

Spectrum Sensing in Cognitive Radio using Random Forest with Wavelet and Empirical Mode Decomposition

Gourav Kumar Gole¹, Dr. Shivangini Morya², Dr. Rajesh Kumar Nagar³

¹PhD Scholar, SAGE University, Indore, golegourav21@gmail.com

²Associate Professor, Department Of ECE, SAGE University, Indore,
drshivangini.morya@sageuniversity.in

³Associate Professor, Department of ECE, SAGE University, Indore, Errajesh973@gmail.com

Abstract

Employing a stochastic approach to cognitive radio, this study meticulously assesses the efficacy of cooperative spectrum sensing strategies, specifically leveraging the Random Forest machine learning paradigm. Comparative analysis is undertaken across two distinct feature extraction methodologies: wavelet transformation and empirical mode decomposition. The primary user signal is modelled as a 5G-compliant orthogonal frequency-division multiplexing (OFDM) waveform. The cooperative sensing framework employs a majority voting fusion rule. Simulation results demonstrate the effectiveness of both feature extraction techniques combined with Random Forest for robust spectrum sensing, highlighting their comparative performance under varying noise conditions.

Keywords: Cognitive Radio, Spectrum Sensing, Random Forest, Wavelet Transform, Empirical Mode Decomposition (EMD), 5G, OFDM, Cooperative Sensing, Machine Learning.

1. INTRODUCTION

The incessant escalation in the exigency for wireless communication services has precipitated a profound scarcity of spectral resources. Cognitive Radio (CR) technology emerges as an avant-garde paradigm poised to ameliorate this conundrum by permitting unlicensed Secondary Users (SUs) to opportunistically exploit licensed spectral bands during intervals of non-utilization by Primary Users (PUs) [1-3]. A quintessential function within the CR framework is spectrum sensing—a critical process entailing the precise detection of PU activity within specified frequency domains. Erroneous sensing outcomes can engender deleterious interference with PUs or engender suboptimal spectral utilization.

Conventional spectrum sensing methodologies, includes Energy Detection, Matched-Filtering, Cyclo-stationary Feature Detection etc [4-8]. These are encumbered by inherent limitations, particularly under situations of diminished Signal-to-Noise Ratio (SNR) or when the intrinsic characteristics of PU signals remain elusive [9]. To surmount these impediments, the deployment of sophisticated Artificial intelligence (AI) algorithms has garnered burgeoning interest. These avant-garde techniques possess the capacity to decipher intricate patterns embedded within received signals, thereby engendering more resilient, adaptive, and efficacious spectrum sensing mechanisms. This paper focuses on leveraging advanced machine learning for spectrum sensing, specifically employing the Random Forest (RF) classifier. We investigate and compare two powerful signal decomposition techniques for feature extraction: The Empirical Mode Decomposition (EMD) and Wavelet Transform (WT). The Wavelet Transform provides a time-frequency representation, allowing for analysis at different scales, while EMD adaptively decomposes a signal into Intrinsic Mode Functions (IMFs) depends on local characteristics. The performance of these methods will be evaluated in a cooperative sensing scenario, where multiple SUs collaborate to make a more reliable decision.

The subsequent discourse is structured as follows: Section 2 furnishes a contextual backdrop on cognitive radio, Random Forest, Wavelet Transform, and Empirical Mode Decomposition. Section 3 explicates the proposed systemic model and methodological approach. Section 4 delineates the simulation parameters. Section 5 gives analysis of simulation outcomes. The paper is summarized in section 6.

2. Background

2.1. Cognitive Radio and Spectrum Sensing

Cognitive Radio epitomizes an advanced, perceptive wireless communication paradigm capable of environmental awareness, adaptive learning, and dynamic modulation of transmission parameters to maximize spectral efficiency. At the core of this paradigm lies spectrum sensing, the inaugural and

quintessential phase within the cognitive cycle. This process entails discerning the presence or absence of a licensed Primary User (PU) within a designated frequency band. Such a challenge can be rigorously modelled as a binary hypothesis testing problem:

- H_0 : PU absent (noise only)
- H_1 : PU present (signal + noise)

Cooperative sensing enhances reliability by allowing multiple SUs to share their local sensing results with a fusion centre, which provide a combined decision. This approach mitigates the effects of fading, shadowing, and hidden node problems, leading to improved detection performance.

