Cognitive and cooperative wireless networks

Sergio Palazzo (✉), Davide Dardari, Mischa Dohler, Sinan Gezici,
Lorenza Giupponi, Marco Luise, Jordi Pérez Romero, Shlomo Shamai,
Dominique Noguet, Christophe Moy, and Gerd Asheid

1 Cognitive radio networks

The traditional approach of dealing with spectrum management in wireless communications has been the definition of a licensed user granted with exclusive exploitation rights for a specific frequency. While it is relatively easy in this case to ensure that excessive interference does not occur, this approach is unlikely to achieve the objective to maximize the value of spectrum, and in fact recent spectrum measurements carried out worldwide have revealed a significant spectrum underutilization, in spite of the fact that spectrum scarcity is claimed when trying to find bands where new systems can be allocated. Just to mention some examples of measurements, different studies can be found in [1–6], revealing that overall occupation in some studies for frequencies up to 7 GHz could be in the order of only 18%.

As a result of the above, one of the current research trends in the spectrum management are the so-called Dynamic Spectrum Access Networks (DSANs), in which unlicensed radios, denoted in this context as Secondary Users (SUs) are allowed to operate in licensed bands provided that no harmful interference is caused to the licensees, denoted in this context as Primary Users (PU). The proposition of the TV band Notice of Proposed Rule Making (NPRM) [7], allowing this secondary operation in the TV broadcast bands if no interference is caused to TV receivers, was a first milestone in this direction. In this approach, SUs will require to properly detecting the existence of PU transmissions and should be able to adapt to the varying spectrum conditions, ensuring that the primary rights are preserved. These events culminated in the creation of the IEEE 802.22, developing a cognitive radio-based physical and medium access control layer for use by license-exempt devices on a non-interfering basis in spectrum that is allocated to the TV broadcast service. Based on these developments it is reasonable to think that the trend towards DSANs has just started and that given the requirements for a more efficient spectrum usage, it can become one of the important revolutions in the wireless networks for the next decades, since it completely breaks the way how spectrum has been traditionally managed.

The primary-secondary (P-S) spectrum sharing can take the form of cooperation or coexistence. Cooperation involves explicit communication and coordination between primary and secondary systems, and coexistence means there is none. When sharing is based on coexistence, secondary devices are essentially invisible to the primary. Thus, all of the complexity of sharing is born by the secondary and no changes to the primary system are needed. There can be different forms of coexistence, such as spectrum underlay (e.g. UWB) or spectrum overlay (e.g. opportunistic exploitation of white spaces in spatial-temporal domain sustained on spectrum sensing, coordination with peers and fast spectrum handover). As for cooperation, again different forms of P-S interactions are possible. For example, spatial-temporal white spaces that can be exploited by SUs can be signalled through appropriate channels. In addition, the interaction between PUs and SUs provides an opportunity for the license-holder to demand payment according to the different quality of service grades offered to SUs.

One of the key enabling technologies for DSAN development is Cognitive Radio (CR), which has been claimed to be an adequate solution to the existing conflicts between spectrum demand growth and spectrum underutilization. The term Cognitive Radio was originally coined by J. Mitola III in [8] and envisaged a radio able to sense and be aware of its operational environment so that it can dynamically and autonomously adjust its radio operating parameters accordingly to adapt to the different situations. CR concept was built in turn upon the Software Defined Radio (SDR) concept, which can be understood as a multiband radio supporting multiple air interfaces and protocols and being reconfigurable through software run on DSP or general-purpose microprocessors. Consequently, SDR constituted the basis for the physical implementation of CR concepts.

Thanks to this capability of being aware of actual transmissions across a wide bandwidth and to adapt their own transmissions to the characteristics of the spectrum, CRs offer great potential for bringing DSANs to reality, and in fact DSANs are usually referred to as Cognitive Radio Networks (CRN). The operating principle of a CR in the context of a DSAN is to identify spatial and temporal spectrum gaps not occupied by primary/licensed users, usually referred to as *spectrum holes* or *white spaces*, place secondary/unlicensed transmissions during such spaces and vacate the channel as soon as primary users return. The CR concept therefore implicitly relies on two basic premises: the existence of enough white spaces caused by primary spectrum underutilization and the ability of secondary users to effectively detect and identify the presence of different licensed technologies in order not to cause harmful interference.

From a general operation perspective, a CR follows the so-called *cognition cycle* to enable the interaction with the environment and the corresponding adaptation. It consists in the *observation* of the environment, the *orientation* and *planning* that leads to making the appropriate *decisions* pursuing specific operation goals, and finally *acting* over the environment. Decisions on the other hand can be reinforced by *learning* procedures based on the analysis of prior observations and on the corresponding results of the prior actuations. Then, when particularizing the cognition cycle to the dynamic spectrum access for a secondary user, the observation turns

out to be the spectrum sensing in order to identify the potential white spaces, the orientation and planning steps would be associated with the analysis of the available white spaces, and finally the acting step would be in charge of selecting the adequate white space to make the secondary transmission, together with the setting of the appropriate radio parameters such as transmit power, modulation formats, etc.

There are a number of techniques to be developed for an implementation of efficient secondary spectrum usage through cognitive radio networks, and are classified in [9] as spectrum sensing, spectrum management, spectrum mobility and spectrum sharing mechanisms. These techniques are briefly discussed in the following:

- *Spectrum sensing*. It consists in detecting the unused spectrum bands that can be potentially exploited for secondary communications. A lot of different spectrum sensing techniques have been studied in the last years, such as the energy detector, which does not include any specific knowledge about the primary signal to be detected, the matched filter detection, which requires the knowledge of the specific primary signal formats, or the cyclostationarity feature detection. Also the possibility of combining sensing measurements from different sensors through appropriate fusion schemes has been considered in the so-called cooperative sensing. Even from a more general perspective, the possibility that the network provides knowledge about the current spectrum bands available through some control channel has also been considered. This is the case of e.g. the development of the so-called Cognitive Pilot Channel (CPC) in [10]. From this perspective, and having in mind the possibility of combining the knowledge provided by the network with the knowledge acquired by the sensing process, spectrum sensing concept can be generalised to the concept of spectrum awareness.

