tation Research Record 699, 1979, pp. 41-49.
9. J. Taylor and R. Carter. Photogrammetric Data Acquisition for a Freeway Ramp Operations Study. HRB, Highway Research Record 319, 1970, pp. 78-87.
10. L.J. Mountain. The Potential of Small-Format Photography for Traffic Data Collection. Univ. of Leeds, Leeds, England, Ph.D. thesis, 1979.
11. J. Uren. Urban Traffic Data Collection from Conventional Aerial Photography. Proc., Conference on Traffic Data Collection, Univ. of Leeds, Leeds, England, Jan. 1979, pp. 143-166.

Decentralized Control of Congested Street Networks

RUDOLF F. DRENICK, SAMIR A. AHMED, AND WILLIAM R. McSHANE

A mathematical model for traffic flow in city streets and its control is presented. The model is thought to be appropriate to the kind of control systems anticipated over the next 10-20 years. These systems are expected to rely to a considerable extent on communications that are less binding on driver behavior than the traffic signals that are now virtually the sole communication-control devices. The model is accordingly probabilistic. It is not limited to the problem of optimal signalization. On the contrary, optimal as well as satisfactory traffic control can be based on it. In either case, the problem develops into one of a special kind of nonlinear programming and of very large scale. A scheme is described for its decomposition into decentralized control, and several algorithms for its computational execution are outlined.

The most important technological development during the past decade may have been that of the very-large-scale-integration (VLSI) circuit chip. By all present indications, its impact will be widespread and profound. The control of traffic in city streets is likely to be affected. One can perhaps anticipate that current control systems will be supplemented and perhaps even supplanted by others that use the new technology. The new generation of such systems might seek to induce desirable traffic patterns by greater flexibility in its adjustment of red-green splits and in its use of turn signals at intersections. It might also combine the conventional signals with traffic advisories broadcast over general, citizens band, and perhaps even dedicated radio transmission channels, all in an effort to create a more satisfactory traffic-flow pattern.

The common characteristic of most of these strategies is that they are not binding on the driver in the way that conventional traffic signals now are. One can accordingly expect drivers' reactions to them to be of an even less deterministic nature than their reactions to the present ones. A model of the traffic system under these conditions seems most appropriate if formulated stochastically, and this is what has in fact been done in this study. The control variables of the system are, roughly speaking, the probabilities with which vehicles can be induced to make right turns, left turns, or no turns at the intersections of the street network in response to the various signals to which their drivers are exposed. The traffic signals, of course, remain as a set of control variables as well.

The problem of designing a traffic control system can then be viewed as that of making the best, or at least a satisfactory, choice of those probabilities. Formulated in this way, it develops into a constrained mathematical programming problem whose solution is made difficult partly by its nonlinearity and partly by its large scale. The nonlinearity is admittedly of a very special kind. The constraints, as well as the objective functions, are typically multilinear in the control variables. This is a feature that should be exploited in the solution procedure, and the several solutions that have been considered do so.

The large scale of the problem, on the other hand, suggests decentralized control schemes. Such schemes are at least intuitively most appropriate when the controlled system is made up of many subsystems geographically distributed over a wide area. It is then an appealing idea to exercise control over each subsystem separately, perhaps based mainly on inputs obtained locally, and to perform the necessary coordination through a hierarchy of supervisory controllers. Decentralized control schemes have the potential of reducing the cost of data communications, providing a high level of fail-safe capability, and allowing greater flexibility in the design and implementation of control strategies.

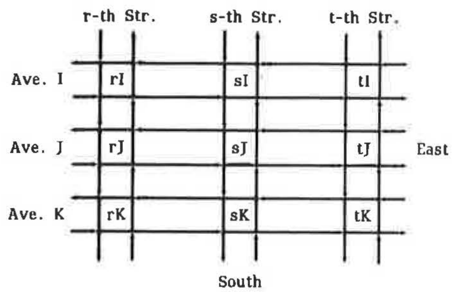
Surprisingly, however, little work has been reported on the application of decentralized control concepts to optimize the operations of large-scale urban transportation systems and even less work that treats these systems stochastically (1). The currently accepted versions of these concepts are explained in a recent article by Barry (2). Chu (3) explored the optimal decentralized control of a string of high-speed, densely packed vehicles using on-board controllers. Looze and others (4) and Kumar and others (5) proposed decentralized control schemes for regulating traffic on urban freeway corridors. Saridis and Lee (6) discussed the general problem of hierarchical control and management of traffic systems, while Chu (7), Sarachik (8), Singh and Titli (9), and Gershwin and others (10), among others, suggested decentralized control algorithms for large street networks, in nonstochastic formulations. A comprehensive survey of decentralized control methods and their applications to large-scale systems has recently appeared (11).

