Resource Allocation Games in Connection-Oriented Networks under Imperfect Information

Aristos Aresti Bobby M. Ninan Mihail Devetsikiotis
Dept. of Electrical and Computer Engineering
North Carolina State University
Raleigh, NC 27695-7911
{aiaresti, bmninan, mdevets}@ncsu.edu

***Abstract*— Game-theoretic formulations of the resource allocation problem have existed for a while. However, the issue of perfect knowledge continues to be a significant hurdle on the path to realistic implementations. In this paper, the notion of playing bandwidth allocation games is investigated under *imperfect information*. Specifically we look at the case of connection oriented networks regulated by resource pricing. We devise a distributed adaptive control strategy based on dynamic estimation in order to cope up with the uncertainty of noise and delay. Simulation results illustrate the scalability and accuracy of the algorithms under multiple scenarios. Potential applications include teletraffic and optical networks, as well as *ad hoc* wireless networks, enabling users to partition bandwidth without the need of a centralized synchronizing entity.**

I. Introduction

Game theory has established itself to be an important tool for analyzing the problem of resource allocation in computer networks. The behavior of competing users in a shared environment comes under the aegis of noncooperative games. When each user maximizes her profit given the strategies of other users, the game settles down at the Nash equilibrium [1] providing a stable operating point for network services. Flow control and the related problem of congestion alleviation has been analyzed using both cooperative [2] and noncooperative flavors e.g. [3], [4]. In [5], the authors consider a Stackelberg game in which users choose routes in a wired network after the leader has chosen routes for its own traffic; in choosing, the leader controls user behavior to optimize network utility. In [6], the authors formulate a CDMA power data-rate control games for which the equilibrium point is studied. Rate-based flow control is also studied in [7]. Kelly in [8], [9] developed a pricing framework based on explicit congestion feedback. [10] examined TCP users in a wired Internet in a similar setting.

The existence of Nash equilibria in an environment populated by self-interested users rests on the assumption of perfect information. This requires every user to be aware of the state of all others at any point in time. Errors in measurement coupled with delay in information propagation can violate this condition. Users might also try to conceal their behavior so as to deny competitors of any unfair advantage. Hence any application of game theory based flow control strategies in a real world setting would have to address the issue of imperfect information. Alpcan et al. in [11] acknowledge the inherent restrictions of implementing cost functions in Internet-style networks and propose a scheme based on the variations in queuing delay of the individual user. In contrast, we derive our inspiration from optimal estimation based adaptive control techniques for stochastic dynamical systems [12], [13]. An analogous approach involving the Kalman filter was employed by Alouf et al. in [14] for the on-line estimation of dynamic multicast groups.

In this paper, we extend our bandwidth allocation model developed for connection oriented networks in [15], [16] to encompass an imperfect information regime, characteristic of real world settings. We propose a distributed adaptive algorithm to be used by each user to attain her desired optimum in the presence of uncertainty. The instantaneous bandwidth utilization observed while entering the network is employed to infer the system state through recursive estimations. Users then modify their arrival rates to maximize their individual utilities. We also develop two variants of the original algorithm to account for the curvilinearity of the input-output response. These adaptive control schemes were simulated for making comparisons of scalability and performance under various system parameters. Results indicate the algorithms converge even in the presence of uncertainty about the number of other players and their strategies. Applications of this algorithm include teletraffic, wireless and optical networks, enabling users to partition bandwidth without the need of a centralized synchronizing entity.

We use the following notation throughout this paper. Vectors are represented by bold-face letters. If Λ is a vector, $\Lambda(i)$ represents its i^{th} iterate, λ_i its i^{th} component and λ a generic component. Optimal entities are marked by an asterisk, for example Λ^*.

The rest of the paper is organized as follows: In the next section, we present a bandwidth allocation model to represent user behavior in an optical network setting. We then develop our observer based feedback control in section III. After motivating a dynamic estimation algorithm in IV, we propose our distributed adaptive control strategies in V. Simulation results for different scenarios are presented in section VI. We conclude by providing a summary and a discussion on future work.

