

Furthermore, although absolute priority helps somewhat to reduce running time variability by minimizing signal delays for all buses, conditional priority helps even more because it targets the problematic late bus and speeds it up compared to the (unaided) early bus.

CURRENT APPLICATIONS OF BUS SIGNAL PRIORITY

Numerous bus signal priority applications are being used in the United States. Although many older systems use absolute priority, more recent systems tend to use conditional priority by making use of smarter buses, signal controllers, or operating centers. Two typical installations are described here.

The city of Portland, Oregon, provides priority at 300+ intersections for buses operated by the Tri-County Metropolitan Transportation District of Oregon (TriMet), the local transit authority. An onboard automatic vehicle location system on each bus compares the real-time location of the bus with predetermined bus schedules along the route. If a bus is more than 30 s late according to the schedule, a signal priority request is transmitted by the bus, invoking a signal priority plan in the local signal controller. TriMet uses the same schedules for its priority system that were developed for service and advertised to the public. TriMet's practice is to set scheduled run times to about the 30th percentile running times, implying that about 70% of the buses will arrive late at a time point (*5*).

Los Angeles, California, runs a conditional priority system based on headways, not departure times. With the use of real-time headway information obtained from subsequent bus departures at intersections, if a bus is more than 1.5 scheduled headways behind its leader, it is given priority.

RESEARCH OBJECTIVES

This paper focuses on developing a methodology for generating optimal running time schedules for routes with conditional priority. It is expected that the optimal running time schedule will be a balance between two extremes: mean running time with absolute (uncondi-

S. Z. Altun, Kittelson & Associates, Inc., Suite 800, 33 North Stone Avenue, Tucson, AZ 85701. P. G. Furth, Department of Civil and Environmental Engineering, Northeastern University, 400 SN, 360 Huntington Avenue, Boston, MA 02115. Corresponding author: S. Z. Altun, selman.altun@aecom.com.

Transportation Research Record: Journal of the Transportation Research Board, No. 2111, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 50–59.
DOI: 10.3141/2111-07

tional) priority and mean running time without priority. It aims to demonstrate the significant impacts of having optimal schedules on the success of the overall bus priority system and provide a comparative benefit analysis of absolute and conditional priority. Finally, the research will try to find out the benefits of holding buses at bus stops until a scheduled departure time, in case conditional priority itself does not offer the desired level of operational control.

RESEARCH APPROACH

The research approach can be divided into two major parts. The first is development of a spreadsheet tool by using Monte Carlo simulation to test different scheduling options for conditional priority with and without holding and to test the performance of signal priority under the following scenarios:

- Availability of green extension and early green (red truncation) options with varying (aggressive versus mild) priority parameters;
- High, low, and zero dispatching irregularity (starting the run late or early); and
- Varying amount of green splits at the signals.

The second part is a test of various hypotheses, based on results obtained from the spreadsheet model, in a microsimulation environment developed in VISSIM for Avenida Ponce de León in San Juan, Puerto Rico.

SPREADSHEET MODEL

A simulation model of a hypothetical route was built in the spreadsheet environment to help analyze service performance by using different signal priority tactics. It accounts for dispatching delay irregularities at the beginning of the route, signal cycle lengths and green splits, and the effect of crowding effect on dwell time.

Base Case

The route simulated with the spreadsheet had 10 bus stops, nine signalized intersections, and nine road segments. For the base case, dispatching irregularity varied between half a minute late and half a minute early drawn from a uniform distribution. Bus headway was 5 min.

At each stop and for each bus, the spreadsheet calculates the arrival headway (the time difference between bus arrival times at stops). The expected number of boarding passengers is calculated by multiplying arrival headway by the passenger arrival rate. Passenger arrivals were then drawn from a Poisson distribution.

The number of passengers alighting at each stop was taken as a percentage of the arrival load, varying by stop. Table 1 shows the expected bus load profile along the route (per bus, assuming a 5-min headway). The load profile is fairly typical with transit centers at both ends and with regular boardings and alightings along the route. The peak load is expected to be 40 passengers. A bus was not allowed to depart from the bus stop earlier than its predecessor (no overtaking).

Road segment running time was assumed to be normally distributed with an expected running time of 2 min with a standard deviation of only 10 s. This way, the focus could be kept on the running time disturbance that stems from the effects of traffic signals.

Each of the nine intersections had the same cycle length, 100 s, and the same fraction green time, 50%. The saturation flow rate was assumed to be 1,600 vehicles per hour. Volumes on the bus street were set to 680 vehicles per hour, so that volume-to-capacity ratio was 0.85.

Two types of priority strategy are used for the spreadsheet model:

- Green extension (represented as X):
  - Aggressive: maximum green extension = 25 s and
  - Mild: maximum green extension = 10 s and
- Early green (also known as red truncation; represented as e):
  - Aggressive: as early as possible after a 15-s minimum red criteria is satisfied (e.g., 35 s early for a 50-s red period) and
  - Mild: 10 s early.

Calculation of Signal Delays: Base Case

Control delay at each intersection is modeled by assuming uniform vehicle arrivals and uniform saturation flow. Delay experienced by buses at an intersection is calculated as a function of arrival time at the intersection during the cycle. Bus arrival time within a given signal cycle was selected at random by drawing a uniform random variable r on the range (0,1) and multiplying it by the cycle length C. Figure 1 shows the bus delay with only green extension as a function of the bus arrival time within the cycle and the maximum allowed extension x. Note that the colors shown in Figures 1 and 2 (red, orange, and green) do not correspond to the actual signal timing indications of red, yellow, and green and are used merely to differentiate three different intervals within a signal cycle in which a bus arrival can occur.

As shown in Figure 1, a bus arriving near the start of red normally would have to wait for the entire red period to get the green light. However, as a result of green extension, the bus delay within this region shown in red (from 0 to x s) is zero. In the orange region, the bus experiences the regular signal delay it would normally experience, because it is detected during red and no green extension is possible. The green region is the unsaturated flow period; a bus arriving in this period has no delay and needs no priority treatment.

