

RESEARCH

Open Access

Discovering and exploiting spectrum power correlations in cognitive radio networks: an experimentally driven approach

Stamatios Arkoulis^{1*}, Evangelos Anifantis¹, Vasileios Karyotis^{1,2}, Symeon Papavassiliou^{1,2}
and Nikolaos Mitrou¹

Abstract

In this paper, we focus on increasing the spectrum awareness of cognitive radio users through statistical processing of spectrum sensing data, obtained via wideband energy-detection-based sensing techniques. Based on observations over real spectrum power measurements, we advocate the existence of correlation properties in the sensed power of measured neighboring channels and propose an inference methodology for exploiting them towards acquiring a more accurate view of the underlying wireless environment. To highlight the benefits of the proposed correlation-based inference methodology on cognitive radio systems, we emphasize on enhancements of white space discovery and channel selection processes, while we thoroughly discuss the impact of our findings on existing relevant approaches. Based on a systematic wireless spectrum survey in the metropolitan area of Athens, Greece, we validate our proposed methodology and assess the achieved performance improvements through the obtained measurement data, confirming its potential value in future cognitive radio networks.

Keywords: Cognitive radios; Spectrum sensing; Energy detection; Power correlations; Coarse-fine sensing; Spectrum occupancy survey

1 Introduction

The traditional static allocation of wireless spectrum applied by federal committees around the world has been lately questioned due to the nowadays documented spectrum under-utilization [1,2]. Cognitive radio (CR) technology was developed as a remedy to the aforementioned problem, conceptually increasing the utilization factor of the existing wireless resources. Specifically, CR-enabled secondary users (SUs) capitalize on spectrum sensing techniques for detecting unoccupied spectrum portions (i.e., white spaces) and leverage their collected information for opportunistically exploiting them, without causing harmful interference to the entrenched primary users (PUs).

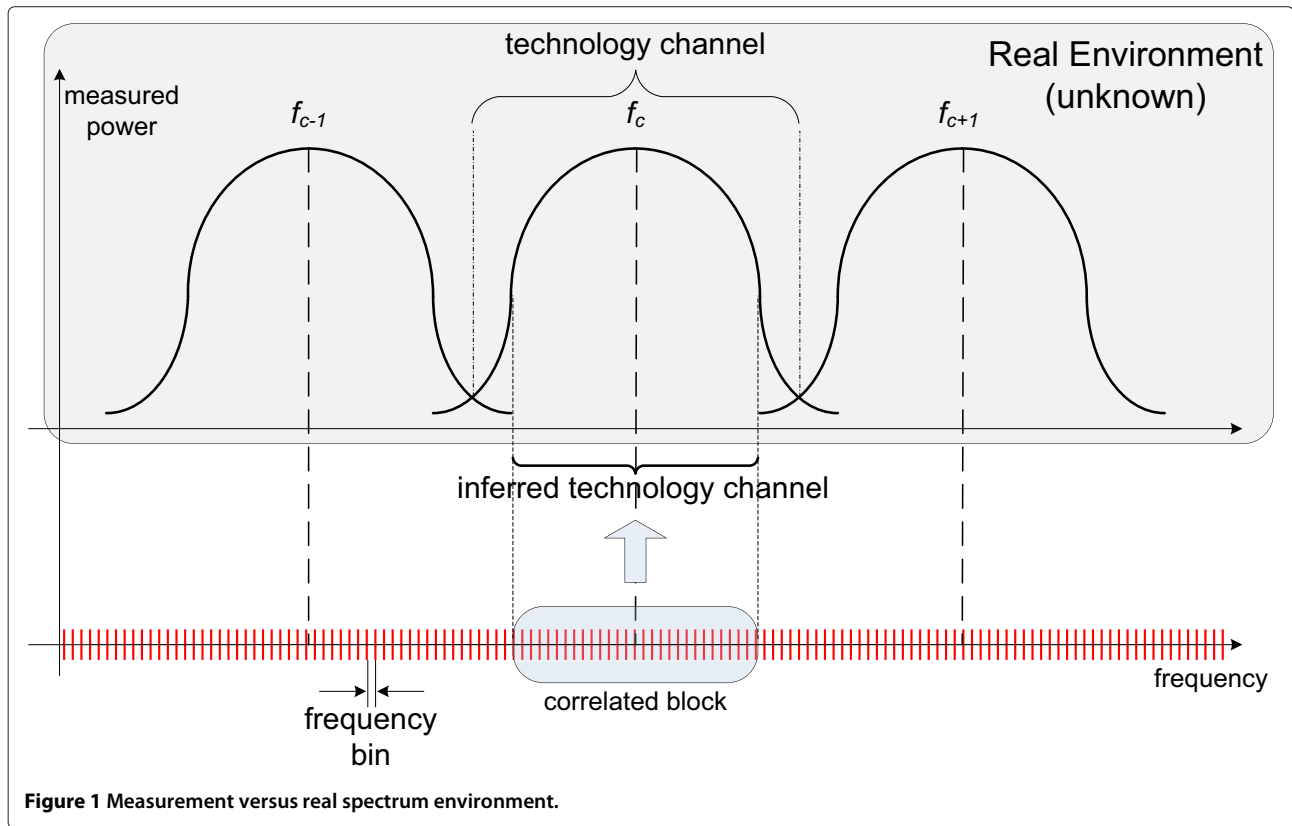
In the currently established CR context, SUs have to examine wide frequency ranges in order to accurately

discover appropriate white spaces. Towards that, various sensing techniques have been proposed for examining holistically these ranges [3]. However, their application has been proven complex, expensive and, thus, prohibitive in practical terms [2]. Alternative approaches split the wide frequency bands into multiple narrower ones and examine them through narrowband sensing techniques [4]. Among them, approaches based on energy detection constitute the most preferable sensing techniques, regarding practical implementation, since they do not take for granted any knowledge on the specific characteristics of underlying PU signals and/or the formation of secondary channels. According to that technique, a wide frequency range of interest is divided into narrower bands, namely, frequency bins (Figure 1), which are sequentially examined (swept) to determine their occupancy, i.e., presence or absence of PU signal transmissions. However, splitting wide frequency bands into narrower ones and exhaustively examining them alone is not sufficient for sustaining the performance requirements of current and future networks.

*Correspondence: stark@cn.ntua.gr

¹School of Electrical and Computer Engineering, National Technical University of Athens (NTUA), Zografou, Athens 15780, Greece

Full list of author information is available at the end of the article



1.1 Paper contributions and outline

In this study, we present and experimentally validate a generic methodology which could significantly enhance approaches that adopt energy detection as their sensing technique. The main objective of this work is to enable SUs to infer valuable characteristics of their otherwise unknown operating spectrum environment through statistical processing of acquired sensing data and without relying on knowledge obtained from external entities such as spectrum databases. As will be explained later (Section 5), this can have significant benefits in the operation and performance of cognitive radio networks.

The cornerstone for our study is the observed correlated behavior of frequency bins spanned by a single technology-specific PU channel (Figure 1). In particular, when a PU channel (e.g., a GSM one) spans several frequency bins, their measured power exhibit correlated behavior. Although important, such correlation properties have not been sufficiently studied and exploited in the literature where relevant, but simplistic and not practically validated, assumptions dominate. In this paper, we advocate that SUs can raise and exploit their awareness on their underlying transmission environment by estimating the statistical correlation of acquired power measurements of adjacent frequency bins, and then clustering the highly correlated ones into distinct groups. In this manner, SUs

will be capable of efficiently detecting frequency regions that present coherent behavior and define sets of frequency bins that are uniformly affected by a single PU. Therefore, SUs become able to infer the ranges of typically unknown technology-specific channels that are defined by primary systems, thus managing to reconstruct the landscape of their underlying spectrum environment, with the benefits that will be explained in Section 5.

The detection of frequency ranges that are affected by the primary system in a specific manner can be practically exploited at large. In this work, we highlight practical sensing techniques, like the well-known coarse-fine ones [5,6], that can significantly benefit from the resulting SU ability to extrapolate the occupancy decision of one bin to the entire coherent group it belongs to. Furthermore, we show how the added knowledge regarding cohesive frequency regions can enable SUs to select more stable white spaces, yielding eventually reduced reconfiguration overhead and additional benefits. Finally, our proposed methodology and obtained results can act supplementary to the existing and future studies that incorporate relevant theoretic assumptions, rendering them more viable for practical applications.