- *Spectrum management*. This refers to the selection of the most adequate spectrum band to carry out the transmission in accordance with the secondary user requirements. This selection should be made based on the characteristics of the channel in terms of e.g. the maximum capacity that can be obtained by the secondary users, and also taking into consideration the maximum interference that can be tolerated by primary receivers. The decision making process here can be benefited from the application of learning strategies, that, based on experience acquired from prior decisions, can orient the decisions towards the selection of some channels in front of others. As an example, in case that in some channels the primary user activity is higher, it is more likely that primary users appear forcing the secondary transmitter to evacuate the channel, so if such knowledge was available, this could prevent the secondary network from selecting these channels.

- *Spectrum mobility*. This functionality consists in establishing appropriate mechanisms to ensure that an on-going secondary communication can be continued whenever a primary user appears in the occupied bandwidth. This will thus involve the ability to detect the appearance of this primary user, which requires some continuous monitoring of the channel, e.g. through sensing mechanisms. Then, when the primary user appears, the occupied channel has to be evacuated, and an alternative spectrum piece has to be found where the communication can be reassumed, which is usually called *spectrum handover*.

- *Spectrum sharing.* This function targets the provision an efficient mechanism so that coexisting secondary users can share the available spectrum holes. Adequate Medium Access Control (MAC) protocols and scheduling mechanisms would be needed, and they would be very much dependant on how the secondary network is deployed, e.g. if it is infrastructure or infrastructure-less based, etc.

Although all the above functions have become a hot research topic during the last few years, there is still a lot of work to do before CRNs can become a reality. This will involve not only technical aspects, but also significant regulatory changes will be needed. In addition, this will also have implications from the techno-economic perspective, with the appearance of new business models to exploit the capabilities offered by CRNs, involving different possibilities ranging from secondary cellular operators that could offer services at cheaper prices at the expense of a somehow reduced quality, to the deployment infrastructure-less secondary networks that would enable the communication of short range devices. Clearly, all these elements put a quite long-term perspective, maybe of several decades, before the final implementation of CRNs.

2 Cognitive positioning

2.1 Introduction

According to the definition above, cognitive systems strive for optimum spectrum efficiency by allocating capacity as requested in different, possibly disjoint frequency bands. Such approach is naturally enabled, by the adoption of flexible MultiCarrier (MC) technologies, in all of its flavors: traditional OFDM, Filter-Bank Multicarrier Modulation (FBMCM) [20], and possibly non-orthogonal formats with full time/frequency resource allocation. Most current and forthcoming wideband standards for wireless communications (3GPP's Long-Term-Evolution is a paradigmatic example in this respect) are based on such multicarrier signalling technology, so that the signal allocated to each terminal is formed as the collection of multiple data symbols intentionally scattered across non-contiguous spectral chunks.

On the other hand, modern wireless networks more and more expect availability of location information about the wireless terminals, driven by requirements coming from applications, or just for better network resources allocation. Thus, signal-intrinsic capability for accurate localization is a goal of 4th Generation (4G) as well as Beyond-4G (B4G) networks. All signal processing techniques that can contribute to the provision of accurate location information are welcome in this respect. Such techniques can pair the ones that a cognitive terminal adopts to establish a reliable, high-capacity link.

The most accurate techniques to perform localization of a wireless terminal are based on time-of-arrival (TOA) estimation of a few radio ranging signals, followed by ranging and appropriate triangulation. Therefore, the precision of positioning is strictly related to the accuracy that can be attained in the estimation of the propagation delay of the radio signal [16, 18]. In the following, we will see that a multicar-

rier signal format, possibly split in (two or more) non-contiguous bands, gives new opportunities in terms of enhanced-accuracy time delay estimation (that ultimately translates into enhanced accuracy positioning). By chance, the two issues of super-accurate signal TOA estimation and sparse multicarrier resource allocation, we can say, "marry in heaven".

To let the reader understand how this could be done, we will start with a review of the Modified Cramér-Rao bound (MCRB), its frequency-domain computation, as well as the study of the impact on the bound of the location of the received signal spectrum within the receiver bandwidth. After this, we will discuss how to optimize an MC signal format through minimization of the MCRB, to come to the description of Cognitive Positioning (CP) opportunities [12, 21].

2.2 Criteria to optimize the function of ranging

The Cramér-Rao bound (CRB) is a fundamental lower bound on the variance of any estimator [13, 19] and, as such, it serves as a benchmark for the performance of actual parameter estimators [11, 15, 17]. It is well known and widely adopted for its simple computation, but its close-form evaluation becomes mathematically intractable when the vector of observables contains, in addition to the parameter to be estimated, also some *nuisance parameters*, i.e., other unknown random quantities whose values are not the subject of the estimator (information data, random chips of the code of a ranging signal, etc.), but that concur to shape the actual observed waveform. The MCRB for estimation of the TOA (delay) τ for a received signal $x(t - \tau)$ embedded in complex-valued AWGN $n(t)$ with two-sided psd $2N_0$ is found to be [14]

$$\mathrm{MCRB}\,(\tau) = \frac{N_0}{\mathbb{E}_{\mathbf{c}}\left\{ \int_{T_{\mathrm{obs}}} \left| \frac{\partial x(t - \tau)}{\partial \tau} \right|^2 dt \right\}}, \qquad (1)$$

where $\mathbf{c}$ is a vector collecting all of the nuisance parameters, T_{obs} is the observation time-interval, and $\mathbb{E}_{\mathbf{c}}$ indicates statistical expectation wrt $\mathbf{c}$. This expression holds in the case of ideal coherent demodulation, i.e., assuming that during signal tracking the carrier frequency and the carrier phase are known to a sufficient accuracy. From this expression, we can devise a simple criterion for optimal signal design: finding that specific waveform that, across a pre-set bandwidth and for a certain signal-to-noise-ratio (SNR), gives the minimum MCRB value. We will not consider here any aspects related to a possible *bias* of the estimator arising in a severe multipath propagation environment, to concentrate on the main issue that we have just stated.

