Applications of Machine Learning in Ultra wide band Communication Systems

by

Hossein Soleimani

A thesis submitted
in partial fulfillment of the requrements for the degree of
Doctor of Philosophy in Information and Communication Engineering
at
Sapienza University of Rome

Certified by:
Maria-Gabriella DI BENEDETTO
Scientist
Thesis Supervisor

May 2016
XXVIII

I would like to dedicate this thesis to my loving parents. …

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains including bibliography, tables and equations and has fewer than 50 figures.

Hossein Soleimani
May 2016
XXVIII

Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Professor Maria Gabriella Di Benedetto for the continuous support of my Ph.D study and research, for his patience, motivation, enthusiasm, and immense knowledge. Without her advice, this work would have not have shaped up the way it has. Throughout my stay in Sapienza University, she has been the most supportive and understanding person, and has helped me achieve.
My sincere thanks also goes to Prof, Denardis and my friends leading me working on diverse exciting projects.

Last but not the least, I would like to thank my family for giving birth to me at the first place and supporting me spiritually throughout my life.

Abstract

This thesis describes a novel structure for the design and analysis of communication systems based on machine learning algorithms. We show that various applications of machine learning algorithms can be used in communication systems, especially in Ultra wide system (UWB) and Bluetooth cases. This derives from the properties of UWB and Bluetooth systems and also enables us to discover the new system structure.

Detection, classification, clustering and recognition data could be a solution to the communication system problems. The classification of Ultra Wide Band (UWB) singles vs. signals emitted in the Industrial, Scientific and Medical (ISM) radio bands, such as Bluetooth could be the first possible issue to prevent from interference. The second possible issue is the clustering multipath channel for feedback to transmitter for using, transmitting UWB pulses as a pre-filter.

The first issue is analyzing the behavior of UWB versus Bluetooth signals in various noisy environments. It can be divided to identifying robust feature extraction and classification algorithms, that would enable to classify the UWB and Bluetooth signals. The UWB pulse duration in comparison respect to other technology signals are extremely short time. Hence, the energy of signal in short time discriminate UWB pulses from others. Thus, simulation analysis shows the short-time energy of UWB over small overlapping time windows had acceptable discriminative performance. The short time energy and its derivative feature is tested by Support Vector Machine with related kernel methods, Probabilistic Neural Networks, K Nearest Neighborhood, and Naive Bayes were tested in order to select the best option towards detection performance in different noisy conditions.

In the second issue, we introduce a new digital feedback architecture for Ultra-wide band (UWB) impulsive radio (IR) transceivers when use time reversal prefiltering. We discuss two possible feedback system designs, i.e., with offline or online learning. In the former, a channel impulse response (CIRs) is extracted from the statistical and multi-path channel models for a known environment in the learning process. However, in the latter CIRs are obtained in a learning phase during the launching time of the transceiver. Channel cluster heads (CCHs) are discovered from CIRs clustering. The main focus is an investigation of offline learning in which CCHs and their code are discovered in the learning step and stored in the transceiver system for using in Pre-filter. In the implementation, stored CCHs are compared with the estimated channel to find the most similar CCH in the receiver and then the selected CCH code is sent back to the transmitter. Finally, the time reversed conjugate of decoded CCH is applied to the transmitted pulses as prefilter. Our analysis and simulations show a considerable reduction in bits number to send back.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

During the last decade in the fields of communication systems proposed a different model to transmit the information from transmitter to receiver. However, a number of challenges remain which can be topics of current research. These challenges are targeted to improve the performance of communication systems from various aspects.

Machine learning which is one of the fastest-growing and interesting fields is could be helpful to resolve communication problems. These techniques are used to figure out face and object recognition, speech recognition and optical character recognition (OCR) and etc. It is also a subject of large research because of the great potentials and wide usage in many applications. In spite of that, applying these techniques usually both require to discover the real world applications problems in order to integrate the machine learning algorithms to solve them as well.

1.1 Communication Systems and Challenges

Nowadays, new communication technologies and services are increasing rapidly that lead to several problems. This section offers a focused and concise summary of several challenges facing current communication systems problem.

i) **Signal Classification and Recognition:** In the last decade, wireless applications have been grown in the different technologies which can be transmitted information in licensed or unlicensed allocated spectrum. Licensed spectrum is the portion of the radio spectrum that needs to allocate by the Federal Communications Committee (FCC).

Instead, unlicensed spectrum, called license-free spectrum, is publicly owned. The Industrial, Medical and Scientific (ISM) band is such an unlicensed bandwidth that has been used in a huge number of new technologies like Blue-tooth, Wi-Fi(IEEE 802.11b) and etc. These technologies act in very narrow bandwidth (large scale time-domain) that offers several unique features to discriminate from Ultra-wide band (UWB) signals.

The use of different methods to detect signals operating in the ISM radio bands may prove to be a successful way forward. Since the UWB systems is interfered with the different technologies channel. Thus, the investigation of the UWB signal behaviour, respect to other technology may be helpful to find the solution for interference. On the other hand, identification and classification of the wireless transmitter and receiver devices at the physical layer would also enable many applications including attack detection, authentication, forensic data collection, Radio Frequency (RF) fingerprinting in the physical layer and defect detection monitoring.

ii) **Data Compression**: In the data transmission using fewer bits than the original data leads to requirement less band width, less energy and fast transmission. Data Compression mainly is used for compress data compression such as images, video and speech with different methods like vector quantization. It divides a set of data points (vectors) into clusters having approximately the same number of points closest to them. Each cluster is represented by its cluster head, which is exploited to encode and decode data point. Thus, it is a useful method to make systems faster and more efficient.

For example, in the UWB case, Pre-filtering techniques like Time Reversal (TR) have been proposed to reduce the rake receiver complexity by moving part of it to the transmitter. It needs that receiver periodically estimates the channel impulse response (CIR) and sends it back to the transmitter. The transmitter applies, then the time-reversed CIR to transmit pulses. Thus, it needs to estimate and send back channel state information (CSI). However, this requires a larger bandwidth for feeding back information. Hence, data compression plays a critical role in the feedback systems by taking into account Multi-user and Inter-symbol Interference (ISI).

1.2 Machine Learning

Machine learning is the science of study and construction of algorithms that learn from storage data and make predictions for new data. This subsection describes machine learning algorithm which can be categorized in supervised and unsupervised learning.

(i) **supervised learning**: Supervised learning is the task of inferring a function from labelled training data to achieve a target function. Target function outputs can be discrete which is called a classifier. Hence, the classifier should predict the correct output label for any valid input object. Unlike, if target function outputs become continuous, is called regression. Hence, the target function should predict the correct output value for any valid input object.

(ii) **unsupervised learning**: Unsupervised learning is the task of inferring a function from unlabelled training data to achieve target function. Unsupervised is also called clustering that categorizing data into groups based on inherent similarity or distance.

1.3 Machine Learning and Solutions for Communication Systems Challenges

The classify signals, cluster data and prediction could be a possible application for machine learning methods in communication systems. In the following we will describe some example for machine learning application. Different technologies have their own specific physical signal methods and the behavior of them may serve to discriminate technologies. Machine learning based methods are useful to discover different technology based on particular signals features. It becomes important when there is no explicit relationship between the parameters of wireless systems. Therefore, mathematical relations cannot be effectively applied in order to find the optimal solution. Instead, machine learning algorithms can be applied to computationally hard problems to search the solution.

One of the possible task of unsupervised learning for communication systems is inferring cluster channel that label and number of channel clusters are unknown. In the case of channel clustering, need to the hidden structure of unlabelled channel for using TR pre-filtering. Clustering channel leads to reduce the amount of information because instead of all channel information to send back transmitter, the corresponding channel cluster head label will be feedback

1.4 Contributions

The main contributions of this thesis as explained are simplified into 3 categories: *classification and analysis of signals*, *Digital feedback design for USB systems using TR* and *localization and positioning based on neural network* algorithms.
The first contribution is classification of signals in this thesis which are listed below:

- Selection of signals, in our case UWB, Blue-tooth and white Gaussian noise.

- Extract features based on energy profile and its derivative.
- Selection of features using minimum Redundancy Maximum Relevance (mRMR) and Genetic Algorithm with Information Theory (GA).
- Classifying extracted features using Support vector machine with different kernels, K-nearest neighbourhood (KNN), parabolic neural network (PNN) and Naïve Bayes.
- Comparison of performances.

The second contribution is clustering channel in this thesis which are listed below:

- The introducing idea of digital feedback for TR based on clustering channel.
- Selection of the multipath channel.
- Defining different similarity and distance cretraia to compare similarity and distance the channels.
- Choosing clustering algorithm.
- Decision about a number of channel cluster heads (CCHs).
- Comparison of the proposed algorithm performance with rake receiver with and without TR.
- Comparison of the performance in the condition of without inter-symbol interference, with inter-symbol interference and multi-user interference.

1.5 Thesis Structure

The main subject of this thesis is the exploration of advantages for applying Machine learning algorithm to resolve the communication systems problem. The rest of the thesis is organized as follows;

In chapter 2, we review related work regarding this thesis.

In chapter 3, we review some of the communication systems such as the UWB, Bluetooth and Wi-fi structure. Another hand the main characterize of Blue-tooth and UWB will be discussed in this chapter. Also, short time pulses in IR-UWB causes to the multipath diversity propagation and received the reflections of pulse form different materials such as bricks and cement. Thus, we review the IEEE 802.15.3.a channel model which is used for the realization of channel impulse response (CIR). Furthermore, Multipath diversity pulse reception with and without TR which is accomplished by a RAKE receiver will be explained as well in this chapter. We also review different modulation of UWB case. Finally, we will explain the positioning with UWB case.

In chapter 4, mandatory and useful background topics of machine learning is studied in the first part of the section which is feature extraction and then we study supervise and unsupervised learning. In the feature extraction is investigated the concept of feature extraction which can be used the comparison of similarity criteria of the different algorithm.

In chapter 5, we examine different classification methods to discriminate ultra wideband VS narrowband signals, in the time domain.

In chapter 6, Since multipath in rake receiver leads to the various issues. We address TR pre-filtering for resolve rake receiver problems and then we test TR with the proposed digital feedback versus pour TR and without TR. The performance of the proposed algorithm evaluates in term of bit-error-rates (BER). We investigate the number of bits required for sending back the channel information Section.

We conclude the thesis in chapter 7.

Chapter 2

State-of-the-art related Works

The problems of the physical layer must be addressed to improve the future communication systems. Physical layer issues are connected to the model of the transmitting systems and applications. The goal of this chapter review of the related problems and applications by highlighting their limitations.

2.1 Classification

The pervasive use of Wi-Fi (IEEE 802.11b), Bluetooth (IEEE 802.15.1) and other standard, has exploited considerable amount of the ISM band (unlicensed bandwidth) because of rapid growth of different technologies. Another hand, as a Federal Communication Commission (FCC) report [32], large part of licensed bands are free at a specific time and location. Hence, in the next generation of wireless communication system will need to exploit the licensed bandwidth.

Cognitive Radios (CR) anticipate that future devices able to use licensed frequency with spectrum sensing algorithms. Spectrum sensing means discovering free radio channels. It is divided into two research areas, the first is frequency usage and the second is air interface classification at a specific frequency [38].