II. Bandwidth Allocation Model

We consider a loss network scenario where N users compete for a finite amount of bandwidth K as expostulated in [15], [16]. Each user has a utility function $U_i(\theta)$ which is maximized at an optimal bandwidth θ_i^*. Requests for bandwidth arrive as a Poisson process with rate λ_i. Queuing is not a concern here as unfulfilled user requests depart the system immediately. The behavior of competing users can then be modelled as a noncooperative game wherein each player strives to attain her optimal θ_i^* by regulating her sending rate λ_i. Under the assumption that total demand for bandwidth is not greater than supply $\sum_{i=1}^{N} \theta_i^* \leq K$, the mean bandwidth allocated to the i^{th} user is given by Little's formula as

$$\theta_i(\Lambda) = \frac{\lambda_i}{\mu_i}(1 - \mathcal{E}(\rho(\Lambda))). \tag{1}$$

Here Λ is the user arrival rate vector $[\lambda_1, \ldots, \lambda_N]$ and μ_i is the exponential service rate for user i. The traffic intensity is defined as $\rho \equiv \sum_{i=1}^{N} \frac{\lambda_i}{\mu_i}$ and blocking probability as $\mathcal{E}(\rho, K) \equiv \frac{\rho^K/K!}{\sum_{k=0}^{K} \rho^k/k!}$. As derived in [16], in a perfect information regime where each user is aware of every other's sending rates, the user arrival rate vector converges to the non-cooperative Nash equilibrium Λ^*

$$\lambda_i^* \equiv \frac{\mu_i \theta_i^*}{1 - \mathcal{E}(\rho(\Lambda^*))} \quad \forall \quad i = 1, \ldots, N \tag{2}$$

The postulation of perfect information is often violated in real networks. The requirement that all the players need to be updated with the current information simultaneously places an enormous onus on the system. A distributed version of such an update protocol would involve $O(N^2)$ packets to be exchanged. Centralized broadcast of updates can reduce this to $O(N)$ but at the cost of increased synchronicity. Propagation and queuing delays cause staleness of information when user behaviors are highly variable. Internet connections are very often bursty and short lived thereby forcing us to examine the less tractable imperfect information scenario.

III. Feedback based rate control

We now detail an observer based control scheme for the bandwidth allocation model system defined above. Under imperfect information, each user i is aware of only her individual tuple of variables - $(\mu_i, \lambda_i, \theta_i(k))$. While the service rate μ_i is characteristic to the user, the mean bandwidth consumed θ_i can be ascertained from its instantaneous value $\theta_i(k)$ at the k^{th} time step. Since the information about the available bandwidth is obtained only when a new arrival enters the network, the system is modelled using difference equations. Each user sees the network as a dynamical system evolving in time with ρ as the system state. She tries to estimate the state using observations $\theta_i(k)$ which she employs to compute an observer based feedback control λ_i designed to attain her optimal bandwidth θ_i^*. The effect of other users on the system is modelled as noise which can be observed as perturbations of ρ from the hypothesized value. The state equation thus becomes

$$\rho(k+1) = \rho(k) - \frac{\lambda_i(k)}{\mu} + \frac{\lambda_i(k+1)}{\mu} + v(\rho(k)) \tag{3}$$

where $v(\rho(k))$ is the state dependent noise. Note that the noise is not statistically independent and hence cannot be modelled as Gaussian. In the absence of estimation errors due to uncertainty, the measured output can be related to the state and control using (1). The user calculates her new feedback control using the estimate of state $\hat{\rho}(k+1)$,

$$\lambda_i^{k+1} = \frac{\alpha_i(k+1)\theta_i^* \mu_i}{1 - \mathcal{E}(\hat{\rho}(k+1))} \tag{4}$$

Since α_i determines how fast each user tries to capture her optimal bandwidth, it can be considered as an indicator of user "aggressiveness". We thus define

$$\alpha_i(k) \equiv 1 - (\beta_i)^k, \quad 0 \leq \beta_i < 1. \tag{5}$$

In a perfect information scenario, the user would eventually attain her optimal bandwidth as $\lim_{i \to \infty} \alpha_i = 1$.

