Optimal and Heuristic DSA Policies for Cellular Networks with Coordinated Access Band

M. Coupechoux, H. Kamal, Ph. Godlewski, J. M. Kélif

▶ **To cite this version:**

HAL Id: hal-01547184

https://hal-imt.archives-ouvertes.fr/hal-01547184

Submitted on 26 Jun 2017

Optimal and Heuristic DSA Policies for Cellular Networks with Coordinated Access Band

Marceau Coupechoux, Hany Kamal, Philippe Godlewski
TELECOM ParisTech & CNRS LTCI
46, rue Barrault, Paris, France
{firstname.lastname}@telecom-paristech.fr

Jean-Marc Kelif
France Telecom R&D
Issy-Les-Moulineaux, France
jeanmarc.kelif@orange-ftgroup.com

Abstract—Due to the increasing demands for higher data rate applications, also due to the actual spectrum crowd situation, DSA (Dynamic Spectrum Access) turned into an active research topic. In this paper, we analyze DSA in cellular networks context, where a CAB (Coordinated Access Band) is shared between RANs (Radio Access Networks). We propose an SMDP (Semi Markov Decision Process) approach to derive the optimal DSA policies in terms of operator reward. In order to overcome the limitations induced by optimal policy implementation, we also propose a simple, though sub-optimal, DSA heuristic. The achieved reward is shown to be very close to the optimal case and thus to significantly exceed the reward obtained with FSA (Fixed Spectrum Access). Higher rewards and better spectrum utilization with DSA optimal and heuristic methods are however obtained at the price of a reduced average user throughput.

I. Introduction

Wireless networks are facing increasing demand for high data rate applications, and hence their demand for spectral resource increases. Researchers have started working on DSA algorithms as a solution to the spectrum scarcity problem encouraged by the rapid progress in SDR (Software Defined Radio) systems that enable the required reconfigurability for DSA and cognitive radio equipments.

In [1], the spectrum management models are divided into four main axis: command and control, exclusive-use, primary/secondary usage, and commons. The exclusive-use model includes a dynamic mode, where spectrum is owned by a single operator at any given point in space or time; owner and usage of the spectrum can however dynamically change. This model is thus particularly adapted to cellular networks. In this context, the IEEE P1900.4 working group has detailed three use cases [2] [3] with increasing levels of reconfigurability and joint management of resources. In this paper, we focus on the first one and consider a single operator with several RANs, able to dynamically distribute its frequency bands between its RANs.

Several papers are dealing with DSA for cellular networks. For example, in [4], authors propose a coordinated DSA system where a pool of resources (CAB or Coordinated Access Band) is shared and controlled by a regional spectrum broker. In [5], authors made use of the genetic algorithm to analyze the DSA in WCDMA networks. In [6], authors propose a MAC protocol enabling ad-hoc secondary users to utilize the unused resources of a GSM system.

It is however difficult to separate technical from pricing aspects when DSA is considered, especially for cellular operators who pay very high prices for the license. The wide interest in DSA is indeed mainly driven by the expected benefits resulting from sharing the spectrum [7]. Reference [8] analyzes a network model where the service providers base stations are sharing a common amount of spectrum. A distributed DSA algorithm is proposed where each user maximizes his utility (bit rate) minus the payment for the spectrum. In [9], authors have considered a spectrum market, where they propose a Rubinstein-Stahl method for the spectrum trading.

In this paper, we present an approach based on SMDP to analyze DSA in a cellular context. We analyze a network model, where different RANs are sharing a CAB, inspired by the idea of resources sharing proposed in [4] and by the single operator use case presented in [3]. We take into account the spectrum price, and maximizing the operator revenue is our main concern.

