COGNITIVE RESOURCE MANAGEMENT

For All Wireless Access Layers

Vuk Marojevic,
Ismael Gomez,
and Antoni Gelonch

The radio communications spectrum has been commonly considered a scarce resource. Measurements, however, revealed that spectral resources are actually underused most of the time. Therefore, advanced radio resource management (RRM) techniques try to balance the spectrum occupation. Increasing the system capacity further requires the coordination of several factors in the wireless business and the convergence of standards and technologies. We identify the need for

- specifying and adopting radio access technologies (RATs) that enable a truly flexible spectrum allocation
- introducing a set of cognitive resource managers with autonomous or semiautonomous decision capabilities.

Radio operators wish to reuse the existing communications infrastructure for integrating new communications systems. Wireless subscribers appreciate the introduction of new services if these

Digital Object Identifier 10.1109/MVT.2012.2190190
Date of publication: 13 April 2012

BACKGROUND IMAGE:
©DIGITAL VISION

services are compatible with their mobile terminals. Designing multipurpose infrastructure and user devices for integrating existing and future radio systems and services is challenging though [1]. Multimode hardware designs may be economically feasible today, but their limited flexibility will soon become a bottleneck. Software solutions should then be considered.

Software radio or software-defined radio (SDR) [2] defines radio transceiver processing chains that are essentially implemented in software (SDR applications) running on general-purpose hardware (SDR platforms). SDR enables infrastructure or hardware upgrades and signal processing or software updates. The possibility of dynamically reconfiguring an SDR platform facilitates adapting the transmission mode to the given environmental conditions.

Advances in optical networking and hardware virtualization enable to connect a set of distributed antennas to a data center for real-time digital signal processing (DSP) [3]. Sharing resources from a virtual resource pool increases the system flexibility but requires sophisticated resource control and management. Therefore, we envisage radio operators renting computing resources from computing operators, which deploy, maintain, and manage the radio infrastructure and distributed computing sites [13]. Despite the split responsibilities, only a coordinated management and allocation of radio and computing resources can ensure smooth system operation and real-time service provisioning. The increasing system complexity is an opportunity for cognitive radio.

Cognitive radio [4] is commonly associated with dynamic spectrum access (DSA) [5], [6]. Marshall [10], however, views DSA as not the only application of cognitive radio. He considers that the radio operation will be dynamically determined as a function of the environmental perception. The European Telecommunications Standards Institute expects a gradual deployment of reconfigurable radio systems toward full SDR deployment, enabling resource sharing and real-time resource management [11]. The concept of opportunistic networking has been recently extended to opportunistic computing [12], where mobile nodes make use of the surrounding computing and communications resources. A mobile device [cell phone, personal digital assistant (PDA), or the like] may then be considered as a mobile computing resource or ad hoc communication relay rather than a sheer personal device.

We consider a cognitive radio system as an ambient-aware wireless communications system that intelligently adjusts itself to its operating environment. The system can dynamically modify the radio transmission modes or parameters as a function of resource constraints, end-to-end quality-of-service (QoS) requirements, business models, or the like. Therefore, this article proposes a cognitive radio system that embraces all wireless access layers. The system enables opportunistic resource allocations and radio against computing resource tradeoffs. Rather than guaranteeing radio and computing resources in advance, they will be dynamically allocated from shared resource pools. We first identify the resource environments before describing the management framework and methodology. Finally, simulation results illustrate the potential benefits of our approach.

QOS REQUIREMENTS AND TOLERABLE INTERFERENCE LEVELS LIMIT THE USE OF SPECTRAL RESOURCES, WHEREAS THE COMPUTING CAPABILITIES RESTRAIN THE EXECUTION OF RADIO AND USER APPLICATIONS.

Resource Environments

Establishing a wireless communications link requires coordinating several processes and allocating the necessary resources from different resource pools. Table 1 identifies five resource environments. Note that resource does not necessarily refer to a physical resource here; it may symbolize capability, capacity, or service.

The spectral resources, essentially, represent the usable radio frequency (RF) bands. Service resources refer to the available user and communications support services. The latter facilitate software downloads and DSP updates. Application resources embrace radio and user applications. A radio application resource stands for a software-defined DSP function, such as digital filter implementation, or a set of functions (SDR application) implementing part of a radio transmitter, receiver, or both. On the other hand, a user application defines the user interface for presenting data in a user-friendly format. The infrastructure resources provide an execution environment for radio and user applications. Finally, the hardware environment includes the RF circuitry, processors, and power supplies (Table 1).