Both the described inference methodology and its potential benefits are experimentally validated using a state-of-the-art sensing architecture, which however, can

be commercially realized as well. Prior to any experiment, we thoroughly evaluated the suitability of the local spectrum environment for CR implementation by conducting a spectrum occupancy survey in the Athens' metropolitan area in Greece. Then, focusing on the more challenging cases for CR spectrum exploitation, we first experimentally demonstrate high correlations between neighboring frequency bins. By applying the proposed methodology, we highlight that the inferred correlated blocks align with the underlying PU technology channels allocation. Finally, we demonstrate the performance improvements that can be achieved by coarse-fine sensing techniques when exploiting the acquired knowledge, as well as the obtained benefits in terms of reconfiguration overhead during the channel selection process.

The rest of this paper is organized as follows. In Section 2 we present some background technical details of the employed framework, while Sections 3 and 4 describe and validate through real experimentation, respectively, the suggested correlation-based inference methodology. In Section 5 exploitation scenarios for our proposed methodology are presented and evaluated. Finally, the past and current state-of-the-art approaches in each of the involved fields are described in Section 6, while Section 7 concludes the paper.

2 Background details of our experimentally driven approach

Energy-detection-based spectrum sensing is characterized by low implementation and computational complexity, as well as high detection speed and independence from the detected signal, i.e., it does not require prior knowledge of the PU signal. Despite some restricted cases of documented poor performance [4], this spectrum sensing approach is currently supported by the majority of the existing cognitive radio platforms [7], hence, allowing new outcomes to be exploited in a practical manner.

Towards employing energy-detection-based spectrum sensing in practice, the utilized equipment should attain high-speed frequency sweep capabilities that enable rapid collection of power measurements in frequency bins of wide frequency ranges. In this paper, we use a vector signal analyzer (provided in Figure 2) connected to a wideband omnidirectional discone antenna via a low-loss coaxial cable. We have selected such a high-end accuracy equipment for providing the most accurate picture of the measured spectrum, thus revealing the realistic properties of the wireless environment that can be further exploited. Even though our used spectrum analyzer can obtain very accurate measurements, it should be noted that the outcomes of our study are not restricted by its specifications, but can be straightforwardly exploited by any energy detection-capable device, perhaps at the cost of some tolerable accuracy loss. For instance, we could

have well used the less heavy and bulky CR-enabled USRP devices, or even experimental hardware implementations of spectrum analyzer substitutes (e.g., [8]) characterized by almost handheld dimensions and low cost (US\$200 to US\$500).

Figure 3 depicts the block diagram of the employed energy detection implementation for a continuous-time received signal $r(t)$. In particular, the anti-aliasing filter is incorporated to eliminate the effect of aliasing and attenuate the frequency components at and above half the sampling frequency when the continuous-time signal is digitized by sampling. Next, a finite duration 7-term Blackman-Harris window is applied to $r(n)$ selecting finite-length segments (K samples), denoted by $y(n)$, and the corresponding frequency coefficients $Y(n)$, as well as the spectrum energy (or the power spectrum density measured in watt per Hertz), are estimated through the fast Fourier transform (FFT) and squaring module [9]. Finally, averaging is used to improve the estimation of power measurement due to the presence of noise. More specifically, within an observation time interval N , a data sequence $y(n)$ is divided into M data segments (by windowing) such that $N = K \times M$, and the averaged spectrum power for each frequency coefficient is computed by averaging over the M data segments. Based on the described implementation, the number of K FFT points (defining the frequency bin resolution) and the number of averages M have been appropriately configured, as noted in Table 1, to improve the energy detector performance.

Mathematically, the problem of energy detection can be formulated as a binary test of the following two hypotheses:

$$\begin{cases} H_0 : y(n) = \vartheta(n) & \text{if signal is absent} \\ H_1 : y(n) = s(n) + \vartheta(n) & \text{if signal is present,} \end{cases} \quad (1)$$

where $s(n)$ and $\vartheta(n)$ denote the discrete-time PU's signal and the additive noise respectively ($n = 1, \dots, N$). As proven by Parseval's theorem, the energy evaluation can be performed both in time and frequency domains. Thus, the averaged periodogram, namely, the averaged power spectrum density (PSD), for each frequency bin k can be estimated as follows:

$$P(k) = \frac{1}{M} \sum_{m=1}^M \frac{1}{K} |Y(k)|^2 \quad k = 0, \dots, K-1 \quad (2)$$

and a decision problem is formulated with test statistic $T_k = P(k)$ for each frequency bin k , i.e.,

$$\begin{cases} T_k \leq \gamma & \text{determines that the frequency bin is idle} \\ T_k > \gamma & \text{determines that the frequency bin is busy} \end{cases} \quad (3)$$

where γ denotes a pre-selected decision threshold.

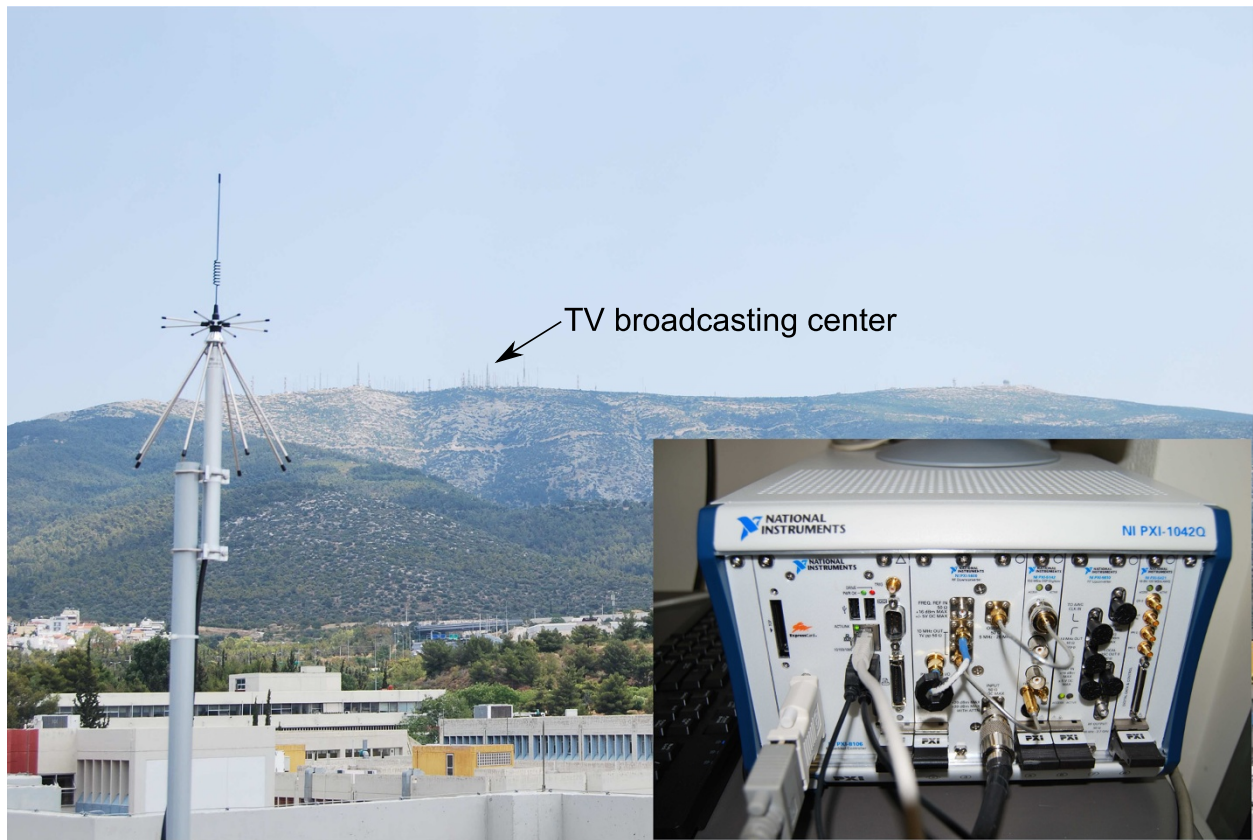


Figure 2 Terrestrial sensing and spectrum analyzer infrastructure.