The air interfaces classification is proposed for improving the quality of spectrum sensing [38]. Fig.2.1 shows the structure of air interfaces classification. Awareness of signals is a

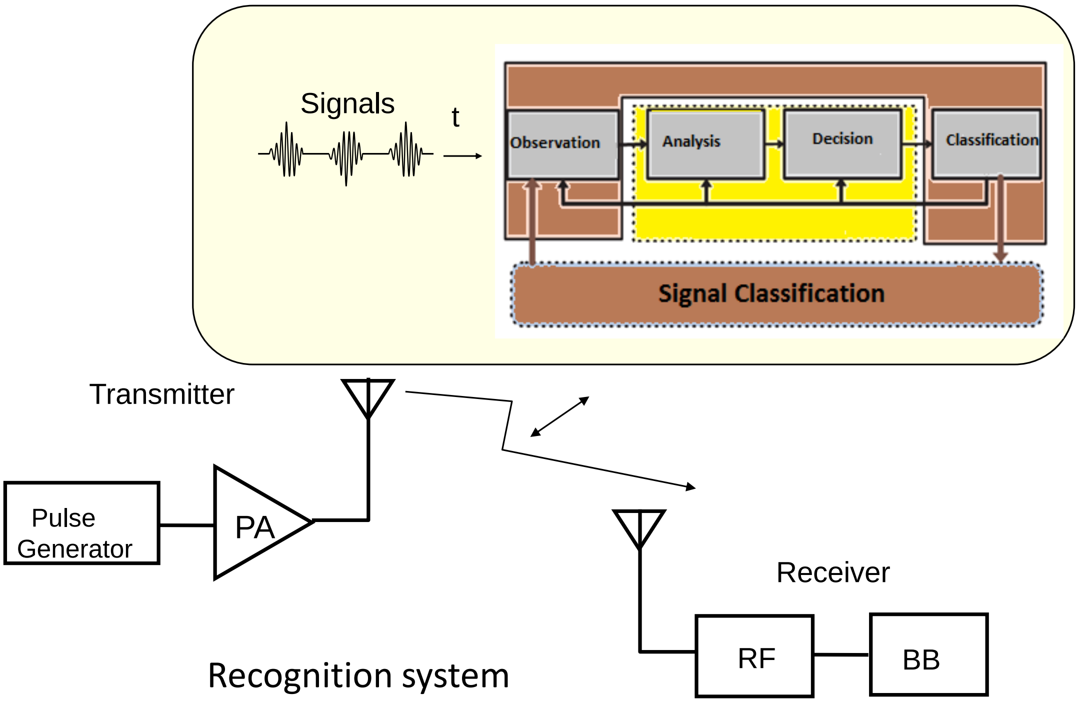


Fig. 2.1 Physical layer awareness

significant part of air interfaces. Radio awareness is investigated in [38, 70, 10, 5] for ISM band technologies such as Bluetooth, Wi-Fi and, Ultra wide band (UWB) systems.

Ultra wide band (UWB) system is an emerging technology to transmit the information to receivers that is targeted to short range and high speed wireless indoor and outdoor system. UWB signals are formed by very narrow time pulses; as such, they correspond to a wide frequency range and overlap with many other signals. Previous work showed that MAC sub-layer features of UWB a peculiar behavior [70, 10, 5], and this can be used to discover UWB in the radio environment. In particular, UWB has specific short time energy profiles.

Identification of a wireless transmitter and receiver devices at the physical layer, would also enable many applications including attack detection [16], authentication, forensic data collection, and defect detection monitoring [17].

Radio Frequency (RF) fingerprinting in physical layer is one answer to find out why signals identifications are so important here as well. RF fingerprinting is the detection of signal (transmitter) features that form a valid device RF fingerprint, based on the relations observed between signals and their transmitters [47]. The wireless network fingerprinting includes a number of valid wireless nodes and a mimic. All the valid nodes are furnished with RF fingerprinting system, which provides two phases. In the first step, the feature RF fingerprint of each valid wireless node is generated by extracting features from the received signal and is stored in the database. In the network operation phase, all valid wireless devices extract the features from the received signal and match it with the profile RF fingerprint stored in the database. The goal of the mimic is to replay the valid wireless nodes' signals in such a way that the physical layer RF fingerprinting scheme is compromised and wireless devices are unable to discriminate between a valid and a malicious device. Each transmitter has its own unique features in the radio waveform. The characteristic of a feature should be such that the variability in the feature should be low for intra transmitter and high for inter transmitters. Using features with this property will provide a consistent and stable identification mechanism and will create a true RF fingerprint of the transmitter. Radio frequency (RF) fingerprinting is a technique which is used to identify wireless devices by extracting a unique identifier from wireless signal transmissions so as to perform device identification or verification[66]. RF fingerprinting attempts to exploit the total effect of differences introduced during transmitter industrialized [72, 75, 81, 83].

Machine learning methods based methods are useful, because some time relationship between the wireless system parameters and the desired performance metrics can not be determined.Thus, mathematical relations cannot be found effectively to find the optimal parameters with respect to the performance metrics. Instead, machine learning algorithms [14, 6, 46] can be applied to computationally hard problems to search the solution and relationships [9].

In the chapter 5, we will analysis UWB and bluetooth signal classification.

2.2 Rake Receiver Problems

In 2002 has been approved by the Federal Communications Commission (FCC) in the United States [32] where a UWB signal either has a fractional bandwidth more than 20 percent or occupies more than 500 MHz bandwidth during transmission period. They allocated 7.5 GHz spectrum band between 3.1 and 10.6 GHz for UWB communications. There are a number of challenges in UWB which can be topics of current research. This technology is formed by very narrow time pulses that offers several unique features to discriminate from other communication technology such as robustness to fading, large channel capacity and strong anti-jamming ability [40].

In UWB short time pulses transit from different paths and receive with multiple replicas at the receiver with specific different time delays and amplitude attenuations. Fig. 2.2 shows the multipath effect of UWB communication systems.

The multipath diversity in Impulse Radio Ultra Wideband (IR-UWB) systems are effective because of extremely short duration of pulses involved in the communication process [78]. It can be modelled with Channel Impulse Response (CIR). The optimum multipath diversity pulse reception in absence of interference is accomplished by rake receivers [27] which is able to recover different paths pulses [79]. Design and implementation of the UWB rake receiver are, therefore, extremely challenging [50].

Prefiltering techniques have been proposed to reduce the rake receiver complexity, by moving part of complexity to the transmitter. For example, the so called Pre-Rake diversity combination technique [31, 48], is able to reduce complexity and cost of the rake receiver.

Time Reversal (TR) prefiltering is other technique to reduce the complexity of rake receiver. It also is helpful to reduce the interference. It is originally applied in acoustics [34], ultrasonic multi user communications [23] underwater acoustics [29], and more recently in wireless communications in general [30] and UWB [64, 63, 33].

A long delay spread [77, 55, 13] lead to high complexity Inter-Symbol Interference (ISI) [12] and Multiuser Interference (MUI) [39, 59], in rake receiver design [11] especially in high data rate scenarios. Time reversal prefiltering can be mitigated interference of users and symbols because it transmits signals via time-reversed version of CIR which is leading to temporal focusing of transmitted signals that, in turn, reduces ISI and MUI, since the received power concentrates within the strongest taps making the task of signal detection easier [71]. Hence, TR simplifies the rake receiver structure and increases the performance of UWB systems because a few channel taps carry most of the energy of the transmitted signals. TR can also positively impact the performance of positioning systems using Direction of Arrival (DOA) [21].

In TR, the receiver periodically estimates the CIR and sends the estimate back to the transmitter, with a period equal to the coherence time, that is the time interval during which channel characteristics can be supposed to be non-varying. The requirement to estimate and send back channel state information (CSI) under the form of quantized versions of CIRs is the main disadvantage of the TR pre-filtering Because it large bandwidth needs to send back channel information.

Higher channel diversity and noise can be increased fading and corruption of feeding back information. The accuracy of CSI also can be increased with the number of quantization levels; however, this requires a larger bandwidth since more information has to be transmitted on the feedback channel [51].

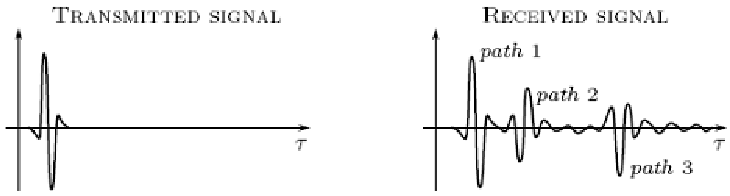


Fig. 2.2 Effect of Multipath

Over the last few years, several methods have been proposed to overcome TR problems by decreasing as much as possible the amount of information that the receiver has to send back to the transmitter. One simple method is to quantize the CIR with coarser steps so that the requirement for sending back large amounts of information to the transmitter is mitigated. When brought to the extreme, this approach leads to the use of only 1 bit for quantization methods. Several works have indeed investigated the use of one bit quantization method applied to the sign of the channel tap coefficients, leading to a pre-filter which can take a value of +1 or -1, [12], [41, 58, 1, 2, 24, 49, 57]. The loss of information in one bit quantization and its effect on the quantization error is investigated in [58, 1, 2, 24, 49, 57, 26, 25].
The main drawback of one bit quantization approach is that it increases the perceived delay spread of the channel [57]. One of the solution is the Phase compensation method [26], which signals are pre-filtered with phase of channel taps. This type of transmission with a pre-filter reduces the Inter-symbol Interference (ISI) since the amplitude of the side lobes of the transmitting signals is decreased with respect to the side lobes of TR method. This in turn reduces the delay spread and multipath dispersion. In [25] the impact of the number of quantization levels for the channel taps phases is evaluated. In parallel in [84], the authors proposed a technique for optimized vector quantization of the CIR in order to reduce information transmission in the feedback channel and hence to decrease the impact of noise.

In the chapter 6, we will explain our solution to feedback data through channel for using TR.

Chapter 3

Communication Technologies

This section summarizes the main characterizes of Bluetooth and UWB. These technologies have their own specific physical signals and the behaviour of them may serve to discriminate technologies. The Bluetooth is working in the unlicensed ISM 2.4 GHz band by the adjusting carrier signal in specific frequency that is used to modulate signals of systems. Carrier signals are wasted the useful energy of transmitting signal. Instead, in UWB, there is no carrier signal as in traditional transmitters. Therefore, from the power consumption aspect, UWB seems to be a better choice.

UWB pulses are very short time. Consequently, they cover the huge bandwidth of frequency which leads to overlap with different technologies signals. The FCC authorized that Power Spectral Density (PSD) of UWB should not exceed emission masks specifying quite low power emission. This low energy allows transmission in bandwidth without licenses. Since low energy signals are similar to noise, their identification becomes difficult [27].

UWB systems have several unique advantages over the conventional narrow band systems like High data rates, Multipath immunity, Fading robustness and High precision for location and ranging applications [40]. We will discuss some of them in the next Section.

3.1 Impulse-radio Ultra-wide band

Impulse radio UWB refers to the IEEE 802.15.4a protocol. On the Impulse radio (IR) UWB systems are used to transmit a sequence of bit streams. In the first step, sequence of bits map to different level, for example in the binary case to 1,-1. In the second step each symbol repeats N_s times and then each repeated symbol formed by the different shape factor. Thus, the single UWB pulse does not contain information of itself. It must add digital information to the analogue pulse, by means of modulation. Symbols may be formed by Pulse position modulation (PPM) or Pulse amplitude modulation (PAM) schemes to transmit information from the Transmitter (Tx) to Receiver (Rx). The transmission rate of the system could be varied by use of PAM and PPM schemes. Although the PPM and PAM constitute the major approaches to modulation in IR-UWB systems, other approaches have been proposed as well [27].

Common multiple access techniques implemented for pulses based on UWB technology are Time Hopping (TH) and Direct Sequence (DS). These techniques are proposed to prevent multi user collisions. A given UWB communication system will be a mixture of these techniques, for example, TH-PPM, TH-PAM or DS-PAM. In TH-mode, the pulse transmission instant is defined by the pseudo-random code [27].

