Mobility Independent and Dependent Predictive Services Management in Wireless/Mobile Multimedia Network

F. De Rango[1], P. Fazio[2], S. Marano[1]
D.E.I.S. Dept, University of Calabria,
Rende (CS), 87036, Italy
[1]{derango, marano}@deis.unical.it
[2]pepfaz77@tiscali.it

Abstract— **In this paper the attention is focused on the management of services independent and dependent from mobility in the adaptive multimedia networking, where the bandwidth of an ongoing multimedia flow can be dynamically adjusted. Two classes of services have been considered: MIP and MDP classes defined in Integrated Services for mobile wireless environment. In order to offer an adaptive QoS (soft QoS) increasing the total wireless system utilization, a rate adaptation algorithm has been considered. The valued algorithm is based on user utility function and its target is to maximize this function. The mobility of host can impact on number of handoff and on channel states (fading), so the rate adaptation accounts the degradation channel state conditions. An algorithm of admission control has been considered too in order to regulate the access to wireless multimedia network. The admission control can use the pre-reservation phase among potentially visited cells from mobile hosts for MIP services and consider only the bandwidth availability on current cells for MDP services. The performance evaluations of the wireless system have been evaluated in terms of total bandwidth utilization for MIP and MDP services, average user utility perceived by mobile users and system outage probability.**

Keywords- *Admission control, adaptive bandwidth allocation, utility function, Markov model, fading channels, pre-reservation.*

I. INTRODUCTION

Recently there is a growing interest in the adaptive multimedia networking, where the bandwidth of an ongoing multimedia flow can be dynamically adjusted. Wireless communications pose special problems, such as limited bandwidth and high error rate, that do not exist in wired networks; in addition, the dynamics of physical channels cause variations in received signals; for these reasons, deterministic service guarantees and bandwidth allocation, commonly used in wired networks, become inadequate for dealing with wireless communications. User mobility has an important impact on QoS parameters of real-time applications and the existing protocols, as proposed in [1], must be extended in order to manage host's mobility. This can be handled by using the MRSVP protocol, which is based on active and passive reservations and capable to pre-reserve an amount of bandwidth for MIP flows to guarantee the desired QoS during hand-off events, while serving MDP requests; in this way, the effects of mobility for MIP connections are minimized. Moreover, different users may experience different link capacities due to different locations, and we believe that bandwidth should be allocated in an adaptive and link-state dependent way. To consider the heterogeneity of different applications and to have a consistent performance measure, we adopt utility functions in our proposed adaptive QoS model. In this paper, we propose an utility-oriented bandwidth allocation scheme and admission control policy, which account for the users' QoS requirements and actively adapts to the dynamics of the physical channel. There has been much work on wireless resource management, focusing on multiple access and channel allocation, however there is less research on adding explicit adaptive mechanisms to bandwidth allocation schemes to deal with the variations of wireless channels. The rest of the paper is organized as follows. Section II discusses the user mobility effect, the wireless variations and the different kinds of wireless service class. In Section III the utility-oriented rate adaptation algorithm is discussed. Section IV describes the simulation results and Section V concludes the paper.

II. USER CLASSES IN ISPNS AND MOBILITY EFFECTS

A. Mobility independent and dependent service classes

Internet best-effort service does not offer any guarantee about available bandwidth, network propagation delays, jitter and packet delivery. As consequence, there have been different reserch groups that tried to define some service model, in order to deal with applications variety in packet networks. Integrated Services networks are the results of such kind of works, as described in [2,3,4,5].

In a real network, resources reservations can be made by protocols, in order to satisfy QoS requirements and to offer to mobile hosts a service "better than best-effort", accounting the inherent time varying environmental conditions evident in radio communications (e.g. fading). In IS networks, each flow can receive different QoS, which must be negotiated at the beginning of sessions, between flows and net, by the RSVP protocol [6] or the MRSVP protocol, in mobile scenarios [7]. There are three provided service classes [8]: Mobility Independent Guaranteed (MIG, for hard and intollerant applications, that need absolute guarantees on packet delays), Mobility Independent Predictive (MIP, for tollerant real-time applications, that can suffer limited bounds on packet delays) and Mobility Dependent Predictive (MDP, for applications that can suffer continuous QoS degradations or connection droppings). In this paper, only MDP and MIP classes have been considered. MRSVP protocol is used for exchanging state information of wireless networks and it can offer soft QoS (adaptive QoS) for MIP and MDP services; so they have two different management in terms of admission control and bandwidth assignments. MIP services use a pre-reservation phase to reserve bandwidth for mobile host in the current cells and in the cells that mobile host probably will visit (passive and active reservations) and MDP services, instead, can reserve the bandwidth only on the current cell (figure 1). According to adaptive multimedia wireless framework [9], MIP and MDP can reserve a bandwidth level that can change during call holding time. This behaviour can guarantee a more flexible resource management, increasing the system utilization.