The performance of the energy detector depends on the probabilities of miss detection P_{md} and false alarm P_{fa} , which are expressed as follows:

$$P_{md} = Pr(T_k \leq \gamma | H_1) \quad (4)$$

$$P_{fa} = Pr(T_k > \gamma | H_0). \quad (5)$$

The decision threshold γ can be theoretically estimated based on (4), (5), and Neyman-Pearson lemma [10]. However, for practical cases other alternative approaches, such as the m -dB and probability of false alarm (PFA) criterion, are considered leveraging real measurements for γ -selection [1]. Both approaches examine γ for each frequency bin independently, requiring the prior collection of corresponding noise samples through antenna replacement with a matched (passive) load. In the former, γ is

chosen equal to m -dB above the average noise level measured to the corresponding frequency bin, while in the latter and more advanced method, the detection threshold is selected such that the maximum fraction of noise samples found above the value of γ for the corresponding bin is equal to P_{fa} .

3 Correlation discovery as an inference methodology

Although the generality of energy-detection-based sensing is appropriate and convenient for various practical implementations, its simplicity poses additional challenges that should be addressed in a more sophisticated manner. More specifically, it is a fact that a cognitive radio user is inherently unaware of the underlying real spectrum environment that is mostly determined by primary

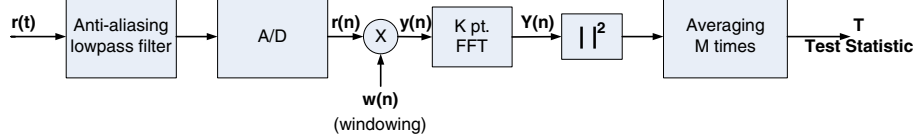


Figure 3 Block diagram of energy detector.

Table 1 Equipment operation parameters

| Parameter | Values |
|----------------------|-----------------------|
| Frequency span | 50 MHz |
| Central frequencies | from 325 to 2,675 MHz |
| Resolution bandwidth | 10 kHz |
| Frequency bin size | 4.02 kHz |
| Detection type | RMS |
| No. of averages | 10 |
| Reference level | −30 dBm |
| Sensing interval | 30 s |

systems. Especially, in case of autonomous infrastructureless networks, SUs might be unable to communicate with external sources that retain information about PU characteristics. Even if this could be feasible, full knowledge of real characteristics is difficult to be obtained and requires information fusion (i.e., official frequency allocation on a per country/region case, applied protocols per band, protocol specifications) from multiple heterogeneous sources. Moreover, the unknown margins of available secondary channels adversely affect the efficiency of energy detection sensing in practice, since it results in exhaustively examining the occupancy of a large amount of frequency bins. Therefore, since it is a basic CR technology principle that spectrum knowledge is not expected in the general case to be obtained by external sources, SUs are solely responsible to create such valuable knowledge through intelligently processing their observations (sensing data). In this manner, they can exploit this knowledge towards driving their decisions for fast sensing, efficient resource allocation, optimal frequency band selection, etc.

As depicted in Figure 1, the real spectrum environment entails PU transmissions (unknown to SUs) at technology-specific channels, e.g., a GSM cellular band, which typically span over a number of measured frequency bins. Consequently, the corresponding frequency bins are expected to exhibit high correlation and similar availability characteristics, as perceived by an independent observer. In fact, it is exactly these correlation properties that can well be exploited by SUs towards inferring characteristics of the underlying spectrum environment. Motivated by this fact, in this study we propose a framework that enables SUs to gain knowledge on their unknown real environment by inferring coherent frequency regions that constitute potential technology-specific channels through appropriate frequency bin grouping (clustering). Towards this direction, each SU leverages historical sensing data, and thus, the proposed methodology does not actually intervene with the SU operation, nor does it decrease the performance of their devices.

In our suggested inference methodology, the set of power sensing data collected independently by a SU is considered as a series of observations that has been obtained for each frequency bin within the respective examined measurement interval \mathcal{I} . In a practical operational scenario, these data are collected by each SU during its conventional energy-detection-based spectrum sensing (for detecting spectrum holes in the underlying spectrum) and are representatives of PUs behavior as well as robust to incidental factors. We assume that for every frequency bin, there is a continuous random variable X_f , characterizing its energy, with value $x_f \in \mathbb{R}$. We also point out that correlation properties are estimated in the raw form of power measurements instead of frequency bins occupancy, thus, remaining robust to the employed γ threshold (and as a result, to the miss detection and false alarm errors) which could jeopardize the accuracy of the obtained results. In order to investigate the dependency between two random variables X_{f_1} and X_{f_2} , the Pearson product-moment correlation coefficient (ρ) is used and it is expressed by

$$\rho_{X_{f_1} X_{f_2}} = \frac{\text{cov}(X_{f_1}, X_{f_2})}{\sigma_{X_{f_1}} \cdot \sigma_{X_{f_2}}}. \quad (6)$$

It is noted that $\text{cov}(X_{f_1}, X_{f_2}) = E[(X_{f_1} - m_{X_{f_1}}) \cdot (X_{f_2} - m_{X_{f_2}})]$ is the covariance of the corresponding random variables, while m_{X_f} and σ_{X_f} denote the expected value (mean) and the standard deviation of the random variable X_f , respectively.

The Pearson correlation coefficient can sufficiently measure the linear dependence between two random variables when the probability distributions of the respective populations are known. However, for practical cases, as in our study, where only a sequence of measurements are known regarding a frequency bin X_f (denoted by the time series $\{x_f^{(1)}, x_f^{(2)}, \dots, x_f^{(T)}\}$), the sample correlation coefficient r is used as an estimation of the population correlation:

$$r_{f_1 f_2} = \frac{\sum_i (x_{f_1}^{(i)} - \bar{x}_{f_1}) \cdot (x_{f_2}^{(i)} - \bar{x}_{f_2})}{\sqrt{\sum_i (x_{f_1}^{(i)} - \bar{x}_{f_1})^2 \cdot \sum_i (x_{f_2}^{(i)} - \bar{x}_{f_2})^2}} \quad (7)$$

and $\bar{x}_f = \frac{x_f^{(1)} + x_f^{(2)} + \dots + x_f^{(T)}}{T}$. By abuse of notation, we hereafter use the symbols ρ and r to denote, respectively, the correlation coefficient (Equation 6) and the sample correlation coefficient (Equation 7) for any pair of frequency bins. The correlation coefficient value can range from -1 (perfect negative correlation) to $+1$ (perfect positive correlation), while $r = 0$ implies that the examined random variables are uncorrelated (but not necessarily

independent), and thus the corresponding frequency bins are not possible to belong to the same technology-specific channel.

By computing the correlation coefficient between any frequency bins pair of interest, a SU can next cluster sequential frequency bins together into discrete bands based on their correlation value and, thus, create groups of frequency bins that present coherent behavior. As explained earlier, each such group essentially corresponds to a technology-specific channel and thus the underlying spectrum environment picture can be successfully and effectively reconstructed with various benefits explained in the sequel.

4 Real experimentation and validation

Towards validating the aforementioned statements in the real-world, we use a spectrum sensing architecture whose main component is a National Instruments PXI-5661 RF Vector Signal Analyzer (National Instruments Corporation, Austin, TX, USA) [11] with sensing range that spans the 9 kHz to 2.7 GHz frequency band, hosted by the NI PXI-1042Q chassis and controlled by the Labview 8.5 software. Our spectrum analyzer was placed in a protected room and connected to a high-performance wideband (300 to 3,000 MHz) omnidirectional discone antenna (specifically the Sirio SD 3000N model; Sirio Antenne, Volta Mantovana, Verona, Italy) using a 10-m low-loss RG213-U coaxial cable (see snapshot in Figure 2). The mounting location of our antenna (Figure 2) was completely unobstructed from all directions and with unhindered view to the city center as well as one of the main TV broadcasting centers of Athens (located in Ymetous mountain) and various transmitters (e.g., GSM, 3G, etc.). Table 1 summarizes the operation parameters of the utilized sensing architecture.

To avoid the weaknesses of relevant works (e.g., [12,13]), where frequency bins span parts of different technology channels and, thus, fail to acceptably reproduce the real spectrum environment, we employ high spectral resolution sensing by setting the resolution bandwidth parameter in the utilized spectrum analyzer equal to 10 kHz. This value sets correspondingly the frequency bin size equal to 4 kHz, which combined with the increased sensitivity of our spectrum analyzer (thermal noise floor below -140 dBm/10 kHz), allows for the detection of even narrowband practical transmitted signals. To balance the trade-off between overestimating and underestimating spectrum utilization in our measurement setup, γ threshold is determined for each frequency bin independently, by combining the 5-dB criterion approach [1] and the PFA 1% criterion [1] (see also Section 2). In addition, we should note that the 300 to 2,700 MHz frequency band was split to 50 MHz sub-bands that were swept sequentially every 30 s.

