Queuing Models for Analysis of Traffic Adaptive Signal Control

Pitu B. Mirchandani, *Senior Member, IEEE*, and Ning Zou

***Abstract*—Microsimulation models are normally used to evaluate traffic-adaptive signal control systems. This paper develops an analytical approach for this evaluation based on queuing models. In particular, a queuing model is developed for a simplified adaptive control strategy that is based on rolling horizon scheme, in which a signal serves two movements alternatively. In this strategy, the first movement is served until the queue dissipates, then the second movement is served until the queue dissipates, then the signal goes back to serving the first movement, and this cyclic process repeats. A numerical algorithm is developed in a stochastic context to compute steady-state performance measures such as average delays and expected queue lengths. These results are compared with simulation-based results, and indeed, the analytically derived numerical method predicts well the simulation results.**

***Index Terms*—Adaptive control, busy period, delay, queuing modeling.**

I. Introduction

SEVERAL performance criteria can be used to evaluate a signalized intersection, such as average number of stops per vehicle, vehicle throughput, average queue length, and average stop delay. Among these indicators, delay is perhaps the most important since it is easily understood, related to travel costs and vehicle emission. For example, in [10], Highway Capacity Manual uses it as the only measurement to determine the level of service of signalized intersections.

Over the last 50 years, there have been many queuing models reported in the literature for analyzing delays and queues at a signalized intersection. The most well known model is that due to Webster [11], who came up with analytical expressions for delays and queues under the assumption of Poisson arrivals. Miller [4] and Newell [5] proposed analytical models to estimate residual queues that exist at the end of green phase at traffic signal on fixed-time plan. These are steady-state results that assume that the arrival pattern remains the same throughout a period large enough to assume steady-state conditions. Hurdle [2] has discussed the limitations of such steady-state models.

The Highway Capacity Manual [10] includes a generalized time-dependent queuing model for computing average delay for fixed-time signal control. Although the expressions for delays are similar to these from queuing models, they need to be calibrated for locations and regions, and in fact, Akcelik [1] has studied these and corresponding Australian and Canadian expressions and their calibration. Olszewski [7], [8] has also proposed a stochastic queuing model to analyze queue and delay distribution at intersections operating under fixed-time plans.

In this paper, a stochastic queuing model is developed for a two-phase simplified adaptive control strategy where the signal is switched to the other phase when the served queues just vanish, and after the queues next served are depleted, it switches back to the current phase. A numerical method is developed to obtain the steady-state distribution of delays, queues, etc., using this queuing model. These steady-state distributions are compared with those obtained through a simulation model. We note that such an adaptive control strategy was first presented and studied by Newell [6].

In Section II of this paper, the green phase distribution, or the conditional "busy period" of this adaptive control strategy, given multiple vehicles waiting at the beginning of green phase, is calculated using transform methods. The result shows that when k vehicles are queued up at the beginning of green phase, then the expected time to clear the queue is k times the busy period of the corresponding queuing system with same parameters. This result is used in Section III to build the queuing model for the adaptive control analyzed in this paper. In Section IV, numerical results are given and compared with simulation-based results. Several loading patterns are studied, including symmetric intersection with equal arrival rates and asymmetric intersection with unequal arrival rates. The results show that the analytical model predicts well the simulation results.

II. Busy Period Analysis

In a general first-in first-out single-server queuing system, a busy period begins when a customer arrives and finds no queue and the server is free to serve him immediately. A busy period ends when the server completes the service of a customer and finds there are no customers in queue waiting for service (i.e., it is a zero queue again). The interval between the end of a busy period and the beginning of the next constitutes an idle period. The intersection can be viewed as an $M/G/1$ system, where M means the arrival process is Poisson; G means the service time follows general distribution, and 1 means there is single server.

Manuscript received March 5, 2006; revised July 3, 2006, September 9, 2006, and September 14, 2006. This work was supported in part by Grant CMS-0231458 from the National Science Foundation and by the Intelligent Transportation System (ITS) Partnership Agreement with Federal Highway Administration (FHWA) and the Arizona Department of Transportation for the Advanced Traffic and Logistics: Algorithms and Systems (ATLAS) Research Center at the University of Arizona. The Associate Editor for this paper was B. de Schutter.

P. B. Mirchandani is with the Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721 USA (e-mail: pitu@sie.arizona.edu).

N. Zou is with the Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721 USA, and also with Kittelson and Associates, Inc., Tucson, AZ 85701 USA (e-mail: zoun@email.arizona.edu).

Digital Object Identifier 10.1109/TITS.2006.888619

Since arrivals are Poisson in an $M/G/1$ system, the idle period, which ends with the next arrival, has a negative exponential distribution. The distribution of the busy period of an $M/G/1$ system is available in many queuing theory textbooks (e.g., [3]).

In this paper, we study different types of busy periods that occur due to the competition of the green phases for the various vehicle streams approaching an intersection.

Assume vehicles arrive in a Poisson manner with rate λ. Let the service rate (or the saturation flow rate) of the intersection be μ. Let ν denote the time to serve one vehicle, its distribution as $b(\nu)$ with expectation $\overline{\nu} = 1/\mu$, and $B^e(s)$ is its s transform. Let q be the busy period when there is only one vehicle waiting at the beginning of the green phase. Define $f(q)$ as the distribution function of q and $F^e(s)$ as its s transform, which is defined as

$$F^e(s) = \int_0^\infty e^{-sq} f(q) dq. \tag{2.1}$$

The expectation of q is (e.g., [3])

$$E(q) = \frac{1/\mu}{1 - \lambda/\mu}. \tag{2.2}$$

We will now compute the expectation of busy period, say, q' when k vehicles are waiting at the beginning of green phase. Define $f(q')$ as the distribution function of q', $F'^e(s)$ is its s transform. The length of the busy period q' consists of the length of service $k\nu$ for the first k vehicles, plus the length of the busy period for the subsequently arriving customers. If n vehicles arrive during the period $k\nu$, then these n customers may be visualized as standing aside for service. The first of the n customers gets serviced immediately after the termination of service of the initial k vehicles; if there are any new arrivals during this service time, then without loss of generality in computing total service time, they may be considered to be served before another customer of the $(n-1)$ remaining in queue is admitted for service. The same modified priority rule may be applied when the next customer goes into service. Therefore, the total busy period is

$$q' = k\nu + q_1 + q_2 + \cdots + q_n \tag{2.3}$$

where the q_i are independent (since the arrival process is Poisson) and have the same distribution as q. Consequently, the conditional Laplace-transform of q', given there are n arrivals in time $k\nu$, is

$$F'^e(s/n, k\nu) = E[e^{-sq'}] = e^{-sk\nu} \left[F^e(s)\right]^n. \tag{2.4}$$

To find the unconditional s transform, we note that the unconditional distribution may be derived by

$$f(q') = \int_0^\infty \sum_{n=0}^\infty f(q'/n, k\nu) P\left[n/(k\nu)\right] b(\nu) d\nu \tag{2.5}$$

where

$$P\left[n/(k\nu)\right] = (\lambda k\nu)^n e^{-\lambda k\nu}/n!, \qquad n = 0, 1, 2, \ldots \tag{2.6}$$

and $b(\nu)$ is the service time distribution. The s transform of (2.5) gives

$$F'^e(s) = \int_0^\infty \sum_{n=0}^\infty F'^e(s/n, k\nu) P\left[n/(k\nu)\right] b(\nu) d\nu. \tag{2.7}$$

