Rate Based ABR Flow Control using Two Timescale SPSA

Shalabh Bhatnagar, Michael C. Fu and Steven I. Marcus *

Institute for Systems Research, University of Maryland, College Park, MD 20742

ABSTRACT

In this paper, a two timescale simultaneous perturbation stochastic approximation (SPSA) algorithm is developed and applied to closed loop rate based available bit rate (ABR) flow control. The relevant convergence results are stated and explained. Numerical experiments demonstrate fast convergence even in the presence of significant delays and a large number of parameterized policy levels.

Keywords: Rate-based ABR flow control, Two timescale SPSA, Optimal parameterized feedback policies.

1. INTRODUCTION

The available bit rate (ABR) service in asynchronous transfer mode (ATM) networks is used primarily for data traffic. As the name suggests, bandwidth allocation for ABR service is done only after the higher priority services such as constant bit rate (CBR) and variable bit rate (VBR) have been allocated bandwidth. The available bandwidth is a time-varying quantity, and for proper utilization, the network requires the ABR sources to flow control their traffic. Two main proposals discussed at the ATM forum for flow control in ABR service[16] were the rate based and the credit based schemes. The rate based scheme was recently accepted by the ATM forum primarily because of the higher hardware complexity and costs involved in the latter scheme. Several heuristic algorithms for computation of explicit ABR rates have since been proposed at the ATM forum.[10] Most other approaches[11,13,8] use fluid models of the network. Even though of obvious practical interest, not much is so far available in terms of stochastic control approaches in the queueing framework largely because of the inherent difficulties with such approaches. For instance, the problem has been studied as a team problem in the discrete time queueing framework but with linearized queue dynamics.[1] In another paper,[15] a continuous time queueing model has been studied and stability conditions for various feedback policies have been obtained. However, performance analysis there has been done only under the assumptions of no delays and continuous observations. In this paper, we also study the rate based ABR flow control problem in the continuous time queueing framework. We formulate the problem as a parameter optimization problem where we use an efficient simulation based stochastic approximation scheme for computing optimal ABR rates in the presence of delays with observations and information feedback done at discrete time instants. We essentially view the ABR problem as a stochastic optimization problem wherein we consider parameterized policies that have several levels of control. We employ a two timescale simultaneous perturbation stochastic approximation (SPSA)[17] algorithm for this purpose. This algorithm is a variant of a two timescale stochastic approximation algorithm[4] originally developed for simulation based parametric optimization of hidden Markov models. The two timescale algorithm had the advantage that it updates the parameter vector at increasing deterministic instants[2] unlike several other schemes which do so at regeneration epochs.[7,9] However, even though it uses two parallel simulations for any N-vector parameter, it had the disadvantage that it updates the parameter once every N-cycles. Using the SPSA variant of the earlier two timescale algorithm, we still need only two parallel simulations, but we update the parameter now once every cycle (instead of every N-cycles as earlier). As a result, in the earlier case, the speed of convergence deteriorates considerably as N becomes large, which is not so using our scheme. The gradient for the SPSA version of the two timescale algorithm is obtained from two performance measurements taken at randomly perturbed settings of the parameter components, most commonly by using i.i.d. symmetric Bernoulli random variables. Our numerical experiments indicate that the two timescale SPSA algorithm is particularly effective in obtaining optimal multilevel feedback policies (with as many levels as one wants) under non-zero delays and sampling instants. This problem was earlier studied with the uncontrolled stream Poisson[3] using the original two timescale algorithm.[4]

The rest of the paper is organized as follows: In Section 2, we formulate the problem, define the multilevel parameterized feedback policies, present the SPSA algorithm for obtaining an optimal such (multilevel) policy and finally state and briefly explain the relevant convergence theorems. In Section 3, numerical examples are presented

Part of the SPIE Conference on Performance and Control of Network Systems III
Boston, Massachusetts • September 1999
SPIE Vol. 3841 • 0277-786X/99/$10.00

for various parameter settings and the superiority of closed loop feedback optimal parameterized policies over optimal open loop policies is illustrated, in addition to obtaining fast convergence using this scheme.