Assume we are receiving a generic pilot ranging signal $x(t; \mathbf{c})$ bearing no information data, but containing a pseudo-random ranging code $\mathbf{c}$, whose chips $\in \{\pm 1\}$ are considered as binary iid nuisance parameters. This signal turns out to be a *parametric random process*, for which each time-unlimited sample function is a signal with finite power P_x and chip rate $R_c = 1/T_c$.

Assuming that the observation time is sufficiently large, the MCRB can be also derived through frequency-domain computation [21]:

$$\text{MCRB}\,(\tau) = \frac{N_0}{T_{\text{obs}}4\pi^2 \int_{-\infty}^{+\infty} f^2 \mathscr{P}_x(f)\,df} = \frac{1}{8\pi^2 N \beta^2 \cdot E_c/N_0}, \tag{2}$$

where $\mathscr{P}_x(f)$ is the power spectral density (PSD) of $x(t)$, E_c is the received signal energy per chip, $NT_c = T_{\text{obs}}$, and β^2 is the normalized second-order moment of the PSD,

$$\beta^2 = \frac{1}{2P_x} \int_{-\infty}^{+\infty} f^2 \mathscr{P}_x(f)\,df. \tag{3}$$

We conclude that the MCRB depends on the second-order moment β^2 of the PSD of the signal x, irrespective of the type of signal format (modulation, spreading, etc.) that is adopted.

2.3 Ranging multicarrier signals

A ranging MC signal can be constructed following the general arrangement of multi-carrier (MC) modulation: the input chip stream $c[i]$ of the ranging code is parallelized into N substreams with an MC symbol rate $R_s = R_c/N = 1/(NT_c) = 1/T_s$, where T_s is the time duration of the "slow-motion" ranging chips in the N parallel substreams. We can use a "polyphase" notation for the k-th ranging subcode ($k = 0, 1, \ldots, N-1$) in the k-th substream (subcarrier) as $c^{(k)}[n] \triangleq c[nN+k]$, where k is the subcode identifier and/or the subcarrier index, whilst n is a time index that addresses the n-th MC symbol (block) of time length $T_s = NT_c$. The substreams are then modulated onto a raster of evenly-spaced subcarriers with frequency spacing f_{sc} and the resulting modulated signals are added to give the (baseband equivalent of the) overall ranging signal. In Filter-Bank Multicarrier Modulation (FBMCM) (also called Filtered Multi-Tone, FMT) the spectra on each subcarrier are strictly bandlimited and nonoverlapping, akin to conventional single channel per carrier (SCPC). The resulting signal is [20]

$$x(t) = \sqrt{\frac{2P_x}{N}} \sum_{n} \sum_{k=0}^{N-1} p_k c^{(k)}[n] g\,(t - nT_s) \exp\left[j2\pi k f_{\text{sc}}t\right], \tag{4}$$

where $g(t)$ is a bandlimited pulse, for instance a square-root raised cosine pulse with roll-off factor α. In this case, the subcarrier spacing is $f_{\text{sc}} = (1+\alpha)/T_s$. It is well known that this arrangement has an efficient realization based on IFFT processing followed by a suited polyphase filterbank based on the prototype filter $g(t)$. The real-valued coefficient p_k in (4), $0 \leq p_k \leq N$, $\sum_{k=0}^{N-1} p_k^2 = N$, allows us to perform the function of *power allocation* by allowing different amounts of signal power to be placed on the different subcarriers. Some p_k's can also be 0 indicating that the relevant subcarriers (or even a whole subband) is not being used. We will see how this feature is essential for the characteristics of cognitive positioning.

When T_{obs} is sufficiently large, $T_{\text{obs}} = N_m T_s$, $N_m \gg 1$, the MCRB for the MC signal can be easily computed:

$$\text{MCRB}\,(\tau) = \frac{T_c^2}{8\pi^2 \cdot \frac{E_c}{N_0} \cdot \frac{N_m}{N} \cdot \left[\xi_g + \frac{(1+\alpha)^2}{N}\sum_k p_k^2 k^2\right]}, \tag{5}$$

where ξ_g is the so-called pulse shape factor (PSF),[1] a normalized version of the Gabor bandwidth of pulse $g(t)$:

$$\xi_g \triangleq \frac{T_s^2 \int_{-\infty}^{+\infty} f^2 |G(f)|^2\, df}{\int_{-\infty}^{+\infty} f^2 |G(f)|^2\, df}. \tag{6}$$

2.4 Cognitive signals for cognitive positioning

An MC signal with uneven, adaptive power distribution can be adopted to implement a *cognitive positioning system*. In our envisioned FBMCM scheme for positioning, the proper power allocation allows to reach the desired positioning accuracy, not only in an additive white Gaussian noise (AWGN) channel, but also in an additive coloured Gaussian noise (ACGN) channel. Coloured noise arises from variable levels of interference produced by co-existing (possibly primary) systems on different frequency bands. The key assumption is that such interference can be modelled as a Gaussian process. This is certainly justified in wireless networks with unregulated multiple access techniques such as CDMA and/or UWB.