Using (2.4) and (2.6), and summing appropriate elements, we get

$$F'^e(s) = \int_0^\infty e^{-[ks + k\lambda - k\lambda F^e(s)]\nu} b(\nu) d\nu. \tag{2.8}$$

By definition of s transform, the above relation becomes

$$F'^e(s) = B^e\left[ks + k\lambda - k\lambda F^e(s)\right]. \tag{2.9}$$

By differentiation, the expected length of a busy period caused by the k vehicles in queue is

$$\begin{aligned} E(q') &= -\frac{d}{ds} F'^e(s)|_{s=0} \\ &= -\frac{d}{ds} B^e(0) * \frac{d}{ds}\left[ks + k\lambda - k\lambda F^e(s)\right]|_{s=0}. \end{aligned} \tag{2.10}$$

Substituting $\overline{\nu} = -(d/ds)B^e(0)$ in (2.10), we have

$$E(q') = \bar{\nu} * \left[k + k\lambda * E(q)\right]. \tag{2.11}$$

From (2.2) and (2.11), we get

$$E(q') = \frac{k/\mu}{1 - \lambda/\mu}. \tag{2.12}$$

This result means that the expectation of busy period caused by k vehicles in queue is k times the expectation of the busy period caused by one vehicle in queue. Intuitively, if one treats the queuing discipline as a last-in–first-out system (noting that the length of busy period remains unchanged, because the number of vehicles being served is irrelevant), then all k vehicles can be served one by one at times when the busy period of the previous vehicle is just terminated. Hence, each of these k vehicles constitutes an independent busy period q. The total busy period caused by these k vehicles is thus k times the individual expected busy period q.

III. Analysis of Adaptive Signal Control

Consider an isolated two-phase intersection with one-way traffic streams without turning movements. Let phase Φ_1 serve the W–E flow and Φ_2 the N–S flow. As illustrated in Fig. 1, let the arrival rate and intersection service rate of the W–E (N–S) direction be $\lambda_1(\lambda_2)$ veh/s and μ_1 (μ_2) veh/s, respectively. We will assume that there is sufficient intersection capacity so that flow ratio $\rho_i = \lambda_i/\mu_i$ is less than one. Furthermore, for the existence of a steady state for the signal control, we will assume that total effective service rate, which includes start-up loss during queue discharge, is greater than total arrival rate. In the figure, the time axis is divided into half cycles for

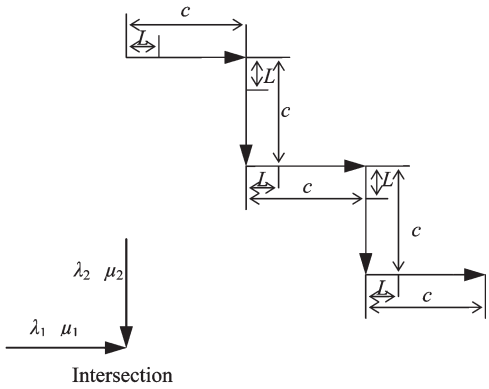


Fig. 1. Adaptive traffic phases at an intersection. L: Time lost during initial queue discharge. c: Half cycle.

the convenience of analysis. Each half cycle c_i is composed of a constant queue discharge lost time L plus a green phase. The odd (even) subscript on c means it serves the W–E (N–S) direction (see Fig. 1).

In the adaptive control scheme analyzed in this paper, phase Φ_1 serves the W–E queue and new arrivals until there is no queue. Then, Φ_2 serves N–S queue and arrivals until there is no queue. Then, it goes back to Φ_1, and so on.

If vehicles arrive at deterministically constant rates, then the optimal signal times (with minimum average delay) for adaptive control can be easily computed as follows. During phase Φ_2, exactly $(c_1 + c_2) \cdot \lambda_2$ vehicles, which arrived in the last half cycle, and the current one need to be served, while total service capacity during this half cycle is $(c_2 - L) \cdot \mu_2$. Analysis for the next phase Φ_1, for time c_3, is similar. That is

$$\begin{aligned}(c_1 + c_2) * \lambda_2 &= (c_2 - L) * \mu_2 \\ (c_2 + c_3) * \lambda_1 &= (c_3 - L) * \mu_1. \end{aligned} \tag{3.1}$$

At steady state, $c_1 = c_3 = c_5 \ldots$, and $c_2 = c_4 = c_6 \ldots$.

Then, optimal half-cycle lengths of two directions are

$$c_1 = \frac{1 + \rho_1 - \rho_2}{1 - \rho_1 - \rho_2} L \tag{3.2a}$$

$$c_2 = \frac{1 - \rho_1 + \rho_2}{1 - \rho_1 - \rho_2} L. \tag{3.2b}$$

Now, consider Poisson arrivals. For a reasonable initialization of the process needed for our numerical analysis, let us assume we start with half cycle c_1 and c_2 given by (3.2). Starting at the beginning of c_2, the W–E vehicle arrivals become Poisson with rate λ_1. Beginning of c_3, N–S vehicles arrive in Poisson manner with rate λ_2. This initialization does not affect the steady state.

Since the adaptive signal serves one direction until queue dissipates, then switches to the other direction, and the switching continues, the $c_3, c_4, \ldots, c_n$ are random variables, whose distributions depend on time needed to serve the queue. Based on the distribution of the half cycles, one can calculate the distribution of delays, cycle lengths, and number of vehicles served during each cycle, as will be shown below.

Let X_{2i} denote the number of vehicles that arrive in the W–E direction during $(c_{2i} + L)$ and X_{2i+1} denote the N–S vehicles during $(c_{2i+1} + L)$. Then, the probability mass function (PMF) of X_2 is

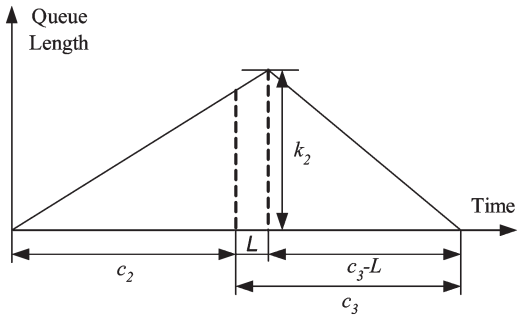


Fig. 2. Queue length curve in one cycle.

$$P_{k_2} = P(X_2 = k_2) = \frac{[\lambda_1 * (c_2 + L)]^{k_2}}{k_2!} e^{-\lambda_1 * (c_2 + L)}. \tag{3.3}$$