4.1 Athens metropolitan area spectrum occupancy survey

To obtain a better insight on the local spectrum environment, as well as identify frequency bands which are most appropriate and challenging for opportunistic secondary exploitation, we first conducted a wideband spectrum occupancy survey in the metropolitan area of Athens, Greece. In line with relevant surveys performed in various locations worldwide [1], the occupancy state of certain bands of interest was periodically determined (for 7 days) through energy detection, using the aforementioned sensing architecture. Special attention was paid on the 300 to 2,700 MHz frequency range, usually referred to as sweet spot due to its intrinsic characteristics and significant services already deployed, which is attractive for SUs and poses significant challenges to the operation of CR networks, mainly regarding spectrum usage and available white space opportunities.

Considering the spectrum allocation formally enacted in Greece by the Hellenic Telecommunication and Post Commission [14], we determined the minimum, maximum, and average occupancy percentage of each nominal band through processing the entire 7-day measurements set. As clearly depicted in Figure 4, spectrum is significantly under-utilized even in the crowded metropolitan area of Athens (of a 4-million population), hence, enabling the potential application of CR systems in a feasible and attractive manner for users and operators. However, two distinct patterns in the utilization of the examined bands were identified. On the one hand, there are bands - like the TV ones - whose sub-bands were found either completely unused or constantly occupied over time. On the contrary, bands like those allocated to the downlink of the Global System for Mobile communications (GSM-900), the Digital Cellular System (DCS-1800) and the Universal Mobile Telecommunication System (UMTS/3G) present significantly more volatile and challenging behavior. Specifically, some base stations are switched off at night - when user demands are not so intensive - for protecting the corresponding equipment, setting idle significant spectrum blocks. This disruptive nature of spectrum occupancy fits better the CR technology principles, while rendering the secondary exploitation of the resulting white spaces really challenging in practice. In light of these results, in the rest of this paper, we pay extra attention on the latter bands, and in particular to the GSM-900, for developing and validating the presented correlation-based inference methodology and its benefits.

4.2 Demonstration of the correlation-based inference methodology

Focusing on frequencies that according to our conducted spectrum survey present volatile and, thus, challenging behavior, hereinafter, we examine a 5-MHz spectrum chunk owned by GSM providers. Figure 5 exhibits the

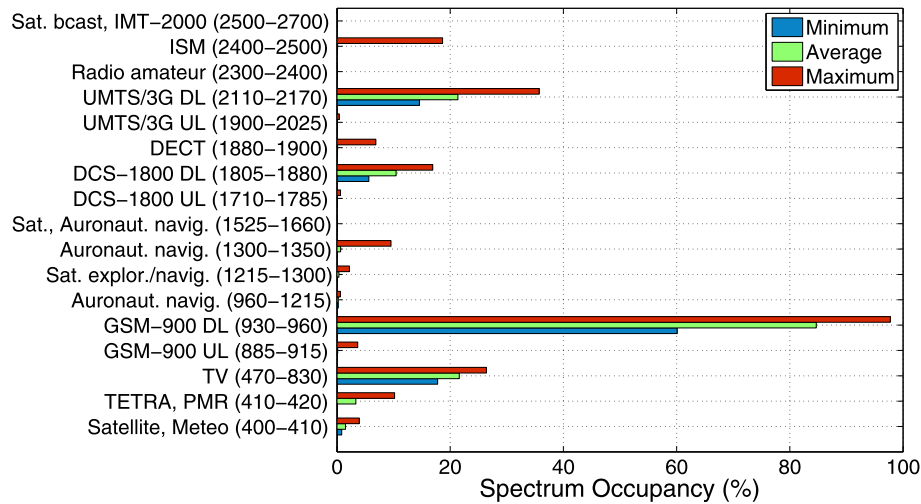


Figure 4 Spectrum occupancy obtained for bands of interest (bands expressed in megahertz).

observed correlation among the corresponding frequency bins, assuming power measurement data collection for a single day (i.e., 5,760 samples per bin). Clearly, significantly high power correlation is observed within some sequential frequency bins, thus, validating our initial expectations as described in Figure 1. Precisely, high correlation occurs in consecutive frequency bins spanned by a single technology specific PU channel, whereas the observed correlation sharply decreases at its boundaries. This behavior can be attributed to the fact that the existing licensed transmissions are in their majority served by contiguously allocated frequency bands, while the computed correlation of power measurements sharply decreases (tends to zero) for frequency bins at which the impact of noise is significant. In this manner, consecutive frequency bins can be clustered together into discrete groups by properly inferring their underlying dependency.

The interpretation of the correlation coefficient and the respective threshold r_{thr} , above which strong dependency

can be inferred, mainly relies on the context and purposes of each examined measurement set. In our case, this dependency does appear since each PU signal spans a superset of frequency bins (i.e., a 200-kHz GSM channel spans over multiple 4 kHz bins), with each such block constituting a nominal channel assigned to a PU. Therefore, as shown in the example of Figure 5, the power values of frequency bins that correspond to the same technology channel present extremely high correlation values and thus, the selection of a high threshold r_{thr} , e.g., 0.75, allows for inferring correlated blocks of consecutive frequency bins with well-defined boundaries.

As already stated, among the key objectives of this work is to infer and decide stable channels consisting of multiple frequency bins with coherent availability behavior clustered together. Based on the already acquired power measurements within the previously examined 5-MHz spectrum chunk, Figure 6 presents the power values detected at the corresponding frequency bins within the

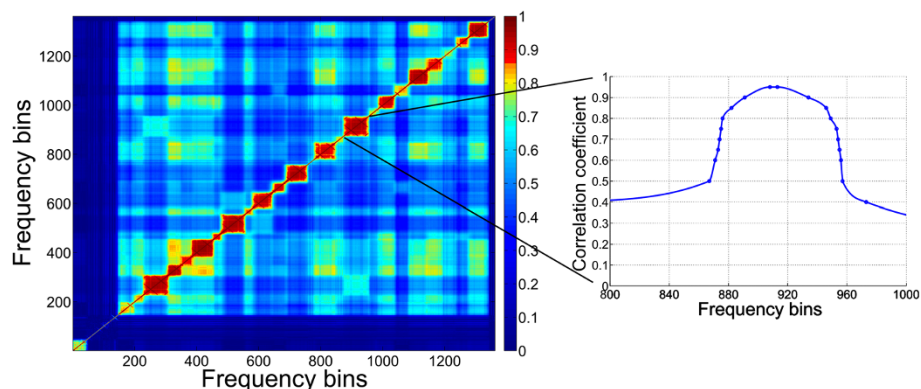


Figure 5 Correlation of power measurements of frequency bins.

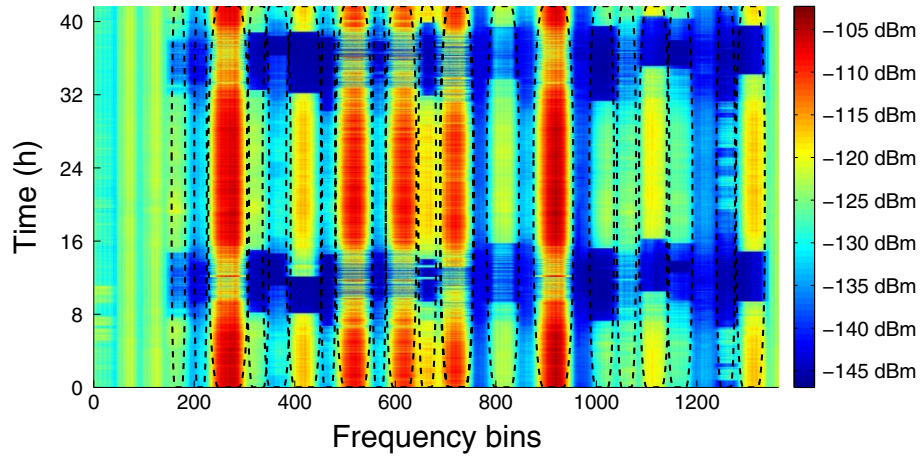


Figure 6 Correlated blocks of frequency bins based on power measurements. Dashed lines denote clustered frequency bins constituting potential SU-friendly white spaces.

sensing interval, while the dashed lines highlight groups of bins where the power correlation coefficient between each pair is greater than 0.75. Notably, the discovered correlated datasets are in line with the underlying primary system technology channel allocation with some boundary overlaps due to adjacent channel power leakages. Therefore, the analyzed correlation-based inference methodology can eventually serve as a tool for SUs to raise their wireless environment awareness and accurately reconstruct the spectrum landscape at their specific local areas.

