Resource Management in IEEE 802.11 Multiple Access Networks with Price-based Service Provisioning

Leonardo Badia, *Member, IEEE*, Simone Merlin, *Student Member, IEEE*, Michele Zorzi, *Fellow, IEEE*

***Abstract*—In this paper we analyze the provisioning of multimedia services over a wireless LAN hot-spot, based on the IEEE 802.11 protocol. We address the issue of defining proper pricing strategies, from the perspective of both evaluating the technical performance and quantifying the economic revenues. We take into account a model for users' behavior that describes all users' choices in a decentralized manner, so that the transmission rate of each node is driven both by multimedia service requirements and by the customer's willingness to pay. The multiple users' medium access mechanism is studied through a simulation analysis based on *ns-2*. Within this model, the network performance is evaluated and discussed, presenting numerical results which can provide practical insight for pricing setup in a wireless LAN hot-spot. We observe that the impact of the pricing policy on the provider's income and on the satisfaction of the users is critical and especially depends on the shape of the pricing function (flat, linear or hybrid). Additionally, we investigate the provider's task of having a suitable price policy which properly tunes the trade-off between the two contrasting factors of achieving high revenue and obtaining high satisfaction of the users.**

***Index Terms*—Resource management and QoS provisioning, Multiple access techniques, Utility functions, Pricing, IEEE 802.11.**

I. Introduction

The diffusion of Wireless Local Area Networks (WLANs) based on the IEEE 802.11 standard [1] is rapidly increasing. At first limited to laptops, it now involves palmtops as well as other kinds of portable devices. Hot-spots, where a set of mobile terminals is connected to a central access point, are emerging as a widespread application of this standard. Nowadays, these kinds of systems are present in business areas such as conference rooms or airport and hotel lounges, where users are interested in easily and rapidly establishing a network connection. Following such a wide diffusion of WLAN devices and coverage availability, also the offered services are going to comprise a broader set of applications, including audio, video and multimedia services.

Current implementations of IEEE 802.11 systems adopt the Distributed Coordination Function (DCF) using Carrier-Sense Multiple Access with Collision Avoidance (CSMA/CA). In particular, we focus on IEEE 802.11b, though our conclusions can easily be extended to any other version of the IEEE 802.11 standard. It is well known [2] that the performance of CSMA/CA-based multiple access networks is heavily affected by the network operating conditions. Thus, the provider is interested in efficiently managing the bandwidth resource. Reasonably, this could mean aiming at achieving a satisfactory income from the network management operation, while providing as many users as possible with a satisfactory service [3]. These two requirements are likely mandatory in order to have a sustainable economic model. For this reason, the investigations on how to properly allocate the radio resource, as well as to set up an appropriate pricing strategy, are key issues for the network operator [4].

To explore these aspects, we refer to the application of economic models to Radio Resource Management, an open field of research on which several contributions have appeared in the recent literature [5]–[7]. In particular, the concept of utility functions and issues taken from game-theory have been employed to represent a tunable Quality of Service (QoS), e.g., obtained through variations of the terminal's data rate [8], [9].

An example of application of micro-economic issues to the management of a WLAN hot-spot is given in [10]. However, note that the micro-economic control performed there refers to the definition of a virtual price that has the effect of regulating the access and is negotiated dynamically [11]. Instead, in the present contribution we are interested in considering more directly the real price established by the operator for the service tariff. This significantly distinguishes our work from other related approaches. Also, such a price is bound to be fixed *a priori* and known in advance by the users.

To quantify these economic concepts and perform numerical evaluations, we adopt the micro-economic model for multiple access wireless networks presented in [12], where users' choices are described as driven by their appreciation of the service, which is in turn influenced by the price paid. In fact, some users may refuse to enter the system due to what they consider to be too high a price. On the other hand, pricing the system usage also allows a better coordination and a more efficient utilization of resources. Finally, dynamically changing network conditions may cause dissatisfaction in the users that do not achieve what they consider an adequate quality of service, according to the price initially negotiated.

For this reason, we evaluate the ratio of *satisfied users*

Manuscript received June 14, 2007; revised September 11, 2007; accepted October 27, 2007. The associate editor coordinating the review of this paper and approving it for publication was M. Guizani. Part of this work has been presented at IWCMC 2006.

L. Badia is with IMT Lucca Institute for Advanced Studies, piazza S. Ponziano 6, 55100 Lucca (email: lbadia@ing.unife.it).

S. Merlin, and M. Zorzi are with the Department of Information Engineering (DEI), University of Padova, via Gradenigo 6/B, 35131 Padova, Italy (email: merlo@dei.unipd.it, zorzi@ing.unife.it).

Digital Object Identifier 10.1109/T-WC.2008.070598

as another indicator of good management that a provider of a real system needs to take into consideration in the long run. Moreover, as will be discussed in the following, certain economic metrics, such as the provider's revenue, are meaningful if they are evaluated on satisfied users only. In addition, this model can even be used as an instrument to identify what a suitable pricing policy for a WLAN hot-spot could be. Note that the investigation of these issues requires to characterize IEEE 802.11 multiple access within the network; to this end, we make use of the *ns-2* simulator [13] which allows a direct and simple integration of these issues.

The rest of this paper is organized as follows. After a short summary of the micro-economic model that will be used to evaluate the behavior of the users of the WLAN hot-spot, presented in Section II, we describe the case study implemented by means of an extended *ns-2* simulator in Section III. In Section IV we show a possible instantiation of the model for what concerns utility and pricing functions and we show in Section V the numerical results of our extensive simulation campaign. Finally, we conclude in Section VI.

II. The Model for Users' Behavior

The IEEE 802.11 infrastructure-based implementation of a WLAN obtained through DCF realizes a centralized hot-spot, where mobile terminals can connect to an Access Point (AP). It is envisioned that the increasing diffusion of such structures will allow users to access the Internet through different kinds of terminals and enjoy a plethora of services, e.g., data download, web-browsing, voice and video call, access to multimedia content and so on.

In such a scenario, characterizing the overall network behavior is very challenging. WLAN terminals have heterogeneous features and the requirements of the demanded services may be extremely variable. Moreover, the issue of QoS provisioning is particularly complicated for protocols which, like IEEE 802.11, are intrinsically best effort (at least in their original concept), i.e., there is no guarantee about the achieved QoS.

Thus, under the perspective of QoS provision, resource management techniques are difficult to investigate, since users' appreciation of the service is often hard to represent with analytical tools. Therefore, an approach commonly followed in the recent technical literature is to employ a utility-based description of users' preferences [5], [8].

The key assumption of this methodology is the availability of *utility functions*, mapping the subjective preferences of the users into numerical values. The absolute values of the utilities can even be arbitrary, as they do not need to have a meaning *per se*, but they should respect certain order relationships, so as to reflect that higher utility values are given to choices which are more preferable for the user.

Since we want to include also pricing in our analysis [4], in this work we specialize these assumptions by describing the users' behavior as driven by two factors: the quality of the service itself, which is assumed to be estimated from a quantitative point of view via subjective testing and is therefore represented by means of a utility function $u(r)$, and the price paid for accessing the service, described by a pricing function $p(r)$. Both of them are non decreasing functions of the allocated resource r. In the WLAN analysis performed in this paper, we identify r with the *achieved data rate*. For the sake of simplicity, in this paper the price paid by a user is determined only depending on this quantity, through a one-shot application of the function $p(r)$. However, the reasonings presented in the following could be promptly extended in order to account for other aspects, such as call priority, duration of the connection and so on, with a multi-dimensional analysis, where r is replaced by a n-tuple of input variables.