From (2.11), we can get the conditional (given k_2) expected green phaselength

$$c_3 - L = k_2 * \frac{1/\mu_1}{1 - \rho_1}. \tag{3.4}$$

The traffic signal switches as soon as the queue becomes zero. Assuming vehicles depart at saturation flow rate during the green phase, the conditional expected number of vehicles served during $(c_3 - L)$ is

$$n_3 = (c_3 - L) * \mu_1 = \frac{k_2}{1 - \rho_1}. \tag{3.5}$$

To calculate the total delay of W–E vehicles in cycle $(c_2 + c_3)$, we need the distribution of the vehicle arrival times in this cycle. Suppose n vehicles arrived in $(0, t)$. Given the Poisson arrival process, the arrival times of these n vehicles $s_1, s_2, \ldots, s_n$ are independently and uniformly distributed in this interval (e.g., [9]). Therefore, the conditional expected total delay D_3 can be viewed as the area of the triangle in Fig. 2

$$D_3 = \frac{1}{2} k_2 * (c_2 + c_3) = \frac{1}{2} k_2 * \left(c_2 + L + \frac{k_2/\mu_1}{1 - \rho_1} \right). \tag{3.6}$$

Hence, the conditional average delay d_3 of the W–E vehicles during $(c_2 + c_3)$ is

$$\begin{aligned} d_3 = D_3/n_3 &= \frac{1}{2} k_2 * \left(c_2 + L + \frac{k_2/\mu_1}{1 - \rho_1} \right) \Big/ \left(\frac{k_2}{1 - \rho_1} \right) \\ &= \frac{1}{2} \left[(c_2 + L) * (1 - \rho_1) + k_2/\mu_1 \right]. \end{aligned} \tag{3.7}$$

The (unconditional) expectation of d_3 is

$$\begin{aligned} E(d_3) &= \sum_{k_2=1}^{\infty} P_{k_2} * D_3/n_3 \\ &= \frac{1}{2}[1 - P(k_2 = 0)] * (c_2 + L) * (1 - \rho_1) + \frac{1}{2} \sum_{k_2=1}^{\infty} P_{k_2} * \frac{k_2}{\mu_1}. \end{aligned} \tag{3.8}$$

The next step is the calculation of the distribution of X_3, which is the number of N–S vehicles arriving in the time interval $(c_3 + L)$. Since c_3 is a discrete random variable, whose PMF is given by (3.3) and (3.4), the calculation of PMF of X_3 is different from that of X_2. For any realization of $(c_3 + L)$, we can get the probability of exactly k_3 N–S vehicle arrivals. The summation of these probabilities over all possible $(c_3 + L)$ gives the probability of exactly k_3 vehicle arrivals during $(c_3 + L)$, that is, the following PMF of X_3:

$$\begin{aligned} P_{k_3} &= P(X_3 = k_3) \\ &= \sum_{\text{all}(c_3+L)} P\left[X_3 = k_3 | (c_3 + L)\right] * P\left[(c_3 + L)\right] \\ &= \sum_{k_2=0}^{\infty} P_{k_2} * P\left\{ X_3 = k_3 \middle| \left[(c_3 + L) = 2L + \frac{k_2/\mu_1}{1-\rho_1} \right] \right\} \\ &= \sum_{k_2=0}^{\infty} P_{k_2} * \frac{[\lambda_2 * (c_3 + L)]^{k_3}}{k_3!} e^{-\lambda_2 * (c_3+L)} \\ &= \sum_{k_2=0}^{\infty} P_{k_2} * \frac{\left[\lambda_2 * \left(2L + \frac{k_2/\mu_1}{1-\rho_1}\right)\right]^{k_3}}{k_3!} e^{-\lambda_2 * \left(2L + \frac{k_2/\mu_1}{1-\rho_1}\right)}. \end{aligned} \tag{3.9}$$

From (2.12), the conditional (given k_3) green phase $(c_4 - L)$, which serves the N–S direction, is

$$c_4 - L = k_3 * \frac{1/\mu_2}{1-\rho_2}. \tag{3.10}$$

The conditional expected number of vehicles served during this green phase is

$$n_4 = (c_4 - L) * \mu_2 = k_3 * \frac{1/\mu_2}{1-\rho_2} * u_2 = \frac{k_3}{1-\rho_2}. \tag{3.11}$$

Similarly, the conditional expected total delay of the N–S vehicle during $(c_3 + c_4)$ is

$$\begin{aligned} D_4 &= \frac{1}{2} k_3 * (c_3 + c_4) \\ &= \frac{1}{2} k_3 * \left(c_3 + L + \frac{k_3/\mu_2}{1-\rho_2} \right) \\ &= \frac{1}{2} k_3 * \left(\frac{k_2/\mu_1}{1-\rho_1} + \frac{k_3/\mu_2}{1-\rho_2} + 2L \right). \end{aligned} \tag{3.12}$$

Hence, the conditional average delay d_4 during $(c_3 + c_4)$ is

$$\begin{aligned} d_4 &= D_4/n_4 \\ &= \frac{1}{2} k_3 * \left(k_2 * \frac{1/\mu_1}{1-\rho_1} + 2L + \frac{k_3/\mu_2}{1-\rho_2} \right) \Big/ \left(\frac{k_3}{1-\rho_2} \right) \\ &= \frac{1}{2}(1-\rho_2) * \left(k_2 * \frac{1/\mu_1}{1-\rho_1} + 2L \right) + \frac{1}{2} \left(\frac{k_3}{\mu_2} \right). \end{aligned} \tag{3.13}$$

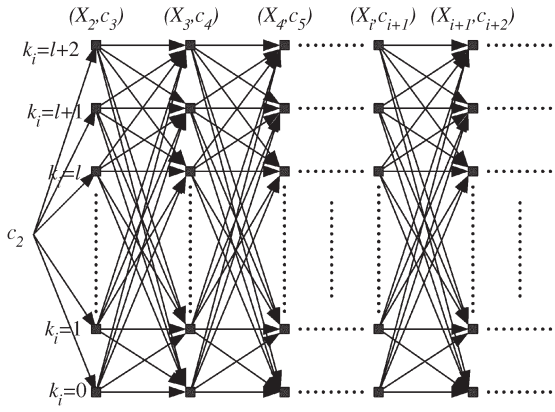


Fig. 3. Scenario tree of the system.

The (unconditional) expectation of d_4 is

$$E(d_4) = \frac{1}{2}(1-\rho_2) * \left\{ 2L * [1 - P(k_3 = 0)] + \sum_{k_2=1}^{\infty} P_{k_2} * \frac{1/\mu_1}{1-\rho_1} k_2 \right\} + \frac{1}{2} \sum_{k_3=1}^{\infty} P_{k_3} * \frac{k_3}{\mu_2}. \tag{3.14}$$

Using such expressions on succeeding half cycles until steady state is reached, one can get the steady-state distribution of 1) the number of vehicles served in a cycle, 2) green phaselengths, and 3) vehicular delays, including average vehicular delays in each direction and in each cycle.

In our initialization of half-cycle lengths, uniform arrivals were assumed. Other initial half cycles could have been used. The reason we selected these half cycles was because they are close enough to steady state so that steady state can be calculated faster.

The final issue in this section concerns the numerical computation of the steady-state probabilities and associated performance measures. Since there are infinite possible realizations of $X_2, X_3, \ldots$ and the number of possible scenarios goes up exponentially, it may appear that the computational load is unmanageable, but if we truncate the state space by neglecting extremely low probability scenarios, the computational load is very manageable. At any stage, we fathom a node in the scenario tree when its probability is below a given threshold. We then can get a relationship of the various measures of interest from one half cycle to another.

