Intelligent Discovery of the Capabilities of Reconfiguration Options in a Cognitive Wireless B3G Context

K. Demestichas*, E. Adamopoulou, M. Theologou

National Technical University of Athens,

9 Iroon Polytechniou Str., Zographou GR-157 73, Athens, Greece

Abstract. Beyond 3G (B3G) wireless connectivity can efficiently be realized by exploiting cognitive networking concepts. Cognitive systems dynamically reconfigure the Radio Access Technologies (RATs) and the spectrum they use, based on experience, in order to adapt to the changing environment conditions. However, dynamic reconfiguration decisions call for robust discovery, i.e. radio-scene analysis and channel identification schemes. This paper intends to contribute in the areas of radio-scene analysis and channel identification: firstly, by providing an overview of interference estimation methods, and explaining how capacity estimations can be derived based on the measured interference levels; secondly, by specifying the information flow for the radio-scene analysis process of a cognitive radio system; and thirdly, by enhancing the above with a learning system, which is essential for obtaining a truly cognitive process. The proposed approach lies in the introduction of a robust probabilistic model for optimal prediction of the capabilities of alternative configurations, in terms of capacity.

Keywords. B3G wireless infrastructures, Capacity estimation, Cognitive radio, Learning and adaptation, SINR estimation

1. INTRODUCTION

Today's wireless access landscape comprises various access technologies: (a) 2G and 3G cellular systems (e.g., GSM, UMTS); (b) wireless access networks of different range capabilities (WLANs/WMANs/WPANs); and (c) broadcast networks, such as Digital Audio Broadcasting (DAB) and Digital Video Broadcasting (DVB). The evolution of the abovementioned wireless communication systems over the past years demonstrates a clear trend towards architectures which will support multiple access technologies, and multimode mobile terminal devices, i.e. capable of alternatively operating in the diverse radio segments available in the infrastructure. This trend is often referred to as 'systems beyond 3G' (B3G), and its main notion is that a network operator can rely on multiple Radio Access Technologies (RATs) for achieving the desired capacity and Quality of Service (QoS) levels, in a cost efficient manner [1]-[2].

A major issue within the B3G context is that the electromagnetic radio spectrum is a limited natural resource, w[*]hich is currently underutilized due to its fragmentation into separate Radio Frequency (RF) bands [3]-[5]. This fragmentation leads to the over-utilization of some frequency bands, whereas others remain largely unoccupied. Thus, there is need for the development of a robust spectrum management scheme, capable of exploiting available frequency bands as efficiently as possible.

This can be achieved by detecting and exploiting the so-called 'spectrum holes', namely the 'white' and 'grey' spaces of the RF spectrum [11]. By using the term 'white spaces', we refer to portions of the RF spectrum that are free of RF interferers, except for ambient noise. Portions that are partially occupied by low-power interferers are referred to as 'grey spaces', while portions that are occupied by high-power interferers are known as 'black spaces'.

The detection and exploitation of spectrum holes raises a great engineering challenge that can be confronted by 'Cognitive Radio' [6]-[10], or a cognitive wireless network in general. Cognitive radio systems are systems capable of sensing and learning the RF environment, and also easily adjusting to current conditions, by making intelligent decisions on how to best utilize spectrum.

As part of the adaptation to the changing conditions, at the physical and MAC layers, B3G cognitive radios may consist of elements (hardware components, such as transceivers) that dynamically change the RATs they operate and the spectrum they use, in order to improve capacity and QoS levels. In other words, a hardware component (transceiver) will be able to change either the RAT in use or the selected frequency or even both, in order to adapt to new conditions and requirements.

As an example of this, let us assume a B3G service area (i.e., a cell) comprising three reconfigurable elements (transceivers) e_1, e_2 and e_3, each one capable of alternatively operating in one of three RATs, namely r_1, r_2 and r_3 (e.g., UMTS, WLAN and WiMax, respectively). During a specific time period, the best configuration, in terms of capacity, coverage, mobility support and cost, might be configuration (r_1, r_3, r_2). However, environmental conditions and requirements change dynamically, so in a next time period configuration (r_1, r_1, r_2) might be the most suitable one. In this example, element e_2 was reconfigured to operate in RAT r_1 rather than r_3. Another possibility of reconfiguration is the use of a different carrier frequency, either maintaining the same RAT or not.

[*] Corresponding author. cdemest@cn.ntua.gr, Tel. +30 210 772 1493, Fax. +30 210 772 2530

As stated in [6] and [11], three fundamental cognitive tasks, tightly interconnected, are identified within the framework of cognitive radio.

(a) Radio-scene analysis, which involves tuning to a frequency and measuring the interference levels perceived.

(b) Channel identification, which encompasses channel capacity estimation, based on the levels of interference measured in (a).

(c) Transmit-power control and dynamic spectrum management.

This paper adopts the general terminology that cognitive tasks (a) and (b) can be referred to by the term *discovery*. Discovery is targeted to the identification of capabilities, e.g. capacity and coverage, of alternate configurations of a reconfigurable element. This paper intends to contribute in the areas of discovery (radio-scene analysis and channel identification):

(a) by providing an overview of interference estimation methods, and explaining how capacity estimations can be derived based on the measured interference levels;

(b) by specifying the information flow for the radio-scene analysis process of a cognitive radio system;

(c) by enhancing the above with a learning system, which is essential for obtaining a truly cognitive process. The proposed approach lies in the introduction of a robust (stable and reliable) probabilistic model for optimal prediction of the capabilities of alternative configurations, in terms of capacity. The latter can then serve as input for reaching optimal configuration decisions, within the cognitive radio framework. The overall idea is depicted in Figure 1.

The remainder of this paper is structured as follows: Section 2 presents the channel quality metrics that can be used for radio-scene analysis, namely the signal-to-interference-plus-noise ratio and the interference temperature. Section 3 is an overview of interference sensing methodology. Section 4 identifies how to derive capacity estimations from the measured interference levels. Section 5 describes the message exchange sequence for the radio-scene analysis process of a cognitive radio system. Section 6 introduces the proposed probabilistic model for achieving robust capacity estimation. Section 7 provides results that demonstrate the efficiency of the proposed model, and section 8 concludes the paper.

2. CHANNEL QUALITY METRICS FOR RADIO-SCENE ANALYSIS

Interference is one of the most constraining factors that wireless communication systems have to deal with, since it is responsible for limiting both the capacity and the coverage area of such a system. By interference we mean any undesired signals perceived by the receiver [12]. It can either be due to natural causes, such as thermal noise in the receiver, or be system related. The latter case occurs when other senders, other than the user, are also emitting energy in the same frequency band.

Since interference is inherent to all radio communications, there is need to utilize an appropriate metric reflecting its effects in signal transmission. The signal-to-interference-plus-noise ratio (SINR) and, more recently, the interference temperature (IT) constitute the proposed channel quality metrics in radio-scene analysis.

SINR

A suitable measure of the received signal's quality [12] is represented by the ratio between the received wanted carrier signal power and the total received interference power, as depicted in the following equation:

$$\frac{S}{I} = \frac{S}{\left(\sum_n I_n + N\right)} \qquad (1)$$

where S denotes the desired signal's power, I_n the interference power originating from interferer n, and N the perceived thermal noise. The aforementioned ratio is known as SINR.

SINR estimation is of particular importance for wireless access systems, where resources are shared dynamically amongst users. Sample applications of SINR estimates are: (a) power control in CDMA systems; (b) data rate adaptation; (c) hand-off decisions; (d) dynamic channel allocation; and (e) diversity combining.

In the cognitive radio context, the need to estimate SINR is equally important, since the SINR is the basis for, firstly, spectrum holes detection and, secondly, capacity estimation. In more detail, a high SINR allows for less redundancy (i.e., parity symbols for error detection and correction) and more efficient modulation schemes, leading to an increase of the capabilities of wireless systems in terms of capacity. What is more, a specific SINR must be maintained at the receiver to guarantee a sufficient QoS.