2. THE OPTIMIZATION PROBLEM

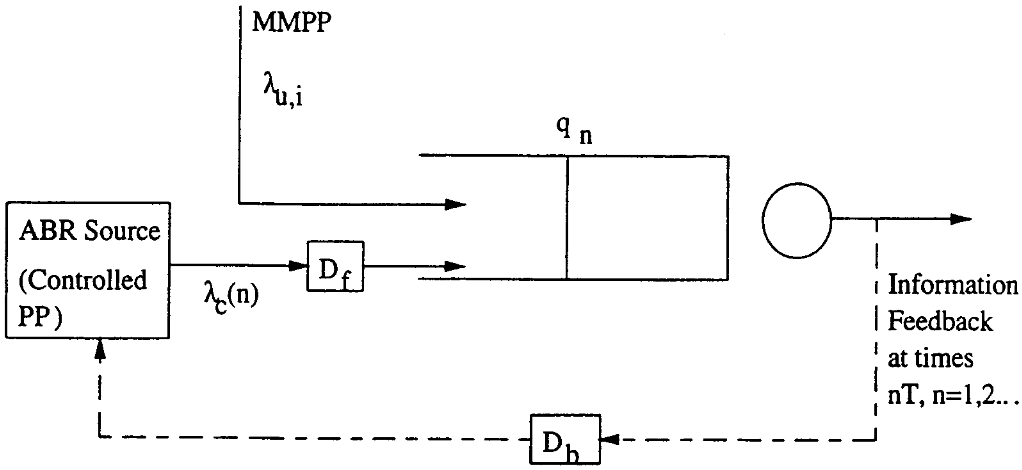

Figure 1. The ABR Model

A single queue with buffer size B is fed with two input streams. This queue may represent a bottleneck node on the virtual circuit of an ABR source. One of the input streams is controlled (representing the traffic from the ABR source) and the other one is uncontrolled (representing all the other traffic in the network passing through this node). The uncontrolled stream is modelled as a Markov modulated Poisson process (MMPP). The service time process is assumed to be i.i.d. with a general distribution which is same for both the streams. Generalization to the case where this is different for the two streams is straightforward. The size of the bottleneck buffer is B which we assume to be a large quantity. The ABR stream is modelled as a controlled Poisson process with instantaneous intensity $\lambda_c(t)$ at time t and which in turn is defined by a feedback controlled law defined in (1) below. The ABR rate is computed at the node and is fed back to the source with a delay D_b. The source then starts sending packets with the new rate. These packets experience a propagation delay D_f in arriving at the node. The rate $\lambda_c(t)$ of the ABR source is held fixed in time intervals $[nT,(n+1)T)$ (as $\lambda_c(nT)$), $n \geq 0$, for some given fixed $T > 0$. By abuse of notation, we shall represent $\lambda_c(nT)$ by just $\lambda_c(n)$. The queue length process $\{q(t), t \geq 0\}$ is observed at instants nT ($q_n \triangleq q(nT)$), $n = 0,1,...$, and from it the desired ABR rate for the next time interval is computed at the node using the following feedback law: Let $S = \{0,1,..,B\}$ represent the set of all queuelength values. Let a_i, $i = 0,1,..,N$, be integers such that $-1 = a_0 < a_1 < a_2 < .. < a_{N-1} < a_N = B$. Define the following subsets of S: $S_i = \{a_{i-1}+1,..,a_i\}$, $i = 1,..,N$. Then for $i = 1,..,N$,

$$\lambda_c(n) = \lambda_i \text{ if } q_n \in S_i, \tag{1}$$

for some $\lambda_1,..,\lambda_N$. Let $\theta \triangleq (\lambda_1,\lambda_2,...,\lambda_N)^T$ represent the N - dimensional parameter vector to be optimized that takes values in a compact set $C \subset \mathcal{R}^N$. We shall assume in particular that C is of the form $\prod_{i=1}^{N}[\lambda_{i,\min},\lambda_{i,\max}]$, with $\lambda_{i,\min} > 0$, for all $i = 1,..,N$. Let X_n represent the state of the modulating chain of the uncontrolled MMPP at time nT that takes values in a finite set S_u. Further, let r_n represent the residual service time of the customer in service at time nT. Then under the type of policies (1), for $D_b = D_f = 0$, it can be seen that $\{(q_n, X_n, r_n)\}$, $n \geq 0$, is ergodic Markov. When D_b, D_f are non-zero, we will assume for simplicity that $D_b + D_f = MT$ for some integer $M > 0$. Simple modifications can however take care of the case when $D_b + D_f \neq MT$ for any $M > 0$. It can also be seen that when $D_b + D_f = MT$ for some $M > 0$, the joint process $\{ (q_n, X_n, r_n, q_{n-1}, X_{n-1}, r_{n-1},..., q_{n-M}, X_{n-M}, r_{n-M}) \}$, $n \geq 0$, is ergodic Markov under the above policies. For ease of exposition, we will consider the case $D_b = D_f = 0$, in detail from now on and explain the changes necessary for nonzero D_b, D_f, as we proceed. Thus for $D_b = D_f = 0$, for any given θ, $\{(q_n, X_n, r_n)\}$ is ergodic Markov with $\{\lambda_c(n)\}$ given by (1). Let $h : S \to \mathcal{Z}^+$ be a given bounded and nonnegative cost function ($\mathcal{Z}^+$ is the space of nonnegative integers). Our aim is to minimize the long-run average cost