Suppose at step 1 of the computational process, the finite PMF for X_2 is calculated (by ignoring very low possibility scenarios). Then, one can easily compute the distribution of $(c_3 + L)$ since it is linearly related to the PMF of X_2. This allows the computation of the conditional probability of k_3 arrivals for each realization of $(c_3 + L)$, which gives us the PMF of X_3 in step 2 of the process. This scenario tree is schematically represented in Fig. 3; note that now, the number of scenarios does not blow up exponentially. The formulas to compute various measures from step to step are derived below. In most cases, less than 30 steps were needed for the

expressions of steady-state probabilities to converge within a threshold of 10^{-6}.

When the distribution of X_{i-1}, which is the number of vehicle arrivals in $(c_{i-1}+L)$, is known, we can get the distribution of the expected length of (c_i+L), since it is linearly related to k_{i-1} as $c_i+L=2L+(k_{i-1}/\mu_j/1-\rho_j)$, where $j=1$ when i is odd and $j=2$ when i is even.

Then, the distribution of X_i is given by

$$\begin{aligned}P_{k_i} &= P(X_i=k_i)\\ &= \sum_{\text{all}(c_i+L)} P\left[X_i=k_i \mid (c_i+L)\right] * P(c_i+L)\\ &= \begin{cases} \sum_{k_{i-1}=0}^{\infty} P_{k_{i-1}} \\ * P\left\{X_i=k_i \mid \left[(c_i+L)=2L+\frac{k_{i-1}/\mu_1}{1-\rho_1}\right]\right\}, & i \text{ is odd} \\ \sum_{k_{i-1}=0}^{\infty} P_{k_{i-1}} \\ * P\left\{X_i=k_i \mid \left[(c_i+L)=2L+\frac{k_{i-1}/\mu_2}{1-\rho_2}\right]\right\}, & i \text{ is even.}\end{cases}\end{aligned} \tag{3.15a}$$

Since X_i arrivals are Poisson, (3.15a) can be simplified to

$$P_{k_i} = \begin{cases} \sum_{k_{i-1}=0}^{\infty} P_{k_{i-1}} * \frac{\left[\lambda_2 * \left(2L+\frac{k_{i-1}/\mu_1}{1-\rho_1}\right)\right]^{k_i}}{k_i!} \\ \quad \times e^{-\lambda_2 * \left(2L+\frac{k_{i-1}/\mu_1}{1-\rho_1}\right)}, & i \text{ is odd} \\ \sum_{k_{i-1}=0}^{\infty} P_{k_{i-1}} * \frac{\left[\lambda_1 * \left(2L+\frac{k_{i-1}/\mu_2}{1-\rho_2}\right)\right]^{k_i}}{k_i!} \\ \quad \times e^{-\lambda_1 * \left(2L+\frac{k_{i-1}/\mu_2}{1-\rho_2}\right)}, & i \text{ is even.}\end{cases} \tag{3.15b}$$

Since the relationship between the conditional green phaselength $(c_{i+1}-L)$ and k_i is

$$c_{i+1}-L = \begin{cases} \frac{k_i/\mu_2}{1-\rho_2}, & i \text{ is odd} \\ \frac{k_i/\mu_1}{1-\rho_1}, & i \text{ is even}\end{cases} \tag{3.16}$$

the conditional expected number of vehicles served during this green phase, given by the product of the green phaselength and the service rate, is

$$n_{i+1} = \begin{cases} (c_{i+1}-L)*\mu_2 = \frac{k_i/\mu_2}{1-\rho_2}*\mu_2 = \frac{k_i}{1-\rho_2}, & i \text{ is odd} \\ (c_{i+1}-L)*\mu_1 = \frac{k_i/\mu_1}{1-\rho_1}*\mu_1 = \frac{k_i}{1-\rho_1}, & i \text{ is odd.}\end{cases} \tag{3.17}$$

Using the queue length trajectories as shown in Fig. 2, the conditional expected total delay during the two half cycles (c_i+c_{i+1}) is

$$\begin{aligned}D_{i+1} &= \frac{1}{2}k_i * (c_i+c_{i+1})\\ &= \begin{cases} \frac{1}{2}k_i * \left(c_i+L+\frac{k_i/\mu_2}{1-\rho_2}\right), & i \text{ is odd} \\ \frac{1}{2}k_i * \left(c_i+L+\frac{k_i/\mu_1}{1-\rho_1}\right), & i \text{ is even}\end{cases}\\ &= \begin{cases} \frac{1}{2}k_i * \left(\frac{k_{i-1}/\mu_1}{1-\rho_1}+\frac{k_i/\mu_2}{1-\rho_2}+2L\right), & i \text{ is odd} \\ \frac{1}{2}k_i * \left(\frac{k_{i-1}/\mu_2}{1-\rho_2}+\frac{k_i/\mu_1}{1-\rho_1}+2L\right), & i \text{ is even.}\end{cases}\end{aligned} \tag{3.18}$$

The conditional average delay d_{i+1} during (c_i+c_{i+1}) is

$$\begin{aligned}d_{i+1} &= \frac{D_{i+1}}{n_{i+1}}\\ &= \begin{cases} \frac{1}{2}k_i * \left(\frac{k_{i-1}/\mu_1}{1-\rho_1}+2L+\frac{k_i/\mu_2}{1-\rho_2}\right) \Big/ \frac{k_i}{1-\rho_2} \\ \frac{1}{2}k_i * \left(\frac{k_{i-1}/\mu_2}{1-\rho_2}+2L+\frac{k_i/\mu_1}{1-\rho_1}\right) \Big/ \frac{k_i}{1-\rho_1}\end{cases}\\ &= \begin{cases} \frac{1}{2}(1-\rho_2) * \left(\frac{k_{i-1}/\mu_1}{1-\rho_1}+2L\right)+\frac{1}{2}\left(\frac{k_i}{\mu_2}\right), & i \text{ is odd} \\ \frac{1}{2}(1-\rho_1) * \left(\frac{k_{i-1}/\mu_2}{1-\rho_2}+2L\right)+\frac{1}{2}\left(\frac{k_i}{\mu_1}\right), & i \text{ is even.}\end{cases}\end{aligned} \tag{3.19}$$

Finally, the (unconditional) expectation of d_{i+1} is

$$\begin{aligned}&E(d_{i+1})\\ &= \begin{cases} (1-\rho_2)*[1-P(k_i=0)]*L+\frac{(1-\rho_2)}{2(1-\rho_1)} \\ \times \sum_{k_{i-1}=0}^{\infty} P_{k_{i-1}} * \frac{k_{i-1}}{\mu_1}+\frac{1}{2}\sum_{k_i=1}^{\infty} P_{k_i} * \left(\frac{k_i}{\mu_2}\right), & i \text{ is odd} \\ (1-\rho_1)*[1-P(k_i=0)]*L+\frac{(1-\rho_1)}{2(1-\rho_2)} \\ \times \sum_{k_{i-1}=0}^{\infty} P_{k_{i-1}} * \frac{k_{i-1}}{\mu_2}+\frac{1}{2}\sum_{k_i=1}^{\infty} P_{k_i} * \left(\frac{k_i}{\mu_1}\right), & i \text{ is even.}\end{cases}\end{aligned} \tag{3.20}$$

Equation (3.20) was used to propagate expected delay from half cycle to half cycle until it converged.

