# Evaluation and Improvement of Traffic Signal Settings by Simulation

STEPHEN B. MILLER, Aerospace Corporation, El Segundo, California, and
JOHN D. C. LITTLE, Operations Research Center,
Massachusetts Institute of Technology

Several methods of setting traffic signals are studied by means of traffic simulation. A test problem, consisting of a six-signal network in Boston, is used to compare three methods: (a) maximal bandwidth progression, (b) "simswitch" (in which the center of red is simultaneous along all parallel streets), and (c) random search (choosing the best of a number of randomly selected settings). Performance is measured as a weighted combination of average delay per vehicle and average number of stops per vehicle. Two systematic procedures for improving a given setting are explored. Both involve one-variable-at-a-time search in the neighborhood of the original setting. In the first, the absolute offset of an individual signal is the changed variable; in the second, the relative offset between a pair of adjacent signals is changed.

The conclusions are reached that (a) operationally significant differences exist among settings that a priori might be expected to be good ones; (b) the criterion adopted for evaluating performance substantially affects which setting will be chosen; (c) many traffic situations do not conform to the simplest models and it is difficult to predict good settings without detailed examination of traffic movement as through simulation; (d) an effective way to obtain good settings through simulation may be to test out several settings considered in advance to be good and then improve the best one or two by systematic search; (e) performance is flow-dependent, i. e., a setting good at one level may be poor at another even though the patterns of flow are similar; and (f) settings with progressive timing seem more likely to be degraded in performance by turning vehicles than are systems with simultaneous switching.

•THE PROBLEM of traffic congestion in urban areas shows little sign of early abatement or easy solution. To a considerable degree this is because improved facilities generate increased use, but as long as benefits exceed total cost, the gains can be welcomed. Traffic congestion can be reduced by building new roads, by improving public transit facilities, or by improving the utilization of the current road system. The present paper concentrates on this last possibility. Even small percentage improvements in existing systems can have substantial significance when one considers alternative ways to achieve the same effect through new facilities.

In trying to improve traffic operations, we are concerned with the driver's convenience and safety. Convenience is represented by few stops, little wait when stopped, short trip times, and, as a means to an end, by high flow rates in the system. Safety

Paper sponsored by Committee on Traffic Control Devices and presented at the 46th Annual Meeting.

is affected by many factors, including a number of characteristics of traffic flow. Several studies (1, 2, 3) have shown that accident rates increase with traffic flow and congestion and generally with the demands imposed on the driver.

Traffic signals usually do not by themselves reduce accidents at an intersection (4). In fact, certain types of accidents, e. g., rear-end collisions, are increased. However, studies have shown that fatal accidents usually decrease when signals are installed and also that the accident rate for high vehicular flow is less with signals than without them (5). Since traffic signals are obviously necessary at many intersections, it is reasonable to ask if some signal settings result in fewer accidents than others. Studies (1, 3) have shown that rear-end collision rate is proportional to the number of stops required. A successful progression system will reduce stops and is therefore likely to reduce the rear-end collision rate. Two studies (6, 7) have shown that a high quality of flow and a low accident rate go together. High quality of flow is associated with a high average speed, infrequent changes of speed, and, when changes occur, changes of moderate size. These flow characteristics are closely related to the driver's desired traffic conditions so that improvements in signal settings from the point of view of driver convenience may frequently reduce accidents as well.

Perhaps the ideal way to set traffic signals would be to instrument a street network so as to measure stops, waits and trip times continuously and then, by extensive experimentation, find the control procedure that maximizes the effectiveness of the system as a whole. Some modern traffic control systems with detectors in the streets and computer control of the signals begin to approach the desired instrumentation (8, 9). However, even where good instrumentation exists, it is doubtful that on-street experimentation could search through the tremendous number of control possibilities without some sort of off-street theory to identify the relevant choices.

From the standpoint of realism the next best thing to instrumentation would be an accurate simulation of traffic movement on the street network. With simulation, traffic control systems can be evaluated on a computer instead of on the street. As simulated cars pass through a street network, their stops and delays are easily recorded for evaluation purposes. However, simulations frequently consume large amounts of computation time. Efficient search methods for evaluating the many possible control procedures need to be devised. Some possible methods are investigated in this paper.

Another class of models used to determine traffic signal settings is aimed specifically toward optimization. Detail in the traffic flow assumptions is sacrificed to make possible efficient search for the best setting. The model and the criteria to be optimized are usually fairly simple and are selected with the objective of capturing the essential features of the problem. An example of this approach is the maximal bandwidth algorithms of Morgan and Little (10) and Little (11). Such models permit finding exact optima, but their traffic assumptions and criteria must be examined for relevance in the particular situation. The value of these solutions can be examined by observing them on the street, or by testing them in a simulation. Examples of the latter are given here.

The decreasing cost of computation coupled with the importance of the traffic control problem has brought about computer-controlled traffic systems in a number of cities (8, 9). Such systems permit real-time control on both an individual intersection and area-wide basis; i. e., vehicle detectors in the streets provide information with which to set the signals more or less immediately and continuously. In the present paper we shall not explore the potentialities of this type of control but rather will concentrate on fixed-time systems. The reasons are twofold. In the first place many cities, by virtue of size, traffic conditions, or financial conditions, will not justify computer-controlled systems very soon. Secondly, computer-controlled systems presently work to a considerable extent from tables of fixed-time settings, an appropriate table being selected by the computer according to traffic conditions.

## A SIMULATION MODEL

Many different levels of detail are possible in a simulation of traffic. For computational efficiency, one seeks the least detail that will still reproduce the most

important features of traffic. Just what model does this is probably yet to be determined. Among the simulation models so far reported are those of Gerlough and Wagner (12), Katz (13), Kell (14), and Blum (15). The model here is that of Schwartz (16), modified slightly. The model is a fairly simple one. Its general features are outlined below. A more detailed description and a discussion of model predictions versus observations may be found in the original report (16).