$$J(\theta) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} h(q_i).$$

It can be shown[5] that if the underlying process is ergodic Markov under all policies, the limit above exists w.p. 1 and is deterministic for each θ. Let $\delta > 0$ be a fixed small constant. Define sequences $\{a(n)\}$ and $\{b(n)\}$ in $(0,1]$ as follows: $a(0) = b(0) = 1$, $a(i) = i^{-1}$, $b(i) = i^{-\alpha}$, $i \geq 1$, for some $\frac{1}{2} < \alpha < 1$. Then it is clear that

$$\frac{a(n+1)}{a(n)}, \frac{b(n+1)}{b(n)} \to 1, \text{ as } n \to \infty, \tag{2}$$

$$\sum_n a(n) = \sum_n b(n) = \infty, \sum_n a(n)^2, \sum_n b(n)^2 < \infty, \ a(n) = o(b(n)). \tag{3}$$

Define $\{n_m, m \geq 0\}$ as follows: $n_0 = 1$ and $n_{m+1} = \min\{j > n_m \mid \sum_{i=n_m+1}^{j} a(i) \geq b(m)\}$, $m \geq 1$. Let $\pi_i(x)$ denote the point closest to $x \in \mathcal{R}$ in the interval $[\lambda_{i,\min}, \lambda_{i,\max}] \subset \mathcal{R}$ (defined earlier) and $\pi(\theta)$ be defined by $\pi(\theta) = (\pi_1(\lambda_1), \pi_2(\lambda_2), .., \pi_N(\lambda_N))^T$. Let $\Delta(i) = (\Delta_{i,1}, .., \Delta_{i,N})^T$, $i = 1, 2, ...$, be such that for any i, $\Delta_{i,1}, .., \Delta_{i,N}$ are symmetric Bernoulli distributed i.i.d. random variables (i.e., $\Delta_{i,m} = \pm 1$ w.p. $1/2$, $\forall i = 1, 2, .., m = 1, 2, .., N$) and the random vector $\Delta(i)$ is independent of $\Delta(j)$ for any $i \neq j$. Define parallel processes $\{(q_j^1, X_j^1, r_j^1)\}$ and $\{(q_j^2, X_j^2, r_j^2)\}$ such that for $n_m < j \leq n_{m+1}$, $\{(q_j^1, X_j^1, r_j^1)\}$ is governed by $\pi(\theta(m) - \delta\Delta(m)) \stackrel{\Delta}{=} (\pi_1(\lambda_1(m) - \delta\Delta_{m,1}), ..., \pi_N(\lambda_N(m) - \delta\Delta_{m,N}))^T$. Similarly, $\{(q_j^2, X_j^2, r_j^2)\}$ is governed by $\pi(\theta(m) + \delta\Delta(m))$ (for $n_m < j \leq n_{m+1}$) defined analogously. In the above, $\theta(m) \stackrel{\Delta}{=} (\lambda_1(m), ..., \lambda_N(m))^T$ represents the mth update of θ and is governed by the following recursion equations: For $i = 1, .., N$,

$$\lambda_i(m+1) = \pi_i\left(\lambda_i(m) + \sum_{j=n_m+1}^{n_{m+1}} a(j)\left(\frac{h(q_j^1) - h(q_j^2)}{2\delta\Delta_{m,i}}\right)\right). \tag{4}$$

Let us now briefly mention the original two timescale algorithm[4] for comparison purposes. Let $\{(q_j, X_j, r_j)\}$ and $\{(\bar{q}_j, \bar{X}_j, \bar{r}_j)\}$, be processes that are respectively governed by $\{\bar{\theta}_j\}$, $\{\hat{\theta}_j\}$, where $\bar{\theta}_j = \theta(m)$ for $n_{Nm} < j \leq n_{N(m+1)}$, $\hat{\theta}_j = \pi(\theta(m) + \delta e_i)$, $i = 1, .., N$, for $j = n_{Nm+i-1}, n_{Nm+i-1}+1, .., n_{Nm+i}-1$, $i = 1, .., N$, $m \geq 0$, and where e_i is the unit vector with 1 in the ith place and 0 elsewhere. Then the algorithm[4] is as follows: For $i = 1, .., N$,

$$\lambda_i(m+1) = \pi_i\left(\lambda_i(m) + \sum_{j=n_{Nm+i-1}+1}^{n_{Nm+i}} a(j)\left(\frac{h(q_j) - h(\bar{q}_j)}{\delta}\right)\right). \tag{5}$$

The SPSA algorithm (4) has faster rate of convergence than (5). In fact, (4) tracks trajectories of the ordinary differential equation (o.d.e) (6) below.[5] It can be shown[4] that (5) tracks trajectories of an o.d.e that is similar to (6) but has a factor of $1/N$ multiplying its RHS. Thus even though the qualitative behaviour of the two algorithms is similar, the factor of $1/N$ in effect slows down the rate of convergence of (5). Lemma 1 and Theorem 1 below establish the preliminary hypotheses for convergence of the SPSA scheme and Theorem 2 gives the main convergence result. We have skipped detailed proofs[5] of the results presented here. Let $p_\theta(i, x, r; i', x', dr')$, $i, i' \in S$, $x, x' \in S_u$, $r, r' \in [0, \infty)$, represent the transition probabilities of the Markov chain $\{(q_n, X_n, r_n)\}$ for given θ. Lemma 1 establishes continuity of the transition kernel and can be shown using sample path arguments.

