Signal Design for Congested Networks Based on Metering

AYELET GAL-TZUR, DAVID MAHALEL, AND JOSEPH N. PRASHKER

A method is presented for designing signal programs for congested urban networks in which one or more intersections are bound to become oversaturated before the others. The idea on which the method is based is to limit the volumes entering the network, to adjust them to the capacity of the critical intersection, and to avoid intersection blockage. This metering procedure enables the designer to determine the location of the queues and to direct them to long links that can act as buffers. This strategy is called queue location management. The procedure is based on integrating a mathematical programming model with TRANSYT. The model is used to determine the green splits and the level of metering, whereas TRANSYT is used for simulating the dynamic processes within the cycle and for offset optimization. A numerical example demonstrates the procedure.

Congestion in urban networks is one of the most urgent problems with which traffic engineers are confronted. McShane and Roess (*1*) emphasized that oversaturation is not merely another level of the common traffic problem—delays and stops—it is a new kind of problem. The reason has to do with the unique phenomenon typical of such conditions, namely, intersection blockage. Blockage or spillback occurs when a long queue starting at one intersection continues into the one upstream. In addition to causing a deterioration in the capacity of the intersection in which it occurs, blockage eventually leads to a decrease in the capacity of the vicinity. The throughput of the blocked intersections drops rapidly, vehicles accumulate in queues, and routes that regularly face normal conditions become congested. Preventing such situations is the major goal of traffic engineers.

Various researchers have investigated these conditions. Gazis (*2*), one of the pioneers in this field, observed the problem of the critical intersection and tried to optimize its performance. He considered a single intersection with a two-phase control and offered a strategy composed of two programs, one to be implemented during the first part of the rush hour and the other during the second part. Michalopoulos and Stephanopoulos (*3,4*) continued this work and tried to broaden the problem by making modifications in an algorithm. Overflow was handled by limiting queue lengths. They also formulated the interdependencies among intersections in the network and solved the equations that represented all intersections simultaneously. Singh and Tamura (*5*) approached the problem from the network point of view. Assuming that demand exceeds capacity at all links, they described the traffic flow by means of a set of differential equations. Their model tried to minimize queue lengths by adjusting green splits.

Civil Engineering Department, Technion-Israel Institute of Technology, Haifa, Israel.

Lieberman (*6*), McShane and Roess (*1*), and Kirby (*7*) all suggested metering the number of vehicles allowed into the network. The phenomenon that demand levels above capacity exist for short periods during the rush hour was the basis for this remedy. The solution is well known for freeway control (*8*). Lieberman distinguished between the treatment given to a major arterial and that to a critical intersection. He formulated various objective functions for controlling traffic flow along an arterial that consider queues and spillbacks explicitly and that emphasize the goal of efficient use of several parts of the arterial as buffers. For critical intersection (CI) control, Lieberman presented a mathematical model suitable for a network that contains at least one CI. The objective function in this case was the release of the queues in the various approaches to the CI according to predetermined priorities. McShane and Roess listed three types of metering: internal, external, and release. Internal metering controls the discharge rate at internal links of the network in order to prevent too many vehicles from reaching critical intersections. External metering limits the number of vehicles entering the network. Release metering refers to regulating the exit of vehicles from parking areas, garages, and lots. Kirby described research undertaken at the University of Leeds that was concerned with improving and applying metering strategies or, as it is called by the author, gating strategies.

During peak demands, as it is clear that queues are inevitable, there is need to determine their location. Thus, distributing the queues over the network in such a way that the damage that they may cause will be minimal (especially avoiding the damage of spillbacks) should be the explicit goal of the control process. The authors call this planning strategy queue location management (QLM), and it is suited for congested periods, when the capacity of an interior network is lower than the demand at its entry links. The QLM procedure, which integrates a mathematical model with TRANSYT, is described in this paper. The mathematical model is used for metering, whereas TRANSYT serves both as a planning tool and as a simulation tool. TRANSYT is the most widespread tool for planning signal-timing programs (*9*). Its main advantage lies in the overall network approach it takes to determine the cycle time, green splits, and offsets. It does not guarantee a global optimum, but it searches for a good solution while taking many parameters into account (e.g., platoon dispersion characteristics). It also allows the user to input subjective priorities among the various objectives. Despite the intensive use of TRANSYT, it is not suitable for dealing with congested networks.