In Schwartz's simulation each vehicle that enters the system is modeled explicitly. Each is generated at some point on the boundary of the network. There is some flexibility in the method of generating the vehicles. The examples here use either exponentially distributed interarrival times between vehicles or constant interarrival times, but other schemes, including the arrival of platoons, can be used.

As soon as a vehicle enters the network and each time a vehicle reaches an intersection, a probabilistic decision is made to determine whether the vehicle will go straight ahead or turn right or left. The values of these probabilities are determined from field observations on the actual network. After leaving an intersection, a vehicle is held in a delay "store" for the amount of time required for it to move to the next intersection. This time is determined by the vehicle's speed and the distance to be traveled. When the vehicle reaches the intersection, it chooses its next link and, if there is no queue and the signal is green, goes through. If there is a queue, the vehicle takes a place at the rear of the queue. If the signal is red, the vehicle waits at the intersection or at the end of the queue, as the case may be. In order to simulate the startup delay that occurs just after a signal turns green, the program retards the advance of any waiting queue for two seconds.

In practice the simulation program models a single generalized intersection and a single generalized link. (A link is defined as a street segment between two adjacent intersections along with a direction of travel on the segment.) Parameters and state variables for individual intersections are entered in tables. State variables are changed and updated for each link, signal light and intersection in each elemental time period (one second in the work here).

In Schwartz's original program, the time to traverse a link was a constant for the link. This means there is no diffusion of traffic platoons or disruption of smooth flow. In practice, as vehicles move away from a recently changed signal, the vehicles in front usually pull away from the bulk of the platoon. Cars toward the rear lag behind. A simulation model that ignored diffusion may make certain ways of setting the signals, e.g., maximal bandwidth methods, look better than they actually are. Accordingly, it was desired to modify the model to include this phenomenon.

One simple way to model diffusion is to let vehicle speeds, instead of being constant, follow a normal probability distribution (17, 18). Although this implicitly assumes free passing and a speed independent of position in the platoon, actual traffic observations show reasonable agreement between model and fact except in conditions of heavy flow. The coefficient of variation (the standard deviation of speeds divided by mean speed) is almost constant at about 0.15 for moderate flow rates. In the modified version of the program the speed of a vehicle on a link is chosen from a normal distribution with the above coefficient of variation and a mean speed based on observation of traffic on the link.

The output of the simulation program includes the total number of vehicles that have entered and left the network, total number of stops, and the total delay for all vehicles that have left the network. Delay to a vehicle on a link is calculated as the time spent on the link over and above the time that would be required to travel the link at the vehicle's chosen speed. In addition the number of cars to enter each link and the average time these cars spend on the link is available. Other measures of performance can be made available.

The running times of the simulation depend on the number of signals and on the input flows. Schwartz's runs on a 26-signal network were about one-half of real time on an IBM 7094. Our runs on a 6-signal network were about one-sixth of real time on the same computer. Quite likely a considerably more efficient program could be written if the effort were devoted to it.

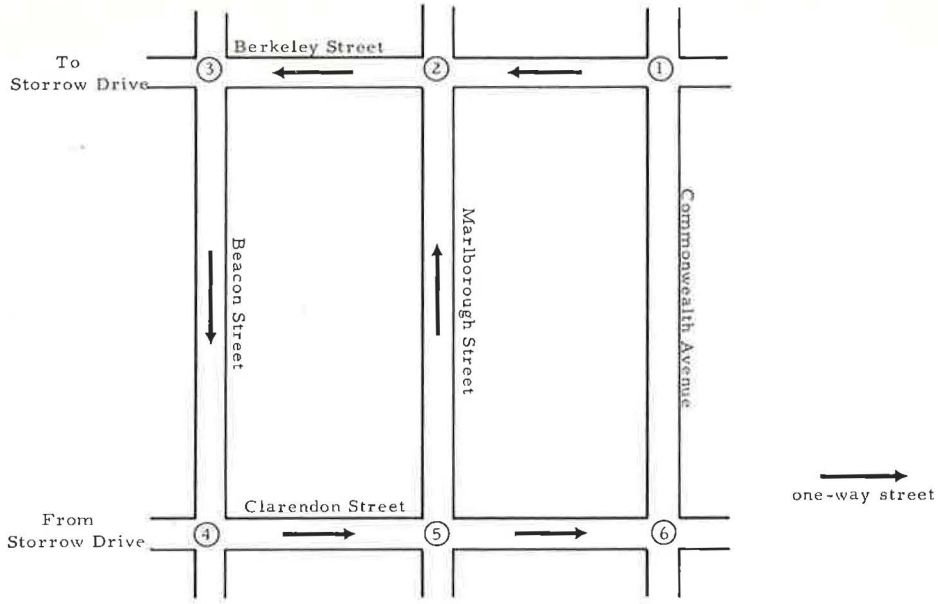


Figure 1. Section of Boston Back Bay used as a test problem (signals are numbered as in text).