3.1.1 PAM UWB signal modulation

PAM is a modulation technique where the amplitude of the pulse varies to contain digital information. The IR-PAM scheme encodes data sequences as follows [27]:

$$S(t) = \sqrt{E_s/N_s} \sum_{i=0}^{m} \sum_{n=0}^{N_s-1} b(i) w(t - (iN_s + n)T_f) \tag{3.1}$$

where $b(i)$ is the given symbol sequence have different level, for example in the binary case has $(-1,1)$ for the i^{th} transmitted symbol that is repeated with N_s pulses. T_f is the frame interval. $\sqrt{E_s/N_s}$ is the amplitude of a transmitted symbol, where E_s is the signal energy per pulse. $w(t)$ can be different derivative of Gaussian waveform but second derivative of Gaussian waveform has good fit to FCC emission mask. The selected waveform for $w(t)$ is the energy normalized second derivative of a Gaussian pulse [80].

$$w(t) = \Big(1 - 4\pi(\frac{t}{\alpha})^2\Big) \exp\Big(-2\pi(\frac{t}{\alpha})^2\Big) \tag{3.2}$$

where α is the shape factor and can be described by following expression:

$$\alpha = 4\pi\sigma^2 \tag{3.3}$$

where σ^2 is the variance.

3.1.2 PPM UWB signal modulation

PPM technique is introduced to build a simple UWB transmitter. It is used a time-hopping code and binary pulse position modulation, with a single reference pulse shape . This system is perhaps the most common in the literature. This system only requires a single template pulse for reception, and most of the complexity of this system resides in providing accurate timing for the generation of the transmitted sequence and subsequent reception. The PPM modulation scheme is most common method of modulation in the where each pulse is delayed or sent in advance of a regular time encodes binary data sequences using Time Hopping (TH) as follows [27]:

$$S(t) = \sqrt{E_s/N_s} \sum_{i=0}^{m} \sum_{n=0}^{N_s-1} w(t-(iN_s+n)T_f - b(i)\varepsilon) \tag{3.4}$$

3.1.3 Channel Model

In UWB systems pulses carried out extremely short time. Thus, transmitting pulses convey from different paths such as obstacle and wall and other building stuff with different reflection, refraction, and reflection as Fig 3.1. Therefore, the multipath leads to pulses reaching the receiving antenna by several replicas from different paths in UWB systems. It makes dependent a transceiver design to the channel between Tx and Rx, determined by the impulse response.

Several models were proposed in the last years, such as IEEE 802.15.3a [35] and IEEE 802.15.4a [56] for simulating UWB systems. In this part the IEEE 802.15.3a channel model with parameters determined based on a measurement campaign in the 2-8 GHz frequency band was adopted [35]. The model is a modified version of Saleh-Valenzuela's model [65]. The model includes reflection, diffraction and scattering effects of the transmitted signal to the receiver through multiple paths with different delays. Equation 3.5 shows the CIR which

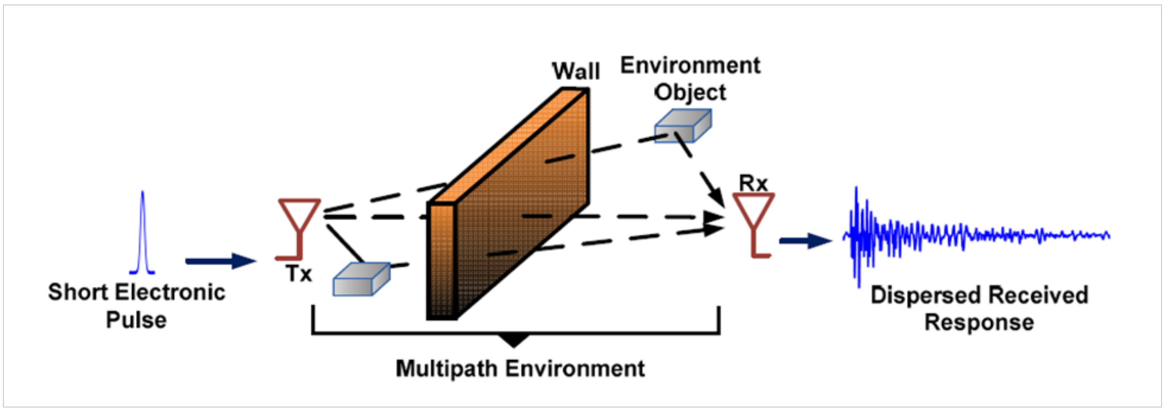


Fig. 3.1 The transmitted pulses are reached to the receiver with multiple replicas.

consists of observed taps of clusters and can be written as:

$$h(t) = X \sum_{l=0}^{L} \sum_{z=0}^{K(l)} \alpha_{z,l} \delta(t - T_l - \tau_{z,l}), \tag{3.5}$$

where $\delta(.)$ is the Dirac delta function, L is the number of clusters, $K(l)$ is the number of rays in the l^{th} cluster, and $\alpha_{z,l}$ are the multipath attenuation coefficients for the z^{th} multipath of l^{th} the cluster. T_l is the arrival time of the l^{th} cluster, $\tau_{z,l}$ is the delay of the z^{th} multipath contribution of the l^{th} cluster, X and is a random variable denoting the amplitude gain due to log-normal shadowing, which is modeled as follows:

$$20\log_{10}(X) \propto Normal(0, \sigma_x^2), \tag{3.6}$$

where σ_x is a log-normal standard deviation of multipath channel realization (dB). Cluster arrival times are generated according to a Poisson process with a probability of cluster arrival time of T_l with respect to previous one given by:

$$p(T_l/T_{l-1}) = \Lambda \exp[-\Lambda(T_l - T_{l-1})], \tag{3.7}$$

where T_l , T_{l-1} and Λ denote respectively the cluster arrival rate time of arrival of the l^{th} , the $(l-1)^{th}$ clusters and cluster arrival rate.

Ray arrival times within cluster are also modeled according to Poison process with ray arrival rate λ, considering a probability of ray arrival times respect to previous one equal to:

$$p(\tau_{z,l}/\tau_{z-1,l}) = \lambda \exp[-\lambda(\tau_{z,l} - \tau_{z-1,l})], \tag{3.8}$$

where $\tau_{z,l}$ and $\tau_{z-1,l}$ denote the time of the z^{th} and the $(z-1)^{th}$ contribution within the cluster l with first starting cluster time is set to $\tau_{0,l} = 0$. The channel coefficients are defined as follows:

$$\alpha_{z,l} = p_{z,l}\xi_l\beta_{z,l} \tag{3.9}$$

$$20\log_{10}(\xi_l\beta_{z,l}) \propto Normal(\mu_{z,l}, \sigma_1^2 + \sigma_2^2)\, or\, |\xi_l\beta_{z,l}| = 10^{(\mu_{z,l}+n_1+n_2)/20} \tag{3.10}$$

where n_1 and n_2 are independent Gaussian random variables with mean $\mu_{n_1} = 0$ and $\mu_{n_2} = 0$ for both case and standard deviation with variance of $\sigma_{n_1}^2$ and $\sigma_{n_2}^2$ that are correspond to the fading on each cluster and ray, respectively. The expected value of $|\xi_l\beta_{z,l}|^2$ can be obtained as:

$$E[|\xi_l\beta_{z,l}|^2] = \beta_0 \exp^{-T_l/\Gamma} \exp^{-\tau_{z,l}/\gamma} \tag{3.11}$$

where β_0 is the mean energy of the first path in the first cluster, and $p_{z,l}$ is an equal probability discrete random variable with values equal to +/-1. ξ_l is the fading related with the l^{th} cluster. $\beta_{z,l}$ is the fading related with the z^{th} ray of the l^{th} cluster. The $\mu_{z,l}$ coefficient is given by:

$$\mu_{z,l} = \frac{10\ln(\beta_0) - 10T_l/\Gamma - 10\tau_{z,l}/\gamma}{\ln(10)} - \frac{(\sigma_{n_1}^2 + \sigma_{n_2}^2)\ln(10)}{20} \tag{3.12}$$

where Γ is a pulse decay factor power within a cluster ray; σ_1 is cluster lognormal fading standard deviation expressed in dB and σ_2 is ray lognormal fading standard deviation also in dB.

Four different configurations are proposed in [35], to match four different indoor environments, which are: line-of-sight (LOS (0-4m)), Non line-of-sight (NLOS (0-4m), and NLOS (4-10m) channels and extreme NLOS environment with 25 nsec RMS delay spread. We call these channel models, Channel Model 1 (CM1), Channel Model 2 (CM2), Channel Model 3 (CM3) and Channel Model 4 (CM4), respectively. We assume that CIRs are generated

Table 3.1 Multipath channel model parameters [35].

$ModelParameters$	Λ	λ	Γ	γ	$\sigma_{n_1}(dB)$	$\sigma_{n_2}(dB)$	$\sigma_x(dB)$
$CM1$	0.0233	2.5	7.1	4.3	3.3941	3.3941	3
$CM2$	0.4	0.5	5.5	6.7	3.3941	3.3941	3
$CM3$	0.0667	2.1	14.00	7.9	3.3941	3.3941	3
$CM4$	0.0067	2.1	24.00	12	3.3941	3.3941	3

according to the IEEE 802.15.3a channel model with parameters defined in [35] for four different indoor environment models. These models provide the best match to significant characteristics of the real world channel. The values of the parameters for these environments are specified in Table 3.1.

3.1.4 UWB Receiver

In this section, we explain the detection of UWB signals that transmit pulses with PPM or PAM modulation from the multipath environments. The signals are taken from a multipath channel that their parameters are known. The receiver performs the opposite operation of the transmitter to recover the data. The receiver has a matched filter to optimized the received signals respect to Additive white Gaussian noise (AWGN) and also detection function of the received pulses amongst the other signals. These information could be helpful to detection of pulses in the UWB systems as we will explain in the following.

The signal at the receiver is determined by:

$$R(t) = S(t) \otimes h(t) + n(t) \tag{3.13}$$

where $\otimes$ and $n(t)$ are, respectively, convolution operation and Additive White Gaussian Noise (AWGN) introduced at the receiver, while $h(t)$ is the multi-path random time-varying channel impulse response.

The optimum wideband signal receiver is a rake receiver, that minimizes the probability of error by using a bank of Single User Matched Filters (SUMFs) to capture pulses and improve the pulse detection performance [27]. The UWB receiver system takes decisions for detection of each binary bit stream as shown in Fig 3.2. The estimated bit is given by:

$$b_i = \int_{\tau}^{\tau+T_i} (R(t) \times (M_i(t) \otimes h(t))dt, \tag{3.14}$$

where $M_i(t) \otimes h(t)$ is the matched filter for i^{th} bit detection that $M_i(t)$ obtain as follows:

$$M_i(t) = \sum_i \sum_{n=0}^{N_s-1} (w(t-(iN_s+n)T_f)) \tag{3.15}$$

3.1.5 Inter-Symbol Interference

The maximum excess delays of IEEE 802.15.3a channel model for CM1, CM2, CM3 and CM4 are around 30 ns till 300 ns [35]. Hence, the IR-UWB systems some portion of the transmitted symbol energy inevitably leaks into the consecutive symbols at high data rates which is called as inter-symbol interference (ISI). ISI is one of the major factors in reducing the detection performance in the UWB systems. Therefore, it is important to consider

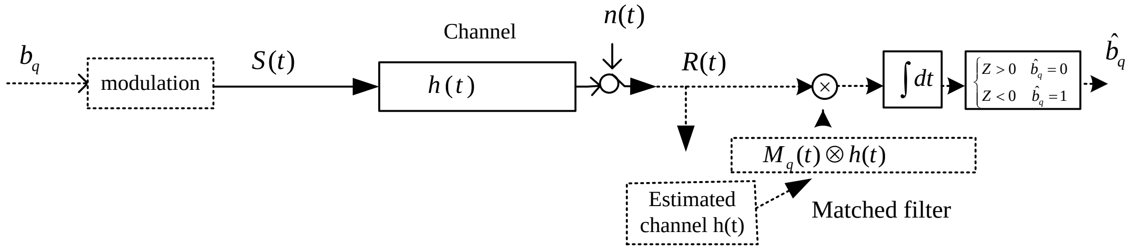


Fig. 3.2 Structure of UWB transceiver that encodes information with multipath channels.

different model to mitigate the ISI in the higher data rate. One of the solution to avoid ISI is a TR prefiltering method. In the next section will discuss about TR.