TRANSYT has two drawbacks in handling oversaturated conditions. The first is that queue lengths are given as an output of the program but are not considered either in the objective function or as a constraint. The other problem, which is more severe, stems from the requirement to input the volumes on each link before the design process. In uncongested networks, when there is a cycle failure, these volumes are simply the sum over the demand of all movements leading to a certain link. As the network becomes more congested and queues accumulate, the volume of certain links is determined by the green duration. In these cases, the volume of the downstream intersection is not the demand but the throughput of upstream intersections. In other words, during congestion volumes are sometimes determined by the throughput and green durations. Thus a vicious circle is created. TRANSYT requires the volumes in order to calculate signal programs, but volumes are determined by the green durations. This vicious circle can be solved by simultaneously determining the volumes and green durations. This can be done by combining TRANSYT with an external program. Integrating TRANSYT with other procedures has already been implemented for arterial signal design (*10*).

A detailed method of applying the ideas of metering to congested urban networks is proposed in this paper. The different stages of the design process are described in the next section. Then the mathematical model is formulated, and in the final section the procedure is illustrated.

PROPOSED APPROACH

The approach proposed combines two modules into a complete algorithm that implements the concept of metering in a way that reflects reality in a reliable manner. The QLM strategy is applied to the entire network by a mathematical model that considers one cycle as the smallest time unit. The dynamic processes within the cycle, such as platoon dispersion and queue formation, are taken into account in the second module, in which TRANSYT is used. The method was designed for networks in which one intersection is known to have a smaller capacity than its demand, and thus is bound to become a CI. Under congested conditions, the throughput of the CI determines the throughput of the network in the vincinity, and therefore the optimal operation of the CI becomes the major goal.

During certain peak periods, experience indicates that demand is greater than the overall network capacity and thus queues are inevitable. Such peak periods are generally not very long, but their consequences affect traffic-flow conditions long after the demand decreases to below capacity. These lasting effects result from the chain reaction of long queues and spillbacks. Therefore, efforts should be made toward minimizing the damage caused by the temporary imbalance between capacity and demand, and special attention should be given to the prevention of spillbacks.

The following approach is recommended for the short periods when demand is higher than capacity. The goal of the design process is to prevent overflows at critical links by metering the amount of vehicles entering them. The residual amount of vehicles can be accumulated at long links at the entrances of the network, which play the part of buffers. Alternatively, vehicles at these entry links are rerouted to alternative paths either by active means of appropriate signs or by passive means (as drivers, encountering long queues, choose a different route on their own).

A prerequisite for the implementation of the foregoing strategy is to define the subnetwork surrounding the CI, for which signal programs should be redesigned. Such a subnetwork is characterized by entry links that are either long enough to be used as buffers or able to lead to alternative routes. The traffic volumes at the entrances of the network are assumed to be constant during the planning horizon considered. It is also assumed that vehicle paths are predetermined. In those cases in which this assumption is unrealistic, an assignment program should be integrated with the control process.

The terms used in this paper are defined as follows:

- Demand—the number of vehicles per time unit arriving at an entry movement;
- Volume—the number of vehicles per time unit arriving at an internal link movement; and
- Throughput—the number of vehicles per time unit discharged from any movement.

The distinction between demand and volume is necessary because the volume of an internal link is determined either by demand or by the capacity of external links.

The calculation algorithm consists of two types of variables: the volumes within the network and the green durations of all intersections. The algorithm is composed of four stages, which are described in Figure 1 and the following paragraphs.

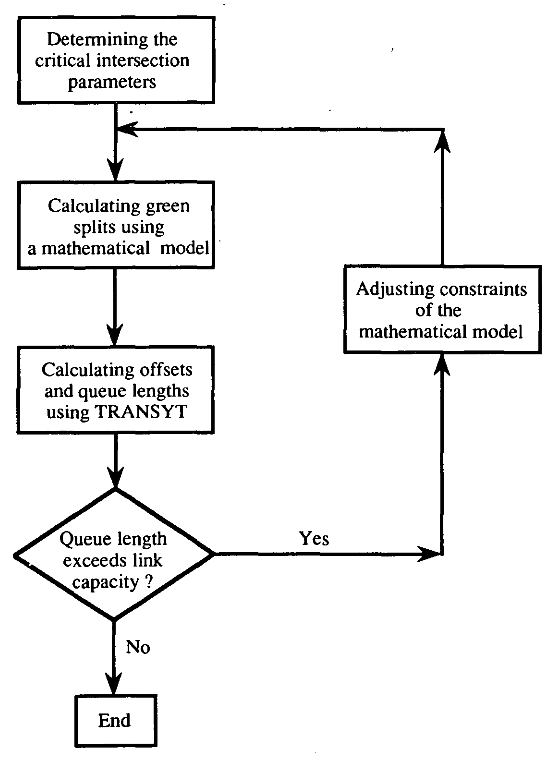

FIGURE 1 Flow chart of the proposed procedure.

