TOWARD A MARKOVIAN TRAFFIC CONTROL EVALUATION SYSTEM

Lonnie E. Haefner*, Washington University; and
John A. Warner III*, R. L. Banks and Associates

The objective of this paper is to present a technique for evaluating traffic control alternatives and resultant performance changes in a complex traffic system. The system can be described as existing in a finite number of states, with some known probability of transition from one state to another in a given time period. Limited field examples from the merging section of the Baltimore Harbor Tunnel are presented and analyzed through a multistage optimization process, termed Markovian decision theory. The technique prescribes an optimal traffic control alternative for each possible state of the system. Changes in flow-density relationships, employment of traffic cones, and a hypothesized metering example are discussed as preliminary tests of the technique.

•MANY TRAFFIC SITUATIONS exist as very complex entities, and there are several ways to improve operation. Control alternatives may be of a permanent type (striping, pretimed versus actuated signalization) or an immediate option within a given control system (metering rate). Such a choice is often based on the criterion of change of flow rate and related to long-term gains of the system and society through amelioration of congestion. The ability to relate an alternative to recognizable long-term gains can be difficult, particularly if the traffic system can operate in a variety of ways at different times.

The objective of this paper is to illustrate preliminary application of a decision algorithm designed to account for these aspects of a complex traffic system by using examples from the merging area of the Baltimore Harbor Tunnel. Through a multistage optimization process over a finite number of recognizable states of the system, the technique prescribes an optimal traffic control alternative for each state. The collection of such optimal alternatives, termed the optimal strategy, maximizes long-run gains to the system through induced changes in the flow-density relationship.

CONCEPTUAL FORM OF THE MODEL

Consider an area of controlled-access highway that requires vehicular traffic to execute merging maneuvers. The state i of the merging area is defined by the level of vehicular density k. There is a probability p_{ij} that can be associated with the transition from one state i to another j. The set of all possible transition probabilities (transition matrix) then describes the behavior of the system. For example, in the four-state ($i = 4$) case, the transition matrix might be

$$p_{ij} = \begin{bmatrix} 0.6 & 0.3 & 0.1 & 0 \\ 0.2 & 0.6 & 0.2 & 0 \\ 0 & 0.1 & 0.4 & 0.5 \\ 0 & 0.1 & 0.3 & 0.6 \end{bmatrix}$$

*This work was performed while the authors were with the Civil Engineering Department, University of Maryland.

Gain of an Ergodic Process

The gain g of an ergodic process can be found from

$$g = \sum_{i=1}^{N} \pi_i q_i$$

where q_i is the expected immediate return in state i and π_i is the steady-state probability of state i. The gain can be visualized as the return per transition of the process.

For the Baltimore Harbor Tunnel case, we get π = (0.33, 0.67) and the gain of the process

$$g = \sum_{i=1}^{2} \pi_i q_i$$

which equals 13.55 vehicles/stage. This result yields total rewards through use of the formula

$$v_i(n) = ng + v_i$$

that are in close accordance with those obtained from the recurrence relationship.

It has been demonstrated how the behavior of the merging area (in its present configuration) can be analyzed. A methodology will now be explored for determining the relative worth of alternate configurations.

ALTERNATE MERGING AREA CONFIGURATIONS

Consider, as an example, the addition of a cone line between lanes 2 and 3 (Fig. 3). Assume that, for this alternative, the area's descriptive parameters become

$$p_{ij} = \begin{bmatrix} 0.55 & 0.45 \\ 0.25 & 0.75 \end{bmatrix} \qquad R_{ij} = \begin{bmatrix} 23 & 22 \\ 20 & 9 \end{bmatrix}$$

which yields

$$q = \begin{bmatrix} 22.55 \\ 11.75 \end{bmatrix}$$

and π = (0.357, 0.643), which gives

$$g = \sum_{i=1}^{2} \pi_i q_i$$

which equals 15.61 vehicles/stage. In this example, we have shown that the hypothetical alternative cone placement would allow a higher rate of vehicular flow.

Up to this point the analysis of permanent changes in the operation of the merging area has been discussed. A methodology for optimizing area traffic behavior through a real-time control procedure is now developed.