IV. Model Evaluation

In this section, the analytical queuing model is evaluated by comparing its numerical solution with simulation-based results. The simulation model of the adaptive control strategy is built in MATLAB. The random number generator of MATLAB was to generate two groups of exponentially distributed random variables for vehicle interarrival times of the two directions. Service rates and lost time were the same as those of the analytical models. For each type of loading pattern, the simulation model was run 10 times, with a length of 500 000 s each. The first 10 000 s of each run were considered warm-up periods and truncated.

TABLE I
COMPARISON OF AVERAGE NUMBER OF VEHICLES SERVED IN ONE CYCLE AND AVERAGE HALF-CYCLE LENGTHS

Flow ratio of N-S	Number of vehicles served		Half cycle length	
	Analytical	Simulation	Analytical	Simulation
0.10	0.50	0.50	5.00	5.00
0.12	0.63	0.63	5.26	5.26
0.14	0.78	0.78	5.56	5.55
0.16	0.94	0.94	5.88	5.88
0.18	0.13	1.12	6.25	6.25
0.20	1.33	1.33	6.67	6.66
0.22	1.57	1.57	7.14	7.14
0.24	1.85	1.84	7.69	7.68
0.26	2.17	2.17	8.33	8.34
0.28	2.55	2.54	9.09	9.08
0.30	3.00	3.00	10.00	9.99
0.32	3.56	3.53	11.11	11.05
0.34	4.25	4.24	12.5	12.49
0.36	5.14	5.12	14.29	14.24
0.38	6.33	6.32	16.67	16.64
0.40	8.00	8.07	20.00	20.13

Several loading patterns were simulated. In each case, we examined 1) the average number of served vehicles in one cycle, 2) the average half-cycle lengths given by the numerical method and simulation, 3) the average delays, and 4) the number of iterations for the numerical method to converge.

Case 1: Symmetric Intersection With Balanced Flow Ratios: In this case, the saturation flow rates of the two directions are 0.5 veh/s, the arrival rates of two traffic streams are equal, and range from 0.05 to 0.2 veh/s. Queue discharge lost time was set to 4 s for all the cases.

Since the system is symmetric, performance measures should be the same for either direction, and therefore, only measures for one direction are listed in Table I, which gives the average number of vehicles served in a cycle and average half-cycle lengths. The table shows that the means of these measures computed by the analytical model match well with those given by the simulation model. Another interesting finding is that the means of these two measures for adaptive signal control are equal to those of a system with deterministically constant arrivals (with same arrival and service rate) and corresponding optimal half cycles.

Fig. 4 shows the average intersection delays calculated by the queuing model and obtained by simulation. The delay plots show that the average delays given by the two methods match well when the total flow ratio is less than 0.6. As the total flow ratio increases beyond this point, the differences increase but are still less than 10%. The analytical model always tends to underestimate the average delays compared to the simulation model. The differences are due to the method of calculating the green phaselength; the analytical model truncates low probability states and uses expected green phaselength while the simulation model uses actual phaselength (i.e., the distribution of the green phaselengths). Since the truncation threshold is very small, say 10^{-6}, decreasing the threshold will not reduce the differences very much. The difference can be reduced by using the distribution of the phaselength.

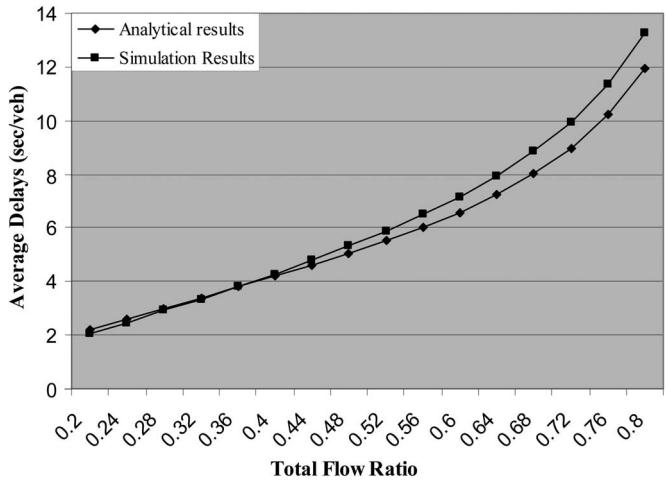


Fig. 4. Average intersection delays.

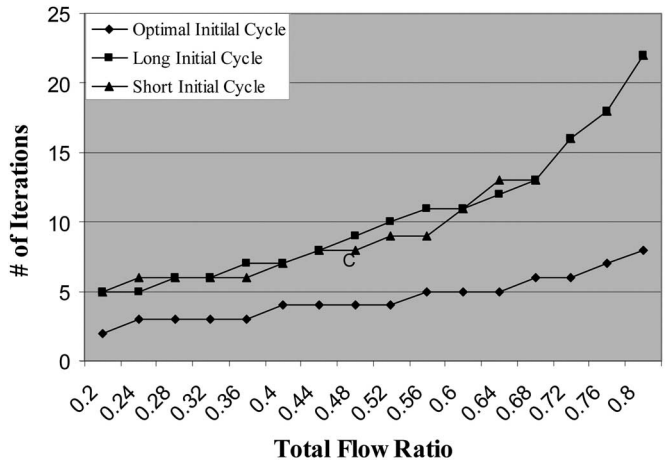


Fig. 5. Number of iterations to converge.

For a numerical method, its convergence, in terms of iterations needed, is a major issue. Fig. 5 shows the numbers of iterations needed to converge, which is defined as when the performance measure does not change by more 10^{-6} from iteration to iteration. Of course, the time to converge depends on the initial cycle lengths. "Optimal initial cycle" in the figure indicates that the initial cycle length determined by (3.2). For comparison, we also used a long initial cycle (where the optimal cycle length is doubled) and a short initial cycle (where the length is halved). While the initial signal settings do not affect the final states at convergence, the system converges much faster with optimal initial cycle where, in most cases, less than ten iterations are needed to reach steady state. The figure also shows that as the total flow ratio increases, the number of iterations needed to converge also increases.

Case 2: Symmetric Intersection With Unbalanced Flow Ratios: The departure flow rates for both directions were set to 0.5 veh/s again. The arrival rates of N–S vehicles ranged from 0.03 to 0.13 veh/s, while the arrival rates of E–W vehicles were twice those of N–S vehicles.