Stage 1—Determining Critical Intersection Parameters

The purpose of the first stage is to find the cycle time and green splits of the CI that would achieve maximum throughput. The output at this stage is the required throughput of each movement of the CI and the suitable green durations that would enable obtaining the throughput. To avoid infinite queues, the volumes of these movements should be equal to their throughput. Several ways of calculating these parameters appear in the literature. For example, the means suggested by Mahalel and Gur (*11*) explicitly considers the phenomenon of a decay in saturation flow as the green duration extends beyond a certain value. This phenomenon calls for a short green duration. There is a trade-off between shortening the green durations in order to achieve high saturation flows and prolonging the cycle in order to minimize the fraction of lost time. Maximum throughput is obtained by determining the green durations in a way that the marginal saturation flow is identical for all phases. At this point a maximum average saturation flow is also achieved.

Stage 2—Calculating Green Splits of All Intersections

After the optimal parameters of the CI are determined, the green splits of all other intersections in the network are calculated so that spillbacks are avoided. This objective is achieved by metering the flow of the links leading to the critical movements in such a way that the volumes arriving at each critical movement equal its predetermined throughput. The calculation process is based on the following assumptions:

- The cycle time determined for the CI at the first stage is common to all intersections in the network;
- The demands and saturation flows of the network are deterministic; and
- The smallest time unit considered by the model is one cycle, and therefore the values of the network parameters (demand, queues, etc.) are calculated for the end of each cycle.

As the demand is assumed to be deterministic and constant, the queue length of each movement might either grow at a fixed rate to infinity or be equal to zero. Defining the queue length of an internal link as zero means that the number of vehicles that can be discharged during the green phase must be greater than or equal to its volume. Thus, the throughput of any internal movement equals its volume. As for entry movements, their goal is to ensure that the number of vehicles entering the network should not exceed its capacity. These are the locations at which queues can accumulate if metering is applied. The throughput of these movements is the minimum between their demand and their capacity (Equation (1). In cases in which the green duration is long enough to discharge all vehicles arriving during the cycle (undersaturated conditions), throughput is identical to the demand; otherwise (in oversaturated conditions), it equals the discharge rate per cycle.

$$\text{throughput } (m) = \min \left[\text{demand } (m);\ \text{capacity } (m)\right] \qquad (1)$$

where m is an entry movement. The volume of each internal movement is the sum of the throughputs of all movements leading to it.

There are two possible means of applying metering. One is to prolong the green duration of undersaturated movements, if there are such movements at the intersection being considered. If metering is needed for all movements, the "all-red" period should be prolonged beyond the minimum duration required for safety considerations. Such an act might cause difficulties if this period becomes long enough to result in a flow during the all-red interval. A constraint for avoiding such cases can be added to the planning algorithm, although it is not included in the formulation given in this paper. A mixed-integer linear programming (MILP) model was developed for this stage and is presented in detail in the next section.

Stage 3—Calculating Offsets and Queue Lengths of All Intersections

After the green splits and volumes of each phase and movement are determined, they are input into TRANSYT and TRANSYT is used for offset calculation. Additional information derived is the maximum queue length on each link during the cycle. This information cannot be obtained from the mathematical model, as it represents conditions in the network only at the end of the cycle. For every internal link, a comparison is to be performed between maximum queue length and link storage capacity in order to track overflows.

Stage 4—Adjusting Parameters of Mathematical Model

In cases for which overflow is tracked, adjustments in the MILP model are required. The modifications are done by enlarging the difference between the volume and the capacity of these movements. After the necessary adjustments are completed, an iterative process is conducted by repeating Stages 2 through 4. A full example demonstrating the execution of the four stages is given subsequently.

MATHEMATICAL MODEL

The mathematical model is used in the second stage of the solution process and is based on the CI parameters (cycle time and green split), which were calculated in the first stage. As was noted in the previous section, the throughput of the whole network is determined by that of the critical intersection, which thus is an input for the mathematical model. The role of the model is to determine the throughput of each intersection surrounding the CI. A MILP model was developed, in which the integer variables distinguish between oversaturated and undersaturated entry movements (see Equation 1). This distinction is necessary in order to calculate the throughput of these movements.

The parameters of the design process are intended to maximize the throughput of the CI. Any feasible solution of the mathematical model that satisfies this purpose is also optimal.

The parameters of the CI appear in the constraints of the model. The terms used in the model are defined as follows:

- Movement—a group of vehicles having the same destination;
- Link—a set of movements sharing the same approach; and
- Phase—a set of movements receiving the right of way simultaneously.

The goal of the mathematical model is to find a feasible solution that achieves the following constraints. The first set of constraints (Equation 2) simply states that the sum of the green durations of all phases should be less than or equal to the difference between the cycle time and the total lost time. For cases in which metering is determined by the model, the constraint satisfies the "<" sign, and a positive slack variable is added.

$$\sum_j G_{ij} \leq C - F_i \cdot L \qquad \forall i \tag{2}$$

where

C = common cycle time of network (sec)
L = lost time for each phase (sec),
G_{ij} = green duration of phase j at intersection i (sec),
F_i = number of phases at intersection i,
i = intersection index, and
j = approach index.