MDP approach has been used to solve several optimization problems in telecommunication networks. In the context of cognitive radio, reference [10] proposes a cognitive medium access protocol that maximizes the throughput while limiting the interference affecting the primary user. The authors formulated the problem within the framework of constrained MDP. In [11], a POMDP (Partially Observable MDP) framework is proposed to optimize the performance of the secondary users while limiting the interference perceived by the primary users. These references however focus on the primary/secondary usage model.

In [12], the SMDP framework is used in a JRRM (Joint Radio Resource Managment) context in order to take an optimal CAC (Call Admission Control) decision, whether to accept a new coming call or to reject it. The reward function in [12] presents the end-user throughput. Different from [12], SMDP is used in this paper to find the optimal spectrum bands allocations. The reward function in this paper takes into account both user satisfaction and spectrum price.

The paper is organised as follows: Section II presents the network model in terms of system model, traffic model and the principle of DSA operation. The SMDP approach is presented in section III. In section IV, we propose a sub-optimal DSA heuristic easier to implement for an operator than optimal policies. The performances of optimal DSA policies

and DSA heuristic are compared to the FSA case in section V. Conclusion is given in section VI.

II. NETWORK MODEL

A. System model

We intend to study cell-by-cell DSA between two RANs. The system is thus made of two cells of two different RANs (in this paper, terms cell and RAN will be used indifferently). RANs do not have their own spectrum bands but rather have to dynamically access to a CAB. The CAB is sub-divided into m_{max} elementary spectrum bands (blocks) that can be used indifferently by any RAN. As traffic grows, a RAN can lease a new elementary band (one block) and as it decreases, the RAN can leave it free for the common pool (section III-C). We assume that the average data rate accessible by users in a RAN is proportional to the bandwidth allocated to the RAN and is equally divided among all users of the RAN (section II-B). The model is shown in Fig. 1. Parameters n_i, $i = 1, 2$ are the number of active users in RAN1 and RAN2. Parameter m_i is the current number of elementary bands leased by RAN i from the CAB.

Both RANs are operated by a single operator responsible for attributing or freeing elementary bands to each RAN. On the one hand, revenue is assumed to be proportional to the satisfaction of the users. On the other hand, it is supposed that spectrum cost follows the law of supply and demand: as free spectrum diminishes, spectrum cost increases (section III-B). We are interested in the optimal policy that assigns bandwidth to the RANs.
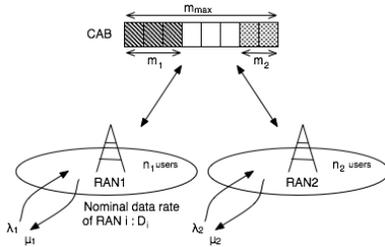

Fig. 1. System model: two RANs access to a CAB according to their needs.

Our model could be coherent with SOFDMA (Scalable Orthogonal Frequency Division Multiple Access) cellular networks (i.e. WiMAX, 3GPP-LTE), where the bandwidth of the system is scalable [13]. In these systems the operator has indeed an additional flexibility in resource allocation through the possibility of scaling the bandwidth.

B. Traffic

We consider a bursty packet traffic, such as web browsing or file downloading on the downlink: a user alternates between packet calls (several packets are transferred in a very short time) and reading times (there is no transfer). In this paper, we focus on the packet call level and so we neglect the details of the packet level.

We assume Poisson arrivals of user downlink packet calls with rate λ_1 in RAN1 and λ_2 in RAN2. Traffic is supposed to

be elastic: the packet call size is exponentially distributed with mean X_{ON} bits in both RANs and so the service rate depends on the available RAN throughput. We assume a fair share of resources between users of a given RAN. For RAN i let D_i be the cell data rate (in bits/s) accessible with an elementary spectrum band. Then, the service rates can be written as:

$$\mu_i = \frac{m_i D_i}{X_{ON}}.$$

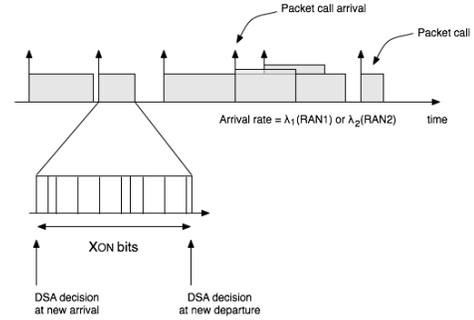An illustration of the traffic model is shown in Fig. 2.

