Novel spectrum sensing schemes for Cognitive Radio Networks

Cantabria University

Santander, May, 2015
The Advanced Signal Processing Group (GTAS, in Spanish) is part of the Communications Engineering Department of the University of Cantabria.

It is located at the E.T.S.I. Industriales y Telecomunicaciones, Avda Los Castros s/n. Santander 39005, SPAIN.
• Signal processing techniques for MIMO communication links.
  • CSI (Channel State Information) estimation techniques, synchronization, detection techniques,...
  • Capacity analysis of MIMO links.
  • Development of hardware MIMO testbeds and performance evaluation.

• Machine-learning techniques and their application to communications.
  • Kernel methods, neural networks and adaptive information processing systems.
  • Multivariate statistical techniques: PCA, CCA and ICA.
  • Nonlinear modeling and nonlinear dynamical systems (chaos).
Novel detection schemes for CR networks

- **Multi-antenna Bayesian Spectrum Sensing**


- **Robust KCCA detector for Cooperative Spectrum Sensing**

Multi-antenna Bayesian Spectrum Sensing

- At each sensing period: a Bayesian inference is applied.
- Priors for the spatial covariance and the probability of channel occupancy.
- Posterior are employed as priors for the next sensing frame.
- Simulations and experimental measurements.

P single-antenna PUs

CR (secondary user) with L antennas
Multi-antenna Bayesian Spectrum Sensing

• The spectrum sensing problem can be formulated as a binary hypothesis test as follows:

\[ H_1 : x_t[n] = H_t s_t[n] + v_t[n], \]
\[ H_0 : x_t[n] = v_t[n], \]

\( x_t \) is the acquired snapshot at time \( n \), \( s_t \) is the primary signal vector.

• Two different structures for the covariance matrix:

\[ H_1 : x_t[n] \sim \mathcal{CN}(0, R_t), \quad n = 0, \ldots, N - 1. \]
\[ H_0 : x_t[n] \sim \mathcal{CN}(0, D_t), \quad n = 0, \ldots, N - 1. \]

Under \( H_1 \), a \( L \times L \) covariance matrix \( R_t \) can be written as \( HH^H + D \), i.e., a rank-\( P \) matrix plus a scaled diagonal matrix.
Bayesian inference

- Prior distributions: a Bernoulli distribution, complex inverse wishart $CW^{-1}$ and the inverse-gamma $g_L^{-1}$.

\[
p(z_t) = \text{Bernoulli}(z_t|\tilde{\pi}_t) = \tilde{\pi}_t^z(1-\tilde{\pi}_t)^{1-z_t}
\]

\[
p(R_t) = CW^{-1}(R_t|\tilde{\eta}_t, \tilde{R}_t) = \frac{|\tilde{R}_t|^\frac{n_t}{2}|R_t|^{-\frac{n_t+L+1}{2}} \exp\left(-\frac{1}{2} \text{trace}(R_t^{-1}\tilde{R}_t)\right)}{2^{\frac{n_L}{2}} \Gamma_L(\tilde{n}_t)}
\]

\[
p(D_t) = G_L^{-1}(D_t|\tilde{\eta}_t, \tilde{D}_t) = \prod_{l=1}^{L} G^{-1}([D_t]_{ll}|\tilde{\eta}_t/2, [\tilde{D}_t]_{ll}/2)
\]

- Parameters of prior distributions: $\tilde{\pi}_t, \tilde{\eta}_t, \tilde{R}_t, \tilde{D}_t$.
- A non-informative prior at $t = 0$. 
Bayesian inference

- Exact posterior distribution of $z_t$, $R_t$ and $D_t$
  - Since the noise is Gaussian, the likelihoods $p(X_t|z_t=0, D_t)$ and $p(X_t|z_t=1, R_t)$ can be written:
    \[
    p(X_t|z_t=0, D_t) = \prod_{n=1}^{N} \mathcal{N}(x[n]|z_t=0, D_t),
    \]
    \[
    p(X_t|z_t=1, R_t) = \prod_{n=1}^{N} \mathcal{N}(x[n]|z_t=1, R_t).
    \]
  - Priors are conjugate and therefore the posterior distributions (conditioned on the channel state) have the same form as the prior
    \[
    p(R_t|X_t, z_t = 1) = \mathcal{W}^{-1}(R_t|\hat{n}_t, \hat{R}_t)
    \]
    \[
    p(D_t|X_t, z_t = 0) = \mathcal{G}_{L}^{-1}(D_t|\hat{m}_t, \hat{D}_t).
    \]
Bayesian inference