Note that G_{ij} is a decision variable, whereas C and L are predetermined parameters.

The second set of constraints refers to the movement of the CI only (Equation 3). The volume of any critical movement (a movement of the CI) should be equal to its throughput, as was determined in the first stage.

$$P_{\mathrm{CI}lm} \cdot \sum_{kr \in \Omega_{\mathrm{CI}l}} O_{ksr} = T_m \qquad \forall\, l \in \mathrm{CI} \qquad \forall m \tag{3}$$

where

T_m = maximum throughput of movement m of CI, as calculated in first stage;
Ω_{il} = set of movements leading to link l of intersection i (thus $\Omega_{\mathrm{CI}l}$, in particular, is set of movements leading to link l of intersection CI);
O_{ksr} = throughput per cycle of movement r of link s of intersection k (notations r, s, and k being used to indicate movements leading to link l of CI);
P_{ilm} = proportion of vehicles in movement m of the total volume of link l of intersection i;
l = link index; and
m = movement index.

Thus $\Sigma_{ksr \in \Omega_{il}} O_{ksr}$ is the volume of link l of intersection i, and $P_{ilm} \cdot \Sigma_{ksr \in \Omega_{il}} O_{ksr}$ is the volume of movement m of link l of intersection i. Note that T_m and P_{ilm} are parameters and O_{ilm} is a decision variable.

For all other internal movements in the network, the volume should equal the throughput in order to avoid queues (Equation 4). In contrast to movements of the CI, however, both the volume and the throughput of each internal movement are unknown.

$$P_{ilm} \cdot \sum_{ksr \in \Omega_{il}} O_{ksr} = O_{ilm} \qquad \forall \text{ internal movements} \tag{4}$$

The next set of constraints (Equation 5) ensures that the green duration of the phase to which movement m belongs is long enough to discharge the throughput determined in Equation 4).

$$S_{ilm} \cdot G_{ij} \geq O_{ilm} \qquad \forall \text{ internal movements (where } m \in \text{ phase } j) \tag{5}$$

where S_{ilm} is the saturation flow of movement m of link l of intersection i, and $G_{ij} \cdot S_{ilm}$ is the maximum number of vehicles that can be discharged from movement m of link l of intersection i per cycle.

As for entry links, the number of vehicles entering each link is a predetermined demand D_{il}, and thus the demand per entry movement is $P_{ilm} \cdot D_{il}$. The next set of constraints allow the planner to limit the queue-building rate per cycle at each link (Equation 6).

$$D_{il} - \sum_{m=1}^{M_l} O_{ilm} = q_{il} \qquad \forall \text{ entry links} \tag{6}$$

where D_{il} is the demand per cycle of link l of intersection i, and q_{il} is the queue-building rate per cycle of link l of intersection i.

Equations 7 and 8 define the 0–1 variables (X_{ilm}), which indicate whether the throughput of entry movement m of link l of intersection i is determined by its demand rate ($P_{ilm} \cdot D_{il}$). In such a case, X_{ilm} receives the value of 0; or is determined by its capacity ($S_{ilm} \cdot G_{ij}$) and receives the value of 1.

$$S_{ilm} \cdot G_{ij} - P_{ilm} \cdot D_{il} + M \cdot X_{ilm} \geq 0 \qquad \forall \text{ entry movements} \tag{7}$$

$$P_{ilm} \cdot D_{il} - G_{ij} \cdot S_{ilm} + M \cdot (1 - X_{ilm}) \geq 0 \qquad \forall \text{ entry movements} \tag{8}$$

where

$$X_{ilm} \begin{cases} = 1 & \text{if } P_{ilm} \cdot D_{il} \geq S_{ilm} \cdot G_{ij} \\ = 0 & \text{otherwise} \end{cases}$$

and

M = a very large number.

The throughput per cycle of each movement (O_{ilm}) is the minimum between its demand and its saturation flow per cycle. The minimum operant can be expressed by the variables X_{ilm} as represented in Equation 9.

$$X_{lm} \cdot S_{ilm} \cdot G_{ij} + (1 - X_{ilm}) \cdot P_{ilm} \cdot D_{il} = O_{ilm} \qquad \forall \text{ entry movements} \tag{9}$$

This equation cannot be part of the MILP model because it is not linear. Thus a new variable is created, according to Equation 10, which is a product of two original variables.

$$SX_{ilm} = X_{ilm} \cdot S_{ilm} \cdot G_{ij} \qquad (10)$$

The new variable turns Equation 9 into Equation 11.

$$SX_{ilm} + (1 - X_{ilm}) \cdot P_{ilm} \cdot D_{il} = O_{ilm} \qquad \forall \text{ entry movements} \qquad (11)$$

Finally Equations 12 to 14 define the variable SX_{ilm} by linear equations.

