Information Processing Techniques for Cognitive Base Transceiver Stations

Luca Bixio, Andrea F. Cattoni, Giacomo Oliveri, Marina Ottonello, Mirco Raffetto, Carlo S. Regazzoni
Department of Biophysical and Electronic Engineering, University of Genoa
Via Opera Pia 11a, 16145 Genoa, Italy
{bixio, cattoni, oliveri, marina, raffetto, carlo}@dibe.unige.it

***Abstract*—In the last decade portable communication devices have become extremely widespread, and the introduction of new wireless technologies is becoming more and more difficult since most of the spectrum is already licensed for some existing services. In the present paper, the joint utilization of smart antennas and cognitive radio is proposed as a possible solution for improving the exploitation of the spectrum resources. To this end a Cognitive Base Transceiver Station based on the above technologies is considered: the capability of such system to establish a communication with a set of mobile stations in a vehicular context will be described.**

I. INTRODUCTION

In the last decade, thanks also to the availability of low-size and low-cost portable terminals, wireless communications have become widespread. However, such a success is making radio resource allocation more and more complex, since almost all frequencies have already been licensed to some existing technologies.

In this context, a recent study carried out by the U.S. Federal Communication Commission (FCC) [1] has shown that despite the fact that almost all radio resources are licensed to some wireless service, the actual exploitation of the licensed spectrum ranges from 15% to 85%. Such an observation has motivated the research activity in the field of opportunistic wireless communications and, in general, of technologies able to improve spectrum utilization.

In order to increase the frequency re-use in wireless cellular networks, the simplest solution is to reduce the dimension of each elementary cell. Such a solution, however, can increase the handover rate and therefore it can reduce the network efficiency, especially in the case of fast mobile (i.e. vehicular) terminals.

When wireless high-mobility terminals are of interest, a possible solution could be based on the utilization of technologies that enable a dynamic and flexible spectrum allocation such as smart antennas [2] and cognitive radio [3]. In this paper, a possible application of such technologies in the context of high-mobility wireless communications is presented. The resulting application, which will be called Cognitive Base Transceiver Station (CBTS), is an intelligent cellular node able to interact with both "standard" and "cognitive" terminals by using smart antennas driven by a cognitive radio software. As a consequence, the CBTS can increase the utilization of the radio spectrum through the joint exploitation of spatial diversity (with low inter-cell handover rate) and of flexible resource allocation allowed by cognitive radio.

In the present paper a description of the information processing techniques necessary for the considered application will be presented from a system-level point of view; moreover, some details regarding reinforcement learning algorithms applied to smart antenna management will be provided and other details on mode identification and spectrum monitoring (MISM) techniques will be discussed. Finally, some preliminary simulation results in order to show the capabilities of the considered approach.

II. COGNITIVE RADIO TECHNOLOGY

Starting from the definition of the cognitive radio paradigm [4], a lot of different approaches have been developed. Recently, Simon Haykin in [5] has stated that cognitive radio is essentially a new emerging approach to wireless communications. Among the proposed applications of such paradigm, one of the most successful is based on the objective of improving the efficiency of the spectrum usage through a dynamic allocation of the available frequencies: this in practice represents the "opportunistic radio" application of the cognitive radio paradigm.

From a practical point of view, in order to reach the considered goal with a cognitive radio terminal, a high level of reconfigurability and flexibility is required [5]. For these reasons the cognitive radio architecture can be thought as an extension of the software defined radio platform, which can be considered as a wireless device characterized by a high level of (software) flexibility in its communication functions (e.g. regarding the modulation, bit rate, code rate etc.) [5]. Although "for reconfigurability, a cognitive radio looks naturally to software-defined radio to perform this task" [5], it has been recently stated that one of the most appreciated capabilities of a cognitive radio system concerns its "cognition" abilities [5]. In this context, recently emerged visions of "intelligence" in artificial systems, which take inspiration from neuro-scientific [6] and robotics [7] studies, could be of great interest. In particular these approaches assure more flexibility than classical rule-based approaches to overcome the unforeseen or emergency situations. Such a property could be of fundamental importance for the proposed application, since its flexibility could allow an enhanced management of wireless terminals in mobile applications.