2.2. Random Forest Classifier

Random Forest [10] constitutes an ensemble learning paradigm for classification and regression, predicated on the aggregation of multiple decision trees constructed during the training phase. It yields the predicted class—via mode—(classification) or the average prediction (regression) across individual trees. This methodology mitigates overfitting and enhances predictive robustness relative to a solitary decision tree through the infusion of stochasticity manifested in two principal mechanisms:

1. Bagging (Bootstrap Aggregating): Each constituent tree is trained on a distinct bootstrap sample, fostering diversity and reducing variance.
 2. Random Feature Subset: During node splitting, only a randomly selected subset of features is evaluated, eschewing the exhaustive consideration of all features and thereby augmenting model generalization.
- This strategic incorporation of randomness fortifies the model's resilience against overfitting while bolstering its overall predictive fidelity. These characteristics make Random Forest a robust and powerful classifier suitable for complex, non-linear classification tasks like spectrum sensing.

2.3. Wavelet Transform for Feature Extraction

The Wavelet Transform (WT) decomposes signals of time domain to its frequency components with the resolution accorded to its scale [11]. Unlike the Fourier analysis, which provides only frequency information, WT provides both time and frequency localization. This multi-resolution analysis makes WT particularly effective for analyzing non-stationary signals, which are common in wireless communication. For spectrum sensing, features such as energy, entropy, or statistical moments can be extracted from the wavelet coefficients (approximation and detail coefficients) at various decomposition levels. These features capture the signal's characteristics across different frequency bands and time scales, providing discriminative information for differentiating between noise-only and signal-plus-noise scenarios.

2.4. Empirical Mode Decomposition (EMD) for Feature Extraction

EMD constitutes a quintessential data-driven, adaptive signal processing methodology designed to decompose a convoluted signal to finite, often minimal, set of IMFs [12]. Each IMF encapsulates a fundamental oscillatory component intrinsic to the original signal, conforming to two rigorous criteria:

1. Extrema-Zero Crossing Balance: Across the entirety of the dataset, the count of local extrema and zero crossings must be either identical or differ by no more than one, ensuring a symmetric oscillatory nature.
2. Mean Envelope Nullity: At any given locus within the signal, the average of envelope delineated by local-maxima and that by local-minima must be zero, guaranteeing the IMF's intrinsic oscillatory symmetry.

EMD is particularly useful for analysing non-linear and non-stationary signals without requiring a priori knowledge of the signal's characteristics. Features derived from IMFs, such as their energy, instantaneous frequency, or entropy, can effectively characterize the underlying signal components and serve as robust indicators for spectrum sensing.

3. System Model and Proposed Methodology

Our proposed cooperative spectrum sensing system, illustrated in Figure 1 (conceptual), integrates advanced signal decomposition with machine learning.