The resource pools contain a wide set of heterogeneous resources that evolve over time: additional spectral resources for radio communications are being explored, radio standards and user services enhanced, signal processing algorithms updated, computing resource management techniques improved, and hardware platforms upgraded. User services require spectral resources for the over-the-air transmission as well as radio and user applications for the corresponding data processing. Applications rely on the infrastructure services that allocate the necessary hardware resources for their execution. Figure 1 indicates these relations and shows the scope of responsibilities of radio and computing operators.

TABLE 1 Environments, resources, and optimization goals.

Environment	Resources	Optimization
Spectrum	Spectrum holes Transmission power	Spectral efficiency Channel capacity Interference reduction Coverage
Service	Available user services: voice, video, streaming, and multimedia QoS levels Communications support services	Service versatility Different QoS levels Personalized services Mean opinion score
Application	Applications Applications' modules (functions)	Implementation efficiency Software portability
Infrastructure	Application deployment and management tools Hardware virtualization Middleware Infrastructure and network discovery	Flexibility Low complexity and resource overhead Real-time management Software portability
Hardware	Analog circuitry Digital processors Energy resources	Flexibility Low cost Reliability

The computing operator essentially manages the limited and distributed computing resources, whereas the radio operator may focus on managing transmission modes, frequency ranges, services, and QoS levels. The radio and computing operators intersect at the application layer. The radio operator specifies the radio-related parameters: tolerable interference level, QoS capability, and security protocols. The computing operator defines the application design methodology and identifies the applications' computing requirements. Thus, the radio operator specifies the necessary communications capabilities, whereas the computing operator deals with application design and deployment issues.

The system capacity is constrained by the limited amount of radio and computing resources. QoS requirements and tolerable interference levels, in particular, limit the use of spectral resources, whereas the computing capabilities restrain the execution of radio and user applications. If not properly managed, these resource limitations may lead to QoS degradations, service interruptions, or lack of service versatility. Although we may formulate individual optimization objectives for each resource type (Table 1, last column), the resource dependencies or correlations call for a joint resource management.

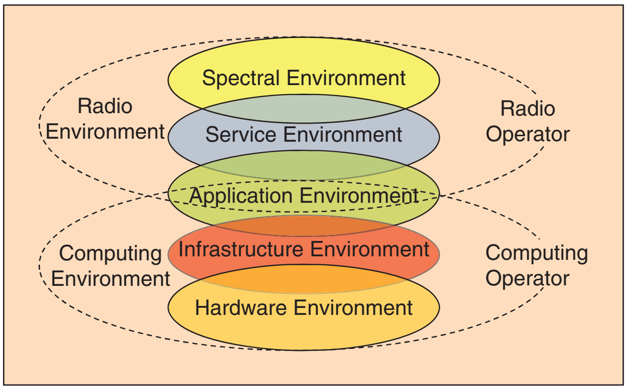

FIGURE 1 Resource relations.

Cognitive Resource Management

When Mitola introduced cognitive radio, he presented the cognition cycle [4] consisting of five fundamental phases. The observe, orient, plan, decide, and act phases facilitate processing environmental state information and deriving appropriate decisions. We suggest five interlaced cognitive cycles in Figure 2. (The five arcs from the decide block back to the environments were omitted for clarity.)

The system's current resource state information is collected (analyze) and combined with the acquired knowledge (learn) to derive the best possible overall decision (decide). These decisions may initiate software downloads and mobile terminal reconfigurations, QoS upgrades or downgrades, or switches to other transmission frequencies. Decisions may have instantaneous or long-term effects. Therefore, the system continuously evaluates its decisions, increasing or refining its knowledge database for effectively responding to repeating or new situations. It may predict system states and schedule actions for avoiding undesirable system situations to occur.

Figure 3 shows a layered base architecture and the five cognitive cycles surrounding it. The base architecture extends Mitola's software radio architecture [7]. It represents a general framework rather than a particular system implementation. The software radio layers correspond to the five resource environments of Figure 1. The sensing and monitoring interface scans the different software radio layers, captures the resource states, and forwards them to an intelligent subsystem. This process runs continuously with a periodicity that may vary from layer to layer.