978-1-4244-1653-0/08/$25.00 © 2008 IEEE

As a consequence, the considered application, a Cognitive Base Transceiver Station, will be developed on the basis of the above described framework, in particular by following an agent-based bio inspired approach. To this end, a reinforcement learning approach will be used for the decision phase, while a support vector machine technique will be used for the mode identification and spectrum monitoring agent.

III. Cognitive Base Transceiver Architecture

In the considered approach the proposed CBTS has to manage the communications with a set of moving wireless terminals exploiting different communication modalities. The framework of the CBTS is designed to guarantee an high level of flexibility during the terminal management; in order to assure this level of adaptivity an *agent-based* approach has been chosen. Therefore the proposed approach can be described as the result of the interaction, cooperation and competition of a set of autonomous agents. In Figure 1 the general framework for the proposed system is shown.

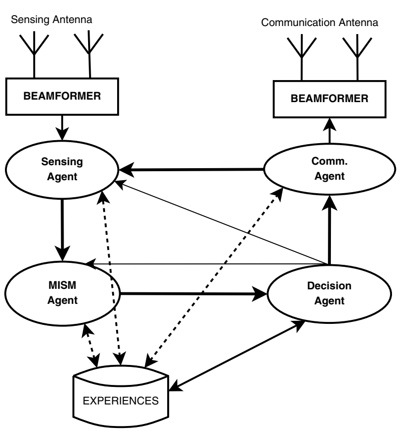


Fig. 1. General framework for the proposed system

The cognitive base station has to perform an accurate sensing of the surrounding environment to localize and track the users belonging to the domain of interest. The sensing agent performs this operation by using a smart antenna system which can be considered as the perceptual organ of the cognitive system.

The MISM agent receives the information collected by the sensing agent and performs the identification of the transmission modality or communication standard employed by each user.

The MISM agent provides the decision agent with this information, the decision agent uses this information and the stored experience to modify the system behaviour in order to reach the prearranged goals (e.g to maximize the number of managed users or to minimize the power consumption). In the proposed approach the decision agent has to reconfigure the communication process but it can be designed to reconfigure the entire system. In a similar way the memory used to store all the relevant experiences in order to perform a continuous learning phase can be used by all the agents (not only by the decision agent) as shown in Figure 1.

At the next stage the information originated by the decision agent are used by the communication agent to reconfigure the smart antenna system in order to guarantee reliable communications with users equipped with different standards. In the proposed system the communication agent has to apply the most appropriate multi-lobe steering strategy for the current context, configuring in a suitable way the communication beamformer.

The agents represented in Figure 1 reflect the so called *cognitive cycle* [4]. It allows to describe the interactions of the cognitive system with the external world as a cycle composed by four main processes: sensing, analysis (MISM), decision and action. This model, used to describe the behaviour of living beings, is executed in a continuous way to dynamically improve the behaviour of the proposed cognitive system: it can be considered as a continuous learning phase.

In the proposed approach the attention is focused on two specific aspects of the cycle: in particular the MISM and the decision phase are considered. It is important however to remark that the development of proposed system is in the early stages: all functions of the CBTS are simplified, in order to test the correctness of the approach.

IV. MISM Agent

The MISM agent is dedicated to identify the transmission modality of each user in order to re-configure the software defined radio device. Let us analyze in detail the role of the MISM agent in the cognitive cycle. (Figure 2).

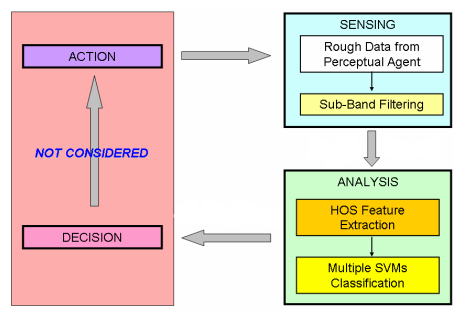


Fig. 2. The role of the MISM agent in the Cognitive Cycle

During sensing phase rough data are acquired: the sensing agent gathers information about the spectrum and hence MISM agent performs a classification in each direction of interest. The sensing agent sends to the MISM agent rough signals partially filtered to reduce the noise. The MISM agent translates the incoming information into a meta-object, enriching the physical signals with all the information needed to describe the context for planning the proper actions. This phase is composed by a feature extractor based on higher order statistics (HOS) of frequency distributions and a bank of multiple support vector machine (SVM) classifiers [8].

