An Integrated Model for Signalized Traffic Intersection Control

Xiao-Hua Yu, *Member, IEEE*

Abstract— **Traffic signal control is an effective way to regulate traffic flow to avoid conflict and reduce congestions. This research investigates a real-time traffic signal control system that integrates a traffic flow prediction model and an adaptive control scheme based on dynamic programming with rolling horizon. The proposed approach estimates the parameter of the arriving traffic flow at the intersection, predicts the state transition probabilities, and then formulates the traffic signal control problem as a decision-making problem of a stochastic system. Two different traffic arrival patterns are considered, including the normal distribution and the Poission distribution.**

I. INTRODUCTION

WITH the ever-increasing traffic demand, congestion has become a serious problem in many major cities around the world. ATMS (advanced traffic management system) is a systematic effort toward the design of an integrated transportation system with new technologies. By regulating the traffic demand at each intersection in the network, the goal is to avoid traffic conflict and shorten the queue length at a stop line.

At a signalized intersection, traffic signals typically operate in one of three different control modes, namely, pre-timed control, semi-actuated control and fully actuated control. Pre-timed control is an open-loop control strategy, in which all the control parameters are fixed and pre-set off-line. It is easy to implement and is well suited for predictable traffic pattern. In actuated control, the control signal is adjusted in accordance with real-time traffic demand obtained from detectors. In general, actuated control performs better than the pre-timed control.

Traffic signal control problem has been studied by many researchers over the years. Some major conventional traffic signal control systems, such as TRANSYT (traffic network study tool) [1], SCOOT (split, cycle and offset optimization technique) [2], and SCATS (Sydney coordinated adaptive traffic system) [3], select the best pre-calculated off-line timing plan based the current traffic conditions on the road. Some recent development on traffic signal control employs artificial intelligent technology, such as neural networks [4] and fuzzy logic [5]. Algorithms using Petri nets [6] and Markov decision control [7] are also investigated.

Markov decision control has been employed to analyze and control many complicated stochastic systems in various areas. In [7], it is shown that when both the state transition probabilities and the one-step reward function are known,

the Markov decision control theory can be successfully applied to solve traffic signal control problem.

A discrete, stationary, Markov control model (also known as a Markov decision process or Markov dynamic programming) is defined on a state space (X, A, P, R) where X, a Borel space, is the state space and every element in the space $x \in X$ is called a state. A is the set of all possible controls (or alternatives), and is also a Borel space. 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. P is a probability measure space in which an element p_{ij}^a denotes the transition probability from state i to state j under control a. Finally, R represents a measurable function called a one-step (immediate) reward [8].

Choosing a particular alternative (control) in a Markov process results in an immediate reward and a transition to the next state. The expected one-step transition reward $\bar{r}(x, a)$, is defined as:

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

If both the state transition probabilities and the reward function are known, then the optimal reward v*, or the supremum (least upper bound) of the total expected discounted reward V, can be obtained by solving a functional dynamic programming equation (or DPE):

$$v* = Tv* \qquad (2)$$

The contraction operator T is defined as:

$$Tv(x) = \max_{a \in A} \left[\bar{r}(x, a) + \beta \sum_{j=1}^{N} v(x) p_{ij}^a \right] \qquad (3)$$

where β ($0 < \beta < 1$) is the discount factor. By using Banach's fixed-point theorem, the unique solution of the above DPE can be calculated iteratively by successive approximation:

$$v_n(x) = \max_{a \in A} \left[\bar{r}(x, a) + \beta \sum_{j=1}^{N} v_{n-1}(x) p_{ij}^a \right] \qquad (4)$$

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 need to be developed to identify the unknown parameter. An optimal adaptive control law can then be generated to maximize the long-term total expected reward based on this estimation. Borkar and Varaiya [9] 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. In this research, we propose an integrated adaptive control model based maximum likelihood estimation/prediction and Markov decision control theory.

Let's consider the case in which every element of the probability transition matrix (which may contain both linear and nonlinear functions of the unknown parameter) is bounded (i.e., $0 \le p_{ij}^u \le 1$, where i, j is the index of the probability matrix; u is the control signal). The maximum likelihood function can be defined as a function of the unknown parameter α which can be obtained from the joint probability of the observations x_0, x_1, …, x_L (where L is also called the "length" of data set):

$$J(\alpha; x_0, x_1, \cdots, x_L) = \prod_{t=0}^{L-1} P_{x_t x_{t+1}}(t, \alpha) \qquad (5)$$

If we take logarithms on both sides, and set its gradient (with respect to α) to 0 to find the maximum value of the likelihood function, i.e.,:

$$\sum_{t=0}^{L-1} \nabla[\log P_{x_t x_{t+1}}(t, \hat{\alpha}_L)] = 0 \qquad (6)$$