For practical reasons, a minimum green-duration constraint is added to the model (Equation 15).

$$SX_{ilm} - M \cdot X_{ilm} \leq 0 \qquad \forall \text{ entry movements} \qquad (12)$$

$$SX_{ilm} - S_{ilm} \cdot G_{ij} \leq 0 \qquad \forall \text{ entry movements} \qquad (13)$$

$$S_{ilm} \cdot G_{ij} - SX_{ilm} \leq M \cdot (1 - X_{ilm}) \qquad \forall \text{ entry movements} \qquad (14)$$

$$G_{ij} \geq \text{Mingreen} \qquad \forall i \ \forall j \qquad (15)$$

The objective function of the mathematical model is to minimize the queue-building rates at the entry links of the network (Equation 16).

$$\text{Min} \sum_{i \in C1} \sum_{j \in C2} W_{il} \cdot q_{il} \qquad (16)$$

where

$C1$ = set of entry intersections,
$C2$ = set of entry links, and
W_{il} = weights that can be added to represent relative importance of each link.

The next section will demonstrate the implementation of the above model to a small network.

NUMERICAL EXAMPLE

The network described in the numerical example consists of six intersections, one of which is critical and marked as Intersection 10 (see Figure 2). The network has 8 entry links. At all intersections, a separate phase is devoted to each approach, and thus the link index and the phase index may be combined. Saturation flows of all traffic movements are 1,800 vehicles per hour (vph) (0.5 vehicles per second); the demand on each entry link is 540 vph; and the lost time per phase is 3 sec. Figure 2 presents a schematic description of the network topology and the proportion of vehicles turning at each movement.

The first stage of the design process, in which the CI optimal parameters are calculated, will not be described here in detail, as it is not the main issue of this paper. The cycle time was

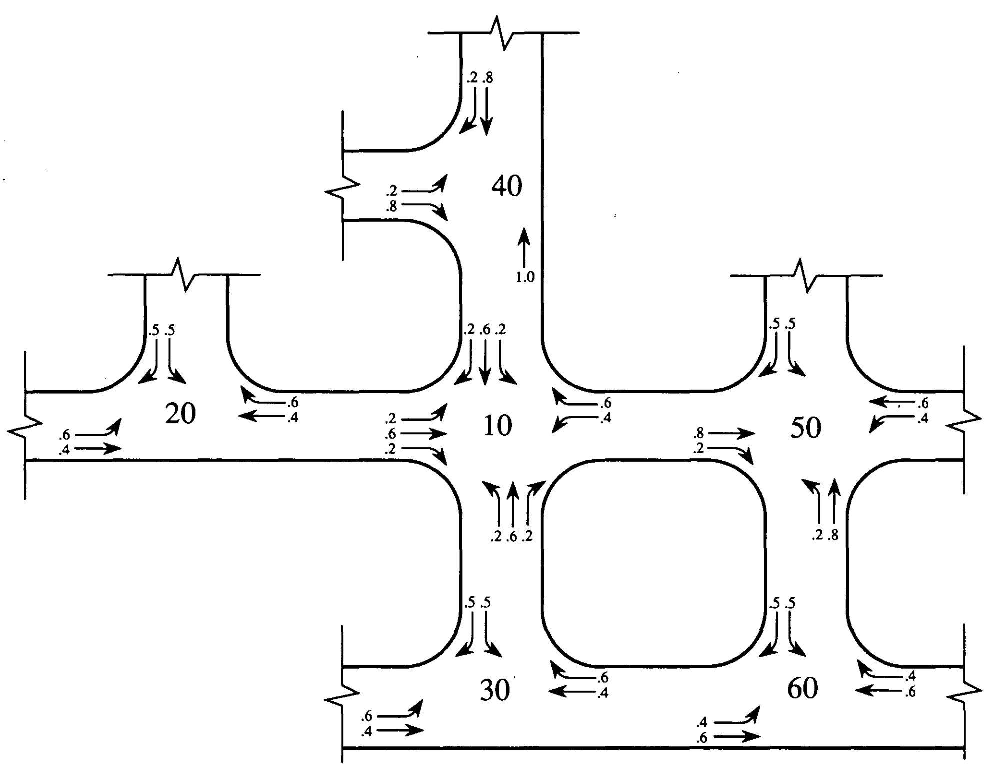

FIGURE 2 Network structure and proportion of turning vehicles of each link.

selected to be 100 sec, and the aggregate throughput of the critical movements is as follows:

- Northbound, 504 vph;
- Westbound, 216 vph;
- Southbound, 504 vph; and
- Eastbound, 360 vph.