The delay distribution with both green extension and early green is shown in Figure 2. Delay with no priority is also included in the graph for comparison. In the red and green regions, buses do not experience any delay. In the orange region, the delay experienced is shown by the equation given.

TABLE 1 Expectations by Stop

	Stop #									
	1	2	3	4	5	6	7	8	9	10
Boardings	30	9	9	9	9	9	9	9	9	0
Alightings	0	6	7	7	7	8	12	11	10	34
Departing load (passengers)	30	33	35	37	39	40	37	35	34	0

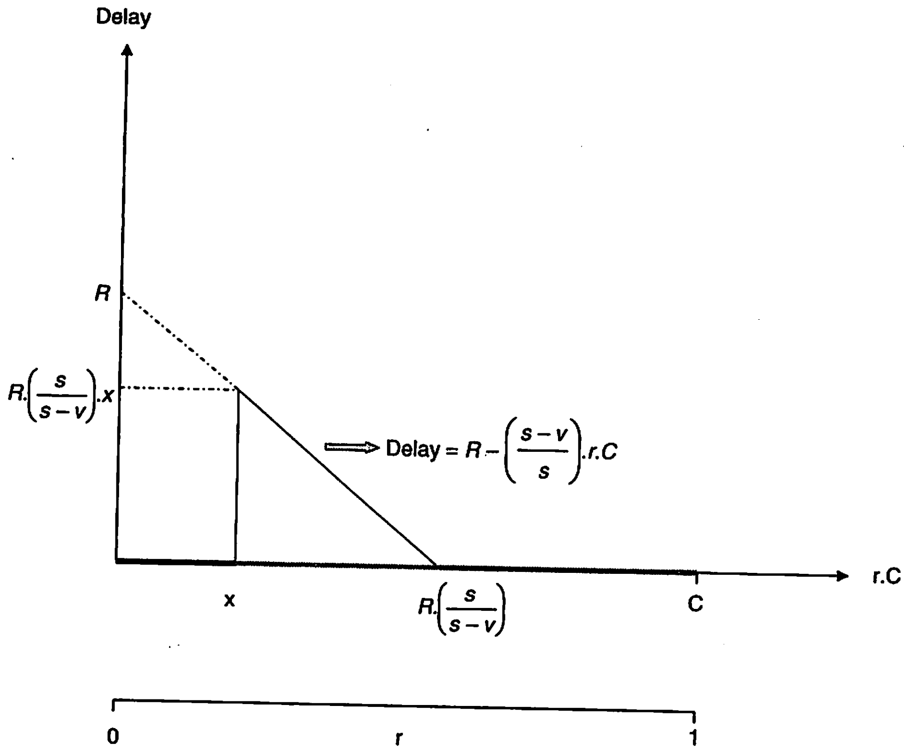

FIGURE 1 Delay with green extension as function of arrival time in cycle, *r.C*.

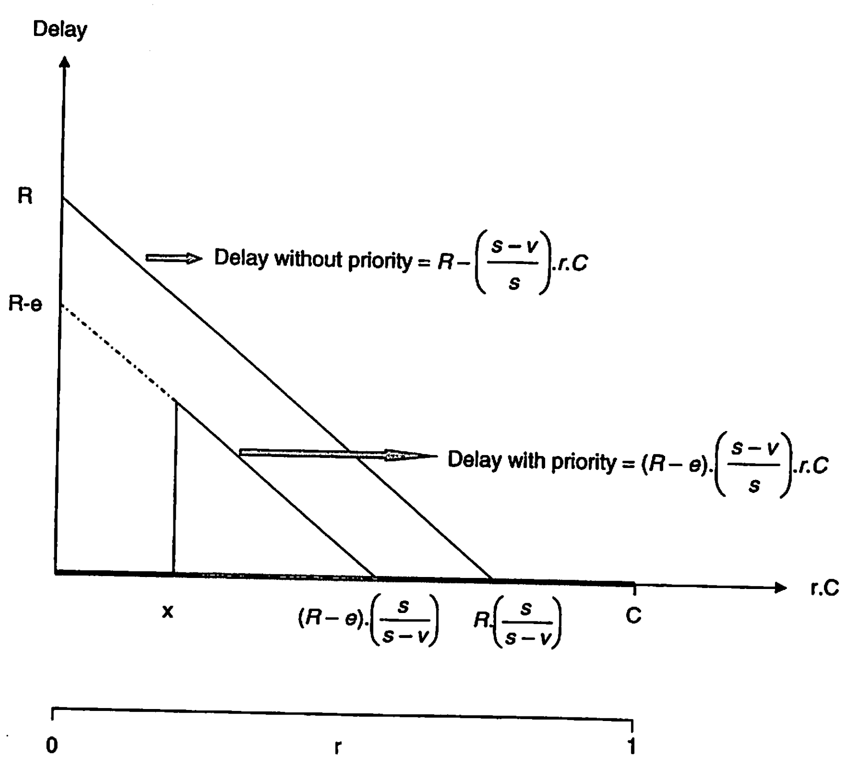

FIGURE 2 Delay with both green extension and early green as function of arrival time in cycle.

Crowding Effect

Once buses reach a certain level of crowding, the time passengers take to board and alight the bus increases. Therefore, as the buses get more crowded, expected dwell time at the stops increases. For the purposes of the spreadsheet model, 3 s per boarding passenger is assumed when the occupancy on the bus is fewer than 40 passengers. Unit boarding time is assumed to increase linearly from 3 s to 6 s as occupancy grows from 40 to 80 passengers.