ALTERNATE REAL-TIME OPERATING PROCEDURES

Consider, again, the geometric alignment of the area described for the Baltimore Harbor Tunnel. Assume that, in addition to operating the system in this "uncontrolled"

Figure 1. General flow-density relationship.

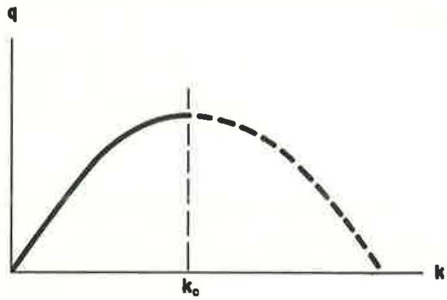


Figure 2. Existing layout of Baltimore Harbor Tunnel merging area.

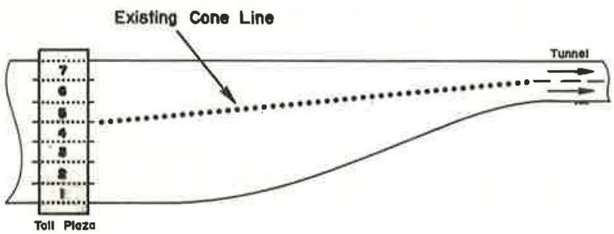
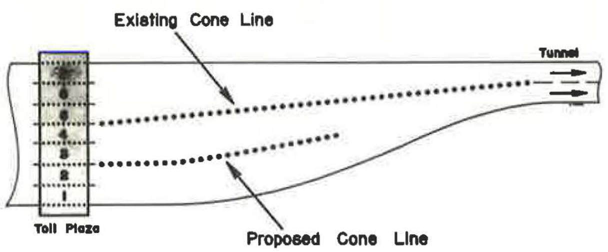


Figure 3. Hypothetical alternative layout of Baltimore Harbor Tunnel merging area.



These transitions may be defined in discrete time (e.g., 15 sec, 30 sec, or 1 min) intervals or as a continuous process (in which case they become transition rates).

Associated with each transition there is a "reward." For purposes of study, the reward is defined as the vehicular flow q measured at the outlet of the merging area. Then a reward matrix (similar in form to the transition matrix) can be constructed, as shown below, in which the elements are changes in q resulting from the system transitions from i to j.

$$R_{ij} = \begin{bmatrix} 10 & 15 & 6 & 3 \\ 12 & 20 & 8 & 5 \\ 8 & 12 & 5 & 4 \\ 6 & 8 & 4 & 2 \end{bmatrix}$$

The general relationship between density and flow is shown in Figure 1 (1, 2).

An objective is to analyze the behavior of this system over time, measured by the number of stages, or state transitions n, that occur. If we assume that the transition matrix is dependent only on the present state of the system, it can be analyzed as a Markov process (3):

$$\sum_{j=1}^{N} p_{ij} = 1$$

and $0 \leq p_{ij} \leq 1$.

Using Markovian decision theory, we can compute various characteristics of the process (4). One characteristic of particular interest is the expected reward from a set of staged decisions, given a starting point in time. In a more sophisticated analysis, it is possible to change both the transition and reward matrices by specifying a set of traffic control alternatives. Then each control alternative has transition and reward matrices associated with it. Given a performance objective (e.g., maximize flow over a period of time), it is possible to define an optimum policy, i.e., the set of optimal traffic control decisions at each stage of the process for each possible state of the system.

If one further assumes that the process is completely ergodic (i.e., after it has been operational for a long time, the probability of the system being in any given state is independent of its starting state), then the long-term average earnings per unit time, defined as the gain, can be found. The optimum policy (set of decisions) is that that maximizes gain.

Application of this optimization technique to a real-time control system in other than the trivial case requires that a decision optimizing immediate return does not necessarily optimize long-range return. It is obvious, if this condition is not satisfied, that the stream of rewards from a series of decision stages could be optimized by simply selecting the decision that optimizes the return at each immediate stage (5).

APPLICATION OF THE MODEL TO THE BALTIMORE HARBOR TUNNEL MERGING AREA

The merging area of the Baltimore Harbor Tunnel (Fig. 2) requires cars to merge into two lanes within a distance of approximately 500 ft. Under present operating discipline, vehicles entering the area in lanes 5, 6, and 7 are separated from those in lanes 1 through 4 by a cone line.