A. Support Vector Machine Classification

SVMs are very powerful tools for classification. They are based on the computation of the optimal hyperplane which separates different classes of training samples in the feature space.

In order to improve the performances of a SVM classifier, it is possible to adopt strategies like bagging or boosting [9]. In particular, in the proposed method the former one is adopted. It consists in a training set $\boldsymbol{X}$ divided into subsets $\boldsymbol{X}_j$, where $\boldsymbol{X} = \bigcup_{j=1}^{J} \boldsymbol{X}_j$ and $\boldsymbol{X}_i \cap \boldsymbol{X}_j = \oslash$ for $i \neq j$. A pool of J one-against-all SVM sets is trained, each one on a different subset, independently. When a new pattern is presented to the system, it is classified by the entire pool of machines and the final classification is taken accordingly to a majority voting algorithm. Each set is composed by M SVMs, one for each mode supported by the CBTS. The input vector $\boldsymbol{x}$ contains a triplet of features $\boldsymbol{x} = \{\overline{f}_S, \sigma_{f_S}, \gamma_{2_S}\}$, i.e. the instantaneous (mean) frequency, its standard deviation, and the kurtosis. The results of the classification procedure are represented by labels that indicate the possible modalities $\boldsymbol{M} = \{M_1, \cdots, M_M\}$. The channel labeling is not considered here, because it can be extracted by applying a-priori information about the classified mode $\boldsymbol{y}$ and the bandwidth occupancy obtained by the sensing agent.

V. Reinforcement Learning Approach for the Decision Agent

The decision agent represents the most important phase of the cognitive cycle as it is responsible for providing to the communication agent a suitable configuration in order to establish and to maintain a reliable communication with the users. In fact the link quality depends on the abilities of the communication agent to modify the antenna radiation pattern at the correct time instant, if possible foreseeing users' movement. To this end the decision agent exploits the output of the previous phases of the cognitive cycle and the acquired experience. In the proposed framework the decision agent is based on a reinforcement learning algorithm that allows to store experience, to exploit it when needed, and to provide a suitable configuration for the communication agent [10].

The chosen reinforcement learning approach is inspired by Damasio's cognition theory, which is based on the fact that biological entities distinguish themselves from the outside world [11]. On the basis of this concept a decision agent has been designed. It stores the global system state into two different sub-states: the internal state, called "proto self status" [11], and the perceived outside world state, called "core self status" [11]. Therefore the task of the decision agent is to exploit the information about the perceived external state obtained from the analysis phase in order to provide a new proto self status using the acquired experience.

Let us denote the vector containing the core self status with $\mathbf{x}_c$ and the vector containing the proto self status with $\mathbf{x}_p$ (it represents the new system's internal state). Furthermore, let us assume that the system interacts with the environment at discrete time steps $t = t_0,\ t_1, \ldots$. The decision agent performs a "trial and error" interaction with the environment in order to obtain the highest reward possible. In particular, at a given time instant t_i, the decision agent has to select an action $\mathbf{x}_p(t_i)$ (among all P possible actions) that leads to the highest "reward" (or reinforcement) $r(t_{i+1})$ [10]. The proposed approach for the decision agent obviously depends on the stored experience that allows the system to acquire the capability of predicting the rewards sent by the environment, learning the best configuration for different situations. Among the proposed reinforcement learning algorithms [10], different choices can be made regarding, as an example, experience exploitation and acquisition [10]. In the considered framework it has been choosen the so-called *Q*-learning [10] algorithm whose application can be described as follows.