• Exact posterior distribution of $z_t$, $R_t$ and $D_t$

  where the posterior parameters depend on the observed data $X_t$ and are given by:

  \[
  \hat{n}_t = \tilde{n}_t + N \\
  \hat{R}_t = \tilde{R}_t + S_t \\
  \hat{m}_t = \tilde{m}_t + N \\
  \hat{D}_t = \tilde{D}_t + \text{diag}(S_t).
  \]

• When $z_t$ is marginalized, each unconditional posteriors becomes a convex combination of the posteriors for each hypotheses, yielding:

  \[
  p(R_t|X_t) = \hat{\pi}_t CW^{-1}(R_t|\tilde{n}_t, \tilde{R}_t) + (1 - \hat{\pi}_t) CW^{-1}(R_t|\tilde{n}_t, \tilde{R}_t) \\
  p(D_t|X_t) = \hat{\pi}_t GL^{-1}(D_t|\tilde{m}_t, \tilde{D}_t) + (1 - \hat{\pi}_t) GL^{-1}(D_t|\tilde{m}_t, \tilde{D}_t), \\
  p(z_t|X_t) = \text{Bernoulli}(z_t|\hat{\pi}_t)
  \]
Bayesian inference

• Exact posterior distribution of $z_t$, $R_t$ and $D_t$

\[
\hat{\pi}_t = \frac{p(X_t|z_t = 1)p(z_t = 1)}{p(X_t|z_t = 1)p(z_t = 1) + p(X_t|z_t = 0)p(z_t = 0)}.
\]

The probability of a transmitter being present given observations $X_t$: $\hat{\pi} = p(z_t = 1|X_t)$.

The channel is occupied when the collision probability $\hat{\pi}$ is below some desired threshold.
Bayesian inference over multiple frames

- Learning from past sensing frames

\[ \begin{align*}
\text{sensing period } t &: X_t \\
\begin{bmatrix} x_t[0] & \ldots & x_t[N-1] \end{bmatrix} & \cdots & \begin{bmatrix} x_{t+1}[0] & \ldots & x_{t+1}[N-1] \end{bmatrix} \\
\text{prior at } t &: p(D_t), p(R_t) \\
p(D_t | X_t), p(R_t | X_t) & \downarrow & \text{posterior at } t \\
p(D_{t+1} | X_{t+1}), p(R_{t+1} | X_{t+1}) & \downarrow & \text{posterior at } t+1 \\
p(Z_t = 1 | X_t) &= \hat{\pi}_t > th \\
p(Z_{t+1} = 1 | X_{t+1}) &= \hat{\pi}_{t+1} > th
\end{align*} \]

- The (unconditional) posteriors after processing the \( t \)-th frame summarize all the information observed so far.
- The posteriors obtained after processing a sensing frame are employed as priors for the next sensing frame.
Bayesian inference over multiple frames

- Problem: The posterior distribution are convex combination of the posterior under each hypotheses

\[
p(R_t | X_t) = \hat{\alpha}_t C \mathcal{W}^{-1}(R_t | \hat{\eta}_t, \hat{\kappa}_t) + (1 - \hat{\alpha}_t) C \mathcal{W}^{-1}(R_t | \hat{\eta}_t, \hat{\kappa}_t),
\]
\[
p(D_t | X_t) = \hat{\alpha}_t \mathcal{G}_L^{-1}(D_t | \hat{\eta}_t, \hat{\kappa}_t) + (1 - \hat{\alpha}_t) \mathcal{G}_L^{-1}(D_t | \hat{\eta}_t, \hat{\kappa}_t).
\]

- Thresholding-based approximation
  Priors can be obtained by truncating \( \hat{\alpha} \) to either 0 or 1 whichever it is closer.

- Kullback-Leibler approximation
  A more rigorous approach is given by minimizing the KL distance

\[
\{ \hat{\eta}_t, \hat{\kappa}_t \} = \underset{\eta_t, \kappa_t}{\text{argmin}} \text{KL}(p(D_t | X_t) \| q(D_t | X_t))
\]
\[
\{ \hat{\eta}_t, \hat{\kappa}_t \} = \underset{\eta_t, \kappa_t}{\text{argmin}} \text{KL}(p(R_t | X_t) \| q(R_t | X_t)).
\]
Bayesian inference over multiple frames

- Forgetting in non-stationary environments

  The channel may vary between consecutive frames, and a mechanism to forget past data is introduced.