It should be noted that assuming a correlated block of \mathcal{N} frequency bins, all entries of the corresponding $\mathcal{N} \times \mathcal{N}$ correlation matrix have to be greater than our pre-defined threshold. Even though the correlation matrices are always symmetric, hence in practice requiring the computation only of half of their entries, the computational overhead for such a procedure is of complexity $O(\mathcal{N}^2)$. This overhead can be further reduced in practice (especially for cases similar to our wideband sensing experimentation where the total number of examined frequency bins, \mathcal{N} , is extremely large) by leveraging problem-specific observations, such as from Figure 5. In particular, considering that a correlated block consists of consecutive frequency bins and assuming a maximum finite number of bins W (dubbed “window”) that a PU signal could span, correlation properties can be intelligently computed following a sliding window approach. In this manner, by progressively scanning the wideband-sensed spectrum, correlation properties are locally estimated for frequency bins residing in the current ‘window,’ thus, reducing the computational overhead to $O(\mathcal{N} \cdot W)$, where $W \ll \mathcal{N}$.

Finally, since the sample correlation coefficient, r , is affected by the sample size, we further analyze our estima-

tions towards ensuring the validity of the inferred statistical correlations. To compute a confidence interval estimation of r with respect to the size of the paired samples and the correlation coefficient ρ (Equation 6), a large sample approximation can be used based on Fisher’s transformation [10]. More specifically, let us consider Fisher’s transformation

$$z = \frac{1}{2} \ln \frac{1+r}{1-r} = \text{artanh}(r) \quad (8)$$

and assume that the number of samples, \mathcal{I} , is large. It can be shown that the distribution of random variable z is approximately normal [10] with mean $m_z \simeq \frac{1}{2} \ln \frac{1+\rho}{1-\rho}$ and variance $\sigma_z^2 \simeq \frac{1}{\mathcal{I}-3}$. Therefore, the γ -confidence interval for the sample correlation coefficient, r , can be approximately estimated by

$$\Pr\{\rho_1 < \hat{r} < \rho_2\} = \gamma, \quad (9)$$

where

$$\rho_1 = \frac{\exp 2z_1 - 1}{\exp 2z_1 + 1} \text{ and } \rho_2 = \frac{\exp 2z_2 - 1}{\exp 2z_2 + 1} \quad (10)$$

$$z_1 = m_z - \frac{z_{1+\gamma/2}}{\sqrt{\mathcal{I}-3}} \text{ and } z_2 = m_z + \frac{z_{1+\gamma/2}}{\sqrt{\mathcal{I}-3}}, \quad (11)$$

and $z_{1+\gamma/2}$ denotes the normal percentile.

Based on Equations 9, 10, and 11 and assuming collected data of a single day (i.e., 5,760 samples), we can calculate the confidence intervals for the sample correlation coefficient at different γ confidence levels. However, even though Fisher’s transformation is commonly used in statistical packages, the asymptotic variance of z is

independent on ρ only for particular structure of the parent bivariate distribution (namely, the underlying distribution of the data), e.g., bivariate normal distributions [15,16]. To validate the accuracy of our estimation, we have also estimated confidence intervals using the bootstrap method [16,17]. Table 2 demonstrates indicative results for correlation confidence intervals when applying both of the aforementioned techniques on the acquired power measurement data. It is observed that our sample size is sufficiently large such that ρ is appropriately estimated by r and therefore, valid results can be obtained.

5 Inference methodology exploitation scenarios

By exploiting correlation properties and assembling groups of frequency bins characterized by coherent behavior, SUs can be efficiently assisted on discovering more and larger white spaces that are also characterized by potentially lower reconfiguration overhead. This in turn can enable cumulative increase of spectrum utilization, as well as higher quality of service capabilities for SU applications. In the following, we explain such obtained benefits within the framework of our proposed inference methodology.

5.1 Improving white space discovery

As already explained, in energy-detection-based sensing, spectrum is typically sensed in a single stage by sequentially and exhaustively examining the occupancy of consecutive frequency bins. In order to avoid falsely estimating spectrum occupancy, the size of these bins is required to be small [1], which however, slows down the overall procedure (as shown in Figure 7). Although the selection of larger frequency bins accelerates sensing, it significantly affects the resulting sensing quality, since larger frequency blocks have higher probability to be determined occupied. For instance, a large frequency block (bin) that includes only a narrowband PU transmission would be determined occupied (since the detected energy is greater than the respective threshold) even if it is idle for most of its frequency range. Coarse-fine sensing approaches [5,6]

emerge as the most appropriate for balancing the aforementioned conflicting objectives. In the general case, these approaches operate by initially sweeping the spectrum with large-sized (coarse) frequency bins for quickly detecting those for which there exist strong indications that they are idle (e.g., by relaxing the decision criterion for spectrum occupancy). Algorithm 1 provides the pseudo-code for a conventional coarse-fine sensing implementation. To regain some lost accuracy, a secondary stage with finer sensing, i.e., sweeping with smaller-sized frequency bins or even applying more advanced feature detection-based sensing schema, is performed on idle coarse bins for obtaining spectrum occupancy estimation with higher confidence. Algorithm 1 describes the basic idea of these techniques, even though slight variations have been proposed to date in the literature (see Section 6.2).

Algorithm 1 Conventional coarse-fine sensing

Input: *Spectrum*, *CoarseBinSize*, *FineBinSize*, *CoarseBinNoiseLevel*, *FineBinNoiseLevel*

Output: Set of *FineBin* found *Idle*

Coarse sensing:

- 1: Divide *Spectrum* to *CoarseBins* acc. to *CoarseBinSize*
- 2: **for each** *CoarseBin* **do**
- 3: measure *Power*
- 4: **if** *Power* < *CoarseBinNoiseLevel* **then**
- 5: *CoarseBin* found *Idle*
- 6: **end if**
- 7: **end for**

Fine sensing:

- 8: **for each** *Idle CoarseBin* **do**
 - 9: Divide *CoarseBin* to *FineBins* acc. to *FineBinSize*
 - 10: **for each** *FineBin* **do**
 - 11: measure *Power*
 - 12: **if** *Power* < *FineBinNoiseLevel* **then**
 - 13: *FineBin* found *Idle*
 - 14: **end if**
 - 15: **end for**
 - 16: **end for**
-

Table 2 Examples of correlation confidence intervals at 95% confidence level

| Measurement bins | r Value | Fisher's transformation | Bootstrap |
|------------------|-----------|--------------------------|---------------------------|
| 1 and 910 | 0.0263 | $0.0004 < \rho < 0.0521$ | $-0.0018 < \rho < 0.0509$ |
| 800 and 910 | 0.5020 | $0.4824 < \rho < 0.5210$ | $0.4795 < \rho < 0.5237$ |
| 900 and 910 | 0.9404 | $0.9374 < \rho < 0.9433$ | $0.9355 < \rho < 0.9448$ |
| 909 and 910 | 0.9894 | $0.9889 < \rho < 0.9900$ | $0.9883 < \rho < 0.9904$ |

Although already efficient, coarse-fine sensing approaches could further benefit from the correlation-based inference methodology proposed in this work. In particular, when a frequency bin is detected idle during the coarse sensing stage and it does fall into a sub-band that is inferred coherent (and thus occupied by a single PU channel) by our mechanism, the fine sensing step should be applied over this entire sub-band (rather than the coarse bin), since it is more likely for the entire sub-band to be