Lemma 1 Under all policies of type (1), the map $(\theta, r) \to p_\theta(i, x_1, r; j, x_2, dr')$ is continuous in θ. □

The following theorem shows that the average cost $J(\theta)$ is continuously differentiable in the parameter θ and has been shown[5] for the case when the service times are exponential. Let $\mu_\theta(q, x, dr)$ represent the stationary distribution of the Markov chain $\{(q_n, X_n, r_n)\}$ when $\theta \in C$ is a given fixed parameter. Let $\nu_\theta(q)$ represent its marginal along the process $\{q_n\}$ alone. Then $J(\theta)$ can be written as

$$J(\theta) = \sum_{q\in S} h(q)\nu_\theta(q).$$

Theorem 1 Under all policies of type (1) and under exponential service times, $J(\theta)$ is continuously differentiable in θ. □

The proof of Theorem 1 is based on Theorem 2[14] from which one can show that the stationary distribution of a parametrized Markov process that is ergodic under all policies is continuously differentiable in the parameter if its stationary transition probabilities are. This result is however valid only for finite state chains. Note that the service time distribution is independent of the parameter θ. Moreover, for exponential service distributions, the residual service times are exponentially distributed as well (because of the memoryless property of exponential distribution), and in fact the subchain $\{(q_n, X_n)\}$, $n \geq 0$, for exponential service distributions and under the type of feedback policies (1), is by itself ergodic Markov for every $\theta \in C$. Also note that this subchain (for any given θ) has a finite state space since both S and S_u are finite. Now the stationary distribution $\mu_\theta(q, x, dr)$ of the original chain $\{(q_n, X_n, r_n)\}$ for given θ can be decomposed as

$$\mu_\theta(q, x, dr) = \phi_\theta(q, x)\psi(dr),$$

where $\phi_\theta(q, x)$ is the stationary distribution of the subchain $\{(q_n, X_n)\}$ and $\psi(dr)$ is exponential with rate μ (the service rate). Moreover, it can be shown[5] that the transition probabilities of the subchain $\{(q_n, X_n)\}$ are continuously differentiable in the parameter. Hence $\phi_\theta(q, x)$ and hence $\mu_\theta(q, x, dr)$ are continuously differentiable in θ. As a result, ν_θ (defined earlier) is continuously differentiable in θ as well and hence the average cost $J(\theta)$ trivially is (since S is finite). The problem of proving Theorem 1 for i.i.d. general service times remains an open problem.

We now present our main convergence theorem. The ordinary differential equation (o.d.e.) technique is commonly used to prove convergence of stochastic approximation algorithms. Theorem 2 shows that the algorithm (4) asymptotically tracks the stable points of the o.d.e. (6) below. Let $\tilde{Z}(t) \stackrel{\Delta}{=} (\tilde{Z}_1(t), .., \tilde{Z}_N(t))$, where $\tilde{Z}_i(t)$, $i = 1, .., N$, satisfy the o.d.e.

$$\dot{\tilde{Z}}_i\,(t) = \tilde{\pi}_i(-\nabla_i J(\tilde{Z}(t))), \quad i = 1, .., N, \ \tilde{Z}(0) \in C, \tag{6}$$

where for any $v(.) \in C_b(\mathcal{R})$,

$$\tilde{\pi}_i(v(y)) = \lim_{0<\Delta\to 0} \left(\frac{\pi_i(y + \Delta v(y)) - \pi_i(y)}{\Delta} \right).$$

For $x = (x_1, .., x_N) \in \mathcal{R}^N$, let $\tilde{\pi}(x) = (\tilde{\pi}_1(x_1), .., \tilde{\pi}_N(x_N))^T$. Let $K = \{\theta \in C \mid \tilde{\pi}(\nabla J(\theta)) = 0\}$. Note that, since $\tilde{\pi}(\nabla J(\theta)).\tilde{\pi}(-\nabla J(\theta)) < 0$, for $\theta \notin K$, by definition, $J(.)$ itself is a strict Liapunov function for (6), and the o.d.e. (6) for arbitrary initial conditions converges to the set K (which serves as the asymptotically stable attractor for the o.d.e[6])). For $\eta > 0$, let $K^\eta = \{\theta \in C \mid \exists\theta' \in K \text{ s.t. } ||\theta - \theta'|| \leq \eta\}$.

