

## Arrival Rate Identification for a Class of Traffic Signal Control Problem

Xiao-Hua Yu\*, Allen R. Stubberud\*\*

\*Department of Electrical Engineering

California Polytechnic State University, San Luis Obispo, CA 93407

\*\*Department of Electrical and Computer Engineering

University of California, Irvine, CA 92697

### ABSTRACT

Setting signals at traffic intersections to reduce congestion is one of the most challenging problems in traffic management. To find the optimal control strategy, specific information of the traffic flows passing through intersections, such as vehicle arrival rates (number of vehicles per hour), must be provided in advance. In most control approaches, this parameter is assumed to be a known constant; however, for an on-line adaptive control in real-time, when this information is not available, or when it fluctuates around its nominal value, parameter estimation/identification becomes crucial.

It has been shown that the Markovian decision control theory can be successfully applied to solve traffic signal control problems, when both the state transition probabilities and the one-step reward function are known. For a class of controlled Markov processes in which each state transition probability is a function of an unknown parameter, an on-line estimation algorithm needs to be developed to identify the unknown parameter first; then an optimal adaptive control law can be generated to maximize the long-term total expected reward based on this estimate. In this case, the choice of the feedback control law "interacts" with the parameter identification, which is also known as the "dual" aspect for adaptive control.

In this paper, an on-line parameter

identification algorithm is investigated for adaptive Markovian decision control at an isolated traffic intersection with unknown vehicle arrival rates. Section 1 gives a brief introduction to Markovian control processes and a maximum likelihood estimation algorithm. Section 2 discusses the traffic dynamic equations and the adaptive Markovian decision control model for an isolated traffic intersection. The proposed algorithm is then tested by computer simulation and the result is shown in section 3.

**Keywords:** Markov decision control, parameter estimation, adaptive control, traffic signal control, maximum likelihood estimation.

## I. MARKOVIAN ADAPTIVE CONTROL PROCESS

### 1.1. Introduction

A discrete, stationary, Markov control model (also known as a Markov decision process or Markov dynamic programming) is defined on  $(X, A, P, R)$  where:

(a)  $X$ , a Borel space, is the state space and every element in the space  $x \in X$  is called a state;

(b)  $A$ , also a Borel space, is defined as the set of all possible controls (or alternatives). Each state  $x \in X$  is associated with a non-empty measurable subset  $A(x)$  of  $A$  whose elements are the admissible controls when the system is in state  $x$ ;

(c)  $P$ , a probability measure space in which an element  $p_{ij}^a$  denotes the transition probability from state  $i$  to state  $j$  under control  $a$ ; and

(d)  $R$ , a measurable function called a one-step (immediate) reward [1].

Choosing a particular alternative (control) in a Markov process results in an immediate reward and a transition to the next state. The total expected discounted reward over an infinite period of time is defined as:

$$V \triangleq E \left[ \sum_{t=0}^{\infty} \beta^t r(x(t), a(t)) \right] \quad (1)$$

where  $r$  is the one-step immediate reward, and  $\beta$  ( $0 < \beta < 1$ ) is the discount factor. The optimal reward  $v^*$ , or the supremum (least upper bound) of  $V$ , is defined as:

$$v^*(x, a^*) = \sup_{a \in A} [V(x, a)] \quad (2)$$

If both the state transition probabilities and the reward function are known, the optimal reward is only a function of the optimal control  $a^*$ . This optimal reward can be obtained by solving a functional dynamic programming equation (or DPE):

$$v^* = T v^* \quad (3)$$

with initial condition  $v(0) = V_0$  and contraction operator  $T$ :

$$T v(x) = \max_{a \in A} \left[ \bar{r}(x, a) + \beta \sum_{j=1}^N v(x) p_{ij}^a \right] \quad (4)$$

where the expected one-step transition reward  $\bar{r}(x, a)$ , is defined as:

$$\bar{r}(x, a) = \sum_{j=1}^N r_{ij}^a p_{ij}^a \quad (5)$$

By using Banach's fixed-point theorem, the unique solution of the above DPE can be calculated iteratively by successive approximation:

$$v_n(x) = \max_{a \in A} \left[ \bar{r}(x, a) + \beta \sum_{j=1}^N v_{n-1}(x) p_{ij}^a \right] \quad (6)$$