In the second stage of the algorithm, the general formulation of the MILP model, as described in the preceding section, turns into a detailed set of equations representing the network. Following are some equations that are part of the full mathematical model formulated for the example given. Each movement is identified by a combination of six digits; for example,

XXYYZZ

where

XX = intersection index (i) (according to the intersection numbering in Figure 2);
YY = link index (l) (10 for northbound links, 20 for westbound links, 30 for southbound links, and 40 for eastbound links); and
ZZ = movement index (m) (10 for a right-turn movement, 20 for a through movement, and 30 for a left-turn movement).

The objective function is as follows:

$$\text{Min}[q_{2030} + q_{2040} + q_{3040} + q_{4030} + q_{4040} + q_{5020} + q_{5030} + q_{6020}] \quad (17)$$

In this equation, equal importance was assigned to all entry links.

Equation 18 is according to Equation 2 and refers to Intersection 30. This intersection has three approaches (phases), each having a lost time of 3 sec, and thus its effective green time should not exceed 91 sec ($100 - 3 \cdot 3 = 91$).

$$G_{3020} + G_{3030} + G_{3040} \leq 91 \quad (18)$$

Equation 19 is a demonstration of Equation 3 for critical movement 101020. Equation 19 sets the volume of Movement 101020, originated from Movements 302010 and 304030, to be equal to the throughput of this critical movement (8.4 vehicles per cycle).

$$0.6 \cdot (O_{302010} + O_{304030}) = 8.4 \quad (19)$$

The next constraints (Equations 20 and 21) refer only to internal movements and are according to Equations 4 and 5. For example, movement 302010 is an internal movement. Equation 20 sets its volume (which is the sum of the throughputs entering it) to equal its throughput, so that no queues build up.

$$0.6 \cdot (O_{602010} + O_{603020}) = O_{302010} \quad (20)$$

Equation 21 states that the capacity of the internal movement is greater than or equal to its volume. The capacity is determined by the green duration multiplied by the saturation flow per second (0.5).

$$0.5 \cdot G_{3020} \geq O_{302010} \quad (21)$$

Movements 304020 and 304030 are entry movements belonging to Link 3040. The link's queue-building rate per cycle is determined by the difference between its demand per cycle (15 vehicles) and the sum throughput of its movements (Equation 22, which is according to Equation 6).

$$15 - (O_{304020} + O_{304030}) = q_{3040} \quad (22)$$

Equations 23 and 24 define the 0–1 variable X_{304030} and are based on the general formulation presented in Equations 7 and 8. This variable distinguishes between the case in which the throughput of Movement 304030 equals its demand ($X_{304030} = 1$) and the case in which it equals its capacity. Note that M was set at 1,000.

$$0.5 \cdot G_{3040} + 1000 \cdot X_{304030} \geq 9 \quad (23)$$

$$0.5 \cdot G_{3040} - 1000 \cdot (1 - X_{304030}) \leq 9 \quad (24)$$

Equation 25 demonstrates the way the throughput of an entry movement (in this case Movement 304030) is determined (according to Equation 11).

$$O_{304030} = SX_{304030} + (1 - X_{304030}) \cdot 0.6 \cdot D_{3040} \quad (25)$$

Equations 26 to 28 define the variables SX_{304030} and OX_{304030} as explained by Equations 12 to 14.

$$SX_{304030} - 1000 \cdot X_{304030} \leq 0 \quad (26)$$

$$SX_{304030} - 0.5 \cdot G_{3040} \leq 0 \quad (27)$$

$$0.5 \cdot G_{3040} - SX_{304030} \leq 1000 \cdot (1 - X_{304030}) \quad (28)$$

For every phase of Intersection 30, a minimum green duration should be obtained (see Equation 15); for example, the constraint for the first phase is Equation 29.

$$G_{3010} \geq 5 \quad (29)$$

The results of the MILP are given in Table 1. In order to summarize these results, link data are presented instead of movements data. These results indicate that queues can be avoided at certain entry links (such as 6020); at others, vehicles are bound to accumulate rapidly (such as at Link 4040). All-red duration was prolonged at Intersection 40 beyond the required minimum, thus the sum of nominal green durations is 56 sec, which is smaller than the effective green time (91 sec).

The third column shows the throughput (in vph) of every movement in the network. Special attention should be given to entry Links 2020, 3040, 4040, 5020, and 5030. Their throughput is lower than the demand rate (540 vph), which reflects the fact that metering is active at those links. In stage three, TRANSYT is used with the output of the MILP model.