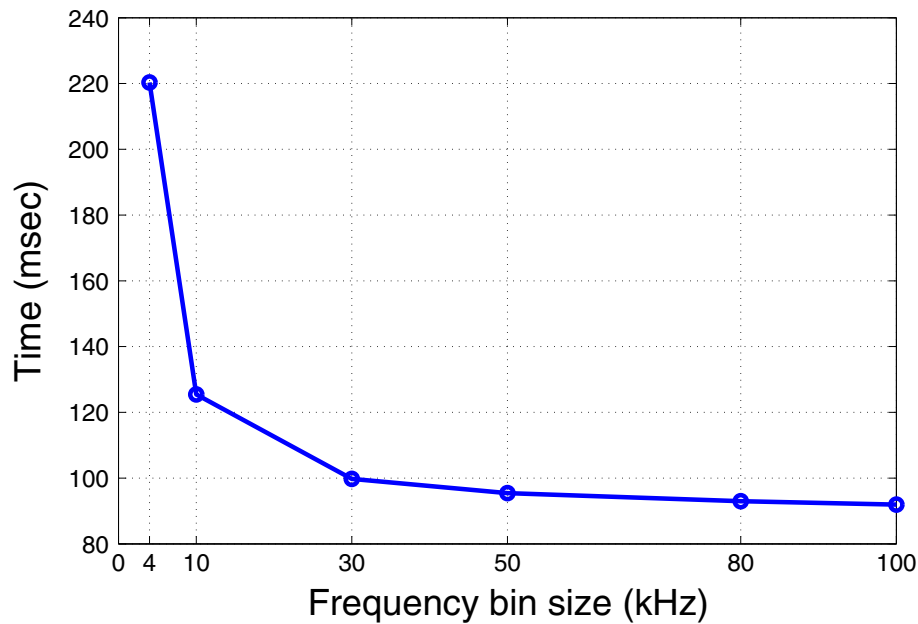


Figure 7 Sweep duration against frequency bin size (examining 50-MHz spectrum chunk using the available sensing architecture).

idle. This modification, presented in detail in Algorithm 2, can provenly enable coarse-fine sensing approaches to detect significantly larger amounts of white spaces, as well as larger contiguous - and thus more appropriate for secondary exploitation - spectrum holes.

Algorithm 2 Enhanced coarse-fine sensing

Input: *Spectrum*, *CoarseBinSize*, *FineBinSize*, *CoarseBinNoiseLevel*, *FineBinNoiseLevel*, *InferredTechChannels*

Output: Set of *FineBin* found *Idle*

Coarse sensing:

as in *Algorithm 1*.

Enhanced fine sensing:

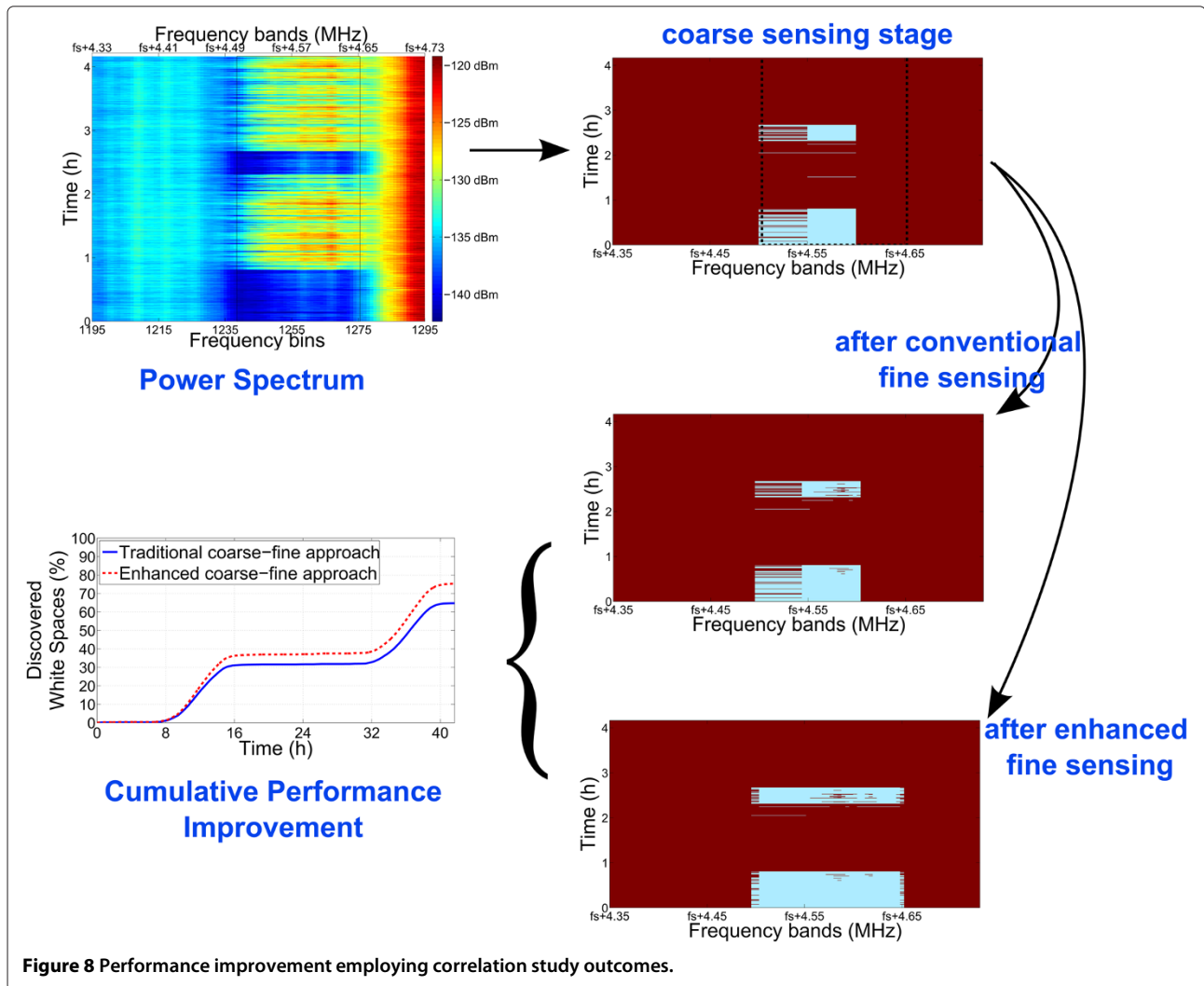
```

8: for each Idle CoarseBin do
9:   if CoarseBin spans inside a InferredTechChannel then
10:    Divide InferredTechChannel to FineBins acc. to FineBinSize
11:   else
12:    Divide CoarseBin to FineBins acc. to FineBinSize
13:   end if
14:   for each FineBin do
15:     measure Power
16:     if Power < FineBinNoiseLevel then
17:       FineBin found Idle
18:     end if
19:   end for
20: end for

```

For demonstration purposes, we focus on a 400-kHz part of the spectrum chunk examined in Section 4.2, at which the proposed correlation-based inference methodology has been employed resulting in frequency bins clustering. Subsequently, we collect new sensing data based on both the conventional and enhanced coarse-fine sensing approaches in order to compare their performance and confirm the improvement in discovery of white spaces induced by the acquired knowledge.

As depicted in Figure 8, the proposed methodology has successfully detected a group of frequency bins that present coherent behavior (represented by the dashed lines) which corresponds to an underlying GSM technology-specific channel that is indeed included in the examined power spectrum block. At first, we apply the conventional coarse-fine sensing technique presented in Algorithm 1, employing frequency bins of size equal to 50 and 4 kHz for the coarse and fine sensing steps, respectively. In parallel, the enhanced version presented in Algorithm 2 is applied (adopting equal frequency bin values) and thus, the entire coherent spectrum band (denoted by dashed lines) is examined, inside which a coarse bin is found idle at the first coarse step. Figure 8 depicts the performance of the involved steps, by which it is evident that in the former case the actual spectrum occupancy is poorly approximated, while in the latter, a more educated view of the underlying spectrum environment is obtained with significantly more white space bands being detected. In this manner, our suggested enhancement can result in less lost transmission opportunities, while the larger sizes of white spaces render them more attractive for practical exploitation by SUs, especially in quality of service demanding



applications. Finally, Figure 8 demonstrates the cumulative amount of detected white spaces (with respect to the total sensing duration), comparing the overall performance of each method when applied to the entire 5-MHz spectrum chunk mentioned above.

5.2 Enhancing the channel selection process

As stated in [18], the time-varying nature of PU activity causes channel availability variations at SUs and, thus, triggers frequent network reconfigurations that entail service disruptions, quality of service and experience reduction, lost transmission opportunities, and spectrum under-utilization as well as packet losses. SUs could avoid such demanding reconfigurations and increase their overall performance by showing preference in selecting secondary channels that are less prone to PU activity alterations.