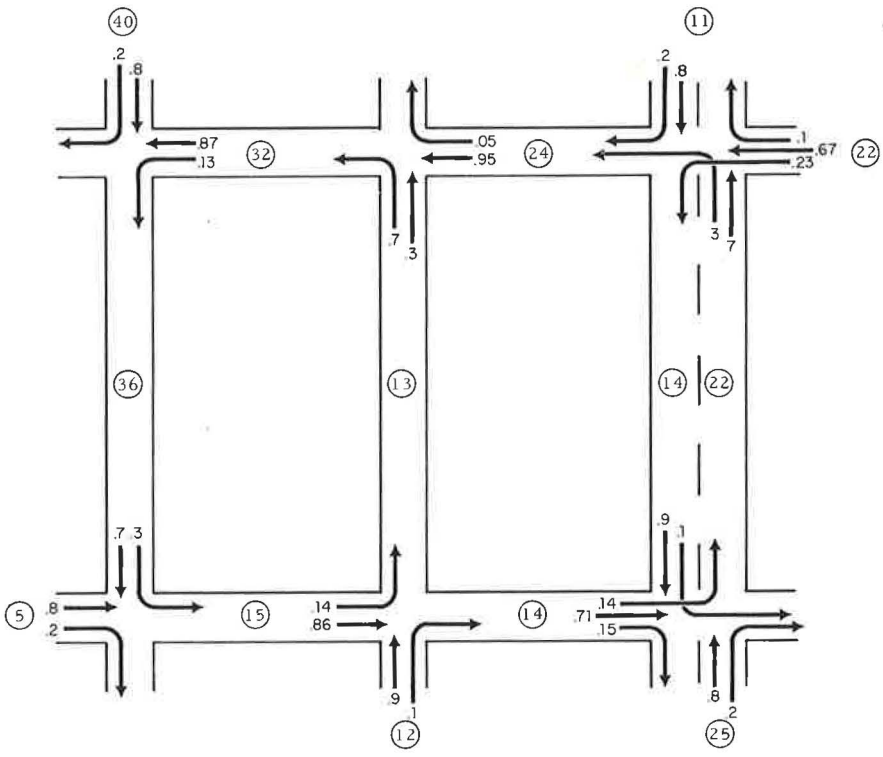


Figure 2. Network parameters for medium traffic flow. Flows are average vehicles/cycle as observed on a weekday early afternoon. Cycle length was 116 seconds.

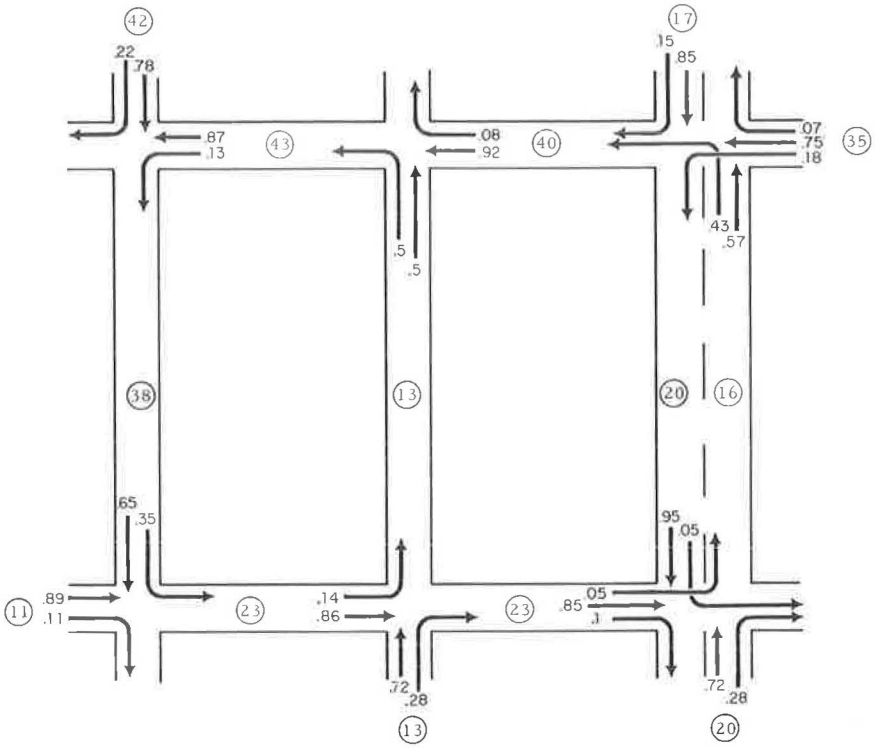


Figure 3. Network parameters for heavy traffic flow. Flows are average vehicles/cycles as observed during a weekday evening rush hour. Cycle length was 105 seconds.

## CRITERIA

In order to evaluate a signal setting, some measure of performance, or criterion, is necessary. There are a number of possibilities; stops, wait while stopped, trip delay, and safety are examples. Helly (19) has discussed "acceleration noise," a measure of unevenness in speed. Another possibility is throughput, particularly throughput at peak periods when congested streets represent a reservoir of unsatisfied demand. Certainly any change that will increase flow at such times is highly desirable. However, any increase in throughput will immediately be reflected in driver-oriented criteria and so these are what we shall use.

The driver's tradeoffs among criteria have not, to our knowledge, been probed in an operational way. Some system changes will improve performance by several criteria. However, as will be demonstrated, it is possible to find examples where number of stops goes up and delay goes down and vice versa. Thus, we really would like to know the driver's tradeoffs. Since these are lacking, we shall do two things: in some cases we shall display multiple criteria, in others we shall make reasonable but arbitrary combinations of criteria, focusing on stops and trip delay.

The model calculates stops under the assumption that the driver, when obstructed by a red light or a queue, maintains his desired speed until the last moment and then stops abruptly. In practice, of course, drivers slow down if they see an obstruction and sometimes may not stop at all, different drivers having different habits. Accordingly, the number of stops calculated by the simulation overestimates the actual number of times a vehicle would become stationary on the street but correctly characterizes the number of times a vehicle is obstructed.

The calculated delay is trip delay, the difference between the time a driver spends on a link and the time he would spend if he traveled at his desired speed. This is also

the wait while stopped under the assumptions of the model although, for the reasons mentioned, it would be an overestimate of the stopped time that would occur on the street.

Our intutitive belief is that drivers would tolerate some added trip delay in order to decrease the irritation represented by stops and obstructions. Accordingly, we have chosen an overall criterion that is a weighted sum of stops and delay. Where a specific number is required, we shall take one stop as equivalent to 15 seconds of trip delay.

Certain traffic phenomena are not included in the present simulation model. Delays that result when a vehicle makes a left turn across opposing traffic are not represented. There is no speed-density relation included nor is there any slowdown process as a vehicle approaches a queue or red signal. However, this model does consider the possibility that a link may become too full for more vehicles to enter.

