COOPERATIVE SPECTRUM SENSING USING MACHINE LEARNING-BASED MODELS

A THESIS SUBMITTED TO
THE FACULTY OF ARCHITECTURE AND ENGINEERING
OF
EPOKA UNIVERSITY

BY

NEVISA PATOSHI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRONICS AND COMMUNICATION ENGINEERING

JULY, 2020

Approval sheet of the Thesis

This is to certify that we have read this thesis entitled **"Cooperative Spectrum Sensing Using Machine Learning-Based Models"** and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Ali Osman Topal
Head of Department
Date: July 24,2020

Examining Committee Members:

Assoc. Prof. Dr. Dimitros A. Karras	(Computer Engineering)
Dr. Maaruf Ali	(Computer Engineering)
Dr. Julian Hoxha	(Computer Engineering)

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Nevisa Patoshi

ABSTRACT

COOPERATIVE SPECTRUM SENSING USING MACHINE LEARNING-BASED MODELS

Patoshi, Nevisa
M.Sc., Department of Computer Engineering
Supervisor: Dr. Julian Hoxha

The unstoppable evolution that has affected mobile telecommunication systems in the last three decades has caused the occupation of the licensed frequencies, but at the same time these frequencies are not being used efficiently. Cognitive Radio is the key technology introduced to overcome the main problems of the spectrum utilization, since it offers the opportunity for other unlicensed users to utilize the licensed band while it is not being used by primary user. Even though it increases the efficiency of spectrum utilization, spectrum sensing in cognitive radios still faces problems for higher-performance and more energy-efficient systems. In this work, are taken in consideration two machine learning algorithms as decision-making tools in the fusion centre of cooperative spectrum sensing network based on energy detection technique. The effectivity of these algorithms is evaluated using Receiver Operating Characteristics (ROC) curve and Area Under The Curve (AUC) values, considering seperately additive white Gaussian noise and Rayleigh fading channel. Moreover, the training period of each algorithm is analyzed to evaluate the execution cost for each of them.

Keywords: *Cognitive Radio, spectrum sensing, machine learning algorithms, energy detection, additive white Gaussian noise, Rayleigh fading*

ABSTRAKT

DETEKTIMI I SPEKTRIT NE SISTEMET KOGNITIVE ME ANE TE ALGORITMEVE TE BAZUAR NE MACHINE LEARNING

Patoshi, Nevisa

Master Shkencor, Departamenti i Inxhinierisë Kompjuterike

Udhëheqësi: Dr. Julian Hoxha

Evolucioni i pandalshëm që ka prekur sistemet e telekomunikacionit celular në tre dekadat e fundit ka shkaktuar zenien e frekuencave të licencuara, por në të njëjtën kohë këto frekuenca nuk janë duke u përdorur në mënyrë efikase. Rrjeti Kognitiv është teknologjia kryesore e prezantuar për të kapërcyer problemet kryesore të përdorimit të spektrit, pasi ofron mundësinë që përdorues të tjerë të palicencuar të përdorin bandën e licencuar në kohën që nuk është duke u përdorur nga përdoruesi parësor. Edhe pse kjo teknologji rrit ndjeshëm efikasitetin e përdorimit të spektrit, kjo teknologji është gjithnjë në kërkim të metodave për të arritur performancë më të lartë dhe përdorim eficent të energjisë. Në këtë tezë, janë marrë në konsideratë dy algoritme të bazuar në Machine Learning për tu përdorur si algoritme vendimmarrës në rrjetin Kognitiv, i cili bazohet në teknikën e detektimit me anë të energjisë. Efektiviteti i këtyre algoritmeve vlerësohet me anë të grafikut të Karakteristikave Operative të Marrësit dhe vlerave të sipërfaqes nën kurbë, duke marrë parasysh një kanal të ndikuar nga zhurma e bardhë e Gausiane dhe shpërndarja Rayleigh . Për më tepër, kohëzgjatja e trajnimit të të dhënave e secilit algoritëm është analizuar për të vlerësuar koston e ekzekutimit për secilin prej tyre.

Fjalët kyçe: *rrjeti Kognitiv, detektimi i spektrit, machine learning, detektimi me anë të energjisë, zhurma e bardhë Gausiane, shpërndarja Rayleigh*

Dedicated to my beloved parents

ACKNOWLEDGEMENTS

First and foremost, I would like to thank God Almighty for giving me the strength, knowledge and ability to undertake this research study and complete it satisfactorily. Without his blessings, this achievement would not have been possible.

I would like to express my gratitude to my supervisor Dr. Julian Hoxha for his guidance and the useful suggestions through the learning process of this master thesis.

I am also thankful to my thesis progress committee members for the thoughtful comments and recommendations on this dissertation.

My acknowledgement would be incomplete without thanking the biggest source of my strength, my family. I would like to thank them for their great love and their continuous support. I am especially thankful to my husband who has been extremely supportive of me throughout this entire process and has made countless sacrifices to help me get to this point. I will always be grateful for your love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CR	Cognitive Radio
SS	Spectrum Sensing
PU	Primary User
SU	Secondary User
ML	Machine Learning
NB	Naive Bayes
CRN	Cognitive Radio Network
SNR	Signal-to-Noise-Ratio
MRC	Maximum Ratio Combining
SVM	Support Vector Machine
GMM	Gaussian Mixture Model
AWGN	Additive White Gaussian Noise

CHAPTER 1

INTRODUCTION

1.1 Introduction to Cognitive Radios

The electromagnetic spectrum is notoriously a limited resource, in particular with reference to frequencies between 3 kHz and 300 GHz that have always been available for the broadcasting of radio waves in telecommunication systems *(Figure 1)*. In order to limit the interference between different services, a spectral resource allocation system has been used through a licensing mechanism regulated by individual states or by international bodies. The unstoppable evolution that has affected mobile telecommunication systems in the last three decades has caused the occupation of the allocation of licenses. However, some services, such as wireless transmissions for local networks (wireless LAN) or data transmissions on third-generation mobile phone networks (HSPA), have undergone an exponential increase in their use with an overload of the frequencies of interest. (Wang & Liu, 2011)

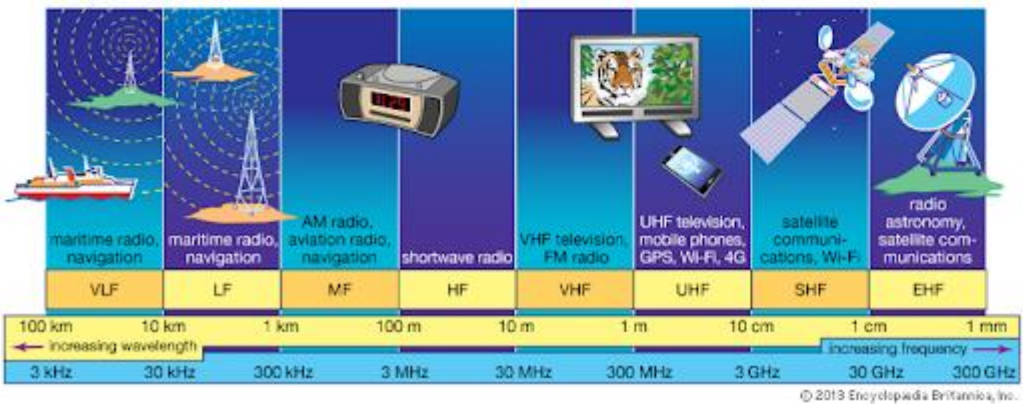

Figure 1. Electromagnetic spectrum used for TLC services

By contrast, other services report little use of the spectral resource assigned to them, according to a latter survey of the Federal Communications Commission (FCC) and by several subsequent studies. In fact, it has been proven through field surveys that the temporal and geographical variance in the use of the assigned spectrum oscillates between 15% and 85% with very low average values and below 20% *(Figure 2)*. It is therefore possible to say that the traditional licensing approach leads to underutilization of the spectrum.

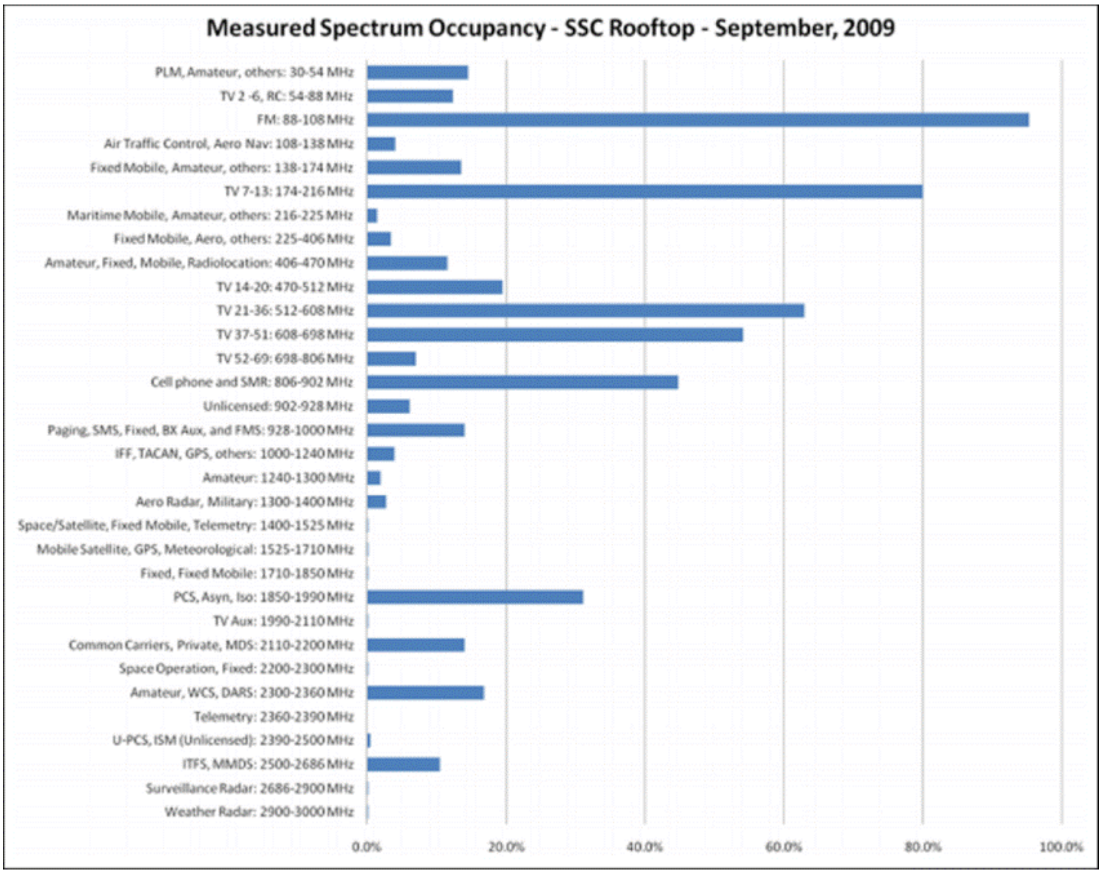

Figure 2. Measured spectrum occupancy in percentage

In order to utilize the spectrum supplies more efficiently, the need to overcome the rigidity of the assignment mechanism through licenses is therefore evident. But, interference between users can happen, in this case. These problems can be avoided through advanced access management of the transmission medium. In the last decade, new technological proposals for devices with a high ability to reconfigure and manage their radio interface have been emerging, such as Software Defined Radio or Cognitive Radio.