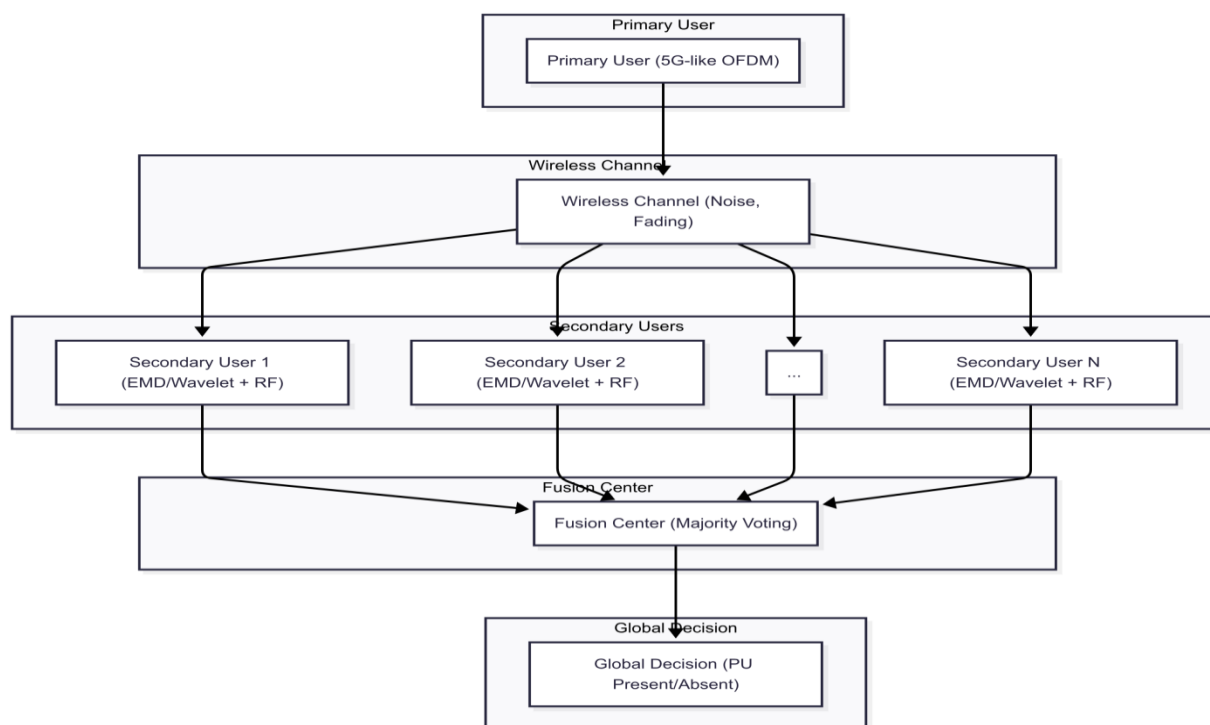


Figure 1: Conceptual Diagram of the Cooperative Spectrum Sensing System

3.1. Primary User Signal

A simplified 5G-like Orthogonal Frequency Division Multiplexing (OFDM) signal is used as primary user's transmitted signal. OFDM is widely adopted in modern wireless standards due to its robustness against multipath fading. The signal generation involves:

- Random bit generation.
- QPSK modulation of bits to symbols.
- IFFT operation to convert frequency-domain symbols to time-domain.
- Adding a cyclic prefix (CP).
- Normalization of the signal power for consistent SNR calculation.

3.2. Secondary User Model

Each secondary user (SU) receives the primary user signal (if present) corrupted by Additive White Gaussian Noise (AWGN). The received signal $y(t)$ can be represented as:

$y(t)=s(t)+n(t)$ under H_1 (signal present)

$y(t)=n(t)$ under H_0 (signal absent)

where $s(t)$ is the PU signal and $n(t)$ is AWGN.

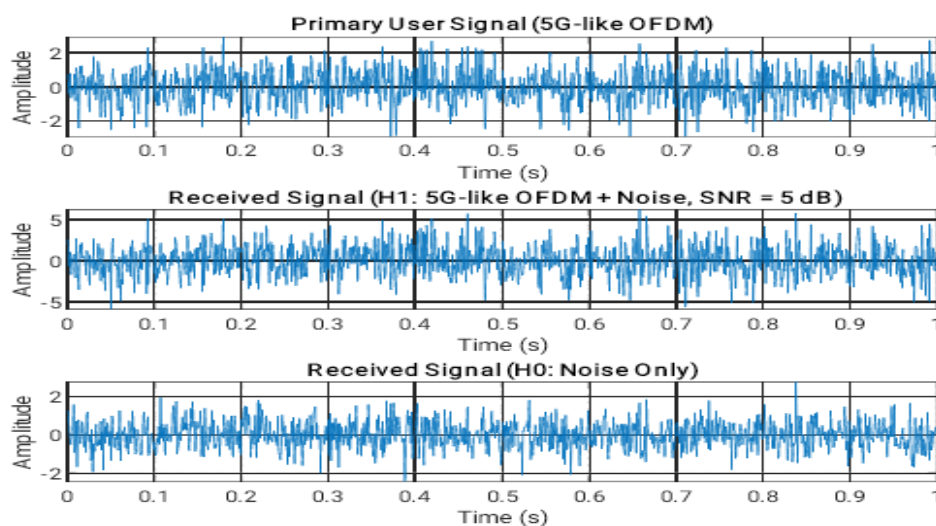


Figure 2: Primary signal and received signal at SNR =5 dB

To visually illustrate the signal characteristics under different hypotheses, Figure 2 presents sample time-domain waveforms. The top subplot displays the generated 5G-like Orthogonal Frequency Division Multiplexing (OFDM) signal, representing the Primary User (PU) transmission. Its noise-like appearance in the time domain is characteristic of OFDM signals. The middle subplot shows a received signal under the H1 hypothesis (PU present), where the PU signal is corrupted by Additive White Gaussian Noise (AWGN) at an SNR of 5 dB. The increased amplitude fluctuations compared to the noise-only case are evident. Conversely, the bottom subplot illustrates a received signal under the H0 hypothesis (PU absent), consisting solely of AWGN. These distinct signal characteristics form the basis for feature extraction and subsequent classification by the machine learning model, enabling the differentiation between occupied and unoccupied spectrum.