Interference temperature

Aside from SINR, an alternative metric called interference temperature (IT) may prove to be considerably useful in radio-scene analysis. The notion of interference temperature has been introduced fairly recently by the Federal Communications Commission (FCC), for "quantifying and managing interference" [22]-[24]. The idea is to regulate received power rather than transmitted power [23], thus yielding a paradigm shift away from the currently applied 'transmitter-centric' model, according to which the transmitted power is designed to approach a prescribed noise floor at a certain distance from the transmitter [11].

According to the FCC's model, for a given frequency band in a given geographic location, an 'interference temperature limit' is defined by some regulatory agency. A transmitter should then ensure that by transmitting it does not raise the current IT above the specified IT limit.

For an antenna, given a bandwidth B in Hertz, interference power I in Watts and Boltzmann's constant $k = 1.38 \cdot 10^{-23}$ Joules per Kelvin degree, the interference temperature T_I in Kelvin degrees is defined as:

$$T_I = \frac{I}{kB} \qquad (2)$$

Let T_L be the IT limit in the geographic region and frequency band under consideration. The difference $T_L - T_I$ will serve as a margin or 'cap' indicating the potential RF energy that could be further introduced into the specified band. Thus, before using the frequency band, it must be guaranteed that the values of transmit power P_S and bandwidth B will not raise the IT over the IT limit, i.e.:

$$P_S / kB + T_I \le T_L \qquad (3)$$

3. INTERFERENCE SENSING METHODOLOGY FOR RADIO-SCENE ANALYSIS

Radio-scene analysis focuses on SINR estimation or, alternatively, IT estimation. The former is a somewhat complex issue that has been dealt with to some extent in the literature, while the latter is relatively more straightforward.

SINR estimation

In SINR estimation, inaccuracies may arise due to the following reasons [13]: (a) interference is additive to the desired signal and cannot be separated; (b) the observation window for estimates is rather small, especially for power-control applications; and (c) the

presence of modulation poses difficulties in directly estimating the SINR from the data symbols.

Over the years, the SINR estimation problem has been discussed, in the literature, in some detail (e.g., [14]-[19]). It has been studied both for analog communication systems (e.g., AMPS) and, more recently, for digital Time Division Multiple Access (TDMA), Code Division Multiple Access (CDMA), and Orthogonal Frequency-Division Multiplexing (OFDM) based systems.

In general, two strategies can be employed for achieving SINR estimations. The first is based on the transmission of pilot symbols (training sequences), whereas the second tries to derive the channel characteristics directly from the data symbols, i.e. the received signal, without the use of training sequences. The two strategies are often referred to in the literature as 'non-blind' and 'blind', respectively.

A training sequence is a priori known to the receiver, thus the task of SINR calculation is made easier, since the receiver knows which symbols it is supposed to receive. Consequently, the use of training sequences allows for greater accuracy, but also introduces a significant overhead, which could be used instead for the transmission of additional data sequences [20]. *In the cognitive radio context*, *the interference estimation part of the discovery procedure can be based solely on pilot transmission*, since we are interested in the potential capabilities of alternate configurations. On the other hand, of course, interference estimation in the frequency band that is currently used for service delivery can follow either of the two strategies.

In the sequel, a widely used non-blind estimation method, known as Minimum Mean Square-Error method (MMSE) [21], will be explained. Let $\mathbf{a} = \{a_1, a_2, ..., a_L\}$ be a training sequence, consisting of L symbols, and $\mathbf{y} = \{y_1, y_2, ..., y_L\}$ be the vector of the received symbols, as depicted in Figure 2. The latter can be expressed as:

$$\mathbf{y} = f\mathbf{a} + \mathbf{w} \qquad (4)$$

where f stands for the attenuation factor, which is assumed constant for the duration of the training sequence, and $\mathbf{w} = \{w_1, w_2, ..., w_L\}$ is the Gaussian noise vector. Hence, the SINR is given by the following expression:

$$SINR = \frac{|f\mathbf{a}|^2}{|\mathbf{w}|^2} = \frac{|f\mathbf{a}|^2}{|\mathbf{y} - f\mathbf{a}|^2} \qquad (5)$$

From (5), it is clear that an estimate for the attenuation factor is needed. According to the MMSE method, the estimation error and the estimate are orthogonal, as follows:

$$(\mathbf{y} - \hat{f}\mathbf{a})(\hat{f}\mathbf{a})^H = 0 \Rightarrow \hat{f} = \frac{\mathbf{y}\mathbf{a}^H}{|\mathbf{a}|^2} \qquad (6)$$

where the superscript H denotes the conjugate transpose. Due to (5) and (6), an SINR estimation is given by:

$$\hat{SINR} = \frac{|y\mathbf{a}^H|^2}{|\mathbf{a}|^2 |\mathbf{y}|^2 - |y\mathbf{a}^H|^2} \qquad (7)$$

Interference temperature estimation

IT estimation is rather straightforward, compared to SINR estimation, as it only involves tuning to the frequency of interest and calculating the received signal energy. However, a significant issue related to the estimation of the current IT is that an IT measurement taken at the transmitter might differ from one taken at the receiver. Thus, in the worst case, the current IT at the receiver's location might be near the IT limit, whilst the current IT at the transmitter's location might be significantly lower, leading the transmitter to wrongly declare the frequency band as usable. Although this unintentional interference is rather rare and small, yet future work will probably need to be conducted, in order to quantify it [23].

As can be deduced from this, a reliable estimation of the current IT can be a challenging task. A more systematic approach is framed in [11], according to which a reliable spectral estimate of the IT can be achieved by: (a) using the 'multitaper' spectral estimation procedure to estimate the power spectrum of the IT; and (b) using a large number of sensors to take measurements from the RF environment.

The method in (a) is based on the use of multiple orthogonal windows ('tapers'), for which more details can be found in [25]-[27]. Considering (b), the use of multiple sensors is feasible in an indoor environment, but may have to be restricted to a single sensor for outdoor environments. In the cognitive radio context, a group of user terminals may undertake the sensing task, i.e. sensing the RF environment and sending the measurements to the reconfigurable element. The latter topic will be further discussed in Section 5.

4. CAPACITY ESTIMATION

SINR estimation and/or IT estimation can consecutively be used, firstly, to identify if the candidate frequency band is usable and, secondly, to determine the capabilities of the corresponding configuration, in terms of capacity. The first task is accomplished in a relatively straightforward manner. If the measured SINR is above a specified threshold, or the measured IT is below the IT limit, then the frequency band is probably usable. Obviously, the measured IT should be significantly lower than the IT limit, and relation (3) should hold, so as to be possible to exploit the band. In [11], [28], a more systematic approach for reliable spectrum holes detection, which is based on the multitaper method combined with singular value decomposition, is described. The remainder of this section elaborates on capacity estimation.

As a general rule, if the measured SINR is high or, equivalently, the measured IT is low, this allows for the employment of a more efficient (in terms of achievable capacity) modulation strategy (i.e., larger number of bits per symbol) and probably less redundancy (i.e., smaller number of parity bits). However, capacity estimation, on the basis of SINR measurement, is not a trivial matter. Nonetheless, reliable results can be reached by employing any of the following strategies: (a) the Shannon-Hartley theorem; (b) the attainable data rate formula; or (c) the bit error probability diagrams of the candidate modulation schemes. Method (a) derives information-theoretic bounds on capacity, for given bandwidth and power constraints. Method (b) goes one step further, by taking into account the constraints of a physically realizable system. Method (c) illustrates how to select the most appropriate and efficient modulation scheme, depending on the circumstances. The aforementioned methods will be described concisely in what follows.

