This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2011 proceedings.

Over-Booking Approach for Dynamic Spectrum Management

Jing Li, Qinghai Yang
State Key Laboratory on ISN,
School of Telecomm. Engineering,
Xidian University, No.2 Taibainan-lu,
Xi'an, 710071, Shaanxi,China.
Email:qhyang@xidian.edu.cn

Lajos Hanzo
School of Electronics and Computer
Science, University of Southampton,
Southampton, SO17 1BJ,
United Kingdom
Email: lh@ecs.soton.ac.uk

Kyung Sup Kwak
Graduate School of IT
and Telecomm., Inha University,
#253 Yonghyun-dong, Nam-gu,
Incheon, 402-751, Korea.
Email: kskwak@inha.ac.kr

***Abstract*—An over-booking based dynamic spectrum management (DSM) scheme is conceived for improving the attainable spectral efficiency. All secondary users (SU) will be categorized into different classes and they borrow spectral resources from the primary users (PU) before data transmission. Under the risk-based policy model, the effects of both booking cancellations and 'no-show' reservations are analyzed. Assuming that the booking demands obey an inhomogeneous Poisson process, we derive the optimal number of excess reservations, while minimizing the total compensation costs. Algorithms are developed for determining the capacity allocation dedicated to each SU class, whilst denying those resource allocations, which would lead to congested bookings.**

***Key words*: over-booking Approach, dynamic spectrum management**

I. Introduction

Recent studies have shown that the traditional fixed resource allocation policy is inadequate for supporting the contemporary high-load spectrum management systems. The potential spectral congestion may be averted by dynamic spectrum management (DSM), which is capable of dynamically accessing the under-utilized segments of the spectrum [1]. Naturally, the primary users (PU) are granted unhindered spectral access and then they may grant spectral access to the secondary users (SU). The devices are allowed to sense and explore a wide range of the availabe frequency band and identify the temporarily unused spectral resource blocks (RB) for their potential communications [2]-[3].

over-booking is an emerging revenue management approach, which is routinely employed in a number of industrial sectors [4]. It operates by allowing the resource provider to accept more reservations than the full capacity. In this treatise, we combine the over-booking strategy with the DSM approach for improving the attainable spectral efficiency. The economic feasibility of the over-booking strategy has been evaluated in previous studies [5]-[6]. Its rationale is that over-booking is efficient, since the SUs may cancel theirs reservations before their transmission is due or simply do not transmit without explicitly demanding a cancellation, due to reasons such as resource or network failures at the other end. Under over-booking, the following two scenarios may be encountered:

1) The number of active SUs is lower than the number of resource blocks (RB) made available by the PU. Then, in contrast to the conventional scenario of operating without over-booking, PUs will earn more revenue.
2) The number of active SUs is higher than the number of available RBs. Then, the PU satisfies the requests of the maximum possible number of SUs with its resources fully exhausted, and dismisses the unsatisfied SUs with some compensation.

The over-booking strategy combined with an efficient revenue management (RM) philosophy has been applied in many areas, such as airline, hospitality, rental car, railway and broadcasting businesses. Several models were constructed in [4] for examining the effects of over-booking policies on airline revenue and costs; Specifically, the simplest single-plane model was studied with respect to the optimal revenue-maximizing over-booking strategy and the arrival probability of ticket holders. Considering the inter-dependence between different spectrum 'consumers', a two-stage subletting mechanism of spectrum usage rights was proposed in [6]. In [7], a novel approach of using RM was presented for determining the reservation price in a grid system. This solution aimed for maximizing the profit by providing the right price for every product and for the different customers, and then by periodically updating the prices in response to market demands. In order to handle unexpected cancellations and 'no-shows' of reservations in the grid system, the authors of [8] investigated several static over-booking policies, such as probability based, risk based, and service-level agreement based approaches. The resource provider opted for accepting more reservations than its capacity.

Our novel contribution is that we extend the above-mentioned risk based model as our optimization tool and categorize the SUs into n classes, while assuming that their session-generation statistics obey the classic Poisson process.

*This research was supported by the RCUK for the UK-China Science Bridges Project: R&D on (B)4G wireless mobile communications, the MKE, Korea, under the ITRC program supervised by the NIPA (NIPA-2011-C1090-1121-0001) and the Fundamental Research Funds for the Central Universities (JY10000901002, JY10000901020), NSF China (60832001, 61001127), the 111 Project (B08038), China.