Let us denote the decision policy with $\pi_t(\mathbf{x}_c, \mathbf{x}_p)$ [10] which represents the probability to select $\mathbf{x}_p(t) = \mathbf{x}_p$ and $\mathbf{x}_c(t) = \mathbf{x}_c$. Moreover we denote the *Q*-function with $Q^{\pi}(\mathbf{x}_c, \mathbf{x}_p)$ [10] which represents the expectation of the reward given certain decision in a given state for a given decision policy. In this approach learning means [10] finding the optimal decision policy, i.e.

$$Q^{\pi^*}(\mathbf{x}_c, \mathbf{x}_p) = \max_{\pi} Q^{\pi}(\mathbf{x}_c, \mathbf{x}_p),\ \forall \mathbf{x}_c,\ \mathbf{x}_p \tag{1}$$

In the proposed reinforcement learning strategy the decision agent has to learn an estimation of the *Q*-function, collecting the necessary experience. In order to perform *Q*-function estimation a triplet composed by the environment status vector $\mathbf{x}_c$, the action vector $\mathbf{x}_p$, and the associated estimate $Q^{\pi}(\mathbf{x}_c, \mathbf{x}_p)$ is stored as entry of a table. Therefore, each time the state $\mathbf{x}_c$ is encountered and a decision $\mathbf{x}_p$ is taken, the estimate of the *Q*-function is updated according to [10]:

$$Q^{\pi}_{k+1}(\mathbf{x}_c, \mathbf{x}_p) = \alpha r_k + (1-\alpha) Q^{\pi}_{k}(\mathbf{x}_c, \mathbf{x}_p) \tag{2}$$

where α is a design parameter. At every step, table entries are updated and the experience is stored and ready to be used. The trade-off needed between exploration and exploitation of the stored experience is obtained using an epsilon-greedy method [10]: with probability $(1-\varepsilon)$ the decision agent exploits experience; with probability ε the decision agent explores a new $\mathbf{x}_p$ choosing a random decision.

It is important to remark that the learning algorithm performances depend on design parameters α and ε and that the number of the table entries can greatly influence the speed of convergence.

VI. Simulation and Results

In order to verify the capabilities of the proposed approach, two software simulators have been developed. In particular, the two software tools are in charge of the simulation of different parts of the overall systems: this choice has been carried out in order to allow detailed analyses of the different phases of the cognitive entity in an independent way. The two software tools are respectively designed in order to

- simulate the learning phase of the smart antenna system (whose objective is to establish and keep a connection with a set of moving terminals) driven by the cognitive radio
- simulate the analysis phase of a cognitive radio whose task in this case is to distinguish between different communication modalities

Due to the preliminary nature of the reported results, both the considered software tools employ some simplified assumptions, which will be clarified next.

Firstly, let us consider the software for simulation of the learning phase of the smart antenna management, which is based on reinforcement learning techniques. In the considered implementation of the such a phase, single-carrier single-modulation terminals at 1.8 GHz are considered. The antenna is composed by a linear equispaced array of 21 dipoles each one 6 cm long, while the spacing between adjacent elements is equal to 5 cm. The system domain of interest is represented by a 300 m long section of a street. The antenna of the base station is placed 25 m far from the street center. Flat free-space loss channel is considered for the simulations. Different wireless terminals can move in the domain of interest at a maximum velocity of 12 m/s. The system objective is to establish and keep a communication with these radio terminals by continuously adapting the antenna beamforming parameters and transmitted power.

The detection of the environment state is performed by means of signal processing on the space-frequency images acquired by the sensing phase. For memory reasons, the state-action pairs are not memorized in a single tabular form but in several different (dynamically updated) memories which employ linear interpolation in order to generalize the available knowledge.

Different rewards are considered for the system in order to observe different possible evolutions. As a first example, the system has been simulated by defining a feedback reward at time t_j as

$$r(t_j) = \sum_{i=1}^{N_{conn}(t_j)} \frac{\mathrm{SNR}_i(t_j)}{\mathrm{SNR}_{\max}} \qquad (3)$$

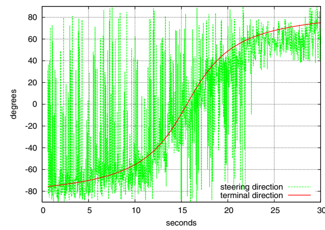
where N_{conn} is the number of established connections, SNR_i is the signal-to-noise ratio for the i-th connection and $\mathrm{SNR}_{\max}$ is a saturation parameter. Since r does not penalize the used power, it is expected that a sufficiently experienced system will try to exploit all the available power while minimizing the steering error for the communications toward the detected terminals.

