Joint Adaptive Rate Control and Randomized Scheduling for Multimedia Wireless Systems

M. A. Haleem and R. Chandramouli
Multimedia System, Networking, and Communications (MSyNC) Laboratory
Department of Electrical and Computer Engineering
Stevens Institute of Technology, Hoboken, NJ 07030
Email: {mhaleem,mouli}@stevens.edu

***Abstract*— A joint channel adaptive rate control and randomized scheduling algorithm based on learning automata (LA) [4] is presented. The scheduling is performed at the medium access control (MAC) layer whereas the rate selection takes place at the Physical/Link (PHY/LINK) layer. The two components residing in the two layers exchange minimal amount of information and adaptively achieve the best throughput and desired quality of service (QoS) in terms of average transmission rates in the prevailing channel conditions. Scheduling is carried out by a LA of *continuous reward penalty* variate, and a *discrete pursuit reward inaction* (DPRI) type [9] is used for adaptive rate selection. While simple to implement, this technique requires no explicit channel estimation phase. The only feedback required are the single bit ACK signal indicating the correct reception of packets. As shown in the convergence theorems, the algorithm achieves optimal performance in "stationary" channels. With slowly varying channels, the MCS selection algorithm sees a "quasi-stationary" channel and adaptively converges to the optimality. Simulation results are provided for parameters as per to HSDPA standard.**

I. Introduction

The primary goal in optimizing a multi-user wireless communication system is to maximize the system throughput (successful bits/second) with limited resources such as transmission power, bandwidth, and hardware complexity. Adaptive rate selection has been of interest in 3G wireless systems [1], [2], and [3]. The requirement therein is to adaptively choose among the set of available modulation and coding schemes (MCS) defining the set of rates, the MCS that maximizes the throughput for the time varying wireless channel. The best performance achievable is of a scheme in which the receiver estimates and feeds back the channel state information to the transmitter prior to transmission of each data packet. For such a system to achieve optimality, the delay involved in the process of channel estimation and feedback must be negligible compared to the time scales of variations of the channel. Estimation errors, errors in feedback, and feedback delays are among the obstacles to overcome to achieve optimal throughput. Adaptive techniques become good alternatives as they can operate when the channel state is unknown or only partially known and they perform excellently in low mobility environments.

Further, user satisfaction requires a guaranteed throughput. Recent work on "throughput optimization subject to fairness in service" such as in [11] and [12] involve feedback of an index representing the received signal to interference plus noise ratio (SINR) by all the users prior to each transmission. The transmitter (base station) then schedules one or more users with the best SINR values during the next time slot. When it comes to fairness of service, one can think of many different metrics such as "equal time", "equal rate", and "proportional fairness" [12]. In a wireless multimedia environment however, the rate requirement of a user depends on the application – streaming video, voice, text messages, and web browsing to name a few. In such a scenario, fairness of service means admission of a user with a throughput guarantee to support the application in concern. Thus the overall requirement of an efficient multi user multimedia system is to maximize the overall throughput subject to the quality of service (QoS) guarantee in terms of the individual throughput values of the users to support the applications of interest. Also worthwhile to note in this context, is the recent interests in the optimization of wireless systems with joint consideration of multiple levels in the layered systems architecture, leading to the phrase, "cross layer optimization" [14].

In this paper, we propose and analyze a two-level stochastic control algorithm based on *learning automata*(LA) [4] for adaptive MCS selection and user scheduling in each "time slot". This algorithm adaptively learns and chooses the best MCS to maximize the throughput in the prevailing channel condition of a user, and computes a randomized schedule to achieve the requested throughput of each user. The adaptive learning of MCS is to be implemented in the physical/Link (PHY/LINK) layer whereas user scheduling is carried out in medium access control (MAC) layer with information exchange between these two layers.

In section II to follow, we present the formulations of the algorithm. Section III establishes the theorems on convergence of the algorithm in stationary and time varying wireless channels. The simulation results illustrating the performance are given in section IV. Conclusions follow in section V.