Likewise, 2 s per alighting passenger is assumed when the bus occupancy is fewer than 40 passengers. Unit alighting time is assumed to increase linearly to from 2 s to 4 s as occupancy grows from 40 to 80 passengers. For unit alighting time, occupancy is simply the total number of people on the bus arriving at the bus stop. For unit boarding time, occupancy was taken to be the continuing load (arriving load minus offs) plus half the number of boarding passengers.

Selecting Running Time Schedules

In the spreadsheet model, running time schedules are calculated by subtracting a certain multiple of the priority push per intersection from no-priority schedules, where priority push equals the expected running time difference between no priority and absolute priority divided by number of intersections. That way, the running time schedule will be a weighted average of the no-priority running time and the absolute-priority running time.

Monte Carlo Simulation Results

The spreadsheet model was used to test several scheduling options. For operations that use conditional priority as a tool for operational control, a good running time schedule is expected to lie between mean running time with no priority and the mean running time with absolute priority; a running time schedule between these two extremes offers the greatest possibility for conditional priority to exercise push–pull control.

Figure 3 shows the mean running time values and the standard deviation of running times of cases with no priority, absolute priority, and conditional priority based on Monte Carlo simulation using the spreadsheet model. Mean plus one standard deviation of running is the chosen measure of performance because it is closely related to allowed time (and therefore operating cost) and because passengers value reliability as well as speed. In the absolute and no-priority cases, scheduled running times are irrelevant. With conditional priority, several scheduled running time options were tested, expressed as the mean running time with absolute priority plus a stated percentage of the difference between the mean running time with no priority and with absolute priority. Thus, for example, Cond 0% uses mean running time with absolute priority as its scheduled running time, and Cond +100% uses mean running time without priority for scheduled running time. Compared to the base case, the priority scenarios show a significant reduction in mean running times. In the case of absolute priority, running time decreases by 3.5 min.

Despite this reduction in running time, the standard deviation for the absolute priority case is almost as high as when there is no priority. As the figure demonstrates, conditional priority helps reduce standard deviation under all scenarios, although the standard deviation is still quite large for the case Cond −150% because the schedule is so tight (tighter than mean running time with absolute priority) that buses nearly always request priority, and so its results differ little from the absolute priority case. The case with the smaller standard deviation is Cond +50%, meaning conditional priority using for scheduled running time halfway between the means of no-priority and absolute priority.

The priority and scheduling options that yield the smallest sum of mean-plus-standard-deviation of running time use conditional priority with a schedule either equal to the mean absolute priority running time or slightly tighter (the Cond −50% case). The former option is preferred because, having a less tight schedule, it generates 14% fewer priority requests at intersections, thus disrupting traffic less. This finding suggests that aggressive running time schedules serve transit better when conditional priority is available.

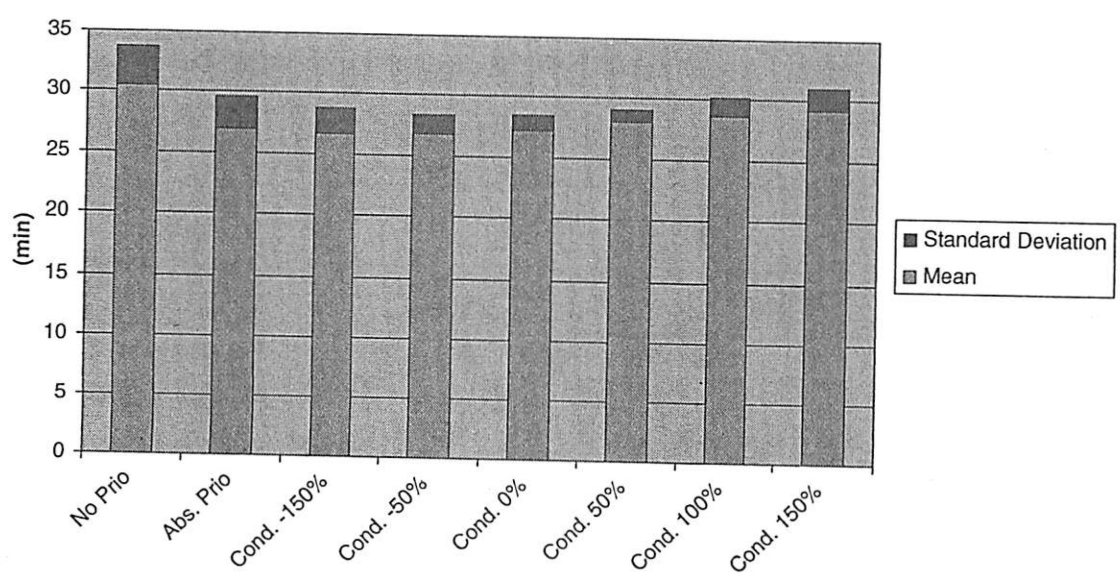

FIGURE 3 Mean plus one standard deviation running time values with different priority and scheduling options (no prio = no priority, abs. prio = absolute priority, cond. = conditional priority).