The service perception is determined by the trade-off between these two parameters, since for every user, qualitatively speaking, the larger the utility and/or the lower the price, the higher the satisfaction. According to the framework described in [12], we represent this with a *service satisfaction function* $A_i(u_i, p_i)$ for every user i belonging to the potential users set $\mathcal{Q}$, where u_i and p_i are user i's utility and price paid. Since both utility and price ultimately depend on the rate r, we will often use a slight abuse of notation, writing $A_i(r_i)$ for short. It is further assumed that the satisfaction function takes values between 0 and 1, so that we can regard it as the probability of the ith user being satisfied.

In the following, we will take $p_i = p(r_i)$ and $A_i = A(r_i)$, because it is reasonable to assume that these functions are homogeneous throughout the whole network (the extension to the case where multiple pricing or QoS classes are present is straightforward). Instead, we assume a different utility function for every user so as to account for the variability of services and terminals. Being a subjective metric, the utility heavily depends on factors which can not be controlled by the resource manager, such as the terminal performance or the users' subjective evaluation of the service quality. Hence, the utility $u_i(\cdot)$ is in general a different function for every user.

We assume that the general objective of the network manager is to have high revenue while at the same time achieving satisfaction of the users. Thus, we evaluate the revenue *on satisfied users only*. The motivation for this is as follows: from an economic point of view, dissatisfied users are expected to abandon the service in the long run and henceforth they can be considered as lost customers. For this reason, generating revenue without satisfying the users appears to be pointless.

When dealing with WLANs based on IEEE 802.11b, the complexity of modeling the MAC protocol makes it difficult to deal with analytical formulations of the capacity. Even though interesting fully closed-form models have been presented in the literature [2], [14], [15], for the sake of simplicity our analysis is numerically evaluated through simulation obtained with *ns-2* [13]. This approach offers in fact good scalability and ease of implementation, and involves only the multiple access of the users to the channel, while the rest of the investigation is founded on analytical reasoning as discussed in the following. Nevertheless, entirely analytical approaches can be envisioned as an interesting evolution of the present paper in future research.

In an IEEE 802.11b scenario, we need a differentiation mechanism to prioritize and coordinate multiple users' requirements in the WLAN environment, which possibly imply a different rate r_i for every user. The value of r_i will then be mapped through the utility and pricing functions to finally determine the probability of accepting the service, $A(r_i)$.

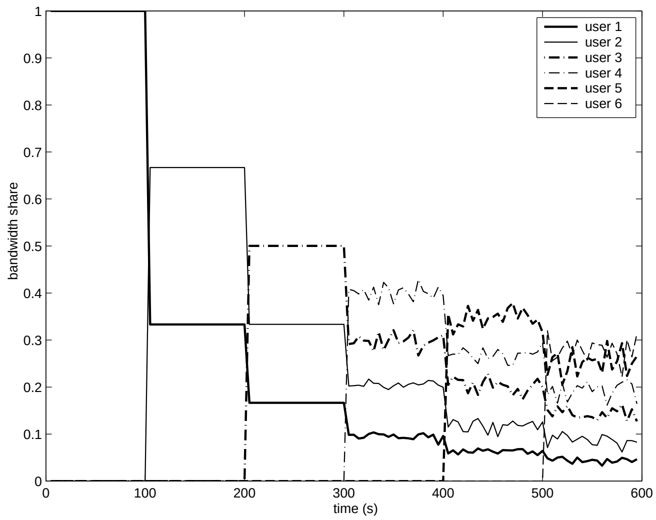


Fig. 1. Proportional share of resource for a WLAN system with six users with increasing priority. This confirms and extends to a larger system the results already obtained by [16].

For investigation purposes only, we suppose that all users generate packets at the same rate, but each user tunes the packet length in order to achieve its requested bit-rate. This mechanism, together with other possibilities, has been proposed in [16]. We remark that other differentiation techniques might be used as well within the same rationale. The choice of this particular strategy is motivated by the fact that, under saturation conditions, the long range average traffic enjoyed by each user is proportional to its packet length. In particular, this mechanism allows for the ratio between the offered traffic of any two users to be the same both in the non saturated and in the saturated case.

To confirm this, Fig. 1 reproduces the results shown in [16], extended to a wider range of number of users. We consider the subsequent allocation of up to 6 users in the WLAN scenario, so that a new user is allocated every 100 seconds. The resource requirements of each user are subsequently increasing, so that the rate requested by the ith user, $i > 1$, is i times that requested by the first one. As shown in the figure, the correctness of the assumption of proportionally fair share of resources holds. However, as the number of users increases the instantaneous variations around the long range average value become more evident.

As shown in Fig. 1, the rate achieved by the users is not constant over time, and in particular varies according to the presence of other users in the network. For this reason, the rate r_i allocated to user i should be regarded as variable over time, i.e., $r_i(t)$. For the purpose of a practical evaluation, and also in order to account for the fact that the users' re-evaluation of their service perception is not instantaneous, we sample the time axis so that each user re-considers its acceptance of the service every ΔT seconds.

If user i enters service at time t_i and its anticipated call duration is T_i seconds, $\lfloor \frac{T_i}{\Delta T} \rfloor$ service evaluations might occur at most, beyond the first one at the time the user is allocated in the system. The call is successfully completed if and only if all these tests are passed. In particular, $r_i^{(0)}$ is the rate requested by user i before entering the system. This rate is assumed equal to the allocation which maximizes its satisfaction probability, i.e.,

$$r_i^{(0)} = \arg\max A(r_i)\,. \tag{1}$$

For $0 < j \leq \lfloor \frac{T_i}{\Delta T} \rfloor$, $r_i^{(j)}$ is instead defined as the average rate perceived up to the jth evaluation, i.e.,

$$r_i^{(j)} = \frac{1}{j\Delta T}\int_{t_i}^{t_i + j\Delta T} r_i(t)\mathrm{d}t\,,$$

where t_i is the start time of user i's service.

The values of $r_i^{(j)}$, with $j > 0$, are meaningful only if the user does not leave the service, since in this case its allocated rate drops to zero. In order to evaluate this aspect, we define, by exploiting the concept of *conditional probability*, the conditional acceptance of r_i' given that r_i was an acceptable assignment as:

$$A(r_i'|r_i) = \begin{cases} \frac{A(r_i')}{A(r_i)} & \text{if } A(r_i') \leq A(r_i) \\ 1 & \text{if } A(r_i') > A(r_i)\,. \end{cases} \tag{2}$$

The call of user i is successfully completed with probability

$$P_i[\text{complete\_service}] = A_i(r_i^{(0)}) \prod_{j=1}^{\lfloor T_i/\Delta T \rfloor} A_i(r_i^{(j)}|r_i^{(j-1)})\,.$$

Otherwise, we distinguish between the case in which the service is evaluated as unacceptable already at the first evaluation, which happens with probability $1 - A_i(r_i^{(0)})$, and the case of service refusal in a subsequent evaluation, when the user is already in the system, whose probability is

$$A_i(r_i^{(0)})\left(1 - \prod_{j=1}^{\lfloor T_i/\Delta T \rfloor} A_i(r_i^{(j)}|r_i^{(j-1)})\right).$$

In the former case the user is said to be *blocked*, in the latter to be *dissatisfied*. In particular, we define

$$P_i[\text{diss\_instant\_}k] = A_i(r_i^{(0)})\left(\prod_{j=1}^{k-1} A_i(r_i^{(j)}|r_i^{(j-1)})\right)\left(1 - A_i(r_i^{(k)}|r_i^{(k-1)})\right)$$

which is the probability that the user is dissatisfied at the kth evaluation.