Theorem 2 Given $\eta > 0$, $\exists\bar{\delta} > 0$ such that for any $\delta \in (0, \bar{\delta}]$, the algorithm (4) converges to K^η a.s. □

Remark Note that K is the set of all critical points of (6), and not just the set of local minima. However, points in K that are not local minima will be unstable equilibria, and because of the presence of noise (randomness) in the algorithm, under fairly general conditions, the algorithm will converge to the η-neighborhood of K_0($\stackrel{\Delta}{=}$ the set of local minima of $J(.)$) $\subset K$.[12] Note that we assumed the cost function to be merely bounded and continuous. If on the other hand, we assume the cost function $h(.)$ to be in addition convex, the average cost $J(.)$ will be convex as well. Moreover, if $J(.)$ is strictly convex, it will have a unique minimum, to which our algorithm will a.s. converge within an η-neighborhood.

3. NUMERICAL EXPERIMENTS

Next, we provide numerical results to illustrate the two timescale SPSA scheme. As mentioned earlier, flow control in ABR service requires balancing various conflicting performance criteria such as mean and variance of delay and throughput. Often this is addressed by minimizing the distance of stationary mean queue length from a given fixed constant N_0.[18,10,15,3] We adopt a similar approach here, i.e., we select our cost function to be $h(x) = |x - N_0|$, where N_0 is assumed given. In other words we assume that the network knows the value of N_0 at the time of setting up the ABR connection. A stochastic approximation scheme for obtaining an optimal such N_0 can also be provided.[5]

We performed experiments with policies that have five and eleven parameter levels. However, we describe here experiments with five level policies only since results for those with eleven level policies are similar.[5] We assume for simplicity that both D_b and D_f are integral multiples of T. The form of the five level policies for obtaining $\lambda_c(n)$, is as follows:

$$\lambda_c(n) = \begin{cases} \lambda_1^* & \text{if} \quad q_n < N_0 - 2\epsilon \\ \lambda_2^* & \text{if} \quad N_0 - 2\epsilon \le q_n < N_0 - \epsilon \\ \lambda_3^* & \text{if} \quad N_0 - \epsilon \le q_n \le N_0 + \epsilon \\ \lambda_4^* & \text{if} \quad N_0 + \epsilon < q_n \le N_0 + 2\epsilon \\ \lambda_5^* & \text{if} \quad q_n > N_0 + 2\epsilon. \end{cases} \tag{7}$$

In the above, we assume that ϵ is a given fixed constant in addition to N_0. The eleven level policies are similarly defined. In the numerical experiments, we actually consider a generalization of the model in Fig.1, where rate feedback is done at instants nF_b, $n \ge 1$, where F_b is a fixed multiple of T. This gives us added flexibility in studying the effect of changes in F_b in addition to those in T. Thus F_b in some sense plays the role of an additional delay. The sequence of events is thus as follows: The ABR rate $\lambda_c(\cdot)$ is computed at times nT, $n \ge 1$, at the node, using feedback policies above. These rates are fed back to the source every F_b units of time. The source receives this rate information with a delay D_b and upon receiving it immediately starts sending packets with the new rate. The packets arrive at the node with a propagation delay D_f.

The parameter vector is computed off-line using the two timescale SPSA algorithm (4). The uncontrolled process is an MMPP with the underlying Markov chain being an irreducible two state chain. For simplicity we assume that the underlying chain undergoes state transitions every T units of time. We make this assumption only to simplify the simulation code. The buffer size B in these experiments is 5×10^5. In what follows, we shall illustrate simulation experiments using our stochastic approximation scheme. We also compare closed loop performance (with delays etc.) with the optimal open loop performance that is obtained using the scalar parameter ($N = 1$) version of (5).

Let θ^* denote the parameter value for the corresponding optimal policy, i.e., $\theta^* = (\lambda_1^*, .., \lambda_5^*)^T$ for the case of 5-level closed loop policies that we show here and $\theta^* = \lambda^*$ for the open loop policy. In the following, subscript θ^* is used in the definition of various performance measures to indicate θ^*-parametrized stationary distributions of the various quantities. Thus, $Var_{\theta^*}(q_n)$ represents the stationary variance of $\{q_n\}$ parameterized by θ^*. Let B_d represent the segment or band (of queue length values) $[N_0 - \epsilon, N_0 + \epsilon]$. We compare performance in terms of parameters of queue length distributions and throughput rate; $\bar{q}$, P_{band}, σ_q, $\bar{\lambda}_c$, P_{idle} and $J(\theta^*)$. These quantities and their estimates are defined as follows:

$$\bar{q} \triangleq E_{\theta^*}[q_n] \approx \frac{1}{N}\sum_{i=0}^{N-1} q_i, \quad \sigma_q \triangleq \mathrm{Var}_{\theta^*}(q_n) \approx \left(\frac{1}{N}\sum_{i=0}^{N-1} q_i^2\right) - (\bar{q})^2,$$

$$P_{band} \triangleq P_{\theta^*}(q_n \in B_d) \approx \frac{1}{N}\sum_{i=0}^{N-1} I\{q_i \in B_d\}, \quad P_{idle} \triangleq P_{\theta^*}(q_n = 0) \approx \frac{1}{N}\sum_{i=0}^{N-1} I\{q_i = 0\},$$

$$\bar{\lambda}_c \triangleq E_{\theta^*}[\lambda_c(n)] \approx \frac{1}{N}\sum_{i=0}^{N-1} \lambda_c(i), \quad J(\theta^*) \approx \frac{1}{N}\sum_{i=0}^{N-1} |q_i - N_0|,$$

where N is taken as 10^5 in our experiments. The last performance measure is the one that the algorithm seeks to minimize, but clearly the others are closely related and are included here because they are often taken as measures of performance in ABR service. One desires P_{band} to be high in order to satisfy the various other performance criteria. One expects as a consequence of the above minimization that this quantity will be maximized. The measure P_{idle} gives the stationary probability of the server lying idle and should be close to zero. The average ABR throughput rate $\bar{\lambda}_c$ is often considered the most important measure of performance in ABR, because it is this measure which tells us whether the available bandwidth has been properly utilized or not. We now illustrate our numerical experiments.

In the simulations for the five-level policies, $\lambda_1, .., \lambda_4 \in [0.10, 3.0]$, with $\lambda_5 \in [0.10, 0.90]$. We consider the following setting for the uncontrolled traffic: $\lambda_{u,1} = 0.05$, $\lambda_{u,2} = 0.15$, $p(1;1) = p(1;2) = p(2;1) = p(2;2) = 0.5$. Thus, the mean uncontrolled traffic rate ($\bar{\lambda}_u$) is 0.1. The service rate $\mu = 1.0$. Further, the constants N_0 and ϵ are chosen to be 10 and 1, respectively. Thus the band B_d is $[9, 11]$. For the SPSA algorithm (4), in the step-size sequence $\{b(n)\}$ in (2)-(3), α is chosen to be 2/3. Thus $b(0) = 1$ and $b(i) = i^{2/3}$, $i \ge 1$. Moreover, the perturbation parameter δ is chosen to be 0.12. Throughout, 'O.L.' represents the optimal open loop policy. We present various experiments for the cases $D_b = D_f = 0$ with increasing T and F_b (Table 1) and increasing D_b, D_f with fixed T and F_b (Table 2). We also show experiments (Table 3) with two controlled sources feeding into the same bottleneck node but with rate

($\lambda_c(n)$) information fed back with different delays (almost without any delay to the first and with a significant delay to the second). The bandwidth in this case is found to be shared equally by the two sources. This amounts to our scheme showing fairness in performance. However, we are aware of the fact that the appropriate framework to study fairness is in tandem queues with different ABR sources feeding packets through different sets of nodes.[16] For small D_b, D_f, T and F_b, our algorithm converges in about 130-150 iterations while for large values of the above parameters, the algorithm takes about 200-250 iterations for convergence. On a Sun Ultra10 workstation with UNIX operating system, the algorithm takes about 5-10 minutes for small D_b, D_f, T and F_b, while for large values of these, it takes about 30-50 minutes for convergence. We also tried running the original two timescale stochastic approximation algorithm[4] (under the same settings of α and δ as the SPSA algorithm) with $D_b = D_f = 0$. It did not converge even after 350 iterations (but was close to it) after running for nearly 200 minutes. This confirms that the SPSA version of the two timescale stochastic approximation scheme shows faster convergence than the original two timescale scheme. Some of our important observations are as follows:

1. We get the best performance for lowest T, F_b, D_b and D_f (see Tables 1 and 2) below. The difference between the lowest and the highest rates (λ_5^* and λ_1^* respectively) decreases as any of T, F_b, D_b or D_f increase.

2. We utilize almost the entire bandwidth, viz., $P_{idle} \approx 0$, and $\bar{\lambda}_c + \bar{\lambda}_u \approx \mu$, even when T and F_b or the delays D_b and D_f are high (Tables 1 and 2).

3. The performance degrades as T and F_b or D_b and D_f increase, but remains better than the optimal open loop case even when these parameters become significantly high. For $T = 1$ and $F_b = 2$ (Table 2), performance is better than the optimal open loop case even for $D_b + D_f = 150$.