Fig. 2. Assumed traffic model.

C. Dynamic spectrum access

In the considered system model, the core issue for the operator of the RANs lies in the trade-off to be found between spectrum cost and revenues obtained from users: more spectrum per RAN means a higher lease cost for the operator but also higher throughputs for users that are encouraged to pay more for the service. As the CAB size is limited and as spectrum cost increases with increasing demand, there is a strong interaction between RANs.

In this paper, a DSA policy is a strategy that dynamically attributes spectrum bands to each RAN from the CAB. We assume that a DSA decision is taken at each new event, i.e., a new packet call arrival or a packet call departure in any RAN (see Fig. 2). A DSA decision is supposed to increase the number of spectrum bands for a RAN by one block, to decrease by one block this number, or to keep constant the spectrum of a RAN. We assume that at least one spectrum block is always available to each RAN, so that starvation is not possible. We are now interested in the optimal DSA policies in terms of operator revenue.

III. SEMI-MARKOV DECISION PROCESS

In order to achieve this goal, we rely on the SMDP framework. We first define the SMDP and the reward function, then use uniformization to obtain an MDP and rely on the policy iteration algorithm to find the optimal DSA policy.

A. State space

The system state is given by all four-tuple (n_1, m_1, n_2, m_2) with constraints $n_1 \leq n_1^{max}$, $n_2 \leq n_2^{max}$ and $m_1 + m_2 \leq m_{max}$. The limitation imposed to the number of active users is equivalent to setting a minimum throughput per RAN. Let S be the state space.

B. Reward function

The reward function is based on the revenue expected by the operator. The higher the satisfaction of users, the higher the operator revenue; the higher the amount of bandwidth leased by RAN, the higher the cost to lease this spectrum band. We define a comfort service rate μ_{com}. The revenue obtained from a given customer in RAN i increases with its satisfaction:

$$\phi_i(n_i, m_i) = K_u(1 - \exp(-\mu_i/n_i\mu_{com})),$$

where K_u is a constant in euros per unit of satisfaction. Satisfaction, defined in [14], is an increasing function of the user data rate and is without unit. Thus the total revenue obtained by the operator in state $s = (n_1, m_1, n_2, m_2)$ is

$$g_1(s) = n_1\phi_1(n_1, m_1) + n_2\phi_2(n_2, m_2).$$

We assume that the spectrum price is increasing when the amount of free spectrum decreases and we define it as:

$$g_2(s) = K_B(m_1 + m_2) \exp\left(-\frac{m_{max} - m_1 - m_2}{m_{com}}\right),$$

where m_{com} is a constant that controls the variation of the price and K_B is a constant in euros per MHz (it is the equivalent spectrum price per cell). If m_{com} is high, the exponential function is close to 1 whatever the state. If m_{com} is small, there is a high discount when the CAB is free. Note that the price paid by the operator for a given elementary band varies with the occupation of the CAB. The global reward function per time unit can thus be written in state s:

$$g(s) = g_1(s) - g_2(s).$$

C. Action space

In each state, the operator is allowed to increase, decrease or leave unchanged the spectrum of each RAN. As shown in Fig. 2, a decision epoch occurs at each packet call arrival, or departure. As state transitions occur only at the arrival or the departure of a single user, we assume that the band assigned to a single RAN can be increased or decreased by a single elementary band. This leads to nine possible actions of the form $a = (a_1, a_2)$, $a_i \in \{0, -1, +1\}$ given in Table. I.