In this paper, a traffic network model based on probabilistic concepts is developed, and the problem of traffic control in such a network is formulated. The reasoning that leads to decentralization as a technique for the solution of that problem is outlined, and, finally, solution algorithms are briefly discussed, those that are already available as well as those that have been developed in this study.

TRAFFIC NETWORK MODEL

A traffic network is treated in this study as a stochastic system that is controlled by influencing the probability that a vehicle arriving at an intersection will make a right or left turn or no turn at all. The discussion presented here is based on a rather simple network of streets. Its extensions to

Figure 1. Segment of a street network.



more general and complex configurations will, however, be readily apparent.

Consider a network that is a rectangular grid of streets numbered 1,2,3, etc., and avenues labeled A,B,C, etc. Figure 1 shows the segment involving the streets numbered r, s, and t and avenues I, J, and K. During every red-green cycle, intersection sJ processes an input quadruple i of arriving platoons into an output quadruple k of departing ones, and it does so with certain "processing" probability p_{ik}^{sJ}. The indices i and k represent the sizes, compositions, and directions of the arriving and departing platoon quadruples at intersection sJ, respectively. A platoon is said to have departed from an intersection when it arrives at the next one. The processing probabilities at the different intersections are the control variables of the system. Signalization is present as a concern but is derived, rather than explicit, in this formulation.

Since every platoon quadruple i is processed into some quadruple k, one must have

$$\Sigma_k p_{ik}^{sJ} = 1 \qquad p_{ik}^{sJ} \geq 0 \tag{1}$$

where i and k range over all of their possible values. In addition, if p_i^{sJ} is the joint probability of arrival of quadruple i and q_k^{sJ} is the joint probability of departure of quadruple k, then

$$q_k^{sJ} = \Sigma_i p_{ik}^{sJ} p_i^{sJ} \tag{2}$$

This equation relates the input and output probabilities at the same intersection. A second relation prevails between inputs and outputs at different intersections. It expresses the fact that input quadruple i at a particular intersection consists of portions of the output quadruples released from the four intersections immediately connected to it. In Figure 1, the input quadruple arriving at intersection sJ is made up of the platoons coming from intersections sI, sK, rJ, and tJ. The joint arrival probability could thus be expressed as

$$p_i^{sJ} = (q_{k_1}^{sI})(q_{k_2}^{sK})(q_{k_3}^{rJ})(q_{k_4}^{tJ}) \tag{3}$$

where the four factors at the right-hand side represent the probabilities of having output platoon patterns from intersections sI, sK, rJ, and tJ that match the input quadruple pattern to intersection sJ.

The processing of quadruple i into k by intersection sJ presumably takes a certain known processing time t_{ik}^{sJ}, which depends on both the input i and the output k. One way of assessing the performance of intersection sJ, therefore, would be by its mean processing time:

$$\tau^{sJ} = \Sigma_{ik} t_{ik}^{sJ} p_{ik}^{sJ} p_i^{sJ} \tag{4}$$

Equations 1-4 written down for all intersections describe the state of traffic flow in the entire network. Control is exercised by adjusting the processing probabilities at the different intersections in such a way that a satisfactory flow pattern results. In this study, a pattern is considered satisfactory if it avoids "overload" throughout the network. The term overload can be given various interpretations; the simplest (though not obviously the most realistic) is that overload is avoided at intersection sJ if

$$\tau^{sJ} \leq \Delta^{sJ} \tag{5}$$

where Δ^{sJ} is the cycle length there. The reason for adapting this definition is that a violation of Equation 5 at any one intersection will eventually lead to unbounded waiting times and queue lengths there. These would entail the same phenomena at neighboring intersections and thus represent an expanding nucleus of congestion in the network.