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:

$$\sum_{t=0}^{L} \nabla[\log P_{x_t x_{t+1}}(t, \hat{\alpha}_{L+1})] = 0 \qquad (7)$$

Applying a Taylor series expansion to (7), we have:

$$\sum_{t=0}^{L} \nabla[\log P_{x_t x_{t+1}}(t, \hat{\alpha}_{L+1})]$$
$$= \sum_{t=0}^{L} \nabla[\log P_{x_t x_{t+1}}(t, \hat{\alpha}_L)]$$
$$+ \left(\sum_{t=0}^{L} \nabla^2[\log P_{x_t x_{t+1}}(t, \hat{\alpha}_L)] \right)(\hat{\alpha}_{L+1} - \hat{\alpha}_L) \qquad (8)$$

where $\nabla^2(\cdot)$ denotes the second order derivative. Consider (6), (7) and (8), and include a step size γ for faster convergence, the parameter estimation after the N-th state transition can be updated as:

$$\hat{\alpha}_{L+1} = \hat{\alpha}_L - \gamma \left(\sum_{t=0}^{L} \nabla^2[\log P_{x_t x_{t+1}}(t, \hat{\alpha}_L)] \right)^{-1} \cdot$$
$$\nabla[\log P_{x_t x_{t+1}}(t, \hat{\alpha}_L)] \qquad (9)$$

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

II. THE MARKOV MODEL FOR TRAFFIC SIGNAL CONTROL PROBLEM

A state space X and a probability measure P must be defined in order to apply the Markov 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.

Assume that at a specific time instant, there are q_g vehicles passing through the intersection if the signal of this direction is green. For each movement j (j = 1, 2, …, 8), the state transition probability can be written as:

$$p_{X^j -> N}^{u^j} = p(\Delta \hat{q}^j + q^j - \delta(u^j)q_g^j \le q_{threshold}^j) \qquad (10)$$

and

$$p_{X^j -> C}^{u^j} = 1 - p_{X^j -> N}^{u^j} \qquad (11)$$

where

$$\delta(u^j) = \begin{cases} 1, & when \ u^j = G \\ 0, & Otherwise \end{cases} \qquad (12)$$

In (10) and (11), 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:

$$p_{C->C}^R = 1, \text{ and } p_{C->N}^R = 0. \qquad (13)$$

The probability matrix can be further specified based on various arrival patterns. In this research, we consider two different situations, i.e., the normal distribution and the Poission distribution.

A. The Normal Distribution

When the arrival of vehicles follows the normal distribution, the probability density function can be written as:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (14)$$

Taking integration, the probability is obtained:

$$p(X \le x_1) = F(x_1) = \int_{-\infty}^{x_1} f(x)dx \qquad (15)$$

If the mean of normal distribution is unknown or the traffic flow fluctuates around its nominal value, we need to apply the maximum likelihood estimation. The first order derivative can be calculated as:

$$\frac{dp}{d\mu} = \frac{d}{d\mu} \left[\int_{-\infty}^{x_1} f(x)dx \right] \qquad (16)$$

Exchange the order of operations, we have:

$$\frac{dp}{d\mu} = \int_{-\infty}^{x_1} \frac{d}{d\mu} \left[\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right] dx$$
$$= \frac{1}{2\sqrt{2\pi}\sigma^3} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \qquad (17)$$

The second order derivative for normal distribution:

$$\frac{d^2 p}{d\mu^2} = \frac{(x_1 - \mu)}{\sqrt{2\pi}\sigma^5} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \qquad (18)$$

Finally, the recursive algorithm for estimation:

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

$$\nabla[\log P_{x_t x_{t+1}}(t, \hat{\mu}_L)] \qquad (19)$$

B. The Poission Distribution

If the arrival of vehicles follows the Poisson distribution, we have:

$$p(n = k) = \frac{(\lambda \Delta t)^k e^{-\lambda \Delta t}}{k!} \qquad (20)$$

where n = 1, 2, ...; λ is the arrival rate and Δt is the time interval. For the Maximum likelihood estimation algorithm, we have:

$$\left. \frac{d[\log J(\lambda; x_0, x_1, \cdots, x_L)]}{d\lambda} \right|_{\hat{\lambda}_L} = 0 \qquad (21)$$

Consider (20), we have:

$$\frac{dp(n=k)}{d\lambda} = \frac{(\lambda \Delta t)^{k-1} \Delta t e^{-\lambda \Delta t}}{k!} (k - \lambda \Delta t) \qquad (22)$$

$$\frac{d^2 p(n=k)}{d\lambda^2}$$
$$= \frac{(\lambda \Delta t)^{k-2} (\Delta t)^2 e^{-\lambda \Delta t}}{k!} [(\lambda \Delta t)^2 - 2(\lambda \Delta t)k + k(k-1)] \qquad (23)$$