## TEST PROBLEM

A specific network of streets has been chosen for study. Although the simulation program is general, runs can only be made after specializing the program to a given network with specific flows, turning percentages, one-way streets, etc. The results we obtain with our network may or may not be representative of others. One might try to construct a "typical" network but it seems doubtful that meaningful general results could be obtained, because real situations tend to have individual peculiarities. Instead we take a specific example. It will demonstrate the feasibility of the techniques and will give some impression of what types of results can occur.

The test problem is a 5-street, 6-signal network drawn from Boston's Back Bay section (Fig. 1). The problem is small enough that multiple exploratory runs can be made within a reasonable amount of computer time and yet the problem still contains certain interesting complexities. The network has two interdependent loops. The six signals, each of which affects the flow on two streets, provide quite a few decision variables. Commonwealth, Berkeley, and Beacon carry traffic to and from the downtown area and have reasonably heavy flows during rush hours. Commonwealth has a pedestrian red-red phase. Both one- and two-way streets are included; in fact, four of the five streets are one-way. While this may seem like a large number, it turns out that almost all the streets in downtown Boston are one-way and this is the case in many cities. Thus, this type of network seems very worthy of study.

Field observations were made to determine the average input flows and turn probabilities needed by the simulation. Data were collected for the evening rush hour, called here "heavy flow" and for weekday early afternoon, called "medium flow." In addition a "light flow" was defined by taking one-half the medium flow values. Figures 2 and 3 show the turn probabilities and flow on each link for high and medium flow.

## METHODS OF SETTING SIGNALS

Finding the exactly optimal fixed-time setting for a complex model of a network of signals is a formidable mathematical and computational problem. The decision variables are numerous—each signal has a green split and an offset and the network as a whole has a period. Some signals may have more complex phases such as left turn arrows or pedestrian signals. We do not know how to solve this problem exactly, but we wish to examine what appear to be sensible approaches. Most of the emphasis will be on finding offsets, although some of our runs concern other aspects and our general search techniques apply to any set of variables.

The traffic signal problem is characterized by multiple local minima. In other words, if one takes a given setting and systematically makes small changes in each control variable, retaining changes that are improvements, and continuing the process until any small change only worsens the solution, a local minimum is, by definition, achieved. However, a different starting setting can lead to a different local minimum, which may be better or worse than the first. Neither is necessarily the global minimum, i. e., the smallest of all local minima.

In the case of maximal bandwidth settings, the criterion and model are sufficiently simple that exact methods are available for sorting through the local minima to find the global minimum. For the relatively complex model of the simulation, no theory presently exists to do this. One might conjecture that a good local minimum would exist in the neighborhood of, say, the maximal bandwidth setting, but such a conjecture would have to be tested.

This suggests the following approach: Find starting settings that would be expected to be good, evaluate them by simulation, and then apply to the best of them some systematic improvement procedure. Accordingly, in this section we discuss methods of determining starting settings and, in the next, improvement techniques.

The methods that we take up for finding starting settings are: maximal bandwidth, simswitch, and random search. This list is not exhaustive; a good setting might come from anywhere. However, each of these techniques is generally applicable.

## Maximal Bandwidth

The concept of a progression is widely used in synchronizing signals. The bandwidth of a progression in a given direction on a given street is that fraction of the cycle during which a vehicle could start at one end and, by traveling at designated speeds, go to the other end without stopping. The concept of a progression was conceived for arteries but it can be extended to networks.

In a network each direction on each street has its own bandwidth. To resolve conflicts among bandwidths, our mathematical program for the problem takes a weighted sum of bandwidths as the objective function. In order to keep the bandwidth on some streets from becoming too small, we shall require all bandwidths to be at least some fraction of the most heavily weighted bandwidth. One must be careful in this not to impose constraints that allow no feasible solution.

In determining maximal bandwidth settings we shall allow the design speed of the progression to vary within certain limits. This permits some recognition of the fact that the simulation does not use a single speed on the street. The period is also a variable of the problem and its value is chosen to maximize the objective function. The mathematical program can be formulated with the green splits as control variables as well, but in the work here we have not done this. Effective green time is divided between streets so that the ratio of green times is the ratio of their values of flow/saturation flow, i. e., Webster's equation (20).

The method for solving for signal settings to give maximal bandwidth is described elsewhere (11). For the Back Bay problem the mathematical program has 36 linear equations and involves 3 integer and 29 continuous variables. Its solution by branch and bound methods required solving 8 ordinary linear programs. The equations and the solution tree are given by Miller (22).

The maximal bandwidth program was solved for the medium traffic case. Since light flow is everywhere half of medium flow, the same solution will apply there. The heavy flow has a slightly different pattern and so possibly a different solution. However, we arbitrarily use the same settings in the work to follow.

## Simswitch

Progressions are generally considered appropriate for low and moderate flows, but at high flows the building up of queues during red tends to obstruct a platoon timed to arrive just as the signal changes to green. One approach to this situation is to make all the signals on a street green at the same time. Then all the platoons travel at once and do not interfere. Of course, the long traverses without stops that might be possible in a progression system cannot be made.

This simultaneous green procedure will be called "simswitch." It is primarily appropriate for rectangular grids. If all the green splits are the same, the signals turn green exactly together (except for pedestrian red-red and other special phases) and the system has only two parameters: green split and period. We shall call this case "simswitch 1." As another possibility, green time may be divided between streets

TABLE 1

MAXIMAL BANDWIDTH AND SIMSWITCH 1 RUNS
(575 seconds, exponential interarrival time input)

| Setting | Flow | Average Delay per Vehicle (seconds) | Average Stops per Vehicle (stops) |
|---|---|---|---|
| Maximal bandwidth (50-second period) | Low | 20 | 1.09 |
| | Medium | 23 | 1.00 |
| | Heavy | 32 | 1.08 |
| Simswitch 1 (50-second period) | Low | 22 | 0.86 |
| | Medium | 25 | 0.97 |
| | Heavy | 40 | 1.27 |
| Simswitch 1 (100-second period) | Low | 31 | 0.83 |
| | Medium | 36 | 0.86 |
| | Heavy | 52 | 1.02 |

proportional to flow/saturation flow. Then period is the only parameter. This case will be called "simswitch 2." The centers of the red periods will be made to coincide for signals on parallel streets.