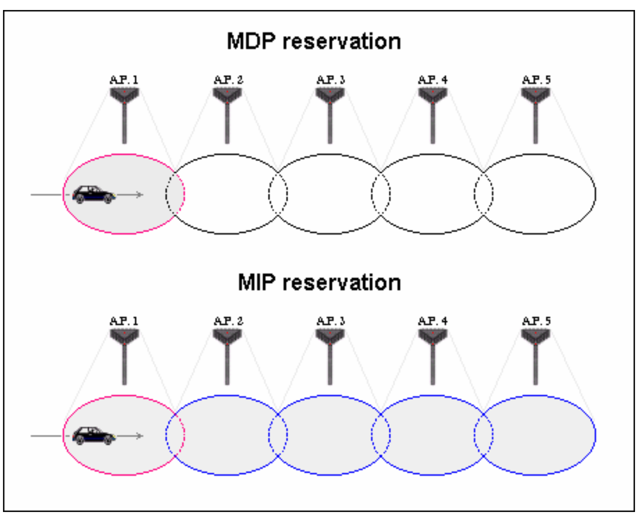

Fig. 1. Different reservation policies for ISPN classes

B. *Mobility and fading effects, radio link model*

Users' mobility causes channel quality and service degradations during a connection: we considered the multipath fading phenomenon, as direct consequence of user mobility and time-varying impulse response of radio links; in particular our work takes in account the *slow-fading* effect, which takes place when the evolution of link changes can be considered to be slower than symbol transmission time. In this way, we give more effectiveness to used rate adaptation algorithm and reservation protocol, because the offered network service always guarantees the minimum delay requests made by users, by allocating more resources to face transmission errors due to absence of ideality in radio channels. Mobility and fading can change system conditions, so it is necessary to use rate adaptations algorithms for varying bandwidth levels, according to channel conditions and host mobility.

In our work we employed a k-states Markov chain model to describe the behavior of a radio-link between users and access-points, as proposed in [10]. We needed to introduce the chain model to consider the absence of ideality in wireless communications and the fluctuations in the received signal level, due to the various propagation phenomenas during a generic connection (shadowing, refraction, fading, etc.). The model can only be used under the assumption of *slow fading: if* f_d is the maximum Doppler shift introduced by the user mobility and T the symbol transmission time, then we can say that the slow-fading condition is verified if $f_d*T<<1$.

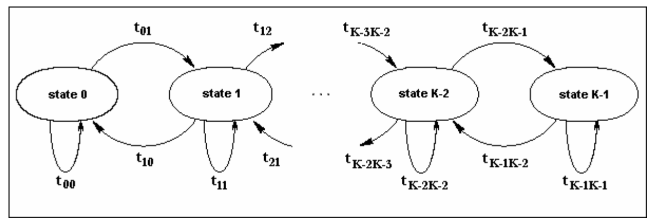

Fig. 2. A k-states Markov chain model for radio links

III. UTILITY-ORIENTED ALGORITHM

We used an *utility-oriented* algorithm for rate-adaptation and admission control, considering the time-varying nature of links, between hosts and access points [11]. In our case, we used monotonically non-decreasing utility functions, describing how the perceived utility changes with the amount of effective bandwidth received by the user. In our service model, each user i can signal its utility function $U_i(r)$ to the network, where r is the amount of effective bandwidth received by the user and the bandwidth allocated to a flow can take its discrete value from the set $B=\{l_1,l_2,\ldots,l_n\}$, where $l_i<l_{i+1}$ for $i=1, .., n-1$. It is assumed that the calls can belong to MIP or MDP class and all of them take (varying) bandwidth values from the same set B.