## 1.2. Parameter Estimation for an Adaptive Markovian Control Process

For a class of adaptive Markov control problems in which the state transition probability is a function of a time-invariant parameter  $\alpha$ , the Markov control model can be defined on  $(X, A, P(\alpha), R)$ . The total expected discounted reward now becomes:

$$V \triangleq E \left[ \sum_{t=0}^{\infty} \beta^t r(x(t), a(t), \alpha) \right] \quad (7)$$

The optimal reward  $v^*$  also becomes a function of parameter  $\alpha$ :

$$v^*(x, a^*, \alpha) = \sup_{a \in A} [V(x, a, \alpha)] \quad (8)$$

A block diagram of the proposed adaptive Markov control strategy is shown in Fig 1. If the parameter  $\alpha$  is known, then the control signal  $a(t)$  can be generated by the controller, based on the current state observation  $x(t)$ , to maximize the total expected reward. This is the standard optimization problem as discussed in section 1.1. However, if  $\alpha$  is unknown, it must be estimated using the current and previous state observations before solving the optimization problem. An adaptive control signal  $a(t, \alpha)$  is then generated based on the estimated value of the parameter. Obviously, if the parameter estimation converges to the true value asymptotically as  $t \rightarrow \infty$ , the total expected reward of Eq. (8) approximates the optimal reward given in Eq. (2).

Borkar and Varaiya [2] showed that when the unknown parameter takes values from a finite set, the maximum likelihood estimate asymptotically converges to a value in the given finite set such that the closed-loop transition probabilities with the estimated value of the unknown parameter are identical to the transition probabilities with the true value.

Let's consider the case in which the element of the probability transition matrix contains both linear and nonlinear functions of the unknown parameter. Note that  $0 \leq p_{ij}^u \leq 1$ , where  $i$  and  $j$  are the indices of the probability matrix;  $u$  is the control signal.

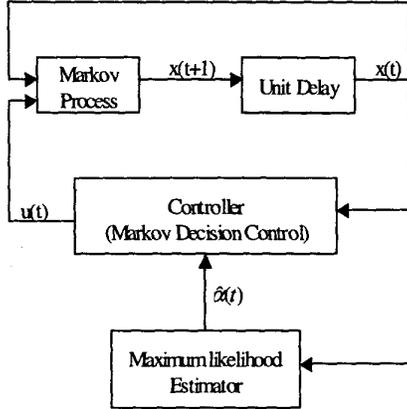


Figure 1. The Adaptive Markov Decision Control Process

The maximum likelihood function can be defined as a function of the unknown parameter  $\alpha$  which can be obtained from the joint probability of the observations  $x_0, x_1, \dots, x_L$  (where  $L$  is also called the "length" of data set):

$$J(\alpha; x_0, x_1, \dots, x_L) = \prod_{t=0}^{L-1} P_{x_t, x_{t+1}}(t, \alpha) \quad (9)$$

If we take logarithms on both sides, then:

$$\begin{aligned} \log J(\alpha; x_0, x_1, \dots, x_L) &= \log \prod_{t=0}^{L-1} P_{x_t, x_{t+1}}(t, \alpha) \\ &= \sum_{t=0}^{L-1} \log [P_{x_t, x_{t+1}}(t, \alpha)] \end{aligned} \quad (10)$$

The maximum value of the likelihood function can be found by setting its gradient (with respect to  $\alpha$ ) to be 0:

$$\left. \frac{d[\log J(\alpha; x_0, x_1, \dots, x_L)]}{d\alpha} \right|_{\hat{\alpha}_L} = 0 \quad (11)$$

that is,

$$\sum_{t=0}^{L-1} \nabla[\log P_{x_t, x_{t+1}}(t, \hat{\alpha}_L)] = 0 \quad (12)$$

where  $\nabla(\cdot)$  is the gradient and  $\hat{\alpha}_L$  is the estimate after  $(L-1)$  state transitions. The maximum likelihood estimate at the next transition also satisfies:

$$\sum_{t=0}^L \nabla[\log P_{x_t, x_{t+1}}(t, \hat{\alpha}_{L+1})] = 0 \quad (13)$$