Preliminary data indicate that the behavior of the area can be approximated by the relationship shown in Figure 1, where q is the traffic flow in vehicles per transition stage interval (taken to be 30 sec) and k is the vehicular density in vehicles/ft.

For an elementary two-stage Markovian model, state one exists when k is less than k_c and state two exists when k is greater than or equal to k_c, which results in the following hypothetical matrices:

$$p_{ij} = \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix} \qquad R_{ij} = \begin{bmatrix} 24 & 21 \\ 21 & 6 \end{bmatrix}$$

Expected Reward of a Policy

The expected reward $v_i(n)$ from a set of staged decisions (policy), given a starting point i, is defined by the recurrence relationship

$$v_i(n) = \sum_{j=1}^{N} p_{ij} \left[r_{ij} + v_j(n-1)\right] \qquad i = 1, 2, \ldots N, \text{ and } n = 1, 2, \ldots$$

By defining q_i, the expected reward from the next stage transition, given the starting state i,

$$q_i = \sum_{j=1}^{N} p_{ij} r_{ij} \qquad i = 1, 2, \ldots N$$

we can write the recurrence relationship in the form

$$v_i(n) = q_i + \sum_{j=1}^{N} p_{ij} v_j (n - 1) \qquad i = 1, 2, \ldots N, \text{ and } n = 1, 2, \ldots$$

As an example, suppose our problem contained two states, with matrices

$$R = \begin{bmatrix} 9 & 3 \\ 3 & -7 \end{bmatrix} \qquad p = \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.5 \end{bmatrix}$$

Then, after we have computed

$$q = \begin{bmatrix} 6 \\ -3 \end{bmatrix}$$

the recurrence relationship can be used to derive the following values:

n	$v_1(n)$	$v_2(n)$
0	0	0
1	6	-3
2	7.5	-2.4
3	8.55	-1.44
4	9.555	-0.444
5	10.5555	0.5556

Therefore, using this recurrence method for the Baltimore Harbor Tunnel case gives the following total expected reward:

n	$v_1(n)$	$v_2(n)$
0	0	0
1	22.8	9.0
2	40.08	20.76
3	55.15	33.63
4	69.34	46.93
5	83.17	60.41
6	96.86	73.96

mode, the option exists at each transition point in time to control density by some traffic operations technique (e.g., metering of traffic input). The set of control decisions at each stage is desired.

For example, assume that behavior under the "control" alternative (denoted by the superscript 2) is described by

$$p_{ij} = \begin{bmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \end{bmatrix} \qquad R_{ij} = \begin{bmatrix} 22 & 21 \\ 21 & 4 \end{bmatrix}$$

The following table then summarizes the pertinent information:

i	a	p^a_{i1}	p^a_{i2}	R^a_{i1}	R^a_{i2}	q^a_i
1 ($k < k_c$)	1	0.6	0.4	24	21	22.8
	2	0.9	0.1	22	21	21.9
2 ($k > k_c$)	1	0.2	0.8	21	6	9.0
	2	0.6	0.4	21	4	14.2

By using the Markovian policy iteration method (see Appendix),[1] we can determine the set of staged decisions that will maximize flow through the area (4). This computation is presented below.

Step 1: Set $v_1 = v_2 = 0$ and enter the policy improvement routine.

Step 2: It will choose the decision that maximizes immediate returns, giving

$$\bar{d} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \bar{p} = \begin{bmatrix} 0.6 & 0.4 \\ 0.6 & 0.4 \end{bmatrix} \qquad \bar{q} = \begin{bmatrix} 22.8 \\ 14.2 \end{bmatrix}$$

Step 3: Entering the value determination routine gives

$$g + v_1 = 22.8 + 0.6v_1 + 0.4v_2$$

$$g + v_2 = 14.2 + 0.6v_1 + 0.4v_2$$

By setting $v_2 = 0$, we get $g = 19.36$, $v_1 = 8.6$, and $v_2 = 0$.