These technologies were firstly designed in the late nineties by Dr. J. Mitola III and are characterized by the ability to readjust their operating parameters dynamically thus having all the potential to identify and use the underutilized portions of the spectrum. This adaptive capacity of Cognitive Radio devices opens up new scenarios in which unauthorized secondary users can utilize the available frequencies even if this is bound by a license and already assigned to other services.

Cognitive radio is the essential interface that allows using the spectral range more optimally since it enables secondary users to sense the portion of the available spectrum, choose the suitable channel, share it with other users and, whenever the primary user shows up in the spectrum, free the channel. All of this operation is performed out considering that the transmission of primary users, approved to be using the frequency range, will never be affected.

Although it increases the efficiency of spectrum utilization, spectrum detection in cognitive radios still faces problems for higher-performance and more energy-efficient systems because the spectrum sensing effectiveness can often be directly proportionate to the duration of spectrum detection.

1.2 Aim of the Study

The aim of this study is to develop a better understanding on the Cognitive Radio Networks and investigate other ways to reach the utilization of the spectrum as efficiently as possible using these networks.

Nowadays, Machine Learning is being used in a great range of application, by training the data from the previous experience to develop a model that can predict the future decisions. Machine Learning models can fully adapt to CR networks since the purpose of these ML-based detection methods is to identify the accessibility of frequency ranges by structuring the procedure as a classification model in which the classifier has to select from two alternatives of each frequency band: available or occupied.

Moreover, this process is done without having any former information about the characteristics CR networks, like the average signal-to-noise ratio of each secondary user or the a priori hypothesis probability.

Thus, the objectives of this thesis are to test the efficiency of several unsupervised ML models, as Naive Bayes classifier, Gaussian Mixture Model and Support Vector Machine, to decide which of these model performs betters, in terms of sensing probability and the execution cost.

1.3 Outline of the Thesis

This thesis is divided into five chapters. Each of them is briefly described below:

In the first chapter is shortly given an introduction of the field chosen for this research and the main objectives of the study.

In the second chapter, the literature review related to this research field is given**.**

The third chapter, gives the methodology used to conduct this study. There are also briefly described the system used as a model for this research, the environment used for the implementation and the metrics used to evaluate the performance of each model.

In the fourth chapter are given the numerical results obtained after the implementation of the models in the selected environment.

Conclusions and future work will be given in ***the last chapter*** of this study.

CHAPTER 2

LITERATURE REVIEW

In this chapter is given the background of Cognitive Radio, including characteristics, functions, architecture and sensing techniques. After that, fusion decision-making models are explained, especially the basic combining and ML-based methods.

2.1 Cognitive Radio (CR) Fundamentals

"A Cognitive Radio (CR) is a radio that can change its transmitter parameters based on interaction with the environment in which it operates. This interaction may involve active negotiation or communications with other spectrum users and/or passive sensing and decision making within the radio. The majority of cognitive radios will probably be SDRs, but neither having software nor being field reprogrammable are requirements of a cognitive radio. (Tavares Azolini & Abrao, DECEMBER 2017)" ~ ***Federal Communications Commission*** *(FCC)*

Cognitive radio is the main technology that makes it possible to use the spectrum more efficiently because it allows secondary users to detect the part of the spectrum disposable, pick the appropriate channel, co-transmit with other users and, free the channel when the primary user re-appears in the band. All this procedure is executed taking into consideration that the transmission of primary users, that are authorized to use the spectrum, will not be interrupted.

While broadcasting at certain periods when primary users are not transmitting, secondary users must observe the availability of the frequency all time long and also assure that the temperature interference is below the bound (Clancy, May 2007). By using several sensing techniques, they stay informed if primary users show up in the spectrum.

To better use the supplies, it is significant for cognitive radio networks to establish more efficient spectrum allocation and sharing models. Moreover, in order to eliminate the interference and crash between two users, strict rules must be determined to regulate the spectrum entries and channel administration. A secure hand-off mechanism is needed to relocate the transmission of secondary users in another available frequency with a minimum level of latency if a primary user shows in the spectrum.

2.1.1 Cognitive Radio Characteristics

Cognitive Radio indicates more flexibility and adaptability in Software Defined Radio because their devices can change their operational parameters like modulation, power of the transmission, frequency, etc., according to the information assembled from the radio medium for more effective usage of the spectrum. Thus, the fundamental features of CR devices can be expressed as (Haykin, February 2005):

1. ***Cognitive capability***

   CR devices stay informed about the radio medium inquiry, such as range of frequencies available, type of modulation, used protocols, location, disposable resources, enabled services, etc.

2. ***Reconfigurability***

   According to the detected data, CR devices are able to continuously change their operational parameters to reach higher performance.

2.1.2 Main Functions

A usual cycle of a CR device, shown in *Figure 3*, involves sensing the holes in the spectrum, choosing the most convenient frequency, arrange the spectrum access among other user and free the frequency when the primary user needs to use it.

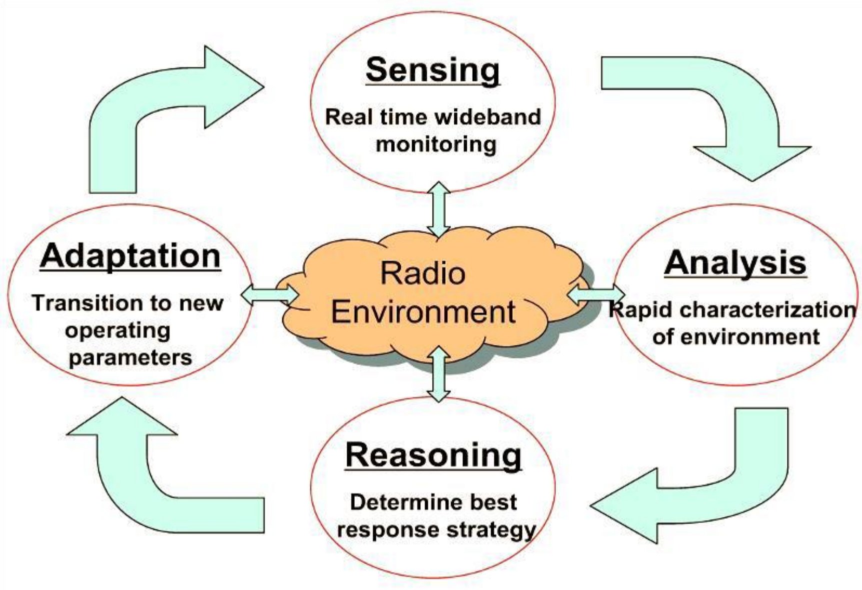

Figure 3. Cognitive Cycle

So the main functions of the cognitive cycle are (Wang & Liu, 2011):

1. ***Spectrum sensing and analysis***

   CR devices are able to sense the spectrum holes, illustrated in *Figure 4*, representing the part of the spectrum unutilized by primary users at e certain time. Moreover, if primary users begin to utilize the spectrum once more, CR devices sense the presence of primary users in order to not interfere in their transmission.

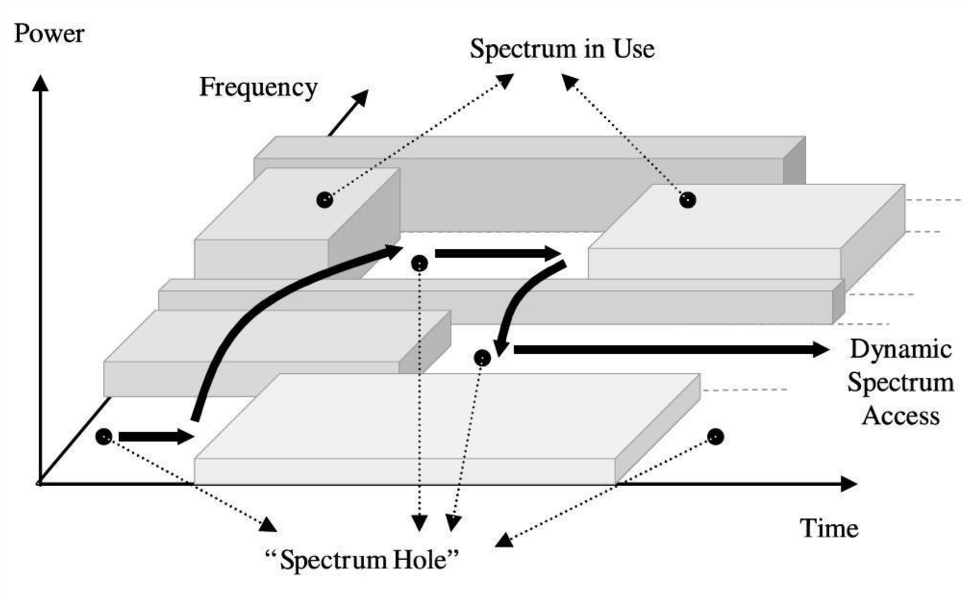

Figure 4. Spectrum hole

2. Spectrum administration and handoff

After sensing the spectrum holes, secondary users decide which frequency to use according to their requirements and they use the spectrum until the primary user shows. In that moment, the handoff mechanism is applied to transfer the transmission of the secondary user in another available frequency, so no interference can happen.

3. Spectrum allocation and sharing

To achieve a better performance and to use the spectrum more efficiently, spectrum allocation and sharing is needed. This means that secondary users are able to divide the available spectrum with primary user and with each other, too. It is important for them to be organized and the interference temperature should be lower than a specified threshold in order to avoid the collisions.

2.1.3 Cognitive Radio Network Architecture

A CR network architecture involves a primary and a secondary network at the same time *(Figure 5)*, since they have to collaborate with each other to use the same band at different moments, without interfering to each other.