The adaptive mechanisms block provides an information exchange interface between the intelligent subsystem and base architecture. This interface has two functionalities: 1) adapting the decisions of the intelligent subsystem to the particular SDR system and 2) triggering the corresponding management layers for realizing necessary actions. Both functionalities are important because they facilitate decoupling the intelligent subsystem from the specific operating environment.

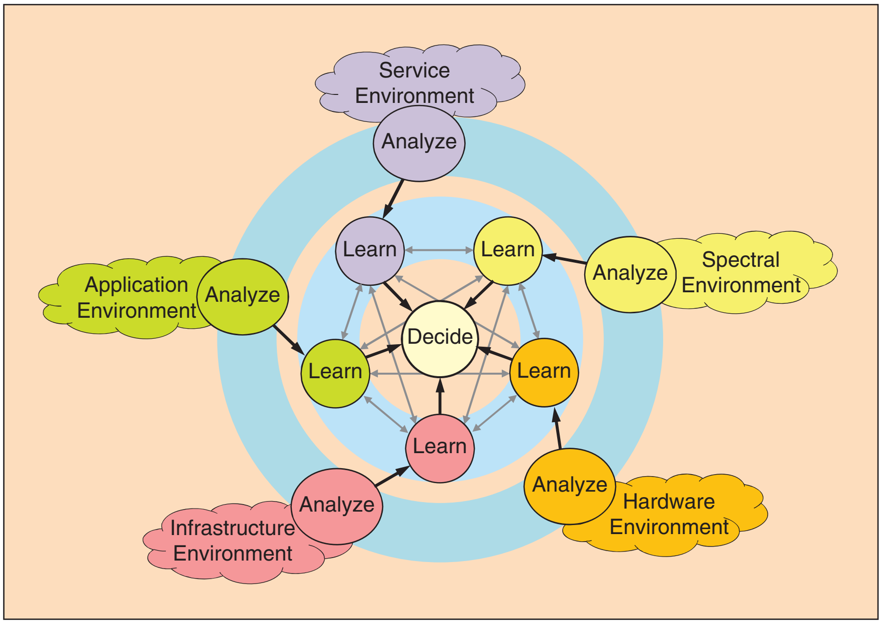

FIGURE 2 Cognitive cycles.

Management Methodology

The proposed framework facilitates a joint management and coordinated reconfiguration of all wireless access layers. The system needs to operate in real time and continuously adapts to its surrounding, where the users opportunistically use spectral, service, application, infrastructure, and hardware resources for fractions of time.

The framework would benefit from a flat management hierarchy and distributed sensing and processing [3]. We suggest two hierarchy levels: the management space and the sensor/actuator space. The sensors continuously sense the resource environments. The actuators respond to high-level management requests with appropriate actions.

Figure 4 shows two radios that are associated with different subsystems. The intelligent subsystem jointly processes the environmental state information and specifies the management rules for each software radio layer. It triggers the allocation and reallocation of resources. Spectral resources will be frequently reassigned for adapting to momentary traffic characteristics. Infrastructure updates and hardware upgrades will be less frequent.

The five managers at the sensor/actuator space represent different logical entities that perform some lower-level management tasks (Table 2). The low-level managers may be further split into submanagers. A modular approach eases the dynamic incorporation of new management elements.

The proposed management methodology leverages flexibility and enables managing the software radio layers in a centralized, semicentralized, or distributed fashion. Cognition may be present not only at the management space but also at the sensor/actuator space. Decisions may be locally derived, and their observations fed back to a central node (intelligent subsystem) for refining the knowledge database [8]. Acquired knowledge may be directly distributed to the surrounding radios, teaching them how to operate in similar situations (docitive networks [9]). Both approaches may eventually be combined for minimizing the signaling overhead.

Analysis

We simulate a simple scenario for demonstrating the feasibility and potential benefits of a cognitive resource management extended to all wireless access layers. SDR mobile terminals are reconfigured to operate in different radio modes, providing different services and QoS levels.