(a) Shannon-Hartley theorem: The concept of capacity estimation may be clarified by utilizing the Shannon-Hartley theorem, as implied in [11] and demonstrated in [24]. According to the celebrated theorem, given the bandwidth B in Hz, and the measured value of $SINR$, capacity C (in bits/s) is related as follows:

$$C = B\log_2(1+SINR) \qquad (8)$$

Alternatively, given the measured value T_I of IT, (5) becomes:

$$C = B\log_2\left(1+\frac{S}{I}\right) = B\log_2\left(1+\frac{LP_S}{kBT_I}\right) \qquad (9)$$

where P_S is the transmit power in Watts and $S = LP_S$ is the power at the receiver's location, with $L \in (0,1)$ representing the signal attenuation due to path loss and shadowing.

(b) Attainable data rate formula: In practice, a physically realizable encoding system must transmit data at a rate R (in bits/s) less than the maximum possible rate C (as given by (8) and (9)) for it to be reliable. Thus, for an implementable system operating at low enough probability of symbol error, a measure called 'signal-to-noise ratio gap', or just 'gap', is utilized [29]. The gap is denoted by Γ and is a function of the permissible probability of symbol error P_e and the encoding system of interest. It provides a measure of the efficiency of an encoding system with respect to the ideal transmission system of equation (8) or (9). The gap is defined as follows:

$$\Gamma = \frac{2^{C/B} - 1}{2^{R/B} - 1} = \frac{SINR}{2^{R/B} - 1} \qquad (10)$$

Hence, equivalently it holds that:

$$R = B \log_2 \left(1 + \frac{SINR}{\Gamma} \right) \qquad (11)$$

As an example, for encoded PAM (Pulse Amplitude Modulation) or QAM (Quadrature Amplitude Modulation) operating at $P_e = 10^{-6}$, the gap Γ is fixed to 8.8 dB. Through the use of codes (e.g., trellis codes or turbo codes), nonetheless, the gap Γ may be reduced to as low as 1 dB.

However, through the use of codes, the actual data rate R_I (in bits/s), i.e. the data rate of the information source, is lower than that given by (11), due to the presence of redundant bits. For a code that uses n-bit code words consisting of k message bits and n-k redundant bits, R_I is given by:

$$R_I = r \cdot R = \frac{k}{n} R \qquad (12)$$

where r represents the dimensionless ratio k/n and is known as the 'code rate'.

(c) Bit error probability diagrams: The goal of this method is to determine the most appropriate modulation scheme among the candidates. The proposed method to accomplish this stands as follows:

Step (i): Based on the measured SINR, the E_b/N_0 ratio, i.e. the energy per bit to noise power spectral density ratio, is determined as follows:

$$\frac{E_b}{N_0} = \begin{cases} \dfrac{1}{\log_2 M} SINR, & \text{M-ary PSK, M-ary QAM} \\ \dfrac{M}{\log_2 M} SINR, & \text{M-ary FSK} \end{cases} \quad (13)$$

Step (ii): Based on the value of E_b/N_0, the bit error probability P_b of each candidate modulation scheme is computed. Figure 3 is a plot of P_b vs. E_b/N_0 for uncoded M-ary FSK (Frequency Shift Keying), PSK (Phase Shift Keying) and QAM signals, in the case of coherent demodulation. Similar curves can be found in the literature for non-coherent demodulation. Also, if error control codes are used, different curves should be utilized, depending on the actual encoding scheme and code rate.

Step (iii): The bit error probability P_b of each candidate modulation scheme is compared to the value of the bit error probability threshold P_b^t (i.e., the maximum permissible bit error probability, e.g. $P_b^t = 10^{-5}$). For a candidate modulation scheme, if $P_b > P_b^t$, then the modulation scheme is rejected.

Step (iv): Among the acceptable modulation schemes, the one that provides the maximum bit rate R, given the bandwidth B in Hz, is selected, according to:

$$R = \begin{cases} B \log_2 M, & \text{M-ary PSK, M-ary QAM} \\ B \dfrac{\log_2 M}{M}, & \text{M-ary FSK} \end{cases} \quad (14)$$

In case that error control codes are in use, relation (12) should consecutively be used, in order to determine the bit rate of the information source.

5. INTERFERENCE SENSING PROCESS IN COGNITIVE RADIO SYSTEMS

In the cognitive radio context, a reconfigurable transceiver of a B3G service area needs to be able to sense the capabilities of alternative configurations. For this reason, a few timeslots during its operation should be devoted to the execution of sensing procedures. Thus, a realistic proposal is to break up a transceiver's operational time into 'service provision timeslots' (during which the transceiver *serves* the network traffic) and 'sensing timeslots' (during which the transceiver *senses* the interference levels of an alternative configuration), as depicted in Figure 4(a).

Assume that a reconfigurable transceiver is currently operating under RAT r_0 and is tuned to carrier frequency f_0. This configuration will be denoted henceforth as (r_0, f_0). In addition, let us assume that the transceiver needs to sense the interference level of an alternative

configuration *(r, f)*, where *(r, f)* ≠ *(r_0, f_0)*. Figure 4(b) and Figure 4(c) illustrate the interference sensing process on the uplink and downlink, respectively. The estimation of the SINR has to be carried out at the receiver, i.e. the transceiver on the uplink and the terminals on the downlink.

In the first case, the transceiver instructs a group of terminals within its service area to temporarily change their configuration *(i)*, and tune to *(r, f)*. The transceiver itself also changes its configuration. Once the tuning has been accomplished *(ii)*, each of the terminals transmits a training sequence back to the transceiver *(iii)*, in order for the latter to estimate the SINR *(iv)*. Once the sensing procedure is complete, both the terminals and the transceiver tune back to their previous configuration *(r_0, f_0)*. Message *(i)* is transmitted under the configuration *(r_0, f_0)*, while all other messages are exchanged under the configuration *(r, f)*.

In the second case, the process is initially similar *(i and ii)*. Then, a training sequence is transmitted by the transceiver to the temporarily reconfigured terminals *(iii)*. Next, the estimation of the SINR takes place at the terminal side *(iv)*, and the results are relayed back to the network side *(v)*.

The number of terminals instructed to temporarily reconfigure may vary. In the simplest case, the sensing process may have to be confined to a single terminal. Moreover, instead of SINR estimation, an interference temperature measurement process may take place. In this case, the transmission of a training sequence is not needed. This leads to a considerable simplification of the overall process, especially for measurements on the uplink.

6. THE PROBLEM OF ROBUST CAPACITY ESTIMATION

The need for robustness

Capacity estimation, on the basis of SINR and/or IT measurement, is a prerequisite for optimal configuration selection. However, the most recently measured (estimated) value of capacity is not necessarily the most accurate one. In practice, the estimated values can fluctuate (oscillate) due to *measurement errors*, as well as *temporary changes* in the environment. This yields the need for a more robust capacity estimation scheme.

An essential requirement for such a scheme is to favour the *autonomic* computing paradigm [31]-[32], which is of high importance for cognitive radio systems. According to this paradigm, each B3G service area (i.e., cell) decides autonomously about the most

appropriate configuration of its transceivers. This notion of 'self-configuration' is an efficient means for tackling complexity and scalability.

Within this framework, the capabilities of candidate configurations, in terms of capacity, can change over time, as they are influenced by the varying conditions in the environment, *especially* the behaviour of "near-by" reconfigurable elements. The goal of 'self-configuration' is to enable all elements to act in a completely distributed (autonomic) manner. This poses a significant engineering challenge: how to increase the degree of assurance that, by assigning a certain configuration *(r, f)* to a reconfigurable transceiver, the resulting capacity will be the expected one (e.g., *x* Mbps). A probabilistic model, as well as a learning and adaptation strategy, should be adopted. The resulting problem is: Given a specific candidate configuration *(r, f)*, how can the most probable value of the random variable "capacity" be predicted?