To be specific, let us investigate the issue of finding the power allocation scheme that gives the *minimum MCRB* for TDE in a Gaussian channel whose (additive) noise has a variable PSD $\mathscr{P}_n(f)$. After some algebra, we find out that

$$\text{MCRB}\,(\tau)|_{\text{ACGN}} = \frac{1}{4\pi^2 T_{\text{obs}}\,(\Delta f)^3}\left(\sum_k k^2 \frac{\mathscr{P}_x(k\Delta f)}{\mathscr{P}_n(k\Delta f)}\right)^{-1}, \tag{7}$$

where $\Delta f = f_{\text{sc}}$ is the subcarrier spacing, and where the PSD of the transmitted signal and of the noise were considered constant across each subband. A fundamental result is obtained if we let $N \to \infty$ (thus $\Delta f \to df$):

$$\text{MCRB}\,(\tau)|_{\text{ACGN}} = \frac{1}{4\pi^2 T_{\text{obs}}}\left(\int_{-B}^{+B} f^2 \frac{\mathscr{P}_x(f)}{\mathscr{P}_n(f)}df\right)^{-1}, \tag{8}$$

where B is the (finite) signal bandwidth.

Coming back to the problem of enhancing TDE accuracy, and sticking for simplicity to the finite-subcarriers version of the problem, we have to minimize the MCRB (7) with a constraint on the total signal power P_x.

---

[1] For a square-root raised-cosine pulse with roll-off factor α, $\xi_g = 1/12 + \alpha^2 \left(1/4 - 2/\pi^2\right)$.

Considering that $\mathscr{P}_x(k\Delta f)$ is proportional to P_k^2, we are to find the power distribution P_k^2 that *maximizes*

$$\sum_k p_k^2 \frac{k^2}{\mathscr{P}_n(k\Delta f)} \tag{9}$$

subject to the constraints $\sum_k p_k^2 = N$ and, of course, $p_k \geq 0$. The optimal distribution is easily found to be

$$\begin{cases} p_k = 0, & k \neq k_M, \\ p_k = \sqrt{N}, & k = k_M \end{cases} \qquad k_M = \arg\max_k \frac{k^2}{\mathscr{P}_n(k\Delta f)} \tag{10}$$

that corresponds to placing all the power onto the sub-band for which the *squared-frequency to noise ratio* (SFNR) $k^2 / \mathscr{P}_n(k\Delta f)$ is maximum. In AWGN, this is the band-edge subcarrier, as is known form Gabor bandwidth analysis.

A more realistic case study for CP in ACGN takes also into account possible power limitations on each subcarrier that prevents from concentrating all of the signal power onto the edge subcarriers (for AWGN) or on the subcarrier with the best SFNR as above. We add thus the further constraint $0 \leq p_k^2 \leq P_{\max} < N$. The solution to this new power allocation problem can be easily found via linear programming:

- order the square-frequency-to-noise-ratios SFNR$_k$ from the highest to the lowest; set the currently allocated power to zero; mark all carriers available;
- find the available power as the difference between the total power N and the currently allocated power. If it is null, then STOP, else, if it is larger than $P_{\max}$, then put the maximum power $P_{\max}$ on the available carrier with the highest SFNR; else put on the same carrier the (residual) available power;
- update the currently allocated power by adding the one just allocated, and remove the just allocated carrier from the list of available carriers. If the list is empty, then STOP, else goto the previous step).

This results in a set of bounded-power subcarriers that gives the optimum power allocation (minimum TDE MCRB) with ACGN. An avenue for research is finding practical algorithms that attain the bounds above in a realistic environment, and extend the results above to cases with strong multipath.

3 Cooperative wireless networks

The concept of cooperation actually emerged in the late sixties with the work of Van Der Meulen about "Three terminal communications". Interestingly, the capacity of this scheme is still an open problem today. More recently, the concept of relaying or cooperation has gained a lot of interest for several reasons:

- In mobile or wireless communications, the potential offered by multi-antenna transmission and/or reception is now clearly established, and this technology, known under the generic name of MIMO, has found its way in a number of stan-

dards. "Classical" MIMO however assumes that the different antennas can be or are colocated on the same site. There are scenarios, however, where this assumption cannot be met for operational reasons and/or cost reasons. Therefore, the idea has emerged as whether different non colocated entities could *form a coalition* to mimic in a distributed manner a multi-antenna system, thereby getting access to the benefits of MIMO in term of rate (multiplexing gain) and/or diversity, according to the well-known trade-off between the two.

- A natural way to exploit this idea is to serve a user by means of two or more base stations. This concept is also known as macro-diversity. Assuming a wired backhaul, the base stations know in the best case all the data and the channel state information of all the users in the cluster or "supercell" served by the coordinated base-stations. There is no issue of *decoding strategy* by the *relay* in such a case because the wired nature of the backhaul makes sure the data are available without any error. The design of linear precoders and decoders for such a scenario, possibly robust to imperfect channel knowledge, is its infancy. Another issue, to avoid very heavy signalling in the backhaul, is that of distributed solutions based on partial data and/or channel knowledge at the coordinating node, yet approaching the performance of a fully *informed* solution.

- While the motivation behind the previous concept is mainly to avoid intercell interference (the COMP approach in LTE), another motivation is associated with the issue of coverage and *badly* located users which might be out of (good) reach by any base station. An emerging concept is that of a "popping up base station" with wireless backhaul, that would help one or many poor users. In this case, because of the wireless nature of the backhaul, issues similar to that of the classical relay channel reappear, which are of course related to the fact that the base station first has to receive properly the data to be relayed. Hence the issue of decoding strategy (in the broad sense) has to be considered. Along these lines there should be an increasing interest for the design of nodes serving as relays and which would be equipped with MIMO capability: MIMO relay schemes.