3.1.6 Time Receiver Pre-filtering

The Time Reversal Pre-filtering (TRP) exploits the time-reversed, complex conjugate channel for transmitting information as shown in Fig 3.3. In this structure, both Transmitter (Tx) and Receiver (Rx) need the perfect knowledge of the channel that can be obtained by executing the following procedure once every coherence time [23]:
It is assumed that transmitter and receiver are perfectly synchronized and signal transmission is carried out as following steps:

1. Tx sends a token pulse enabling Rx to estimate the channel impulse response;
2. Rx sends the estimates CIR to Tx, using a feedback channel;
3. Tx transmits pre-filtered symbols with the channel reverse conjugate for the duration of the coherence time as Equation 3.16.

$$R(t) = S(t) \otimes h^*(-t) \otimes h(t) + n(t), \tag{3.16}$$

where $*$ denotes complex conjugate operation.

The UWB receiver system takes decisions for detection of each binary bit stream as shown in Fig 3.3. The estimated bit is given by:

$$b_i = \int_{\tau}^{\tau+T_i} (R(t) \times (M_i(t) \otimes h^*(-t) \otimes h(t))dt, \tag{3.17}$$

where $M_i(t) \otimes h^*(-t) \otimes h(t)$ is the matched filter for i^{th} bit detection that $M_i(t)$ obtain as follows:

$$M_i(t) = \sum_i \sum_{n=0}^{N_s-1} (w(t - (iN_s + n)T_f)) \tag{3.18}$$

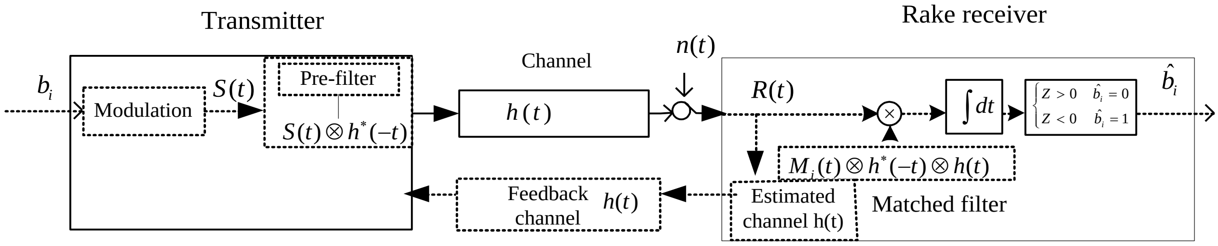


Fig. 3.3 Structure of TR based on UWB transceiver that encodes information with multipath channels.

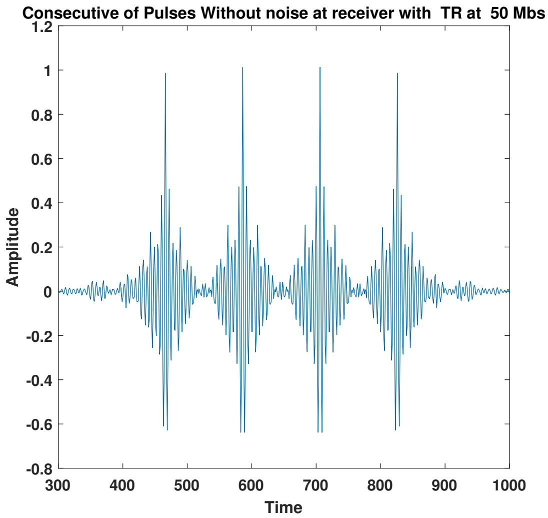
3.1.6.1 Time Receiver Pre-filtering Benefits

TR prefiltering lead to concentration of received signal power in the signalling as shown in Fig. 3.4. It illustrates the detection capability of TR respect to no TR. It is powerful method for mitigating MUI and ISI specially in high data rate. The proposed technique in TR relies on the capability of modeling the UWB channel with high accuracy. As we explained in 3.1.3, accurate channel models are also fundamental in performance evaluation of TR.' The TR technique provides better performance, higher achievable data rates and more secure data transmission with respect to traditional (no time reversal) UWB systems and also reduces the complexity of the receiver.

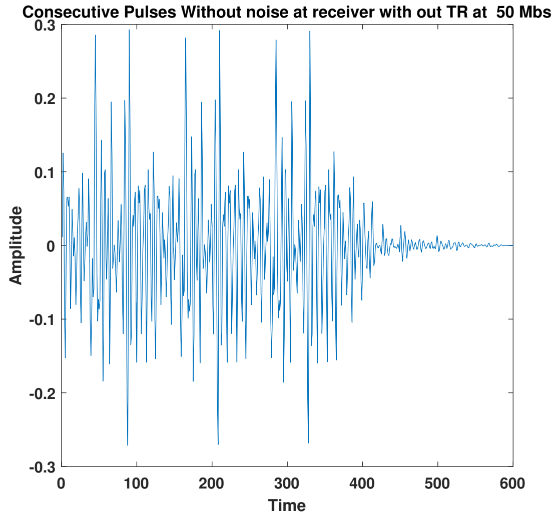
3.1.6.2 Time Receiver Pre-filtering Problems

In theory, the TR pre-filtering technique requires the complete channel information for feedback. However, transferring such a massive information requires a large bandwidth feedback channel. This makes the approach unpractical in UWB systems [30].

The conventional solution is to introduce a digital feedback system that conveys a quantized version of channel taps. The quantization operation introduces however several sources of



(a) TR



(b) Without TR

Fig. 3.4 Four consecutive pulses at receiver with TR and without TR.

errors.

1. The first one is the unrecoverable A/D conversion error that prevents from perfectly recovering the original channel: the lower the number of bits per samples the largest the conversion error.

   For example, let us assume that each CIR has $K(l)+L+1$ samples, requiring $(K(l)+L+1)*b$ bits to be sent back via the feedback channel, where b denotes the bits per sample used in the quantization corresponding to 2^b levels per sample.

2. The second source of error arises from the noise in the feedback channel. The error probability of the received bits from the feedback channel at the transmitter side is not zero and it depends on the noise of the environment.

Since in UWB systems the channel is characterized by a short coherence time ranging from hundred to several hundred nano seconds depending on the environment, it is crucial to increase the number of levels in order to increase the accuracy of the feedback information and to decrease the quantization errors in the feedback system. As result the amount of information to be sent back can become unsustainable.

3.1.7 Multi-User Interference

The effect of multi-user interference (MUI) on IR-UWB Networks are argued in the many real system. The simple way to avoid MUI is using multiple access methods which can be Time Hopping (TH) and Direct Sequence (DS) methods. We will mention TH in the following work. It is also useful to use TR method in parallel to concentrate the transmitting signals over the same link.

3.1.7.1 Time Hopping

TH is a pseudo-random code (C) to avoid collision in multiple access [27] which can be used in the IR-TH-PAM scheme to encodes symbol sequences as follows [27]:

$$S(t) = \sqrt{E_s/N_s} \sum_{i=0}^{m} \sum_{n=0}^{N_s-1} b(i) w(t - (iN_s + n)T_f - C_{(iN_s+n)} T_c) \tag{3.19}$$

The pulse is sent with C_{iN_s+n} delay time which is bounded in the frame interval T_f. The $(C_i)_{i \in Z}$ sequence is generated from a pseudo random TH code taking values in $(0, 1, ..., N_h - 1)$. T_c is the chip time interval. The IR-TH-PPM can be determine by:

$$S(t) = \sqrt{E_s/N_s} \sum_{i=0}^{m} \sum_{n=0}^{N_s-1} w(t - (iN_s + n)T_f + C_{(iN_s+n)} T_c - b(i)\varepsilon) \tag{3.20}$$

where $C_{(iN_s+n)} T_c + \varepsilon < T_f$ condition should be considered for all $C_{(iN_s+n)}$.
As explained in Section 3.1.6, transmitting the signal defined in Equation 3.19 without any additional operation poses several problems. Therefore, the time reversal pre-filtering (TRP) technique was proposed to address such issues.

3.2 Bluetooth

Bluetooth is one of Low power technologies that provide the tools and abilities to design low-cost applications which can operate and communicate in mobile environments, in the optimal case using only a single chip. The Bluetooth was originally designed to provide a useful and powerful technology for indoor wireless communication as [43] , [7] that is widely available today. It argues that Bluetooth is a technology well suited for comparison with UWB signal. The study is demonstrated that 69 million Bluetooth chips were shipped in 2003, and are predicted 720 million units till 2008 [8]. Laptops, cell phones, and PDAs are increasingly shipped with an integrated Bluetooth radio.

3.2.1 Bluetooth Generator

This section presents an overview of Bluetooth, with a special emphasis on the parts that concern a Bluetooth generation. We explain Bluetooth signal generation in the context of signal recognition and shows how most concepts in recognition of UWB are easily discriminate from Bluetooth. It describes how these concepts can be implemented in the Matlab simulink using the Bluetooth protocol. In the setup phase, Bluetooth signals were synthetically generated using MATLAB as Fig 3.5. Following the Bluetooth standard (IEEE 802.15.1), time is divided into fixed slots of. It is used a simple Bluetooth wireless data link, which applies Gaussian Frequency Shift Keying (GFSK) over a radio channel with maximum capacity of 1 Mbps. The method executes a 79 frequency hops for each packet. The duration of each packet was randomly generated according to the IEEE 802.15.1 protocol. It transmits low power signals, which are used often for the indoor applications.