In the following, this problem is solved through a robust learning and adaptation strategy, based on Bayesian networks [33], which are valuable tools for learning and reasoning through probabilistic relationships [34]-**Error! Reference source not found.**. The solution does not violate the autonomy of network elements. In fact, no cooperation (e.g., no message exchange) between the different network elements is needed.

Formulation as a Bayesian network

Figure 5(a) depicts a basic Bayesian network that is proposed for modelling the specified problem. *CAP* is a random variable representing capacity. *CFG* is another random variable, representing a configuration, e.g. configuration *(r, f)* may be an instance of *CFG*. *CFG* is the Bayesian network's predictive attribute (node), while *CAP* is the target attribute.

The goal is the determination of the maximum value of the conditional probability $\Pr[CAP \mid CFG]$. Therefore, a conditional probability table (CPT) is organized, the structure of which is depicted in Figure 5(b). Each CPT refers to a particular RAT. Consequently, if R is the set of possible RATs, then $|R|$ CPTs in total are required for the full information. Each column of the CPT refers to a specific configuration (i.e., RAT and carrier frequency). Each line of the CPT corresponds to a capacity value. In this sense, a discrete set of potential capacity values is defined. Each cell (intersection of line and column) provides the probability (likelihood) that the configuration (corresponding to the column) will achieve the potential capacity value (corresponding to the line). Given a

configuration, the most probable value of capacity is the value that corresponds to the maximum conditional probability.

Figure 5(b) is an example for an arbitrary RAT r_1. With F_{r1} denoting the set of spectrum carriers with which RAT r_1 may operate, the CPT consists of $|F_{r1}|$ columns, corresponding to configurations $c_1=(r_1, f_1)$, …, $c_{|Fr1|}=(r_1, f_{|Fr1|})$, and m lines, corresponding to capacity values cap_1, cap_2, …, cap_m. Without loss of generality, enumeration is done in ascending order, i.e. $cap_1 < cap_2 < \ldots < cap_m$. In other words, cap_m is the maximum value. The cell at the intersection of line i and column j is a probability value. It expresses the likelihood that capacity cap_i will be achieved, given the fact that configuration c_j has been selected. Formally, this is denoted as $\Pr[CAP = cap_i \mid CFG = c_j]$.

Learning and adaptation process

In the previous subsection, it was defined that the capabilities of configurations, in terms of capacity, are modelled through the CPTs. The next step is to describe how to update the CPTs. Figure 5(c) is the general representation of the process. This learning and adaptation process yields the robust methods for discovering the performance capabilities of candidate configurations. The update process takes into account the measurements (estimations) of the cognitive radio system and, more specifically, the "distance" (absolute difference) between each probable value and the measured value.

Let us assume that measurements (obtained through the basic discovery-sensing functionality described in sections 3, 4 and 5) show that a specific configuration can achieve capacity cap_{meas}. This measurement can be exploited, in order to fine-tune (enhance or decrease) the values of the CPTs, so as to increase the degree of assurance of future predictions. Let dif_{max} be the maximum difference between the probable capacity values, i.e. $dif_{max} = cap_m - cap_1$.

Then, the following correction factor, cor_i, can be computed for each candidate capacity value cap_i:

$$cor_i = 1 - \frac{|cap_i - cap_{meas}|}{dif_{\max}} \qquad (15)$$

It holds that $0 \leq cor_i \leq 1$. A value close to 1 reflects that the corresponding candidate value cap_i is close to the measured value cap_{meas}, thus it should be reinforced accordingly. The opposite stands for a value that is close to 0.

The correction of the CPT values $\Pr[CAP = cap_i \mid CFG]$ can then be done as follows, for each candidate capacity value cap_i:

$$\Pr[CAP = cap_i \mid CFG]_{new} = N \cdot cor_i \cdot \Pr[CAP = cap_i \mid CFG]_{old} \qquad (16)$$

The parameter N is a normalizing factor whose value is computed by requiring all the "new" probabilities to sum up to 1.

The system *converges* when the most probable candidate capacity value (i.e., the one with the maximum probability) is being reinforced, while the probabilities of the other candidate capacity values are either being reduced or reinforced less. After convergence, we limit the number of consecutive updates that can be done on the probability values associated with each capacity value. This is done for assisting fast adaptation to new conditions. For the same reason, we do not allow that a probability falls under a certain threshold, a/m, where $0<a<1$ (m is the number of potential capacity values). In such cases, the normalization factor, N, is computed by requiring all the other "new" probabilities to sum up to $1-(k \cdot a/m)$, where k is the number of probabilities that are assigned equal to the threshold.

7. RESULTS

This section exhibits results on the efficiency of the robust discovery method. Three scenarios are realized. The first scenario aids in the comprehension of the proposed technique. The other two focus on the evaluation of the scheme's performance, concerning its adaptation speed when dealing with severe and permanent changes in the values of the measurements taken. In all the scenarios, parameter a has been set equal to 0.1.

Scenario 1: Simple example

This scenario helps the in-depth understanding of the proposed technique. Our focus is on an arbitrary configuration $c_1=(r_1, f_1)$. It is assumed that there are m=5 candidate capacity values (in Mbps): cap_1=0.5, cap_2=1.0, cap_3=1.5, cap_4=2.0, cap_5=2.5. Hence, dif_{max}=2 Mbps. It should be noted that a denser grid of candidate values could be selected (actually, in this case our results would have been favoured). What is more, the distance between two subsequent candidate values needs not be the same. Also, only three consecutive reinforcements are allowed, after convergence.

Figure 6(a)-(d) depicts the distribution of conditional probabilities in four cases. In each case cap_{meas} is (in Mbps) 1.2, 1.5, 2.1 and 0.75, respectively. The algorithm is applied in five phases of runs. Initially, the conditional probabilities are uniformly distributed, i.e. equal to

0.2, in all four cases (phase 1). By using (15), we calculate the correction factors. Then, by using (16), we compute the new (adjusted) conditional probabilities. The results for each case are further analyzed in the following.

Figure 6(a) shows that the model correctly and quickly adapts to the situation, by selecting cap_2 as the most probable value, in phase 2 (phase 1 represents the initial phase). As was expected, there are high values for cap_2 and cap_3, a slight diminishment for cap_1 and cap_4, and a severe degradation for cap_5. As the scheme is further applied, and since cap_{meas} does not change, the most probable value remains the same and is actually further reinforced. The learning and adaptation model accurately adapts to the second case also (Figure 6(b)), in which cap_{meas}=1.5. There is peak at cap_3, whereas cap_2 and cap_4 remain practically the same, and, finally, cap_1 and cap_5 suffer significant diminishment. Figure 6(c) shows the results from the third case, in which cap_{meas}=1.8. The model quickly adapts to cap_4, the probabilities that correspond to cap_3 and cap_5 are increased, while the probabilities of cap_1 and cap_2 drop. The model is also robust in the last, rather unlikely, case, in which cap_{meas}=0.75 (Figure 6(d)). It suggests cap_1 and cap_2 as the most likely values. The probability of cap_3 is slightly increased, while the values of cap_4 and cap_5 are degraded.

Scenario 2: Adaptation speed vs. number of consecutive reinforcements

The main goal of this scenario is to examine how many steps it takes for the scheme to adapt to a new situation, and more specifically to a sudden large change in the environment conditions. Figure 7(a)-(b) shows the speed of the adaptation when there is a sudden degradation of the measured capacity. In the Figure 7(a), we allow only three consecutive reinforcements of the most probable value, after convergence. In Figure 7(b), only one reinforcement is allowed.

Figure 7(a) shows what happens when cap_{meas} suddenly becomes 1.1 Mbps and constantly remains the *same* for all next series of measurements. Our starting point is the case depicted in Figure 6(c) (i.e., our model evolved as depicted in Figure 6(c), before the sudden change in cap_{meas}). The goal is to examine how quickly the system can adapt and converge to cap_2, which is the value that is nearest to cap_{meas}. As can be observed, in 4 steps (phases 2-5) the most probable value drops from cap_4 to cap_3. In another 6 steps (phases 6-11), candidate values cap_2 and cap_3 are suggested as the most likely ones. Finally, in the next step (phase 12), cap_2 becomes the most probable one.