## Simulation Results

Simulation runs have been made for all three flows with the maximal bandwidth settings and with two simswitch 1 settings, the first having a 50-second period, the second a 100-second period. Both simswitches employed equal green times in each direction. The traffic input at the edges of the network was Poisson (i.e., exponentially distributed interarrival times). All runs were for 575 seconds of street time, of which the first 50 seconds were ignored so as to allow the network to fill with cars. The results in terms of average delay per vehicle and average stops per vehicle are given in Table 1 and Figure 4.

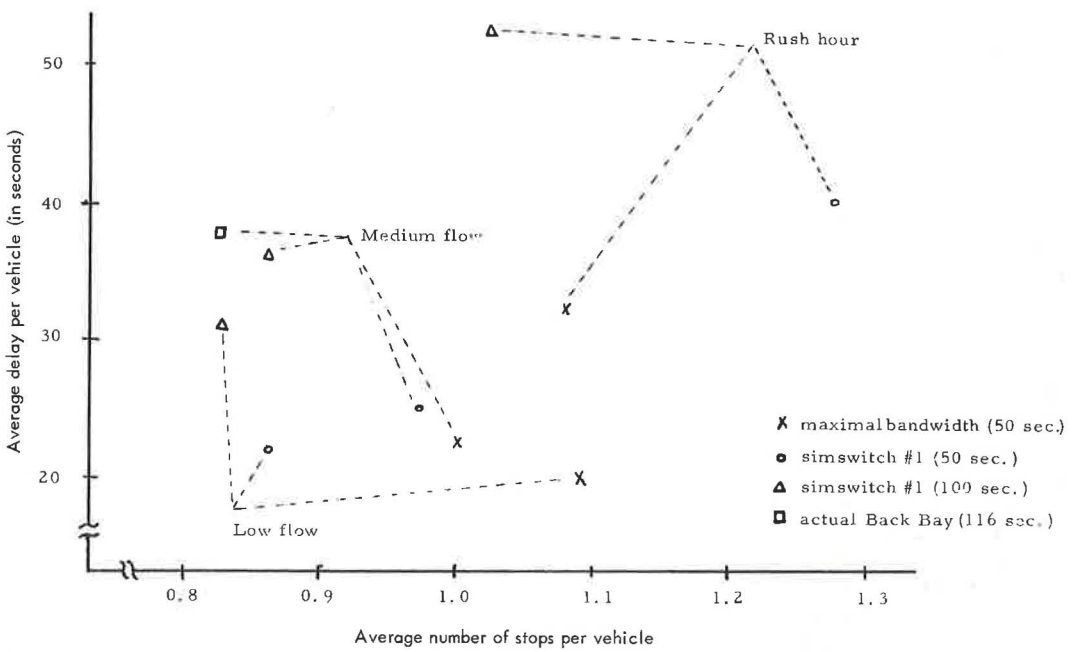


Figure 4. Values of performance measures for various signal settings and traffic flows.

The first observation to make is that the choice of performance criterion affects the choice of setting. For example, at heavy flow, if delay is disregarded and stops are the criterion, the 100-second simswitch 1 is best; whereas for the reverse situation maximal bandwidth is best. Similar situations occur at other flows.

If we adopt the trade-off relation of 1 stop = 15 seconds delay, the best settings are

| | |
|---|---|
| low flow: | 50-second simswitch 1 |
| medium flow: | maximal bandwidth |
| heavy flow: | maximal bandwidth |

These results are not necessarily what one would expect. Maximal bandwidth progressions would be expected to be good at low flows, simultaneous switching at high ones. Some of this can be explained retrospectively by examining the specific situations involved. More important, however, these results suggest that it is difficult to predict good settings without a detailed examination of traffic movement as through a simulation.

A result that conforms to expectations is that the longer period simswitch gives more delay but fewer stops. The driver can go farther during green without stopping but once stopped must wait for a longer red.

## Random Search

Because of the existence of local minima, there may be an advantage to generating a variety of different settings and evaluating them by simulation. A common way to generate a large number of different solutions in mathematical programming problems is to select values for the variables by a random process. In the present situation, offsets (in seconds) for signals with periods of 100 seconds or less can be taken directly from 2-digit random number tables, discarding values that are too large.

Random generation of settings has two interesting advantages. First of all, it is quite free of preconceived notions. It will explore unconventional settings good and bad and, from time to time, will produce rather effective new approaches. Second, after a number of settings have been run, a statistical estimate can be made of the chances of improving the results further.

As an example, 12 runs have been made with random offsets under the following conditions: medium flow, 50-second period, constant interarrival times for entering vehicles, and a run length of 200 seconds of street time, the first 100 ignored in calculating the objective function. The objective function uses the trade-off of one stop equals 15 seconds delay. The calculation is: objective function = (average delay in seconds/vehicle) + (15 seconds/stop) (average stops/vehicle).

Table 2 gives the results. The objective function ranges from 39.0 to 65.1 seconds/vehicle with a mean of 47.2 and a standard deviation of 6.8. The standard deviation makes possible an estimate of how hard it is to improve the run by a given amount through further random search. Thus, if the distribution of objective function values is assumed to be normal with the mean and standard deviation given, there is one chance in 20 that another run would be 36.0 or better.