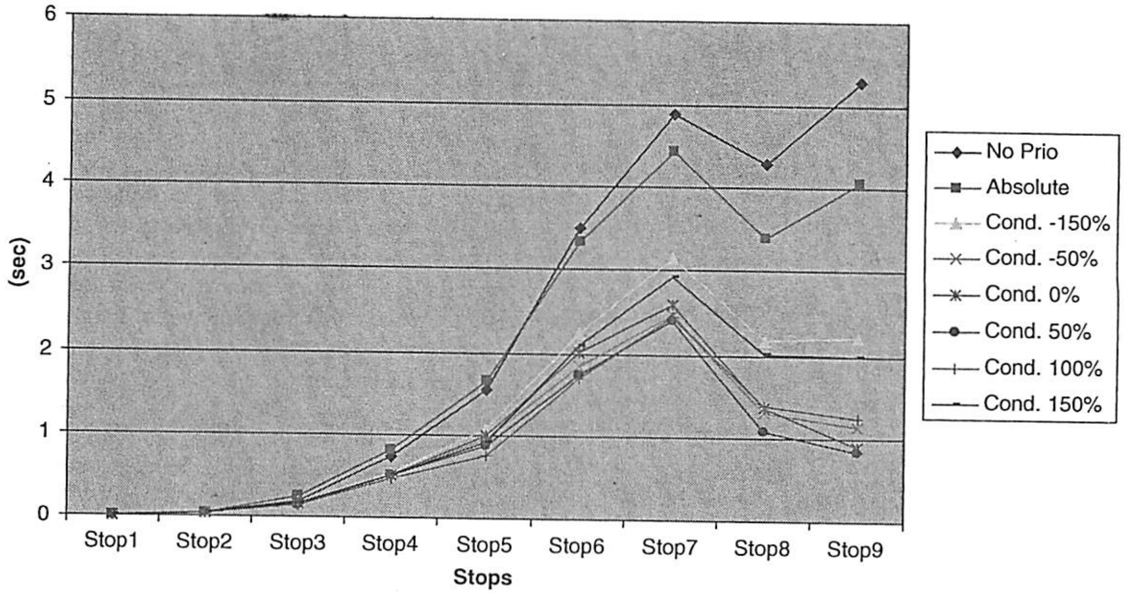

FIGURE 4 Crowding penalty per bus by stop.

Crowding Penalty

"Crowding penalty" is defined as the additional dwell time at a bus stop because of overcrowding, calculated by subtracting dwell time from what it would be without the incremental boarding and alighting time due to overcrowding. Figure 4 shows the crowding penalty per bus at each stop for the different scheduling and priority cases. Although the crowding penalty is rather small overall, ranging from 0 to 6 s per stop, Figure 4 nevertheless clearly demonstrates the advantage of conditional priority, which has smaller crowding penalties than the no-priority and absolute-priority scenarios. There is also a positive correlation between the crowding penalty and the standard deviation of running times.

Conditional Priority and Holding

The benefit of holding buses at bus stops until the scheduled departure time in conjunction with conditional priority was tested as a strategy to improve the level of operational control. Like conditional priority, holding depends on the running time schedule. Figure 5 shows the effects of conditional priority with holding, with the scheduled running time expressed as mean running time under absolute priority plus a multiple of its standard deviation.

When holding is applied with a multiple of 2, results are almost the same as without any holding, because conditional priority has almost the same effect as holding, slowing down buses if they begin to run early. The best results—low mean running time and very low

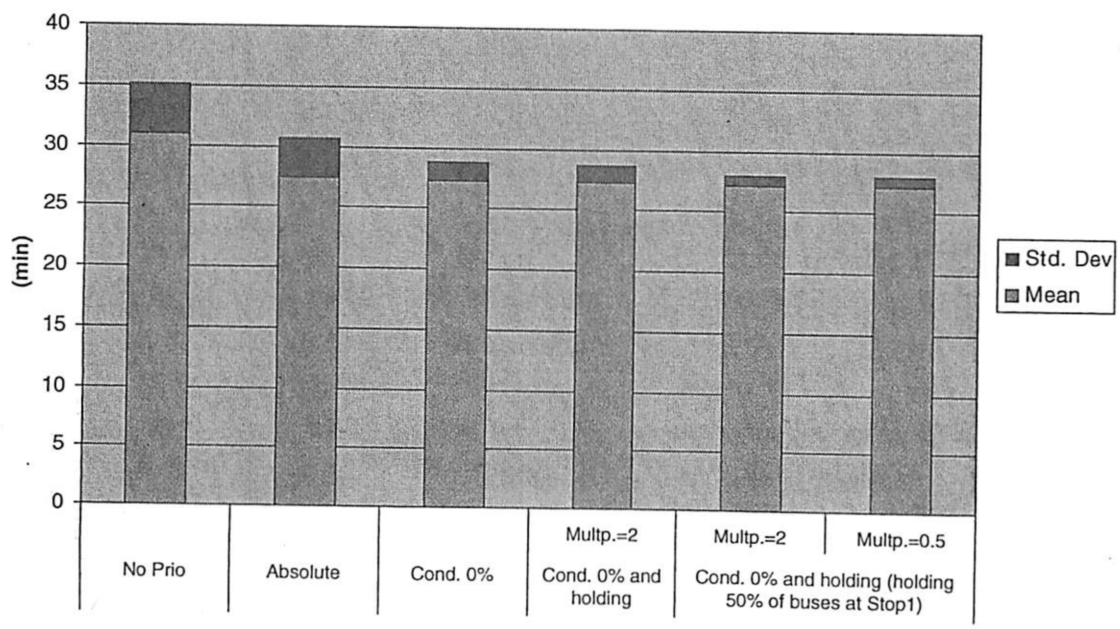

FIGURE 5 Effect of holding on running times (multp. = multiple of SD).

standard deviation—are obtained when the holding schedule for the first stop is the mean departure time and the holding schedule at remaining stops is mean departure time plus a multiple of 2 or 0.5 times the standard deviation of running time for the remainder of the route. This shows that aggressive holding is most beneficial at the beginning of a route; later on, conditional priority and holding only the very early buses is enough to achieve a very regular service.

Impact of Different Degrees of Priority

The simulation model was also used to explore the impact of mild (M) and aggressive (A) priority (defined earlier) by using green extension and early green. Green extension offers a relatively large delay reduction, but only to the small fraction of buses that arrive at an intersection near the end of green. Early green serves all the buses that are caught at red, offering a small benefit to many buses, but with greater disruption to conflicting traffic.

Results, shown in Figure 6, show that mean and standard deviation of running time are both minimized when both aggressive green extension and early green are used. Use of solely mild green extension, a common strategy, was found to be almost as ineffective as having no priority at all. Figure 6 also shows the relatively greater incremental value of early green compared to green extension.