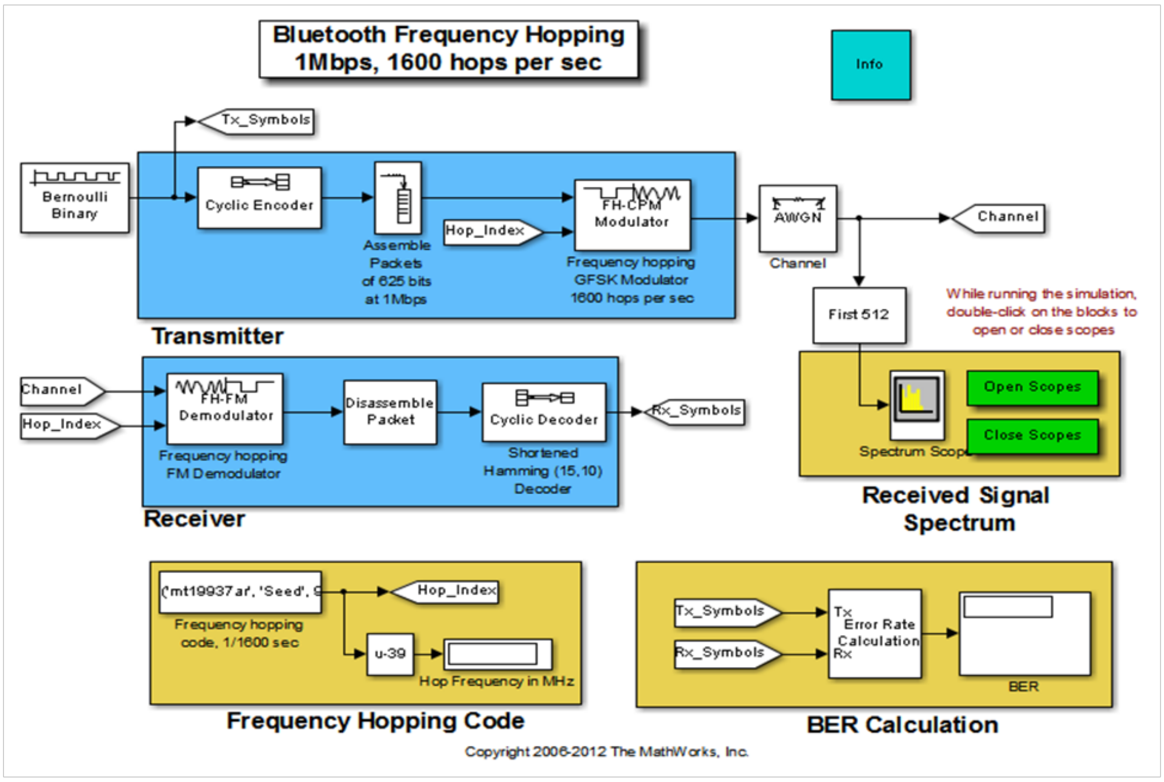


Fig. 3.5 Bluetooth generator structure of Matlab

Chapter 4

Machine Learning

In this chapter, a review of machine learning will present, which can be used for some application of communication systems such as recognizing data points and clusters. Machine Learning consists of three steps, namely data acquisition, feature extraction and classification or clustering.

Two main parts of machine learning are divided into supervised and unsupervised learning. In the supervise learning the labels of data is known, hence these methods are useful to classify data points that will be explained in section 4.3. However, in the unsupervised learning is trying to find unlabeled data structure which is explained in section 4.4. Unsupervised methods are useful for clustering data points which we do not have any information about the number of clusters and labels of data points. Finally, in section 4.5 we discuss about evaluation criteria.

4.1 Feature extraction

In general, the features are significantly important to recognize, classify and cluster data points. They are represented by properties of data points. The compositions of different kinds of patterns, also called feature spaces, define data points in a high dimensional features. In order to classify and cluster data points, an appropriate set of features has to be selected. These features are expected to satisfy several aspects as follows:

1. They should be distinguished the objects clearly;
2. They should be invariant to irrelevant transformations of the given data;
3. They should compact in dimensionality in order to reduce memory consumption and computation time;
4. They should be easy to extract and insensitive to noise.

The feature extraction methods may involve prior knowledge of the problem domain, but appropriate features are mostly not easy to find and often lead to long time evaluations [42].

4.2 Similarity

Different similarity (or dissimilarity) criterion function for computing the distances between each pair of nodes has been proposed in literature [44], [28]. The distance, such as Euclidean, Mahalanobis and Minkowski are good measures of dissimilarity that attempt to minimize the same distance between the points in the same cluster in order to reach the optimal cluster heads. However, the Cosine, Jaccard, Spearman and Pearson Correlation functions compute the similarity of cluster data point.

Weighted Overlap (WO) is another similarity technique which compares two data points based on the relative ranking of their dimensions. Two similar data points are expected to have their dimensions in similar orderings. In comparison to the Spearman correlation, WO provides the advantage that it penalizes the differences in the top-ranking dimensions more than it does for the bottom ones [61].

4.3 Supervise learning

From the several available classifiers, the following ones were selected: Support vector machine with different kernels, K nearest neighborhood (KNN), parabolic neural network

(PNN) and Naïve Bayes [44], [28]. Support vector machine arguably are the most successful classification method in machine learning [69]. SVM are principled derivation methods that use optimization packages to get the solution and classification. Benefit of machine learning in support vector machine is expanded from linear to non linear cases, as described below [69].

4.3.1 Support vector machine

Most of previous works on classification have shown that Support vector machines (SVMs) are the most successful classification methods in machine learning [15]. SVM with kernels are basically derivation methods that use optimization schemes to get the solution and the classification. In the following, we will shortly review the most famous kernels used in SVM classifier. Four of the following kernels are non linear. The benefit of the non linear cases is that, by using them, it is possible to map the data on a high dimensional space, allowing a better data classification, even if, as some experimental work have shown previously, a lower robustness than linear case is possible [69].

(I) **Gaussian RBF** We use RBF Kernel with the Sequential Minimal Optimization (SMO) Method [73], a faster algorithm for training SVM that uses pair wise classification to break a multi-class problem into a set of 2-dimensional sub problems, eliminating the need for numerical optimization.

$$k(x_i,x_j) = exp(-\gamma(x_i^T - x_j)^2) \tag{4.1}$$

(II) **Multilayer Perception** We use MLP Kernel with the least-squares (LS) Method [62].

$$k(x_i,x_j) = S(x_i^T x_j) - 1) \tag{4.2}$$

S is a sigmoid function.

(III) **Linear kernel** We use linear Kernel with the SMO Method .

$$k(x_i, x_j) = x_i^T x_j \tag{4.3}$$

(IV **Polynomial kernel** We use Polynomial Kernel with the least-squares (LS) Method. We consider a polynomial kernel power to degree 4.

$$k(x_i, x_j) = (x_i^T x_j + 1)^4 \tag{4.4}$$

(V) **Quadratic kernel** We use quadratic Kernel with the least-squares (LS) Method.

$$k(x_i, x_j) = (x_i^T x_j - 1)^2 \tag{4.5}$$

4.3.2 Probabilistic Neural Networks (PNNs)

Artificial neural network is composed of a set of neurons which are connected together in different layers. Connection of layers is a mathematical formula, which is like a multidimensional polynomial. Parabolic neural networks are based on biological neural networks processing the information [76]. PNNs are able to classify data. Different signals have their own patterns which are useful for the aim of classification. PNN is an adapted version of radial basis function (RBF) which estimates the probability density functions. PNN, composed of multiple layers, is trained faster and produces more accurate models, compared to the other neural networks. PNNs utilize an input, a hidden, and an output layer and they are suitable for classification problems. The spread parameter is set to 0.1 which yields the best results in PNNs. The network acts as a nearest neighbor classifier if the spread parameter is near zero.

4.3.3 Naive Bayes classifier

Naïve bayes classifiers are discriminate and supervise leaning methods that optimize conditional likelihood. The intuition of naive bases is very simple and it is based on Bayes Rules. In this classifier output of probability condition is prediction of classifier and the likelihood of the estimators is maximized. This classifier is very simple but the performance is quite good.

4.3.4 K nearest neighbours

One of the simplest classification algorithms is the K nearest neighbours [54] [61]. KNN takes new point and classifies it according to the majority vote of the K nearest point in the data set which is called training data. Majority data vote determines that new data belongs to which class. It is a regular method for classification with the optimum number of the closest neighbours and the most suitable distance. New coming data looks at and query from K neighbourhoods then calculates distance from these K neighbourhoods and finally samples map to the nearest group and maps to the nearest one. Four KNNs have been employed with different distance functions; however we use the Euclidean distance because it does not affect the classification accuracy. We choose K=3 neighbourhoods to evaluate the total experiment. In this case, none of the results of the K-NN would be stable and thus valid for classification

4.4 Unsupervised learning.

The main goal is to accelerate these processes by utilizing unsupervised learning algorithm to cluster channels. In this Section, we investigated with respect to this goal. Clustering is an unsupervised machine learning approach adopted in scientific data analysis. A clustering of data partitions the set of unlabeled data in a set of clusters and associates cluster label to data. The number of cluster heads and also the category labels of data points are unknown.

Common clustering approaches discover a set of cluster heads so as to minimize the sum of squared errors between data points within a cluster and their nearest cluster heads.

4.4.1 Clustering algorithms

The first issue is the selection of the clustering algorithm itself. Clustering methods have been organized in several categories [82]. K-centers clustering methods are divided into K-medoids clustering, K-medians clustering and K-means [68]. K-means initializes a set of cluster centers [53] and is the most popular and most commonly used algorithm employing a squared error criterion. It iteratively updates the center of each cluster by minimizing the overall Euclidean distance between the data points and the cluster heads till a convergence criterion is met. However, k-median clustering minimizes the median value of the Euclidean distance of all data points from cluster heads. K-medoids clustering algorithm takes as input from an initial set of exemplars and then iteratively minimizes the maximum value of city block distance of all data points assigned to the closest cluster heads. The k-centers clustering algorithm is sensitive to choosing the initial set. Hence, it needs to be run with many different random initializations to give opportunity to random node to be the cluster heads in order to find the best solution. It works well only when the number of clusters is small and chances are good that at least one restart lies close to a good clustering solution.

In the density-based cluster methods it is supposed that the data of each cluster has a specific probability distribution. Hence, each cluster node is identified based on fitting their probability distribution [4]. Simulated Annealing [67] is suggested to search globally so as to avoid local optima. It is a stochastic search technique which is applied to small sets while learning execution times [82]. Hierarchical Methods cluster the data by partitioning the samples in either a top-down or bottom-up approach recursively, which are called respectively agglomerative hierarchical and Divisive hierarchical clustering algorithm. By giving the chance to all of the nodes to be clustered heads, hierarchical algorithms merge pairs of clusters; therefore, they are independent of the initial cluster number and conditions and priori [82]. We consider 3 algorithm for clustering we are going to explain in detail.

4.4.1.1 Vertex substitution heuristic

The first algorithm, called vertex substitution heuristic (VSH), uses p-median as a resolving method to locate p facilities (clusters) to N customers (data point) in order to minimize the total cost of demands between each customer and the closest facility [45]. VSH finds an optimal replacement for p facilities [45]. VSH is used as a clustering algorithm on the basis of the similarity matrix in [74]. The agorithm minimizes the similarity between channels and channel cluster head candidates. VSH selects randomly columns from similarity that are called exemplar (cluster heads) candidates. is the number of clusters which are set manually in the initial step. VSH is iteratively refined by evaluating the effect of substituting an unselected point for one of the selected cluster heads. The selected columns become cluster heads if the sum of the maximum values of rows is maximized.

4.4.1.2 Expectation maximization

The second clustering algorithm that we consider in our experiments is K log(K) – EM. The algorithm performs an initialization step that starts with the K log(K) initial sets and then prunes small clusters with farthest-first algorithm in order to discover K initial clusters [18, 19]. The initial set is used to implement an EM updating algorithm [22]. The EM algorithm estimates cluster similarities and consequently it updates the cluster heads in each maximization step.

4.4.1.3 Affinity propagation

In most of the clustering algorithms, the number of clusters should be specified before carrying out the partitioning. Determining the number of clusters without prior data knowledge is a however difficult task and significantly affects the clustering algorithm's performance because there are of possible distinct cluster centres. The initial cluster number is typically determined manually or based on different algorithms such as Akaike information criterion (AIC) [3] or Cluster Separation [20] which aim at estimating the optimal number of clusters.