The distinction between blocked and dissatisfied users correctly reflects that they can refuse the service due to their own a priori decision of not entering the system, for example because of the price being too high, or can experience unacceptable service degradation due to a congestion arisen later. As is well known, the impact on the QoS of these events is considerably different. The reason for explicitly classifying also blocked users is that considering pricing implies that the system is admission controlled [6], as discussed in the introduction and as will be numerically shown in the following.

In this way, our previously discussed revenue evaluation can be formalized as follows. We evaluate the revenue R as the sum of paid prices, but subdividing it between the contributions determined by satisfied and dissatisfied users, respectively.

A satisfied user i will pay in the end a price $p\big(r_i^{(F)}\big)$ determined by the average rate $r_i^{(F)}$ perceived during its entire service connection, which is:

$$r_i^{(F)} = \frac{1}{T_i}\int_{t_i}^{t_i+T_i} r_i(t)\mathrm{d}t\,.$$

The revenue generated by satisfied users, $R^{(s)}$, is therefore determined as

$$R^{(s)} = \sum_{i\in\mathcal{Q}} p\big(r_i^{(F)}\big)P_i[\text{complete\_service}] \tag{3}$$

For what concerns users who are not satisfied, blocked ones do not generate revenue at all. The value $R^{(d)}$, which is the potential revenue generated by dissatisfied users, is instead equal to

$$R^{(d)} = \sum_{i\in\mathcal{Q}} \sum_{k=1}^{\lfloor T_i/\Delta T\rfloor} P_i[\text{diss\_instant\_}k]p\big(r_i^{(k)}\big). \tag{4}$$

For our evaluation it does not matter whether in the end dissatisfied users pay or not. Either virtual or real, a high revenue generated by dissatisfied users is an index of inefficiency, since it means that part of the resources have been wasted to be allocated to dissatisfied users. For this reason, a suitable provider's goal could be to maximize $R^{(s)}$ and minimize $R^{(d)}$ at the same time, or at least to trade-off one for the other.

III. Case Study

The aforementioned model is applied in this paper to a Hot-Spot scenario where a single IEEE 802.11b Access Point (AP) is in charge of managing a variable number of users. The performance of this case study is evaluated by means of experiments with the *ns-2* simulator. Note that this way of proceeding could be replaced by considering for example analytical evaluations of medium access sharing in CSMA scenarios [14], [15], which would be the critical point of the analysis. However, they either rely on approximations or are too complicated to be put in close form with the users' satisfaction framework. For this reason, such an integration aimed at obtaining an entirely analytical approach is left for further research.

The main element of the simulator is the so-called wNode, a typical node of *ns-2*, which we provided with some additional features to account for the micro-economic behavior of the WLAN users. We implemented the model of Section II in order to drive the choices of the users in terms of selecting the most suitable transmission rate, according to the trade-off between paid price and gained utility, so as to eventually evaluate the supplied QoS.

Also, in order to support the users' prioritization due to their different rate requirements, we focus on a Constant Bit Rate (CBR) service over User Datagram Protocol (UDP) and we set the packet length proportional to the requested rate. In this way the contention process within nodes is always fair but the time of transmission, and hence the bytes transmitted, are proportional to the *requested rate* $r_i^{(0)}$. In this way, we aim at representing real-time interaction with the traffic. This choice can be easily and directly replaced within the simulator by more complicated medium sharing mechanisms, even though other issues (e.g., about fairness or traffic shaping) would probably arise and would need to be addressed.

The users behave dynamically, coming and going from the Hot-Spot and setting up connections of different types in terms of duration and transfer rate. The arrivals follow a Poisson process. The arrival rate is λ and the service duration is exponentially distributed with parameter μ; however, the users might leave the system if they consider the service dissatisfactory (this is why the exit process is no longer Poisson when dissatisfied users begin to appear). The ratio λ/μ is still useful to understand how many users on average would be under service if the dissatisfied users did not abandon the system.

The micro-economic model previously discussed is used to evaluate the users satisfaction and is implemented in a distributed manner at each node. Essentially, there are three kinds of events that matter in the system: a new node establishing a connection, a node ending its connection due to successful service completion (these two are regulated through parameters λ and μ) and finally the evaluations of the users about their service, which might determine a premature termination due to dissatisfaction in the service received [7].

When a user i arrives to the system, the simulator evaluates at first $r_i^{(0)}$ as described by (1). With probability $A(r_i^{(0)})$, the user accepts to establish a connection at rate $r_i^{(0)}$, and in this case the expected duration T_i of the connection is also determined as a random exponentially distributed value with parameter μ. Then, the user is added to the system directly through the already implemented *ns-2* functionalities. This means that the rate provisioning of $r_i^{(0)}$ is not guaranteed, due to the possible presence of other users. It is possible that the transmission rates no longer match the initial requirements because of congestion, which means that for node i the perceived rate r_i is lower than $r_i^{(0)}$. In general, we can regard $r_i^{(0)}$ as the *requested* rate, and every value $r_i^{(j)}$ with $j > 0$ as the *achieved* rate after $j\Delta T$ seconds, keeping in mind that the achieved rate is not necessarily equal to the requested one (in case of congestion it is indeed lower), but due to the fair sharing property of IEEE 802.11 discussed in Section II the two values are roughly proportional to each other. For this reason, in the simulator, every $\Delta T = 20$ seconds the achieved rate is re-evaluated, based on the conditional probability in (2).

At the end of the simulation run, users can be subdivided in terms of how their transmission ended: as described in Section II there are blocked users that do not accept to establish the connection at all; also, other users may accept their initial transmission rate, but when it decreases due to other arrivals they perceive it as too low and so they exit from the system, i.e., they are dissatisfied users; finally, there are users that finish their transmission in a satisfactory manner.

IV. Utility and Pricing functions

From the mathematical point of view, both utility and pricing functions ($u(r)$ and $p(r)$, respectively) do not decrease as the allocated resource increases. However, whereas the price might even indefinitely increase as the allocation becomes larger and larger, the utility must saturate after a certain point.

In particular, we assume that, after a value $r_{\max}$, further utility increases are negligible. Thus, it is not meaningful to allocate the rate r for a single user outside the finite interval $[0, r_{\max}]$.

A natural limitation for r is the highest data rate which a terminal can achieve, e.g., 11 Mbps for the IEEE 802.11b standard. However, it is not sensible to choose $r_{\max}$ as this value, which can not be achieved in practice, unless a single user is present in the network, and even in this case rate fluctuations are still present. Therefore, in the numerical evaluations we let $r_{\max} = 5$ Mbps, which is a more sensible value. In other words, the data rates enjoyed by the users are capped to 5 Mbps. Indeed, evaluations similar to the ones presented in Section II showed that, if any user of an IEEE 802.11b system required a rate higher than this value, it would be never satisfied, unless it is the only user of the system (in which case our analysis about multiple access would be meaningless).

Analytical expressions of the utility $u(r)$ might be obtained through subjective testing. This will result in a different utility function for each user. Moreover, utilities are assumed to be normalized between 0 and 1, and for every user we impose that $u(0) = 0$ and $u(r_{\max}) = 1$.