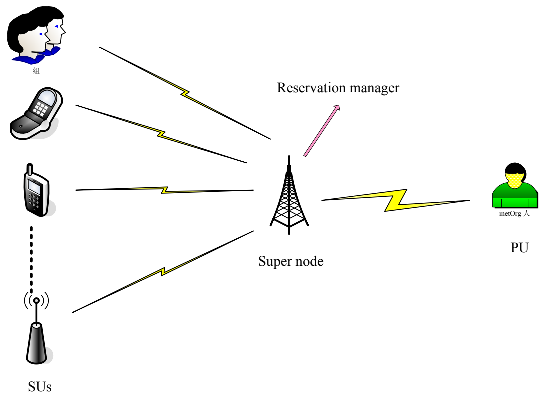

Fig. 1. System model

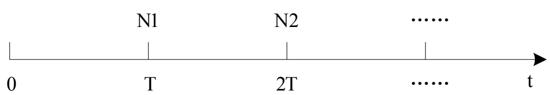

Fig. 2. PU's unused RBs in the time domain

More explicitly, we conceive an over-booking based DSM mechanism for determining the appropriate number of excess reservations, while minimizing the total compensation cost. Novel algorithms are developed for deciding which particular SU will be served and which will be dismissed in each SU class at a certain compensation cost.

II. System Model

The spectrum overlay-based DSM system under our consideration consists of a single PU, M SUs and a super-node, as shown in Fig.1. The PU reserves a set of spectral RBs, but it may use only a given fraction of them in a period, while offering the free RBs for lease.

The PU's unused RBs portrayed in the time domain is shown in Fig.2. At time instant T, we assume that the PU has $N1$ unused spectral RBs for leasing. It shares this information with the super-node, which may be regarded as the reservation manager. Based on this information, the super-node accepts reservations from the SUs.

All the SUs should submit their reservations for using spectral RBs to the super-node before commencing data communication. We assume that a SU reserves a single RB in a time slot. In this paper, we consider the situation within a single time slot, where the SUs submit their reservations during the period $[0, T)$ and actually use the spectral RBs at service instant T. We also assume that spectral RBs are indistinguishable for example in terms of their channel quality, but nevertheless, we charge different prices for different SU classes [9]-[10]. More specifically, we split the SUs into n different classes, based on their flexibility, price sensitivity and booking instant prior to the reserved period.

III. Problem Statement

In this section, we formulate our over-booking strategy. Fig.3 illustrates an example of the number reservations both

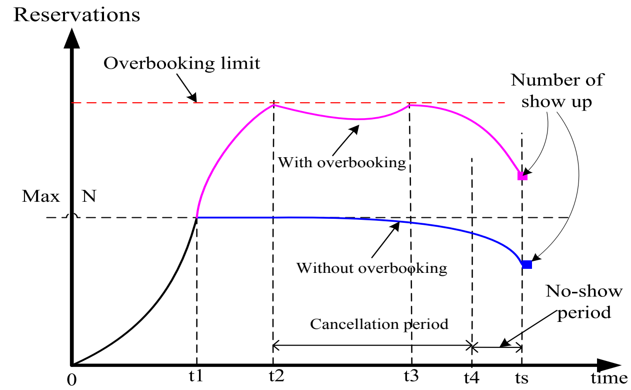

Fig. 3. An example of reservations with and without over-booking

with and without over-booking. A resource cancellation is defined as the event, when a SU terminates its reservation before the commencement of its service period, while a 'no-show' event as failing to commence transmissions without submitting a cancellation notice. The over-booking limit is set higher than the number N of the available spectral RBs, which is saturated at instant t_1 of Fig.3. The PU still accepts reservations after the instant t_1 until it reaches the over-booking limit, which is significantly higher than the actual capacity N. By contrast, the PU operating without over-booking has to deny potential reservations after the instant t_1 of Fig.3. In practice, there will be both cancellation and 'no-show' events between t_1 and the service instant t_s. As a result, the PU relying on over-booking may observe that more SUs commenced their sessions at t_s than without over-booking. Naturally, this may increase the profit.

However, the number of SUs actually occupying the reserved spectral RBs at the service instant t_s of Fig.3 may not accurately match the PU's true capacity N. If it is larger than N, the PU will have to dismiss the excess SUs and pay them compensation. Denoting the revenue accruing from a single reservation by u_i, the general compensation cost by C and the number of total reservations actually occupied at the service time t_s by M, the total revenue of the PU may be expressed as

$$U = \sum_{i}^{M} u_i - C = M\bar{u} - C. \tag{1}$$

To generate as much revenue as possible, while improving the spectral efficiency, we aim for increasing M while minimizing C. In the best-case scenario, there is no need to deny any excess reservations due to having a large number of cancellations and no-shows. However, this scenario imposes three problems:

1) How to find an optimal over-booking limit that exceeds the maximum capacity N, without incurring an excessive compensation cost.
2) How to decide the over-booking limit for each SU class in order to protect the higher-tariff SUs from being denied access during peak booking periods.