The problem of achieving a satisfactory flow pattern, as formulated here, is seen to be a mathematical programming problem, though one without an obvious objective function. A set of processing probabilities p_{ik}^{sJ} that obeys the constraints of Equations 1-5 is to be determined, if it exists. In other words, any feasible solution to the problem is regarded to be satisfactory. It is customary in mathematical programming to prescribe a suitable objective function and to seek a solution that is optimal relative to it. In this case, one can take the same approach on the basis of the argument that overload at even only one intersection, as just explained, represents a highly undesirable traffic phenomenon and that it accordingly should be avoided if at all possible. This view suggests a rather natural choice of the objective function, namely

$$\tau = (1/N)\Sigma\tau^{sJ} \tag{6}$$

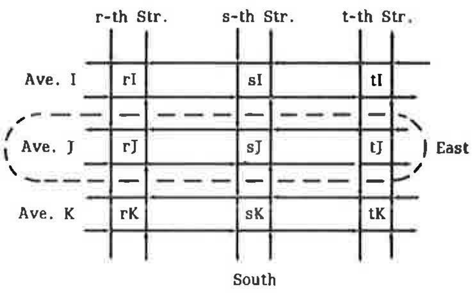
in which the sum extends over all intersections in the network (and the factor 1/N is there mainly for cosmetic reasons). The minimization of τ would of course be subject to the constraints of Equations 1-5. Additional constraints may be needed to impose origin-destination (O-D) specifications, queue-length limits, and other traffic restrictions. These are not mentioned here for simplicity of exposition.

If the Equation 6 choice of an objective function is considered inappropriate, others can be substituted for it, and several that have been considered by traffic engineers are natural candidates. The nature of the problem, and to some extent also that of its solution, are not greatly affected by such modifications.

The problem in any case is of very large scale in general, regardless of whether scale is measured in terms of the number of variables or the number of constraints. The number depends on what variables or constraints are counted. In terms of variables, if one considers only the processing probabilities p_{ik}^{sJ}, there will be as many as there are platoon quadruplets per intersection, squared (to allow for arrivals and departures), and multiplied by the number N of intersections in the network. In terms of constraints, if one considers only the probability and overload constraints (Equation 5 and Figure 1), there will be N(N + 1).

The problem is also nonlinear in its variables, as one readily recognizes. If Equation 6 is used as the objective function, the nonlinearities are of a special nature. The variables enter into the constraints and the objective function in sums of products, but none are raised to powers other than 0 or 1. They are, in other words, multilinear func-

Figure 2. Decentralization by avenue.



tions of those variables. The same is true of the problem if objective functions other than that of Equation 6 are used, provided they are expected values of cost-benefit parameters attached to the traffic flow.

CONTROL DECENTRALIZATION

The large scale of the programming problem described in the preceding section makes it highly desirable to seek a solution by way of a decomposition algorithm or, which is saying the same thing, to effect control by decentralization. Of the various schemes that have been considered so far in this study, one that is patterned on the hierarchical decomposition procedures developed by Mesarovic and his students (9, Chapter 4) currently appears the most promising. A brief description of these procedures and their use in the study is given here.

The idea of the hierarchical decomposition schemes is to resolve the Lagrangian L that corresponds to a large-scale programming problem into a number of terms

$$L = \Sigma_J L^J \tag{7}$$

each of which involves variables associated mainly with a single subsystem rather than with the system as a whole. There evidently is a good deal of latitude in the interpretation of the word "mainly" and, for that matter, the word "subsystem". In fact, success with the approach is often tantamount to a judicious exploitation of that latitude.

The approach that has been used in this study so far is the following. It has been assumed that the traffic in the network is characterized by a "dominant direction of flow," e.g., from north to south in Figure 1. One can then assign the control of traffic along and across each avenue to a subordinate controller, as indicated in Figure 2. The coordination of the signaling among avenues can be assigned to one supervisory controller or to several.