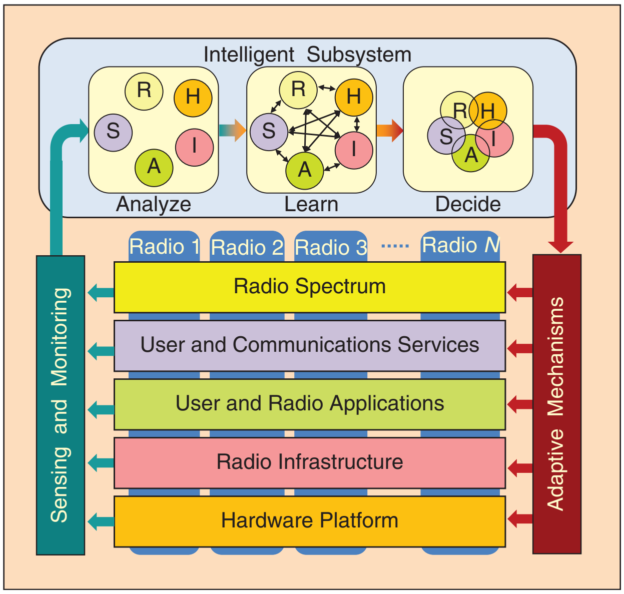

FIGURE 3 Resource management framework.