Table 2 gives a summary of the main results produced by TRANSYT. It should be noted that queue storage capacity is not defined for entry links as these links are assumed to be very long. As was expected, the maximum queue length of each link, within the cycle, as given by TRANSYT, is longer

TABLE 1 Results of First Iteration of MILP Model

Link No.	Direction	Entry Link	Queue Building up rate (per cycle)	Green duration	Throughput
1010	NB	N	0	28	504
1020	WB	N	0	12	216
1030	SB	N	0	28	504
1040	EB	N	0	20	360
2010	NB	Y	0	86	540
2020	WB	Y	10.83	8	150
3020	WB	N	0	21	384
3030	SB	N	0	6	461
3040	EB	Y	2.32	25	456
4010	NB	N	0	28	384
4030	SB	Y	0	30	540
4040	EB	Y	12.5	5	456
5010	NB	N	0	23	418
5020	WB	Y	12.5	5	90
5030	SB	Y	10.64	8	157
5040	EB	N	0	51	418
6020	WB	Y	0	30	540
6030	SB	N	0	7	120
6040	EB	N	0	54	504

than the queue length at the end of the cycle, as calculated by the MILP model. Special attention should be given to internal links where link capacity tends to be low and overflow must be prevented.

If no queue exceeds the link length, the process ends at this point, otherwise Step 4 of the algorithm is applied. In this case, the equations representing the queue in the MILP model should be modified. The difference between the volume entering the link and its capacity is increased. In mathematical terms, Equation 5 turns into Equation 30, in which Diff is a constant defined by the designer.

$$S_{ilm} \cdot G_{ij} - P_{ilm} \cdot \sum_{rsk \in \Omega_{il}} O_{rsk} \geq \text{Diff} \qquad (30)$$

In the example described in this section, spillbacks were detected on Links 1010 (which belongs to the CI) and 5010. For Link 1010, two options are available. The first option is to change the signal program for the CI and to solve this problem from the beginning. The second is to change the right-hand side of the appropriate constraint (Equation 30). The throughput of Link 1010 was reduced from 504 vph to 486 vph, but its green duration remained the same. The difference between the volume of Link 5010 per cycle and its capacity was set at greater than or equal to 1. Table 3 provides a summary of the results of the second iteration of the MILP model and the results of TRANSYT. The changes can be demonstrated for Link 1010. In order to reduce the volume of this link, the green duration of approach 3040 was decreased from 25 to

TABLE 2 Maximum Queue Lengths Resulting from TRANSYT in First Iteration

Link No.	Direction	Entry Link	Saturation degree	Queue storage capacity	Max. queue length
1010	NB	N	1	10	11
1020	WB	N	1	10	3
1030	SB	N	1	14	12
1040	EB	N	1	11	7
2010	NB	Y	0.35	irrelevant	3
2020	WB	Y	3.75	irrelevant	29
3020	WB	N	1	12	10
3030	SB	N	0.99	12	10
3040	EB	Y	1.2	irrelevant	27
4010	NB	N	1	13	11
4030	SB	Y	1	irrelevant	15
4040	EB	Y	6	irrelevant	30
5010	NB	N	0.97	9	10
5020	WB	Y	6	irrelevant	30
5030	SB	Y	3.33	irrelevant	30
5040	EB	N	0.46	9	4
6020	WB	Y	1	irrelevant	15
6030	SB	N	0.95	9	2
6040	EB	N	0.52	11	6

TABLE 3 Results of Second Iteration of MILP Model and TRANSYT

Link No.	Direction	Entry Link	Queue Building-up Rate (per cycle) (MILP)	Green duration (MILP)	Throughput (MILP)	Saturation Level (TRANSYT)	Max. queue Length (TRANSYT)
1010	NB	N	0	28	468	0.93	10
1020	WB	N	0	12	216	1	3
1030	SB	N	0	28	504	1	12
1040	EB	N	0	20	360	1	8
2010	NB	Y	0	86	540	0.35	3
2020	WB	Y	10.83	8	150	3.75	28
3020	WB	N	0	43	383	0.49	4
3030	SB	N	0	26	461	0.98	10
3040	EB	Y	0	22	397	1.36	28
4010	NB	N	0	27	482	0.99	10
4030	SB	Y	0	30	540	1	15
4040	EB	Y	12.5	5	90	6	28
5010	NB	N	0	24	403	0.93	9
5020	WB	Y	10.24	5	90	6	30
5030	SB	Y	12.5	9	162	3.33	30
5040	EB	N	0	50	410	0.46	3
6020	WB	Y	0	30	540	1	15
6030	SB	N	0	7	118	0.94	2
6040	EB	N	0	54	468	0.48	9

22 sec; the flow from this direction thus changed from 456 vph to 397 vph.

Results of TRANSYT after the second iteration (Table 3) show that queues at Links 1010 and 5010 were actually shortened; however, these are not the only changes. The new splits and offsets affected other queues at entry links, and thus all queues should be checked again. Stages 2 to 4 of the algorithm can be repeated until a satisfactory solution is obtained.