However, the definition of a secondary channel (namely, the pair of central frequency (f_c) and bandwidth (BW),

which in turn can be mapped into a set of frequency bins) is not trivial within an opportunistic unlicensed spectrum access model and without assuming *a priori* knowledge of the primary system transmission details. Since the result of spectrum sensing characterizes the instantaneous spectrum availability, it is hard to select/predict sets of bins (i.e., secondary channels) that will be in long term robust to PU alterations.

Towards this direction, our proposed methodology can be directly exploited since the secondary channels can be selected/defined as the inferred blocks of bins that exhibit high correlation and are affected by a single PU. In this manner, each defined secondary channel can be opportunistically exploited by SUs according to the corresponding (unique) PU activity. On the contrary, an alternative secondary channel that spans over multiple technology-specific channels is inevitably affected by multiple independent PUs and thus, it must be released whenever one of the respective PUs is activated. Even though the real

availability depends on the activity of the respective PUs, defining secondary channels at the shadow of multiple PUs contributes, in general, to lower channel availability, aggravating the secondary network performance. It is noted that similar observations have been theoretically proven in [19], where cohesive white spaces are preferred for SUs, since they exhibit more stable behavior.

To validate in practice the benefits of our presented methodology on secondary channels selection/definition, Table 3 demonstrates indicative experimental results for 1-day duration in the previously examined 400-kHz spectrum chunk. In particular, we denote with central frequency f_c and bandwidth BW the secondary channel that is defined according to our inferred correlated group of frequency bins and then, we modify the secondary channel selection by slightly shifting it in the frequency domain as well as re-adjusting its bandwidth. Notably, the complement of the observed duty cycle of the secondary channel, namely, the probability to find it idle, is greater (almost double) when the central frequency is equal to f_c . The main reason is that by slightly shifting the secondary channel in frequency domain, the availability of the involved frequency bins is affected by the activity of more than one independent PU. Additionally, it is obvious that when the bandwidth of a secondary channel is within the ranges of the inferred coherent sub-band, the complement of its observed duty cycle remains significant, while when increasing its width more than the coherent region it is significantly reduced, since the channel is affected by the activity of multiple PUs. Therefore, exploiting the proposed methodology and recognizing the frequency margins of PU transmissions, SUs are able to select coherent secondary channels, hence, achieving increased expected availability and experiencing less reconfiguration costs.

5.3 Impact on theoretical approaches

Several theoretical approaches rely on the existence of secondary channels, which can be opportunistically accessed. However, in practical cases and without assuming external knowledge, spectrum holes are dynamically formatted (both in size and spectrum location), rendering difficult the definition of convenient secondary channels. Such an obstacle can be overcome by obtaining knowledge on the initially unknown PU spectrum allocation and consequently, our methodology can efficiently contribute to this

direction. This knowledge also enables SUs to estimate the expected length of white spaces that could potentially be detected at each band of interest and hence, determine transmission policies based on their communication needs. For instance, inferring technology channels of length equal to 200 kHz (e.g., GSM case) allows SUs to expect larger white spaces only at the cost of significantly higher reconfiguration overhead due to reasons explained in Section 5.2.

In addition, the proposed correlation-based inference methodology can also act supplementary to any existing or future approach that is based on correlation assumptions in frequency domain. Studies like [20] and [21] aim at accelerating the overall spectrum sensing procedure by assuming the occupancy of adjacent frequency bins correlated. However, both works do not deal with practical considerations regarding the methodology for obtaining full knowledge on the respective correlation coefficients values. Similar issues hold even in more advanced studies, like [19], where correlated occupancy is manually validated by observation on adjacent bins for a limited frequency range. To shed some more light on these assumptions, our methodology delves into the correlated behavior of neighboring frequency bins in a more educated and realistic manner and, thus, allows for any future work to exploit relevant observations with higher confidence. Additionally, it drives future approaches towards reflecting reality in a more accurate fashion, as well as avoiding assumptions like those made in [22] regarding the absence of any correlation in the occupancy of neighboring frequency bins.

6 Related work

6.1 Correlation properties in the frequency domain

Although numerous studies in modern literature exploit the correlation properties of channels in the time or space domains, only a few works have actually considered such properties in the frequency domain.

More specifically, in [12] the correlated occupancy state of neighboring channels was practically validated, with the ulterior view to accurately predict their future behavior. Using a spectrum analyzer, employing energy detection, as well as focusing on the GSM band, the occupancy state of subsequent 200-kHz-sized GSM channels was first determined. In the sequel, the correlation coefficients between

Table 3 Probability to find idle various selected/defined secondary channels

| Bandwidth | Central frequency | | | | |
|-------------|------------------------|------------------------|--------|------------------------|------------------------|
| | $f_c - 80 \text{ kHz}$ | $f_c - 40 \text{ kHz}$ | f_c | $f_c + 40 \text{ kHz}$ | $f_c + 80 \text{ kHz}$ |
| BW - 40 kHz | 0.3142 | 0.5046 | 0.5578 | 0.3282 | 0.2590 |
| BW | 0.3062 | 0.3890 | 0.5394 | 0.2606 | 0.2588 |
| BW + 40 kHz | 0.2934 | 0.3140 | 0.3282 | 0.2588 | 0.2586 |

the resulted occupancy time series of every two neighboring such channels were computed, yielding that the occupancy of all GSM channels allocated to the GSM band is highly correlated. In [19] the correlated occupancy of adjacent channels was also examined towards optimizing channel selection procedure. However, the occupancy correlation was only manually determined (i.e., no formal methodology was applied) by employing energy detection and observing real trace data collected over a limited frequency range. In [13] the behavior of the duty cycle (occupied-unoccupied behavior) of neighboring frequency bins that belong to certain parts of PU signals was practically determined as correlated through real-world experiments.

The main weakness of the abovementioned approaches is that both channel occupancy and duty cycle data are sensitive to the employed γ threshold and, thus, prone to miss detection and false-alarm errors jeopardizing the accuracy of the obtained results. A more accurate, and closer to our work, estimation is employed in [23], since correlation properties are estimated over the measured power spectral density values. However, all these approaches have investigated the underlying correlation properties in the frequency domain from a different perspective and not by aiming at frequency bin grouping and inferring the ranges of typically unknown technology-specific channels that are defined by primary systems. In addition, contrary to our study, the majority of the above approaches consider large frequency bins in their experiments, which increases the potential failure to discover correlation properties, since different PU signals can occupy different parts of a single frequency bin.

It should be finally noted that in works like [20] and [21], which propose new sensing schemes of increased performance, the occupancy of adjacent channels is simply assumed correlated, while none of them validates this assumption in practice or even define a formal methodology towards this direction.

6.2 Coarse-fine sensing approaches

The first coarse-fine sensing approaches that emerged in the literature were exclusively based on the energy detection technique and, thus, were called multi-resolution techniques, due to their sole sensing resolution difference. Although based on unrealistic assumptions, the work in [5] proves that a two-stage approach can outperform common one-stage high-resolution sensing techniques, under certain circumstances. In [6] two-stage sensing is further accelerated by using multiple antennas, at the cost of increased implementation and hardware complexity. Works in [24] and [25] were among the first of such multi-resolution studies, where the coarse-fine spectrum sensing was conducted in the analog domain (based on the wavelet transform coefficient). However, FFT-based

approaches are generally more preferable in practice due to their simpler hardware implementation. In line with the latter statement, in [26] a more efficient and less computationally complex FFT is proposed, which is suitable for application in two-stage spectrum sensing procedures. Even the IEEE 802.22 task force has identified the importance and suitability of such techniques, having already evaluated their benefits for possible integration in the respective cognitive radio-enabled protocol [27]. In a similar manner, a relevant technique [28] is proposed in the context of LTE protocol - which realizes practically cognitive radio concepts - for enhancing the white space detection procedure, introducing novel sensing techniques for both coarse and fine stages. Finally, works like [29] aim at determining the optimal frequency block size for both coarse and fine sensing stages while requiring prior knowledge on the parameters of the utilized equipment, the used sensing algorithm, and the expected signals to be sensed. However, the imposed requirements are considered quite restrictive, thus rendering their application difficult in practice.