II. Joint Randomized Scheduling and Stochastic Rate Selection

A LA maintains an action (control) probability vector $p(n) = [p_1(n), p_2(n), \cdots, p_r(n)]$ to select an action among a set of actions at the time (iteration) n. Following an action

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

at time n, the automaton receives a feedback indicating a *reward* (success) or a *penalty* (failure). The change in the probabilities in each update can be continuous in [0,1] or of discrete values. In a *continuous reward penalty* (CRP) LA, the probability $p_i(n)$ of action i is updated based on this feedback. In the class of LA of *discrete, reward-inaction* type, there is no update in $p(n)$ for a penalty while it is updated by a discrete amount for a reward. Such an approach has been observed to have superior convergence properties [6]. As introduced by Oommen and Lanktot in [9], a good policy to update the probability vector $p(n)$ is of a *pursuit* algorithm that always rewards the action with current minimum penalty estimate, or in other words, the one that "pursues" in the direction of best reward. It has been shown that stochastic automaton of *discrete pursuit reward-inaction* (DPRI) type outperforms others in speed of convergence. In the following subsections we present the details of the proposed algorithm for user scheduling and MCS selection.

A. Randomized User Scheduling

The user scheduler assumes a *best effort* queue and a set of K users each with a throughput requirement. Users to be served in each time slot is selected according to a service probability vector, $p^s(n) = [p_0^s(n), p_1^s(n), \cdots, p_K^s(n)]$ where $p_0^s(n)$ is the probability associated with the "best effort queue". Service probability of a user is updated following transmission of P packets or when the time elapsed since the last update is Q slots, whichever occurs first. The update is an increment/decrement in the probability of assignment of the user in concern. The increment/decrement of the probability of assignment is a function of the current value and the difference between the achieved short term throughput and the requested throughput. The increment/decrement of the probability is compensated for, by a corresponding decrement/increment in the probability of the "best effort queue". The user scheduling algorithm is summarized as follows.

Parameters

R^{req} the set of requested throughput values.
$R^{ave}(n)$ the set of achieved throughput values during the time interval $n - Q_0 + 1$ to n.
a, b, η scaling parameters ($0 < \{a, b\} \leq 1$ and $\eta > 0$)

Pseudo-code

Initialize $p_i^s(n) = 1/(K+1)$, for $0 \leq i \leq K$.
Repeat

1) Take control from PHY/LINK layer *rate adaptation* algorithm.
2) Select a user to serve in the next time slot with probabilities $p^s(n) = [p_0^s(n), p_1^s(n), \cdots, p_K^s(n)]$.
3) If $Q_0 = \min\{\text{time for } P, Q\}$ achieved for the selected user i, goto step 4; else goto step 7.
4) Compute $\nu = \frac{R_i^{req} - R_i^{ave}(n)}{\eta R_i^{req}}$.
5) Let

$$\beta_i(n) = \begin{cases} 1, & \text{if} \quad \nu > 1 \\ \nu, & \text{if } 1 > \nu > -1 \\ -1, & \text{if} \quad \nu < -1 \end{cases} \tag{1}$$

6) Update $p^s(n)$ according to the following.
for $\beta_i(n) > 0$,

$$p_i^s(n+1) = p_i^s(n) + a p_i^s(n) \beta_i(n) \tag{2}$$
$$p_0^s(n+1) = p_0^s(n) - a p_i^s(n) \beta_i(n) \tag{3}$$

for $\beta_i(n) < 0$,

$$p_i^s(n+1) = p_i^s(n) - b p_i^s(n) |\beta_i(n)| \tag{4}$$
$$p_0^s(n+1) = p_0^s(n) + b p_i^s(n) |\beta_i(n)| \tag{5}$$

7) Pass control to PHY/LINK layer *rate adaptation* algorithm for the scheduled user.