CONCLUDING REMARKS

The contribution to traffic management in congested urban networks of the method presented in this work compared with other metering strategies lies in its integration of a mathematical tool and a simulation tool. This integration achieves a better reflection of reality in the planning process and thus ensures more reliable results. Reliability is further increased by the possibility of exploiting the capabilities of each tool without violating the basic axioms on which it was grounded. The proposed method follows the principle of adjusting the capacity of the entry links to the one of the CI through metering at the entrances of the network. Thus, queues are accumulated in predetermined approaches and QLM is achieved, meaning that the length and location of queues are controlled by the planner. Because of differences in the geometry of the various intersections (number of lanes, parking arrangements, etc.), not all movements are oversaturated, even in congested conditions. Therefore, the design procedure explicitly deals with congested conditions when and where they exist; it does not, however, consider oversaturation on all links as a preliminary assumption [in contrast to other works (*3,5*)].

As a part of the procedure advanced here, an MILP model was formulated through which the green durations and volumes of all movements are determined; it ensures that no queues remain in internal links at the end of each cycle. This model was combined with TRANSYT in a way that uses the advantages of the latter while overcoming its inefficiencies. The MILP model creates undersaturated conditions within the network, which are the only conditions suitable for TRANSYT. TRANSYT is then used for offset optimization in calculating the optimal offsets and maximum queue lengths during the cycle. The two planning tools, the mathematical model and TRANSYT, complement each other to achieve signal-timing programs that are designed explicitly for the congested conditions that exist in the network.

Despite the advantages of the proposed approach presented in this paper, some limitations of the algorithms suggested should be addressed. The main deficiency of the method stems from the assumption that the route choice within the network is constant and is known in advance. Further research should concentrate on integrating the two existing modules of the process presented with an assignment module. Another disadvantage of this method, which unfortunately is common to most planning methods used today, is the separation of the split and the offset optimization processes. This separation, though meant to simplify the design algorithm, might damage the quality of the solution obtained.

The third limitation stems from the usage of TRANSYT as the simulation tool. TRANSYT performs a deterministic simulation, and thus does not reflect the stochastic nature of reality. Substituting TRANSYT with a different simulation tool that can continue serving the functions of optimizing offsets and estimating expected queue lengths, while explicitly considering stochastic phenomena, should be considered in the future.

It is hoped that implementation of the method described in this paper in a real network will be tested soon. The conclusions of such an experiment will serve as guidelines for improving the method.

REFERENCES

1. R. W. McShane and R. P. Roess. *Traffic Engineering*. Prentice-Hall, Englewood Cliffs, N.J., 1990, Chapter 30.
2. D. C. Gazis. Optimum Control of a System of Oversaturated Intersections. *Operations Research*, No. 12, 1964, pp. 815–831.
3. P. G. Michalopoulos and G. Stephanopoulos. Optimal Control of Oversaturated Intersections: Theoretical and Practical Considerations. *Traffic Engineering and Control*, Vol. 19, No. 5, 1978, pp. 216–222.
4. P. G. Michalopoulos and G. Stephanopoulos. An Algorithm for Real-Time Control of Critical Intersections. *Traffic Engineering and Control*, Vol. 20, No. 1, 1979, pp. 9–15.
5. M. G. Singh and H. Tamura. Modelling and Hierarchical Optimization for Oversaturated Urban Road Traffic Network. *International Journal of Control*, Vol. 20, No. 6, 1974, pp. 913–934.
6. E. B. Lieberman. Concepts of Control for Oversaturated Networks. Presented at TRB Traffic Signal Systems Committee Meeting, Minneapolis, Minn., 1990.
7. H. R. Kirby. Transport Research at the University of Leeds: Report for 1991. *Traffic Engineering and Control*, Vol. 33, No. 3, 1992, pp. 171–178.
8. A. J. Lindley and D. G. Cappele. Freeway Surveillance and Control. *Traffic Engineering Handbook*, 4th ed. (J. L. Pline, ed.). Prentice-Hall, Englewood Cliffs, N.J., 1992.
9. D. I. Robertson. *TRANSYT: A Traffic Network Study Tool.* Laboratory Report 253. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1969.
10. M. J. Moskaluk and P. S. Parsonson. Arterial Priority Option for the TRANSYT-7F Traffic-Signal-Timing Program. In *Transportation Research Record 1181*, TRB, National Research Council, Washington, D.C., 1988, pp. 57–60.
11. D. Mahalel and Y. Gur. *Optimal Operation of a Critical Intersection* (in Hebrew). Research Report. Transportation Research Institute, Technion, Haifa, Israel, 1991.

Publication of this paper sponsored by Committee on Traffic Flow Theory and Characteristics.