TABLE I
LIST OF POSSIBLE ACTIONS

Action	a vector	action index
Band1 constant and Band2 constant	$(0, 0)$	1
Band1 constant and Band2 increases	$(0, +1)$	2
Band1 constant and Band2 decreases	$(0, -1)$	3
Band1 increases and Band2 constant	$(+1, 0)$	4
Band1 increases and Band2 increases	$(+1, +1)$	5
Band1 increases and Band2 decreases	$(+1, -1)$	6
Band1 decreases and Band2 constant	$(-1, 0)$	7
Band1 decreases and Band2 increases	$(-1, +1)$	8
Band1 decreases and Band2 decreases	$(-1, -1)$	9

The effective action space depends on the state. If $m_i = 1$ the spectrum band of RAN i cannot decrease. If the CAB is blocked, i.e., if $m_1 + m_2 = m_{max}$, no band can increase.

D. Transition probabilities

Let $p_{s,s'}(a)$ be the probability that at the next decision epoch, the system will be in state $s' = (n_1', m_1', n_2', m_2')$ if a is chosen in state $s = (n_1, m_1, n_2, m_2)$. Let $1/\nu_s(a)$ be the expected time until next decision epoch if action a is chosen in state s:

$$\nu_s(a) = \mathbb{1}_{\{n_1 < n_1^{max}\}}\lambda_1 + \mathbb{1}_{\{n_2 < n_2^{max}\}}\lambda_2$$
$$+ \mathbb{1}_{\{n_1 > 0\}}\mu_1 + \mathbb{1}_{\{n_2 > 0\}}\mu_2.$$

Transition probabilities are given by:

$$p_{s,s'}(a) = \lambda_i/\nu_s(a) \text{ if } (n_i' = n_i + 1)$$
$$\text{and } (\forall j \ m_j' = m_j + a_j),$$
$$\mu_i/\nu_s(a) \text{ if } (n_i' = n_i - 1)$$
$$\text{and } (\forall j \ m_j' = m_j + a_j).$$

E. Uniformization

A step of uniformization is now needed in order to transform the continuous time Markov chain into an equivalent discrete time Markov chain. This is done by choosing a sufficiently small transition step $1/\nu$ ($\forall s, a, \ \nu_s(a) \leq \nu$) and allowing self transitions from a state to itself.

Transition probabilities are modified in the following way:

$$\tilde{p}_{s,s'}(a) = p_{s,s'}(a)\nu_s(a)/\nu \text{ for } s \neq s',$$
$$\tilde{p}_{s,s}(a) = 1 - \sum_{s' \neq s} \tilde{p}_{s,s'}(a) \text{ otherwise.}$$

A DSA policy R associates to each system state s, an action $R(s)$ from the action space of s.

F. Policy iteration

We are interested in finding the optimal policy R^* of the continuous-time average cost problem described above. For that, we apply the policy iteration algorithm to the auxiliary discrete-time average cost problem obtained after uniformization (see [15], vol.2, p.315). The iterative algorithm is now succinctly described.

- **Step 0 (initialization)**: Let R be an arbitrary stationary policy.
- **Step 1 (value-determination)**: For the current policy R, we solve the system of linear equations whose unknowns are the variables $\{J_R, h_R(s)\}$: $h_R(1) = 0$ and

$$h_R(s) = g(s) - J_R + \sum_{s' \in S} \tilde{p}_{s,s'}(R(s))h_R(s').$$

- **Step 2 (policy improvement)**: For each $s \in S$, we find:

$$R'(s) = arg \max_{a \in A(s)} \left\{ g(s) - J_R + \sum_{s' \in S} \tilde{p}_{s,s'}(a)h_R(s') \right\}.$$

- **Step 3 (convergence test)**: if $R' = R$, the algorithm is stopped, otherwise, we go to step 1 with $R := R'$.