To do so, one collects in each of the terms L^J all those making up the Lagrangian L that can be associated with the traffic on and across avenue J. With the objective function of Equation 6, this makes

$$L^J = (1/N)\,\Sigma_s \tau^{sJ} + \Sigma_s \lambda^{sJ} (\tau^{sJ} - \Delta) \tag{8}$$

disregarding any additional terms attributable to O-D constraints and others. The mean processing times τ^{sJ} in Equation 8 are explicitly multilinear functions of the processing probabilities $p_{ik}{}^{sJ}$ of intersection sJ but implicitly also of those of other intersections. Strictly speaking, τ^{sJ} would be a function of the processing probabilities of all other intersections but, with the dominant direction of flow, only of those lying north of avenue J.

One can now seek an optimal set of processing probabilities in the usual way--namely, among those that satisfy the following equations:

$$\partial L/\partial p_{ik}{}^{sA} = \partial L^A/\partial p_{ik}{}^{sA} + \partial L^B/\partial p_{ik}{}^{sA} + \ldots + \partial L^Z/\partial p_{ik}{}^{sA} = 0 \tag{9}$$

$$\partial L/\partial p_{ik}{}^{sB} = \partial L^B/\partial p_{ik}{}^{sB} + \ldots + \partial L^Z/\partial p_{ik}{}^{sB} = 0 \tag{10}$$

$$\ldots$$

$$\partial L/\partial p_{ik}{}^{sZ} = \partial L^Z/p_{ik}{}^Z = 0 \tag{11}$$

and so on, along with the overload and other constraints. The form of these equations is due to the assumption of a dominant direction of flow that implies that L^A depends only on the $p_{ik}{}^{sA}$, L^B only on $p_{ik}{}^{sA}$ and $p_{ik}{}^{sB}$, etc.

The form of Equations 9-11 further suggests that the controller for avenue J be assigned the solution of the optimization problem represented by the equation

$$\partial L^J/\partial p_{ik}{}^{sJ} + \alpha_{ik}{}^{sJ} = 0 \quad (J = A, B, \ldots, Z) \tag{12}$$

along with overload and other constraints pertinent to avenue J. This optimization would be for the processing probabilities $p_{ik}{}^{sJ}$ of the intersections along that avenue. The quantity

$$\alpha_{ik}{}^{sJ} = \partial L^K/\partial p_{ik}{}^{sK} + \ldots + \partial L^Z/\partial p_{ik}{}^{sZ} \tag{13}$$

would be supplied to the avenue J controller by a superior to be treated as an additive constant during the optimization. It would be updated by the superior and resupplied to the subordinates for a new optimization.

The actual computational procedure, as in most nonlinear programming problems, would not, however, seek a direct simultaneous solution of Equations 12 and 13 but would use a search algorithm selected from among the several candidates briefly discussed in the next section of this paper.

SOLUTION TECHNIQUES

It has been pointed out that the problem of devising satisfactory or optimal traffic control, as formulated in this study, is a constrained nonlinear programming problem. It is, in fact, of a special kind that has been referred to here as "multilinear". This is true regardless of whether or not the control scheme is decentralized.

Several existing solution techniques suggest themselves. Among them are, first of all, a number of well-tested algorithms for the solution of constrained nonlinear programming in general (12,13). Their very generality, however, tends to be a disadvantage in that their convergence may be slow and uncertain in practice. Another group of techniques goes under the term of geometric programming (14,15). They apply to a more restricted class of problems but one that includes many of the multilinear ones. Unfortunately, they do not readily accommodate equality constraints of the kind that are inherent in the traffic-control problem formulation discussed here and thus are not particularly appropriate.

An effort was accordingly made in this study to investigate solution algorithms that are tailored to multilinear problems. One such algorithm exploits the fact that the multilinear programming problem is a natural generalization of the linear ones and hence is a rather direct analog to the well-known simplex algorithm (16).

As of this writing, however, a number of specialized techniques patterned on the gradient method are being favored because they seem most amenable to the kind of extensions in the traffic-control problem that are expected to become necessary in the near future.

DISCUSSION OF FORMULATION

This paper has described the present status of a study that seeks to formulate the problem of traffic control in city streets, with a view to the way in which the control might be executed in another decade or two. It may be of interest to add some brief remarks on the thinking that led to the current formulation, the features of it that are now thought to be undesirable, and the developments in it that are anticipated for the near future.