  **Bayesian $\lambda$ forgetting**: prior distributions for frame $t+1$ is given by a “smooth” version of the posterior distribution (after processing frame $t$) and the original distribution for $R_t$ and $D_t$

  \[
  p(D_{t+1}|X_t) \propto p(D_t|X_t)^\lambda p(D_0)^{1-\lambda},
  \]
  \[
  p(R_{t+1}|X_t) \propto p(R_t|X_t)^\lambda p(R_0)^{1-\lambda}.
  \]

  With this forgetting step, the parameters of the prior distributions to be used for Bayesian inference at $t+1$ are given by,

  \[
  \hat{n}_{t+1} = \lambda \hat{n}_t + (1 - \lambda) \hat{n}_0
  \]
  \[
  \hat{R}_{t+1} = \lambda \hat{R}_t + (1 - \lambda) \hat{R}_0
  \]
  \[
  \hat{m}_{t+1} = \lambda \hat{m}_t + (1 - \lambda) \hat{m}_0
  \]
  \[
  \hat{D}_{t+1} = \lambda \hat{D}_t + (1 - \lambda) \hat{D}_0.
  \]
The algorithm only requires updating and storing $\hat{\mathbf{R}}_t$, $\hat{n}_t$, $\hat{\mathbf{D}}_t$, $\hat{m}_t$, from one frame to the next, it requires a fixed amount of memory and computation per sensing frame.

**Algorithm 1** Online Bayesian Multiantenna Sensing

1. Initialize Parameters: $\lambda$, $\hat{\mathbf{R}}_0$, $\hat{n}_0$, $\hat{\mathbf{D}}_0$, $\hat{m}_0$
2. for Frame $t = 1, 2, \ldots$ do
3.  Sense the medium $N$ times through $L$ antennas to get $\mathbf{X}_t$
4.  Exact posterior: Compute $\hat{\mathbf{R}}_t$, $\hat{n}_t$, $\hat{\mathbf{D}}_t$, $\hat{m}_t$ and $\hat{n}_t$
5.  Output $\hat{n}_t$, probability of a PU being present during $t$
6.  Compute the approximated posterior parameters using KL minimization or thresholding
7.  Forget: Compute $\tilde{\mathbf{R}}_t$, $\tilde{n}_t$, $\tilde{\mathbf{D}}_t$, $\tilde{m}_t$
8. end for
Simulation results

- PD for the Bayesian detector (using the two posterior approximations Bayes-KL and Bayes-T) and the GLRT vs. the number of sensing frames.

Stationary channel: $N=50$, $SNR=-8dB$

For a slowly time-varying channel
Simulation results

- PD for the Bayesian detector and the GLRT vs. the number of sensing frames

For a fast time-varying channel
Simulation results

- ROC curve for the Bayesian and GLRT detector

For stationary channel with SNR=−8dB and $\lambda = 1.0$

For slowly stationary channel with SNR=−8dB and $\lambda = 0.97$
Simulation results

- ROC curve for the Bayesian and GLRT detector

For fast time-varying channel with SNR=-8dB
\[ \lambda=0.97, \lambda_{ch}=0.10 \]
Simulation results

- Detection probability and false alarm probability versus SNR

For stationary channel with SNR=-8dB, $\lambda=1.0$, and $\lambda_{ch}=1$

For slowly stationary channel with SNR=-8dB, $\lambda=0.97$, $\lambda_{ch}=0.95$
Simulation results

- Detection probability and false alarm probability versus the SNR

For fast time-varying channel with SNR=-8dB and $\lambda = 0.95$ and $\lambda_{ch} = 0.1$
Experimental evaluation

Laboratory equipment: signal generators, oscilloscopes, spectrum analyzers

N210 Ettus devices with XCVR2450 daughterboard, two-antenna cognitive receiver compose of two N210 boards connected through a MIMO cable
Experimental evaluation

- Bayesian spectrum sensing

A PU accesses the channel according to a predefined pattern, and a SU (a CR user with two antennas) senses periodically the medium.

\[ N \] samples are acquired at each sensing period and stored in a \( 2 \times N \) matrix format:
Experimental measurements

- ROC curves for the Bayesian and GLRT detectors using the CR platform in a realistic indoor channel at 5.6GHz

Bayesian (squares), Sphericity (circles) and Hadamard (crosses) detectors, in a static environment. $N = 50$ and a SNR = $-7.3\,\text{dB}$, $\lambda=1.0$.