The SMDP approach has the advantage of providing optimal policies and an upper bound on the achievable reward. The policy iteration algorithm takes into account not only RANs loads, the number of active users and RANs interactions but

also the whole dynamics of the system. Optimal policies are thus strongly dependent on the system parameters and simple examples cannot be easily generalized when the number of system states increases. In the next section, we propose a sub-optimal DSA heuristic that overcomes these limitations for an operator, while still providing a high reward.

IV. DSA HEURISTIC

A. DSA policies implementation

In order to implement optimal policies, an operator would have to run the policy iteration algorithm for all possible system parameter sets and store results to be dynamically used according to the context. Running policy iteration on a real-time basis is indeed unthinkable, especially when the number of system states increases (for example if many cells or users are considered). The proposed DSA heuristic intends to ease DSA implementation for an operator. With this heuristic, massive storage of data is not needed and computations can be done on the fly.

B. Proposed DSA heuristic

In order to obtain a simple heuristic, each of the RANs can be considered as M/M/1/n_i^{max} system. Each has a service rate μ_i. We then fix the allocated bandwidth to both RANs for a given load: (m_1, m_2) is thus independent from (n_1, n_2).

With these assumptions, the average heuristic reward for the operator, g_H, can be easily computed for all possible combinations of allocated bands (m_1, m_2), along with the corresponding λ_i values. The average reward is the sum of the rewards obtained from the two RANs. For a given $(\lambda_1, \lambda_2, m_1, m_2)$:

$$
g_H(\lambda_1, \lambda_2, m_1, m_2) = \sum_{i=1}^{2} \sum_{n_i=0}^{n_i^{max}} \pi_{n_i}(\lambda_i) n_i \phi_i(n_i, m_i) \\
- g_2(n_1, m_1, n_2, m_2), \quad (1)
$$

where the $\pi_{n_i}(\lambda_i)$, $i \in \{1, 2\}$, $n_i \in \{0, ..., n_i^{max}\}$ are the steady state probabilities of a M/M/1/n_i^{max} with arrival rate λ_i and service rate μ_i. We use this result for the proposed DSA heuristic:

1) Estimate arrival rates λ_1 and λ_2;
2) Consider all possible tuples (m_1, m_2);
3) Compute for each tuple (m_1, m_2) the average reward according to Eq. 1;
4) Allocate bandwidth according to the tuple (m_1, m_2) that maximizes the average reward g_H.

The estimation of the λ_i is not in the scope of this paper, we will thus assume that they are known. Eq. 1 can be instantaneously computed for realistic values of the n_i^{max} and can be easily extended to several cells.

V. RESULTS

In this section, we compare the results obtained with optimal DSA policies, the proposed DSA heuristic and FSA in terms of operator reward, CAB utilization, and average user throughput.

A. Parameters

The CAB is assumed to have a size of 6 MHz, the elementary band ($m_i = 1$) has a size of 1 MHz, and m_{com} = 4 MHz. For the sake of simplicity, we assume that both RANs have the same characteristics: the average cell data rates D_i are considered to be 1250 Kbps, $X_{ON} = 3$ Mbits, $\lambda_1 = \lambda_2 = \lambda$, and $n_1^{max} = n_2^{max} = 8$. The pricing constants are fixed as follows: $K_u = 100$ euros, $K_B = 1$ euro, and $\mu_{com} = 0.167$ s^{-1}(which corresponds to a comfort throughput of 500 Kbps).

B. Arrival rate thresholds

For the considered parameter set, Fig. 3 shows the average reward g_H as a function of the arrival rate λ for different combinations of the allocated bands.
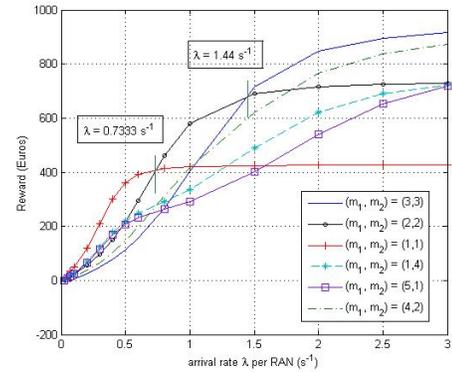

Fig. 3. Operator reward obtained for different allocated bands combinations.