3) If SU congestion is encountered at the service time t_s, how to protect higher-tariff bookings from being denied in the first instance.

These problems will be addressed below.

IV. Risk-Based Policy

In this section, a risk-based policy is proposed for determining the optimal over-booking limit, while taking into account both the profit and the compensation cost. Denoting the price of a spectral RB as p, we can define the compensation cost to be paid to each SU as $c = p(1+k)$, where k is a positive constant. Obviously, we have $c > p$. To calculate the over-booking limit, we have to know the probability of SU session initiation and the true number of SUs requiring a RB.

Let us assume that the SU session initiations obey a Poisson process having a parameter λ. Within a booking interval T, on average, there will be λT users submitting their reservations.

Let $A(x)$ denote the probability that the number of reservations A is less than or equal to x, where x is the number of bookings, which obeys:

$$A(x) = \Pr(A \le x) = \sum_{n=0}^{x} \frac{(\lambda T)^n}{n!} e^{-\lambda T}. \tag{2}$$

Denote the number of bookings that will indeed be occupied at the serving time instant t_s of Fig.3 by B. We define furthermore $F_x(y)$ as the probability that B is lower than or equal to a certain number y, yielding:

$$F_x(y) = \Pr(B <= y \,|x). \tag{3}$$

Let us now derive the expected revenue difference encountered upon increasing the booking limit from b to $(b+1)$. We decompose the problem into the following three cases:

1) $\lambda T < (b+1)$, which implies that increasing the limit does not affect the actual RB demand imposed by the number of SUs requiring a RB, hence no revenue change is encountered. The probability of this scenario is $A(b)$.
2) When we have $\lambda T > (b+1)$, the true number of SUs requiring a RB is lower than N. Then, an extra profit of p is obtained, since this demand may indeed be satisfied.
3) If $\lambda T > (b+1)$ and the number of 'shows' is larger than N, the PU has to deny access for a SU at the cost of compensation. A loss of $c-p$ is incured, where $c > p$.

Thus, we can formulate the expected revenue difference as:

$$\begin{aligned} &\mathcal{E}[R(b)] \\ &= [1-A(b)]\,[pF_{b+1}(N)+(p-c)(1-F_{b+1}(N))] \\ &= [1-A(b)]\,[p-c(1-F_{b+1}(N))]. \end{aligned} \tag{4}$$

As long as $\mathcal{E}[R(b)]$ is larger than zero, the over-booking limit will be increased. This process is illustrated in **Algorithm 1**, where the optimal booking limit b_o has a minimum given by the maximum number of RBs N, and increases until $\mathcal{E}[R(b)]$ becomes zero or negative.

In multi-class scenarios, given an additional booking, the increased revenue cannot be so readily calculated. We thus employ an alternative approach relying on the weighted average price of

Algorithm 1: Over-booking Limit using a Risk-Based Policy

$b_o \leftarrow N$;
$R \leftarrow [1 - A(b_o)] \cdot [p - c \cdot (1 - F_{b_o+1}(N))]$;
while $R > 0$ do
  $b_o \leftarrow b_o + 1$;
  $R \leftarrow [1 - A(b_o)] \cdot [p - c \cdot (1 - F_{b_o+1}(N))]$;
end
return b_o.

$$\hat{p} = \sum_{i=0}^{n} \mu_i p_i. \tag{5}$$

Note that different SU classes obey dissimilar Poissonian processes. Here we resort to a weighted average density of

$$\hat{\lambda} = \sum_{i=0}^{n} \mu_i \lambda_i. \tag{6}$$

Upon replacing p and λ by their weighted versions in Eq. (5) and (6) in **Algorithm 1**, respectively, we arrive at the optimal over-booking limit b_o suited for multi-class SU scenarios.