End Repeat

B. Adaptive Rate Assignement

Once the scheduler selects a user i based on the current probability vector $p^s(n)$, the rate adaptation algorithm randomly selects a rate (corresponding to an MCS) from the set of r rates, $R = \{R_j : j = 1, 2, \cdots, r\}$ following the probability vector $p_i^R(n) = [p_{i1}^R(n), \cdots, p_{ir}^R(n)]$ of user i. The DPRI algorithm for adaptive MCS selection for user $i (0 \leq i \leq K)$ maintains a vector of running estimates of throughput values $\hat{D}_i(n) = [\hat{D}_{i0}(n), \hat{D}_{i1}(n), \cdots, \hat{D}_{ir}(n)]$ where $\hat{D}_{ij}(n) = R_j(1 - \hat{P}_{ij}^e(n))$ for $1 \leq j \leq r$. In this, $\hat{P}_{ij}^e(n)$ is the estimate of probability of frame error when j^{th} MCS is selected with the available SINR. This is computed as the ratio of the count of absent ACK signals to the count of instances the rate is selected, during the estimation "window". Fair estimation requires a sufficiently long estimation window. On the other hand, too long of an estimation window may result in sub-optimal performance as it fails to detect short term peaks and valleys in the time varying channel envelop. Initially, all the rates are assigned an equal probability of $1/r$. Then the rate selection proceeds with the fixed $p_i^R(n)$ until every MCS is selected at least M (a tunable parameter) number of times. Then the estimation of $\hat{P}_{ij}^e(n)$ and hence the update of $p_i^R(n)$ starts and continues. At a given time n, the frame error probabilities are computed to be $\hat{P}_{ij}^e(n) = \frac{1}{M} \sum_{k=L_{ij}(n)-M+1}^{L_{ij}(n)} I_{ij}(k)$ where $I_{ij}(k)$ is an indicator function s.t. $I_{ij}(k) = 0$ if an ACK is received following the transmission of a packet at k^{th} attempt and $I_{ij}(k) = 0$ otherwise. $L_{ij}(n)$ is the number of times the rate R_j is selected from $n = 0$ till time n. Following the transmission of each data frame, the probability $p_{im}(n)$ of the best rate R_m of user i is incremented by $(r-1)\Delta$ where $\Delta = \frac{1}{rN}$ is the smallest step size. N here is the resolution parameter. The proposed user rate selection algorithm can be summarized as follows.

Parameters

$L_{ij}(n)$ Number of times the rate R_j is selected from time 0 till n for user i.
$I_{ij}(k)$ "0" or "1" on receiving or not receiving ACK following the kth use ($1 \leq k \leq L_{ij}(n)$) of rate R_j.
B a bias to prevent $p_{ij}^R(n) = 0$ $\forall i$ and $\forall j$ to facilitate tracking of time varying channel.

Pseudo-code
Initialize $p_{ij}^R(n) = 1/r$, for all i and for all j.
Repeat

1) Take control from *user scheduler* at time n.
2) Pick a rate $R_j (1 \leq j \leq r)$ according to probability distribution $p_i^R(n)$.
3) Update $I_{ij}(k)$ and $L_{ij}(n)$ on receiving or not receiving an ACK signal.
4) Update $\hat{P}_{ij}^e(n+1)$ and thus $\hat{D}_{ij}(n)$according to the following.

$$\hat{P}_{ij}^e(n+1) = \frac{1}{M} \sum_{k=L_{ij}(n)-M+1}^{L_{ij}(n)} I_{ij}(k) \tag{6}$$

$$m = \arg\max_j \hat{D}_{ij}(n) \tag{7}$$

5) If $L_{ij}(n) \geq M$ for all j (initialization phase completed) goto step 6; else goto step 7.
6) Update $p_i^R(n)$ according to the following equations:

$$p_{ij}^R(n+1) = \max\{p_{ij}^R(n) - \Delta, B\}, \forall j \neq m \tag{8}$$

$$p_{im}^R(n+1) = 1 - \sum_{j \neq m} p_{ij}^R(n+1). \tag{9}$$

7) Pass control to *user scheduler* with the R_i^{ave} information.