The network tries to dynamically allocate bandwidth such that each user's instant utility is maintained above the minimum level and, in the long run, the bandwidth is allocated fairly and utilized efficiently. As described earlier, the communication link of each user can be modelled by a k-state Markov chain: we can indicate the average state-holding time of each state m with t_m and the bandwidth degradation ratio of the state with D_m, where $0\leq D_m<1, \ \forall 1\leq m\leq k$. If, at a particular time instance, r_i is the amount of bandwidth that the network is allocating to user i, we define the *received instant utility* as $u_i = U_i((1-D_{i,m})*r_i)$. One of the objectives of the bandwidth allocation scheme is to guarantee the minimum utility level for each user i; if we define *utility outage* as the event that user i's instant utility level falls below the minimum, the scheme should guarantee that the probability of an utility outage is smaller than a certain threshold p_{outage}. In addition, the fairness criterion should also be based on utility: considering users i and j with average utility $u_{i,avg}$ and $u_{j,avg}$ respectively, we can define the *normalized gap* of the average utility received and the minimum level $u_{*,min}$ as: $G_i=(u_{i,avg}-u_{i,min})/u_{i,min}$, so we want all users to have the same normalized gap in the long run ($G_i \approx G_j$, $\forall i,j$).

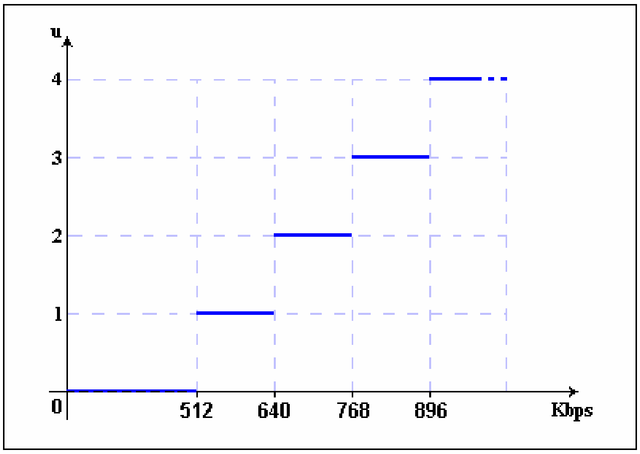

Fig. 3. An example of utility function

When a user's link *degrades*, it may surrender some bandwidth to another user, with a smaller normalized gap; when a link *upgrades* the user may receive some bandwidth from another one, with a larger normalized gap; in this way there is a net gain in the combined instant utility. If at a particular time user *i*'s link state changes to state *p*, the following steps are performed: 1) all users' average utility level and normalized gap are updated; 2) users are sorted in increasing order of normalized gap; 3) if the instant utility level of user *i* is below the minimum, some users' bandwidth will be reduced and reallocated to user *i* to meet its $u_{i,min}$; 4) if there is no step three, user *i* may give up part of its bandwidth to another user if the link degrades, whereas it may receive some bandwidth if the link upgrades. We call the user who gives up part of its bandwidth to others the *benefactor*, and the user who receives bandwidth from others the *beneficiary*. In the third step, to satisfy user i's $u_{i,min}$, the scheme searches for *benefactor(s)* starting from the user with the largest normalized gap. Suppose the user with the largest normalized gap is user j, whose link is currently in state q and it is above $u_{j,min}$. User j will yield

$$\min\left(\frac{r_{i,\min}}{1-D_{i,p}}-r_i, r_j-\frac{r_{j,\min}}{1-D_{j,q}}\right)$$

amount of bandwidth to user *i*, where r_i and r_j are the bandwidth allocated to users *i* and *j*, respectively, before the link state transition. This procedure will be repeated until $u_{i,min}$ is reached or all the users have been checked.

When user i's link degrades, the scheme will search for an appropriate *beneficiary*, checking the users in increasing order of normalized gap; when the *beneficiary* is found, the scheme decides the amount of bandwidth to transfer between the users, trying to maximize the combined utility of them. This procedure is repeated until one *beneficiary* is found or all users with smaller normalized gap than user *i*'s have been checked. Similarly, when user *i*'s link upgrades, user *i* becomes the beneficiary and users with larger normalized gap are the candidates for *benefactor*. The scheme checks the candidates in decreasing order of normalized gap and, when the *benefactor* is found, the scheme decides the amount of bandwidth to exchange, maximizing the combined utility of the two users. Besides the link state changes, adjustments in bandwidth allocation are also needed when an user arrives (new user) or departs. If *r* is the amount of bandwidth which needs to be collected from current users because of an user arrival, user *j* with largest normalized gap G_j is to give up

$$\min\left(\max\left(0, r_j-\frac{r_{j,\min}}{1-D_{j,q}}\right), r\right)$$

amount of bandwidth, where q is the current link state of user j. This procedure will be repeated until enough bandwidth has been collected or all the current users have been serched. If after searching all the current users, the collected bandwidth is still not enough, the scheme will start a second round of collection, again starting from the user with the largest normalized gap, but, this time, each chosen user will be dropped out from the network. Similarly, if there is surplus bandwidth, the users with the first k smallest normalized gap are chosen to receive the surplus bandwidth. Each user can increase its effective bandwidth up to the maximum effective bandwidth level.