TABLE II
COMPARISON OF AVERAGE NUMBERS OF VEHICLES SERVED IN ONE CYCLE AND AVERAGE HALF-CYCLE LENGTHS IN N–S DIRECTION

Flow ratio of N-S	Number of vehicles served		Half cycle length	
	Analytical	Simulation	Analytical	Simulation
0.06	0.29	0.29	4.59	4.59
0.08	0.42	0.42	4.84	4.84
0.10	0.57	0.57	5.14	5.14
0.12	0.75	0.75	5.50	5.49
0.14	0.97	0.97	5.93	5.94
0.16	1.23	1.23	6.46	6.45
0.18	1.57	1.56	7.13	7.13
0.20	2.00	2.00	8.00	8.00
0.22	2.59	2.57	9.18	9.14
0.24	3.43	3.44	10.86	10.88
0.26	4.72	4.75	13.45	13.50

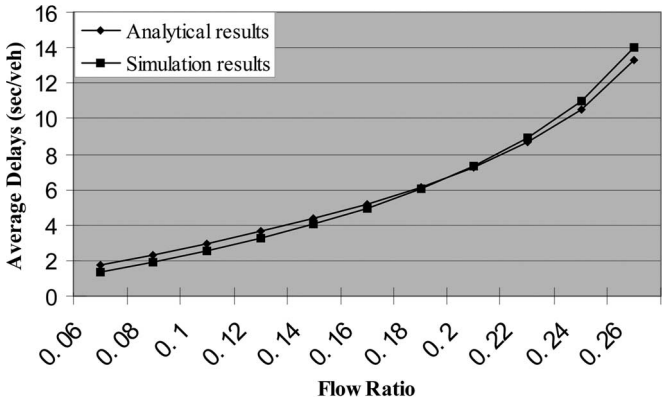


Fig. 6. Average delays of N–S vehicles.

Table II gives the average numbers of vehicles served in one cycle and average half-cycle lengths in the N–S direction, given by the analytical model and the simulation model. The results match well. As in Case 1, the performance measures given by the adaptive signals are equal to those of a system with deterministically constant arrivals and using corresponding optimal half-cycle lengths. The results for the E–W vehicles lead to the same conclusions.

Figs. 6 and 7 show the average delays in the two directions given by the two methods. The delay plots show that the average delays given by the two methods are very close when the total flow ratio is relatively low, say, less than 0.6. As the total flow ratio increases, the differences increase. As before, the numerical queuing model tends to underestimate the average delay as compared with the simulation model. The differences in the E–W direction (which has a higher flow ratio) are larger. Fig. 8 shows the average intersection delays.

Fig. 9 plots the numbers of iterations needed to converge at various flow ratios. Again, when optimal initial cycle lengths are used, the numerical method converges much faster; with a "bad" initial cycle, it still converges to the same steady state, but it is much slower. As before, more iterations are needed to converge when the flow ratio increases.

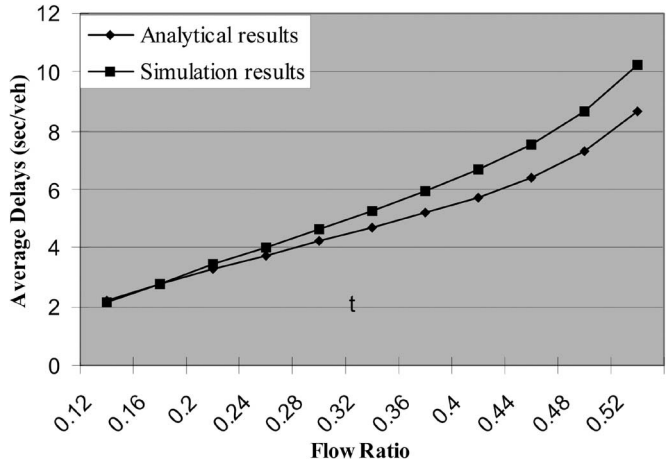


Fig. 7. Average delays of E–W vehicles.

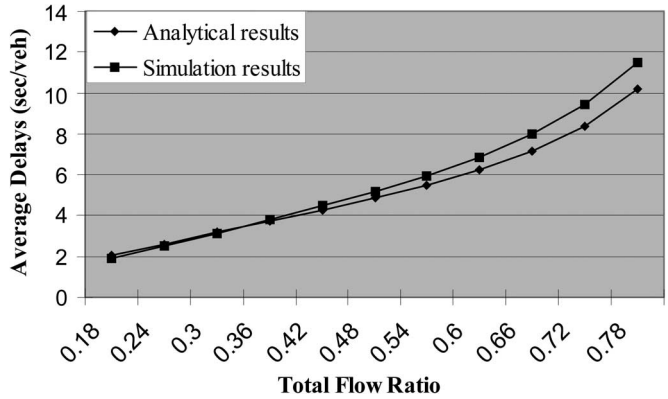


Fig. 8. Average intersection delays.

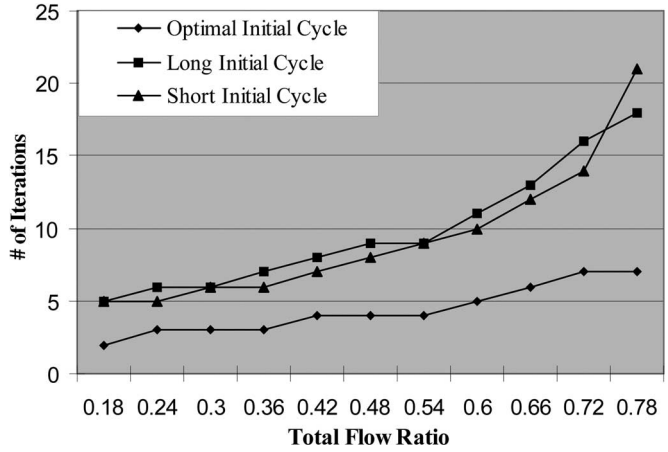


Fig. 9. Number of iterations to converge.

Case 3: Asymmetric Intersection With Balanced Flow Ratios: In this case, the departure rate of N–S vehicles is 0.5 veh/s and of E–W vehicles is 1.0 veh/s. The arrival rates of N–S direction ranges from 0.05 to 0.20 veh/s, and the arrival rates of E–W direction are twice of those of the N–S vehicles.

Table III gives the average numbers of vehicles served in one cycle and average half-cycle lengths in the E–W direction given by the analytical model and the simulation. Again, the

TABLE III
COMPARISON OF AVERAGE NUMBERS OF VEHICLES SERVED IN ONE CYCLE AND AVERAGE HALF-CYCLE LENGTHS OF E–W DIRECTION

Flow ratio of E-W	Number of vehicles served		Half cycle length	
	Analytical	Simulation	Analytical	Simulation
0.10	1.00	1.00	5.00	5.00
0.12	1.26	1.26	5.26	5.26
0.14	1.56	1.57	5.56	5.57
0.16	1.88	1.88	5.88	5.88
0.18	2.25	2.25	6.25	6.25
0.20	2.66	2.66	6.67	6.66
0.22	3.14	3.14	7.14	7.14
0.24	3.69	3.72	7.69	7.72
0.26	4.33	4.34	8.33	8.34
0.28	5.09	5.10	9.09	9.1
0.30	6.00	6.02	10.00	10.02
0.32	7.11	7.11	11.11	11.11
0.34	8.50	8.49	12.50	12.49
0.36	10.28	10.29	14.29	14.29
0.38	12.67	12.7	16.67	16.7
0.40	16.00	16.10	20.00	20.09

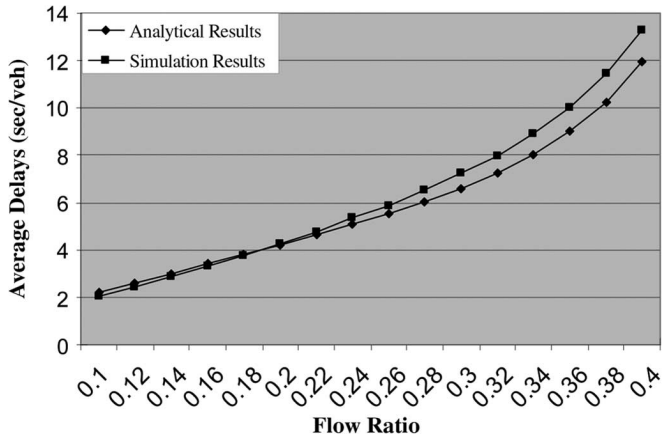


Fig. 10. Average delays of N–S vehicles.

results of the two models match well. Also, as for Case 1, results of the adaptive signals equal to those of a system with deterministically constant arrivals and optimal half cycles. The results for the E–W vehicles lead to the same conclusions.