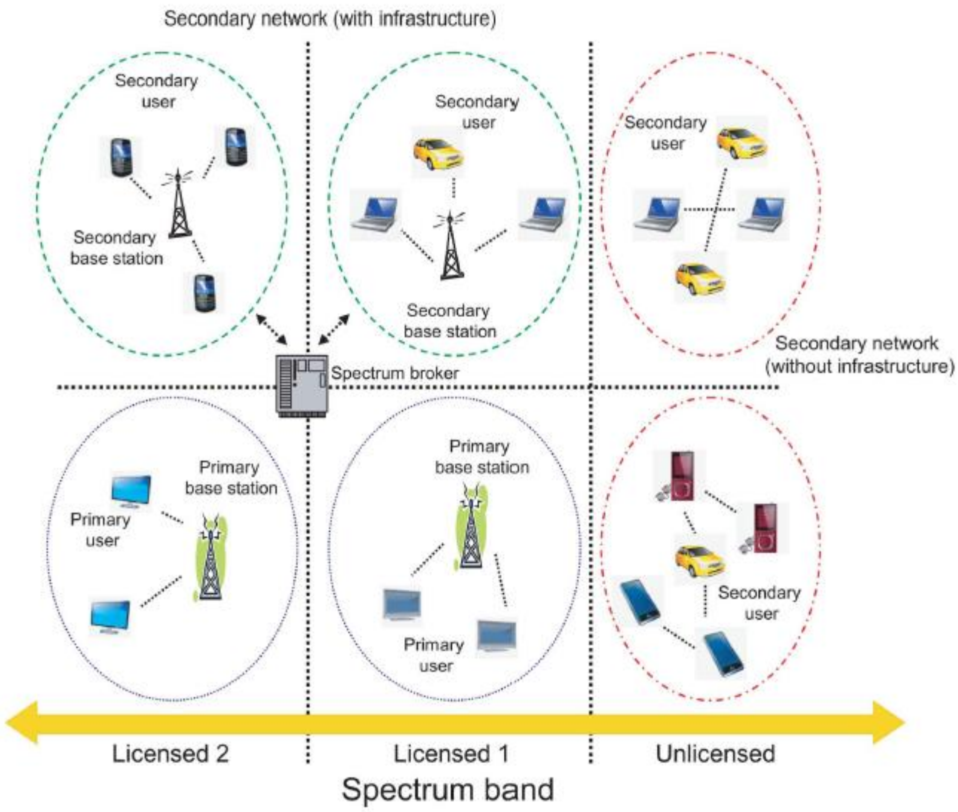

Figure 5. Cognitive Radio Network Architecture

A primary network consists of several primary users managed by one or several primary base stations. Since primary users are authorized to utilize the spectrum, their devices do not contain CR functions.

On the other hand, secondary networks consists of secondary users, usually managed by a secondary base station, that are not licensed to utilize the spectrum, unless it is not being utilized by a primary user. Secondary base stations control the spectrum access traffic of secondary users that use the same range of frequencies and if there is more than a secondary network using the same band, the traffic is controlled by spectrum broker (Raman, Yates, & Mandayam, November 2005)

2.2 Cooperative Spectrum Sensing

Spectrum sensing is the ability of the CR devices to gather the essential data from the medium in order to reach higher performance. It is the first step to have efficient spectrum usage. Cooperative sensing includes the co-operation of the secondary users with each other in order to increase the sensing probability (Kaabouch & Hu, OCTOBER 2014) . It takes advantage of the spatial and multiuser diversity to increase the spectrum sensing reliability and the probability of detection to lower the false alarms and to use the spectrum more efficiently.

Even though it has high efficiency, its performance is constrained by several effects including shadowing, noise uncertainty, etc. Moreover, in case of a high number of secondary users that need to examine a lot of spectrum channels, they have to distribute the detection information to each other. This requires huge amount of data swapping, resulting in a lot of energy consumption that means increased costs of the network.

2.2.1 Spectrum Detection Techniques

Spectrum detection is the first function of a CR device and it can be executed in several domains, such as time, frequency and spatial. In order to change their parameters, it is important for these devices to assemble the necessary information from the surroundings and based on the needed a priori data, spectrum detection techniques are divided in several groups.

- **Energy Detection Sensing**

Energy detection is based on the recognition of the signal by calculating its energy. This technique is the most used because its simple application and also no former information is needed for primary user's signal.

For the received signal, the hypothesis model can be written as:

$$\begin{aligned} \boldsymbol{H_0}&: \boldsymbol{y(k) = n(k)} \\ \boldsymbol{H_1}&: \boldsymbol{y(k) = gx(k) + n(k)} \end{aligned} \quad [1]$$

where H_0 is the null hypothesis denoting that the primary user is not using the spectrum, H_1 denotes that the spectrum is being used by primary user, $n(k)$ is the Gaussian noise, $x(k)$ is the signal of the primary user and g is the channel gain. The detection statistics is calculated as:

$$\boldsymbol{T = \frac{1}{M}\sum_{k=1}^{M}|y(k)|^2} \quad [2]$$

where M is the number of the executed samples.

The detection statistics T is compared to a given threshold λ to determine whether the spectrum is available. The probability of detection P_T and probability of false alarm P_F define the efficiency of the detector.

$$\boldsymbol{P_T = Pr(T > \lambda | H_1)} \quad [3]$$

$$\boldsymbol{P_F = Pr(T > \lambda | H_0)} \quad [4]$$

In order to reach a higher performance, the detector must keep the probability of false alarms in low level. The disadvantages of this technique are the incapacity to differentiate the primary user from other sources, especially in low levels of signal-to-noise ratio, and the difficulty in defining a proper threshold.

- **Cyclostationary Based Sensing**

In this technique, the signals are modulated with sine waves carriers in order to easily differentiate primary user's signal from noise (Cabric & Brodersen, October 2005) (N., Cordeiro, & Challapali, December 2005)

The detection statistics is calculated using the cyclic autocorrelation function (CAF) that verifies when the primary user re-apears in the spectrum:

$$\boldsymbol{R_y^{\alpha}(\tau)} = \boldsymbol{E\left[y(k+\tau)y^{*}(k-\tau)e^{j2\pi\alpha k}\right]} \quad [5]$$

where α represents the cyclic frequency, * is the complex conjugation and $E[\cdot]$ represents the expectation operation.

Taking in consideration sensing in low SNR levels, cyclostationary detector performs better than energy detector. However, this detector requires some prior data from primary user's signal.

➢ **Matched Filtering**

Matched filtering performs better than the other detectors to define the primary user's signal in low SNR levels, but only if it is fully informed about the features of the primary user's signal. Otherwise the sensing performance will be reduced a lot.

The key benefit of this technique is the fact that the number of samples needed for the analysis is inversely related to SNR. So to reach higher performance, a small number of samples is necessary.

On the other hand, matched filtering is very expensive to apply because it requires dedicated receivers for every CR device.

2.2.2 Spectrum Allocation and Sharing Topologies

According to the management of the traffic, cooperative sensing is divided in two categories:

1. Centralized Topology

In centralized architecture, the user traffic is totally managed by a fusion centre. According to the accumulated detection data from each user, this fusion centre localizes the spectrum holes and transmits the processed data to secondary users. Centralized sensing networks are classified in two groups (Akyildiz, Lee, Vuran, & Mohanty, 2006):

a) Partly Cooperative Networks

The users do not exchange information with each other. They sense the channel separately and after that broadcast the data to the fusion centre.

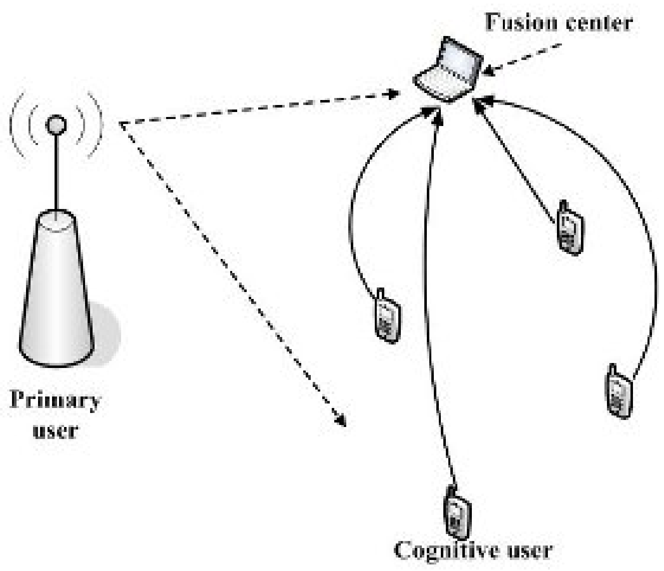

Figure 6. Design of partly cooperative networks

b) Fully Cooperative Networks

The users share with each other the detection data and after that they transmit this information to the fusion centre.

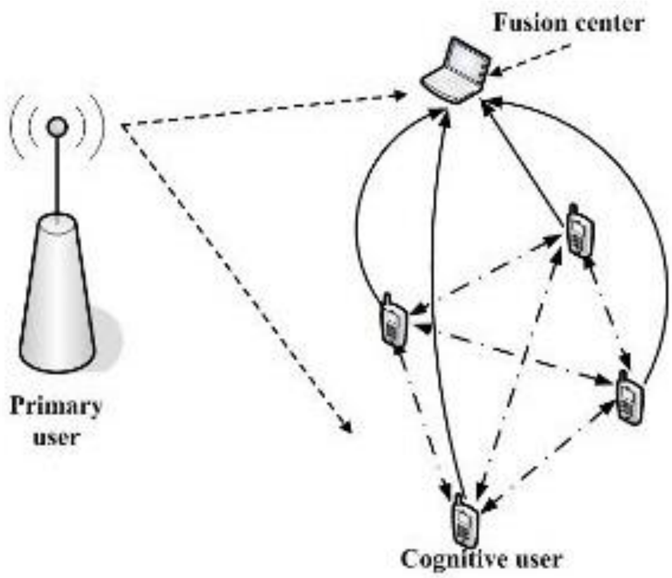

Figure 7. Design of fully cooperative networks

2. Distributed Topology

In distributed architecture, there is not a centre that controls the traffic, but the users communicate with each other and decide based on the gathered information. One of the advantages of the distributed sensing networks is that they do not require a backbone infrastructure.

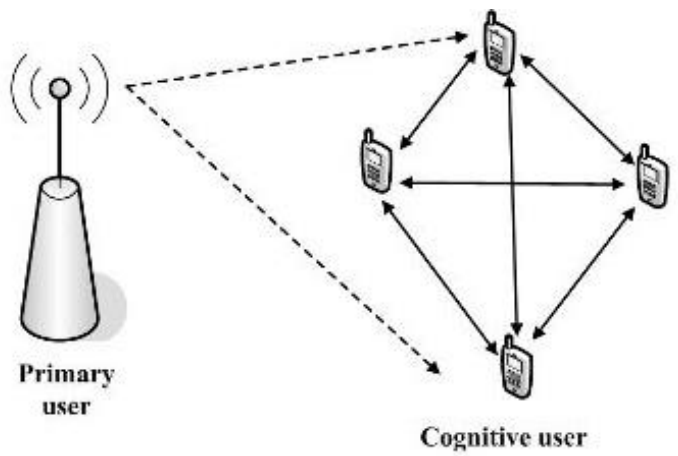

Figure 8. Design of distributed sensing architecture

2.3 Fusion Centre

In centralized sensing, the fusion centre is the main component that manages the traffic between SUs. It processes all the data gathered from every user, comes up with a final decision which of the hypothesis H_1 or H_0 is accurate and communicates it to each user of the network. There exist a lot of decision models applied to fusion center, but the ones used in this research are described in the sections below.