For comparison, a maximal bandwidth run under the same conditions gives 42.4 seconds/vehicle. Thus, the best of the 12 random settings is about 8 percent better. This will not always happen, however. Under another set of operating conditions, the maximal bandwidth setting gave 4 percent better than the best of (a different) 12 random settings.

Random search is not a very elegant method and its usefulness is strongly dependent on the speed of the simulation because it is necessary to pick through many bad settings to find the good ones. However, random search is easy to do and with a fast simulation may well be competitive with other methods.

TABLE 2

12 RANDOM OFFSET RUNS
(200 seconds, medium flow, constant interarrival time input, 50-second period)

| Run | Objective Function (seconds/vehicle) | Run | Objective Function (seconds/vehicle) |
|---|---|---|---|
| 1 | 49.2 | 7 | 44.6 |
| 2 | 39.5 | 8 | 65.1 |
| 3 | 44.7 | 9 | 39.0 |
| 4 | 46.5 | 10 | 44.0 |
| 5 | 52.6 | 11 | 50.0 |
| 6 | 45.1 | 12 | 46.0 |

Mean: 47.2 seconds/vehicle Standard deviation: 6.8 seconds/vehicle

## Other Runs

A number of additional runs have been made.

Actual Back Bay Settings—For completeness, the green splits, period, and offsets currently in use at medium flow in Back Bay have been run. The result is plotted in Figure 4. The point is of interest as another possible setting in the test problem, but is not necessarily a good evaluation of the setting in actual practice. The reason is that our study condition isolates the test area with respect to timing of flows (i. e., platoon arrivals) although not their magnitude. This was necessary because otherwise it would be almost meaningless to vary period in the test problem.

Sensitivity of Maximal Bandwidth Calculation—If the maximal bandwidth calculation is sensitive to its input constants and if inaccurate constants are used, the resulting settings might be very poor. As some test of this possibility, a maximal bandwidth calculation has been made with green splits changed 20 percent and the speed range shifted from 30-36 ft/sec to 33-39 ft/sec. (The mean speed in the simulation is 33 ft/sec.) Other constants have been left alone. A very different setting emerges, although the objective function of the mathematical program is not much changed. When the setting is used in the simulation, the number of stops is unchanged and the delay increases only 5 percent. Thus, at least in this one case, although the settings are sensitive, the criterion does not change much.

Effect of Modeling Diffusion—Several runs have been made comparing the original Schwartz simulation to the present version, which include platoon dispersion. Little effect on average stops or delay is observed. However, the dispersion model has been kept in the program.

Simswitch 2—As described earlier, this setting was green splits determined by Webster's equation. The results are inferior to those of simswitch 1 on our test problem. We note, however, that Schwartz (16) in his simulation of Back Bay as a whole found simswitch 2 somewhat better.

Effect of Turning Vehicles—Progression systems cater to vehicles that go straight through the intersections. In our network an average of about 20 percent of the vehicles turn upon entering an intersection. This may degrade the performance of the maximal bandwidth setting. To investigate this, two runs were made at low flow with turns eliminated, one with the maximal bandwidth setting and the other with simswitch 1. The maximal bandwidth run showed a 25 percent reduction in average delay and 29 percent reduction in average stops. In contrast, simswitch 1 showed 9 percent reduction in each. This helps explain the somewhat surprising result found earlier that simswitch 1 was superior to maximal bandwidth at low flows.

## IMPROVEMENT TECHNIQUES

We consider next the possibility of improving a given setting. One important approach is to inspect the computer output, looking for trouble spots and making changes.

Beyond this, however, it seems desirable to develop systematic improvement methods. Many techniques have been proposed for optimizing a function of several variables. One of the simplest techniques and frequently a rather effective one is one-variable-at-a-time search, called by Leon (21) "univar." We shall describe a univar algorithm, discuss two ways of applying it to the improvement of offsets, and report our computational experience.

## A Univar Algorithm

Let $x_1, \ldots, x_n$ be the variables to be chosen and let the goal be to minimize an objection function, $f(x_1, \ldots, x_n)$ (here calculated by simulation). The steps to be taken are:

1. Set the $x_j$ to their starting values.
2. Calculate $f(x_1, \ldots, x_n)$. Set BVSF (= best value so far) to $f(x_1, \ldots, x_n)$ and IBVSF (= index of variable whose change caused BVSF to be reset) to 1.
3. Set $i = 1$.
4. Increase $x_i$ to $x_i + \Delta_i$, where $\Delta_i$ is the increment chosen for increasing and decreasing $x_i$. Calculate $f(x_1, \ldots, x_n)$. If its value is less than BVSF, reset BVSF to the new value, reset IBVSF to i, and restart step 4. Otherwise reduce $x_i$ by $\Delta_i$ (i.e., to its previous value) and continue.
5. If any improvement was made in step 4, go to step 7. Otherwise continue.
6. Decrease $x_i$ to $x_i - \Delta_i$. Calculate $f(x_1, \ldots, x_n)$. If its value is less than BVSF, reset BVSF to the new value, reset IBVSF to i and restart step 6. Otherwise increase $x_i$ by $\Delta_i$ and continue.
7. If $i < n$, set i to $i + 1$. Otherwise set $i = 1$.
8. If $i \neq$ IBVSF, return to step 4. Otherwise, there has been no improvement in one full cycle through all n variables and we are finished.

A few remarks may be added. No upper and lower bounds have been placed on the $x_i$ but this would not be difficult to do. Bounds are not necessary for offsets because offsets are treated modulo the period. Although the algorithm could be programmed, it is so simple that our practice has been to perform it by hand, doing only the simulation on the computer. The process can be stopped before the end if improvements become infrequent. Since the simulation uses random numbers to determine certain car movements and since only a finite time sample is run, it is possible that a local minimum could be produced by a statistical fluctuation. This does not appear to have been a problem in our runs, and it has been our choice to accept this possibility rather than, say, extend the region of search beyond the point where the objective function first starts to increase.