Affinity propagation (AP) is an exemplar-based unsupervised hierarchical clustering algorithm [36] which evaluates all the data points as possible cluster heads. Thus, the algorithm does not need to know the number of clusters in advance. Instead, AP automatically determines the number of cluster heads by adjusting preference values. Preference is a real number in the similarity matrix . A low value of preference results in low number of clusters. Therefore, the main reasons to opt for the AP clustering algorithm is its capability of selecting cluster heads without having any prior knowledge of the number of cluster heads. A different approach is adopted in Affinity propagation (AP), an exemplar-based unsupervised hierarchical iterative clustering algorithm [36]. AP evaluates in fact all the data points as possible cluster heads: thus, the algorithm neither needs to set the number of clusters in advance, nor to perform an initial random selection. Instead, AP automatically determines the number of cluster heads as result of a set of preference values, defined as the value of similarity $S(i,i)$ of each point i to itself.

AP is based on messages exchanges between points until a set of cluster heads emerges. More specifically, the AP algorithm takes the similarity between pairs of points as input and its output is a set of clusters. In its basic form, the AP algorithm is implemented by exchanging and updating Responsibility and Availability messages between the channels. Responsibility $R(i,j)$ denotes how well-appropriate channel j is to be as the cluster head for i. Availability $A(i,j)$ estimates how well-suited data point i is to substitute j as a cluster head. The damping factor ρ_d is used in AP algorithm when the Responsibilities and availabilities update in each step by the damping factor ρ_d that leads to avoiding numerical oscillations caused by overshooting [36].

The AP clustering algorithm is presented in Algorithm 1. At each iteration, each channel sends a message which reflects the current affinity that the channel has for choosing another channel as its cluster head. Specifically, the data point j is the cluster head of the data point i when the value of $A(j,j)+R(j,j)$ in $i=j$ converges after a fixed number of iterations. Upon convergence, $\Phi(j)$, i.e., the index of the cluster head assigned for each channel, is determined. These iterations are minimum number of iterations $I_{t_{min}}$ and maximum value $I_{t_{max}}$. The minimum value is applied if no change is observed message results till iteration of message passing stop. In any case, AP is terminated after a maximum of $I_{t_{max}}$ iterations.

Algorithm 1 Discover cluster heads using AP algorithm in learning step.

Input: Realize N data point.
Output: Cluster labels of each data points $\Phi(j)$
Initialization:
$S(j,j)= \left\{ \begin{array}{ll} \text{use equations } Section 4.2 & \forall i,j \in k, i \neq j, \\ p, & \forall i,j \in k, i = j, \end{array} \right\}$
$A(i,j) \leftarrow 0, R(i,j) \leftarrow 0, \forall i,j \in k$
while $T \leq Threshold$ **do**
    updating responsibility and availability till converge
    $R(i,j) \leftarrow S(i,j) - \max_{j' \neq j} \{A(i,j') + S(i,j')\}, \forall i,j \in k$
    $R^{new} = (\rho_d)R^{new} + (1-\rho_d)R^{old}$
    $A(i,j) \leftarrow \left\{ \begin{array}{ll} \min\left[0; R(j,j) + \sum_{i' \notin (i,j)} \max\left[0; R(i',j)\right]\right], & \text{for } i \neq j, \forall i,j \in k, \\ \sum_{i' \notin j} \max[0, R(j,j)], & \text{for } i = j, \forall i,j \in k, \end{array} \right\}$
    $A^{new} = (\rho_d)A^{new} + (1-\rho_d)A^{old}$
end while
Calculate cluster head's $j \leftarrow [A(j,j)+R(j,j)] > 0$,
Return selected labels ($\Phi(j)$) for each channel based on maximum similarity from column of similarity

4.5 Evaluation Criteria.

We used different metrics to measure the performance of our proposed algorithm. These metrics are useful to analyze data without prior assumption about the data. For signals classification, Confusion matrix is a particular table layout that allows visualization of the efficiency of an algorithm in Machine Learning. Each column of the matrix denotes the instances in a predicted class, while each row is the instance in an actual class. The accuracy (AC) is the proportion of total true positives divided by the sum of the total true positives pulse false negatives and false positives each class. The recall (correct classification rate) is the proportion of true positive each class divided true positive pulse false negatives. The precision is the proportion of the predicted true positive each class divided true positive same class and false positives classes. The F-measure is the harmonic mean of precision and recall and calculated as: 2*recall*precision / (recall+ precision). For clustering channels, the optimal solution must demonstrate

Chapter 5

Analysis in noisy channels and selected feature analysis

One of the most important issues in UWB is that there exists interference with other signals and noise which sometimes is too much. UWB overlaps with different technologies, especially in unlicensed bands such as ISM, therefore we need to explore the behavior of UWB signals versus other signals. In our configuration the receiver has four different states, UWB, Bluetooth, UWB plus Bluetooth and noise. Figure 5.1 shows the system model.

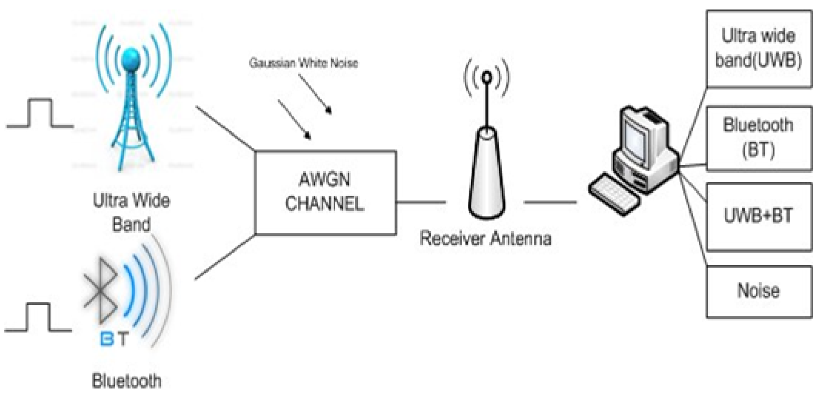


Fig. 5.1 System model Schematic

5.1 Simulating parameters

Data is analyzed to recognize and detect UWB and Bluetooth signals, as mentioned in previous sections. One of the most important issue in UWB is that there exists interference with other signals as well as noise.

We consider that UWB signals used 2PPM-TH modulation technique in transmitter as section 3.1.7.1. UWB signals were generated form Equation 3.20 with the following parameters:
Number of pulses per bit was $N_s = 1$, Frame time was $T_f = 10ns$, Chip time was $T_c = 1ns$, PPM time shift was $\varepsilon = 0.5ns$, Pulse shaping factor was $\alpha = 0.25ns$ and Pulse duration was $TM = 0.5ns$.
Bluetooth data under analysis in this section was synthetically generated using MATLAB as explained in Section 3.2. The time window was set to 5 ns that is a very short interval respect to the standardized Bluetooth time slot ($625\mu s$). For this reason, the Bluetooth frequency hopping nature does not affect the evaluation of the chosen features. Additive Gaussian white noise is used in the simulations.

5.2 Energy profile features

As explained in Section 4.1, extracted features are significant to classify the signals. In this section, we explain the feature extraction based on energy profile.
Received signals are illustrated in Fig. 5.2. The left-hand part shows signals for $10ns$, with $40dB$ SNR. Intuitively, shapes of singletons are similar to noise, especially in high SNR, and as a result are difficult to identify. The right-hand part of Fig. 5.2 shows the short-term energy of the signals for overlapping time windows. It indicates that the energy variation of

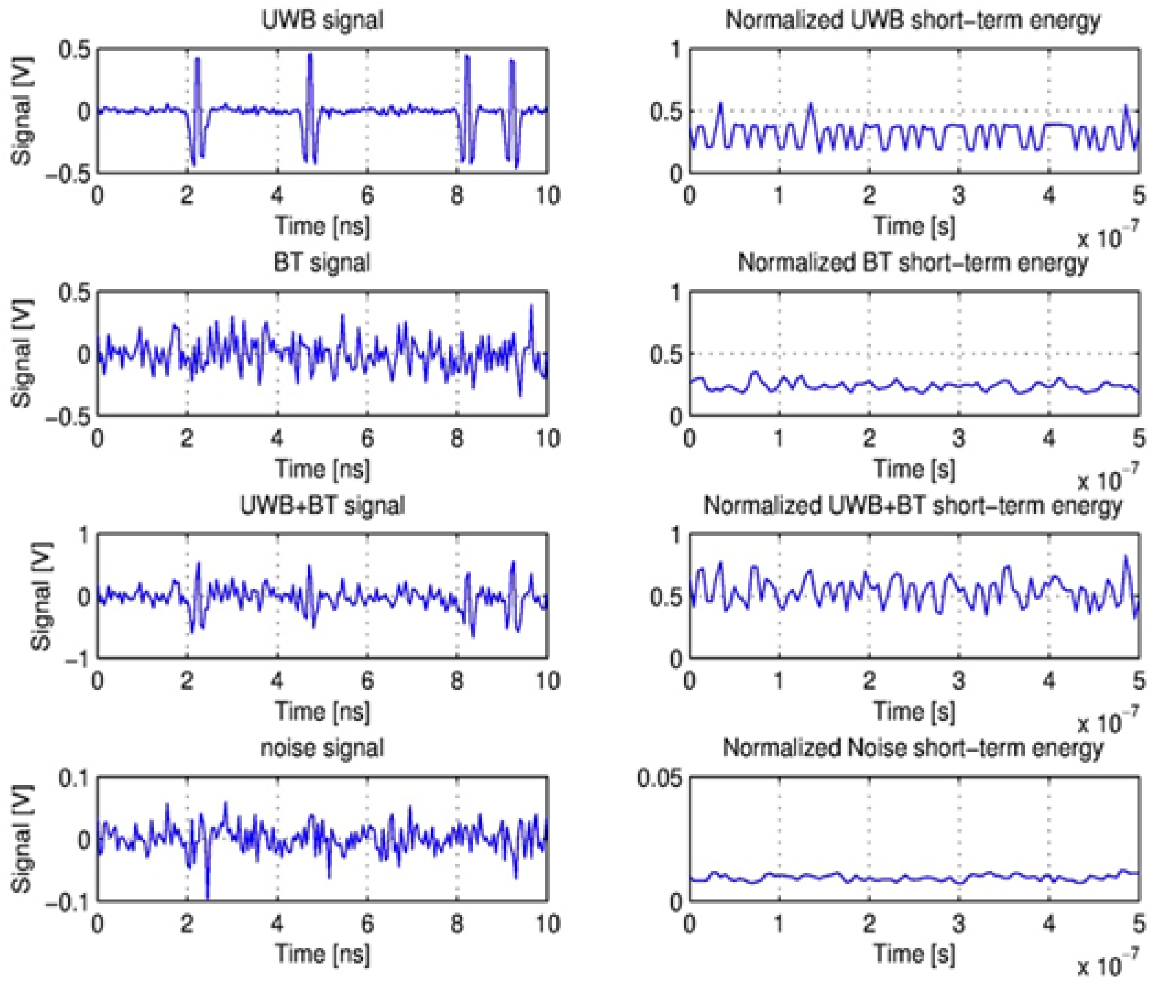


Fig. 5.2 Different Signals and their energy with 40 dB SNR

UWB is smooth, which is a helpful behavior to recognize it from other signals.

In order to extract features from signals, the whole signal was subdivided into overlapping time windows. Then, we calculated energy for each of the obtained time window. In our experiments, we selected 5 ns for the length of time windows with 50 percent overlap. We extract features from energy profile and derivative of energy signals that are generated with Matlab. As shown in Fig. 5.2 right part, we calculate energy in the small time window of capturing in which the length of windows are $5ns$ and 50 percent have overlapping. As we know all of signals have 4 class labels: UWB, Bluetooth, overlapping UWB and Bluetooth signals and noise, respectively.