We consider sigmoid-shaped utility functions, as commonly done in many related papers [7], [8], [12]. In our mathematical representation, the middle-point and the slope of the sigmoid are regulated by means of numerical parameters, indicated with K and ζ, respectively. We take them as adjustable parameters for which $0 < K < 1$ and $\zeta > 1$, so that

$$u(r) = \begin{cases} K^{1-\zeta}\left(\dfrac{r}{r_{\max}}\right)^{\zeta} & \text{for } 0 \le r \le K r_{\max} \\ 1-(1-K)^{1-\zeta}\left(\dfrac{r_{\max}-r}{r_{\max}}\right)^{\zeta} & \text{for } K r_{\max} < r \le r_{\max} \end{cases} \tag{5}$$

We stress that this specific choice is reported here only in order to clarify how numerical evaluations are obtained; however, it is not mandatory at all for the model, whose validity is still preserved for different choices of utility functions, as long as they satisfy the general properties outlined in Section II, e.g., in terms of monotonicity. Moreover, we also believe our choice to be realistic, as it correctly describes a reasonable user behavior and includes tunable parameters in order to depict a wide range of subjective preferences. The provider can not affect these functions, but it may estimate them in a database of users' requests. In this sense, a parametric representation may even be used by the provider to map the preferences of each user via a finite number of parameters.

The choice of the pricing function is instead made by the provider, and determining a good pricing function is key to manage and regulate the medium access in an efficient manner. In this paper, we adopt a simple tunable framework which allows us to investigate and understand what a suitable pricing function for the WLAN service may be. In particular, we need to capture the trade-off between flat and linear pricing [17].

Observe that not only do the tariffs paid impact on the obtained revenue and, more in general, on the satisfaction of paying users, but they also determine the network access for new connections, both in the sense of offering an acceptable price and under the point of view of how the resources are shared, as discussed in Section II. This means that pricing has a two-fold effect on revenue collection: it determines the amount that every user pays, but also allows for an implicit form of admission control, which, if properly performed, may additionally adjust the revenue and/or other goals.

In this respect, the two aforementioned pricing strategies have opposite behaviors, which justifies the need for the correct regulation of their trade-off. Flat pricing policies are easy to understand for the users, and in fact are widely employed for this reason, and they additionally encourage high resource utilization, as they let the users have a large amount of resources without raising the price. However, when the network load is high, flat pricing strategies cause congestion more often, since the users are not forced in any way to regulate their resource demand according to the price, which is fixed anyway. On the other hand, linear pricing solves this problem since it obtains a self-regulation of the users, due to the fact that the higher the request, the higher the price. For this reason, unnecessarily high demands are avoided. The negative consequence of this principle is that this kind of pricing results in a lower revenue for the provider, as will be numerically shown in the next section.

Since we are interested in quantitatively evaluating this trade-off, we will focus in the following on a general pricing policy framework, where we consider pricing functions being a mixture of a flat and a linearly increasing behavior. By appropriately tuning the steepness of the pricing function via a parameter, called *pricing shape factor* q, it is possible to switch from a fully linear ($q \to \infty$) to a fully flat price ($q \to 0$). This framework can be useful to also identify intermediate functions which mediate between flat and pricing strategies. Also, in a network design view, we can even search for an optimal pricing function, which achieves the best design trade-off.

Our choice in this paper is to consider the following $p(r)$:

$$p(r) = r_{\max}^{-1}\left(1 - q\ln\left(1+\frac{1}{q}\right)\right)^{-1}\frac{r\bar{p}}{r_{\max}\,q + r}, \tag{6}$$

where the pricing shape factor q determines whether the pricing policy is flat, or linear, or a hybrid between these two, whereas the *pricing scaling factor* $\bar{p}$ is defined as

$$\bar{p} = \int_0^{r_{\max}} p(r)dr\ . \tag{7}$$

Note that $\bar{p}$ is by definition independent of q. Thus, q turns the pricing from linear to flat, whereas $\bar{p}$ determines, to some extent, whether the pricing curve is overall "high" or "low," regardless of its shape.

The choice of this framework allows to decouple the evaluations where the price is increased for a given shape of the pricing function, i.e., q is constant and $\bar{p}$ varies, and the analysis of the most preferable pricing shape (i.e., of the optimal q) for a given scaling factor $\bar{p}$. This tunability, which will be exploited in the results, also reflects reality in the sense that the network provider can indeed regulate both aspects of the pricing function, i.e., not only the absolute value but also the shape of the tariff mechanism. Thus, our proposed

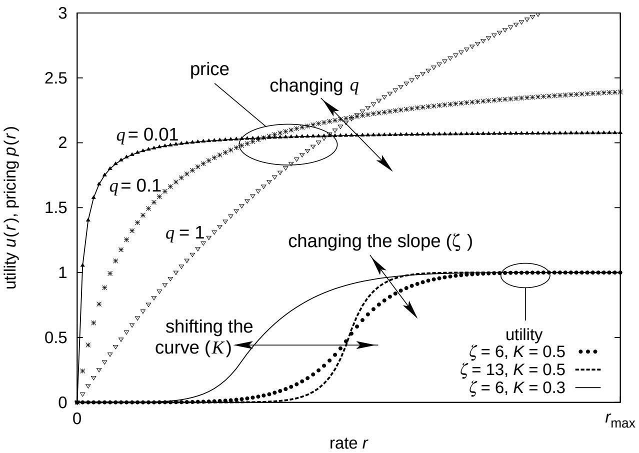


Fig. 2. Examples of sigmoid utilities chosen as $u(r)$ to represent the QoS perceived by the users, and of different pricing functions $p(r)$, from approximately flat to approximately linear, obtained by tuning q (just for graphical representation, $\bar{p}$ is set to 2).

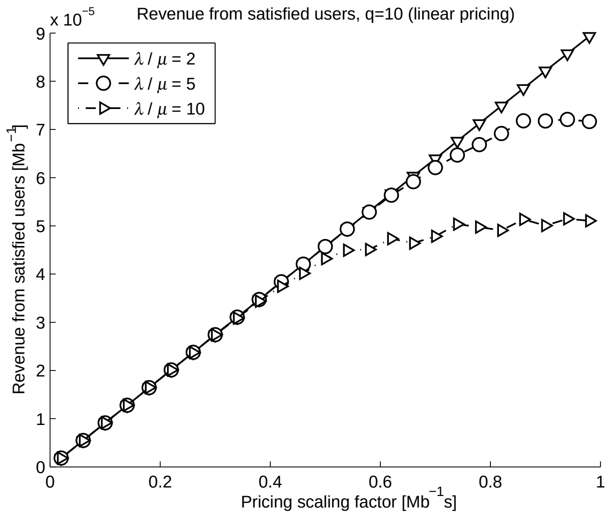


Fig. 3. Linear pricing policy, obtained with $q = 10$: revenue coming from satisfied users only, for different choices of the network load λ/μ, as a function of the pricing scaling factor $\bar{p}$.

framework can serve as an effective guideline to quantitatively estimate the proper pricing policy.

Examples of utility and pricing curves with diverse values of their inner parameters are reported in Fig. 2. Finally, for what concerns the analytical expression of the satisfaction function $A(u,p)$, we take the expression adopted in [12], which is:

$$A(u,p) = 1 - \exp(-ku^{\mu}p^{-\varepsilon}), \quad (8)$$

where k, ε and μ are to be chosen as proper positive constants, which tune the shape of the function and regulate the aforementioned QoS/price trade-off. For example, increasing μ makes the users more sensitive to the utility, whereas increasing ε does the same for the price. The last value, k, is simply a normalization constant. Anyway, observe that the choice reported in (8) is not restrictive, as the behavior of $A(u,p)$ can be easily adjusted by tuning the parameters, k, ε and μ, so this is indeed a very general choice. Moreover, this particular $A(u,p)$ could be replaced by any function respecting the general rationale of taking values in $[0,1]$ and being not decreasing in u and not increasing in p, respectively.