In centralized architecture, the user traffic is totally managed by a fusion centre. According to the accumulated detection data from each user, this centre localizes the spectrum holes and transmits the processed data to secondary users. Every detection node j, detects the energy level, compares it to a given threshold λ and gives as feedback a binary number $Xj = 1$, if it assumes that H_1 is true, and $Xj = 0$ if not. An elementary solution is the Voting Rule that is based over some logic operations, such as:

a) OR Rule

It is a hard-combining method and the detection performance is calulated as:

$$\begin{cases} P_{OR\_T} = 1 - \prod_{j=1}^{N} \left(1 - P_{T_j}\right) \\ P_{OR\_F} = 1 - \prod_{j=1}^{N} \left(1 - P_{F_j}\right) \end{cases} \quad [6]$$

b) AND Rule

Using the AND rule, also a hard-combining method, the detection performance is calculated as:

$$\begin{cases} P_{AND\_T} = \prod_{j=1}^{N} P_{T_j} \\ P_{AND\_F} = \prod_{j=1}^{N} P_{F_j} \end{cases} \quad [7]$$

c) Maximum Ratio Combining (MRC)

In this technique the hypothesis test is given as:

$$\begin{cases} H_1, & if \ \sum_{j=1}^{M} w_j y_j \geq \lambda \\ H_0, & otherwise \end{cases} \quad [8]$$

where M is the number of SUs and w_j is the weighted energy level for the j_{th} SU.

In order to increase the efficiency and to reduce the implementation cost, maching learning based models are being used as decision-making tools in the fusion centre of cooperative spectrum sensing network. These methods are described in the section below.

2.4 Machine Learning Based Models

Machine Learning is the analysis of the computational models that upgrade dynamically via observation and previous experience. It is often referred as artificial intelligence. Machine learning techniques develop a mathematical model based on the available data, defined as "training instances", to make predictions about the future without even being explicitly programmed for doing so. These algorithms are included in a diverse range of applications, like phishing emails and image processing, in which it is hard to build algorithms for the execution of some required dutties.

Numerous networking requirements are fulfilled from the ML aspect, such as a) traffic prediction; b) traffic segmentation; c) traffic forwarding; d) resource planning; etc., while in CRN, the purpose of these ML-based detection methods is to identify the accessibility of frequency ranges by structuring the procedure as a classification model in which the classifier has to select from two alternatives of each frequency band: available or occupied.

The advantage of the ML-based models in CRN is that they do not demand for previous information on channel, including a priori probabiblity of channel availability and SNR of SUs. A vector with channel condition results is utilized by fusion centre to train the ML methods.

The training procedure is a computerized method to closely resemble a function that schedules the expected samples of energy on the SUs to a PU tag. The methods are capable of deciding the channel condition after the training period, depending on unknown energy samples. It enables the structure of the cognitive radio network to take into account the greatest-effort strategy, whether by refusing primary user collaboration or by providing secondary users with channel knowledge.

2.4.1 Unsupervised Learning

In this case, for the training of the classifier are needed only the energy matrices (i.e., y = {y(1),..., y(M)}) and this makes it more quickly adapted in practice in comparison to the supervised learning. Since there is not a specific instructor involved in training, unsupervised learning needs to depend on the clustering framework of the energy vectors.

A cluster of the energy matrices in relation to the corresponding multivariate Gaussian distribution is generated for any PU status. In particular, observations of the energy matrices are obtained from the Gaussian mixture distribution whose pdf is given as:

$$\boldsymbol{f(x) = \sum_{S} p(s) \cdot \varphi((X|\mu_Y), \Sigma_Y)} \qquad [9]$$

where $p(s)$ is the probability of the PU status, μ_Y is the mean vector, Σ_Y is the covariance matrix and $\varphi((X|\mu_Y), \Sigma_Y))$ is the pdf of the Gaussian mixture distribution given as:

$$\boldsymbol{\varphi((X|\mu_Y), \Sigma_Y)) = \frac{1}{2\pi^{(N/2)}|\Sigma_Y|^{1/2}}\, exp\left\{-\frac{1}{2}\,(x-\mu_Y)^T \Sigma_Y^{-1} (x-\mu_Y)\right\}} \qquad [10]$$

After the classifier has been instructed to use clustering, it gets the classification test energy vector and finds out the cluster in which it belongs to.

2.4.2 Supervised Learning

Moreover, in practice the PU has to notify the cognitive network of channel occupancy for some of the energy vectors for training purposes. So comparing to the unsupervised learning, it is more difficult to execute. Yet, supervised learning appears to have higher performance because of the prior knowledge about the quality of the platform.

2.4.3 System Model

The system model used for this research consists of a cognitive radio network (CRN) with a single PU and k SUs, where all secondary users utilize energy detection as sensing technique during a detection period τ along the bandwidth ω. They provide $K = 2\omega\tau$ samples for every detection period, since the considered sampling frequency is at the Nyquist rate, $f_s = 2\omega$. Each SU calculates the level of energy and communicates it to the fusion centre. Depending on energy rates recorded by all SUs, the fusion system decides the state of the channel.

Throughout this work, it is followed a very common PU model, in which PU changes between active and passive status. Let P determine the PU status:

$$\boldsymbol{S} = \begin{cases} \mathbf{1}, & \boldsymbol{if\ PU\ is\ transmitting} \\ \mathbf{0}, & \boldsymbol{otherwise} \end{cases} \quad [11]$$

So according to the status of PU, is determined the occupancy of the channel:

$$\boldsymbol{O} = \begin{cases} \mathbf{1}, & \boldsymbol{if\ \ S = 0\ \ \ (available)} \\ \mathbf{-1}, & \boldsymbol{if\ \ S = 1\ (unavailable)} \end{cases} \quad [12]$$

The sensing process at the ith SU is done according to two hypothesis:

$$\begin{cases} \boldsymbol{H_1}: \boldsymbol{z_j(m) = h_j x(m) + n_j(m), if\ the\ PU\ is\ active} \\ \boldsymbol{H_2}: \boldsymbol{z_j(m) = n_j(m), otherwise} \end{cases} \quad [13]$$

where h_j is the channel gain from the PU's transmitter to the j_{th} SU's receiver, $x(l)$ is the transmitted signal from PU and $n_j(l)$ is AWG noise at the j_{th} SU receiver. The channel gain h_j is calculated as:

$$\boldsymbol{h_j = g_j \sqrt{d_i^{-L}}} \quad [14]$$

where g_j is the fading coefficient, considered uniform in this research; d_j is the Euclidean distance between the j_{th} SU and the PU; and L is the path-loss coefficient calculated as 4. So, in this research is taken in consideration a non-line-of-sight (NLOS) channel.

The normalized energy calculated at the j_{th} SU, after the detection period, is:

$$\boldsymbol{y_j = \frac{1}{\sigma_n^2}\sum_{m=1}^{M} z_j^2(m)} \quad [15]$$

where M is the number of samples considered for every detection period.

According to H_0 hypothesis, $z_j(m) = n_j(m) \sim N(0, \sigma_n^2)$, so, y_j will have a Chi-squared distribution with k degrees of freedom:

$$\boldsymbol{y_j = \sum_{m=1}^{M}\left[\frac{n_j(m)}{\sigma_n}\right]^2 = \sum_{m=1}^{M} \hat{z}_j^2(m)} \quad [16]$$

where $z_j(m) \sim N(0, \sigma_n^2) \;\therefore\; y_j \sim \chi_k^2$

While, according to hypothesis H_1, the normalized energy will have Gamma distribution with shape $\frac{k}{2}$ and scale $2(1 + \gamma i)$:

$$\boldsymbol{y_j = \sum_{m=1}^{M}\left[\frac{h_j x(m) + n_j(m)}{\sigma_n}\right]^2 = \sum_{m=1}^{M} \hat{z}_j^2(m)} \quad [17]$$

where $z_j(m) \sim N(0, 1 + \frac{h_j^2\sigma_s^2}{\sigma_n^2}) \therefore y_j \sim \Gamma\left(\frac{k}{2}, 2\left(1 + \gamma_j\right)\right)$,

where γ_j is the average SNR calculated as:

$$\boldsymbol{\gamma_j = \left(\frac{h_j\sigma_s}{\sigma_n}\right)^2} \quad [18]$$

In the energy detection model, the availability of the spectrum is defined by comparing every SU's estimated energy to a certain threshold λ:

$$A_j = \begin{cases} H_1, & if\ y_j \geq \lambda \\ H_0, & if\ y_j < \lambda \end{cases} \quad [19]$$

So, false alarm and detection probabilities are calculated respectively as follows:

$$\begin{cases} P_F = P(y \geq \lambda\,|H_0) \\ P_D = P(y \geq \lambda|H_1) \end{cases} \quad [20]$$

Consiquently, taking into account the hypothesis H_0, P_F of the j_{th} SU can be written as the right-tail probability of a central Chi-squared random variable:

$$P_{F_j} = \int_{\lambda}^{\infty} f(y_j|H_0)dy \triangleq Q_{\chi_N^2}(\lambda) \quad [21]$$

where $f(y_j|H_i)$ stands for the conditional PDF of the normalized energy calculated at the the j_{th} SU, given hypothesis H_i for i = 0,1.

By fixing a desired probability of false alarm (P_F^*), the threshold parameter λ is derived from [21] as:

$$\lambda = Q_{\chi_N^2}^{-1}(P_F^*) \quad [22]$$

Taking in consideration the incomplete Gamma function, [22] is formulated as:

$$\lambda = 2\Gamma_u^{-1}\left(P_F^*, \frac{k}{2}\right) \quad [23]$$

where $\Gamma_u(x, n)$ represents the incomplete Gamma function, calculated as:

$$\Gamma_u(x, n) = \frac{1}{\Gamma(x)} \int_n^{\infty} t^{x-1}\, e^{-t} dt \quad [24]$$

where $\Gamma(x)$ represents the Gamma function.

In the same way is calculated the probability of detection:

$$P_{D_j} = \int_{\lambda}^{\infty} f(y_j|H_1)dy \triangleq Q_{\Gamma}\left(\lambda; \frac{k}{2}, 2(1+\gamma_j)\right) \quad [25]$$

that can be expressed as:

$$P_{D_j} = \Gamma_u\left(\frac{k}{2}, 2(1+\gamma_j)\right) \quad [26]$$

To conclude, these are the derivation of the false alarm and detection probabilities for the system model used for this research.