If the solution of the fixed point equation (2) is known to all the users, the user rates get decoupled. Thus the players move to the Nash equilibrium by the modified state equation

$$\lambda_i(k+1) = \lambda_i(k) + \eta(\lambda_i^* - \lambda_i(k)), \quad 0 < \eta \leq 1 \tag{6}$$

However in a decentralized game, each user is only aware of the dynamics of his individual state and observer equations. Thus the final equilibrium if attained may not be the Nash Equilibrium Point (NEP) as above.

IV. Dynamic estimation

The key ingredient in calculating an observer based feedback control is the estimation algorithm which provides an accurate estimate of the system state with minimal computational and storage requirements. The user collects observations $\{\theta(0), \theta(1), \ldots, \theta(k)\}$ which she employs to estimate the time varying system state ρ^k. A recursive estimation procedure is vital in reducing the observation history to be maintained at each point in time. We make a first order approximation of the relationship between the inputs and measured outputs by assuming θ to be linear in ρ i.e $\theta(k) \simeq h(k)\rho(k) + w(k)$ where $h(k)$ is the design parameter and $w(k)$ is the unknown noise value. Each user then strives to reduce her least squares error

$$L(\rho, k) = \frac{1}{2k} \sum_{j=1}^{k} \gamma^{k-j}(\theta(j) - h(j)\rho(j))^2 \tag{7}$$

for the weight $0 \leq \gamma < 1$. Eqn. (7) exponentially de-weights past measurements indicating that greater importance is placed on current measurements. $L(\rho, k)$ is minimized by the classic Recursive Least Squares (RLS) algorithm which produces a time-varying estimate of the system state ρ [13] as

$$\hat{\rho}(k) = \hat{\rho}(k-1) - K_k(\lambda^k \hat{\rho}(k-1) - \theta(k)) \tag{8}$$

where $K_k = P_k \lambda(k)$ and $P_k = \frac{P_{k-1}}{\gamma}(1 - \frac{P_{k-1}^2(\lambda(k))^2}{\gamma + (\lambda(k))^2 P_{k-1}})$. The algorithm is initiated with $\hat{\rho}(0) = 0$, $P_0 \gg 0$.

0-7803-8533-0/04/$20.00 (c) 2004 IEEE

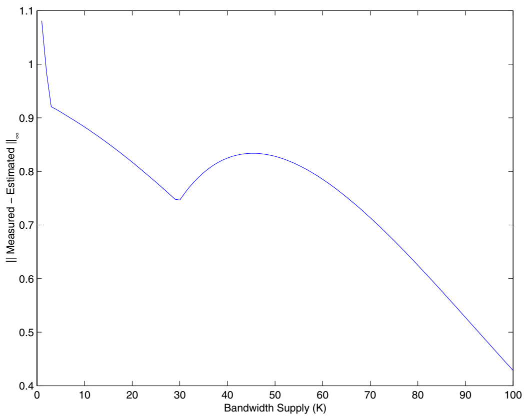

Fig. 1. Comparison of measured and estimated outputs for 2 users.

V. Algorithms

We detail three algorithms below which would then be compared with respect to their convergence and accuracy among others.

A. Original algorithm

This method estimates system state using (8) while the control is calculated by approximating (4) as

$$\lambda_i(k+1) = \frac{\alpha_i(k+1)\theta_i^*\mu_i}{1-\mathcal{E}(\hat{\rho}(k))} \tag{9}$$

We denote it as the "original algorithm".