Figure 7(b) shows what happens if only one reinforcement is allowed after convergence. Again, Figure 6(c) is our starting point and cap_{meas} is 1.1 Mbps. In just 2 steps (phases 2-3), the most probable value drops from cap_4 to cap_3. In another 3 steps (phases 4-6), cap_2 almost reaches cap_3, and in the next step (phase 7) cap_2 becomes the most probable value.

Another conclusion can be deduced from the simulations described in this subsection. The number of consecutive reinforcements, after convergence, clearly affects the model's adaptation speed. High number of consecutive reinforcements reduces the adaptation speed.

Scenario 3: Complex example

Our focus is on a second arbitrary configuration $c_2=(r_2, f_2)$. The goal of this subsection is to provide a more detailed and complex scenario, so as to demonstrate both the efficiency and the adaptation speed of the proposed probabilistic model. Especially regarding the model's adaptation speed, it should be noted that the scenario presented below can be classified as one of the worst case scenarios, due to the radical changes of the measured values.

It is assumed that there are 11 candidate values (in Mbps): cap_1=1, cap_2=2, …, cap_{11}=11. Then, dif_{max}=10 Mbps. Again, it is assumed that only three consecutive reinforcements are allowed after convergence. The scenario consists of three phases, described in the following. Figure 8(a)-(c) depicts the distribution of conditional probabilities for each phase.

Figure 8(a) shows the behaviour when cap_{meas}=7.3 Mbps. At first, the conditional probabilities of the candidate values are uniformly distributed, as depicted in phase 1. The robust discovery functionality correctly adjusts (phase 2), by selecting cap_7 as the most probable value. During the next steps (phases 3-5), the measured value does not change, so the most probable value remains the same, and is actually further reinforced. As has been explained, the number of allowed further reinforcements depends on the implementation. In this scenario, three further consecutive reinforcements (phases 3-5), after convergence, are allowed. Higher values help the system to avoid oscillations, while lower values help it adjust more quickly to severe and relatively permanent changes.

Figure 8(b) shows what happens when, at some time, the measured value changes to cap_{meas}=9.8 Mbps. The last phase of Figure 8(a) is the starting point for this phase, i.e. phase 1 of Figure 8(b) is the same as phase 5 of Figure 8(a). As depicted in Figure 8(b), the model's performance is quite satisfactory, as in only 2 steps (phase 3) the most probable value shifts from cap_7 to cap_8. In another 3 steps (phase 6), the model selects cap_9 as the most probable value, and, finally, in another 4 steps (phase 10), cap_{10} (which is the nearest

to the measured value) is selected. Thus, in this phase, the system proves to be both fast in adaptation, and also careful (the entire adaptation process takes more than just 1 or 2 steps), in order to avoid oscillations. It is also worth noting that, as may be observed in Figure 8(b), the conditional probabilities of cap_1 (phases 2-10), cap_2 (phases 2-10), cap_3 (phases 3-10), cap_4 (phases 4-10) and cap_5 (phases 7-10) have been set by the model to the minimum allowed value (i.e., the threshold's value). This can help the model in future adaptations, as happens in the case of the next phase of the scenario.

In Figure 8(c), the system is challenged to adapt to an extremely radical change. The capacity's measured value, cap_{meas}, drops from 9.8 to 3.8. Figure 8(c) depicts how the model adapts to the situation. Phase 1 represents the distribution of the final step of Figure 8(b). As may be observed, in just 1 step (phase 2) the most probable value drops from cap_{10} to cap_9. In another 3 steps (phase 5), the most probable value shifts from cap_9 to cap_8. In the next step (phase 6), candidate value cap_4 (which is the closest to the measured value), as well as the candidate values in the neighbourhood of cap_4 increase significantly. In the next step (phase 7), value cap_4 outperforms the others. Thus, the model managed to adapt in only 6 steps, which is a remarkable performance. The "threshold" mechanism has also contributed in the model's satisfactory adaptation speed, by not allowing the conditional probability of cap_4 and its neighbours to drop below an appropriate minimum value.

8. CONCLUSIONS

B3G wireless connectivity can efficiently be realized by exploiting cognitive networking concepts. Cognitive systems dynamically reconfigure the RATs and the spectrum they use, in order to adapt to the changing environment conditions. However, dynamic reconfiguration decisions call for robust radio-scene analysis and channel identification schemes. This paper has contributed in the areas of radio-scene analysis and channel identification: firstly, by providing an overview of interference estimation methods (section 3), and explaining how capacity estimations can be derived based on the measured interference levels (section 4); secondly, by specifying the information flow for the radio-scene analysis process of a cognitive radio system (section 5); and thirdly, by enhancing the above with a learning system, which is essential for obtaining a truly cognitive process (sections 6-7). The proposed approach was to develop a robust probabilistic model for optimal prediction of the capabilities of alternative configurations, in terms of capacity. The approach relied on the use of basic Bayesian logic, in combination with a learning and

adaptation strategy. Results that expose the behaviour and efficiency of the proposed scheme were also presented.

The short-term future plan is to enrich the basic Bayesian model that has been described, by adding more nodes (random variables), including ‘coverage’ and ‘context’ (i.e., ‘traffic’ and ‘user mobility’). The overall future plan is to further employ probabilistic relationships and autonomic computing principles in the direction of realizing cognitive, wireless access, infrastructures. The goal is to develop an autonomic manager which will encompass the robust estimation scheme. The manager will consist of policies, context perception capabilities, reasoning algorithms, learning functionality and knowledge engineering, technologies for the representation of ontologies and semantics. All these will yield a system that hypothesises on causes to a problem, and subsequently validates or falsifies the hypothesis.

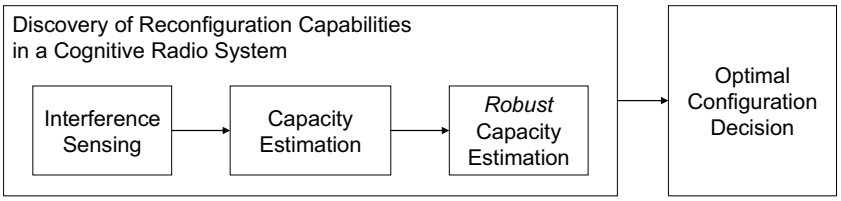


Figure 1: The process of discovering reconfiguration capabilities in a cognitive radio system