The idea of the formulation arose from a recent effort at developing a mathematical approach to organization theory (17). The parallel between the control of traffic in city streets and organizations may seem rather remote, but there are in fact a number of important analogies. For one, both can be designed with the aim of avoiding overload among junctions--i.e., among the intersections of city streets and among the members of an organization. Moreover, in both the control variables develop into probabilities of the kind that have here been called processing probabilities. There are, however, substantial distinctions as well. Most important may be the fact that the flow of traffic in a street network is a much more involved phenomenon than the flow of information and material in a well-functioning organization. In fact, it may be safe to say that an organization that had as disorderly a flow pattern as city traffic would be virtually unmanageable.

The main shortcoming of the current formulation of the control problem is felt to be its nondynamic character. It is a tacit assumption in a solution by mathematical programming that, once obtained, the solution will also be promptly adopted. In traffic control, however, and especially in the kind of control scheme envisioned here, this is unlikely to be so. The control system will thus have to monitor its own success with the traffic pattern by means of suitably placed sensors and adjust its control signals accordingly. The dynamics of the resulting feedback loops will have to be combined with the driver characteristics in order to achieve satisfactory operation. At this time, the gradient-like methods mentioned in this paper seem the most amenable of those considered or developed so far.

REFERENCES

1. Traffic Control Systems Handbook. Federal Highway Administration, U.S. Department of Transportation, 1976.
2. A.J. Barry. Distributed Intelligence in Traffic Control Systems. Proc., Engineering Foundation Conference on Research Directions in Computer Control of Urban Traffic Systems, Asilomar, CA, ASCE, New York, 1979, pp. 58-69.
3. K.C. Chu. Decentralized Control of High-Speed Vehicular Strings. Transportation Science, Vol. 8, No. 4, 1974, pp. 361-384.
4. D.P. Looze and others. On Decentralized Estimation and Control with Application to Freeway Ramp Metering. IEEE Trans. on Automatic Control, Vol. AC-23, No. 2, 1978, pp. 268-275.
5. K.S. Kumar and others. Reliable Decentralized Control Strategies for Freeway Regulation. U.S. Department of Transportation, April 1980.
6. G.N. Saridis and C.S. Lee. On Hierarchically Intelligent Control and Management of Traffic Systems. Proc., Engineering Foundation Conference on Research Directions in Computer Control of Urban Traffic Systems, Asilomar, CA, ASCE, New York, 1979, pp. 209-218.
7. K.C. Chu. Decentralized Real-Time Control of Congested Traffic Networks. International Business Machines Corp., Armonk, NY, Res. Rept. RC-6337, Dec. 1976.
8. P.E. Sarachik. Clearing of Congested Multi-Destination Networks. Proc., Engineering Foundation Conference on Research Directions in Computer Control of Urban Traffic Systems, Asilomar, CA, ASCE, New York, 1979, pp. 231-250.
9. M.G. Singh and A. Titli. Systems Decomposition, Optimization, and Control. Pergamon Press, Oxford, England, 1978.
10. S.B. Gershwin and others. Hybrid Optimization in Urban Traffic Networks. U.S. Department of Transportation, Rept. DOT-TSC-RSPA-79-7, 1979.
11. N.R. Sandell and others. Survey of Decentralized Control Methods for Large-Scale Systems. IEEE Trans. on Automatic Control, Vol. AC-23, No. 2, 1978, pp. 108-128.
12. D.G. Luenberger. Introduction to Linear and Nonlinear Programming. Addison-Wesley, Reading, MA, 1965.
13. E. Polak. Computational Methods in Optimization: A Unified Approach. Academic Press, New York, 1971.
14. R.J. Duffin, E.L. Peterson, and C. Zener. Geometric Programming. Wiley, New York, 1967.
15. C.S. Beightler and D.T. Philips. Applied Geometric Programming. Wiley, New York, 1976.
16. R.F. Drenick. Decentralized Control of Street Traffic. U.S. Department of Transportation, 1980.
17. R.F. Drenick. Organization and Control. In Directions in Large-Scale Systems (Y.C. Ho and S.K. Mitter, eds.), Plenum Press, New York, 1976, pp. 279-302.