B. Logarithmic variant

In microeconomics, the relationship between an economic output (y) and its inputs (x) is often described by a Cobb-Douglas type of production function $y = C\prod_{i=1}^{n} x_i^{a_i}$. Thus there exists a linear relationship between the logarithmic values of input and output namely $\log y = \log C + \sum_{i=1}^{n} a_i \log x_i$. The correspondence between user inputs and observations for a two player scenario was analyzed as follows. For symmetric users, we collected multiple input-output data and computed the exponents for equation $\theta = C\lambda_1^{a_1}\lambda_2^{a_2}$ by least squares fitting.

Fig. 1 illustrates the sup-norm of the error between the measured and estimated outputs for various values of K. It is clear that the input-output relationship can be approximated by a log-linear one. The state estimation equation can then be rewritten as

$$\log\hat{\rho}(k) = \log\hat{\rho}(k-1) - K_k(\lambda(k)\log\hat{\rho}(k-1) - \log\theta(k))$$

The time varying feedback control is computed using (9). We denote the above algorithm as the "logarithmic variant".

C. Newton-Raphson variant

Another approximation of (4) could be carried out by ignoring the effect of the external noise $v(\rho(k))$ on the evolution of system state. Thus

$$\rho(k+1) = \hat{\rho}(k) - \frac{\lambda(k)}{\mu} + \frac{\lambda(k+1)}{\mu} \tag{10}$$

Substituting (10) in (4), we obtain

$$\lambda(k+1) = \frac{\alpha(k+1)\theta^*\mu}{1-\mathcal{E}(\rho(\hat{k}) - \frac{\lambda(k)}{\mu} + \frac{\lambda(k+1)}{\mu})}$$

The task of computing the new control thus reduces to finding the root of $g(\lambda(k+1))$

$$g(\lambda(k+1)) \equiv \lambda(k+1) - \frac{\alpha(k+1)\theta^*\mu}{1-\mathcal{E}(\hat{\rho}(k) - \frac{\lambda(k)}{\mu} + \frac{\lambda(k+1)}{\mu})}.$$

We employ the damped Newton-Raphson method for this purpose. Denoting $\delta(n)$ as the n^{th} iterate of $\lambda(k+1)$ and κ $(0 \le \kappa < 1)$ as the damping coefficient, the iterations are

$$\delta(n+1) = \delta(n) - \kappa\ g'(\delta(n))^{-1} g(\delta(n)) \tag{11}$$

where

$$g'(\delta(n)) = 1 - \alpha(k+1)\theta^*[((\frac{K}{\hat{\rho}(k) - \frac{\lambda(k)}{\mu} + \frac{\delta(n)}{\mu}} - 1)\mathcal{E}(\hat{\rho}(k) - \frac{\lambda(k)}{\mu} + \frac{\delta(n)}{\mu}) + \mathcal{E}^2(\hat{\rho}(k) - \frac{\lambda(k)}{\mu} + \frac{\delta(n)}{\mu})]/(1-\mathcal{E}(\hat{\rho}(k) - \frac{\lambda(k)}{\mu} + \frac{\delta(n)}{\mu}))^2$$

and

$$g(\delta(n)) = \delta(n) - \frac{\alpha(k+1)\theta^*\mu}{1-\mathcal{E}(\hat{\rho}(k) - \frac{\lambda(k)}{\mu} + \frac{\delta(n)}{\mu})}$$

Starting from $\delta(0) = \frac{\alpha(k+1)\theta^*\mu}{1-\mathcal{E}(\hat{\rho}(k))}$, we compute the successive approximations of $\lambda(k+1)$ using (11). Termination of the iterations is contingent on the condition of the error falling below the tolerance threshold ξ i.e $|\delta(n+1) - \delta(n)| \le \xi$. The "Newton-Raphson variant" thus estimates system state using (10) and computes the corresponding control by (11).