V. RESULTS

We consider a scenario where a single AP is located in the center of a 32 m $\times$ 32 m square area, surrounded by a variable number of mobile users with heterogeneous requirements in terms of service, which are mapped through different utility functions. Propagation effects and mobility have been implemented with already available *ns-2* modules, determining a radio scenario affected by slow fading with pedestrian mobility of the terminals, which move with an average speed of 0.5 m/s.

A total operation duration of 2000 s is evaluated. During this time, calls are generated according to a Poisson process with intensity λ and call durations are exponential with average duration $1/\mu = 1/150\,\mathrm{s}^{-1}$. The ratio λ/μ is tuned in order to obtain different traffic conditions, where an increasing value leads to a more congested system.

For what concerns propagation values, we consider the path gain of terminals as determined by the product between a path loss term and a lognormal shadowing term. The former is taken proportional to $d^{-3.5}$, where d is the distance between transmitter and receiver. The latter, expressed in dB, has zero mean and standard deviation equal to 6. We consider IEEE 802.11b implementation, so the maximum signalling rate of the terminals is 11 Mbps (this is not to be confused with the maximum data rate, capped at 5 Mbps as previously explained). Indeed, in the considered scenario, due to the small size of the network, this value is always available to all terminals. However, we also performed similar evaluations for more complicated scenarios where different signalling rates coexist and they exhibit similar trends to the ones shown.

The utilities of arriving users are generated with ζ and K randomly distributed in [6,20] and [0,0.85], respectively. The user satisfaction parameters are $\mu = 2$, $\varepsilon = 4$, $k = -\ln 0.9$. All these values are given as input to the *ns-2* simulator, and other parameters simply reflect the implementation of the IEEE 802.11b standard in this simulator.

We show the results of our evaluation in Figs. 3–8 for what concerns two different pricing policies, i.e., flat and linear pricing. These strategies are obtained within the tunable function $p(r)$ reported in (6), by assigning respectively a very low value and a very high value to the parameter q. Recall that this function tends to a flat or linear behavior for q tending to 0 or to infinity, respectively. In the numerical evaluations, the flat and linear policies are obtained with $q = 0.001$ and $q = 10$, respectively. Users arriving in the system are classified considering the three categories already explained. The generated revenue is evaluated as split between satisfied and dissatisfied users (blocked users are not considered as they do not generate revenue, not even virtually). The x-axes of all these figures report the pricing scaling factor $\bar{p}$; dimensionally, this is the inverse of a rate, thus it is measured in Mbps^{-1}, i.e., Mb^{-1}·s. When the y-axis reports a revenue value, this refers to the instantaneous value of the revenue, normalized to the time unit (in order to abstract from the duration of the simulation run). Thus, the unit of measure is Mb^{-1}.

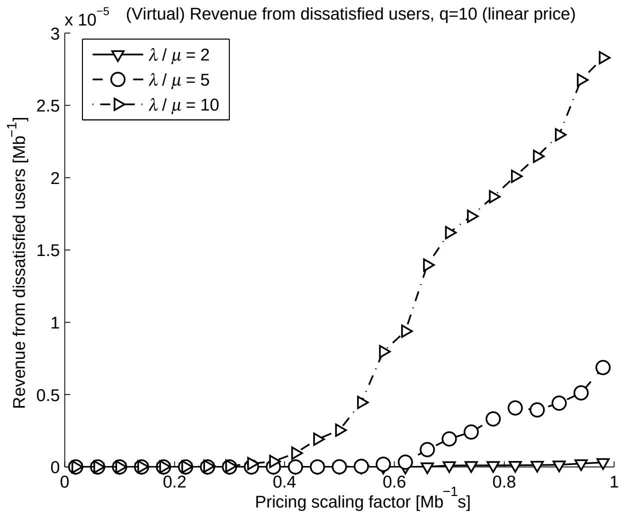


Fig. 4. Linear pricing policy, obtained with $q = 10$: (virtual) revenue coming from dissatisfied users, for different choices of the network load λ/μ, as a function of the pricing scaling factor $\bar{p}$.

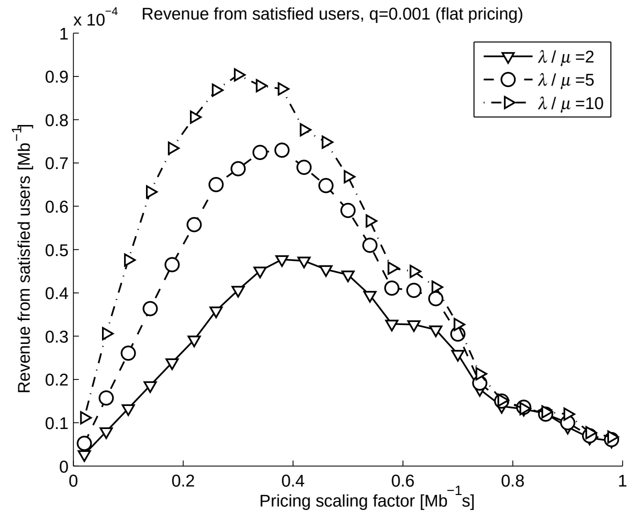


Fig. 6. Flat pricing policy, obtained with $q = 0.001$: revenue coming from satisfied users only, for different choices of the network load λ/μ, as a function of the pricing scaling factor $\bar{p}$.

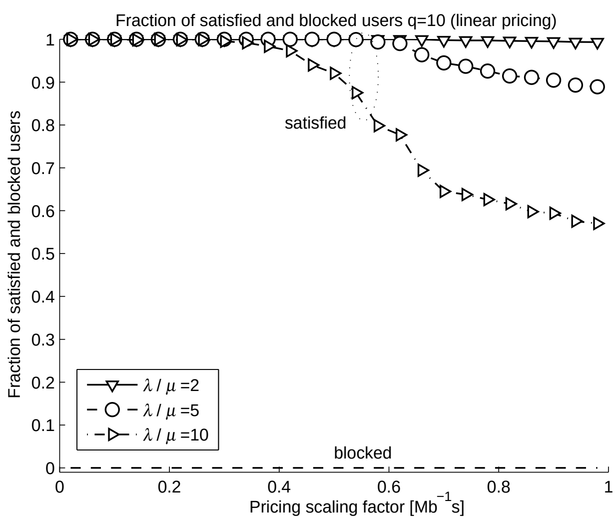


Fig. 5. Linear pricing policy, obtained with $q = 10$: satisfaction and blocking rate, for different choices of the network load λ/μ, as a function of the pricing scaling factor $\bar{p}$.

Fig. 3 shows that a linear pricing policy is able to efficiently regulate users' access to the system. In fact, when either the price or the network load is low, the revenue from satisfied users exhibits a linear increase, which means that almost all the users are able to satisfactorily end their service period. This is confirmed in Fig. 4, which shows low values of the (lost) revenue from dissatisfied users, and Fig. 5, where it is emphasized that almost all users are satisfied. The fraction of blocked users is very close to 0 for any value of the price (observe that blocking does not depend on the load, as it is evaluated before the users enter the system). This happens because, with this policy, the users *self-adapt* their demand to the price. A further consequence of this self adaptation is that users are better able to coordinate their resource sharing, since due to the linear behavior of $p(r)$, the most demanding users are charged the most. Increasing the price scaling factor $\bar{p}$, the penalty for the users with high requested rate also increases, which causes the users to decrease their rate request.