End Repeat

III. Convergence of the Stochastic Algorithm

The convergence properties of CRP and DPRI algorithms are analyzed in [7] and [9] respectively. The proof of convergence therein are in the context of stationary channels i.e., for fixed $D_{ij}(n)$ so that the probability vectors $p^s(n)$ and $p_i^R(n)\forall i$ converge arbitrarily close to the optimum when allowed to run for sufficiently long time. We postulate that if the channel variations are sufficiently low relative to the speed of convergence, or in other words if the channel is *quasi-stationary*, the algorithm can adaptively optimize $p^s(n)$ and $p_i^R(n)\forall i$. Each time the channel state changes $p^s(n)$ and $p_i^R(n)\forall i$ undergo changes until a new optimum set of probabilities are achieved. Such changes require us to avoid any element of $p_i^R(n)\forall i$ from getting set to null. Thus we introduce the bias parameter B in the second-level automaton. Then the rate adaptation automaton learns the best rate for a user i such that $p_{im}^R(n)$ is arbitrarily close to $1-(r-1)B$. Optimality requires that the convergence completes within a time duration small compared to the duration the channel would stay in each state before a transition. Further, the user scheduling algorithm should augment the selection probability $p_i^s(n)$ along with $p_0^s(n)$, so that to bring the throughput of the user to the requested value with sufficient rapidity.

In the sequel we first show that the first level automaton of user scheduling algorithm can achieve the requested throughput of each user. Next we state the theorems on the convergence of the second level automaton of the channel adaptive rate assignment algorithm to the optimal rates.

Lemma 1: For any user i, with the optimal PHY layer transmission rate of R_m in a given channel state, and a sufficiently slowly varying channel let $p_i^{req,m}$ be the selection probability required by the randomized user scheduling algorithm to achieve the requested throughput, $R_i^{req} = R_m(1-P_{im}^e)p_i^{req,m}$. The iterative updating process converges such that $p_i^s(n) \rightarrow p_i^{req,m}$ w.p. 1 as $n \rightarrow \infty$.

Proof: From (2)-(5), the conditional expectation of the change in $p_i^s(n)$ can be expressed as

$$E[\Delta p_i^s(n)|p_i^s(n)] = \{aPr(\beta_i(n) \geq 0) \sum_{k=0}^{k_0} |\beta_i^K(n)| Pr(K=k) - b(1-Pr(\beta_i(n) \geq 0)) \sum_{k=k_0+1}^{Q_0} |\beta_i^K(n)| Pr(K=k)\} p_i^s(n)$$

where $\beta_i^K(n)$ stands for the value of $\beta_i(n)$ conditioned on the assignment of k out of Q_0 slots to user i. Further $Pr(K = k)$ is the probability of assigning k out of P last slots to user i (which follows a binomial distribution with success probability $p_i^s(n)$). k_0 is the number of slots required to be assigned to user i to achieve the requested rate R_i^{req}. Thus we have $Pr(\beta_i(n) \geq 0) = Pr(K \leq k_0)$ which monotonically decreases as $p_i^s(n)$ increases. Further $|\beta_i^{k_0}(n)| = 0$ and $|\beta_i^k(n)|$ is symmetric around $k = k_0$ and increases proportional to $|k-k_0|$. It can be easily seen that for any given k_0, the part of the expression for $E[\Delta p_i^s(n)|p_i^s(n)]$ within the curly brackets is a monotonically decreasing function of $p_i^s(n)$. Thus the parameters a and b can be tuned so that $E[\Delta p_i^s(n)|p_i^s(n) < p_i^{req,m}]$ is positive and $E[\Delta p_i^s(n)|p_i^s(n) > p_i^{req,m}]$ is negative. Thus by the sub/super-martingale convergence theorems [4] we conclude that the update process converges with probability one as $n \rightarrow \infty$ and can be made convergent to the desired value of $p_i^{req,m}$with the proper choice of parameters a and b. ■

In the above proof there is an underlying assumption that $p_0^s(n)$ of (3)is sufficiently large to achieve the desired $p_i^s(n)$ for all users $i = 1, \cdots, K$. Having established *Lemma* 1, on the convergence of user selection automaton, it remains to show that the rate selection automaton converges to optimality with quasi-stationary channels. To this end we state the following theorems without proof which are found in [9]. For simplicity, the user index i is omitted in the writing.

Theorem 1: Suppose there exists an index m and a time instance $n_0 < \infty$ such that $\hat{D}_m(n) > \hat{D}_j(n), \forall j \neq m$ and all $n \geq n_0$. Then there exists an integer N_0 such that for all resolution parameters $N > N_0$, $p_m(n) \rightarrow 1$ with probability one as $n \rightarrow \infty$ and $B \rightarrow 0$.