Applying a Taylor series expansion to Eq. (13), we have:

$$\begin{aligned} &\sum_{t=0}^L \nabla[\log P_{x_t, x_{t+1}}(t, \hat{\alpha}_{L+1})] \\ &= \sum_{t=0}^L \nabla[\log P_{x_t, x_{t+1}}(t, \hat{\alpha}_L)] \\ &+ \left( \sum_{t=0}^L \nabla^2[\log P_{x_t, x_{t+1}}(t, \hat{\alpha}_L)] \right) (\hat{\alpha}_{L+1} - \hat{\alpha}_L) \end{aligned} \quad (14)$$

where  $\nabla^2(\cdot)$  denotes the second order derivative. Consider Eqs. (12), (13) and (14), the parameter estimation after the  $N$ -th state transition can be updated as:

$$\hat{\alpha}_{L+1} = \hat{\alpha}_L - \left( \sum_{t=0}^L \nabla^2[\log P_{x_t, x_{t+1}}(t, \hat{\alpha}_L)] \right)^{-1} \nabla[\log P_{x_t, x_{t+1}}(t, \hat{\alpha}_L)] \quad (15)$$

when  $\left( \sum_{t=0}^L \nabla^2[\log P_{x_t, x_{t+1}}(t, \hat{\alpha}_L)] \right)^{-1}$  exists. A step size  $\gamma$  can be included in Eq. (15) for faster convergence:

$$\hat{\alpha}_{L+1} = \hat{\alpha}_L - \gamma \left( \sum_{t=0}^L \nabla^2[\log P_{x_t, x_{t+1}}(t, \hat{\alpha}_L)] \right)^{-1} \nabla[\log P_{x_t, x_{t+1}}(t, \hat{\alpha}_L)] \quad (16)$$

Note that the convergence condition for this algorithm is that the initial estimation of the parameter is not too far away from its true value so that the remainder of Taylor expansion is small. In the next session, this algorithm will be applied to estimate the vehicle arrival rate on-line at a traffic intersection when this arrival rate fluctuates

around its nominal value.

## II. MARKOVIAN CONTROL FOR AN ISOLATED TRAFFIC INTERSECTION

### 2.1. The Dynamic Model of an Isolated Traffic Intersection

A typical four-legged intersection is shown in Fig. 2. Assume that the traffic flows move along two directions, i.e., north/south (denoted by 1) and east/west (denoted by 2); and it is sampled every  $\Delta t$  time interval with the discrete time index,  $k$ . The relationship between the current queue length  $q(k)$ , the queue length at the previous time instant  $q(k-1)$ , the input  $q_{in}(k)$  and the output  $q_{out}(k)$  during time interval  $[k-1, k)$ , can be written as:

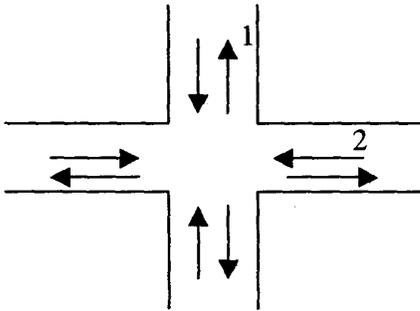


Figure 2. An isolated traffic intersection with through movement only

$$q(k) = q(k-1) + q_{in}(k) - q_{out}(k) \quad (17)$$

where  $q_{out}(k)$  can be further expressed as a nonlinear function:

$$q_{out}(k) = f_{out}(u(k), q(k), \Delta t, h_{min}, t_{hr}) \quad (18)$$

and

$$f_{out}^j(u^j(k), q^j(k), \Delta t, h_{min}, t_{hr}) = \begin{cases} \min \left[ q^j(k), \frac{\Delta t}{(h_{min} + t_{hr})} \right] & \text{when } u^j(k) = \text{Green} \\ 0 & \text{when } u^j(k) = \text{Red} \end{cases} \quad (19)$$

where  $h_{min}$  is the minimum headway (headway is the time or distance spacing between two successive vehicles in a given traffic lane measured from front to front);  $t_{hr}$  is the human reaction time to the green signal;  $u^j(k)$  is the control signal for the  $j$ -th movement ( $j = 1, 2$ ) and similarly,  $q^j(k)$  is the queue for the  $j$ -th movement at time instant  $k$ .