The average delays in the two directions and the combined intersection delays are given in Figs. 10–12. The number of iterations needed for convergence for different flow ratios and initial signal cycles are given in Fig. 13. In summary, the numerical solution of the queuing model approximates well the simulation results and converges quickly.

Case 4: Asymmetric Intersection With Equal Arrival Rates: In this last case, the departure rates in the two directions are same as those of the Case 3. The arrival rates in the two directions are equal and range from 0.06 to 0.28.

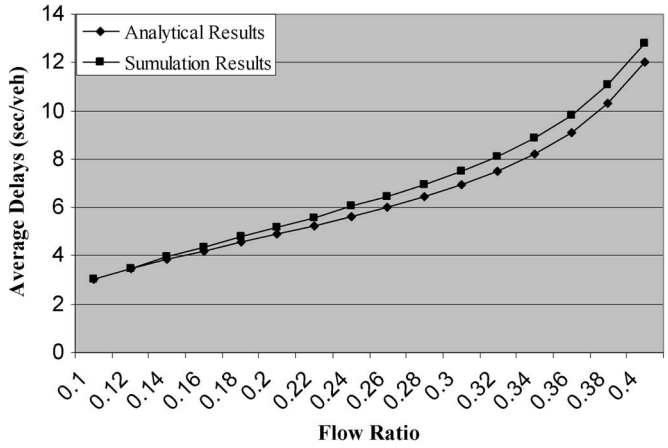


Fig. 11. Average delays of E–W vehicles.

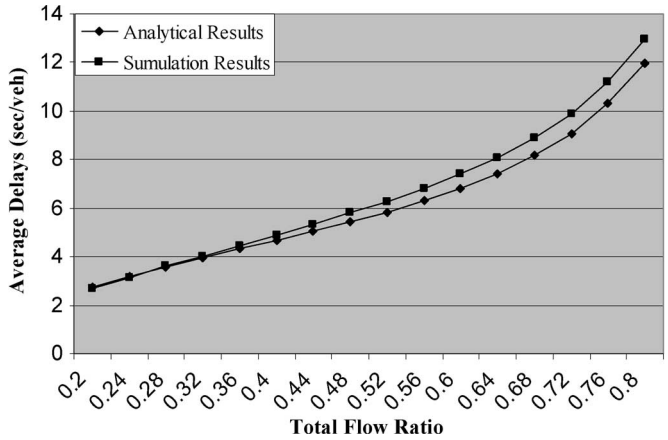


Fig. 12. Average intersection delays.

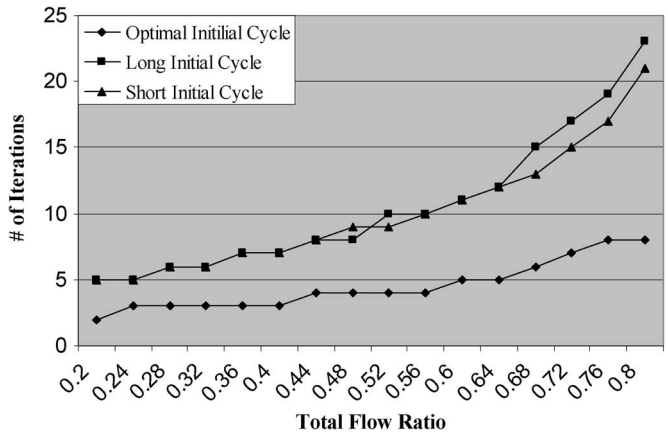


Fig. 13. Number of iterations to converge.

Table IV gives the average numbers of vehicles served during one cycle and average half-cycle lengths in the N–S direction given by analytical and simulation models.

The average delays in the two directions and for the whole intersection are given in Figs. 14–16. The number of iterations needed for convergence for difference flow ratios and initial signal cycles are given in Fig. 17. Again, the analytical results match well with the simulation results, and the numerical solution of the analytical model converges quickly.

TABLE IV
COMPARISON OF AVERAGE NUMBERS OF VEHICLES SERVED IN ONE CYCLE AND AVERAGE HALF-CYCLE LENGTHS OF N–S DIRECTION

Flow ratio of N-S	Number of vehicles served		Half cycle length	
	Analytical	Simulation	Analytical	Simulation
0.12	0.59	0.59	5.17	5.18
0.16	0.84	0.84	5.68	5.68
0.2	1.14	1.15	6.29	6.29
0.24	1.50	1.50	7.00	7.00
0.28	1.93	1.93	7.86	7.87
0.32	2.46	2.46	8.92	8.91
0.36	3.13	3.14	10.26	10.29
0.4	4.00	3.97	12.00	11.95
0.44	5.17	5.16	14.35	14.32
0.48	6.86	6.86	17.71	17.72
0.52	9.45	9.47	22.91	22.95
0.56	14.00	13.99	32.00	31.97

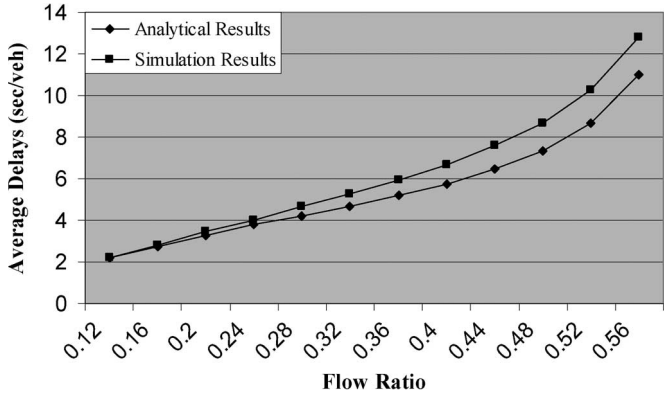


Fig. 14. Average delays of N–S vehicles.

Summarizing the evaluation of the queuing model of the four different cases of intersection loading, we conclude the following.

1) The numerical solution of the queuing model estimates very well the average number of vehicles served in each cycle and average half-cycle length of the strategy.
2) The adaptive signal control strategy, for Poisson vehicle arrivals, performs as well as the optimal strategy for a system with deterministically constant arrivals, where the arrival rates are the same for both systems.
3) When total flow ratio is lower than 0.6, the estimations of average delays given by the analytical model are very close to the simulation-based results. The queuing model underestimates the average delay slightly when the total flow ratio is higher than 0.60.
4) When the initial cycle lengths are optimal, the numerical method converges very quickly: in less than ten iterations. With other initial cycle lengths (twice or half the optimal lengths), the model still converges but at a slower rate. Also, in general, more iterations are needed when the flow ratio increases.