Impact of Dispatching Delay

In the base case, dispatching irregularity is uniform between 30 seconds late and 30 seconds early. In case of high dispatching delay, where the irregularity band is increased to ±1 minute, Figure 7 shows how the mean and standard deviation of running time increase substantially. Because dispatching irregularity cannot be corrected with absolute priority, the standard deviation stays almost as large as the case with no priority. In contrast, conditional priority decreases the mean plus standard deviation running time by more than 10 minutes.

Impact of Different Green Splits

Figure 8 shows that the running time benefits of priority are more than twice as large when the bus movement has a split of 30% green, 70% red, compared to a favorable (70% green, 30% red). When splits are less favorable, buses are more likely to be delayed by traffic signals, and delays are longer, giving priority more of an impact when applied.

Summary of Results from Spreadsheet Simulation Analysis

Conditional priority with aggressive priority tactics and with scheduled running time based on the mean running time using absolute priority proves to be the most efficient combination. It reduces the mean plus one standard deviation of running times value to a low level, while requesting priority 72% of the time. This scheduling option is chosen to be carried forward and used in a microsimulation model of a real street.

MICROSCOPIC SIMULATION TOOL

Avenida Ponce de León in San Juan, Puerto Rico, is simulated by using VISSIM, a microscopic time step and behavior-based simulation software developed to model urban traffic and public transit operations. It has been applied in different projects, and studies have shown that it is equally capable of simulating traffic as other software packages, such as CORSIM (*6*).

VISSIM was chosen because of its ability to model transit operations, including tracking the vehicle load, and its flexibility for signal control logic (*7*). Its companion programming language, VAP, allows user-defined signal control logic, which can be manipulated to model the overcrowding effect as well. The simulation model development and calibration took significant time and effort. The simulation was used as a test bed to confirm the findings of the spreadsheet

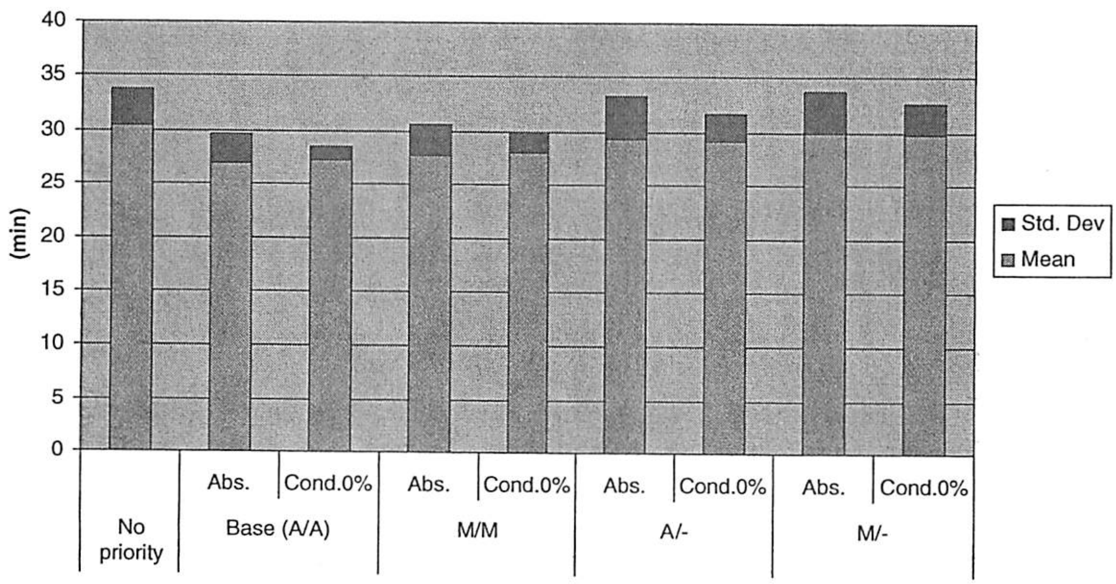

FIGURE 6 Running-time results for different degrees of priority. First indicator is for green extension, second for early green (M = mild, A = aggressive, – = none).

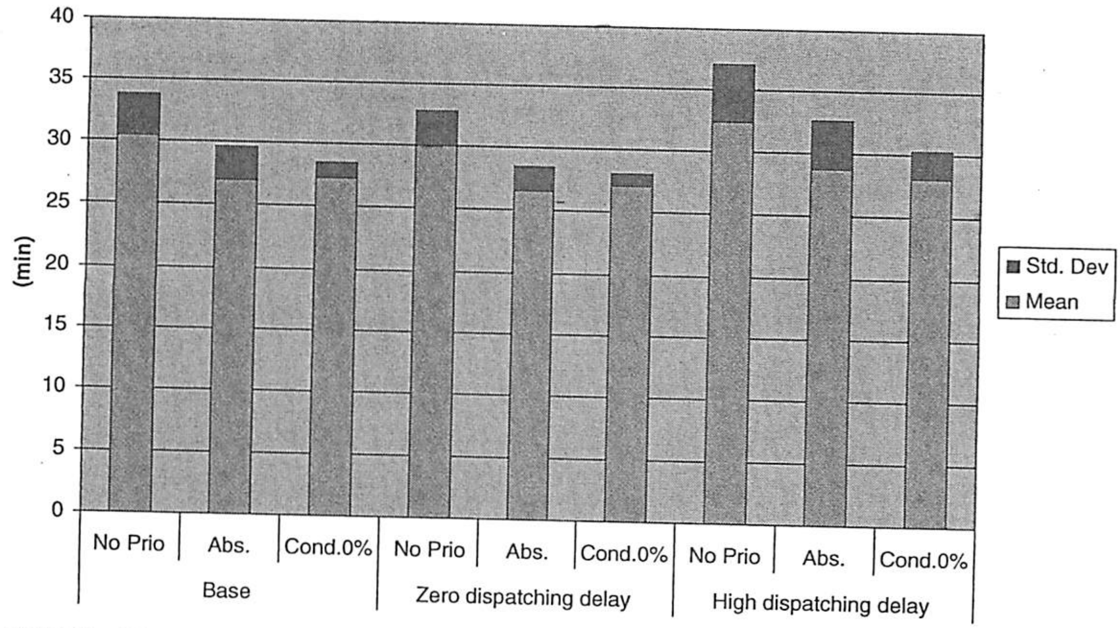

FIGURE 7 Effect of dispatching delay on running times.

simulation model and should be considered generic for the purposes of this paper. Complete details of the VISSIM model can be found in the paper by Janos and Furth (*8*).