In another major category, coarse sensing is based on energy detection, while fine stage on more advanced feature detection techniques. For instance, in [30] spectrum is divided in pre-defined channels and the channel with the lowest energy (as determined in the coarse stage) is examined by applying an advanced cyclostationary feature detection sensing technique. In [31] channels are examined sequentially, and every channel whose energy in the coarse stage does not exceed a pre-defined threshold is immediately examined in detail using a cyclostationary feature detection sensing technique. In a similar manner, in [32] energy detection and cyclostationary feature detection are applied in the coarse and fine sensing stages, respectively, but the latter is omitted when the decision at the former is sufficient to characterize the channel's occupancy with significant confidence.

7 Conclusions

In this paper we presented and experimentally validated a methodology for improving SU operation by raising knowledge regarding the real spectrum environment. The correlation properties in power measurements among neighboring frequency bins are the cornerstone of our study, while their intelligent exploitation can enhance significantly white space discovery and improve the selection of appropriate secondary channels. Conducting a real spectrum occupancy survey, we utilized real data towards validating the proof-of-concept of our methodology and evaluating its potential benefits. According to this real experimentation, our study aims at raising practical issues in the field of cognitive radio networks, while allowing for future approaches to further utilize the obtained outcomes.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

S. Arkoulis and E. Anifantis were partially funded by NTUA (ELKE Research Funds) and the Foundation for Education and European Culture (IPEP), respectively.

Author details

¹School of Electrical and Computer Engineering, National Technical University of Athens (NTUA), Zografou, Athens 15780, Greece. ²Institute of Communications and Computer Systems (ICCS), Zografou, Athens 15780, Greece.

Received: 14 September 2013 Accepted: 10 January 2014

Published: 28 January 2014

References

- Lopez-M Benitez, F Casadevall, Methodological aspects of spectrum occupancy evaluation in the context of cognitive radio, in *European Wireless Conference 2009*, Aalborg, 17–20 May 2009, pp. 199–204
- Y Zhang, J Zheng, H-H Chen, *Cognitive Radio Networks: Architectures, Protocols, and Standards*. (CRC, Boca Raton, 2010)
- H Sun, A Nallanathan, C-X Wang, Y Chen, Wideband spectrum sensing for cognitive radio networks: a survey. *Wireless Commun. IEEE*. **20**(2), 74–81 (2013)
- T Yucek, H Arslan, A survey of spectrum sensing algorithms for cognitive radio applications. *Commun. Surv. Tutorials IEEE*. **11**(1), 116–130 (2009)
- L Luo, NM Neihart, S Roy, DJ Allstot, A two-stage sensing technique for dynamic spectrum access. *Wireless Commun. IEEE Trans.* **8**(6), 3028–3037 (2009)
- NM Neihart, S Roy, DJ Allstot, A parallel, multi-resolution sensing technique for multiple antenna cognitive radios, in *IEEE International Symposium on Circuits and Systems*, (IEEE Piscataway, 2007), pp. 2530–2533
- P Pawelczak, K Nolan, L Doyle, SW Oh, D Cabric, Cognitive radio: ten years of experimentation and development. *Commun. Mag. IEEE*. **49**(3), 90–100 (2011)
- H Kim, M Suzuki, M Minami, H Morikawa, Spectrum sensor for distributed spectrum sensing, in *IEEE Consumer Communications and Networking Conference (CCNC)*, (IEEE Piscataway, 2012), pp. 44–45
- AV Oppenheim, RW Schaffer, *Discrete-Time Signal Processing*, 3rd edn. (Prentice Hall, Englewood Cliffs, 2009)
- A Papoulis, *Probability & Statistics*, vol. 2. (Prentice-Hall, Englewood Cliffs, 1990)
- National Instruments Corp., 2.7 GHz RF Vector Signal Analyzer with Digital Downconversion. www.ni.com/pdf/products/us/cat_vectorsignalanalyzer.pdf. Accessed 15 Aug 2013
- S Yin, D Chen, Q Zhang, M Liu, S Li, Mining spectrum usage data: a large-scale spectrum measurement study. *Mobile Comput. IEEE Trans.* **11**(6), 1033–1046 (2012)
- M Wellens, J Riihijärvi, P Mähönen, Empirical time and frequency domain models of spectrum use. *Phys. Commun.* **2**(1), 10–32 (2009)
- Hellenic Telecommunications and Post Commission. www.eett.gr. Accessed 15 Aug 2013
- D Hawkins, Using U statistics to derive the asymptotic distribution of Fisher's Z statistic. *Am. Stat.* **43**(4), 235–237 (1989)
- NJ Cox, Speaking stata: correlation with confidence, or Fisher's Z revisited. *Stat. J.* **8**(3), 413–439 (2008)
- TJ DiCiccio, B Efron, Bootstrap confidence intervals. *Stat. Sci.* 189–212 (1996)
- S Arkoulis, E Anifantis, V Karyotis, S Papavassiliou, N Mitrou, On the optimal, fair and channel-aware cognitive radio network reconfiguration. *Comput. Netw.* **57**, 1739–1757 (2013)
- D Xu, E Jung, X Liu, Optimal bandwidth selection in multi-channel cognitive radio networks: how much is too much?, in *3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (IEEE Piscataway, 2008), pp. 1–11
- L Luo, S Roy, Modelling and analysis of detection time trade-offs for channel searching in cognitive radio networks. *IET Commun.* **6**(8), 819–827 (2012)
- K Hossain, B Champagne, Wideband spectrum sensing for cognitive radios with correlated subband occupancy. *Signal Process. Lett. IEEE*. **18**(1), 35–38 (2011)
- M Wellens, A de Baynast, P Mahonen, Performance of dynamic spectrum access based on spectrum occupancy statistics. *Commun. IET*. **2**(6), 772–782 (2008)
- A Palaos, J Riihijärvi, O Holland, A Achtzehn, P Mahonen, Measurements of spectrum use in London: exploratory data analysis and study of temporal, spatial and frequency-domain dynamics, in *IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN)*, (IEEE Piscataway, 2012), pp. 154–165
- Y Hur, J Park, W Woo, K Lim, C-H Lee, H Kim, J Laskar, A wideband analog multi-resolution spectrum sensing (mrss) technique for cognitive radio (cr) systems, in *Proceedings of IEEE International Symposium on Circuits and Systems*, (IEEE Piscataway, 2006), p. 4
- Y Hur, G Park, C-H Lee, H Kim, J Laskar, A cognitive radio (cr)-based mobile interactive digital broadcasting application adopting a multi-resolution spectrum-sensing (mrss) technique, in *IEEE 66th Vehicular Technology Conference*, (IEEE Piscataway, 2007), pp. 1912–1916
- Q Zhang, AB Kokkeler, GJ Smit, An efficient multi-resolution spectrum sensing method for cognitive radio, in *Third International Conference on Communications and Networking in China*, (IEEE Piscataway, 2008), pp. 1226–1229
- SJ Shellhammer, Spectrum sensing in IEEE 802.22, in *IAPR Workshop on Cognitive Information Process*, Santorini, 9–10 June 2008
- MA Abdelmonem, M Nafie, MH Ismail, MS El-Soudani, Optimized spectrum sensing algorithms for cognitive LTE femtocells. *EURASIP J. Wireless Commun. Netw.* **2012**(1), 1–19 (2012)
- H Zamat, B Natarajan, Optimization of sensing receiver for cognitive radio applications. *EURASIP J. Adv. Signal Process.* **2009**, 5 (2009)
- W Yue, B Zheng, A two-stage spectrum sensing technique in cognitive radio systems based on combining energy detection and one-order cyclostationary feature detection, in *Proceedings of the International Symposium on Web Information Systems and Applications (WISA'09)*, Nanchang, 22–24 May 2009 pp. 327–330
- S Maleki, A Pandharipande, G Leus, Two-stage spectrum sensing for cognitive radios, in *2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, (IEEE Piscataway, 2010), pp. 2946–2949
- PR Nair, AP Vinod, KG Smitha, AK Krishna, Fast two-stage spectrum detector for cognitive radios in uncertain noise channels. *IET Commun.* **6**(11), 1341–1348 (2012)

doi:10.1186/1687-1499-2014-17

Cite this article as: Arkoulis et al.: Discovering and exploiting spectrum power correlations in cognitive radio networks: an experimentally driven approach. *EURASIP Journal on Wireless Communications and Networking* 2014 **2014**:17.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com