2.4.4 Proposed Scheme

The aim of CSS techniques in this research is to properly decide the occupancy of channel O assumed from the energy vector Y. This is similar to developing a classifier to match the energy vector Y appropriately with the channel occupancy O. In terms of machine learning, an energy vector is similar to a feature. In order to build the classification method, firsty enough training energy vectors need to be gathered.

Have y(m) identify the m-th energy vector of training and o(m) identify channel occupancy correlating to y(m). After that, the structure of the energy vectors of training is placed into the training classification, i.e. y={y(1),,..., y(M)}, where M represents the number of training instances. Two cases are known: supervised learning, where the energy vectors of training should be mapped with the correlating channel ocupancy, and unsupervised learning, where this process is not necessary. The classifier is then instructed using the energy vectors. The training process is different for every machine learning method.

After the classifier is instructed, an energy testing vector y* is taken for classification, including its correlating channel occupancy o*. Furthermore, the accessibility of channel $\widehat{o}$ determined by the classifier must be indicated. The energy vector y* mentioned below is classified as either “channel availability” ($\widehat{o}$ = 1) or “channel unavailability” ($\widehat{o} = -1$).

The architecture of the proposed plan is given in *(Figure 9)*. It involves the training unit and the classification unit that work separately. When the cognitive network requires information about channel occupancy, it produces the testing energy matrix. This matrix is used in the classification unit to define the accessibility of the channel utilizing the classifier.

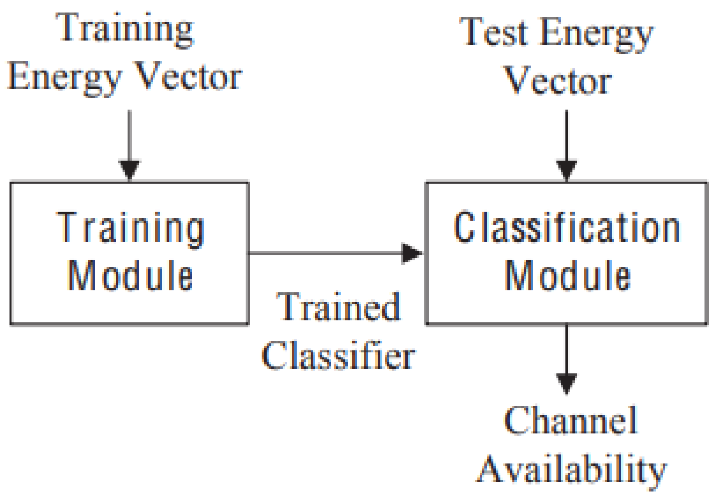

Figure 9. Proposed Scheme

The training unit instructs the classifier utilizing the training energy matrix and generates a trained classifier for the classification unit. This unit can be operated in the first implementation of the cognitive network and when the network of the PU changes. In addition, the training process can be done in the backstory while the classification unit works as usual.

2.4.5 Machine Learning Based Models

Some of the unsupervised machine learning-based models used in CR networks are:

a) Gaussian Mixture Model

The model of the Gaussian mixture is a statistical method built on the prediction of Gaussianity and autonomy among all the energy samples collected from every independent secondary user on the same scheme. These two concepts enable us to identify the mean *m**, variance Σ^2, and the vector of mixing proportions as follows:

$$\boldsymbol{m*} = \begin{pmatrix} \boldsymbol{M \dots M} \\ \boldsymbol{M(1+\overline{\gamma_i}) \dots M(1+\overline{\gamma_N})} \end{pmatrix} \quad [27]$$

$$\boldsymbol{\Sigma^2} = \begin{pmatrix} \boldsymbol{2M \dots 2K} \\ \boldsymbol{2M(1+\overline{\gamma_i})^2 \dots 2M(1+\overline{\gamma_N})^2} \end{pmatrix} \quad [28]$$

$$\boldsymbol{\pi} = [\boldsymbol{P(H_0)\ P(H_1)}]^{\boldsymbol{T}} \quad [29]$$

where *M* refers to the taken number of samples through detection time.

GMM is an unsupervised machine learning technique that can also be defied as a weighted sum of multivariate Gaussian probability densities given by (Choi, Saquib, & Hossain, NOVEMBER 2013):

$$\boldsymbol{f(x|\theta)} = \sum_{\boldsymbol{m=1}}^{\boldsymbol{M}} \boldsymbol{p_m \cdot \varphi((X|\mu_m), \textstyle\sum_m)} \quad [30]$$

where p_m is the probability of the PU status, μ_Y is the mean vector, $\sum_Y$ is the covariance matrix and $\varphi((X|\mu_m), \sum_m))$ is the Gaussian density given as:

$$\varphi((X|\mu_m), \textstyle\sum_m)) = \dfrac{1}{2\pi^{(N/2)}|\sum_m|^{1/2}}\ exp\left\{-\dfrac{1}{2}\ (x-\mu_m)^T {\textstyle\sum}_m^{-1}(x-\mu_m)\right\}$$

[31]

and θ includes every parameter for the GMM such as: p_m, μ_m, and $\sum_m$for all *m = 1,...,M*. All these parameters can be calculated utilizing the maximum-likelihood (ML) estimation for the given group of the evergy matrices (i.e. *y={y(1),,..., y(M)}*) that is given as:

$$\boldsymbol{w(y|\theta)} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{I}} \boldsymbol{ln}\left(\sum_{\boldsymbol{m=1}}^{\boldsymbol{M}} \boldsymbol{p_m} \cdot \boldsymbol{\varphi}\left(\left(\boldsymbol{y^{(i)}}\middle|\boldsymbol{\mu_m}\right), \boldsymbol{\Sigma_m}\right)\right) \quad [32]$$

The ML estimator is the optimizing variable for this log-likelihood function. The variables which optimize the log-likelihood function is achieved utilizing the EM algorithm. This algorithm changes the variable θ maximizing the given function:

$$Q(\theta'|\theta) = E\left\{\sum_{i=1}^{I} ln\left(\sum_{m=1}^{M} p'_{l(i)} \cdot \varphi\left(\left(y^{(i)}\middle|\mu'_{l(i)}\right), \Sigma'_{l(i)}\right)\right) \middle| y, \theta\right\}$$

$$= \sum_{\boldsymbol{i=1}}^{\boldsymbol{I}} \left\{\sum_{\boldsymbol{m=1}}^{\boldsymbol{M}} \boldsymbol{u_m^{(i)}} \cdot \boldsymbol{ln\, p'_m} + \sum_{\boldsymbol{m=1}}^{\boldsymbol{M}} \boldsymbol{u_m^{(i)}} \cdot \boldsymbol{ln\, \varphi}\left(\left(\boldsymbol{y^{(i)}}\middle|\boldsymbol{\mu'_m}\right), \boldsymbol{\Sigma'_m}\right)\right\} \quad [33]$$

When the optimal parameter θ^* is calculated, a testing energy matrice y^* is taken for classification and the classifier determines to which cluster it belongs by comparing it to e given threshold $\boldsymbol{\lambda}$**.** The channel is considered unavailable if:

$$\boldsymbol{ln}\left(\sum_{\boldsymbol{m=2}}^{\boldsymbol{M}} \boldsymbol{p_m^*} \cdot \boldsymbol{\varphi}\left(\left(\boldsymbol{y^*}\middle|\boldsymbol{\mu_m^*}\right), \boldsymbol{\Sigma_m^*}\right)\right) - \boldsymbol{ln}\left(\boldsymbol{p_m^*} \cdot \boldsymbol{\varphi}\left(\left(\boldsymbol{y^*}\middle|\boldsymbol{\mu_1^*}\right), \boldsymbol{\Sigma_1^*}\right)\right) \geq \boldsymbol{\lambda} \quad [34]$$

b) Support Vector Machine

The SVM is a supervised machine-learning approach designed to define a linear hyperplane with a maximum margin between groups by using a kernel function *k(x, x')* to the input matrix in order to maximize its range from input space to feature space. So according to the status of PU, is determined the occupancy of the channel as in [12].

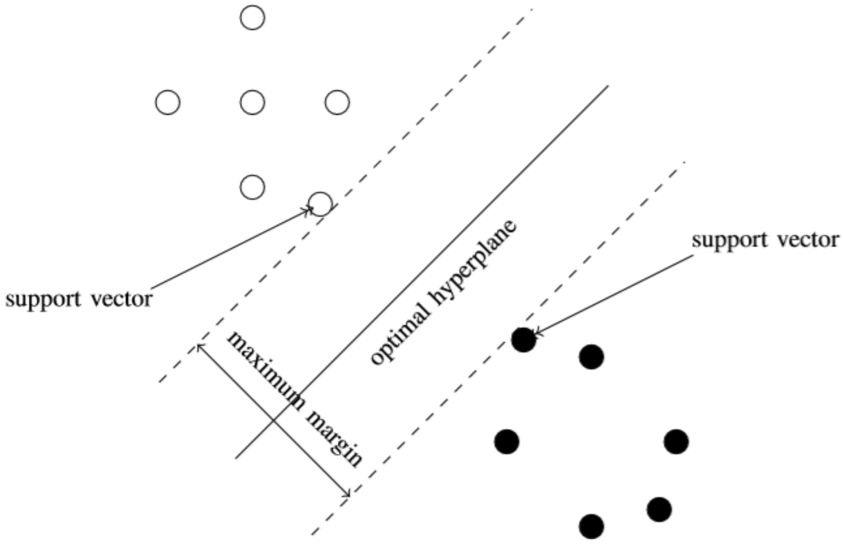

Figure 10. Support vector machine decision plane

For the given linear problem *(Figure 10)*:

$$\boldsymbol{h}(\boldsymbol{y}) = \boldsymbol{w}^{\boldsymbol{T}} \phi(\boldsymbol{y}) + \boldsymbol{b} \quad [35]$$

where $\phi(\cdot)$ is a function for the transformation from feature to space, the aim is to get w and b so that $h(y_m) > 0$ if $O = 1$ and $h(y_m) < 0$ if $O = -1$. Furthermore, to increase the decision margin of the hyperplane h(y) = 0 we can use [35]. Considering that the space of any point y to the decision surface is:

$$\frac{\boldsymbol{O}_{\boldsymbol{m}}(\boldsymbol{w}^{\boldsymbol{T}}\ \phi(\boldsymbol{y}_{\boldsymbol{m}}) + \boldsymbol{b})}{||\boldsymbol{w}||} \quad [36]$$

it is assumed that the nearest point to decision surface is $O_m(w^T\ \phi(y_m) + b) = 1$. So the optimization problem given as:

$$\underset{\boldsymbol{w},\boldsymbol{b}}{\boldsymbol{minimize}} \frac{\boldsymbol{1}}{\boldsymbol{2}}\ ||\boldsymbol{w}||^2 \quad [37]$$

s.t. $O_m(w^T\ \phi(y_m) + b) \geq 1, \quad m = 1, \ldots, M$

assures that each sample is classified correctly and the classes are excellently distuinguished. The positions y_m are referred as *support vectors*. In order to avoid the problem of the clogging, slack variable δ_m and *overlap budget* ξ are added. So [37] is written as:

$$\underset{w,b}{\boldsymbol{minimize}} \frac{\mathbf{1}}{\mathbf{2}} \, ||\boldsymbol{w}||^2 \quad [38]$$

$$\text{s.t. (1)} \;\; O_m(w^T \, \phi(y_m) + b) \geq 1 - \delta_m, \qquad m = 1, \dots, M$$

$$\text{(2)} \;\; \delta_m \geq 0, \qquad m = 1, \dots, M$$

$$\text{(3)} \;\; \sum_{m=1}^{M} \delta_m \leq \xi$$

where ξ is utilized to lower training errors and manage the complexity.