Short-time energy features, since calculated in small time windows, provide a robust tool for detecting UWB signals. This because, given the faster rate of change of the total energy in UWB signals due to their impulsiveness, the feature provides means for discriminating UWB from other signals. On the other hand, Bluetooth signals have continuous behavior similar to sinusoids within packet duration, and for this reason Bluetooth short time energy has smoother variation with respect to the UWB one.
We leveraged seven features in our experiments: short time energy, amplitude of derivative of short time energy, phase of derivative of short time energy, fast Fourier transform short time energy, short Fourier transform of short time energy, short time energy gradient, and Gaussian normalize window short time energy.

Data were divided so that 70 percent were used to train and the remaining was used to classify. We listen to the noisy data received from the antenna.

5.3 Feature Selection

According to analyzed data we applied energy profile as feature. Mentioned features are useful to classify data. Next step, two feature selection algorithms evaluated features.

When large amounts of data are available, feature selection methods are useful to reduce the amount of features. The low rate of variation of the efficiency should be considered as well. To this end, two feature selection algorithms were applied:

1. minimum Redundancy Maximum Relevance (mRMR) (applied with three different methods) [60];
2. Genetic Algorithm with Information Theory (GA) [52].

The mRMR and GA algorithms were used in order to rank the features which are able to better classify the samples for each scenario. In general, mRMR algorithm calculates features

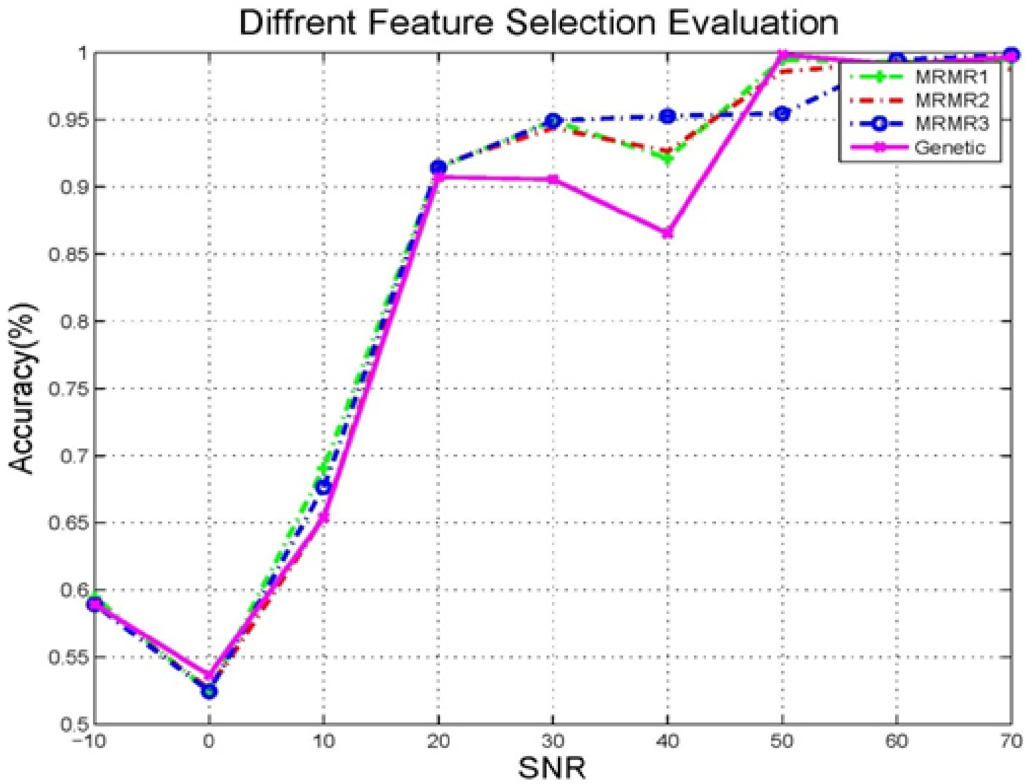


Fig. 5.3 Evaluation feature selection in different SNR

relevance and redundancy using Mutual Information (MI). Following [60], this algorithm uses three different criteria to select features from the sorted subsets:

1. Mutual Information Difference (MID (mRMR1)), defined as the difference between relevance and redundancy;
2. basic scheme (mRMR2), in which no particular combination of relevance and redundancy is defined;
3. Mutual Information Quotient (MIQ (mRMR3)), defined as the ratio between relevance and redundancy.

In any case, the output of each criteria is a vector with the indexes of the features that composes the optimal set of features.

We show in Fig. 5.3 different feature selection methods that are applied with SVM RBF classifier on several SNRs, when we use feature selection methods. We have selected four best features using feature selection algorithm slot that were sorted with the mRMR algorithm using 3 different criteria and genetic algorithm. mRMR algorithm looks at the relevance and redundancy of features by using mutual information measure between features and labels. We compare the behavior of different feature selection methods in Fig. 5.3, in which the x axis is SNR. We can select which of the extracted features are important. All results are experimental and here we want to have fewer features. Fig 5.3 demonstrates the effect of feature selection showing that with fewer data we have the same performance.
The best 4 features are chosen by the different feature selection algorithms as following:

1. mRMR1: Short window time energy, Gaussian normalizes window time energy, and Short Fourier transform of short time energy and Fast Fourier transform short time energy.

2. mRMR2: Short time window energy; Fast Fourier transform of short time energy, short Fourier transform of short time energy and Gaussian normalize window short time energy.

3. mRMR1: Short window time energy, short Fourier transform of short time energy, fast Fourier transform short time energy and Gaussian normalize window time energy.

4. GA: Short Fourier transform of short time energy, amplitude of derivative of short time energy, Gaussian normalize window short time energy and short time energy gradient.

5.4 Performance of different classifiers

We carried out experiments with support vector machine based on five different kernels, K nearest Neighborhoods, parabolic neural network and naïve bayes classifier as explained

in section 4.3. In order to find the optimal separating hyper plane in SVM, we performed training using the least-squares, Sequential Minimal Optimization and quadratic programming methods. The basic principle of SVM is to construct the optimal separating hyper plane which maximizes the distance between the closest sample data points in the (reduced) convex hulls for each class, in an n-dimensional feature space.
We selected 1916 samples for training and testing. The output of the classifiers is a prediction value of the actual samples of classes. In order to evaluate the performance of a classifier, the repeated held-out cross-validation method is used. According to this method the samples of each class in the data collection are divided into a training set containing 70 available data and a disjoint test set containing the remaining 30 of the data. The training and the test set are selected randomly. The classifier is trained using the training set and the recall and accuracy is estimated on the test set.

Accuracy is an advantageous metric to measure the systems in many applications which is calculated form confusion matrix. Fig. 5.4 is accuracy performance of signals detection with classifiers. Accuracy illustrates total performance, which means the capability of classifiers to discriminate classes' samples. High accuracy demonstrates power of classifier to separate data of 4 classes. Signals corresponding evaluation, represented in Fig. 5.4, is the total performance of the approaches with regards to the all 4 classes. Fig. 5.4 demonstrates the accuracy of instances in -20 dB SNR up to 60 dB. SVM Linear and MLP have massive misclassification data, meaning that SVM MLP has worst result in whole SNR. If we look at all of SNR, SVM RBF has acceptable efficiency. Other classifiers have very similar results; hence, classifier selection in different SNR is imposed according to the maximum accuracy. For example, in SNR -20 dB we should choose the best classifier, which is the SVM quadratic and has 52 percent detection rate. SVM linear has the best performance in -10 dB SNR, however, SVM RBF has best classification rate in more than 20 dB SNR.

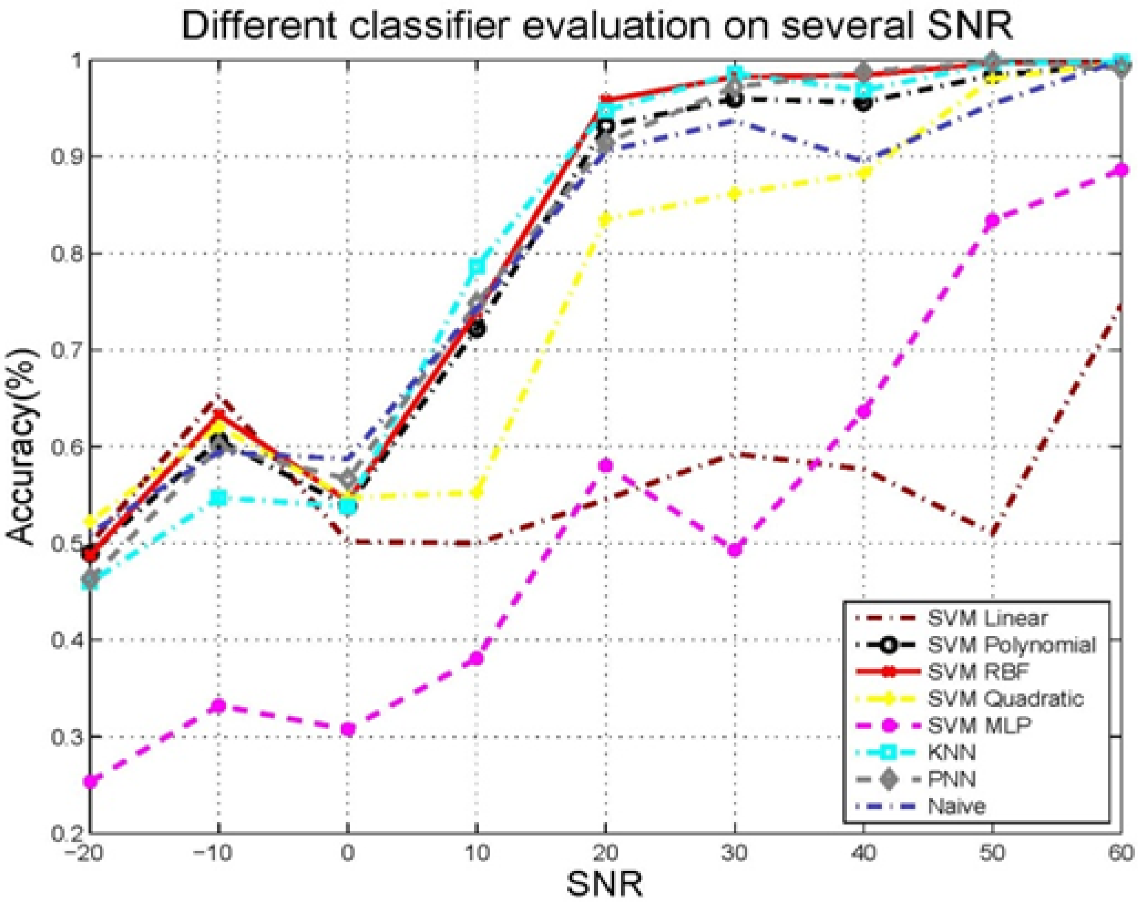


Fig. 5.4 Evaluation classifier in different SNR

5.5 Result with Confusion matrix

We have investigated only 4 classes in the database in which they always had 20s dB SNR. We suppose that the distances between transmitters and receivers are constant. The performance of the predictive model is investigated using the confusion matrix in Table 5.1. Each class has 143 test observations that are nominated of energy profile and its derivative which are samples of the corresponding class signals. Basically, the highest number in diagonal show that test observation is correctly classified; on the other hand, any number in non-diagonal part means that it is not correctly classified. Table 5.1, provides in detail, the confusion matrix measured for the SVM Polynomial and shows that it has the best performance to detect UWB compared to the other classifiers. For detection of Bluetooth, we leveraged K nearest neighbors that have 136 true positive. Also detection rate of SVM RBF for detecting UWB+ Bluetooth is the best, between the others. In the detection of noise, majority of classifiers have acceptable prediction in 20 signals to noise ratio. Our proposed method has good performance to recognize noise in 20 dB SNR.