- Moving to a totally different scenario like wireless sensor networks, there is also a clear interest for cooperative solutions. As a matter of fact, sensing nodes may usually be equipped with a battery which is supposed to last as long as possible, may not be easily reachable and limits the communication capability of the node in term of available power. Therefore not all nodes are necessarily able to directly establish a connexion with the *fusion center* or the *collecting point*. The network is then of the mesh types rather than of the star type. The data has to reach the collecting node by means of multiple hops where some sensing nodes basically relay the information of their neighbours. The choice of the cooperating nodes may be based on several criteria or utility functions, incorporating not only rate and/or bit/packet error measures but also penalty depending on the power used and/or the status of the battery of the possibly cooperating nodes.

- These days there is an increasing interest for *cognitive communications systems*. The basic idea behind is the capability to sense the spectrum, to detect possible holes and to establish communication in *free frequency slots*. While the concept of *cognition* has mainly been discussed for the PHY layers, there is an emergence

of the concept of cognitive networks, where the load or presence of *available resources* would even be exploited at different layers of the systems. Coming back to the PHY layer, cooperation is a natural tool and a desirable feature in order to be able to properly sense the spectrum and address concerns like the hidden terminal problem. Therefore, obtaining *spectrum maps* naturally leads to a joint cooperative-cognitive approach.

Wireless systems are often networks in the sense that multiple entities or users compete for the available resource(s). For instance, a node in a multihop setup may be expected to relay the signals of several adjacent nodes. In that case, the strategy to be chosen by the relaying node to serve the neighbouring nodes has to be properly addressed. An important concept that emerged recently and deserves further investigation is that of *network coding* which shows promises but also needs to properly encode and simultaneously relay the information of several users at the same time.

A final remark has to do with energy saving and *green communications*. For a prescribed performance metric to be achieved at the receiving end, the combination of transmitter and relay might require a lower total transmission power than if the transmitter alone is sending this information. There are results clearly indicating this. However it should also be noted that transmission power is only one part of the global picture. Associated with any communicating nodes, there are additional power prices like those associated with computation, security, etc. It would be highly interesting and motivating to investigate the potential of relaying or cooperative communications at the light of a holistic analysis of power consumption.

4 Docitive radios & networks

4.1 Introduction

As already evidenced throughout this paper, cognitive radios and networks are perceived as a facilitator for improved efficiency of scarce spectrum access and management. Cognition, from *cognoscere = to know* in Latin, is typically defined as "a process involved in gaining knowledge and comprehension, including thinking, knowing, remembering, judging, and problem solving" [22]. Cognition has been the focus of numerous disciplines in the past, such as biology, biomedicine, telecommunications, computer science, etc. Across all these domains, emphasis has clearly been on a certain degree of intelligence which allows a cognitive system "to work properly under conditions it was initially not designed for" [23].

Said intelligence is typically accomplished by profoundly sensing the surroundings of the cognitive node and learning upon the acquired information [8]. This learning process is often a lengthy and complex process in itself, with complexity increasing with an increasing observation space. It is however needed to truly realize a cognitive radio as otherwise only opportunistic access is guaranteed at best. And whilst cognition and learning have received a considerable attention from various communities in the past, the process of knowledge transfer, i.e. teaching, however has received fairly little attention to date. This contribution thus aims at introducing

a novel framework referred to as docitive radios, from *docere = to teach* in Latin, which relates to radios (or general entities) which teach other radios. These radios are not (only) supposed to teach them the end-result (e.g. in form of "the spectrum is occupied") but rather elements of the methods to getting there. This concept mimics well our society-driven pupil-teacher paradigm and is expected to yield significant benefits for cognitive and thus more efficient network operation. Important and unprecedented questions arise in this context, such as which information ought to be taught, what is the optimum ratio between docitive and cognitive radios, etc.

An illustrative example, which will be corroborated in more depth in a subsequent section, models a cognitive radio system as a multi-agent system, where the radios learn through the paradigm of multi-agent learning. A typical learning mechanism for single agent systems is Q-Learning, belonging to the class of Reinforcement Learning. When it comes to multi-agent systems, Q-learning can be adapted to this setting, by implementing decentralized Q-learning. In this case, each node has to build a state-action space where it needs to learn the optimal policy for taking actions in each state. Depending on the dimension of the state-action space, the training process may be extremely time consuming and complex. However, if nodes are instructed to learn some disjoint or random parts of the state-action space, then they can share the acquired knowledge with their neighboring nodes. This facilitates learning but does not yield the end-result per se.

4.2 Brief taxonomy

A high-level operational cycle of docitive radios is depicted in Fig. 1. It essentially extends the typical cognitive radio cycle [8] through the docitive teaching element, where each of these elements typically pertains to the following high-level issues:

- *Acquisition.* The acquisition of data is quintessential in obtaining sufficient information of the surrounding environment. This data can be obtained by means of numerous methods, such as sensing performed by the node itself and/or in
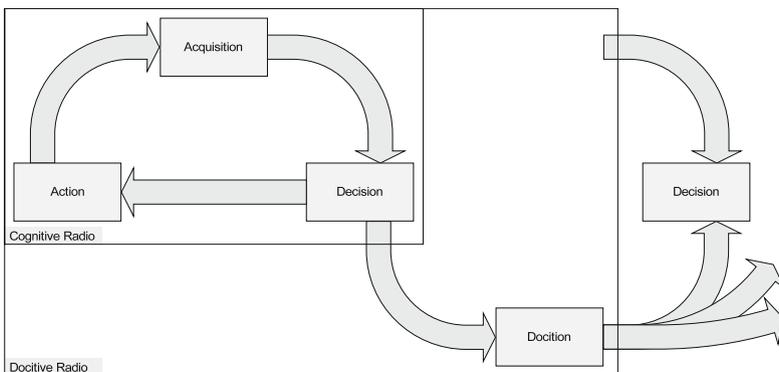
Fig. 1. Docitive cycle which extends the cognitive cycle through the teaching element

conjunction with spatially adjacent cooperating nodes; docitive information from neighboring nodes; environmental/docitive information from databases; etc.