### 2.2. The Markovian Control Model

A state space  $X$  and a probability measure  $P$  must be defined in order to apply the above adaptive Markovian control theory to traffic systems. Since the queue length is the state variable in the traffic dynamics equation, one may want to choose the number of vehicles to be the state of the Markov control model. However, the resulting total number of states is very large. In order to reduce the number of states (and thus reduce both the computational time and memory space), a threshold (number of vehicles) is chosen for the queue of each movement at an intersection. If the queue length of a specific movement is greater than the threshold value, then this movement is defined in the congested mode; otherwise it is in the non-congested mode. These two modes (congestion/non-congestion) are defined as the two states in the binary state space  $X$ .

The state space is discrete, thus the probability measure  $P$  defines a discrete transition law. In the traffic control problem, the probability matrix  $P$  is time varying due to the time-varying traffic flow,

therefore:

$$\underline{p}(k) = f_p[\underline{q}(k), \Delta\hat{q}(k+1), \underline{u}(k)] \quad (20)$$

where  $\underline{q}(k)$  is the current queue,  $\Delta\hat{q}(k+1)$  is the estimated number of arrivals in the next time interval, and  $\underline{u}(k)$  is the control signal. The probability matrix can be further specified based on different arrival patterns. Under most circumstances, the arrival of vehicles at an isolated intersection follows the Poisson distribution, i.e.,

$$p(n=k) = \frac{(\lambda \Delta t)^k e^{-\lambda \Delta t}}{k!} \quad (21)$$

where  $n = 1, 2, \dots$ ;  $\lambda$  is the arrival rate and  $\Delta t$  is the time interval. Assuming that at a specific time instant there are  $q_g$  vehicles passing through the intersection if the signal of this direction is green, then for each movement  $j$ :

$$p_{X^j \rightarrow N}^{u^j} = p(\Delta\hat{q}^j + q^j - \delta(u^j)q_g^j \leq q_{threshold}^j) \quad (22)$$

and

$$p_{X^j \rightarrow C}^{u^j} = 1 - p_{X^j \rightarrow N}^{u^j} \quad (23)$$

where

$$\delta(u^j) = \begin{cases} 1, & \text{when } u^j = G \\ 0, & \text{Otherwise} \end{cases} \quad (24)$$

and  $X^j = N$  or  $C$  is the current state ( $N$  for non-congestion and  $C$  for congestion);  $u^j = G$  or  $R$  is the control signal ( $G$  for green signal and  $R$  for red signal). Two special cases are noted that:

$$p_{C \rightarrow C}^R = 1, \text{ and } p_{C \rightarrow N}^R = 0. \quad (25)$$

The reward matrix  $R$  has the same dimension and a definition similar to that of the probability matrix. The control objective herein is to minimize the queue length, so the functions of queue length corresponding to different states are chosen to generate the reward matrix:

$$R_{X^1, X^2}^{u^j} = f_u(q^j, q_{threshold}^j, u^j) \quad (26)$$

### 2.3. Parameter Estimation and Adaptive Control

In the above discussion, it is assumed that the vehicle arrival rates (for both directions) at the intersection are known constants. When the vehicle arrival rates are unknown, or when they fluctuate around their nominal values, it is necessary to estimate these parameters on line to make our adaptive controls more accurate. The maximum likelihood estimation algorithm discussed in section 1.2 can be employed:

$$\left. \frac{d[\log J(\lambda; x_0, x_1, \dots, x_L)]}{d\lambda} \right|_{\hat{\lambda}_L} = 0 \quad (27)$$

Consider Eq. (21), we have:

$$\frac{\partial p(n=k)}{\partial \lambda} = \frac{(\lambda \Delta t)^{k-1} \Delta t e^{-\lambda \Delta t}}{k!} (k - \lambda \Delta t) \quad (28)$$

$$\begin{aligned} \frac{\partial^2 p(n=k)}{\partial \lambda^2} &= \frac{(\lambda \Delta t)^{k-2} (\Delta t)^2 e^{-\lambda \Delta t}}{k!} [(\lambda \Delta t)^2 - 2(\lambda \Delta t)k + k(k-1)] \\ & \quad (29) \end{aligned}$$