TRITAPT

TRITAPT (trip-time analysis in public transport) is an application for analyzing data collected by onboard computers of public transit vehicles (*9*). TRITAPT can be used to find locations on the routes where delays occur and quantify the severity of the delays. It can also be used to design optimal timetables.

VISSIM outputs were modified to emulate data from an automatic vehicle location system and then fed into TRITAPT. The conversion tool was created by Peter Knoppers of Delft University of Technology, Transportation Planning and Traffic Engineering Section.

Base Case Results

Without priority, microsimulation analysis resulted in an average running time of 36 min along the Avenida Ponce de León corridor. Simulation results closely matched real-world observations, with traffic at a volume-to-capacity ratio close to 1.0 (highly congested),

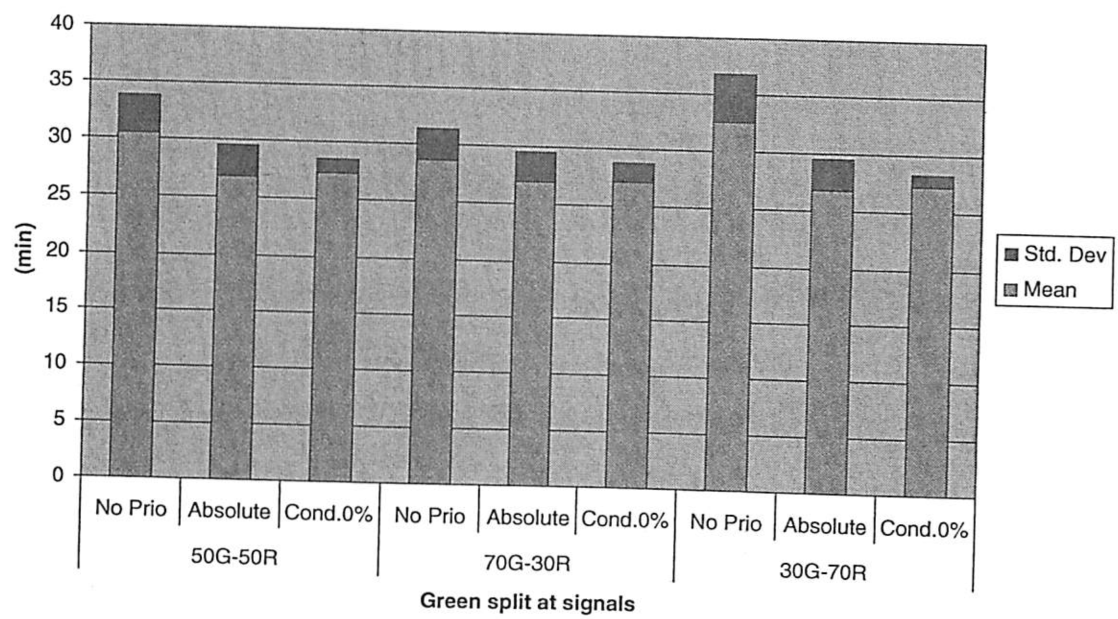

FIGURE 8 Effect of different green splits on running times.

considerable delays for buses, and a large and progressive (worsening along the route) variation in running time.

Running Time Spread with Absolute Priority

Figure 9 shows each bus trip's cumulative running times from the start of the route until departure from each stop when operating with absolute priority. The blue stars indicate the average running time to for each stop. For a perfectly regular service, all the trips should be passing through the blue stars.

With absolute priority, the mean running time is approximately 31.7 min, a 4.3-min reduction compared to the base case. However, the individual trip times vary considerably from this mean.

Running Time Spread with Conditional Priority

Conditional priority gives almost an identical mean running time value compared to the case with absolute priority; however, the band of variation of running times is smaller and more trips are clustered around the mean, as shown in Figure 10. On the other hand, there remain many early trips that could not be helped only with conditional priority.

Running Time Spread with Conditional Priority and Holding

Figure 11 shows that holding improves on-time performance even more by pulling the earlier trips close to the mean. While improving schedule adherence significantly, conditional priority with holding increases the mean running time by only about 20 s, to 32 min. The clustering around mean running time is especially strong in this case, not only because early buses have largely disappeared, but because the elimination of early buses has helped limit the development of late buses due to the bunching effect.

SUMMARY AND CONCLUSIONS

This paper explored how to design running time schedules and manage transit operations to take advantage of transit signal priority. It developed a methodology for generating optimal schedules for conditional priority as a weighted average of mean running time with and without absolute priority. It demonstrated that the effectiveness of conditional priority depends on carefully constructed schedules. Schedules close to the mean of running times with absolute priority are found to be most effective. The model predicts substantial reductions in mean running time, standard deviation of running time, headway irregularity, and crowding. The results indicate that conditional priority, if designed and managed effectively, significantly improves service reliability. Finally, the results show that holding buses at bus stops until a scheduled departure time furthers reliability improvement, although the effectiveness of holding, like that of conditional priority, depends on having a well-tuned schedule. The benefit of holding is greatest when applied early in the route with an aggressive running time schedule.