We can notice that the (m_1, m_2) values that give the maximum reward are: (1,1), (2,2), and (3,3) depending on the arrival rate λ. The maximum reward can then be obtained by dynamically allocating symmetric numbers of elementary bands to the RANs according to the cell load. Threshold values for λ are given on Fig. 3.

C. Operator reward, CAB utilization, user throughput

Fig. 4 compares operator rewards obtained respectively with optimal DSA policy, the proposed heuristic and FSA. By definition, FSA allocates $m_i = 3$ elementary bands to each RAN whatever the system state. It can be seen that optimal policies provide significant increases of the reward for low to intermediate values of λ (for example +229% at $\lambda = 0.5$ s^{-1}). At high load, FSA and optimal DSA policy converge as expected. The proposed heuristic provides the optimal reward at low load and converges also to FSA at high load: only for intermediate values of λ, there is a small degradation of the reward (for example, -21% at $\lambda = 0.7$ s^{-1}).

These results can be explained by a better utilization of the sepctrum. CAB utilization is illustrated in Fig. 5 as a function of the arrival rate λ. Optimal DSA policy smoothly increases the CAB utilization as arrival rate increases. The proposed heuristic follows this trend with a step function. It is worth
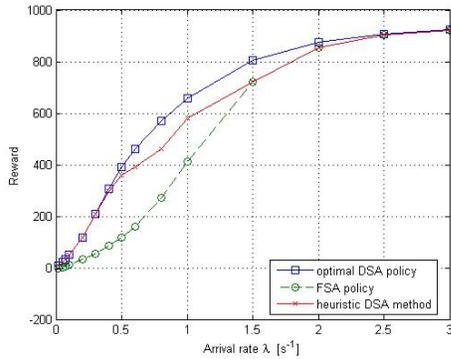
Fig. 4. Operator reward obtained with optimal DSA policy, heuristic DSA and FSA.

Fig. 6. Average user throughput with optimal DSA policy, heuristic DSA and FSA.

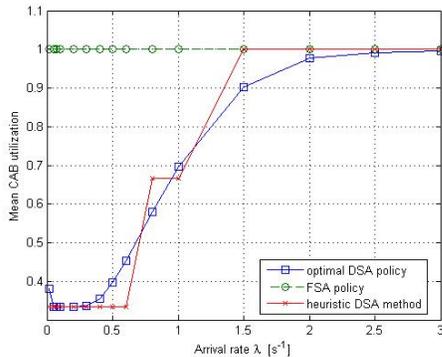mentioning the DSA gain in terms of spectral resource usage with respect to FSA.

Fig. 5. CAB utilization with optimal DSA policy, heuristic DSA and FSA.

Operator reward and better spectrum utilization are however obtained at the price of a degradation of the average user throughput. Fig. 6 illustrates the average user throughput as a function of the RANs load λ. Optimal DSA policy and proposed heuristic show again similar results. The achieved average user throughput with FSA is however much higher, especially at low loads. According to the traffic assumptions (see section II-B), a single user is indeed allowed to take advantage of the whole bandwidth allocated to a RAN. At low loads, FSA allocates 3 MHz to each RAN, while DSA methods allocates only 1 MHz leading to lower user throughputs.