By using the Lagrange primal function:

$$\boldsymbol{\mathcal{L}}(\boldsymbol{w}, \boldsymbol{b}, \boldsymbol{\delta}, \boldsymbol{\alpha}, \boldsymbol{\mu}) = \frac{\mathbf{1}}{\mathbf{2}} \, ||\boldsymbol{w}||^2 + \, \boldsymbol{\xi} \sum_{\boldsymbol{m}=\mathbf{1}}^{\boldsymbol{M}} \boldsymbol{\delta_m} - \sum_{\boldsymbol{m}=\mathbf{1}}^{\boldsymbol{M}} \boldsymbol{\alpha_m}[\boldsymbol{d_m h}(\boldsymbol{y_m}) - \mathbf{1} + \boldsymbol{\delta_m}] - \sum_{\boldsymbol{m}=\mathbf{1}}^{\boldsymbol{M}} \boldsymbol{\mu_m \delta_m} \quad [39]$$

where μ_m and α_m are assumed as Lagrange multipliers.

If the derivatives are set to zero, is taken:

$$\frac{\boldsymbol{\partial}}{\boldsymbol{\partial w}} = \mathbf{0} \rightarrow \boldsymbol{w} = \sum_{\boldsymbol{m}=\mathbf{1}}^{\boldsymbol{M}} \boldsymbol{\alpha_m d_m} \phi(\boldsymbol{y_m}) \qquad [40]$$

$$\frac{\boldsymbol{\partial}}{\boldsymbol{\partial b}} = \mathbf{0} \rightarrow \sum_{\boldsymbol{m}=\mathbf{1}}^{\boldsymbol{M}} \boldsymbol{\alpha_m d_m} = \mathbf{0} \quad [41]$$

$$\frac{\boldsymbol{\partial}}{\boldsymbol{\partial \delta_m}} = \mathbf{0} \rightarrow \; \boldsymbol{\alpha_m} = \boldsymbol{\xi} - \boldsymbol{\mu_m} \quad [42]$$

So from the equations above we have set KKT conditions:

$$\boldsymbol{\alpha_m \geq 0} \quad [43]$$

$$\boldsymbol{d_m h(y_m) - 1 + \delta_m \geq 0} \quad [44]$$

$$\boldsymbol{\alpha_m (d_m h(y_m) - 1 + \delta_m) = 0} \quad [45]$$

$$\boldsymbol{\mu_m \geq 0} \quad [46]$$

$$\boldsymbol{\delta_m \geq 0} \quad [47]$$

$$\boldsymbol{\mu_m \delta_m = 0} \quad [48]$$

The dual problem is derived if we substitute [40]–[48] in [39] :

$$\boldsymbol{\widehat{\mathcal{L}}(\alpha) = \sum_{m=1}^{M} \alpha_m - \frac{1}{2} \sum_{m=1}^{M} \sum_{n=1}^{M} \alpha_m \alpha_n \, d_m d_n k(y_m, y_n)} \quad [49]$$

where $k(y_m, y_n) = \phi(y_m)^T \phi(y_n)$ is a kernel function, so it calculates the inner product of the variables obtained in the feature space under the embedding ϕ of two locations in the input space.

So the optimization problem is formulated as:

$$\boldsymbol{\underset{\alpha}{maximize}\, \widehat{\mathcal{L}}(\alpha)} \quad [50]$$

s.t. (c.1) $0 \leq \alpha_m \leq \xi$

(c.2) $\sum_{m=1}^{M} \alpha_m d_n = 0$

and it is calculated utilizing standard quadratic techniques.

Considering α^* as the solving of the dual problem and b^* as the solving of the primal problem the output of SVM will be:

$$\boldsymbol{h}(\boldsymbol{y}) = \sum_{\boldsymbol{m}=\boldsymbol{1}}^{\boldsymbol{M}} \boldsymbol{\alpha_m}^* \boldsymbol{O_m} \boldsymbol{k}(\boldsymbol{y}, \boldsymbol{y_m}) + \boldsymbol{b}^* \quad [51]$$

Channel state can now be easily distinguished if we convert the output of SVM into an calculated a posteriori probability $\widehat{P}(\mathrm{H}_1|y)$:

$$\widehat{\boldsymbol{P}}(\boldsymbol{h(y)}) = \frac{\boldsymbol{1}}{\boldsymbol{1} + \boldsymbol{e}^{(\boldsymbol{Ah}(\boldsymbol{y})+\boldsymbol{B})}} \quad [52]$$

and the paramters A and B can be calculated if the negative log likelihood function (LLF) of the training energy vectors is minimized:

$$\boldsymbol{minimize} - \textstyle\sum_{\boldsymbol{m}} \boldsymbol{t_m} \, \boldsymbol{log}(\boldsymbol{p_m}) + (\boldsymbol{1} - \boldsymbol{t_m}) \, \boldsymbol{log}(\boldsymbol{1} - \boldsymbol{p_m}) \quad [53]$$

$$\boldsymbol{where} \quad \boldsymbol{p_m} = \frac{\boldsymbol{1}}{\boldsymbol{1}+\boldsymbol{e}^{(\boldsymbol{Ah}(\boldsymbol{y_m})+\boldsymbol{B})}} \quad [54] \quad \boldsymbol{and} \quad \boldsymbol{t_m} = \frac{\boldsymbol{d_m}+\boldsymbol{1}}{\boldsymbol{2}} \quad [55]$$

Utilizing SVM results the channel condition can be evaluated as:

$$\widehat{\boldsymbol{S}}_{\boldsymbol{SVM}} = \begin{cases} \boldsymbol{H_1}, & \boldsymbol{if} \ \widehat{\boldsymbol{P}}(\boldsymbol{h(y)}) \geq \boldsymbol{1} - \boldsymbol{P_{fa}}^* \\ \boldsymbol{H_0}, & \boldsymbol{otherwise} \end{cases} \quad [56]$$

c) *Naive Bayes Classifier*

The Naive Bayes Classifier, assuming that the energy levels calculated at every SU are reciprocally independent and the sensed energy is specified, evaluates the a posteriori probability of the channel availability using Bayes Theorem as:

$$\boldsymbol{P}(\boldsymbol{H_1}|\boldsymbol{y}) = \frac{\boldsymbol{f}(\boldsymbol{y}|\boldsymbol{H_1})\boldsymbol{P}(\boldsymbol{H_1})}{\boldsymbol{f}(\boldsymbol{y}|\boldsymbol{H_1})\boldsymbol{P}(\boldsymbol{H_1}) + \boldsymbol{f}(\boldsymbol{y}|\boldsymbol{H_0})\boldsymbol{P}(\boldsymbol{H_0})} \quad [57]$$

where $P(H_1)$ and $P(H_0)$ represent the *a priori* probabilities of every hypothesis, calculated as below:

$$\boldsymbol{P}(\mathbf{H_0}) = \frac{\boldsymbol{M}^{\{\boldsymbol{H_0}\}}}{\boldsymbol{M}} \quad [58]$$

$$\boldsymbol{P}(\mathbf{H_1}) = \frac{\boldsymbol{M}^{\{\boldsymbol{H_1}\}}}{\boldsymbol{M}} \quad [59]$$

where $M^{\{\mathrm{H}_1\}}$, $i = 0,1$ is the number of the ith occurrence of the hypothesis.

To conclude, the Naive Bayes channel availability is evaluated as:

$$\hat{\boldsymbol{S}}_{\mathbf{NB}} = \begin{cases} \mathbf{H_1}, & \boldsymbol{if}\ \boldsymbol{P}(\boldsymbol{H_1}|\boldsymbol{y}) \geq \mathbf{1} - \boldsymbol{P_{fa}}^{*} \\ \mathbf{H_0}, & \boldsymbol{otherwise} \end{cases} \quad [60]$$

CHAPTER 3

METHODOLOGY

In this chapter is given the methodology used to conduct this study. There are also briefly described the environment used for the implementation and the metrics used to evaluate the performance of each model.

3.1 Research Approach

The objective of this work is making a comperative survey of the ML-based algorithms used by fusion centre of cognitive radio networks. So, if various methods need to be examined and compared, then a case study is the most suitable research strategy. A case study research is chosen for this reason, as the most useful research approach to this master thesis.

3.2 Selected Environment

The platform used to test the efficiency of the analytical and machine-learning techniques presented in this research is ***MATLAB***. The algorithms are developed in *MATLAB* and *R language*.

MATLAB *(matrix laboratory)* is software for numerical computing, firstly released at 1984. It uses MATLAB language, but also adapts *object-oriented programming models*. This platform merges *computation*, *visualization* and *programming* into a single environment (Houcque, August 2005). MATLAB is generally utilized for:

- Numerical Computations
- Simulation
- Designing Models

- Data Analyzing
- Algorithm and Application Developments

R is a programming language, applied in statistical platforms and data analysis. It is widely used for development of machine learning models.

3.3 Evaluation Metrics

The efficiency evaluation of each technique used in this research is done based on these parameters:

- ***Receiver Operating Characteristic (ROC) Curve*** **–** It is probability curve, used to evaluate the efficiency of classification methods at fixed threshold values.
- ***Area under The Curve (AuC)*** **–** It represents the coefficient of separability. So it shows the capability of the method to distinguish between classes.
- ***Training time*** **–** It shows the period that the method needs to train a specific dataset.

3.4 Implemented Machine Learning Based Models

Machine Learning-based models tested in this research are described below:

- Naive Bayes Classifier
- Gaussian Mixture Model
- Support Vector Machine with Linear kernel function
- Support Vector Machine with Gaussian kernel function

These algorithms are compared with basic analytical decision-making techniques OR, AND and MRC.