Theorem 2: For each rate, R_i, assume $p_i(0) \neq 0$. Then for any given constants $\delta > 0$ and $M < \infty$, there exist $N_o < \infty$ and $n_0 < \infty$ such that under the DPRI algorithm, for all learning parameters $N > N_0$ and all time $n > n_0$: P{each rate chosen more than M times at time n}$\geqslant 1-\delta$.

Theorem 3: In every stationary channel, the DPRI is ϵ-optimal. More explicitly, given any $\epsilon > 0$ and $\delta > 0$, there exists a $N_0 < \infty$ and a $n_0 < \infty$ such that for all $n \geqslant n_0$ and $N > N_0$: $P[|P_m(n) - 1| < \epsilon] > 1 - \delta$ as $B \to 0$.

In the original form of *Theorem 1*, the bias parameter B is not present as it has been for stationary "environments". Nevertheless, the proof with B requires only minor changes. *Theorem 1* states that convergence is achieved as $n \to \infty$. However, it is the *Theorem 3* that establishes the convergence to the optimality with a required degree within a finite time. *Theorem 2* is to prove that the condition for *Theorem 3* is achievable within finite time.

IV. Simulation Results

The simulation was carried out for a frequency flat fading radio link with a single antenna at the transmitter and a single antenna at the receiver. We present here the results for two specific channels namely a "stationary" channel and a "time varying" channel. The realizations of channels were generated for 20000 time slots (frames). We consider a set of six rates, $\{0.12, 0.24, 0.36, 0.48, 0.60, 0.72\}$ (Mb/s) which are typical in 3G wireless systems and wireless LANS such as 802.11b. The frame duration was taken to be 1 transmission time interval (TTI) which was 0.667ms, and the lengths of bit streams in each "data frame" was selected to match the transmission rate in each time slot. The parameter settings of user selection automaton were $P = 50, Q = 100, a = 0.2, b = 0.1,$ and $\eta = 0.2$. Those of rate selection automaton were $M = 10, N = 10$. The bias parameter B was set to "0" for static channels and to 0.01 for time varying channels. For a given SINR, the frame error probability, P_{ij}^e of the selected MCS j of user i were read out from a pre derived set of curves for transmission through additive white Gaussian noise (AWGN) channel. The ACK/NACK signals were generated based on such error probabilities.

The simulation results for the stationary channel as mentioned above are presented in Fig. 1 and 2. Included in the results are the performance of a MCS selection scheme based on perfect channel state information (PCSI) at the transmitter with no feed back delays or errors. In such a scheme the transmitter is considered to know the channel so

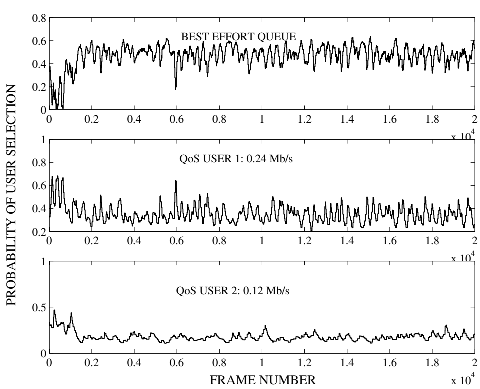

Fig. 1. Time evolution of scheduling probabilities in stationary channel; frame duration = 1 TTI (0.667ms).

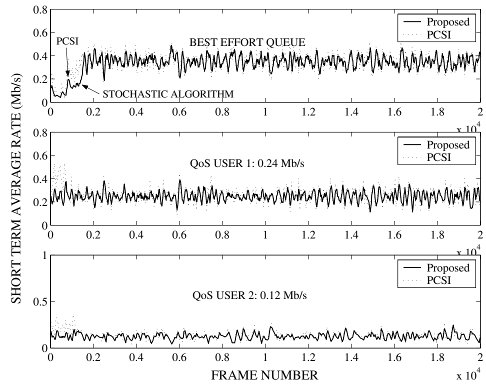

Fig. 2. Throughput comparison of DPRI and PCSI in stationary channel (averaged over consecutive 50 frames); frame duration = 1 TTI (0.667ms).

that to select the rate that maximizes the throughput given by $D_i(n) = R_j(1 - P_{ij}^e(n))$ over all $j = 1, 2, \cdots, r$ for each user $0 \leq i \leq r$. Shown in Fig. 1 are the evolution of user selection probabilities when there are two QoS users each with an average rate requirements of 0.24 and 0.12 Mb/s and a best effort queue. The best effort user is assumed to have a fixed SINR of 15dB and the QoS users 12 and 9 dB. Fig. 2 shows the convergence of the average rates of QoS users to the required average rates. Note that there are short term fluctuations due to frame error probabilities resulting from AWGN.