Let us now derive the SU 'show-up' probability $F_x(y)$ at time instant t, under the assumption that each SU's probability of appearance is independent of that of the other. Denote the 'show-up' probability of a single SU at time instant t as q. The number of 'shows' follows a Binomial distribution with the probability density function (pdf) of

$$f_x(y) = \sum_{k=0}^{y} \binom{x}{k} q^k (1-q)^{x-k}. \tag{7}$$

Accordingly, we arrive at:

$$F_x(y) = \sum_{k=0}^{y} \frac{x!}{(x-k)! * k!} q^k (1-q)^{1-k}. \tag{8}$$

In the rest of this treatise, simulation results are provided for illustrating the optimal over-booking level b_o versus the following parameters: spectral RB price p, ratio of compensation to price $(1+k)$, the Poisson density λ, 'show-up' probability q and the number of blocks N. The curves are averaged over 10,000 simulation runs. Our results demonstrate that the value of b_o is independent of the former three parameters. In other words, it is only dependent on the show-up probability q and the number of RBs N. The impact of the 'show-up' probability is characterized in Fig.4, when using the parameters of p=200, $(1+k)$=2, λ=6 and N=100.

These values are representative of relevant practical situations. We observe that b_o monotonically decreases as the 'show-up' probability increases, where we arrive at a piecewise-linear curve, because b_o is an integer corresponding to different values of q. The small inset in the upper right corner shows the details of a specific segment

Likewise, the impact of the number of blocks N on b_o is depicted in Fig. 5 for the parameter values of p=200, $(1+k)$=2, λ=6 and q=0.85. We observe that b_o monotonically increases with N.

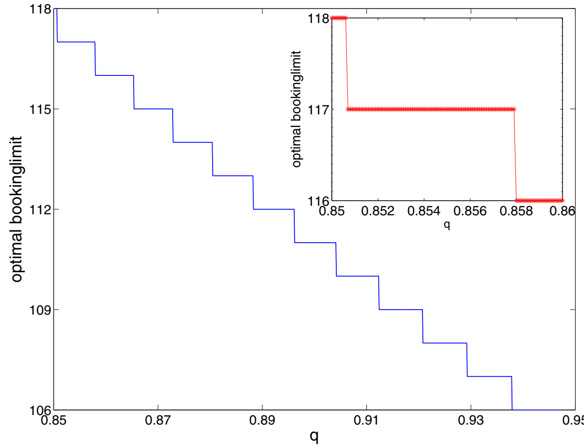

Fig. 4. Optimal over-booking limit b_o with different q

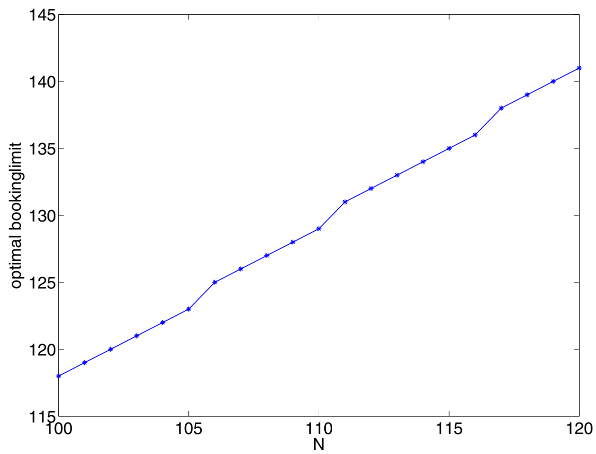

Fig. 5. Optimal over-booking limit b_o versus N

V. Capacity Allocation with over-booking

Our capacity allocation problem is that of finding the booking limit for each SU class, in order to maximize the overall expected revenue. If too many spectral RBs are allocated to lower-tariff SUs during peak periods, we may lose the chance of earning more revenue from accepting prospective high-tariff SUs. By contrast, having an insufficient quota for the lower-tariff SUs in off-peak periods may lead to poor spectral efficiency and less revenue. Thus, striking an attractive compromise for each class at different time periods is a key issue.

TABLE I
Example of SU segment

Class	User Category	Constraints
1	Best-effort	Non-refundable, only for a part of RBs
2	Emergency service	Allow cancellation

Here, we present an example, where the SUs are split into two classes. As listed in TABLE 1, each class has a set of constraints, which correspond to dissimilar prices. In this paper, we assume that Class 1 (e.g., best-effort) SUs reserve their RBs prior to Class 2 (e.g., emergency service) SUs. An example of the reservation process is shown in Fig.6. To prevent higher-class SU bookings from being rejected, an increased protection level of b_2 is required, i.e. once the booking limit b_1 of Class 1 is reached, future SUs will be relegated to the second class. Here we set $p_2 > p_1$.