Thus, when the load is low, all users are able to access the system. In this case, all the bandwidth is used and the revenue is a linear function of the pricing scaling factor $\bar{p}$. For higher load values, the satisfaction rate and the revenue generated by satisfied users decreases. In this case, increasing $\bar{p}$ causes the revenue from satisfied users to saturate, as seen in Fig. 3. Because of the higher chance of collision, which implies a lower perceived quality, the higher the load, the lower the saturation value of the curve. Moreover, in view of the previous observation of almost all users being admitted in the system, Fig. 5 implies that applying a high price in a congested system would result in more than $1/3$ of dissatisfied users.

To sum up, when linear pricing policies are considered, the users have more freedom in their allocation choice. Thus, in general a linearly-dependent pricing achieves high satisfaction rate in absolute terms. However, in case of congestion, there is a risk of dissatisfying a large share of the users that try to enter the system but do not achieve a satisfactory service. Moreover, compared to other policies, a linear pricing does not generally obtain very high revenue.

Finally, we observe that price variations influence the provider's revenue in two ways. Besides determining directly the revenue proportionally to the unit price, they also affect users' requests and their satisfaction. This relation implies an *implicit admission control*, created by pricing the resource and therefore allowing the users to self-manage system access. A better understanding of this relationship between pricing and users' satisfaction can be gained considering another pricing strategy, i.e., flat pricing, shown in Figs. 6–8.

The flat case also applies implicit admission control, but in the same way to all users, whereas the linear case was adaptive to their requests. This implies that every user simply asks for the transmission rate which gives the highest utility, since the maxima of utility and acceptance probability occur for the same rate. Due to the increasing behavior of the utilities, this is obtained when the rate is as high as possible, i.e., at $r_{\max} = 5$

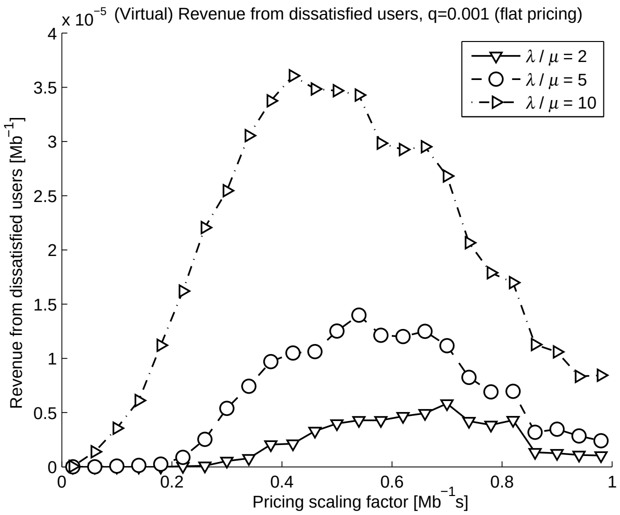


Fig. 7. Flat pricing policy, obtained with $q = 0.001$: (virtual) revenue coming from dissatisfied users, for different choices of the network load λ/μ, as a function of the pricing scaling factor $\bar{p}$.

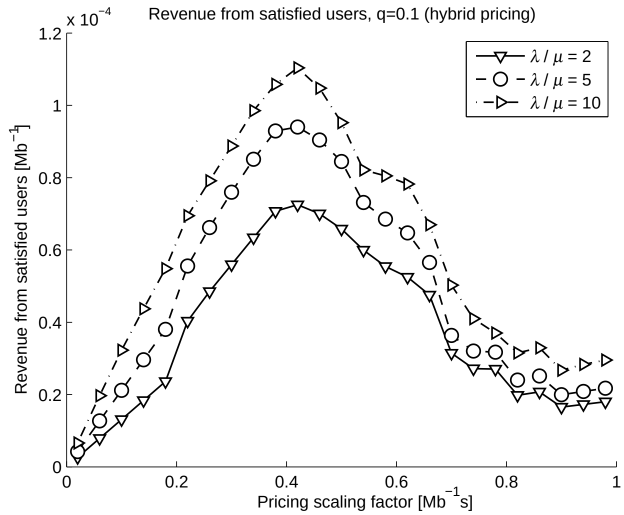


Fig. 9. Hybrid pricing policy, obtained with $q = 0.1$: revenue coming from satisfied users only, for different choices of the network load λ/μ, as a function of the pricing scaling factor $\bar{p}$.

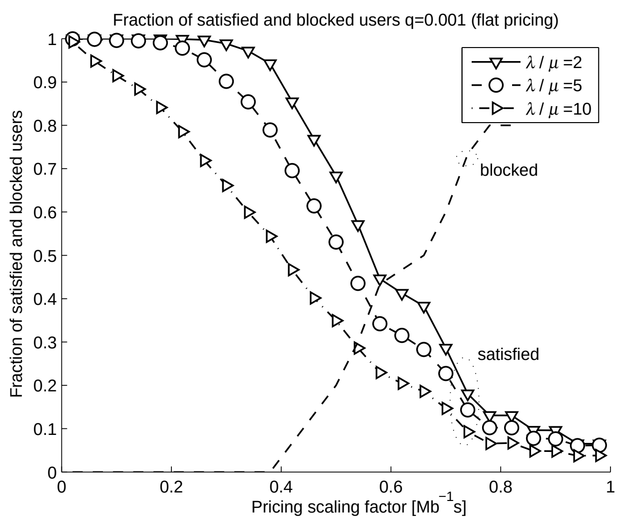


Fig. 8. Flat pricing policy, obtained with $q = 0.001$: satisfaction and blocking rate, for different choices of the network load λ/μ, as a function of the pricing scaling factor $\bar{p}$.

Mbps. Whereas in the linear case a self-adaptation of users' requests was observed, this is no longer possible for the flat pricing. Thus, there is a significant number of blocked user, as reported in Fig. 8. Again, note that this value does not depend on the load, but only on the price.

We remark that flat pricing leads both to higher rate requests, and also to generally increased revenue (see Fig. 6). Compared to the linear pricing case in fact, from the individual perspective of the users the fixed price to pay regardless of the rate seems to be relatively cheaper (especially if they request a high rate) although it is higher in absolute terms. However, the downside faced by the users (and hence also by the provider) due to these improvements is an increased congestion, which leads to an overall decrease in the users' satisfaction.

In particular, observe that in the linear price case the revenue from satisfied users did not change very much with the load, and the highest revenue was obtained for low values of the load. Fig. 6 shows instead that the flat pricing revenue is an increasing function of the load. On the other hand, also the revenue coming from dissatisfied users increases with the load, more rapidly than the revenue from satisfied users (compare the trend in Fig. 7, which is steeper, with the one of Fig. 6). The overall satisfaction rate is significantly lower for the flat price than for the linear pricing case (compare Fig. 8 with Fig. 5). On the other hand, with respect to the linear pricing case where almost all users were admitted, here most of the users are blocked, not dissatisfied. Even though in this paper we concentrate for simplicity on the satisfaction rate itself, it is also true that there is an important difference at the Admission Control level between blocked and dissatisfied users. In this sense, the performance of the flat pricing could be considered better than the linear policy, as it trades dissatisfied users for blocked ones. On the other hand, the overall number of satisfied users is extremely low.