VI. Results

Computer simulation was employed to investigate the behavior of the algorithm under several scenarios. For ease of analysis, the system considered was composed of a homogenous population of identical users. Unless mentioned otherwise, the default values used throughout the simulation were: $\theta^* = 20$, $\mu = 1$, $\beta = 0.1$, $\mathsf{P}_0 = 10^3$, $\xi = 10^{-3}$, $\kappa = 0.1$. Each different scenario was executed for 1000 repetitions to obtain statistically significant results. For stochastic optimization methods like the ones used above, measurement noise makes it unrealistic to expect the algorithms to converge to a single value. We thus employed the following stopping criterion in our experiments. The algorithm is presumed to have stabilized at iteration n when 10 successive values of the system state are at most 0.1 apart from each other

$$|\rho(n) - \rho(n-j)| \le 0.1 \quad \forall \quad 1 \le j \le 10$$

The maximum number of iterations until convergence was set to 10^5 beyond which it was terminated and considered as not converging.

The cardinal outcome of the experiments was the convergence of the algorithms in a noisy environment, ruling out the occurrence of system instabilities such as oscillations and finite time singularities. This is impressive considering the fact that each user pursues her utility optimization oblivious to the number of competitors or the algorithms adopted by them to attain their objectives. Further the equilibria under uncertainty are close to the Nash equilibrium in a perfect information regime. Other characteristics of the algorithms are detailed below.

A. Scalability

The scalability of an algorithm determines the speed at which it stabilizes the system for increasing number of users. The original algorithm does not scale well and does not converge when the number of users are greater than 5. Further as seen in Fig. 4, the number of iterations required for convergence are two orders of magnitude more than the two variants. This instability thus rules out its deployment in a real world setting forcing us to exclude it from further consideration. The logarithmic and Newton-Raphson variants scale well with N and are compared in Fig. 2 indicating the superiority of the former over the latter. The effect of other users on the constitution of system state becomes prominent as the number of users proliferate. This affects the validity of assumption (10) and the scalability of the Newton-Raphson version.

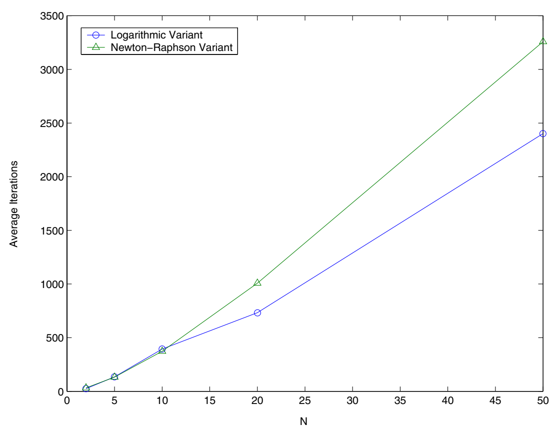

Fig. 2. Iterations vs Number of users: Logarithmic and Newton-Raphson.

B. Accuracy

The effect of noise and delay in an imperfect regime is to move the system away from the non-cooperative Nash equilibrium. Hence a suitable metric to quantify the superiority of the algorithms would be the deviation of their equilibria (Λ') from zero in the fixed point iteration of (2). Accuracy is then quantified by computing the residuals $||\Lambda' - F(\Lambda')||_\infty$

$$F_i(\Lambda') \equiv \frac{\mu_i \theta_i^*}{1 - \mathcal{E}(\rho(\Lambda'))} \quad \forall \quad i = 1, \ldots, N$$

The error is proportional to the residuals and is depicted in Fig. 3. The error values indicate that the variants provide a reasonable approximation to the Nash Equilibrium point. As opposed to the Newton-Raphson version, the logarithmic algorithm displays increasing accuracy with respect to number of users . This is due to the progressively better log-linear approximation to the control-observation function as evident in Fig. 1. The increasing accuracy leads us to suspect that the logarithmic variant might asymptotically converge to the Nash equilibrium under certain limits.

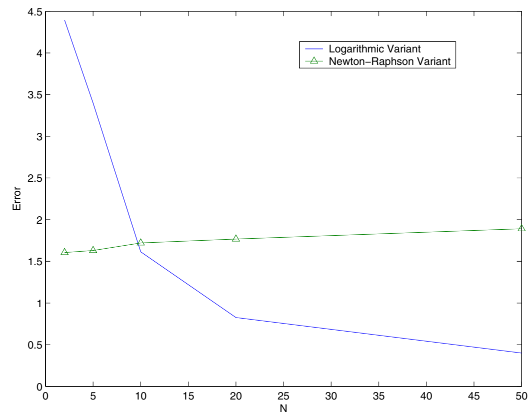

Fig. 3. Error vs Number of users: Logarithmic and Newton-Raphson variants.