- *(Intelligent) Decision*. The core of a cognitive radio is without doubt the intelligent decision engine which learns and draws decisions based on the provided information from the acquisition unit. The majority of cognitive devices today run some simple opportunistic decision-making algorithms; however, some more sophisticated learning and decision algorithms in form of e.g. unsupervised, supervised or reinforcement learning are available too.
- *Action*. With the decision taken, an important aspect of the cognitive radio is to ensure that the intelligent decisions are actually carried out, which is typically handled by a suitably reconfigurable software defined radio (SDR), some policy enforcement protocols, among others.
- *Docition*. An extension of the cognitive networking part is realized by means of an entity which facilitates knowledge dissemination and propagation, where so far rather end results have been shared (e.g. through cooperative sensing). A significant and non-trivial extension to this docitive paradigm comprises dissemination of information which facilitates learning.

4.3 Docitive example: Wireless multi-agent systems

We subsequently exemplify the operation of a docitive radio by means of wireless multi-agent learning systems. To use prior notion from the machine learning community, we will use the concept of agents which are defined as a computational mechanism that exhibits a high degree of autonomy performing actions, based on information from the environment. As a result of that, a multi-agent system is a complex system where multiple agents interact with one another, where the actions of each agent have impact on the environment of the others and where each agent has only partial information of the overall system [24]. The cognitive radio scenario can be easily mapped onto a multi-agent system [25], since it consists of multiple intelligent and autonomous agents, i.e. the cognitive radios, with the following characteristics: 1) the aggregated interference on a primary receiver depends on the independent decisions made by the multiple agents; 2) there is no central entity in charge of providing a global control of interference at the primary receivers coming from the multiple cognitive radios, so that the system architecture is decentralized; 3) a solution based on a centralized agent would not be scalable with respect to the number of cognitive radios; 4) the data based on which the cognitive radio system makes decisions about resource allocation come from spatially distributed sources of information and the decision making process is asynchronous for the multiple agents; 5) the individual decisions of each agent have to be self-adaptive depending on the decisions made by the other agents and on the surrounding environment. The above mentioned self-adaptation has to be achieved progressively by directly interacting with the environment and by properly utilizing the past experience, which is obtained through real-time operations. As a result, the common objective of the multiple agents is to distributively learn an optimal policy to achieve a common objective.

In case of single agent, the environment can be modeled as Markov Decision Process (MDP) [26], which is a tuple $\langle S, A, T, R \rangle$, where S is a finite set of environment states, A is a set of actions, $T : S \times A \times S \rightarrow [0, 1]$ is the Markovian transition function that describes the probability $p(s'|s, a)$ of ending up in state s' when performing action a in state s, and $R : S \times A \rightarrow \mathbb{R}$ is a reward function that returns the reward obtained after taking action a in state s. An agent's policy is defined as a mapping $\pi : S \rightarrow A$. The objective is to find the optimal policy π^* that maximizes the expected discounted future reward $U^*(S) = max_\pi \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t)|\pi, s_0 = s\right]$, for each state s, and where s_t indicates the state at time step t. The expectation operator averages over reward and stochastic transitions and $\gamma \in [0, 1)$ is the discount factor. We can also represent this using Q-values, which store the expected discounted future reward for each state s and possible action a:

$$Q^*(s, a) = R(s, a) + \gamma \sum_{s'} p(s'|s, a) \max_{a'} Q(s', a'), \qquad (11)$$

where a' is the action to take in state s'. The optimal policy for a state s is the action $\arg\max_a Q^*(s, a)$ that maximizes the expected future discounted reward. Reinforcement learning can be applied to estimate $Q^*(s, a)$; in particular, Q-learning is a widely used reinforcement learning method when the transition model is unavailable [27, 28]. This method starts with an initial estimate $Q(s, a)$ for each state action pair. When an action a is taken in state s, reward $R(s, a)$ is received and the next state s' is observed, the corresponding Q-value is updated by:

$$Q(S, a) = Q(s, a) + \alpha \left[R(s, a) + \gamma \cdot \max_{a'} Q(s', a') - Q(s, a) \right], \qquad (12)$$

where $\alpha \in (0, 1)$ is an appropriate learning rate. Q-learning is known to converge to the optimal $Q^*(s, a)$.

In case of multi-agent systems, all knowledge is not available locally in a single agent, but relevant knowledge, such as training experience and background information, is distributed among the agents within the system. In this case, we talk about distributed reinforcement learning, or multi-agent learning. The problem is how to ensure that individual decisions of the agents result in jointly optimal decisions for the group and how to reliably propagate this information over the wireless channel to spatially adjacent nodes. In principle, it should be possible to treat a multi-agent system as a single agent with complete information about the other agents, and learn the optimal joint policy using single-agent reinforcement learning techniques. However, both the state and action space scale exponentially with the number of agents, rendering this approach infeasible for most problems. Alternatively, we can let each agent learn its policy independently of the other agents, but then the transition model depends on the policy of the other learning agents, which may result in oscillatory behavior. This introduces game-theoretic issues to the learning process, which are not yet fully understood [29].

Contributions in literature [30] suggest that the performances of a multi-agent system can be improved by using cooperation among learners in a variety of ways.