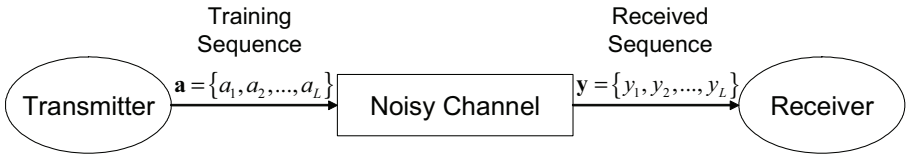


Figure 2: Model for simple SINR estimation

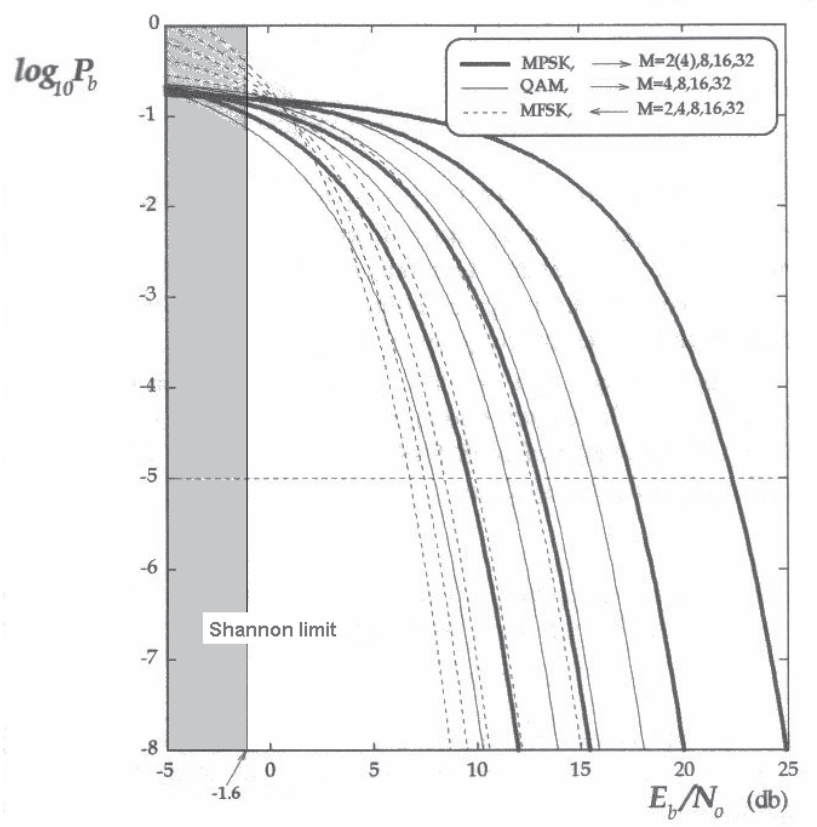


Figure 3: Bit error probabilities for uncoded M-ary FSK, PSK and QAM, coherent demodulation - Adapted from [30]

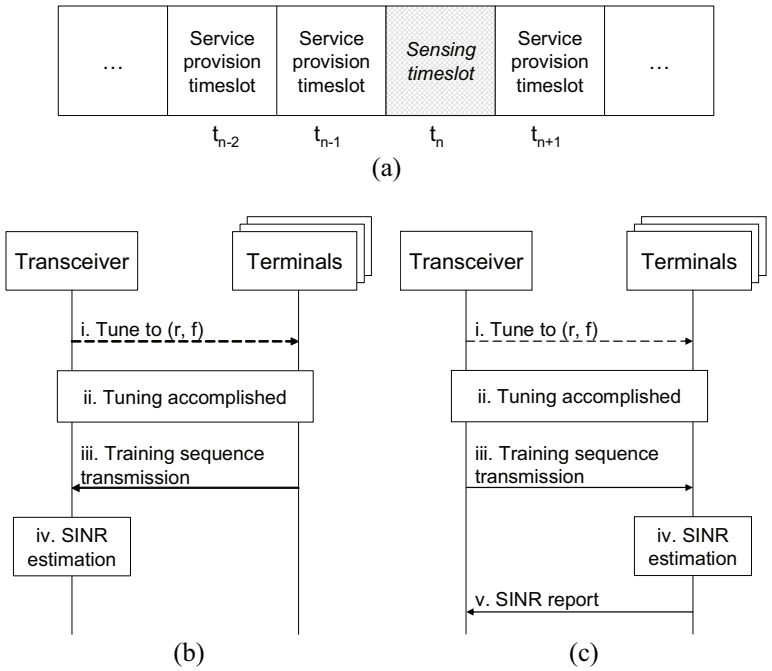
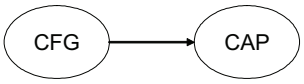


Figure 4: Interference sensing process in a cognitive radio system: (a) notion of service provision timeslots and sensing timeslots; (b) process for the uplink; (c) process for the downlink

(a)



(b)

CAP	$RAT = r_1$		
	$FRQ = f_1$ ($CFG = c_1 = (r_1, f_1)$)	…	$FRQ = f_{\lvert F_n \rvert}$ ($CFG = c_{\lvert F_n \rvert} = (r_1, f_{\lvert F_n \rvert})$)
cap_1	$\Pr[CAP = cap_1 \lvert CFG = c_1]$	…	$\Pr[CAP = cap_1 \lvert CFG = c_{\lvert F_n \rvert}]$
cap_2	$\Pr[CAP = cap_2 \lvert CFG = c_1]$	…	$\Pr[CAP = cap_2 \lvert CFG = c_{\lvert F_n \rvert}]$
…	…	…	…
cap_m	$\Pr[CAP = cap_m \lvert CFG = c_1]$	…	$\Pr[CAP = cap_m \lvert CFG = c_{\lvert F_n \rvert}]$

(c)

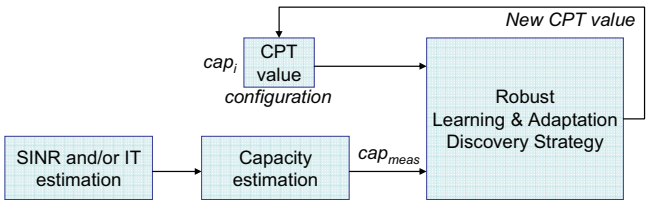
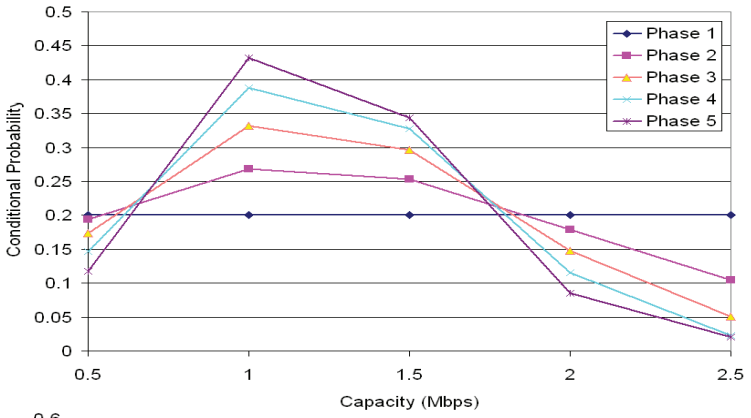
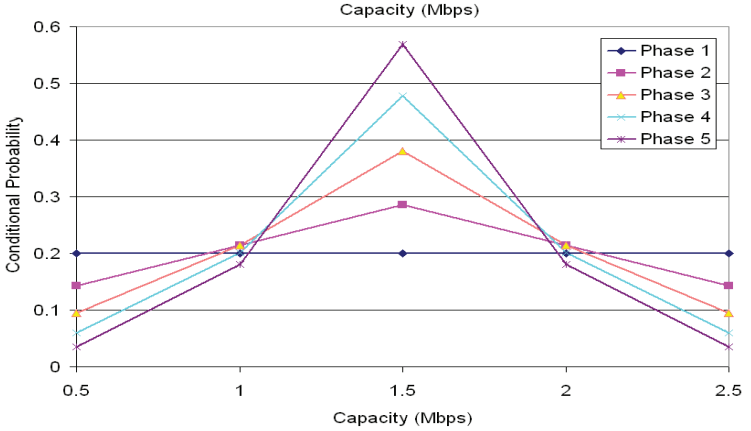


Figure 5: (a) Basic Bayesian network for robust capacity estimation; (b) Sample of the conditional probability table for capacity; (c) Process for updating the CPT values

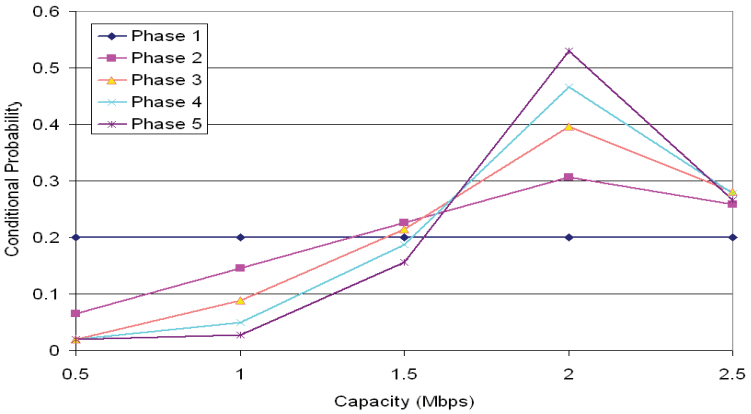
(a)