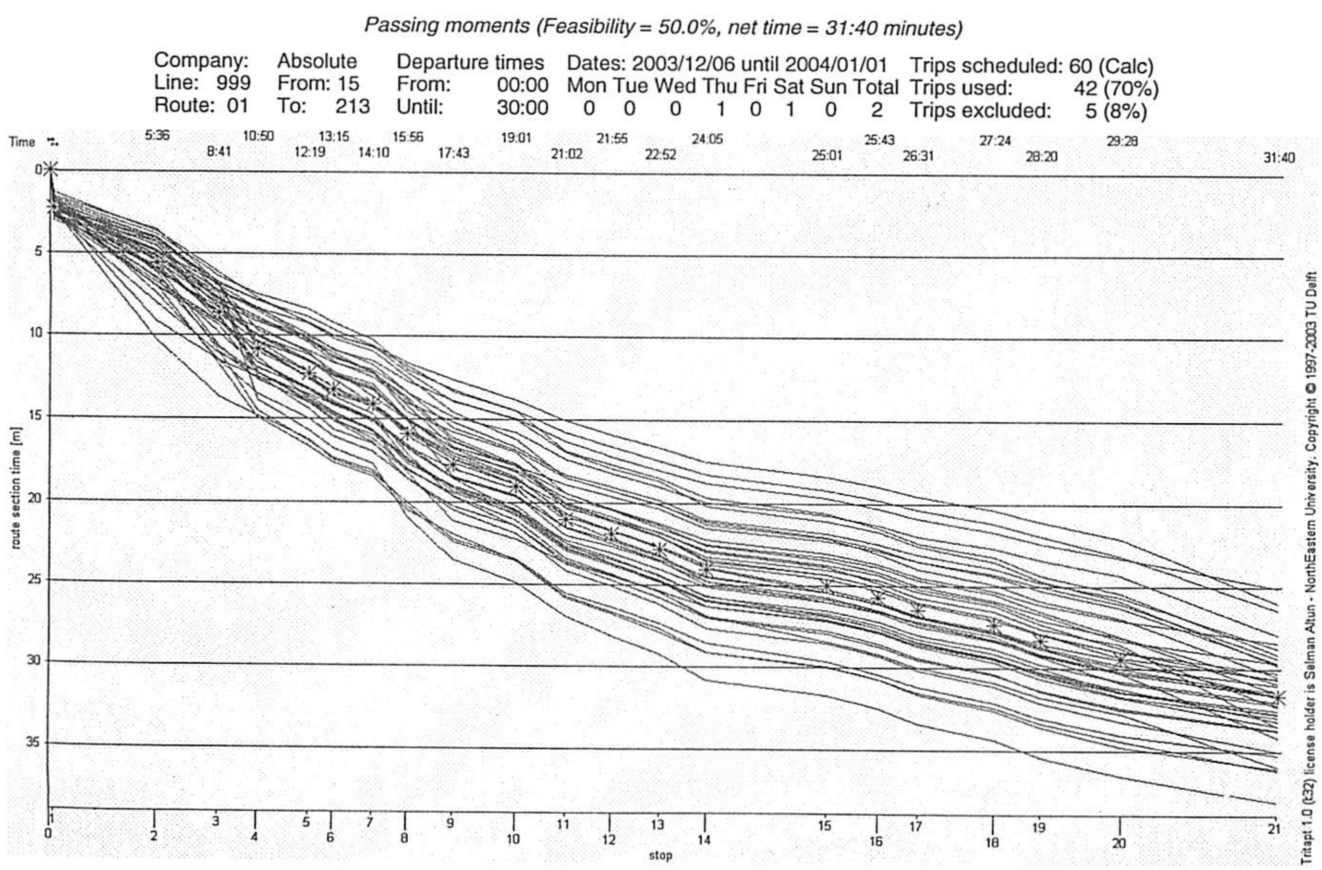

FIGURE 9 Cumulative running time with absolute priority.