CHAPTER 4

NUMERICAL RESULTS

In this chapter are included the numerical results obtained after the implementation of the models in the selected environment.

4.1 Tested Scenarios

The evaluation of the ML-based methods, described in this thesis, is done using two different scenarios, based on the system model described in chapter four.

The first scenario takes in consideration a CR network, consisting of a single PU and three SUs, located in different distances from the PU position.

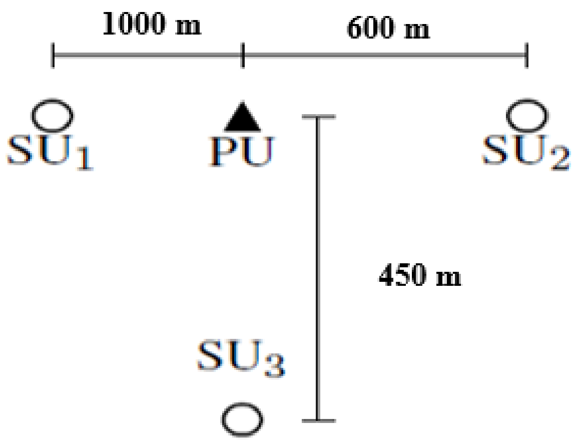

Figure 11. Scenario 1

While, the second scenario takes in consideration a CR network, consisting of a single PU and four SUs, interacting with each other.

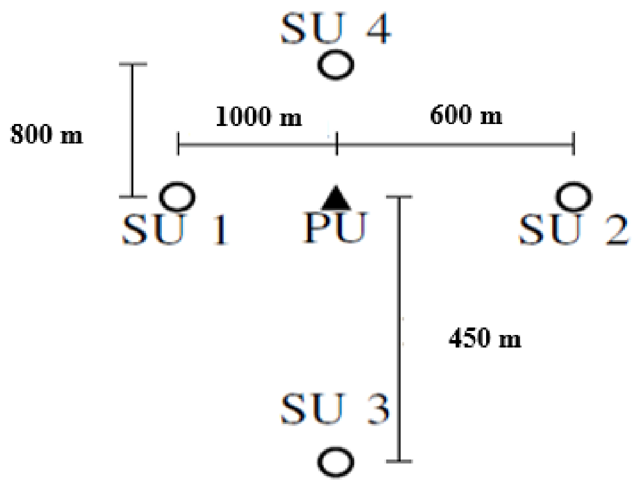

Figure 12. Scenario 2

4.2 System Parameters

To analyse the performance of the unsupervised ML models taken in consideration in this thesis, Monte-Carlo simulations (MSC) with a number of 5×10^4 realization are performed. The given values of the system parameters are concluded in the table below:

Table 1. System Parameters for Monte Carlo Simulations

Bandwidth (w)	5MHz
Noise PSD (η_0)	-150 dBm
PU Transmission Power (σ_s^2)	0.1 mW
Sensing Period (τ)	5µs
Sampling Frequency (f_s)	10 MHz
PU active probability ($P(H_1)$)	0.5
Number of SUs (N)	3, 4
Number of samples (K)	50

Training Set Size (Ω_{train})	1000

4.3 Detection Performance

The detection performance of each method discussed is evaluated through the *receiver operating characteristic (ROC) curve* and the *area under the ROC curve (AUC)* value. The simulations are done considering separately the *AWGN channel* and the *Rayleigh fading channel* for each scenario.

4.3.1 Detection Performance for Scenario 1

As described in the sections above, the first scenario consists of three differently positioned SUs and a single PU. The ROC curve used to define the performance of each method is shown above, taking in consideration AWGN channel and Rayleigh fading channel, respectively.

- ***Performance Considering AWGN Channel***

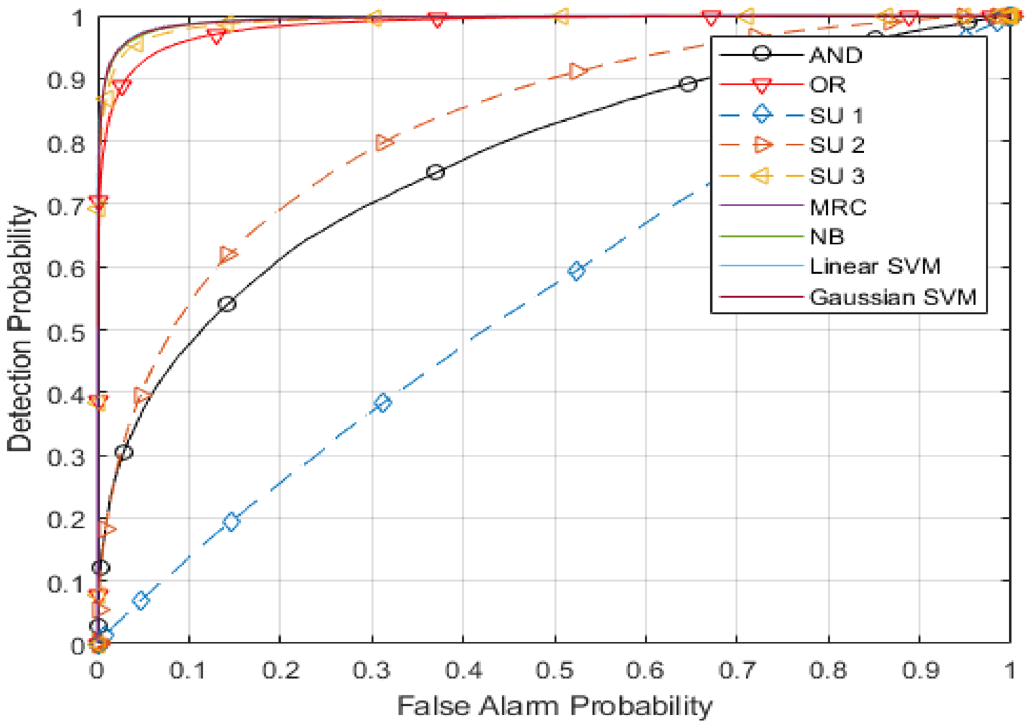

Figure 13. ROC curve for Scenario 1 considering AWGN channel

In *Figure 13* is shown the ROC curve for the first scenario, considering AWGN channel, where it is seen that the MRC technique performs better than other techniques and then followed by GMM and SVM with linear kernel function.

➢ ***Performance Considering Rayleigh Fading Channel***

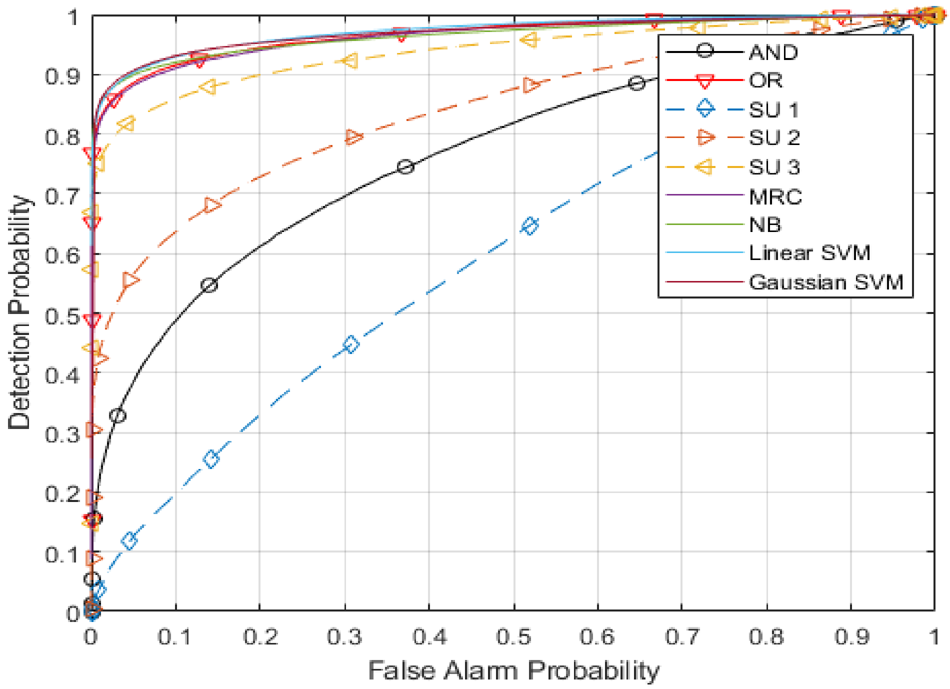

Figure 14. ROC curve for Scenario 1 considering Rayleigh Fading

While in Rayleigh fading, ML-based models, especially GMM and SVM with gaussian kernel function techniques, have higher performance that MRC technique, as shown in *Figure 14*Figure 14, because the MRC technique is based on the mean SNR of every user that in a Rayleigh fading channel, it changes in time.

The results are better noticed by the AUC values concluded in the *Table 2* and it is understood that Naive Bayes Model is less effective than other techniques under both, AWGN and Rayleigh fading channel. On the other hand, the techniques that have higher performance are Gaussian Mixture Model and SVM with linear kernel.

Table 2. AUC Results for Scenario 1

	TECHNIQUES				
CHANNEL	***MRC***	***NB***	***SVM-Linear***	***SVM-Gaussian***	***GMM***
AWGN	0.9941	0.9930	0.9938	0.9934	0.9936
Rayleigh	0.9650	0.9650	0.9728	0.9688	0.9731

4.3.2 Detection Performance for Scenario 2

The same procedure is followed for the second scenario that, unlike the first scenario, has an extra SU in its network.