(b)



(c)



(d)

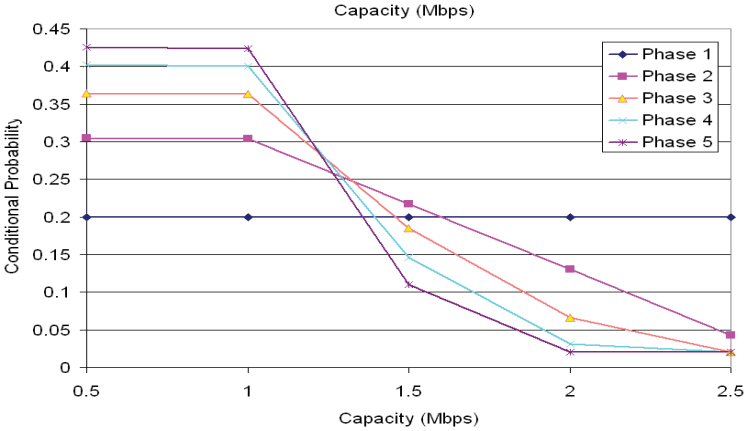
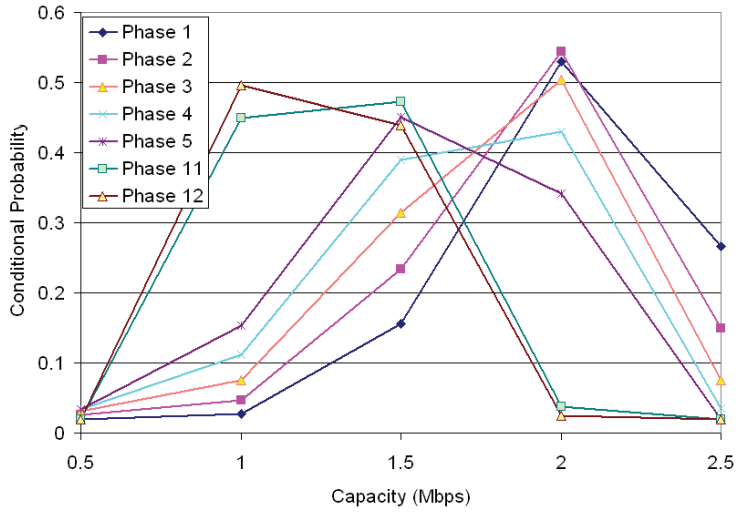


Figure 6: Scenario 1 - Arbitrary configuration c_1. Conditional probabilities corresponding to candidate capacity values, when cap_{meas} is (in Mbps): (a) 1.2; (b) 1.5; (c) 2.1; (d) 0.75

(a)



(b)

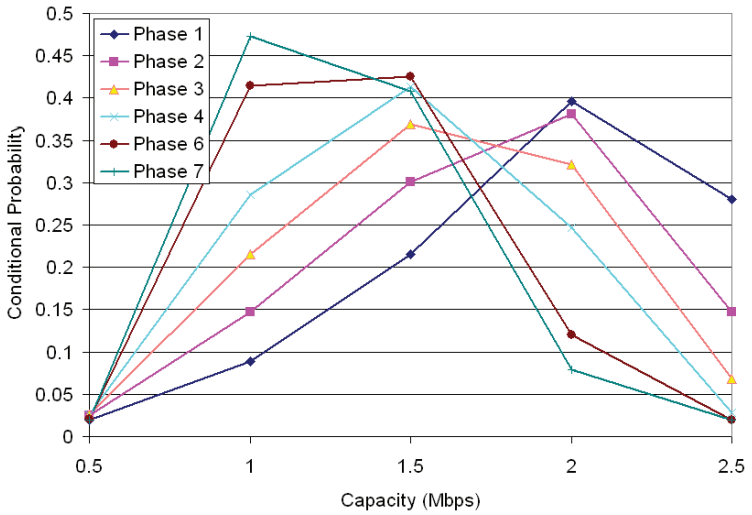
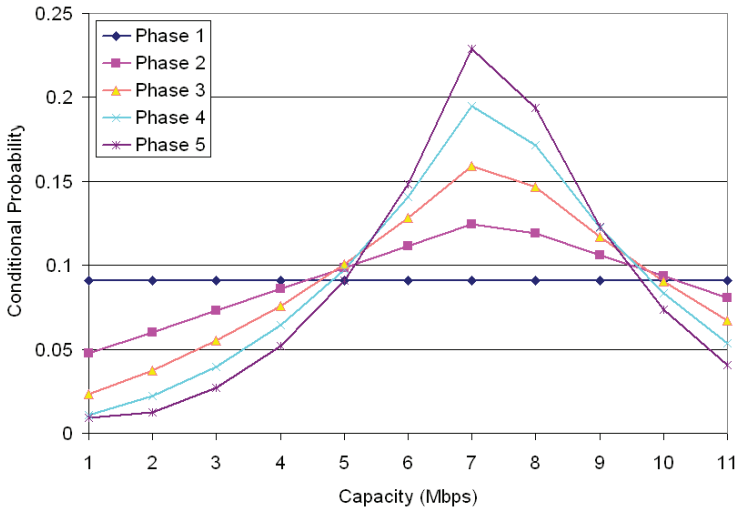
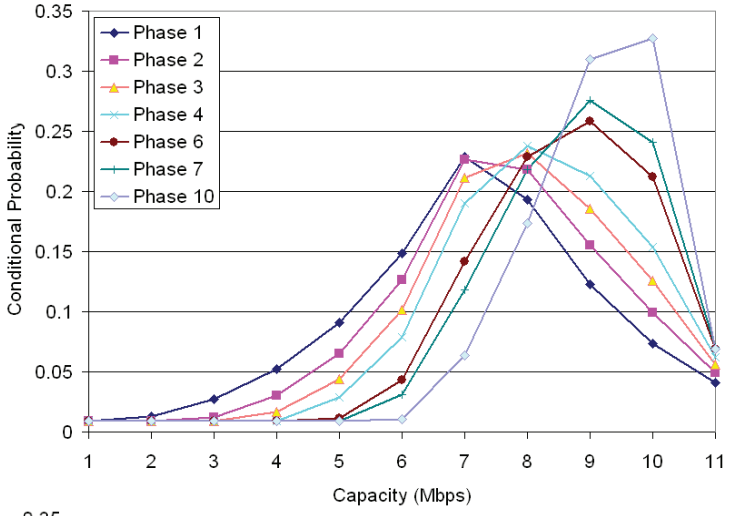


Figure 7: Scenario 2 - Arbitrary configuration c_1. Learning and adaptation process when cap_{meas} suddenly degrades from 2.1 to 1.1 Mbps. Speed of adaptation when the number of consecutive updates, after convergence, is: (a) three; (b) one

(a)



(b)



(c)

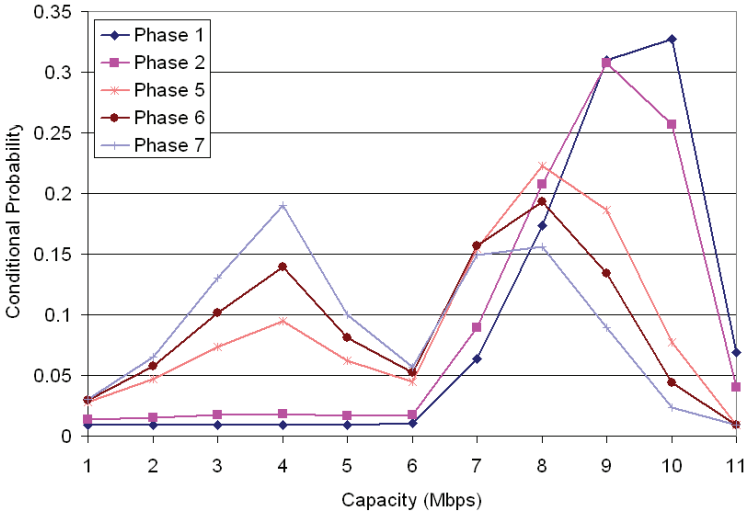


Figure 8: Scenario 3 - Arbitrary configuration c_2. Conditional probabilities that correspond to candidate capacity values, when cap_{meas} is sensed at (in Mbps): (a) 7.3; (b) 9.8; (c) 3.8

ACKNOWLEDGEMENT

The work presented herein is conducted in the framework of Ph.D. research performed by K. Demestichas and E. Adamopoulou, under the supervision of Prof. M. Theologou. The work is funded by the General Secretariat of Research and Technology (GSRT) of the Greek Ministry of Development, in the context of the ARIADNE project (03ED235).