Assume the four traffic movements at intersection are independent, then:

$$\begin{aligned} \nabla[\log P_{x_t, x_{t+1}}(t, \hat{\lambda}_L)] &= [1/P_{x_t, x_{t+1}}(t, \hat{\lambda}_L)] \sum_{i=1}^N \frac{\partial p_i}{\partial \lambda} \\ & \quad (30) \\ \nabla^2[\log P_{x_t, x_{t+1}}(t, \hat{\lambda}_L)] &= -[1/P_{x_t, x_{t+1}}(t, \hat{\lambda}_L)^2] \sum_{i=1}^N \frac{\partial p_i}{\partial \lambda_i} \\ & \quad + [1/P_{x_t, x_{t+1}}(t, \hat{\lambda}_L)] \sum_{i=1}^N \frac{\partial^2 p_i}{\partial \lambda_i^2} \quad (31) \end{aligned}$$

where  $N=4$ . Finally, the vehicle arrival rate can be estimated by:

$$\begin{aligned} \hat{\lambda}_{L+1} &= \hat{\lambda}_L - \gamma \left( \sum_{i=0}^L \nabla^2[\log P_{x_t, x_{t+1}}(t, \hat{\lambda}_L)] \right)^{-1} \\ & \quad \nabla[\log P_{x_t, x_{t+1}}(t, \hat{\lambda}_L)] \quad (32) \end{aligned}$$

A set of optimal traffic control signals will be obtained using Eq.(1) to Eq.(5) based on this estimated value of vehicle

arrival rate. The flow chart for this adaptive control algorithm is shown in Fig. 3.

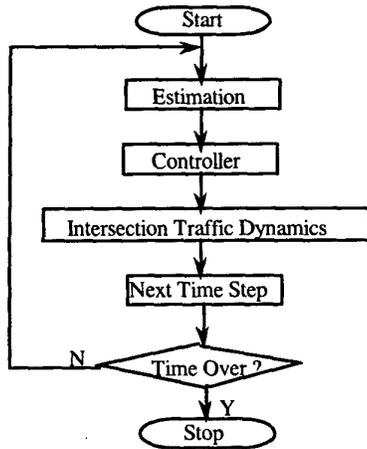


Fig. 3. Flow chart for adaptive Markovian decision control at a traffic intersection

### III. SIMULATION RESULT

The proposed adaptive control algorithm with on-line parameter identification is tested by computer simulation. Assume that at an isolated intersection, the traffic flows move along two directions (as shown in Fig. 2) with two sets of traffic control signals (green for east/west or green for north/south). Since there are four traffic movements under consideration, the total number of states for this intersection is 16. The arrival vehicles are assumed to follow Poisson distribution.

The following figure illustrates the result when the initial value of the arrival rate is 350 (vehicle/hour) while the actual arrival rate is 400 (vehicle/hour). The x-axis shows the time (in seconds) and the y-axis shows the estimated arrival rate. The estimated value approaches to the true value in 200 seconds, with the steady state error of 1.1%.

### IV. CONCLUSION

In this paper, an on-line parameter identification algorithm based on maximum likelihood estimation for adaptive Markov control theory is discussed and applied to estimate the vehicle arrival rates at an isolated traffic intersection. Further evaluation and testing on this algorithm will be conducted.

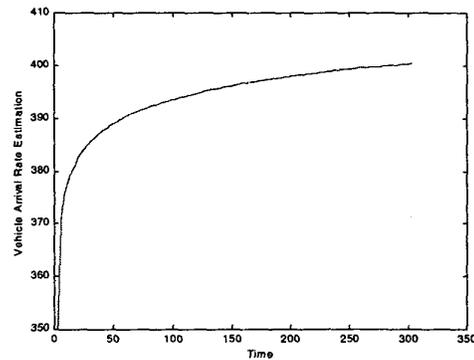


Fig. 4. Computer Simulation Result

### REFERENCES

- [1] O. Hernandez-Lerma, *Adaptive Markov control process*, applied mathematical sciences 7, New York: Springer-Verlag, 1989.
- [2] V. Borkar and P. Varaiya, "Adaptive control of Markov chains, I: finite parameter set", *IEEE Transaction on Automatic Control*, Vol. AC-24, No.6, 1979, pp. 953-957.
- [3] X.-H. Yu, and A. Stubberud, "Markovian decision control for traffic signal systems", *Proceedings of the 36th IEEE conference on decision and control*, 1997.
- [4] R. Wilshire, R. Black, et al., *Traffic control systems handbook*, FHWA-IP-85-12, 1985