In fact, it can be assumed that each agent does not need to learn everything by its own discover, but can take advantage of the exchange of information and knowledge with other agents or with more expert agents, thus leading to a teaching paradigm. It is demonstrated in [30] that if cooperation is done intelligently, each agent can benefit from other agents' information. Depending on the degree of cooperation among agents, we propose to consider the following cases for future studies:

- *Independent learners*. The agents do not cooperate, ignore the actions and rewards of the other agents in the system and learn their strategies independently. The standard convergence proof for Q-learning, in case of single agent system, does not hold in this case, since the transition model depends on the unknown policy of the other learning agents. In particular, each agent's adaptation to the environment can change the environment itself in a way that makes the other agents' adaptations invalid. Despite that, this method has been applied successfully in multiple cases.
- *Cooperative learners sharing state information*. The agents follow the paradigm of independent learning, but can share instantaneous information about their state. It is expected that sharing state information is beneficial in case that it is relevant and sufficient for learning.
- *Cooperative learners share policies or episodes*. The agents follow the paradigm of independent learning, but can share information about sequences of state, action and reward and learned decision policies corresponding to specific states. These episodes can be exchanged either among peers, or with more expert peers, i.e. teachers. It is expected that such cooperative agents can speed up learning, measured by the average number of learning iterations, and reduce the time for exploration, even though the asymptotic convergence can be reached also by independent agents.
- *Cooperative learners performing joint tasks*. Agents can share all the information required to cooperatively carry out a certain task. In this case the learning process may be longer, since the state-action space is bigger, but oscillatory behaviors are expected to be reduced.
- *Team learners*. The multi-agent system can be regarded as a single agent in which each joint action is represented as a single action. The optimal Q-values for the joint actions can be learned using standard single-agent Q-learning. In order to apply this approach, a central controller should model the MDP and communicate to each agent its individual actions, or all agents should model the complete MDP separately and select their individual actions. In this case, no communication is needed between the agents but they all have to observe the joint action and all individual rewards. The problem of exploration can be solved by using the same random number generator and the same seed for all agents. Although this approach leads to the optimal solution, it is infeasible for problems with many agents, since the joint action space, which is exponential with the number of agents, becomes intractable.

From the above, we incur that the concept of joint learning has received attention in recent years in the machine learning and artificial intelligence community; how-

ever, its application to cognitive radios operating primarily over a wireless broadcast channel has not been addressed by any study yet and, coupled with the potential gains, essentially inspired the concept of docitive radios.

4.4 Vision and challenges

Docitive radios and networks emphasize on the teaching mechanisms and capabilities of cognitive networks, and are understood to be a general framework encompassing prior and emerging mechanisms in this domain. Whilst the exchange of end-results among cooperatively sensing nodes has been explored in the wireless communication domain and the joint learning via exchange of states has been known in the machine learning community, no viable framework is available to date which quantifies the gains of a docitive system operating in a wireless setting. Numerous problems hence remain, some of which are listed below:

- *Information theory*. One of the core problems is how to quantify the degree of intelligence of a cognitive algorithm. With this information at hand, intelligence gradients can be established where docition should primarily happen along the strongest gradient. This would also allow one to quantify the tradeoff between providing docitive information versus the cost to deliver it via the wireless interface. Some other pertinent questions are how much information should be taught, can it be encoded such that learning radios with differing degrees of intelligence can profit from a single multicast transmission, how much feedback is needed, how often should be taught, etc?
- *Wireless channel*. Whilst not vital to the operation of docitive engines, it is of importance to quantify the coherence times of the wireless medium. This, in turn, allows one to estimate whether the channel dynamics allows for sufficient time for the knowledge dissemination and propagation.
- *PHY layer*. At this layer, as well as all OSI layers above, a pertinent question is which of the states should be learned individually, and which are advantageously taught? Another open issue is how much rate/energy should go into docition versus cognition?
- *MAC layer*. Open challenges relate to the problematic of optimal broad/multicast protocols which allow a single docitive radio to disseminate to as many as possible cognitive entities, all of which could have a different degree of intelligence.
- *Docitive system*. At system level, numerous questions remain open, such as what is the optimal ratio of docitive versus cognitive entities; what is the optimal docition schedule; should every cognitive entity also be a docitive one; what is the docition overhead versus the cognitive gains; etc.
- *Distributed learning*. More specifically to docition, scalability is a problem for many learning techniques and especially for multi-agent learning. The dimension of the search space grows rapidly with the number and complexity of agent behaviors, the number of agents involved and the size of the network of interactions among them. In addition to that, multi-agent systems are typically dynamic environments where the agents learn and the adaptation to one another changes the

environment itself. For this co-adaptation of learners the literature has recently focused on demonstrating the achievement of suboptimal Nash equilibriums, but the convergence to optima is still a wide-open issue.

We believe that we just touched the tip of an iceberg as preliminary investigations have shown that docitive networks are a true facilitator for utmost efficient utilization of scarce resources and thus an enabler for emerging as well as unprecedented wireless applications.

5 Conclusions

Cognitive and cooperative wireless networks are addressed in this section, providing a view on efficient ways of setting networks.

When addressing Cognitive Radio Networks, several aspects of spectrum usage and management are discussed. A number of techniques to be developed for the implementation of efficient spectrum usage through cognitive radio networks are dealt with: spectrum sensing, spectrum management, spectrum mobility, and spectrum sharing mechanisms.

Cognitive Positioning is then addressed, in relation with cognitive radio, Multi Carrier systems being taken as an example. Signal-intrinsic capability for localisation is discussed, namely on the criteria to optimise the function of ranging, and on the characteristics of the signals.

A brief discussion on cooperative networks follows. Colocation of base stations and MIMO relay schemes are among the topics listed in the subsection.