3.3. Feature Extraction

The core of the sensing mechanism relies on extracting discriminative features from the received signal.

3.3.1. Wavelet Transform (WT) Features

For WT-based sensing, the received signal is decomposed using a db4 Daubechies wavelet at level 4. The features extracted are:

- **Energy of Approximation Coefficients** : Sum of squared absolute values of the approximation coefficients at the highest decomposition level.
- **Total Energy of Detail Coefficients**: Sum of squared absolute values of all detail coefficients across all decomposition levels.
- **Shannon Entropy of Approximation Coefficients** : Measures the information content/randomness of the approximation coefficients.
- **Total Shannon Entropy of Detail Coefficients** : Measures the information content/randomness of all detail coefficients.

These four features form the input vector to the Random Forest classifier.

3.3.2. Empirical Mode Decomposition (EMD) Features

For EMD-based sensing, the received signal is decomposed into Intrinsic Mode Functions (IMFs). The features extracted are:

- **Total Energy of All IMFs**: Sum of squared absolute values of all elements across all generated IMFs.
- **Total Shannon Entropy of All IMFs**: Sum of Shannon entropies calculated for each IMF.
- **Energy of the First IMF**: Sum of squared absolute values of the first IMF. The first IMF typically captures the highest frequency components of the signal.
- **Shannon Entropy of the First IMF**: Shannon entropy of the first IMF.

These four features also form the input vector to the Random Forest classifier, maintaining a consistent feature vector size for fair comparison.

3.4. Machine Learning Classification (Random Forest)

A Random Forest classifier is trained using a dataset comprising feature vectors extracted under both H0 and H1 hypotheses. The dataset is split into training and testing sets. The trained RF model then makes a local decision (0 for H0, 1 for H1) for each SU based on its extracted features.

3.5. Cooperative Sensing

A fusion center collects the local decisions from all participating SUs. A **Majority Voting** rule is employed for the global decision: if more than half of the SUs detect the presence of a PU (i.e., vote for H1), the fusion center declares the PU as present; otherwise, it declares the PU as absent.

4. Simulation Setup

The simulations are conducted using MATLAB. The key parameters are summarized in Table 1.

Parameter	Value
Sampling Frequency (F _s)	1000 Hz
Signal Duration (T)	1 second
Number of Samples (N _{samples})	1000
SNR Range	-15 dB to 5 dB (2 dB step)

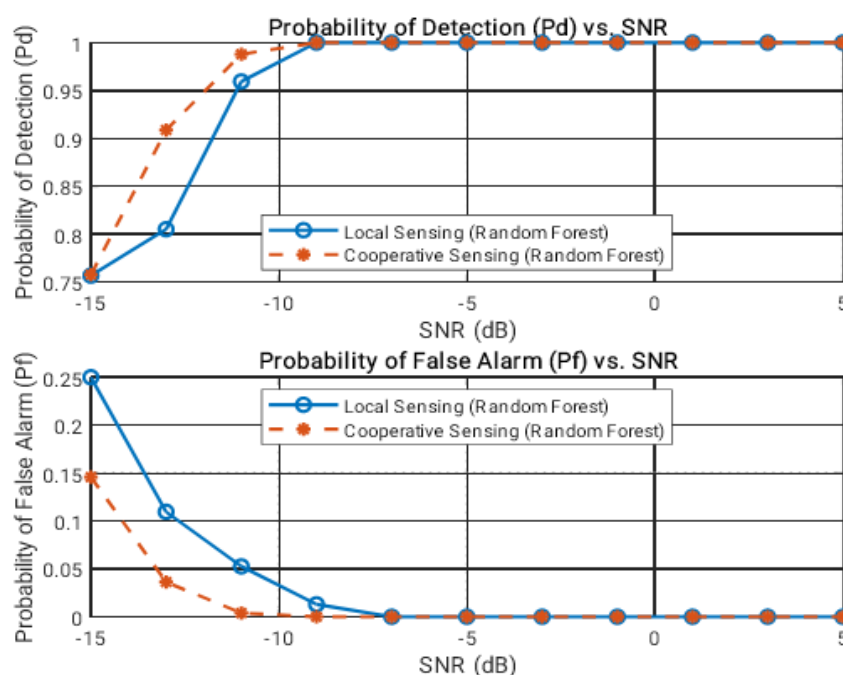
Number of Sensing Iterations	500 per SNR point
Number of Secondary Users	3
Wavelet Type (for WT)	Daubechies 'db4'
Wavelet Decomposition Level	4
OFDM Subcarriers	64
OFDM Cyclic Prefix Length	16
OFDM Modulation Order	4 (QPSK)
ML Classifier	Random Forest
Random Forest Trees	100
Cooperative Fusion Rule	Majority Voting