To guarantee users' minimum utility level, an admission control policy should be enforced to limit the number of users in the system. Recalling that when an user's instant utility falls below its minimum utility level there is an utility *outage* for the user, the probability p_0 of such event at any time is:

$$p_0 = P_r\left\{\sum_{i=1}^{n}\frac{r_{i,\min}}{1-D_{i,m_i}} > R\right\},$$

where m_i is user *i*'s link state at the time istance, p_{mi} is the probability of the user *i*'s link to being in state m at a particular time, n is the total number of users including the new one and *R* represents the bandwidth associated to the wireless cell *c*. Modelling the wireless channel through the Finite State Markov chain (FSMC) where the state of the chain represents the channel degradation, the transition probabilities represent the probability to change degradation state,as shown in figure 2, it is possible to know the probability of the link to being in the state *i* :

$$p_i = \frac{\pi_i t_i}{\sum_{i=1}^{k}\pi_i t_i},$$

with π_i representing the steady state probability associated to channel state *i*, and t_i representing the *i*'s average holding time. Through p_i it is possible calculate in the worst case p_0 accounting the channel state conditions in the following way

$$p_0 = \sum_{A} \prod_{1 \le i \le n} p_{m_i} \,,$$

where

$$A = \{m_1, m_2, ..., m_n \mid 1 \le m_1, ..., m_n \le k, \sum \frac{r_{l,1}}{1 - D_{l,m_l}} > R\} \,,$$

with m_i representing the user i's link state at the time instance. The admission control algorithm works differently for two classes of service (MIP, MDP). For MIP class, the flow is admitted if:

$$\sum_{c=1}^{C} p_{0,c} \le C \cdot p_{outage} \,,$$

where C is the number of cells that mobile host will visit and p_{outage} is the outage probability of the wireless system (for MDP class, the condition must be verified only for the current cell). So, when a new user arrives, the scheme calculates $p_{0,c}$ as described and if $p_{0,c} \le p_{outage}$ for each cell, the new user is admitted, otherwise it is rejected. If a user *j* is admitted, it is initially allocated $r_{j,min}/(1-D_{j,q})$, where q is user *j*'s current link state. The assigned amount of bandwidth to *j* is contributed by the users currently in the network following the algorithm we described previously. The MIP services can under-utilize the system resource for passive reservation, so it is permitted to MDP flows to use the *passive bandwidth* of MIP flows. When a flow needs of its passive bandwidth (e.g. for handoff), the passive bandwidth is switched in active reservation and the MDP is pre-empted. For avoiding the increasing of MIP flows cut, rate adaptation is used for MDP. So the available bandwidth obtained by rate adaptation algorithm is assigned to an MDP flow, that used passive bandwidth. If there is no availability of bandwidth, the MDP flow is cut.

IV. SIMULATIONS

In order to evaluate the performance of rate adaptation scheme, call admission control for MIP and MDP and the conformance to QoS parameters (outage probability, minimum utility function and an high system utilization), different simulations have been lead out. Our simulated net consists of 5 wireless cell, each one covered by an access point (as illustrated in figure 4) with bandwidth capacity of 5.5Mbps; the access points are wired connected, by a switching subnet, to the net-sender.

	p_i	e_i	D_i (%)	t_{mi} (s)
0	0.30233	0.150588	41.8	0.1
1	0.095	0.0670038	24.58	0.1
2	0.25267	0.016459	9.5	0.1
3	0.35	5.92112e-5	0	0.1