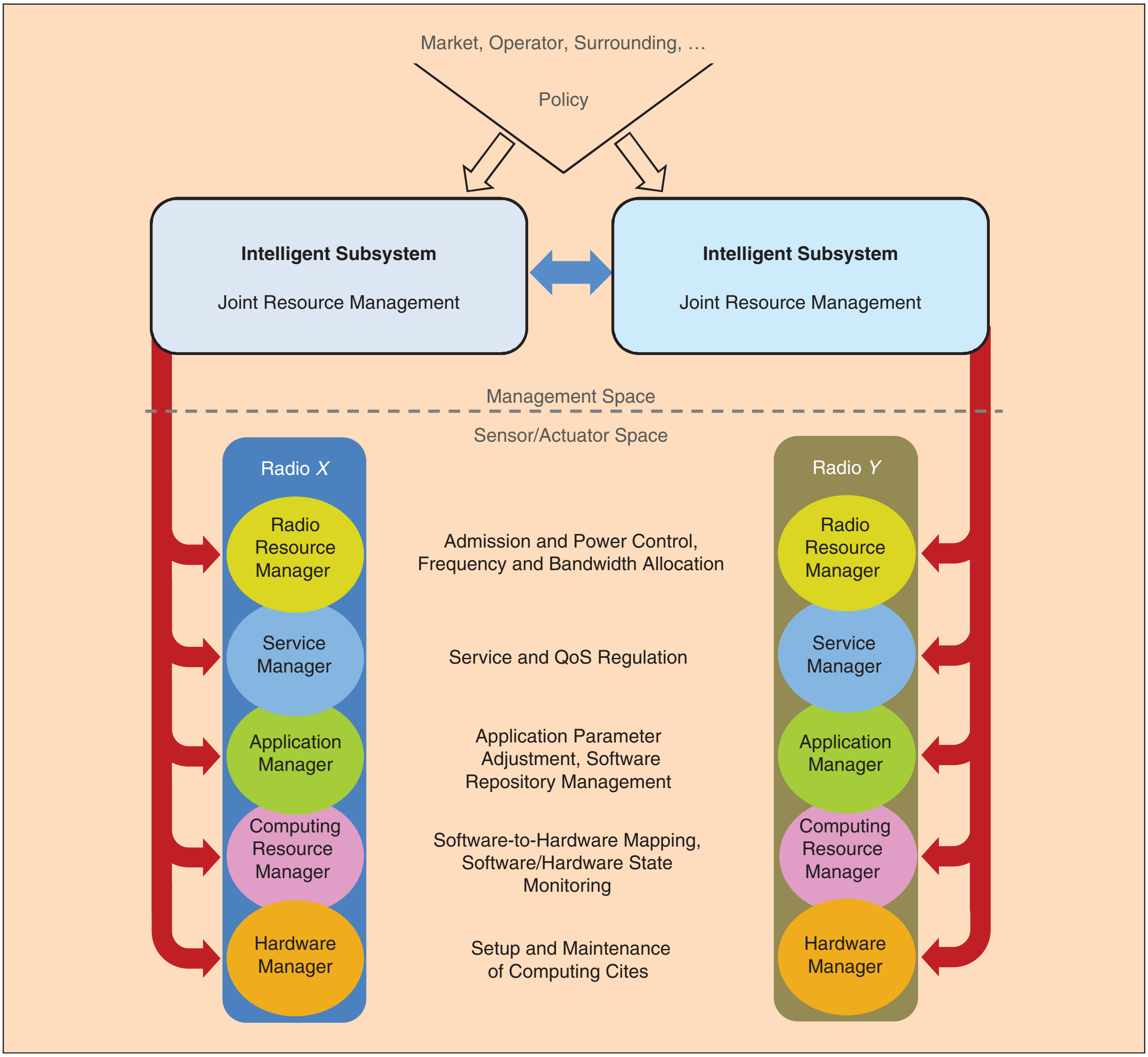

FIGURE 4 Management methodology.

The different managers observe the environment and process the information within the intelligent subsystem. Here, the radio resource manager decisions are characterized by the signal-to-noise ratio (SNR) target. The simulation considers uniformly distributed SNR values between 1 and 10 dB. The service manager provides one out of four QoS levels and manages the users' service requests and profiles. All users are willing to accept a QoS downgrade if the desired level is not feasible. The application manager correspondingly offers four SDR processing chains, enabling data rates of 1,024, 384, 128, and 64 kb/s. We assume the Universal Mobile Telecommunications System receiver processing models consisting of 20 processing blocks each [8, Section 7.4.2-C]. The four data rates, or QoS levels, are requested with a probability of 10, 20, 30, and 40%, respectively.

The receiver's turbo decoder incorporates an early stopping mechanism. The number of necessary decoding iterations and the processing chain's computing demand are a function of the SNR target. The mobile terminals' processors are dual cores. The processing capacities are randomly distributed between 3,800 and 7,200 million operations per second per core. Terminals are reconfigurable and may run more than one wireless transceiver at a time. Therefore, the instantaneous computing resource status and SNR figure limit the terminal's capability for providing certain QoS levels. The local hardware managers inform the computing resource manager at the base station about the terminals' available computing resources.

Upon service request, the intelligent subsystem evaluates the available environmental parameters before deciding which SDR application should be downloaded at each terminal. The computing resource manager therefore tests the application-to-platform mapping, employing the t_1-mapping algorithm [8]. A candidate application is chosen if the mapping trial satisfies all the computing resource demands with the available resources (feasible mapping). Otherwise, the application that offers the next lower QoS level is tested. Because of the limited computing resources for processing numerous session requests per second, not more than two or four mapping trials per user may be allowed. Knowledge is dynamically acquired by accumulating the results of the mapping trials as a function of the computing state information only (Comp. Only), assuming the lowest SNR value, or the computing state and SNR information (Comp. and SNR). We compare these two methods with a system that does not process any computing resource information (None). Figure 5 shows the results.

We observe that, without knowledge of the terminals' computing resource capabilities, 45% of user sessions cannot be established [Figure 5(a)]. A system that processes the terminals' computing capabilities is able to provide a feasible service to many of these users, although at a lower QoS [Figure 5(b)]. The decaying curves in Figure 5(a) and the saturating curves in Figure 5 (b) indicate the learning process. The learning speed is a function of the number of trials and the information to be processed. The more information the system acquires (computing and SNR here), the longer it takes for the system to learn, but the better the decisions—2% rather than 6 or 45% infeasible sessions [Figure 5(a)] and 34 instead of 41% degraded sessions [Figure 5(b)].

***TABLE 2* Resource managers.**

Manager	Management tasks
Radio resource manager	It processes environmental state information for allocating or reallocating radio resources.
Service manager	It manages service and QoS profiles and their dynamic deployment as a function of environmental conditions and other managers' actuations.
Application manager	It selects software-operational modes and parameters as a function of user service and QoS request, terminal characteristics, user preferences, and hardware and software compatibilities.
Computing resource manager	It dynamically maps different radio and user applications to the available hardware resources, ensuring real-time execution under the given system constraints.
Hardware manager	It automatically discovers and sets up and maintains RF, processing, and networking components.

Additional radio parameters or more versatile terminal architectures and SDR applications will require longer learning processes. More sophisticated cognitive engines are needed to deal with the diversity of the wireless scenario.