This happens since users have no incentive to decrease their unnecessarily high rate requirements and henceforth the network is brought toward a low-performance operation point, where all users request the highest rate. In this case, a relatively higher revenue is achieved by serving few users which utilize the resource at the maximum level.

We therefore conclude that, even though flat pricing is often adopted in WLAN hot-spots, where the payment of a fixed fee guarantees the access for a given time (but without quality constraints), it is likely not to be suitable for heavily loaded scenarios, due to excessively low satisfaction rate. Instead, it might be a good choice in a class-based scenario for business customers, i.e., for a limited number of users with top requirements and therefore willing to pay more. Instead, to manage the majority of the customers without strong quality requirements, a usage-based linear pricing is more efficient.

Indeed, there is a trade-off between flat and linear pricing strategies. To further investigate this point, we can consider Figs. 9–11, where a hybrid pricing policy, obtained by setting $q = 0.1$ in (6), has been investigated.

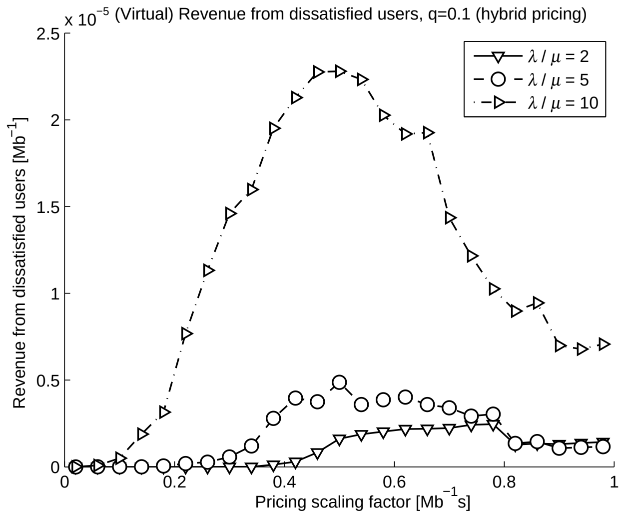


Fig. 10. Hybrid pricing policy, obtained with $q = 0.1$: (virtual) revenue coming from dissatisfied users, for different choices of the network load λ/μ, as a function of the pricing scaling factor $\bar{p}$.

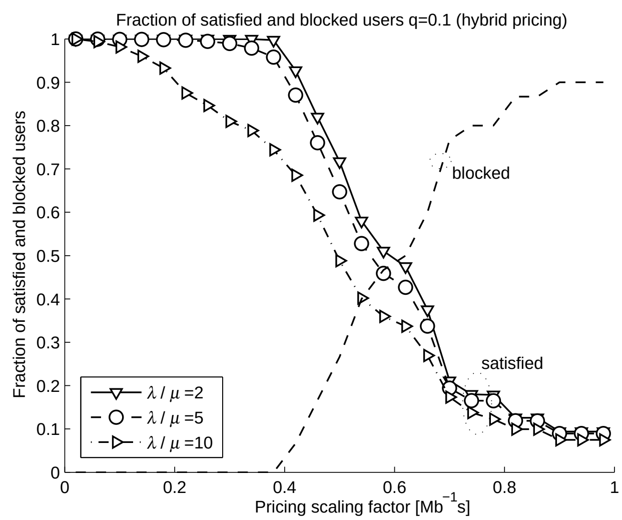


Fig. 11. Hybrid pricing policy, obtained with $q = 0.1$: satisfaction and blocking rate, for different choices of the network load λ/μ, as a function of the pricing scaling factor $\bar{p}$.

The behavior of the curves for the hybrid policy is qualitatively similar to the flat pricing curves. As Fig. 11 shows, the admission control is blocking a significant fraction of the users. However, with respect to the performance of the flat pricing reported in Figs. 6–7, this hybrid pricing obtains both a higher revenue from satisfied users (see Fig. 9) and a lower (lost) revenue from dissatisfied users (Fig. 10). In particular, this latter (negative) performance index is kept low when the network load is moderate. To sum up, it seems that a hybrid pricing with $q = 0.1$ could be preferable than a pure flat or a pure linear strategy. In particular, it obtains a slightly higher satisfaction rate than the flat pricing and a higher revenue than both policies.

Motivated by these results, we thus aim at generalizing them and exploring the aforementioned trade-off between users' satisfaction and provider's revenue by varying the parameter q in the tunable pricing policy reported in (6). To this end, we

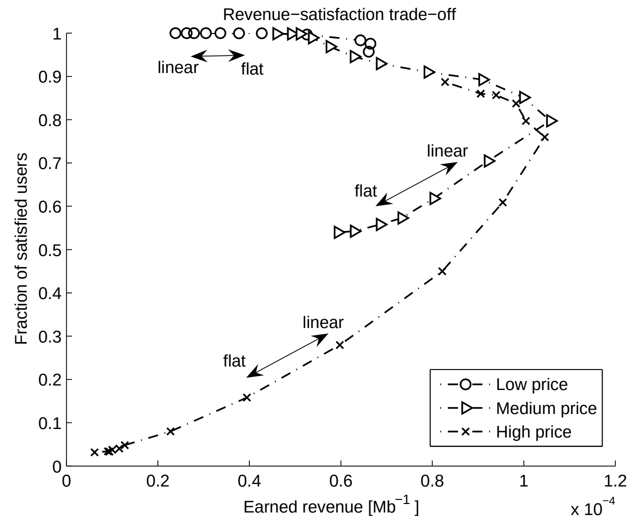


Fig. 12. Investigation of the suitable pricing policy. Trade-off between revenue and satisfaction. The value of $\bar{p}$ is 0.25 for Low Price, 0.5 for Intermediate Price, 1.0 for High Price).

focus on the case with $\lambda/\mu = 5$ and we consider three different values of the price scaling factor, i.e., $\bar{p} = 0.25, 0.5, 1.0$. The result of this investigation, reported in Fig. 12, can actually work as a strategy for the provider to take an appropriate choice between the two contrasting objectives. In other words, the shape of the pricing policy can be determined by looking at one suitable point in Fig. 12. To properly read the curves, note that points belong to a geometric sequence with q ranging from 10 to 0.0228 with 1.5 as ratio between adjacent samples. The point $q = 0.001$ is also added for completeness.

From the figure, the previously discussed behaviors of linear and flat policies are confirmed, i.e., linear pricing achieves higher satisfaction but also lower revenue. However, purely flat or purely linear strategies do not offer generally a good tradeoff, since the curves tend to wrap, and hybrid solutions are preferable. In fact, in many cases hybrid pricing interestingly achieves better revenue than both flat and linear pricing, and the resulting trade-off may be appealing for the provider, as the revenue is greatly increased at the cost of a small degradation of the users' satisfaction.

Thus, we emphasize the need for an appropriate investigation of all pricing policies by allowing more factors than the simple average price in order to tune the price not only quantitatively but also qualitatively (i.e., changing the shape itself of the pricing function). Moreover, our approach can be useful as an effective guideline to explore the trade-off in this sense. For example, according to the relative weight given in the provider's goal to the revenue versus the users' satisfaction, Fig. 12 allows in general to properly set q, and hence the pricing policy.

VI. CONCLUSIONS

We studied network management and pricing policies for a WLAN hot-spot, considering both technical and economic perspectives. The goal of the network manager includes different aspects, such as good network efficiency and high appreciation by the users, which concur to determining a satisfactory revenue. Moreover, several contrasting trends occur and need to be jointly addressed.