5. RESULTS AND DISCUSSION

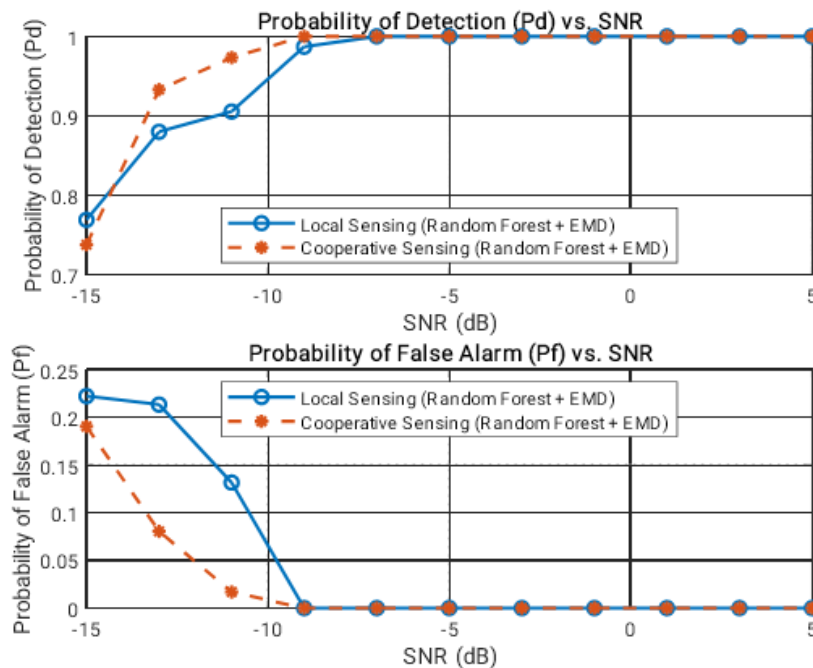
The efficacy of spectrum sensing methodologies is appraised through two pivotal metrics: Probability of Detection (P_d) and Probability of False Alarm (P_f).

- Probability of Detection (P_d): This metric quantifies the likelihood of accurately discerning the presence of a primary user (PU), thereby reflecting the system's sensitivity and reliability in identifying actual signal occupancy.
- Probability of False Alarm (P_f): Conversely, this parameter measures the propensity of erroneously signaling the presence of a PU when, in fact, the spectrum remains unoccupied, thus indicating the propensity for spurious detections.

Together, these metrics provide a comprehensive assessment of the spectrum sensing technique's precision and robustness in dynamic wireless environments. The simulation results, presented in Figures 2 and 3, illustrate the P_d and P_f curves versus SNR for both local and cooperative sensing, comparing the Wavelet-based and EMD-based feature extraction methods.



Figures 4: Probability of Detection (P_d) and Probability of False Alarm (P_f) versus SNR for Wavelet + Random Forest, for both local and cooperative sensing.



Figures 4: Pd vs SNR and Pf vs SNR for EMD+RF, for both local and cooperative sensing

Figures 3 and 4 present the Probability of Detection (Pd) and Probability of False Alarm (Pf) versus SNR for local and cooperative spectrum sensing, comparing Wavelet Transform (WT) and Empirical Mode Decomposition (EMD) feature extraction methods, both with Random Forest (RF) classification.