Conclusions

Future DSP chains will essentially be implemented in software. Software repositories will then provide the processing modules and entire applications, which can be dynamically deployed as a function of the spectral, service, infrastructure, and hardware environments. Rather than guaranteeing radio and computing resources in advance, they will be dynamically allocated from shared resource pools.

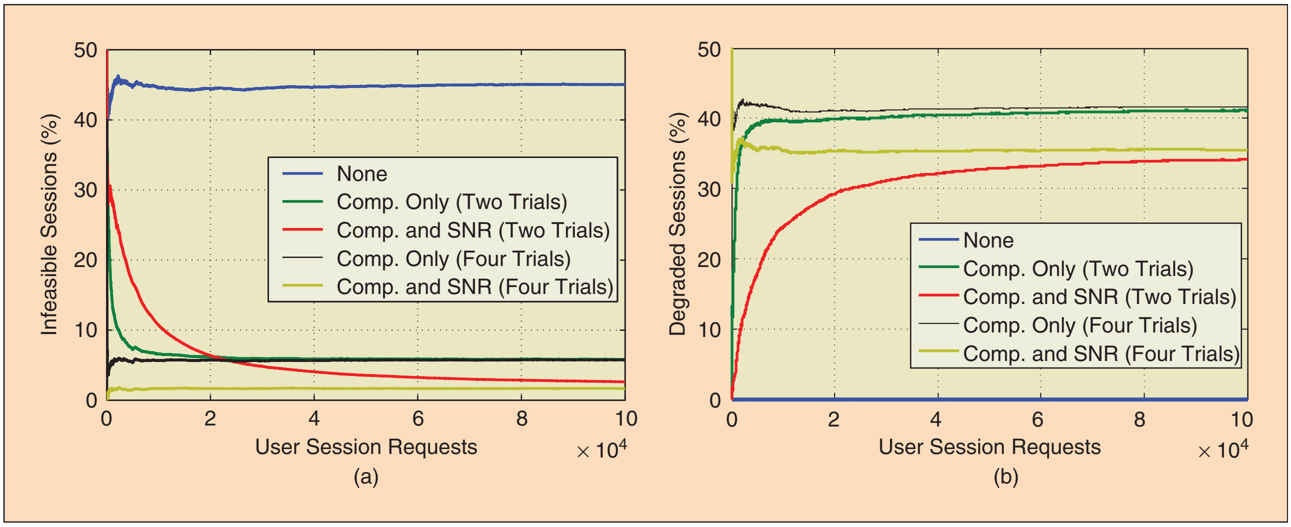

FIGURE 5 Simulation results: percentage of (a) infeasible and (b) degraded sessions.

TABLE 3 Different parties and their benefits.

Party	Benefits
Service provider	Easy distribution of user services. Possibility of introducing value-added services. Wider business scope (communications support services).
Wireless subscriber	Innovative, personalized, and ubiquitous services. Service on demand (virtually unlimited resources). Flexible service deals.
Hardware manufacturer	New product and technology opportunities: energy-scalable equipment, flexible RF components, wideband antennas and data converters, and real-time capable multiprocessing platforms.
Application developer	Hardware-independent application design. Faster time to market. New business opportunities (third parties).
Radio operator	Risk outsourcing for the deployment of new infrastructure. On demand use of infrastructure (pay per use). Focus on RRM issues. New value-added services.
Computing operator	Economy of scale advantages: e.g., lower cost of purchase, better specialization, spread of operational costs (cooling, power conversion, R&D, maintenance, auxiliary staff, and real estate) over multiple hardware devices and sites.

The increasing flexibility and complexity of wireless communications systems suggest a cognitive resource management that embraces all wireless access layers for coordinating the use of resources. The system we envisage facilitates

- automatically selecting radio applications and operational modes
- evaluating RATs and operational modes
- dynamically testing and improving the radio infrastructure
- evaluating hardware platforms
- monitoring and updating user profiles
- supervising the fixed and mobile communications sites for dynamically creating ad hoc or relay networks.

The benefits of a cognitive resource management extended to all wireless access layers are manifold, essentially enabling better specializations and higher resource efficiencies (Table 3). These benefits come at the price of additional resource overheads for acquiring and processing the environmental state information in real time. Additional hardware resources are particularly needed for carrying out the necessary management activities. The sensing, collaborative learning, and decision-making processes also rely on additional signaling. We, therefore, envisage a gradual implementation of cognitive management services.