Finally, the concept of Docitive Radios & Networks is introduced, i.e., a novel framework on radios and networks that teach other radios and networks. These radios and networks are not (only) supposed to teach them the end-result, but rather elements of the methods to getting there. A taxonomy is presented, together with an example, a vision, and challenges.

References

1. Sanders F.H.: Broadband spectrum surveys in Denver, CO, San Diego, CA, and Los Angeles, CA: Methodology, analysis, and comparative results. In: Proc. IEEE Int. Symp. Electromagnetic Compatibility, Rome, Italy (1998).
2. McHenry M.A., Tenhula P.A., McCloskey D., Roberson D.A., Hood C.S.: Chicago spectrum occupancy measurements & analysis and a long-term studies proposal. In: Int. W. Technology and Policy for Accessing Spectrum, Boston, MA (2006).
3. Petrin A., Steffes P.G.: Analysis and comparison of spectrum measurements performed in urban and rural areas to determine the total amount of spectrum usage. In: Proc. Int. Symp. Advanced Radio Technologies, Boulder, CO (2005).
4. Chiang R.I.C., Rowe G.B., Sowerby K.W.: A quantitative analysis of spectral occupancy measurements for cognitive radio. In: Proc. IEEE Veh. Technol. Conf., Dublin, Ireland (2007).
5. Wellens M., Wu J., Mähönen P.: Evaluation of spectrum occupancy in indoor and outdoor scenario in the context of cognitive radio. In: Proc. Int. Conf. Cognitive Radio Oriented Wireless Networks and Communications, Orlando, FL (2007).

6. López-Benítez M., Casadevall F., Umbert A., Pérez-Romero J., Hachemani R., Palicot J., Moy C.: Spectral occupation measurements and blind standard recognition sensor for cognitive radio networks. In: Proc. Int. Conf. Cognitive Radio Oriented Wireless Networks and Communications, Hannover, Germany (2009).

7. Federal Communications Commission (FCC), Notice of Proposed Rule Making, ET Docket no. 04-113 (2004).

8. Mitola J. III, Maguire G.Q.: Cognitive radio: Making software radios more personal. IEEE Pers. Commun. **6**(4), pp. 13–18 (1999).

9. Akyildiz I.F., Lee W.-Y., Vuran M.C., Mohanty S.: Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey. Comput. Networks **50**(13), pp. 2127–2159 (2006).

10. Bourse D., et al.: The E2R II Flexible Spectrum Management (FSM) – Framework and Cognitive Pilot Channel (CPC) Concept – Technical and Business Analysis and Recommendations. E2R II White Paper (2007).

11. Alberty T.: Frequency domain interpretation of the Cramér-Rao bound for carrier and clock synchronization. IEEE Pers. Commun. **43**(2-3-4), pp. 1185–1191 (1995).

12. Celebi H., Arslan H.: Cognitive positioning systems. IEEE Trans. Wireless Commun. **6**(12), pp. 4475–4483 (2007).

13. Cramér H.: Mathematical methods of statistics. Princeton Univ. Press, Princeton, NJ (1946).

14. D'Andrea A.N., Mengali U., Reggiannini R.: The modified Cramér-Rao bound and its application to synchronization problems. IEEE Pers. Commun. **42**(2-4), pp. 1391–1399 (1994).

15. Kay S.M.: Fundamentals of statistical signal processing: Estimation theory. Prentice-Hall, Englewood Cliffs, NJ (1993).

16. Mallinckrodt A., Sollenberger T.: Optimum pulse-time determination. IRE Transactions PGIT-3, pp. 151–159 (1954).

17. Mengali U., D'Andrea A.N.: Synchronization Techniques for Digital Receivers. Plenum Press, New York, NY (1997).

18. Skolnik M.: Introduction to radar systems. McGraw-Hill, New York, NY (1980).

19. Van Trees H.L.: Detection, Estimation and Modulation Theory. Wiley, New York, NY (1968).

20. Lottici V., Luise M., Saccomando C., Spalla F.: Non-data aided timing recovery for filter-bank multicarrier wireless communications. IEEE Trans. Commun. **54**(11), pp. 4365–4375 (2006).

21. Luise M., Zanier F.: Multicarrier signals: A natural enabler for cognitive positioning systems. In: Proc. Int. Workshop on Multi-Carrier Systems & Solutions (MC-SS), Herrsching, Germany (2009).

22. (About.com definitions) Cognition – What is cognition? [Online]. Available: http://psychology.about.com/od/cindex/g/def_cognition.htm

23. Private communication with Apostolos Kontouris, Orange Labs, France (2007).

24. Sycara K.P.: Multiagent systems. AI Magazine **19**(2), pp. 79–92 (1998).

25. Galindo-Serrano A., Giupponi L.: Aggregated interference control for cognitive radio networks based on multi-agent learning. In: Proc. Int. Conf. Cognitive Radio Oriented Wireless Networks and Communications, Hannover, Germany (2009).

26. Bellman R.: Dynamic Programming. Princeton Univ. Press, Princeton, NJ (1957).

27. Harmon M.E., Harmon S.S.: Reinforcement learning: A tutorial (2000).

28. Watkins J., Dayan P.: Technical note: Q-learning. Machine Learning **8**, pp. 279–292 (1992).

29. Hoen P., Tuyls K.: Analyzing multi-agent reinforcement learning using evolutionary dynamics. In: Proc. European Conference on Machine Learning (ECML), Pisa, Italy (2004).

30. Tan M.: Multi-agent reinforcement learning: Independent vs. cooperative learning. In: Huhns, M.N., Singh M.P. (Eds.): Reading in Agents. Morgan Kauffman, San Francisco, CA, pp. 487–494 (1993).