VI. CONCLUSION

In this paper, we have studied DSA in cellular networks context. We have used the SMDP framework to derive optimal DSA policies in terms of the operator reward. We have proposed a simple heuristic DSA method to defeat the generalization difficulty of the optimal policies over realistic systems. The achieved reward using the heuristic DSA policy gives a very close reward to the optimal obtained by SMDP and thus significantly exceeds the reward obtained with FSA. Operator revenue increases but better spectrum utilization is obtained at the price of a user throughput degradation.
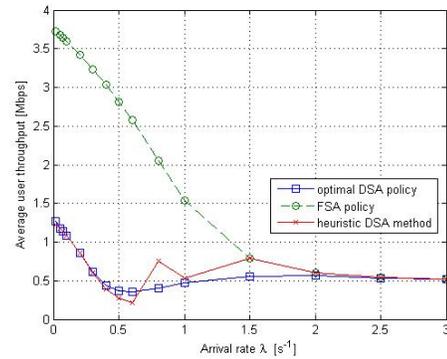
REFERENCES

[1] M. Buddhikot, "Understanding Dynamic Spectrum Allocation: Models, Taxonomy and Challenges," in *Proc. IEEE DySPAN'07*, pp. 649-663, 2007.
[2] S. Buljore et al., "IEEE P1900.4 Standard: Reconfiguration of Multi-Radio Systems," in *Proc. IEEE SIBIRCON'08*, pp. 413-417, 2008.
[3] S. Filin et al., "Dynamic Spectrum Assignment and Access Scenarios, System Architecture, Functional Architecture and Procedures for IEEE P1900.4 Management System," in *Proc. IEEE CrownCom'08*, pp. 1-7, 2008.
[4] M. Buddhikot, P. Kolodzy, K. Ryan, J. Evans, and S. Miller, "DIM-SUMNet: New Directions in Wireless Networking Using Coordinated Dynamic Spectrum Access," in *Proc. IEEE WoWMoM'05*, pp. 78-85, 2005.
[5] D. Thilakawardana, K. Moessner, and R. Tafazolli, "Darwinian Approach for Dynamic Spectrum Allocation in Next Generation Systems," *IET Communications*, vol. 2, no. 6, pp. 827-836, 2008.
[6] S. Sankaranarayanan, P. Papadimitratos, A. Mishra, S. Hershey, "A Bandwidth Sharing Approach to Improve Licensed Spectrum Utilization," in *Proc. DySPAN'05*, pp. 279-288, 2005.
[7] J.M. Chapin and W.H. Lehr, "Cognitive Radios for Dynamic Spectrum Access-The Path to Market Success for Dynamic Spectrum Access Technology," *IEEE Communications Magazine*, vol 45, no. 5, pp. 96-103, 2007.
[8] J. Acharya, and R.D. Yates, "A Price Based Dynamic Spectrum Allocation Scheme," in *Proc. ACSSC'07*, pp. 797-801, 2007.
[9] L. Vanbien, L. Yuewei, W. Xiaomeng, F. Zhiyong, and Z. Ping, "A Cell Based Dynamic Spectrum Management Scheme with Interference Mitigation for Cognitive Networks," in *Proc. VTC'08*, pp. 1594-1598, 2008.
[10] S. Geirhofer, L. Tong, and B.M. Sadler, "Cognitive Medium Access: A Protocol for Enhancing Coexistence in WLAN Bands," in *Proc. IEEE GLOBECOM'07*, pp. 3558-3562, 2007.
[11] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *IEEE J. on Select. Areas in Commun.*, vol. 25, no. 3, pp. 589-600, 2007.
[12] M. Coupechoux, J.M. Kelif, and Ph. Godlewski, "SMDP Approach for JRRM Analysis in Heterogeneous Networks," in *Proc. European Wireless'08*, pp. 1-7, 2008.
[13] IEEE 802.16-2005 standard for local and metropolitain area networks.
[14] N. Enderlé and X. Lagrange, "User Satisfaction Models and Scheduling Algorithms for Packet-Switched Services in UMTS," in *Proc. VTC'03*, vol. 3, pp. 1704-1709, 2003.
[15] D. P. Bertsekas, "Dynamic Programming an Optimal Control," third edition, Athena Scientific, 2007.