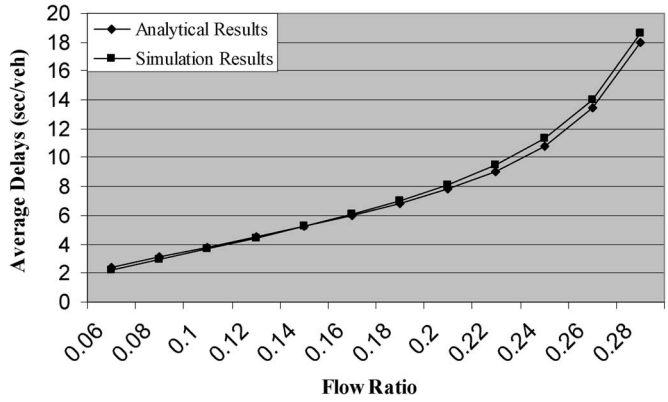


Fig. 15. Average delays of E–W vehicles.

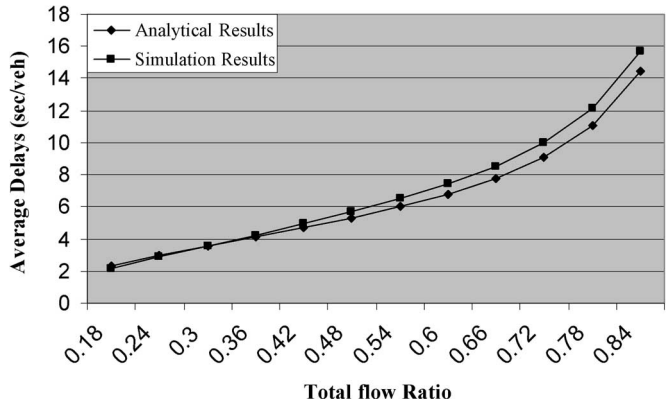


Fig. 16. Average intersection delays.

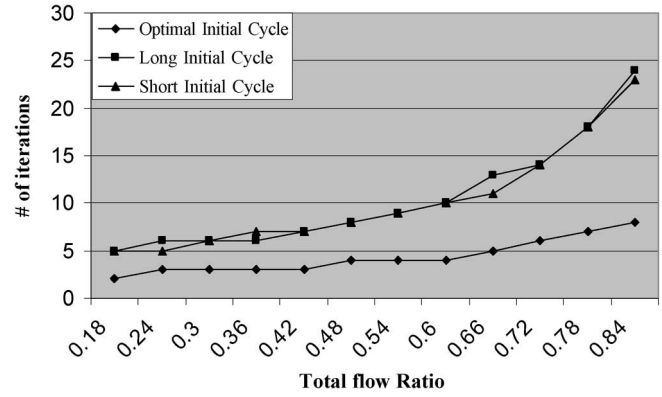


Fig. 17. Number of iterations to converge.

V. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

This paper develops an analytical queuing model to analyze the queues and delays at an isolated two-phase intersection control that uses a simple adaptive signal strategy originally proposed by Newell [6]. In this strategy, the first movement is served until the queue dissipates, then the second movement is served until the queue dissipates, and, then, the signal goes back to serving the first movement, and this cyclic process repeats. A numerical algorithm was developed to compute steady-state performance measures such as average delays and expected queue lengths.

Numerical solution of the model was compared with results from a simulation model. The comparison shows that

the analytical model approximates very well the simulation results when the flow ratio is less than 0.60. The model slightly underestimates the average delays when flow ratio increases.

We are currently investigating the use of this modeling methodology to model an adaptive control strategy for a four-phase intersection control, where the signal switches to the next phase after queues being served in the current phase are depleted. We also plan to develop such queuing models for an intersection operating under fixed signal timing. The major complication in modeling this scenario is the consideration of residual queues, when they occur, from one cycle to another. (In our adaptive strategy, there are no residual queues.) In order to reduce the differences between the results of numerical method and those of the simulation, the exact distribution of the busy period is being studied. We also plan to model queues at ramp meters using this approach.

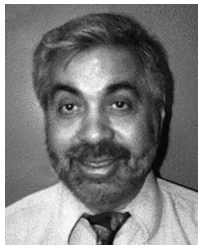
A major assumption, and perhaps a weakness, of the model is the assumption of Poisson arrivals. We plan to investigate if such a model can be used for other arrival distributions, in particular, an arrival pattern that approximates batch (platoon) arrivals.

ACKNOWLEDGMENT

The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views of the sponsors of this work.

REFERENCES

[1] R. Akcelik, "The highway capacity manual delay formula for signalized intersections," *ITE J.*, vol. 58, no. 3, pp. 23–28, Mar. 1988.
[2] V. F. Hurdle, "Signalized intersection delay models—A primer for the uninitiated," *Transp. Res. Rec.*, no. 971, pp. 96–105, 1984.
[3] L. Kleinrock, *Queueing Systems. Volume 1. Theory*. Hoboken, NJ: Wiley, 1975.
[4] A. J. Miller, "Settings for fixed-cycle traffic signals," *Oper. Res. Q.*, vol. 14, no. 4, pp. 373–386, Dec. 1963.
[5] G. F. Newell, "Approximation methods for queues with applications to the fixed-cycle traffic light," *SIAM Rev.*, vol. 7, no. 2, pp. 223–240, Apr. 1965.
[6] ——, "The rolling horizon scheme of traffic signal control," *Transp. Res. Part A*, vol. 32, no. 1, pp. 39–43, Jan. 1998.
[7] P. S. Olszewski, "Modeling of queue probability distribution at traffic signals," in *Proc. 11th Int. Symp. Transp. and Traffic Theory*, Tokyo, Japan, 1990, pp. 569–588.
[8] ——, "Modeling probability distribution of delay at signalized intersections," *J. Adv. Transp.*, vol. 28, no. 3, pp. 253–274, fall/winter 1994.
[9] S. Ross, *Stochastic Process, 2nd Edition*. Hoboken, NJ: Wiley, 1996.
[10] TRB, *Special Report 209: Highway Capacity Manual*, 1985, Washington, DC: TRB, Nat. Res. Council.
[11] F. V. Webster, "Traffic signal settings," Road Res. Lab., Ministry Transport, HMSO, London, U.K., pp. 1–43, Road Research Technical Paper 39, 1958.



Pitu B. Mirchandani (M'80–SM'82) received the B.S. and M.S. degrees in engineering from the University of California, Los Angeles, and the S.M. degree in aeronautics and astronautics and the Sc.D. degree in operations research from the Massachusetts Institute of Technology, Cambridge.

He is a Professor of systems and industrial engineering with the University of Arizona, Tucson. His interests span a broad variety of areas of operation research, including optimization, control of stochastic systems, and logistics, routing, location, and scheduling.

Ning Zou received the B.S. and M.S. degrees in mechanical and electrical engineering from Zhejiang University, Hangzhou, China, in 1997 and 2000, respectively, and the M.S. degree in industrial engineering from the University of Arizona, Tucson, in 2004. He is currently working toward the Ph.D. degree in systems and industrial engineering at the University of Arizona.

He is a Transportation Analyst with Kittelson and Associates, Inc., Tucson. His research focuses on mathematical modeling of traffic systems.