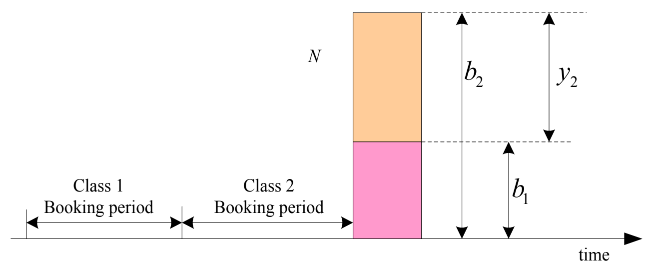

Fig. 6. Protection levels and booking limits for two classes

Let us denote the cumulative distribution function of the RBs requested in class i as $D_i(x)$. Note that our capacity allocation analysis is based on future bookings. We assume that the RB request function of SUs obeys the Poisson distribution with parameters λ_1 and λ_2, respectively. Upon representing the unit time interval by τ, we arrive at:

$$D_i(x) = \sum_{n=0}^{x} \frac{(\lambda_i \tau)^n}{n!} e^{-\lambda_i \tau}. \tag{9}$$

Likewise, upon increasing the booking limit from b_1 to (b_1+1), the expected revenue changes. We denote the difference as $R(b_1)$, which also depends on the RB demand d_i of the two SU classes. Similarly to the single-class case, we shall consider the following three scenarios:

1) If $d_1 < (b_1 + 1)$, then the expected revenue remains the same;
2) If $d_1 > (b_1 + 1)$, then the revenue depends on d_2; If $d_2 < (N - b_1)$, the revenue will be increased by p_1;
3) If $d_1 > (b_1 + 1)$ and $d_2 > (N - b_1)$, the PU will lose $(p_2 - p_1)$.

Accordingly, the expected revenue difference may be formulated as

$$\begin{aligned} \mathcal{E}[R(b_1)] &= [1 - D_1(b_1)][p_1 D_2(N - b_1) \\ &\quad + (p_1 - p_2)(1 - D_2(N - b_1))] \\ &= [1 - D_1(b_1)][p_1 - p_2(1 - D_2(N - b_1))]. \end{aligned} \tag{10}$$

The process of computing b_1 is illustrated in **Algorithm 2**. As a direct result, the protection level of Class 2 is given by $(N - b_1)$.

Let us now consider our simulation results provided for illustrating the capacity allocation of each class. In Fig.7, we observe that as N increases from 40 to 46, b_1 is gradually increased, while b_2 remains fixed. By contrast, for $N > 46$, b_1 converges to the limit value 34 and b_2 increases gradually.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2011 proceedings.

Algorithm 2: Booking Limit (N, p_1,p_2,d_2)

$b_1 \leftarrow 0$;
while $b_1 < N$ do
  $b_1 \leftarrow (b_1 + 1)$;
  $\mathcal{E}[R(b_1)] \leftarrow [1 - D_1(b_1)] \cdot [p_1 - p_2(1 - D_2(N - b_1))]$;
  if $\mathcal{E}[R(b_1)] <= 0$, **return** $b_1 - 1$;
end
return b_1.

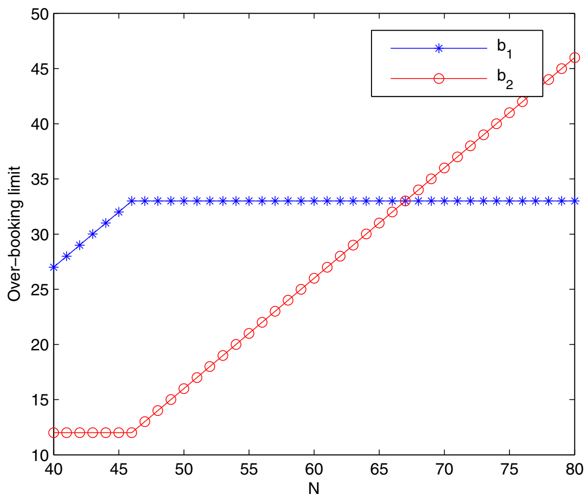

Fig. 7. Each SU class over-booking limit for different N

This implies that for a given number of RBs, there is a channel access limitation for Class 1 SUs, which creates room for future potential higher-class SUs, promising an increased spectral efficiency and increased revenue. Note that the sum of b_1 and b_2 equals to b_o in Fig.5.