4. We considered the case of two controllers feeding arrivals into the same bottleneck node in addition to the uncontrolled MMPP stream (Table 3). Explicit rate information was fed back to the two sources with different delays D_{b1} and D_{b2}. Further there were different delays D_{f1} and D_{f2} in customers arriving to the bottleneck node from the two sources. We observed that the stationary mean rates $\bar{\lambda}_{c1}$ and $\bar{\lambda}_{c2}$, for the two sources are almost the same even when the difference in the delays is significantly high. This amounts to our scheme showing 'fairness' in performance. We also observed the other performance metrics in this case and found that the performance here is not as good as that of a single source with the lower delays (in the setting of Fig.1) and also it is not as bad as that of a single source with the higher delays. Thus in some sense the performance of the two sources here is getting averaged and which is possibly the reason for getting fairness in performance.

We also performed several other experiments with varying ϵ, N_0 and uncontrolled MMPP settings (in addition to those with eleven level policies).[5] These have not been reported here.

4. CONCLUSIONS

We studied the problem of rate based ABR flow control in the presence of information and propagation delays, in the continuous time queueing framework by developing and applying a numerically efficient two timescale SPSA algorithm. Relevant convergence theorems were stated and explained. Numerical experiments were conducted to verify and illustrate the theoretical results. In the numerical experiments, we addressed the problem (an approach common in the literature) by minimizing the distance of stationary mean queue length from a given constant N_0. Our experiments indicate that closed loop policies lead to a significant improvement in performance for reasonable values of information and propagation delays. The algorithm readily computes optimal policies amongst multilevel parametrized policies, even when there are significant delays, by using just two parallel simulations for any number of parameter levels. It was found that the scheme converges orders of magnitude faster than a previously proposed two timescale stochastic approximation scheme.[4] We also considered experiments with two ABR sources sharing the same bottleneck node but with the two sources experiencing significantly different propagation and information delays. We found that the sources under stationarity share the bandwidth equally between them. This is a very interesting observation and it amounts to our scheme exhibiting fairness in performance. The same should however be tried on tandem queues[16] to conclusively demonstrate the above claim.

ACKNOWLEDGEMENTS

This research was supported by the NSF under Grant DMI-9713720, by the Semiconductor Research Corporation under Grant 97-FJ-491 and by DoD contract Number MDA90497C3015.

REFERENCES

1. E. Altman, T. Basar, and R. Srikant. Robust rate control for ABR sources. In *Proceedings INFOCOM 98, San Francisco, CA*, March 1998.
2. J. D. Bartusek and A. M. Makowski. On stochastic approximation driven by sample averages: convergence results via the ODE method. Technical Report, Institute for Systems Research, University of Maryland, http://www.isr.umd.edu/TechReports/ISR/1994/TR_94-4/, 1994.
3. S. Bhatnagar. Multiscale stochastic approximation schemes with applications to ABR service in ATM networks. Doctoral dissertation, Dept. of Electrical Engineering, Indian Institute of Science, Bangalore, India, July 1997.
4. S. Bhatnagar and V. S. Borkar. Multiscale stochastic approximation for parametric optimization of hidden Markov models. *Probability in the Engineering and Informational Sciences*, 11:509–522, 1997.
5. S. Bhatnagar, M. C. Fu, and S. I. Marcus. Optimal multilevel feedback policies for ABR flow control using two timescale SPSA. Technical Report, Institute for Systems Research, University of Maryland, http://www.isr.umd.edu/TechReports/ISR/1999/TR_99-18/ (shortened version submitted for journal publication), 1999.
6. V. S. Borkar. Stochastic approximation with two time scales. *Systems and Control Letters*, 29:291–294, 1997.
7. E. K. P. Chong and P. J. Ramadge. Stochastic optimization of regenerative systems using infinitesimal perturbation analysis. *IEEE Trans. on Autom. Contr.*, 39(7):1400–1410, 1994.
8. A. I. Elwalid and D. Mitra. Statistical multiplexing with loss priorities in rate based congestion control of high speed networks. *IEEE Trans. on Comm.*, 42(11):2989–3002, 1994.
9. M. C. Fu. Convergence of a stochastic approximation algorithm for the $GI/G/1$ queue using infinitesimal perturbation analysis. *J. Optim. Theo. Appl.*, 65:149–160, 1990.
10. S. Kalyanaraman. Traffic management for the available bit rate (ABR) service in asynchronous transfer mode (ATM) networks. Ph.D. dissertation, Dept. of Computer and Information Sciences, The Ohio State University, August 1997.
11. A. Kolarov and G. Ramamurthy. A control theoretic approach for high speed ATM networks. In *Proceedings of IEEE Infocom97*, pages 293–301, Apr 1997.
12. H. J. Kushner and G. G. Yin. *Stochastic Approximation Algorithms and Applications.* Springer Verlag, New York, 1997.
13. R. S. Pazhyannur and R. Agrawal. Feed-back based flow control of B-ISDN/ ATM networks. *IEEE JSAC*, pages 1252–1266, 1995.
14. P. J. Schweitzer. Perturbation theory and finite Markov chains. *J. Appl. Prob.*, 5:401–413, 1968.
15. V. Sharma and J. Kuri. Stability and performance analysis of rate based feedback flow control of ATM networks. *To appear in QUESTA*, 1999.
16. K. Y. Siu and H. Y. Tzeng. Intelligent congestion control for ABR service in ATM networks. *ACM SIGCOMM Computer Communication Review*, pages 81–106, 1995.
17. J. C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Autom. Cont.*, 37(3):332–341, 1992.
18. Y. T. Wang and B. Sengupta. Performance analysis of a feed back congestion control policy. *ACM SIGCOMM Computer Communication*, pages 149–157, 1991.