The simulation results for this case are reported in Fig. 3. In the reported example, there is at most one moving terminal in the domain of interest. It is shown that the considered system is effectively able to learn that the strategy described above.

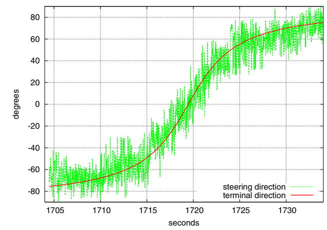
As a second example, the system has been simulated by defining

$$r(t_j) = \sum_{i=1}^{N_{conn}(t_j)} \frac{\mathrm{SNR}_i(t_j)}{\mathrm{SNR}_{\max}} - \sum_{i=1}^{N_{com}(t_j)} \frac{p_i(t_j)}{p_{\max}} \qquad (4)$$

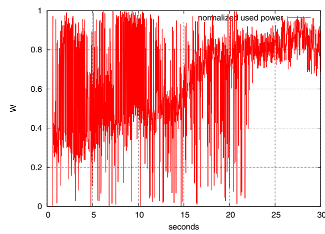
where N_{com} is the number of attempted communications. The learning and optimization procedure in this case is more complex since contrasting goals (minimize used power and maximize received signal-to-noise ratio) are imposed. In this case, r penalizes the used power, and therefore the system is expected to reach a balanced strategy between minimization



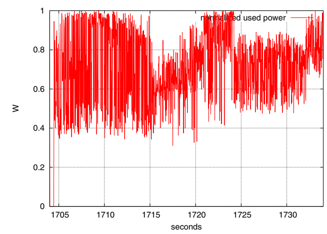
(a) Steering direction at the startup.



(b) Steering direction after 1 hour.



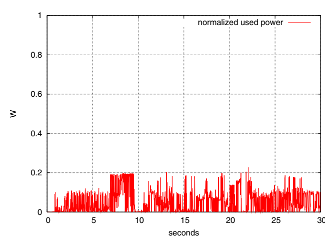
(c) Normalized power at the startup.



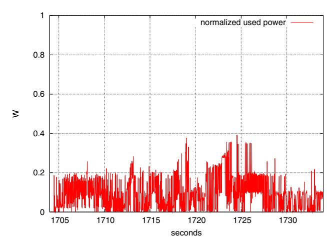
(d) Normalized power after 1 hour.

Fig. 3. System performances during the learning phase with r as in (3). Single moving terminal is considered.

of used power and maximization of received signal-to-noise ratio.



(a) At the startup.



(b) After 1 hour.

Fig. 4. System used power during the learning phase with r as in (4). Single moving terminal is considered.

The obtained results regarding the used power in this case are reported in Fig. 4 (the steering performances are comparable those of the previous case and thus are not reported). It is possible to see that the system reaches a balanced operation point in which the steering error is minimized while a smaller (with respect of the previous case) amount of power is exploited. The outcome of these simulations is that, although in a simplified scenario, the reinforcement learning approach is able to learn the most effective steering strategy.

Let us now consider the simulation of the mode identification and spectrum monitoring function of the Cognitive Base Transceiver Station. In this case, we have considered the case of a single transmitting terminal and a single (passive) receiving antenna belonging to the base station. The task of the mode identification function is to state if the perceived signal contains a certain transmission modality.

The MISM agent has been tested in a simplified communication situation with only two possible communication modalities, i.e. IEEE 802.11b and IEEE 802.11g. A single channel for each standard has been considered. A particular characteristic is that, in both cases, channels have the same central carrier frequency, even if they can be not superimposed. The simulations have been performed in order to test the system in different conditions in terms of signal-to-noise ratio. All the simulations have been performed on an intermediate frequency signal sampled at 1 Gs/s.

The objective of the classifier based on support-vector machines is to distinguish between three classes of signals: noise, IEEE 802.11b and IEEE 802.11g. To perform this task, the features considered are the frequency moments (up to the fourth order) of the incoming signals. The results regarding the classification accuracy which have been obtained for moderate signal-to-noise ratios (SNR) (from 10 to 20 dB) are shown in Fig. 5(a).