Assume each traffic movements at intersection are independent, then:

$$\nabla[\log P_{x_t x_{t+1}}(t, \hat{\lambda}_L)] = [1 / P_{x_t x_{t+1}}(t, \hat{\lambda}_L)] \sum_{i=1}^{N} \frac{\partial p_i}{\partial \lambda} \qquad (24)$$

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

$$+ [1 / P_{x_t x_{t+1}}(t, \hat{\lambda}_L)] \sum_{i=1}^{N} \frac{\partial^2 p_i}{\partial \lambda_i^2} \qquad (25)$$

where N = 8. Finally, the vehicle arrival rate can be estimated by:

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

$$\nabla[\log P_{x_t x_{t+1}}(t, \hat{\lambda}_L)] \qquad (26)$$

III. ADAPTIVE CONTROL

As stated in the previous sections, the proposed adaptive controller consists of two modules, i.e., the estimation (or prediction) module and the control module. Being related to the current state of each traffic movement, both the probability measure and the reward function are time-varying in real-time traffic control. In this paper, the sampling time is chosen as the minimum green extension time, Δt. A rolling horizon approach is used to achieve a real-time adaptive control. Every Δt seconds, the P and R matrices are calculated; then a decision is made to choose the control signal for the next time interval based on the current measurements from detectors and our estimation. Once the optimal solution is found, it is implemented only for Δt seconds. At the next time step, the probability matrix and reward matrix are updated and the whole decision-making process is repeated.

IV. SIMULATION

The proposed adaptive control algorithm with on-line parameter identification is tested by computer simulation. For the sake of simplicity, let's assume that the traffic flows move along two directions (east/west or north/south) at an isolated intersection with two sets of traffic control signals (green for east/west or green for north/south). Assume that the intersection is "clear" when the simulation starts (i.e., zero initial conditions, or no queue at the beginning), and each traffic movement is independent. We choose the maximum green time to be 30 seconds, minimum arrival and departure headway to be 2 seconds, respectively. Loss time (human reaction time) is 0 second. The sampling rate, which is also the minimum green time, is chosen to be 3 seconds. Assume that the arrival pattern is Poission distribution. The test is performed when the initial value of the arrival rate is 400 (vehicle/hour) while the actual arrival rate is 450 (vehicle/hour). The estimated value approaches to the true value in 200 seconds, with the steady state error of 1.5%. By applying the adaptive control algorithm, we found that the total vehicle delay (for the intersection) is 186 seconds.

V. CONCLUSION

In this paper, an integrated model which combines an on-line parameter identification algorithm using maximum likelihood principle and an adaptive Markov decision control is investigated. The proposed algorithm is applied to the traffic signal control problem. Two different vehicle arrival patterns are considered here, including the normal distribution and the Poisson distribution. Further evaluation and testing on this approach will be performed.

REFERENCES

[1] Transportation research center, Traffic network study tool: TRANSYT-7F software summary, University of Florida, 1987.
[2] P. Hunt, and D. Robertson, "The SCOOT on-line traffic signal optimization technique", *Traffic engineering and control*, April 1982.
[3] P. Lowrie, "The Sydney coordinated adaptive control systems – principles, methodology, algorithms", IEE conference publication, vol. 207, 1982.

[4] S. Chien, Y. Ding, "Dynamic Bus Arrival Time Prediction with Artificial Neural Networks," *J. Trans. Engrg.*, vol. 128, 2002.

[5] W. Wei, et. al., "Traffic signal control using fuzzy logic and MOGA," *Proceedings of IEEE conference on systems, man, and cybernetics, 2001*.

[6] G. List, and M. Cetin, "Modeling traffic signal control using Petri nets", IEEE Transaction on intelligent transportation systems, vol. 5, no. 3, 2004.

[7] X.-H. Yu, and A. Stubberud, "Markovian decision control for traffic signal systems", *Proceedings of the 36th IEEE conference on decision and control*, 1997.

[8] O. Hernandez-Lerma, *Adaptive Markov control process*, applied mathematical sciences 7, New York: Springer-Verlag, 1989.

[9] V. Borkar and P. Varaiya, "Adaptive control of Markov chains, I: finite parameter set", IEEE Transaction on Automatic Control, vol. 24, no.6, 1979, pp. 953-957.

[10] R. Wilshire, R. Black, et al., *Traffic control systems handbook*, FHWA-IP-85-12, 1985.