Author for Correspondence: Email: marcus@isr.umd.edu; Fax: (301) 314-9920

Table 1: $D_b = D_f = 0$

T	F_b	λ_5^*	λ_4^*	λ_3^*	λ_2^*	λ_1^*	$\bar{q}$	P_{band}	σ_q	λ_c	P_{idle}	$J(\theta^*)$
1	2	0.1	0.3	0.7	1.7	2.6	10.1	0.45	7.3	0.90	0	2.01
1	5	0.2	0.5	0.8	1.3	1.8	9.5	0.34	12.8	0.90	0	2.71
1	10	0.3	0.6	0.9	1.2	1.5	10.0	0.26	21.3	0.89	0.01	3.52
1	O.L.	-	-	-	-	-	5.3	0.08	32.3	0.76	0.17	6.39
5	10	0.3	0.7	0.9	1.1	1.4	9.8	0.26	21.6	0.89	0.02	3.59
5	25	0.4	0.8	0.8	0.9	1.1	8.2	0.17	34.1	0.85	0.07	4.77
5	50	0.3	0.4	0.6	0.9	0.9	6.4	0.12	38.3	0.75	0.16	5.97
5	O.L.	-	-	-	-	-	5.1	0.09	31.9	0.76	0.16	6.39

Table 2: $T = 1$, $F_b = 2$

D_b	D_f	λ_5^*	λ_4^*	λ_3^*	λ_2^*	λ_1^*	$\bar{q}$	P_{band}	σ_q	λ_c	P_{idle}	$J(\theta^*)$
1	0	0.1	0.3	0.7	1.8	2.2	10.1	0.38	9.7	0.89	0	2.41
1	1	0.1	0.4	0.8	1.3	2.2	9.8	0.38	9.7	0.88	0	2.42
5	5	0.2	0.7	0.9	1.1	1.5	9.7	0.26	20.0	0.88	0.01	3.58
10	10	0.4	0.7	0.9	1.0	1.2	8.8	0.20	28.0	0.87	0.03	4.46
20	10	0.4	0.5	0.6	0.6	1.0	8.4	0.14	39.5	0.84	0.06	5.23
20	40	0.5	0.8	0.9	1.1	1.0	8.2	0.13	44.0	0.80	0.06	5.43
30	20	0.4	0.6	0.8	0.8	0.9	8.1	0.13	48.0	0.79	0.08	5.74
40	10	0.5	0.6	0.8	0.8	0.9	6.4	0.12	52.7	0.78	0.12	5.83
40	30	0.5	0.5	0.6	0.9	0.9	6.9	0.12	56.4	0.78	0.12	5.88
50	50	0.5	0.5	0.9	0.9	0.9	7.1	0.12	65.2	0.78	0.10	5.97
50	100	0.5	0.8	0.8	0.8	0.9	6.7	0.11	59.5	0.77	0.13	6.07
O.L.	-	-	-	-	-	-	5.3	0.08	32.3	0.76	0.17	6.39

Table 3 (Two ABR Sources): $T = 1$, $F_b = 2$

D_{b1}	D_{b2}	D_{f1}	D_{f2}	λ_5^*	λ_4^*	λ_3^*	λ_2^*	λ_1^*	$\bar{q}$	P_{band}	σ_q	λ_{c1}	λ_{c2}	P_{idle}
1	10	4	10	0.1	0.1	0.3	0.6	0.8	8.6	0.26	18.1	0.44	0.45	0.02
1	30	4	10	0.1	0.2	0.3	0.7	0.7	8.9	0.23	29.4	0.44	0.44	0.02
1	50	4	20	0.1	0.2	0.3	0.5	0.7	8.1	0.21	29.1	0.42	0.43	0.03
1	80	4	100	0.1	0.1	0.6	0.5	0.6	8.7	0.18	35.9	0.41	0.42	0.04