A. MISM Agent: Feature Space

The proposed method generates a tri-dimensional feature space. In Figure 5(b) the distribution of features, for the three classes, in the best case, i.e. SNR is equal to 20 dB, 5 paths and 16384 samples, is presented. Two out of three features are normalized with respect to the number of Fast Fourier Transform (FFT) samples in order to be comparable for all the windows.

The noise class is evidently well separated from the others, and this behavior can be noticed for all the simulated parameters. Let us focus the attention on the remaining two classes.

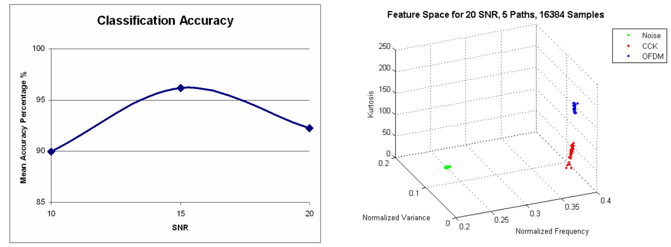
B. MISM Agent: Classification Performances

All the trained SVMs are characterized by radial basis function kernels. Different pools with different numbers of SVMs, from only one to 21 machines, have been trained in order to verify if increasing the number of multiple classifiers can lead to improvements in the classification performances. Training has been made by keeping constant the length of the radio signal. This fact implies that widening the length of the time window, the number of training samples has been reduced.

Before training the machines, all the training samples have been randomly shuffled and hence distributed to each machine of the pool accordingly to what depicted in Section IV-A. This choice has been made in order to decrease the overall complexity of the system. In fact, as an example, it could be possible to train different pools for each value of SNR, but a SNR estimator should be required, increasing the computational load of the agent.

VII. Conclusions

In the current paper an approach for the improvement of the efficiency of the radio resource usage based on the joint exploitation of smart antennas and cognitive radio software has been presented. The information processing techniques for the management of the introduced Cognitive Base Transceiver Stations have been presented, and the first preliminary numerical results of the performances of such application have been presented. In particular, it has been shown that it is possible to learn the optimal steering directions for a smart antenna



(a) Mean classification accuracy obtained by the MISM agent.

(b) Feature Space for 20 SNR, 5channel paths, 16384 FFT samples.

Fig. 5. Performances of the MISM agent.

through the exploitation of reinforcement learning techniques, and that the proposed classification approach can be a starting point for the development of the context analysis system of the base station.

Although the reported results represent only a preliminary attempt to state the effectiveness of the considered approach in simplified hypotheses, the obtained performances show that the considered framework could provide a sufficiently general approach for the design of integrated wireless base stations based on bio-inspired cognitive paradigms.

Acknowledgment

This work has been developed under the framework of the *PRIN-SMART project*, financed by the Italian Ministry of the University and Research (MUR).

References

[1] "Spectrum policy task force report," tech. rep., Federal Communication Commission, 2002.

[2] M. Chryssomallis, "Smart antennas," *IEEE Antennas and Propagation Magazine*, vol. 42, pp. 129–136, 2000.

[3] J. Mitola, *Software Radio Architecture: Object-Oriented Approaches to Wireless Systems Engineering*. John Wiley and Sons, New York, NY, USA, 2000.

[4] J. Mitola, "Cognitive radio: making software radio more personal," *IEEE Pers. Comm.*, vol. 6, pp. 48–52, August 1999.

[5] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 2, pp. 201–220, 2005.

[6] R. Llinas, *I of the Vortex*. Bradford Book, MIT Press, Cambridge, MA, 2001.

[7] L. Steels and R. Brooks, *The Artificial Life Route to Artificial Intelligence: Building Embodied Situated Agents*. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, 1995.

[8] A. F. Cattoni, M. Ottonello, M. Raffetto, and C. S. Regazzoni, "Hos-based mode classification for infomobility framework," in *ICST First International Workshop on Cognitive Wireless Networks*, (Vancouver, British Columbia, Canada), August 14 2007.

[9] D. Opitz and R. Maclin, "Popular ensemble methods: An empirical study," *Journal of Artificial Intelligence Research*, vol. 11, pp. 169–198, 1999.

[10] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 1998.

[11] A. R. Damasio, *The Feeling of What Happens-Body, Emotion and the Making of Consciuousness*. Harvest Books, 2000.