VI. Lower-Class Denial Scheme

In this section, we propose a RB request-rejection scheme for reactively deciding upon which particular excess reservation should be dismissed, based on both c and the user class. The idea behind this approach is to protect higher-class bookings from being denied at first, while there are some lower-class RB requests, which may be rejected at a lower loss of revenue. Hence, we refer to this scheme as the lower-class denial scheme, which is formally presented in **Algorithm 3**.

Initially, we obtain a list of bookings, which are stored in the *bookingList*. We also have to keep track of the excess number of requests, beyond the current number of RBs N, and the total number of reservations obtained from *bookingList*. This value is stored in *overbooked*. Then, we sort the entities stored in *bookingList* based on c for each class. Then, as the number of denied RB reservation attempts, denoted in 'denial' increased, the algorithm continues rejecting more bookings from the list, until we arrive at *denial*=*overbooked*.

The algorithm denies the lower-class SUs in *bookingList* first. If the lower-class *bookingList* is empty, the algorithm

Algorithm 3: Lower-Class Denial Approach

input T **and** N;
$bookingList \leftarrow$ # of booking list (T);
$overbooked \leftarrow$ # of total over-bookings (bookingList)-N;
$denial \leftarrow 0$;
sort($bookingList, class\_denied\_cost$);
while $denial < overbooked$ **do**
  $data \leftarrow$ # of bookings($bookingList, HEAD$);
  remove($data, bookingList$);
  $denial \leftarrow denial+$ # of toal over-bookings (data);
end

continues by denying the higher-class SUs. As a result, the higher-class bookings are better protected from being denied. This will further increase the PU's profit and improve the attainable spectral efficiency.

VII. Conclusions

An over-booking based DSM approach was proposed for improving the attainable spectral efficiency via optimizing the number of excess RB reservations. We applied a risk-based policy model for finding the optimal over-booking limit without incurring an excessive compensation cost. Booking cancellations and 'no-show' reservations were allowed and the RB's reservation demands were assumed to obey an inhomogeneous Poisson process. Furthermore, we analyzed the optimal over-booking limit. Finally, algorithms were developed for determining the resource allocation of each class of SUs and for deciding which paritcular congested bookings should be dismissed at a given compensation cost.

References

[1] J. Whalley and P. Curwen, "Third Generation New Entrants In the European Mobile Telecommunications Industry," *Telecommunications Policy*,vol. 30, pp. 622–632, Dec. 2006.

[2] H. Wang, L. Gao, X.g Gan, X. Wang and E. Hossain, "Cooperative Spectrum Sharing in Cognitive Radio Networks: A Game-Theoretic Approach," *in Proc. IEEE ICC2010*, pp. 1-5, May 2010.

[3] Y. li, X. Wang and M.Guizani, "Resource Pricing with Primary Service Guarantees in Cognitive Radio Networks: A Stackelberg Game Approach," *in Proc. IEEE GLOBECOM 2009*, pp. 1-5, Dec. 2009.

[4] D. P. Kraines, "Optimal Overbooking," *The UMAP Journal*, Vol.23, no.3, pp. 283-299,2002.

[5] B. Urgaonkar, P. Shenoy and T. Roscoe, "Resource Overbooking and Application Profiling in Shared Hosting Platforms," *in Proc. ACM OSDI '02*, vol. 36, no.SI, 2002.

[6] L. Mastroeni and M. Naldi, "Option-based Dynamic Management of Wireless Spectrum," *NGI '09*, pp. 1-8, 2009.

[7] A. Sulistio, K. H. Kim, and R. Buyya, "Using Revenue Management to Determine Pricing of Reservations," *in Proc.the 3rd Intl. Conference on e-Science and Grid Computing*, pp. 396-405, Dec. 2007.

[8] A. Sulistio1, K. H. Kim, and R. Buyya, "Managing Cancellations and No-shows of Reservations with Overbooking to Increase Resource Revenue," *CCGRID 8th IEEE International Symposium*, pp. 267-276, May 2008.

[9] H. Mutlu, M. Alanyali, and D. Starobinski, "Spot Pricing of Secondary Spectrum Access in Wireless Cellular Networks," *IEEE/ACM Trans. Networking*, vol. 17, no. 6, pp. 1794-1804, Dec. 2009.

[10] D. Niyato, E. Hossain, and Z. Han, "Dynamic Spectrum Access in IEEE 802.22-based Cognitive Wireless Networks: a Game Theoretic Model for Competitive Spectrum Bidding and Pricing," *IEEE Wireless Communications*, vol. 16, no. 2, pp. 16-23, Apr. 2009.