Acknowledgment

This work was supported by Spanish government (MICINN) and FEDER under project TEC2011-29126-C03-02.

Author Information

Vuk Marojevic (marojevic@tsc.upc.edu) received his Dipl.-Ing. degree from the University of Hannover in 2003 and Ph.D. degree from the Universitat Politècica de Catalunya (UPC) in 2009, both in electrical engineering. He is currently an assistant professor at the Department of Signal Theory and Communications of the UPC. His research efforts include software-defined and cognitive radio architectures and algorithms for an efficient and flexible computing resource management.

Ismael Gomez is a Ph.D. candidate at the Department of Signal Theory and Communications of UPC. He gained expertise in designing SDR execution environments for high-performance reconfigurable computers. His Ph.D. research is on flexible and energy-efficient wireless communications.

Antoni Gelonch is an associate professor at the Department of Signal Theory and Communications of UPC. His research focuses on implementation methodologies supporting cognitive resource management strategies and algorithms for flexible wireless communications systems and networks.

References

[1] T. H. Han, J. C. Yoo, and H. Lee, "Power, interface, and integration: Handset chipset design issues," *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 172–179, Nov. 2009.

[2] J. Mitola, "The software radio architecture," *IEEE Commun. Mag.*, vol. 33, no. 5, pp. 26–38, May 1995.

[3] S. Zhou, M. Zhao, X. Xu, J. Wang, and Y. Yao, "Distributed wireless communication system: A new architecture for future public wireless access," *IEEE Commun. Mag.*, vol. 41, no. 3, pp. 108–113, Mar. 2003.

[4] J. Mitola, III, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Royal Institute of Technology (KTH), Stockholm, Sweden, May 2000.

[5] S. Haykin, J. H. Reed, G. Y. Li, and M. Shafi, Eds., "Cognitive radio. Part 1: Practical perspectives," *Proc. IEEE (Special Issue)*, vol. 97, no. 4, pp. 608–773, Apr. 2009.

[6] S. Haykin, J. H. Reed, G. Y. Li, and M. Shafi, Eds., "Cognitive radio. Part 2: Fundamental issues," *Proc. IEEE (Special Issue)*, vol. 97, no. 5, pp. 784–939, May 2009.

[7] J. Mitola, III, "Cognitive radio: Agent-based control of software radios," in *Proc. 1st Karlsruhe Workshop Software Radio*, Karlsruhe, Germany, Mar. 2000.

[8] V. Marojevic. (2009, July). Computing resource management in software-defined and cognitive radios, Ph.D. dissertation, Universitat Politècica de Catalunya (UPC), Barcelona, Spain. [Online]. Available: http://flexnets.upc.edu/trac/wiki/Publications

[9] L. Giupponi, A. Galindo, and M. Dohler, "From cognition to docition: The teaching radio paradigm for distributed & autonomous deployments," *Elsevier Comput. Commun.*, vol. 33, no. 7, pp. 2015–2020, Nov. 2010.

[10] P. F. Marshall, "Extending the reach of cognitive radio," *IEEE Proc.*, vol. 97, no. 4, pp. 612–625, Apr. 2009.

[11] M. Mueck, A. Piipponen, K. Kalliojärvi, G. Dimitrakopoulos, K. Tsagkaris, P. Demestichas, F. Casadevall, J. Pérez-Romero, O. Sallent, G. Baldini, S. Filin, H. Harada, M. Debbah, T. Haustein, J. Gebert, B. Deschamps, P. Bender, M. Street, S. Kandeepan, J. Lota, and A. Hayar, "ETSI reconfigurable radio systems: Status and future directions on software defined radio and cognitive radio standards," *IEEE Commun. Mag.*, vol. 48, no. 9, pp. 78–86, Sept. 2010.

[12] M. Conti, S. Giordano, M. May, and A. Passarella, "From opportunistic networks to opportunistic computing," *IEEE Commun. Mag.*, vol. 20, no. 7, pp. 126–139, Sept. 2010.

[13] I. Gomez, V. Marojevic, and A. Gelonch, "Resource management for software-defined radio clouds," *IEEE Micro.*, vol. 32, no. 1, pp. 44–53, Jan./Feb. 2012.

VT