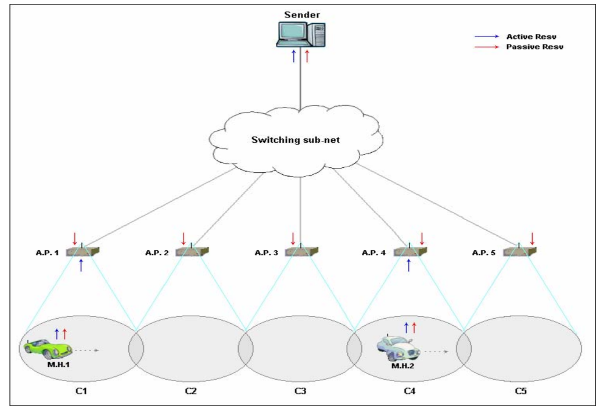

Fig. 4. The simulated net

The set of possible discrete bandwidth levels (Kbps) is: B = {512, 640, 768, 896} and the utility function is the same of figure 4, so the set of instant utility values is: U = {1,2,3,4} for all users. In our simulations the traffic load is composed by MIP and MDP flows in variable percentage. The bandwidth is managed by the policies illustrated in Section III and the outage threshold is the same for both flow classes (conjunctive bandwidth management). The mobile host can move with average speed selected uniformly in range [5,75] Km/h. The following curves illustrates the performances of the utility-oriented algorithm for different values of outage threshold and mobile host speed.

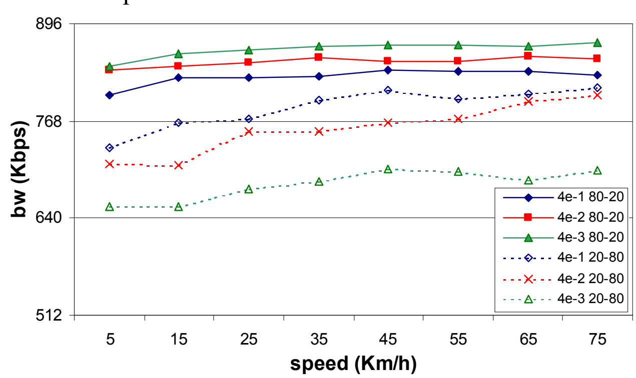

Fig. 5. Average allocated bandwidth for MIP users

In figure 5 it can be observed the improvement in resource allocation for MIP users, by increasing their traffic percentages from 20% to 80%; for higher MIP traffic, more users, belonging to this class, can enter the system, pre-empting MDP flows and, in the case of conjunctive resource management, degrading MDP reservations more frequently; in addition, for high MIP traffic percentages, increasing outage threshold, a decreasing in allocated bandwidth can be observed: in this situation, there are more MIP users sharing the same cells capacities, so they must perceive a lower amount of bandwidth.

Figure 6 shows the average system utilization: the system is lower utilized by increasing hosts' speed. For high speed values there are more link variations and, consequently, system must handle a larger number of bandwidth reallocation; this causes an utilization wastage, which can reach a magnitude of 35%-40%. Increasing MIP traffic percentage (from 20% to

80%) a system under-utilization is visible: for high values of MIP traffic percentages, there are a lot of passive pre-reservations and, even though MDP flows can utilize unused MIP bandwidth, there is a low resource utilization, more evident for high speed, because there are fewer MDP users in the system and the bandwidth wastage, caused by passive reservations, cannot be faced. Varying outage threshold, there are different observed values of resources utilization, for the same reason early discussed, that is to say the admission control is less selective for higher threshold values, so more users can enter the network and a higher utilization can be reached.