- ***Performance Considering AWGN Channel***

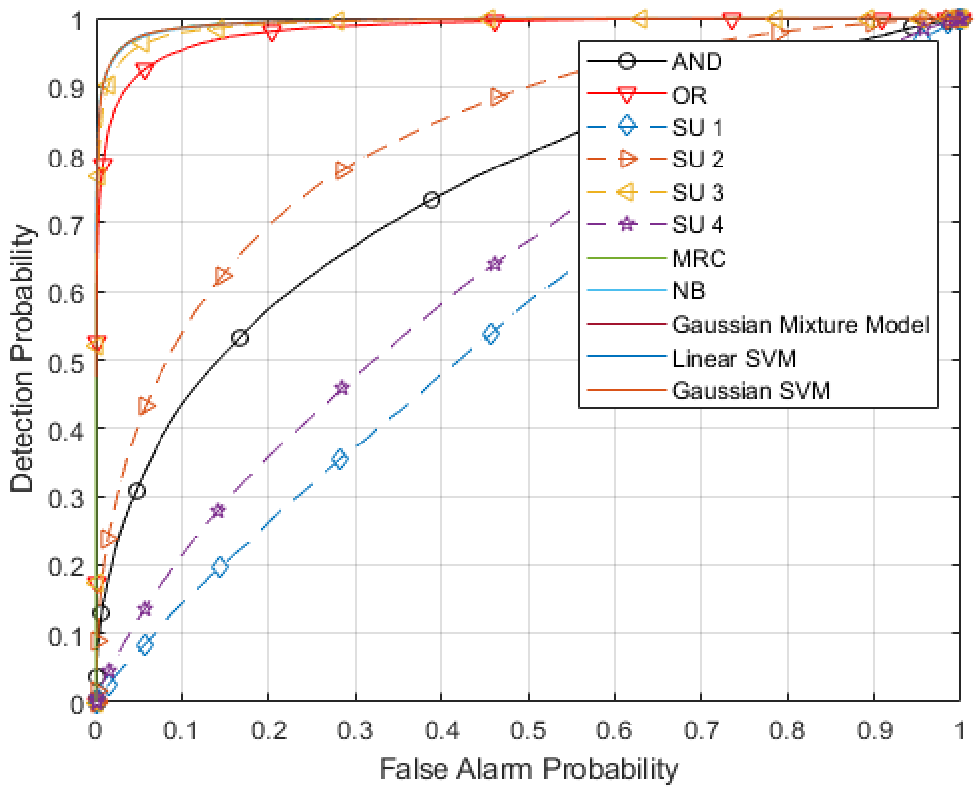

Figure 15. ROC curve for Scenario 2 considering AWGN channel

In *Figure 15* is shown the ROC curve for the second scenario, considering AWGN channel, where it is seen that the MRC technique performs better than other techniques, like in the first scenario. So, the addition of the other users does not affect its performance.

- ***Performance Considering Rayleigh Fading Channel***

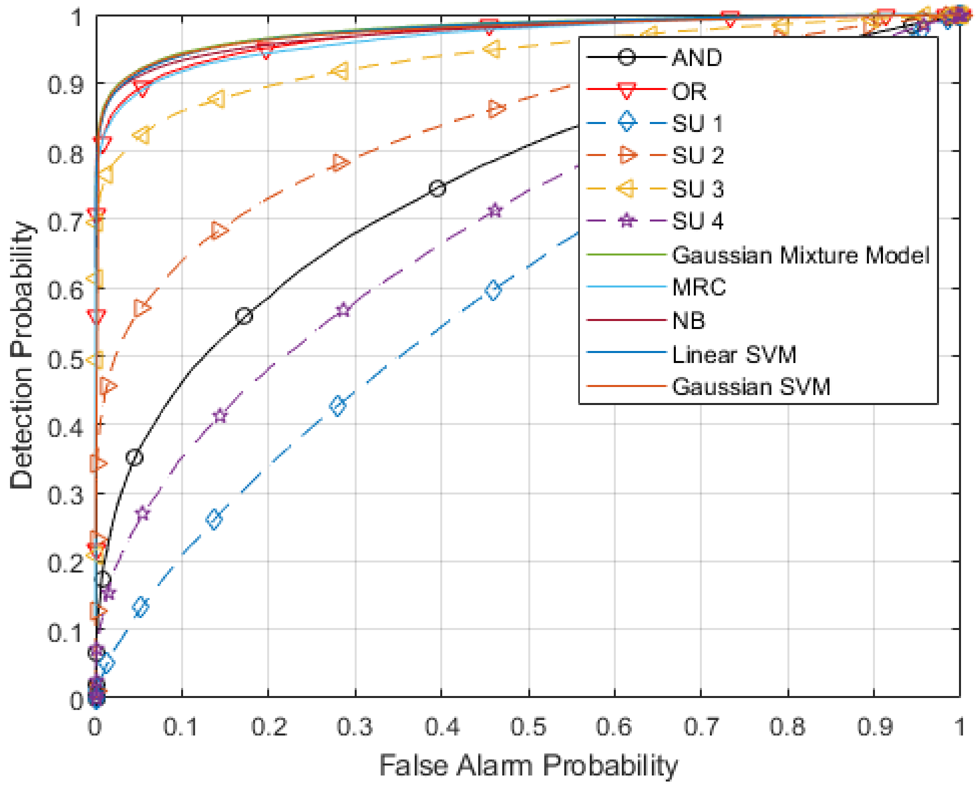

Figure 16. ROC curve for Scenario 2 considering Rayleigh Fading

While in Rayleigh fading, unlike the first scenario, the ML-based models have higher performance that MRC technique, as shown in *Figure 16*Figure 14.

AUC results for the second scenario are concluded in the *Table 3* and it is understood that Naive Bayes Model is less effective than other techniques under both, AWGN and Rayleigh fading channel. On the other hand, the techniques that have higher performance are Gaussian Mixture Model and SVM with linear kernel.

Table 3. AUC Results for Scenario 2

CHANNEL	TECHNIQUES				
	MRC	***NB***	***SVM-Linear***	***SVM-Gaussian***	***GMM***
AWGN	0.9941	0.9927	0.9937	0.9934	0.9934
Rayleigh	0.9663	0.9723	0.9771	0.9747	0.9773

4.4 Execution Cost

For the comparison of the execution cost of all ML-based models analysed in this research, the training period for 1000 training samples is taken in consideration. The results are shown in *Table 4*:

Table 4. Training Results

Techniques	***Training period (s)***
GMM	0.962
NB	0.840
SVM-Linear	2.120
SVM-Gaussian	0.654

From the results is understood that SVM with Gaussian kernel function is more efficient than other ML-based models, regarding training period, since the training period is shorter for the same training dataset.

CHAPTER 5

CONCLUSIONS

5.1 Conclusions

In this thesis, we conducted a comparative analysis of four machine learning algorithms, specifically Gaussian mixture model, naive Bayes and support vector machine with linear and Gaussian kernel function, that were applied as decision-making tools in the fusion centre of cooperative spectrum sensing network based on energy detection technique.

The performance of these algorithms was tested using Receiver Operating Characteristics (ROC) curve and Area Under The Curve (AUC) values, considering seperately additive white Gaussian noise and Rayleigh fading channel. Moreover, two different scenarios were use for evaluation. The first considered scenario was a CR network consisting of a single primary user and three secondary users, located in different distances. While in the second scenario, another secondary user was added to the network.

Numerial results proved that under AWGN channel, the performance of the analytical technique, maximum ratio combining, was higher compared to the performance of the machine learning techniques. On the other hand, under Rayleigh fading channel, it was quite the opposite. All machine learning algorithms demonstrated higher results than MRC because of changing SNR levels in all secondary users for every detection period.

In order to measure the computational cost, the training time of each algorithm was considered and according to the results, the support vector machine with Gaussian kernel function has the lowest execution cost, since the training time was smaller among the other techniques.

REFERENCES

Akyildiz, I. F., Lee, W.-Y., Vuran, M. C., & Mohanty, S. (2006, September 15). NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey. *Computer Networks, 50*(13), pp. 2127-2159.

Benedetto, F. (2012). *Dispense per il corso di SOFTWARE DEFINED RADIO anno academico 2011- 2012*. Rome, Italy.

Cabric, D., & Brodersen, R. (October 2005). Physical layer design issues unique to cognitive radio systems. *Personal, Indoor and Mobile Radio Communications. 2*, pp. 759-763. Berlin: IEEE.

Choi, K. W., Saquib, N., & Hossain, E. (NOVEMBER 2013). Machine Learning Techniques for Cooperative Spectrum Sensing in Cognitive Radio Networks. *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, 31*, 2209-2221.

Clancy, T. (May 2007). Achievable Capacity Under the Interference Temperature Model. *Proc. 26th IEEE Int. Conf. Comput. Commun. (INFOCOM)* (pp. 794–802). Anchorage: IEEE.

D. Cabric & R. W. Brodersen. (2005). *Physical layer design issues unique to cognitive radio systems.* in Proc. IEEE 16th Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC), vol. 2, pp. 759–763.

David Houcque. (August 2005). *INTRODUCTION TO MATLAB FOR ENGINEERING STUDENTS.*

Ethem Alpaydın. (2010). *Introduction to Machine Learning: Second Edition.* The MIT Press.

Ghasemi A. & Sousa E. (2005). *Collaborative Spectrum Sensing for Opportunistic Access in Fading Environments[C]*. IEEE DySPAN.

Haykin, S. (February 2005). Cognitive Radio: Brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications, 23*(2), 201-220.

Henrique, C., & Azolini Tavares & Taufik Abrao. (SEPTEMBER 2019). *Machine Learning Models for Cooperative Spectrum Sensing.* IET COMMUNICATIONS,.

Houcque, D. (August 2005). *Introduction To MATLAB For Engineering Students.* Northwestern University.

Kaabouch, N., & Hu, W.-C. (OCTOBER 2014). *Handbook of Research on Software-Defined and Cognitive Radio Technologies for Dynamic Spectrum Management.* North Dakota: IGI Global.

M. Öner & F. Jondral. (2004). *Air interface recognition for a software radio system exploiting cyclostationarity.* in Proc. 15th IEEE Int. Symp Personal, Indoor, Mobile Radio Commun (PIMRC) vol. 3, pp. 1947–1951.

Mahsa Derakhshani & Tho Le-Ngoc. (2015). *Interference Statistics and Capacity-Outage Analysis in Cognitive Radio Networks".*

N. Shankar C. Cordeiro & K. Challapali. (2005). *Spectrum agile radios: Utilization and sensing architectures.* in Proc. 1st IEEE Int. Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN),.

N., S. S., Cordeiro, C., & Challapali, K. (December 2005). Spectrum agile radios: utilization and sensing architectures. *Int. Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN)* (pp. 160-169). New York: IEEE.

Raman, C., Yates, R., & Mandayam, N. (November 2005). Scheduling variable rate links via a spectrum server. *New Frontiers in Dynamic Spectrum Access Networks (DySPAN).* Baltimore: IEEE.

T. Clancy. (2007). *Achievable capacity under the interference temperature model.* in Proc. 26th IEEE Int. Conf. Comput. Commun. (INFOCOM),.

Tavares Azolini, H. C., & Abrao, T. (DECEMBER 2017). Bayesian Estimators for Cooperative Spectrum Sensing. *2017 IEEE URUCON.*

V., A., & Viswanathan R. (1992). *Asymptotic performance of a distributed detection system in correlated Gaussian noise.* IEEE TRANSACTIONS ON SIGNAL PROCESSING.

Wang, B., & Liu, K. J. (2011). Advances in Cognitive Radio Networks: A Survey. *IEEE Journal Of Selected Topics In Signal Processing*, *5*(1), 5-23.

Wang, J., Wu, Q., & Xueqiang Zheng & Juan Chen. (NOVEMBER 2009). *Cooperative Spectrum Sensing*.