C. Effect of Aggressiveness on System Convergence

User aggressiveness is characterized by β as defined in (5). It modifies the rate of user aggression as

$$\frac{d\alpha_i}{dk} = -\beta^k \log(\beta)$$

Hence as it approaches zero, players try to attain their optimal bandwidth more aggressively and in fewer time steps. The effect of aggression on system convergence for the three algorithms is shown in figs. 4, 5, 6. The number of iterations required for convergence increase as the variants become less aggressive, causing them to move slowly towards their equilibrium points.

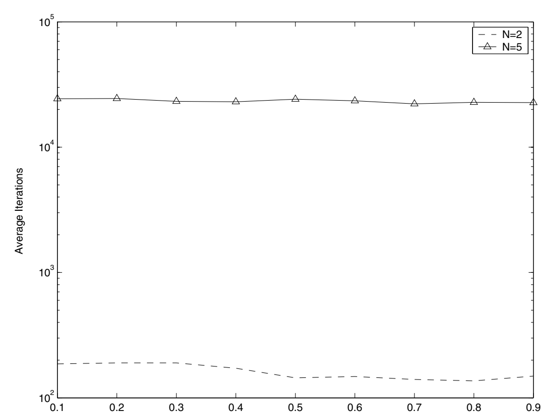

Fig. 4. Iterations vs β, Original algorithm.

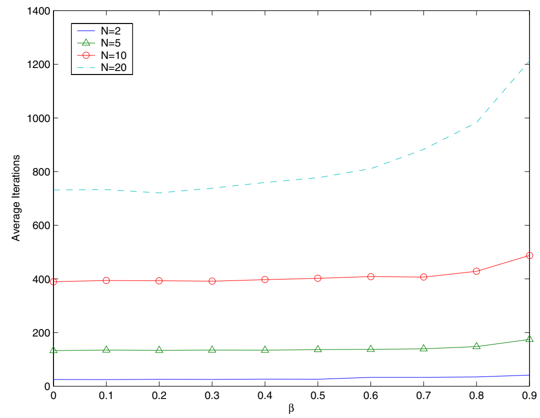

Fig. 5. Iterations vs β, Logarithmic variant.

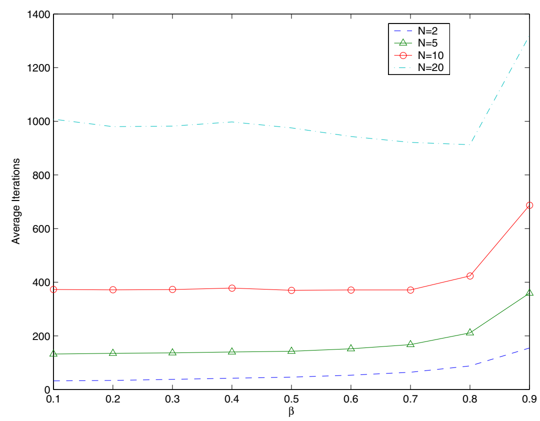

Fig. 6. Iterations vs β, Newton-Raphson variant.

D. Impact of User Demand on System Convergence

User demand is characterized using ϵ defined as $\epsilon \equiv 1 - \frac{\sum_{i=1}^{N} \theta_i^*}{K}$. As we vary ϵ from 1 to 0, user demand approaches the supply limit. We investigated the effect of ϵ on system convergence for the Logarithmic variant and N = 2,5 and 10 users as illustrated in Fig. 7. The Newton-Raphson version also displayed analogous behavior. As the demand for bandwidth increase, it restricts the adaptability of each user. Any slight perturbation in demand due to noise becomes amplified slowing the convergence of the distributed algorithm.