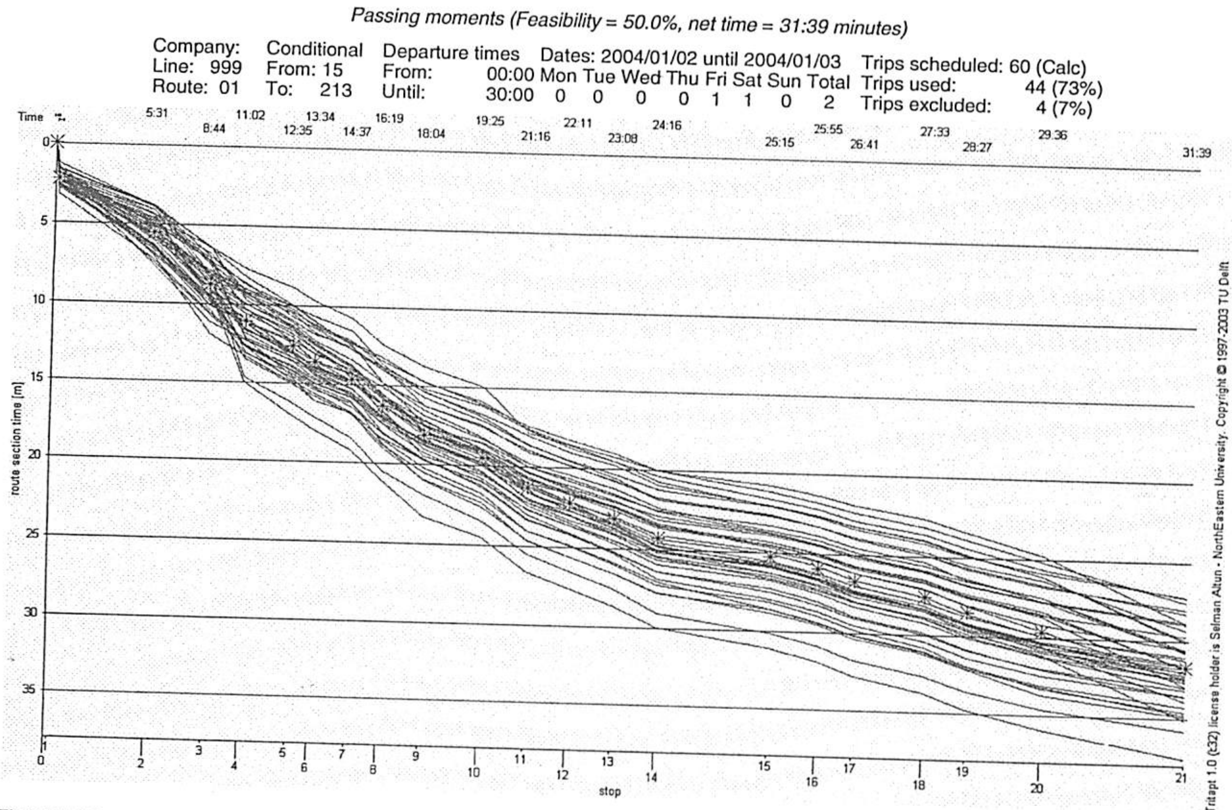

FIGURE 10 Cumulative running time with conditional priority.

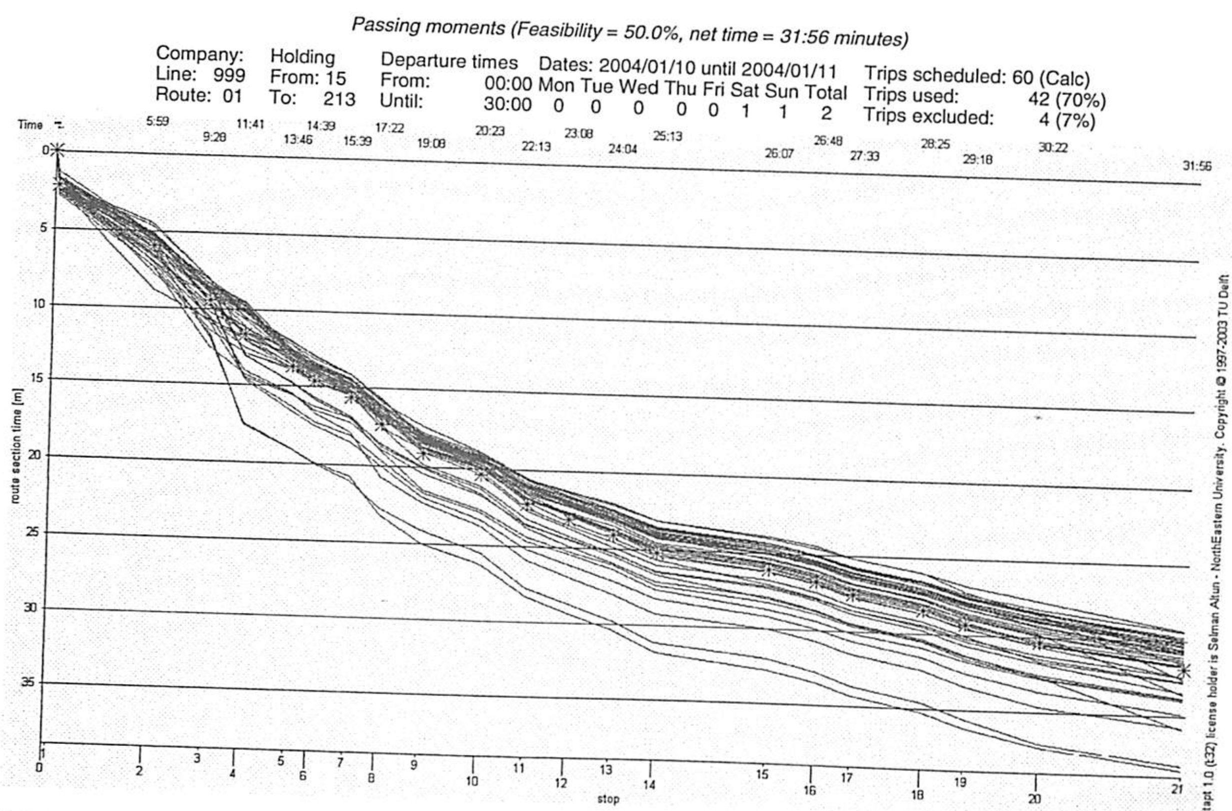

FIGURE 11 Cumulative running time with conditional priority and holding.

ACKNOWLEDGMENTS

This research was supported by the Puerto Rico Highway and Transportation Authority through the Tren Urbano Technology Transfer program. The authors thank Peter Knoppers of the Delft University of Technology, Transportation Planning and Traffic Engineering Section, for his contributions to analyzing microsimulation data and thank Peter J. Koonce of Kittelson & Associates for his review and comments during the development of this paper.

REFERENCES

1. *Transit Signal Priority (TSP): A Planning and Implementation Handbook.* ITS America, Washington, D.C., May 2005.
2. Muller, Th. H. J., and P. G. Furth. Integrating Bus Service Planning with Analysis, Operational Control, and Performance Monitoring. Presented at Annual Conference of the Intelligent Transportation Society of America, May 2000.
3. 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.
4. *An Overview of Transit Signal Priority.* Advanced Traffic Management Systems Committee and Advanced Public Transportation Systems Committee, Intelligent Transportation Society of America, July 2002.
5. Callas, S. *Tri-Met's Transit Signal Priority System and Evaluation.* TriMet, Portland, Ore., 2002.
6. Bloomberg, L., and J. Dale. 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.
7. *VISSIM User Manual: Version 3.60.* PTV, Karlsruhe, Germany, 2001.
8. Janos, M., and P. G. Furth. Bus Priority with Highly Interruptible Traffic Signal Control: Simulation of San Juan's Avenida Ponce de Leon. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1811,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 157–165.
9. TRITAPT. cttrailf.ct.tudelft.nl/verkeerskunde/tritapt/.

The Bus Transit Systems Committee sponsored publication of this paper.