The results of applying univar to the best of the 12 random settings reported in Table 2 are given in Table 3. The offsets have been changed in the order of their signal numbers, although this is not required. Run conditions are the same except for changing offsets. The objective function is reduced from 39.0 seconds/vehicle to 36.5 seconds/vehicle, an improvement of about 6 percent. Twenty-one simulations have been used. We conclude that the method is capable of finding moderate improvements, even in relatively good settings. We conjecture that most of the effect is in the first iteration through all the offsets.

Several other univar runs have been made on various starting settings under various run conditions. These were all stopped after a single iteration and usually showed improvements in the 7-10 percent range.

## Choice of Variables

We wish now to distinguish between two ways of using univar. They will be called changing one signal at a time and changing a group. The distinction and its relevance are clearest in an arterial problem. Suppose the absolute offset of a signal, $S_j$, in the middle of an artery is changed. This has the effect of changing two important relative offsets, the ones relative to the adjacent signals on either side of $S_j$. In a real sense, we have changed two important variables at once, not one. If we wish to vary just

TABLE 3

APPLICATION OF UNIVAR ALGORITHM TO IMPROVE OFFSETS
(One-signal-at-a-time Method)

| Run No. | Offsets (sec) for Signal No. : | | | | | | Objective Function (sec/vehicle) | Is this a new minimum? |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | | |
| Original | 13 | 23 | 29 | 41 | 25 | 36 | 39.0 | |
| 2 | 15 | 23 | 29 | 41 | 25 | 36 | 41.5 | no |
| 3 | 11 | 23 | 29 | 41 | 25 | 36 | 38.1 | yes |
| 4 | 9 | 23 | 29 | 41 | 25 | 36 | 40.8 | no |
| 5 | 11 | 25 | 29 | 41 | 25 | 36 | 42.3 | no |
| 6 | 11 | 21 | 29 | 41 | 25 | 36 | 40.8 | no |
| 7 | 11 | 23 | 31 | 41 | 25 | 36 | 41.6 | no |
| 8 | 11 | 23 | 27 | 41 | 25 | 36 | 36.8 | yes |
| 9 | 11 | 23 | 25 | 41 | 25 | 36 | 40.3 | no |
| 10 | 11 | 23 | 27 | 43 | 25 | 36 | 36.5 | yes (best value) |
| 11 | 11 | 23 | 27 | 45 | 25 | 36 | 37.6 | no |
| 12 | 11 | 23 | 27 | 43 | 27 | 36 | 39.5 | no |
| 13 | 11 | 23 | 27 | 43 | 23 | 36 | 39.7 | no |
| 14 | 11 | 23 | 27 | 43 | 25 | 38 | 37.9 | no |
| 15 | 11 | 23 | 27 | 43 | 25 | 34 | 36.9 | no |
| Second Iteration | | | | | | | | |
| 16 | 13 | 23 | 27 | 43 | 25 | 34 | 38.3 | no |
| 17 | 9 | 23 | 27 | 43 | 25 | 34 | 37.8 | no |
| 18 | 11 | 25 | 27 | 43 | 25 | 34 | 39.7 | no |
| 19 | 11 | 21 | 27 | 43 | 25 | 34 | 40.0 | no |
| 20 | 11 | 23 | 29 | 43 | 25 | 34 | 38.7 | no |
| 21 | 11 | 23 | 25 | 43 | 25 | 34 | 40.4 | no |

$$\text{Improvement} = \frac{39.0 - 36.5}{39.0} = 6.4\,\%$$

Note: There is no reason to continue the second iteration beyond run No. 21 because we have returned to the same set of signal offsets as existed in run No. 10. From No. 22 on the results would be exactly the same as those for runs following run No. 10.

TABLE 4

APPLICATION OF UNIVAR ALGORITHM
(Group Method)

| Run No. | Offsets (sec) for Signal No. : | | | | | | Objective Function (sec/vehicle) | Is this a new minimum? |
|---|---|---|---|---|---|---|---|---|
| | 6 | 1 | 2 | 3 | 4 | 5 | | |
| Original | 36 | 13 | 23 | 29 | 41 | 25 | 39.0 | |
| 2 | 36 | 15 | 25 | 31 | 43 | 27 | 37.5 | yes |
| 3 | 36 | 17 | 27 | 33 | 45 | 29 | 36.7 | yes (best value) |
| 4 | 36 | 19 | 29 | 35 | 47 | 31 | 38.6 | no |
| 5 | 36 | 17 | 29 | 35 | 47 | 31 | 45.2 | no |
| 6 | 36 | 17 | 25 | 31 | 43 | 27 | 50.3 | no |
| 7 | 36 | 17 | 27 | 35 | 47 | 31 | 40.9 | no |
| 8 | 36 | 17 | 27 | 31 | 43 | 27 | 41.0 | no |
| 9 | 36 | 17 | 27 | 33 | 47 | 31 | 39.9 | no |
| 10 | 36 | 17 | 27 | 33 | 43 | 27 | 42.6 | no |
| 11 | 36 | 17 | 27 | 33 | 45 | 31 | 44.8 | no |
| 12 | 36 | 17 | 27 | 33 | 45 | 27 | 46.3 | no |

$$\text{Improvement} = \frac{39.0 - 36.7}{39.0} = 5.9\,\%$$

Note: The starting signal is No. 6 and the order of removal from the remaining group is 1, 2, 3, 4, and 5. The algorithm has been stopped at the end of one complete iteration.

one of them, we should hold fixed all the offsets on one side of $S_j$ and vary together the offsets of $S_j$ and all the signals on the other side. This procedure can be started at one end of the artery and carried through to the other. At the first step, all offsets but the first are changed together; at the second, all but the first two; and so on, with the group being changed shrinking by one at each step. We still call the process univar since, in effect, a single relative change dictates all the offsets.