Figs. 3-6 show the results for time varying channel. The number of users and their rate requirements are set to the same values as in the case of stationary channel. The average SINRs of the users were set to 15, 12, and 9dB. The time variations in the channels were generated using Jakes' model [10] with 13 taps. Fig. 3 shows the signal envelops. Fig. 4 gives the evolution of user selection probabilities. When the SINR of a user increases/decreases, the probability of selection adaptively decreases/increases to maintain the average transmission rate to the requested value. Note that as in Fig. 5, the short term average transmission rates achieves the requested rates within first few iterations and remains at the requested values. There is no effort by the algorithm to regulate the transmission rate of the best effort queue. Shown in Fig. 6 is the evolution of rate assignment probability vector, $p_1^R(n)$ of QoS user 1. The correspondence to the variation in channel SINR is observable.

V. Conclusions

A two-level hierarchy of stochastic learning algorithm for multi user wireless channel was formulated and studied. The algorithm adaptively increases the frequency of selecting the modulation and coding schemes that maximize the throughput for the fading wireless channel and schedules user to be transmitted in each time slot so that to achieve the requested transmission rates. The algorithm is based on stochastic learning automata that maintain vectors of probabilities for random selections of rates (determined by a finite set of modulation and coding schemes) for each user and to select the users to be transmitted in each time slot. These vectors of probabilities

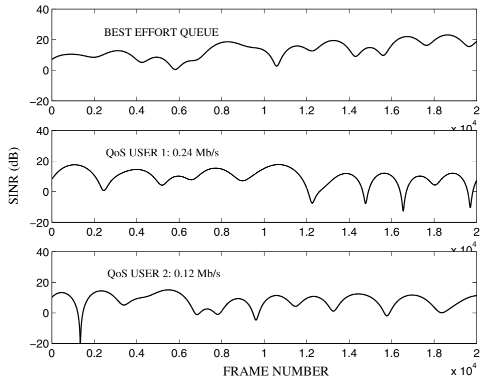

Fig. 3. SINR variation against time for the set of users at a speed of 0.02km/h (based on Jakes' model); frame duration = 1 TTI (0.667ms).

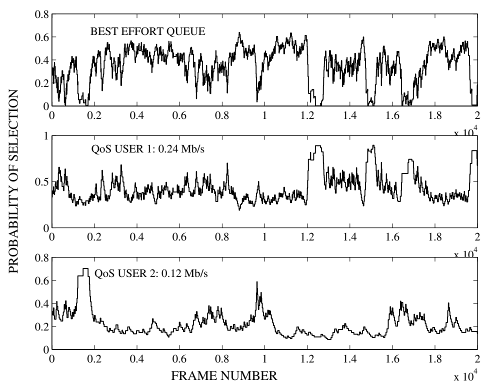

Fig. 4. Time evolution of probabilities in dynmaic channel;frame duration = 1 TTI (0.667ms).

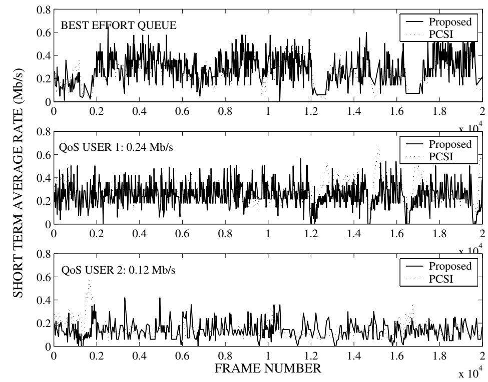

Fig. 5. Throughput comparison of DPRI and PCSI in time varying channel; frame duration = 1 TTI (0.667ms).