Key Observations:

- **SNR Impact:** As expected, Pd consistently increases with SNR, reaching near-perfect detection ($P_d \approx 1$) above -5 dB. Conversely, Pf decreases with increasing SNR, indicating improved noise-signal discrimination.
- **Cooperative Sensing Advantage:** Cooperative sensing consistently outperforms local sensing across the entire SNR range. It achieves significantly higher Pd (e.g., above 0.9 at -10 dB where local sensing is lower) and notably reduces Pf, especially at low to moderate SNRs. This highlights the robustness gained from majority voting.
- **Feature Method Comparison:** Both Wavelet+RF (Figure 3) and EMD+RF (Figure 4) demonstrate strong performance. However, EMD-based feature extraction (Figure 4) shows a slight advantage in achieving higher Pd and comparable or slightly lower Pf at very low SNRs (-15 dB to -10 dB). This suggests EMD's adaptive nature may be more effective for complex 5G-like OFDM signals in highly noisy conditions. Both methods converge to similar high performance at higher SNRs.

In conclusion, cooperative spectrum sensing significantly enhances performance. Both WT and EMD, coupled with Random Forest, provide robust sensing, with EMD showing a marginal edge in challenging low-SNR environments."

6. CONCLUSION AND FUTURE WORK

This paper presented a comparative study of cooperative spectrum sensing utilizing the Random Forest classifier with features extracted using either Wavelet Transform or Empirical Mode Decomposition. A simplified 5G-like OFDM signal was used as the primary user signal. The simulation results demonstrate that both Wavelet-based and EMD-based feature extraction methods, when combined with the Random Forest classifier, achieve high detection probabilities while maintaining low false alarm rates across a range of SNRs. Furthermore, cooperative sensing via majority voting consistently improved the overall sensing performance compared to individual local sensing. The choice between Wavelet and EMD may depend on specific signal characteristics and computational constraints, but both offer viable and effective approaches for robust spectrum sensing. Future work could explore:

- Investigating other advanced ML/DL techniques such as Convolutional Neural Networks (CNNs) directly on raw signal data or spectrograms.
- Evaluating the performance with different types of primary user signals (e.g., LTE, Wi-Fi, other 5G waveforms) and realistic channel models (e.g., fading channels).
- Optimizing the parameters of the Random Forest classifier and the feature extraction methods.

- Exploring other fusion rules for cooperative sensing (e.g., weighted fusion, soft decision fusion).
- Analyzing the computational complexity and energy efficiency of these methods for practical CR implementations.

REFERENCES

- [1] Mitola, J. (2000). Cognitive radio: An integrated agent architecture for software defined radio. Doctoral dissertation, KTH Royal Institute of Technology.
- [2] FCC: ET Docket No 03-222 Notice of Proposed Rule Making and Order. Federal Communications Commission (FCC); 2003.
- [3] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," in *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 116-130, First Quarter 2009, doi: 10.1109/SURV.2009.090109.
- [4] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb 2005.
- [5] S. Schell and W. Gardner, "Detection of the number of cyclostationary signals in unknown interference and noise," in *Proc. ACSSC*, Pacific Grove, CA, USA, Nov. 5-7, 1990.
- [6] A. Martian, B. T. Sandu, O. Fratu, I. Marghescu and R. Craciunescu, "Spectrum sensing based on spectral correlation for cognitive radio systems," 2014 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems (VITAE), Aalborg, Denmark, 2014, pp. 1-4, doi: 10.1109/VITAE.2014.6934448.
- [7] X. Zhang, R. Chai and F. Gao, "Matched filter based spectrum sensing and power level detection for cognitive radio network," 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Atlanta, GA, USA, 2014, pp. 1267-1270, doi: 10.1109/GlobalSIP.2014.7032326.
- [8] Daniela Mercedes Martínez Plata, Ángel Gabriel Andrade Reátiga, Evaluation of energy detection for spectrum sensing based on the dynamic selection of detection-threshold, *Procedia Engineering*, Volume 35, 2012, Pages 135-143, ISSN 1877-7058, <https://doi.org/10.1016/j.proeng.2012.04.174>.
- [9] kyildiz, I. F., Lee, W. Y., Vuran, M. C., & Mohanty, S. (2006). A survey on spectrum sensing in cognitive radio networks. *IEEE Communications Magazine*, 44(4), 32-40.
- [10] Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32.
- [11] Daubechies, I. (1992). Ten lectures on wavelets. SIAM.
- [12] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., ... & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903-995.