REFERENCES

[1] "Management of heterogeneous networks", *feature topic in the IEEE Communications Magazine*, Vol. 36, No. 3, March 1998.

[2] "Fourth generation wireless networks and interconnecting standards", *Special Issue in the IEEE Personal Communications*, Vol. 8, No. 5, Oct. 2001.

[3] P. Kolodzy et al., "Next generation communications: Kickoff meeting", *in Proc. DARPA*, Oct. 2001.

[4] M. McHenry, "Frequency agile spectrum access technologies", *in FCC Workshop Cognitive Radio*, May 2003.

[5] G. Staple and K. Werbach, "The end of spectrum scarcity", *IEEE Spectrum*, Vol. 41, No. 3, pp. 48–52, March 2004.

[6] J. Mitola and G. Maguire Jr., "Cognitive Radio: Making Software Radios More Personal", *IEEE Personal Communications Magazine*, Vol. 6, No. 6, pp. 13-18, August 1999.

[7] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio", Doctor of Technology, Royal Institute of Technology (KTH), Stockholm, Sweden, 2000.

[8] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation Issues in Spectrum Sensing for Cognitive Radios", *Conference Record of the 38th Asilomar Conference on Signals, Systems and Computers*, Vol. 1, pp. 772-776, Nov. 2004.

[9] W. Tuttlebee, "Software Defined Radio: Origins, Drivers, and International Perspectives", *Wiley*, New York, 2002.

[10] M. Milliger et al., "Software Defined Radio: Architectures, Systems and Functions", *Wiley*, New York, 2003.

[11] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications", *IEEE Journal on Selected Areas In Communications*, Vol. 23, No. 2, pp. 201-220, Feb. 2005.

[12] B. Walke, P. Seidenberg, and M. P. Althoff, "UMTS – The Fundamentals", *Wiley*, 2003.

[13] A. Sampath and D. R. Jeske, "Analysis of Signal-to-Interference Ratio Estimation Methods for Wireless Communication Systems", *IEEE International Conference on Communications*, Helsinki, Finland, Vol. 8, pp. 2499-2503, June 2001.

[14] C. E. Gilchriest, "Signal-to-noise Monitoring", *JPL Space Programs Summary*, Vol. IV, No. 32-37, pp. 169-184, June 1966.

[15] J. W. Layland, "On S/N Estimation", *JPL Space Programs Summary*, Vol. III, No. 37-48, pp. 209-212.

[16] C. M. Thomas, "Maximum Likelihood Estimation of Signal-to-Noise Ratio", *Ph.D. Dissertation*, University of Southern California, Los Angeles, 1967.

[17] T. R. Benedict and T. T. Soong, "The joint estimation of signal and noise from the sum envelope", *IEEE Transactions on Information Theory*, Vol. IT-13, No. 3, pp. 447-454, 1967.

[18] B. Shah and S. Hinedi, "The split symbol moments SNR estimator in narrow-band channels", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 25, No. 5, pp. 737-747, 1990.

[19] D. R. Pauluzzi and N. C. Beaulieu, "A Comparison of SNR Estimation Techniques for the AWGN Channel", *IEEE Transactions on Communications*, Vol. 48, No. 10, pp. 1681-1691, Oct. 2000.

[20] T. Petennanny, D. Boss, and K. D. Kammeyer, "Blind GSM Channel Estimation under Channel Coding Conditions", *in Proc. of the 38th Conference on Decision and Control*, Phoenix, Arizona, USA, Dec. 1999.

[21] D. Shin, W. Sung, and I. Kim, "Simple SNR Estimation Methods for QPSK Modulated Short Bursts", *in Proc. of the IEEE Global Telecommunications Conference (GlobeCom 2001)*, Vol. 6, pp 3644-3647, Nov. 2001.

[22] Federal Communications Commission, "Establishment of interference temperature metric to quantify and manage interference and to expand available unlicensed operation in certain fixed mobile and satellite frequency bands", *ET Docket 03-289*, Notice of Inquiry and Proposed Rulemaking.

[23] T. C. Clancy and W. A. Arbaugh, "Interference Temperature Multiple Access: A New Paradigm for Cognitive Radio Networks", *in submission to ACM Symposium on Mobile Ad-hoc Networking and Computing (MobiHOC 2006)*, 2006.

[24] P. J. Kolodzy, "Interference temperature: a metric for dynamic spectrum utilization", *International Journal of Network Management, Special Issue: Management of Interference in Wireless Networks, John Wiley & Sons*, Vol. 16, No. 2, pp. 103-113, March 2006.

[25] D. J. Thomson, "Spectrum estimation and harmonic analysis", *Proc. IEEE*, Vol. 20, pp. 1055-1096, Sept. 1982.

[26] D. J. Thomson, "Multitaper analysis of nonstationary and nonlinear time series data", *Nonlinear and Nonstationary Signal Processing*, W. Fitzgerald, R. Smith, A. Walden, and P. Young, Eds. London, U.K.: *Cambridge University Press*, 2000.

[27] P. Stoica and T. Sundin, "On nonparametric spectral estimation", *Circuits, Syst., Signal Process.*, Vol. 16, pp. 169-181, 1999.

[28] M. E. Mann and J. Park, "Oscillatory spatiotemporal signal detection in climate studies: A multiple-taper spectral domain approach", *in Advances in Geophysics*, R. Dnowska and B. Saltzman, Eds. New York: Academic, Vol. 41, pp. 1–131, 1999.

[29] S. Haykin, "Communication Systems", *Wiley*, 4th Ed., ISBN: 0-471-17869-1, 2001.

[30] K. Schwieger, A. Kumar, and G. Fettweis, "On the Impact of the Physical Layer on Energy Consumption in Sensor Networks", *in Proc. of the Second European Workshop on Wireless Sensor Networks*, pp. 13-24, 31 Jan.-2 Feb. 2005.

[31] J. Strassner, "Policy-based network management", *Morgan Kaufmann Publishers*, U.S.A., 2005.

[32] J. Strassner, "Autonomics – A critical and innovative component of seamless mobility", Technical Report, http://www.motorola.com/mot/doc/5/5978_MotDoc.pdf, Motorola, Dec. 2005.

[33] R. E. Neapolitan, "Learning Bayesian Networks – Series in Artificial Intelligence", *Prentice Hall*, 2002.

[34] E. Adamopoulou, K. Demestichas, A. Koutsorodi, and M. E. Theologou, "Access Selection and User Profiling in Reconfigurable Terminals", *in Proc. of the 15th Wireless World Research Forum (WWRF)*, Paris, Dec. 2005.

[35] J. Cheng and R. Greiner, "Learning Bayesian belief network classifiers: algorithms and system", *in Proc. of the 14th Canadian Conference on Artificial Intelligence*, pp. 141-151, 2001