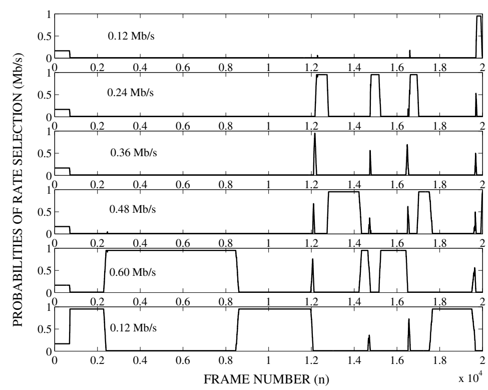

Fig. 6. Time evolution of rate selection probabilities in time varying channel for QoS user 1; frame duration = 1 TTI (0.667ms).

are updated based on the parameters derived using the ACK signals feedback from receivers following the transmission of data packets. This algorithm eliminates explicit channel state estimation and feed back. Theorems were established to show that when the channels are stationary i.e., when the channel signal power to noise (additive) power ratios remain constant, the algorithm is guaranteed to convergence to the optimal modulation and coding scheme of each user and achieves average transmission rates as requested by the users. It is also shown that when the channel variations are sufficiently low, the algorithm can adaptively change the rate selection probabilities and user selection probabilities to converge to the new optimal solution as the channels vary with time. Simulation results using the typical parameters of third generation wireless systems and wireless lans were presented to illustrated the performance.

ACKNOWLEDGMENTS

This work is supported in part by the grants NSF ITR 0082064 and NSF ANI CAREER 0133761.

REFERENCES

[1] K. Balachandran, S. R. Kadaba, and S. Nanda, *Channel Quality Estimation and Rate Adaptation for Cellular Mobile Radio*, *IEEE JSAC*, vol. 17, pp. 1244-1256, July 1999.

[2] 3GPP,"Physical layer aspects of UTRA High Speed Down-link Packet Access (Release 4)," TR25.848 v4.00 (2001-03), March 2001.

[3] TSG-RAN WG1, "Alternative to 8 PSK Modulation/Coding Scheme for HSDPA," TSGR1#20(01)0471 Texas Instruments, May 2001.

[4] K. S. Narendra and M. A. L. Thathachar, *Learning Automata: An introduction*, Englwood Cliffs, NJ:Prentice-Hall, 1989.

[5] M. A. L. Thathachar, and P. S. Sastry, "Estimator algorithms for learning automata," in *Proc. Platinum Jubilee Conf. Syst. Signal Process.*, Bangalore, India, Dec. 1986, pp. 29-32.

[6] M. A. L. Thathachar, and P. S. Sastry, 'A class of rapidly converging algorithms for learning automata," in *Proc. IEEE Int. Conf. Cybern. Soc.*, Bombay, India, Jan. 1984, pp. 602-606.

[7] M. A. L. Thathachar, and K. M. Ramachandran, "Asymptotic behaviour of a learning algorithm," *Int. J. Control*, vol. 39, pp827-838, 1984.

[8] M. .L. Tsetlin, "On the behavior of finite automata in random media," *Automat. Telemek.*, vol. 22, pp. 1345-1354, Oct. 1961.

[9] B. J. Oommen and J. K. Lanctôt, "Discretized pursuit learning automata," *IEEE Trans. Syst., Man, Cyebern.*, vol. 20, pp. 931-938, July/Aug. 1990.

[10] W. C. Jakes, *Microwave Mobile Communications*. New York: IEEE, 1974.

[11] M. Sharif and B. Hassibi, "On the Capacity of MIMO Broadcast Channel with Partial Side Information," *Submitted to IEEE Trans. Info. Theory*.

[12] P. Viswanath, D. N. C. Tse, R. Laroia, "Opportunistic beamforming using dumb antennas," *Proc. of IEEE ISIT 2002*.

[13] J. Hall and P. Mars, "Satisfying QoS with a Learning Based Scheduling Algorithm," *Proc. of IEEE IWQoS '98*.

[14] H. Zheng, "Optimizing wireless multimedia transmissions through cross layer design," *Proc. of IEEE ICME 2003*.