Step 4: Now, applying the policy improvement routine gives

i	a	$q^k_i + \sum_{j=1}^{2} p^k_{ij} v_j$
1	1	22.8 + 0.6(8.6) + 0.4(0) = 27.96
	2	21.9 + 0.9(8.6) + 0.1(0) = 29.64*
2	1	9.0 + 0.2(8.6) + 0.8(0) = 10.72
	2	14.2 + 0.6(8.6) + 0.4(0) = 19.36*

Step 5: This yields

$$\bar{d} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \qquad \bar{p} = \begin{bmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \end{bmatrix} \qquad \bar{q} = \begin{bmatrix} 21.9 \\ 14.2 \end{bmatrix}$$

Step 6: The value determinations are

$$g + v_1 = 21.9 + 0.9v_1 + 0.1v_2$$

$$g + v_2 = 14.2 + 0.6v_1 + 0.4v_2$$

[1] The original manuscript included an appendix entitled Markov Policy Iteration Method available in Xerox form at the cost of reproduction. When ordering, refer to XS-46, Highway Research Record 456.

Step 7: This yields $g = 20.8$, $v_1 = 11.00$, and $v_2 = 0$.
Step 8: Again, by applying policy improvement, we get

i	a	$q_i^k + \sum_{j=1}^{2} p_{ij}^k v_j$
1	1	$22.8 + 0.6(11.00) + 0.4(0.0) = 29.4$
	2	$21.9 + 0.9(11.00) + 0.1(0.0) = 31.8^*$
2	1	$9.0 + 0.2(11.00) + 0.8(0.0) = 11.2$
	2	$14.2 + 0.6(11.00) + 0.4(0.0) = 20.8^*$

Step 9: This gives

$$\bar{d} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

for the second consecutive time, identifying it as the optimal decision. The optimal policy, then, is to choose alternative 2 (employ the control alternative) at each decision point in time.

CONCLUSIONS

Efficient use of this technique requires data sufficient to permit formation of the transition and reward matrices. Preliminary efforts with twin time-lapse cameras (one recording flow and the other density) indicate that this method will provide suitable results.

One shortcoming of the technique is the fundamental assumption that the system under study operates as an ergodic Markov process, with present-state operation purely a function of the state of the system immediately prior to it and long-run operation independent of initial state. In addition, the appropriate definition of states is worthy of detailed study for each individual problem. An excessive number of states will yield a cumbersome and costly algorithm. Likewise, too few states will obscure the traffic flow phenomena at levels relevant for control considerations. Definition of states should be closely related to the sensitivity of the alternatives' ability to meaningfully alter the flow-density relationship. The choice of transition stage time is also an important consideration. Too lengthy a period might allow several important state changes to occur. Too short a period would result in increased cost of data acquisition and system operation. Preliminary operation of the traffic system discussed indicates that a stage length roughly equal to the time required to travel through the merging area at periods of moderate flow, that of 30 sec, is best. In general, the transition stage time chosen for study should relate to realistic needs for monitoring the system and allow for stabilization of short-term perturbations resulting from employment of a control alternative.

In any system where use of this evaluation approach is considered, the data collection phase should include checking to see how closely the system under study operates as a Markov process. Once this assumption is met, a wide range of geometric and control alternatives can be tested for any traffic situation that can be characterized by a flow-density relationship.

ACKNOWLEDGMENTS

The authors wish to acknowledge E. C. Carter, J. W. Hall, and S. Palaniswamy for their helpful discussion and suggestions on analytic approaches to the problem and F. Arzt for data reduction efforts.

Support from the Federal Highway Administration and the Maryland State Department of Transportation is gratefully acknowledged. The opinions, findings, and conclusions expressed in this paper are those of the authors and not necessarily those of the Maryland State Department of Transportation or the Federal Highway Administration.

REFERENCES

1. An Introduction to Traffic Flow Theory. HRB Spec. Rept. 79, 1964.
2. Drew, D. R. Traffic Flow Theory and Control. McGraw-Hill, New York, 1968.
3. Feller, W. An Introduction to Probability Theory and Its Application. John Wiley and Sons, New York, 1957.
4. Howard, R. A. Dynamic Programming and Markov Processes. M.I.T. Press, 1960.
5. Bellman, R. Dynamic Programming. Princeton Univ. Press, 1957.