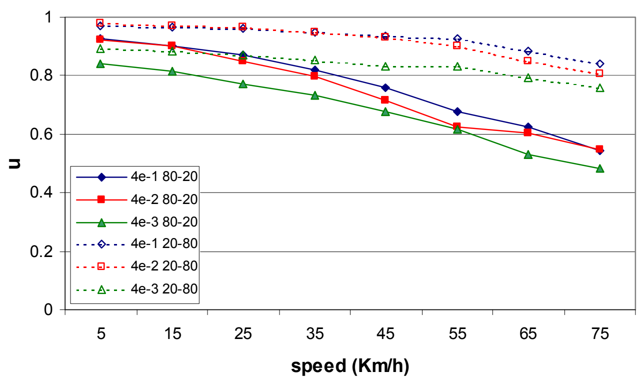

Fig. 6. Average system utilization

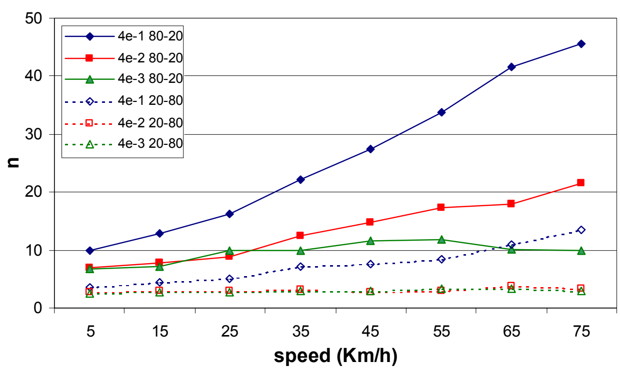

Fig. 7. Average MIP admitted flows

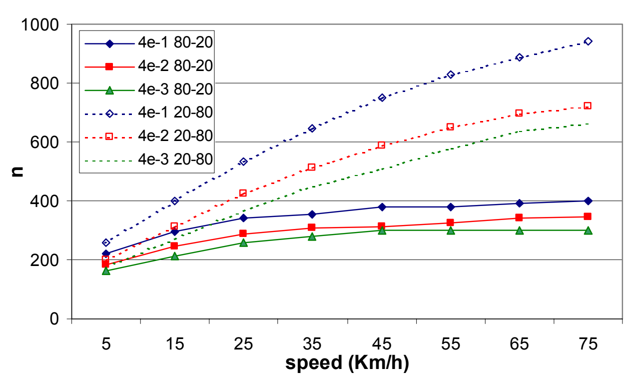

Fig. 8. Average MDP admitted flows

From figures 7 and 8 it can be observed that the maximum values of admitted flows in MIP case do not exceed 15 and 50 (for 20% and 80% percentages), while for MDP case the maximums are near 950 and 400 (for 80% and 20% percentages): as illustrated in figure 1, MDP users make requests only to current cells, while MIP users make reservations over all system cells, so they are subject to a more strictly admission policy and the probability of a system enter is lower than MDP..

Through the simulations it has been verified also that high values of outage threshold (like 0.4) cannot be suffered by MIP flows, cause there are too much dropped flows, while for low values (like 0.004) the phenomenon can be disregarded.

V. CONLCUSIONS

In this paper we analyzed the performances of a proposed utility-oriented algorithm and we also seen that there is need to consider the physical link variability, in order to give more effectiveness to bandwidth allocation schemes and to take in account the absence of ideality during radio connections. In our work two classes of service are provided (MIP and MDP) and the analysis of mobility effects has been made for both classes. From simulation results we observed that the channel degradation becomes more evident for high speed values and the hand-off number increases, but the pre-reservation of MIP classes guarantees a full compliance to QoS parameters (outage probability and maximized user utility function) for low values of outage threshold, so the mobility effects are minimized. High percentages of MIP traffic lead to a system under-utilization, although the MDP users can use passive resources; in addition high values of outage threshold cause a violation of QoS requirements for MIP users.

REFERENCES

[1] R. Braden, L. Zhang, "Resource ReSerVation Protocol (RSVP) - Version 1 Message Processing Rules", September '97.

[2] IETF RFC 1633 – "Integrated Services in the Internet Architecture: an Overview".

[3] J. Wroclawski, "Specification of the Controlled-Load Network Element Service", September '97.

[4] S. Shenker, C. Partridge, R. Guerin, "Specification of Guaranteed Quality of Service", September '97.

[5] S. Shenker, J. Wroclawski, "General Characterization Parameters for Integrated Service Network Elements", September '97.

[6] J. Wroclawski, "The Use of RSVP with IETF Integrated Services", September '97.

[7] Talukdar A. K., Badrinath B.R. and Acharya A., "MRSVP: A Reservation Protocol for an Integrated Services Packet Network with Mobile Hosts".

[8] Talukdar A. K., Badrinath B.R. and Acharya A., "On Accomodating Mobile Hosts in an Integrated Services Packet Network", In the Proceedings of the IEEE INFOCOM '97, April.

[9] V.Bharghavan, K-W.Lee, S.Lu, S.Ha, J-R.Li, D.Dwyer, "The TIMELY Adaptive Resource Management Architecture," *IEEE Personal Communications Mag.*, pp.20-31, Aug.1998.

[10] H.S.Wang, N.Moayeri, "Finite-State Markov Channel-A Useful Model for Radio Communication Channels," IEEE Trans. On Vehicular Technology, vol.44, no.1, pp.163-171, Feb.1995.

[11] Y.Cao, Victor O.K.Li, "Utility-oriented Adaptive QoS and Bandwidth Allocation in Wireless Networks," IEEE Int. Conf. On Communications (ICC2002).