One-shot detectors (Sphericity and Hadamard detectors) show to be almost identical.
Experimental measurements

- A more challenging scenario: the experimental evaluation in a non-stationary environment i.e. slow time-varying and fast time-varying environment

- For time-varying scenarios, a beamforming at the TX side is implemented:

- ROC curves for the Bayesian (squares), Sphericity (circles) and Hadamard (crosses) detectors, in a slowly time-varying environment. $N = 50$, and a SNR $= -1.18$ dB.
Experimental measurements

- ROC curves for the Bayesian (squares), Sphericity (circles) and Hadamard (crosses) detectors, in a fast time-varying environment. \( N = 50 \) and a SNR = \(-2.5\)dB.
Conclusions

• A Bayesian framework employs a forgetting mechanism where the posterior for the unknown parameters $R_t$ and $D_t$ are used as priors for the next Bayesian inference.

• This scheme is evaluated under a stationary channel, slowly time-varying channel, and fast time-varying channel.

**For stationary environments:**

A Bayesian detector provides the best detection performance, since the unknown covariance matrices $(R_t$ and $D_t)$ remains constant.

A KL posterior approximation provides a best performance in comparison to the thresholding-based approximation.

Our simulation results and experimental measurements show to have a significant gain over one-shot GLRT detector by setting a forgetting factor $\lambda = 1.0.$
Conclusions

For non-stationary environments:

Bayesian scheme also shows a better performance over one-shot detectors.

A coarse approximation (thresholding-based approach) attain a better performance.

In this case, a small degradation in its performance is observed by setting a higher value for $\lambda = 1.0$.

- A Bayesian detector show the feasibility of learning efficiently the posteriors parameters to detect a PU signal under stationary and non-stationary environments.
A KCCA for Robust Cooperative Spectrum Sensing

- A kernel canonical correlation analysis (KCCA) technique is performed at the fusion center (FC).
- Statistical tests are extracted: decisions either at each SU (autonomously) or cooperatively at the FC.
- Simulations and experimental measurements.
Robust KCCA Spectrum Sensing Scheme

- We consider $M$ secondary users and a PU in the same area; and the signal model takes into account the presence of local interferences.

\[
p(r|\mathcal{H}_1) \neq \prod_{i=1}^{M} p_i(r_i|\mathcal{H}_1)
\]

\[
p(r|\mathcal{H}_0) = \prod_{i=1}^{M} p_i(r_i|\mathcal{H}_0)
\]

- The optimal detectors at each SU will be highly correlated, i.e. if SUs are either all under the null hypothesis or all under the alternative hypothesis.

- The proposal aims to find the non-linear transformations of the measurements that provides maximal correlation. These non-linear transformations are employed to decide if the measurements come from the distribution $p(r|H_1)$ or from $p(r|H_0)$. 
Operation of the KCCA scheme

• Our scheme starts with an initial cooperative learning stage where the sensors measurements are transmitted to the FC.

• Local statistics (near-optimal local decision functions) are extracted and broadcasted to the SUs, which can operate in one of two modes:

1. **Autonomous testing:**
   
   Each SU takes independent decisions based on its local test statistic.

2. **Cooperative testing:**
   
   Each SU transmits its local test statistic to the FC, where a global decision is finally made by combining the local test statistics.
Operation of the KCCA scheme

- Features extracted during the sensing period

\[ x_{in} = \left( f_{in}^1 f_{in}^2 \ldots f_{in}^{N-1} f_{in}^N \right)^T \]

- For each i-th sensor a data set is collected
Kernel Canonical Correlation Analysis

• Kernel-based learning

The data are transformed into a high-dimensional feature space:

\[ \Phi : x_{in} \rightarrow \Phi(x_{in}). \]

The inner product (in the feature space) can be calculated as positive definite kernel function \( k(.,.) \).

\[ \kappa(x_{ij}, x_{ik}) = \langle \Phi(x_{ij}), \Phi(x_{ik}) \rangle. \]

E.g. standard Gaussian kernel:

\[ \kappa(x_{ij}, x_{ik}) = \exp\left(-\frac{\|x_{ij} - x_{ik}\|^2}{2\sigma_i^2}\right), \]

Given a data set, a Gram matrix (or kernel matrix) \( K_i \) contains all possible inner products

\[ K_i(j, k) = \kappa(x_{ij}, x_{ik}) = \Phi(x_{ij})^\top \Phi(x_{ik}). \]
• Kernel Canonical Correlation Analysis for CSS

The pairwise canonical correlation between the data sets:

\[ \rho_{ij} = z_i^T z_j = \alpha_i^T K_i K_j \alpha_j, \quad \text{where} \; z_j = K_i \alpha_i. \]