The same idea can be extended to networks, but is not so clear there. A starting signal can be defined and the others put in a group with each signal assigned an order of removal. At each stage of univar the relative offset of all those in the shrinking group is varied with respect to the next. However, in the network case, there will often be more than one adjacent pair whose relative offset is being changed. Therefore, some of the intended effect is lost. Nevertheless, if the grouping is arranged so that one of the adjacent pairs is usually an important heavy flow street and the others are usually in less important places, the effect can still be quite strong.

The group change method has been applied to our network. One univar iteration has been performed and the results are given in Table 4. There is reasonably good improvement, again about 6 percent, although actually not quite as much as found by the one-signal-at-a-time method in its first iteration. However, the group change is expected to be at its best where arteries dominate the problem and such is not the case here.

## CONCLUSIONS

A traffic simulation has been used to evaluate several methods of setting signals in a specific test problem. In addition, the simulation has been made an integral part of proposed search procedures for improving settings. The work suggests the following conclusions:

1. There are operationally significant differences among settings that might beforehand be expected to be good. For example, the settings for medium flow plotted in Figure 4 were chosen for their potential effectiveness, yet the range among them is about 20 percent in stops/vehicle, 50 percent in delay/vehicle, and 30 percent in our composite objective function.
2. The criterion adopted for evaluating system performance substantially affects the choice of setting and the operating results. Thus, Figure 4 shows that, at heavy flow, 100-second simswitch 1 would be best if stops were the only criterion, whereas maximal bandwidth would be best if delay were the only criterion.
3. It is difficult to predict what settings will be good without detailed examination of traffic movement. This is illustrated by the reversal of our expectations for simswitch and maximal bandwidth between high and low flows. A number of factors, such as turning percentages, flow levels, and pedestrian signals, may affect performance, whereas most mathematical optimization methods to date work with simplified models.
4. Settings with progressive timing seem more likely to be degraded in performance by turning vehicles than are systems with simultaneous switching. This reasonable result is suggested by runs with and without turns.
5. A promising approach to finding good settings is to develop likely ones by various methods, evaluate them by simulation, and then improve them by systematic procedures. Three methods of finding starting settings have been tried: maximal bandwidth progression, simswitch, and random search. Of these, maximal bandwidth has worked out fairly well in this test problem, but a single problem provides insufficient basis for generalization. Random search may be helpful if the simulation to be used is fast. A systematic improvement procedure using one-variable-at-a-time search is found effective in improving settings. Improvements in the case of this test problem have been on the order of 6-10 percent.

## REFERENCES


1. Versace, John. Factor Analysis of Roadway and Accident Data. HRB Bull. 240, 1960, pp. 24-32.
2. Pennsylvania Turnpike Joint Safety Research Group. Accident Causation. Westinghouse Air Brake Co., Pittsburgh, 1954.
3. Traffic Control and Roadway Elements—Their Relationship to Highway Safety. The Automotive Safety Foundation, 1963.
4. Vey, Arnold H. Effect of Signalization on Motor Vehicle Accident Experience. Inst. of Traffic Eng. Proc., 1933, pp. 56-63.
5. Solomon, David. Traffic Signals and Accidents in Michigan. Public Roads, Vol. 30, No. 10, Oct. 1959, pp. 234-237.
6. Greenshields, Bruce D. Traffic Accidents and the Quality of Traffic Flow. HRB Bull. 208, 1959, pp. 1-15.
7. May, Adolf D. Friction Concept of Traffic Flow. HRB Proc., Vol. 38, 1959, pp. 493-510.
8. Casciato, L., and Cass, S. Progress and Experience in Toronto Traffic Control System. Conference on Traffic Surveillance, Simulation and Control, U. S. Bureau of Public Roads, Washington, 1964.
9. San Jose Traffic Control Project Progress Report. San Jose, Calif., May 1966.
10. Morgan, J. T., and Little, J. D. C. Synchronizing Traffic Signals for Maximal Bandwidth. Operations Research, Vol. 12, Nov. 1964, pp. 896-912.
11. Little, J. D. C. The Synchronization of Traffic Signals by Mixed Integer Linear Programming. Operations Research, Vol. 14, July 1966, pp. 568-594.
12. Gerlough, D. L., and Wagner, F. A. Simulation of Traffic in a Large Network of Signalized Intersections. Proc. Second Internat. Symp. on Theory of Traffic Flow, London, 1963.
13. Katz, J. H. Simulation of a Traffic Network. Communications of the ACM, Vol. 6, Aug. 1963, pp. 480-486.
14. Kell, J. H. Results of Simulation Studies as Related to Traffic Signal Operation. Presented at 33rd Annual Meeting, Inst. Traffic Eng., Toronto, 1963.
15. Blum, A. M. A General Purpose Digital Simulator and Examples of Its Application. Part III—Digital Simulation of Urban Traffic. IBM Systems Journal, Vol. 3, 1964, pp. 41-47.
16. Schwartz, Jesse G. An Urban Street Network Simulation Model. MIT Civil Eng. Dept. Rept., Cambridge, Mass., Nov. 1966.
17. Grace, Muriel J., and Potts, R. B. A Theory of the Diffusion of Traffic Platoons. Operations Research, Vol. 12, 1964, pp. 255-275.
18. Pacey, G. M. Progress of a Bunch of Vehicles Released From a Traffic Signal. Rept. No. RN/2665, Road Res. Lab., England, 1956.
19. Helly, W., and Baker, P. O. Accelerated Noise in a Congested Signalized Environment. Rept. R&D 65-2, Engineering Dept., Port of New York Authority, June 1965.
20. Webster, F. V. Traffic Signal Settings. Road Res. Tech. Paper No. 39, Her Majesty's Stationery Office, London, 1958.
21. Leon, A. Memoranda on Optimization Techniques. Sloan School of Management, MIT, 1966 (multilith).
22. Miller, S. B. The Solution of a Two-Dimensional Traffic Light Problem. Master's Thesis, MIT, 1966.