VII. Conclusions and Future Work

In this paper, we investigated the notion of playing bandwidth allocation games under imperfect information. Specifically we looked at the case of loss networks regulated by resource pricing and devised a distributed adaptive control strategy based on dynamic estimation. Simulation results illustrated the scalability and accuracy of the algorithms under multiple scenarios.

We are currently working towards extending our model to incorporate adaptive pricing as a possible tool for network control. The dynamic estimation algorithm may be enhanced with nonlinear stochastic approximation techniques. The effect of myopic user adaptation on global system behavior could

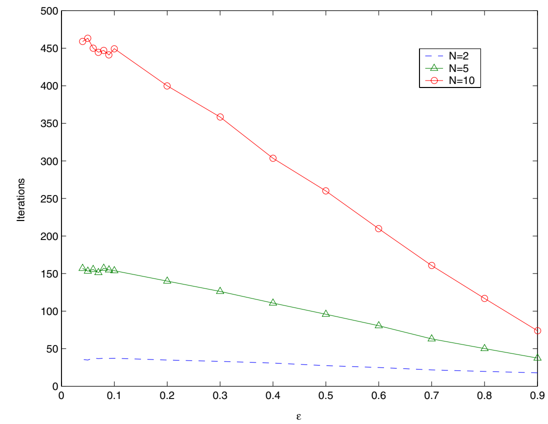

Fig. 7. Iterations vs ϵ, Logarithmic variant.

have implications for adaptive algorithms like the myriad TCP variants populating today's Internet.

References

[1] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd ed. Philadephia, PA: SIAM, 1999.

[2] H. Yaiche, R. R. Mazumdar, and C. Rosenberg, "A game theoretic framework for bandwidth allocation and pricing in broadband networks," *IEEE/ACM Transactions on Networking*, 2000.

[3] C. Douligeris and R. Mazumdar, "Multilevel flow control of queues," in *Johns Hopkins Conference on Information Sciences*, Baltimore, MD, March 1989.

[4] A. Economides and J. Silvester, "Priority load sharing: an approach using Stackelberg games," in *Proc. Allerton Conference on Communication, Control, and Computing*, 1990.

[5] Y. Korilis, A. Lazar, and A. Orda, "Achieving network optima using Stackelberg routing strategies," *IEEE/ACM Transactions on Networking*, vol. 5, no. 1, 1997.

[6] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 291–303, February 2002.

[7] E. Altman and T. Basar, "Mutli-user rate-based flow control," *IEEE Transactions on Communications*, vol. 46, no. 7, pp. 940–949, July 1998.

[8] F. P. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, pp. 33–37, 1997.

[9] F. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.

[10] R. La and V. Anantharam, "Charge sensitive TCP and rate control in the Internet," in *Proc. IEEE INFOCOM*, 2000.

[11] T. Alpcan and T. Basar, "A utility-based congestion control scheme for internet-style networks with delay," in *Proc. of INFOCOM 2003*, June 2003.

[12] R. F. Stengel, *Optimal Control and Estimation*. Dover Publications, 1994.

[13] J. C. Spall, *Introduction to Stochastic Search and Optimization*, ser. Discrete Mathematics and Optimization. Wiley-Interscience, 2003.

[14] S. Alouf, E. Altman, and P. Nain, "Optimal on-line estimation of the size of a dynamic multicast group," in *Proc. of INFOCOM 2002*, vol. 2, June 2002, pp. 1109 –1118.

[15] B. M. Ninan, G. Kesidis, and M. Devetsikiotis, "A simulation study of non-cooperative pricing strategies for circuit-switched optical networks," in *Proc. ACM/IEEE MASCOTS*, 2002.

[16] B. M. Ninan and M. Devetsikiotis, "Pricing mediated bandwidth allocation for the Next Generation Internet," in *GLOBECOM 2003*, San Francisco, 2003.

 0-7803-8533-0/04/$20.00 (c) 2004 IEEE