A measurement of the correlation associated to the \( i \)-th data set:

\[ \rho_i = \frac{1}{M-1} \sum_{j=1, j \neq i}^{M} \rho_{ij}, \]

A generalized canonical correlation can be obtained as

\[ \rho = \frac{1}{M} \sum_{i=1}^{M} \rho_i. \]
Kernel Canonical Correlation Analysis

- The maximization of $\rho$ with respect to the canonical vectors subject to the energy of the canonical variates and the norm of the projectors can be solved yielding the following generalized eigenvalue problem.

$$\frac{1}{M} \mathbf{R} \boldsymbol{\alpha} = \beta \mathbf{D} \boldsymbol{\alpha},$$

where

$$\beta = \frac{1+(M-1)\rho}{M},$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{K}_1 \mathbf{K}_1 & \cdots & \mathbf{K}_1 \mathbf{K}_M \\ \vdots & \ddots & \vdots \\ \mathbf{K}_M \mathbf{K}_1 & \cdots & \mathbf{K}_M \mathbf{K}_M \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \mathbf{K}_1 (\mathbf{K}_1 + c \mathbf{I}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{K}_M (\mathbf{K}_M + c \mathbf{I}) \end{bmatrix}.$$

- Local and Global Tests:

$$T_i(x_i) = \sum_{j=1}^{N} \alpha_{ij} \tilde{k}(x_i, x_{ij})$$

Local test: $\alpha_{ij}$ refers to the j-th element of a canonical vector $\alpha_i$. A weighted sum of similarities.

$$T_i(x) = \sum_{i=1}^{M} T_i(x_i),$$

Global test: best one-dimensional approximation of the canonical variates.
Simulation Results

- Probability density functions (PDF), local statistics and ROC curves: noise and primary signal

PDF for the primary and noise signal at SU 1, and decision function (local statistics). Positive and negative values for the noise and primary signal respectively. SNR -5.3 dB

ROC curves for local decisions (at each SU) and centralized decisions (at the FC) using a KCCA and an energy detector
Simulation Results

- Probability density functions (PDF), local statistics and ROC curves: noise, interference and primary signal.

PDF and decision functions at SU 1: the primary signal is assigned negative values. SINR -8.5 approx.

ROC curves for local decisions (at each SU) and centralized decisions (at the FC) using a KCCA and an energy detector.
Simulation Results

- Probability density functions (PDF), local statistics and ROC curves: noise, interference and signal at SU 1 for SINR 7.45 dB

Two features extracted during the sensing period: the energy (at the left side) and kurtosis (right side).
Simulation Results

The corresponding ROC curves for local decisions (at each SU) and centralized decisions (at the FC) using the energy, the kurtosis, or both of them.
Experimental evaluation

- Testbed: two SUs, an interfering node (INT), a PU and the FC in the middle of them. All USRPs are synchronized by a pulse per second (PPS) signal.

- Measurement procedure: the PU transmits using two bands of frequency channels (2-4 MHz and 4-6 MHz), each SU senses a different band, and the INT node transmit randomly on any of the channels or on both
Experimental measurements

The corresponding ROC: SINR 0.63 dB at both SUs.
Experimental measurements

The received power corresponding to the noise and the PU signal in one of the SU have similar energy values.

The corresponding ROC: SINR -11.4 dB and -9.2 dB at the SU1 and SU2 respectively.
Experimental measurements

The corresponding ROC: SINR -6.3 dB and -5.1 dB at the SU1 and SU2 respectively
Conclusions

• The proposed approach has been evaluated under different scenarios in which noise or noise plus interference are present, and for which different features are extracted during the sensing period.

• Our approach operates in blind manner, and can be applied to time-changing environments, since it adapts itself by retraining from time to time.

• For scenarios with only noise and using only energy measurements: the KCCA and the energy detector attain the same performance, since the obtained tests are close to the optimal NP detector.

• In scenarios with noise plus interference, our KCCA detector obtains a significant gain over an energy detector.
Conclusions

• Regarding the experimental measurements: we corroborate the learning ability to detect the PU signal by exploiting the correlation among the received signals.

• In fact, more challenging cases not taken into account in our simulation environment are also addressed, e.g. different noise variance at each SU as well as the interference power received at each SUs.

• Our technique exhibits a much better performance than that of the energy detector as the interference level increases, since our KCCA framework exploits better the correlation of the received PU signal when more uncorrelated external interference is present.
Thank you for your attention

juliocesar@gtas.dicom.unican.es