To capture these aspects, we applied a micro-economic framework to describe the behavior of the users. This model includes the trade-off between the requirement for a satisfactory QoS represented through the utility function u and the reaction to the pricing function p. We analytically formulated the metrics which impact on the provider's objective, especially focusing on revenue evaluation distinguishing between the contribution generated by *satisfied* and *dissatisfied users*.

To apply this analysis to a practical case, we implemented the micro-economic framework within the well known *ns-2* simulator. Our numerical results show that the overall behavior of the system is strongly affected by the micro-economic management. This is true both for the generated revenue (and especially the relationship between the revenue generated by satisfied vs. dissatisfied users), and the percentage of users which are successfully admitted in the system *and* complete their service in a satisfactory manner.

Thus, an appropriate choice of the pricing policy is key for the provider to obtain good system performance. In particular, the pricing strategy should regulate the users' access in order to prevent users from achieving dissatisfactory service due to congestion. To this end, we also explicitly addressed the trade-off between revenue and users' satisfaction as regulated by the *pricing shape factor*. An important conclusion is that pricing policies which are hybrid between a flat and a linear behavior often perform better or at least in a way which is more suitable for the provider. Our model not only quantitatively validates this statement, but also offers a framework to select the most appropriate shape factor according to this trade-off.

REFERENCES

[1] IEEE Std 802.11b, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Higher-Speed Physical Layer Extension in the 2.4 GHz Band," Supplement to IEEE 802.11 Standard. 1999 Edition.

[2] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Select. Areas Commun., Wireless Ser.*, vol. 18, no. 3, pp. 535–547, 2000.

[3] B. Rong, Y. Qian, and K. Lu, "Integrated downlink resource management for multiservice WiMAX networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 6, pp. 621–632, June 2007

[4] J. K. MacKie-Mason and H. R. Varian, "Pricing Congestible Network Resources," *IEEE J. Select. Areas Commun.*, vol. 13, no. 7, pp. 1141–1149, 1995.

[5] Y. Xue, B. Li, and K. Nahrstedt, "Optimal resource allocation in wireless ad hoc networks: a price-based approach," *IEEE Trans. Mobile Comput.*, vol. 5, no. 4, pp. 347–364, April 2006.

[6] D. Niyato and E. Hossain, "Call admission control for QoS provisioning in 4G wireless networks: issues and approaches," *IEEE Network*, vol. 19, no. 5, pp. 5–11, Sept.-Oct. 2005.

[7] S.K. Das, H. Lin, and M. Chatterjee, "An econometric model for resource management in competitive wireless data networks," *IEEE Network*, vol. 18, no. 6, pp. 20–26, Nov.-Dec. 2004.

[8] G. Bianchi, A. T. Campbell, and R. R.-F. Liao, "On utility-fair adaptive services in wireless networks," in *Proc. International Workshop on Quality of Service (IWQoS)*, pp. 256–267, 1998.

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

[10] R. Battiti, M. Conti, E. Gregori, and M. Sabel, "Price-based Congestion-Control in Wi-Fi Hot Spots", in *Proc. WiOpt*, pp. 91–100, 2003.

[11] J. Liu, M.M.A. Salama, and R.R. Mansour, "Identify the impact of distributed resources on congestion management," *IEEE Trans. Power Delivery*, vol. 20, no. 3, pp. 1998–2005, Jul. 2005.

[12] L. Badia, S. Merlin, A. Zanella, and M. Zorzi, "Pricing VoWLAN Services through a Micro-economic Framework," *IEEE Wireless Commun. Mag.*, vol. 13, no. 1, pp. 5–13, February 2006.

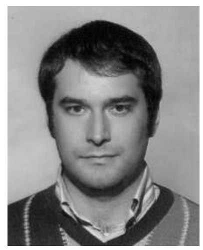
[13] K. Fall and K. Varadhan, *ns Manual* [Online]. Available: http://www.isi.edu/nsnam/ns/ns-documentation.

[14] F. Calì, M. Conti, and E. Gregori, "IEEE 802.11 wireless LAN: capacity analysis and protocol enhancement," in *Proc. IEEE INFOCOM*, vol. 1, pp. 142–149, 1998.

[15] M.H. Manshaei, G.R. Cantieni, C. Barakat, and T. Turletti, "Performance Analysis of the IEEE 802.11 MAC and PHY layer Protocol," in *Proc. IEEE WoWMoM*, 2005.

[16] I. Aad and C. Castelluccia, "Differentiation mechanisms for IEEE 802.11," in *Proc. IEEE INFOCOM*, pp. 209–218, 2001.

[17] C. Courcoubetis, F. P. Kelly, V. A. Siris, and R. Weber, "A study of simple usage-based charging schemes for broadband networks," *Telecommun. Syst.*, vol. 15, pp. 323–343, 2000.



Leonardo Badia (S'02, M'04) received the Laurea Degree (with honors) in Electrical Engineering and the Ph.D. in Information Engineering from the University of Ferrara, Italy, in 2000 and 2004, respectively. During 2002 and 2003, he was on leave at the Radio System Technology Labs (now Wireless@KTH), Royal Institute of Technology of Stockholm, Sweden. After having been with the Engineering Department of the University of Ferrara, he joined in 2006 the "Institutions Markets Technologies" (IMT) Lucca Institute for Advanced Studies, in Lucca, Italy, where he is currently a Research Fellow. He also collaborates with the University of Padova, Italy. His research interests include energy efficient Ad Hoc Networks, transmission protocol modeling, Admission Control and economic modeling of Radio Resource Management for Wireless Networks. Dr. Badia serves also as reviewer for several periodicals in the communication area.

Simone Merlin (S'05) received the Laurea Degree in Telecommunication Engineering (with honors) from the University of Padova, Italy, in 2004. During 2006 and 2007, he was a visiting scholar with the Wireless Networking Group at University of Illinois at Urbana Champaign. He is a Ph.D. student in the wireless networking group at the University of Padova, graduating in 2008. Currently, he is pursuing an internship at Qualcomm Inc., San Diego. His main research interests are on cross layer optimization in wireless ad-hoc and cellular networks.

Michele Zorzi (S'89, M'95, SM'98, F'07) received the Laurea Degree and the Ph.D. in Electrical Engineering from the University of Padova, Italy, in 1990 and 1994, respectively. During the Academic Year 1992/93, he was on leave at the University of California, San Diego (UCSD), attending graduate courses and doing research on multiple access in mobile radio networks. In 1993, he joined the faculty of the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy. After spending three years with the Center for Wireless Communications at UCSD, in 1998 he joined the School of Engineering of the Università di Ferrara, Italy, where he became a Professor in 2000. Since November 2003, he has been on the faculty at the Information Engineering Department of the University of Padova. His present research interests include performance evaluation in mobile communications systems, random access in mobile radio networks, ad hoc and sensor networks, energy constrained communications protocols, and broadband wireless access. Dr. Zorzi was the Editor-In-Chief of the IEEE WIRELESS COMMUNICATIONS MAGAZINE in 2003–2005, is currently the Editor-In-Chief of the IEEE TRANSACTIONS ON COMMUNICATIONS, and serves on the Editorial Boards of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the WILEY JOURNAL OF WIRELESS COMMUNICATIONS AND MOBILE COMPUTING and the ACM/URSI/KLUWER JOURNAL OF WIRELESS NETWORKS. He was also guest editor for special issues in the IEEE PERSONAL COMMUNICATIONS MAGAZINE (Energy Management in Personal Communications Systems) and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Multi-media Network Radios).