5.6 Performance detection of class

In the following, the behavior of the best classifier (SVM with RBF kernel), is investigated against changing the test data with 10 dB SNR. In general we can use different metric that first one recall that is classifier sensitivity. Fig. 5.5 highlights the behavior of the classifier on the four classes and for varying numbers of cross-validation repetitions and varying parts of samples used in testing. As shown in Fig. 5.5, for 10 dB SNR, UWB+BT with RBF kernel provides the best correct classification rate of 99.3 percent, 98.61 percent, 98.9 percent respectively. We have seen precision Noise has the second best performance of 96.7 percent, whereas UWB have the recall with SVM RBF of 93 percent. The recall of noise signals for

Table 5.1: Confusion matrix on the 7 features when 30 percent of the samples of 20 dB SNR data are used for testing.

	mRMR on 20 dB with 7 features								
SVM Polynomial	133	8	1	1	SVM Linear	79	48	11	5
	14	129	0	0		16	102	25	0
	12	3	128	0		37	97	9	0
	0	0	0	143		21	0	0	122
SVM Quadratic	131	12	0	0	SVM RBF	130	12	0	1
	51	92	0	0		9	133	1	0
	13	18	112	0		1	0	142	0
	0	0	0	143		0	0	0	143
SVM MLP	79	42	16	6	KNN	122	21	0	0
	50	83	9	1		7	136	0	0
	68	22	53	0		1	1	141	0
	26	0	0	117		0	0	0	143
PNN	102	39	1	1	Naive	105	32	6	0
	4	139	0	0		11	131	1	0
	1	3	139	0		4	0	139	0
	0	0	0	143		0	0	0	143

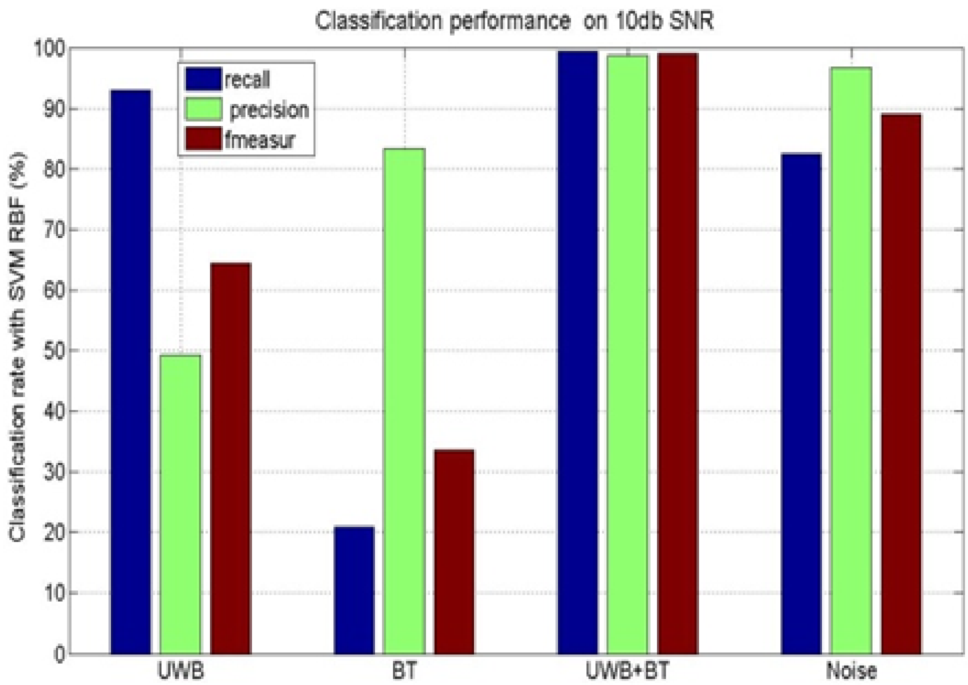


Fig. 5.5 Behavior of difference classifier on difference classes for classification when the size of the test signals is 30 percent

RBF is 82.5 percent. Bluetooth (BT) provides performance of 33.5 percent according to F-measure when RBF is used.

Chapter 6

Digital Feedback

The problem of TR mentioned in Section 3.1.6.2. In this chapter, we propose a new feedback design, based on vector quantization for TRP-UWB that takes advantage of a machine learning algorithm to reduce the amount of information to be sent on the feedback channel.

6.1 Digital Feedback Design

In order to solve TR problem in this Section a new digital feedback system based on Channel Cluster Heads (CCHs) is proposed. As it will be shown later, the proposed feedback system requires less bits to be fed back. Hence, this allows to reliably protect the bits from communication errors and additionally reduces the time required to send feedback information, making the implementation of UWB systems more practical. The proposed is shown in Fig. 6.1 and can be described as follows:

Let us assume we have N energy normalized channel realizations $C = [h_1(t), h_2(t), ..., h_N(t)]$. A set $C_H = [H_0(t), H_1(t), ..., H_x(t), ..., H_m(t)]$, representing the CCHs is selected out of the set C of channels as a result of learning process carried out by means of clustering algorithm. It will be explained how to select CCHs and then describe the distance criteria used in clusters as Sections 4.2. Next, the detailed description of the clustering algorithm is provided as Section 4.4.1.

In detail, every coherence time, the receiver estimate the channel and select the CCH within the set C_H which is the most similar to the estimated channel according to the selected similarity or dissimilarity criterion. The receiver sends then back the corresponding CCH code to the transmitter. Finally, the transmitter decodes the received CCH code and prefilters the data to be transmitted using the time-reversed conjugate of the corresponding CCH $(h_{H_0}(t))$ as follows:

1. The transmitter modulate the bit stream $b(i)$ according to the selected modulation scheme as explained in Equations 3.19 and 3.20 and then applies the CCH (h_{H_0}) to transmitted signals $(S(t))$ as follows:

$$T(t) = S(t) \otimes h^*_{H_0}(-t). \tag{6.1}$$

2. In the receiver, an estimation of the bit is obtained by using a matched filter as follows:

$$b_i = \int_{\tau}^{\tau+T_i} (S(t) \otimes h^*_{H_0}(-t) \otimes h(t) + n(t)) \times (M_i(t) \otimes h^*_{H_0}(-t) \otimes h(t))dt \tag{6.2}$$

where $(M_i(t) \otimes h^*_H(-t) \otimes h(t))$ is the mask of the matched filter that guarantees optimal bit decision. $h(t)$ and $n(t)$ are channel and Additive White Gaussian Noise, respectively. $M_i(t)$ obtain as follows:

$$M_i(t) = \sum_i \sum_{n=0}^{N_s-1} (w(t - (iN_s + n)T_f - C_{(iN_s+n)}T_c)) \tag{6.3}$$

The proposed algorithm foresees thus two main steps:

1. Learning step - Generate or collect CIRs and divide them into clusters, and select CCHs of subgroups using a machine learning process (to be performed once);

2. Selection step - Apply the best matching CCH to channel as a pre-filter for transmitting pulses of UWB system (to be performed once every coherence time).

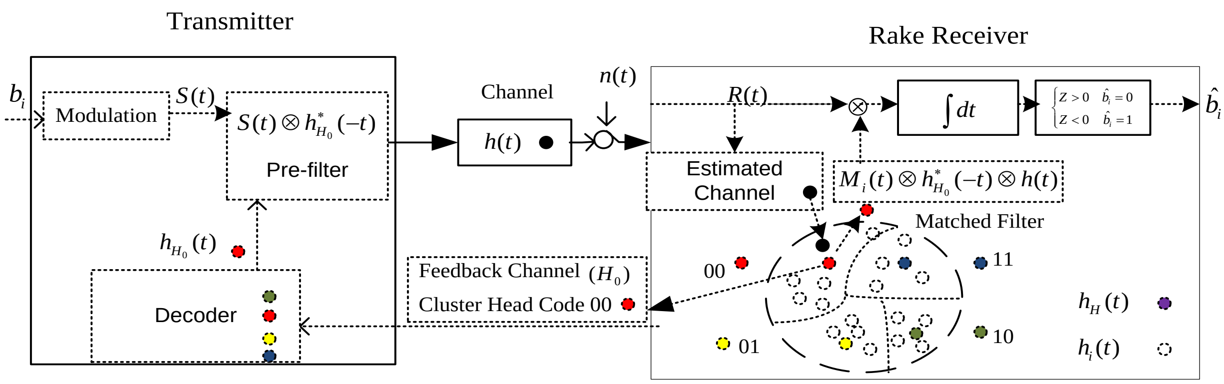


Fig. 6.1 Block diagram of a UWB system which uses the proposed algorithm with offline learning approach.

6.2 Advantage of Proposed Algorithm

The proposed digital feedback provides four main Advantage as following:

1. Lower amount of information transfer from the feedback channel because it requires sending backs only the code of the selected CCH and not the whole channel information as done in traditional time reversal
2. It significantly reduces the dedicated bandwidth required for feedback,
3. It considerably improves the speed of transceivers,
4. It is robust to noise in the feedback channel since the few bytes required to send the codes can be heavily error protected.

We will demonstrate that in the proposed method a maximum of five bits information is required to be sent back.

6.3 Possible Scenarios

One of the important factor to evaluate the proposed algorithm is channel or simulation environment. Two scenarios can be considered for obtaining the channel realizations and for discovering their CCHs in the learning step: unknown vs. known environment [55].

1. *i)* **Unkown environment:** in this scenario, channel information is unavailable at startup. A set-up phase is needed in order to collect the CIRs. During the set-up phase, the transmitter sends several token pulses and the receiver collects the corresponding CIRs. After this, the receiver partitions the captured CIRs and selects the CCHs using a clustering algorithm. Subsequently, it sends back CCHs and their codes to the transmitter [49].

2. *i)* **Kown environment:** in this scenario, the channel characteristics are known, allowing the use of IEEE 802.15.3a [56] as explained in Section 3.1.3 or IEEE 802.15.4a [56] channel models. These channel models can be tuned so to match channel characteristics in particular indoor and outdoor environments, and provide reliable benchmarks to evaluate and discover the CCHs. Hence, the CIRs are generated from the channel model and then they are clustered. In this scenario there is no need for the set-up phase, enabling the systems to start operating immediately.

The focus of this chapter is on a known environment, in which channel information was already measured and modeled for specific indoor and outdoor areas. Channel Impulse Responses (CIRs) can be then generated from a standard channel model like IEEE 802.15.3.a [56].

6.4 Similarity metrics

Both learning and selection steps rely on the definition of a similarity metric for identifying CCHs (learning step) and selecting the best matching CCH (selection step). In The this section reviews some suitable similarity metrics from the literature, while Section 6.4.1 proposes a new metric specifically defined for the proposed algorithm. Finally, possible choices for the clustering algorithms are reviewed in Section 4.2.

Similarity metrics play a key role in finding the groups of data points which are in the same cluster. A good similarity metric should lead to the distance between data points in the same cluster to be significantly less than the distance with data points in other clusters.
Before reviewing possible metrics proposed in the literature and describing in detail those selected in this work, it is wise to clarify what function should be used as input to the similarity metric.
As explained in Section 3.1.6, in TRP the signal after the channel (before considering noise) is the convolution of the pulse $w(t)$ with the time-reversed conjugate of the channel $h_j^*(-t)$ and the channel $h_j(t)$ as follows:

$$\Psi_{jj}(t) = w(t) \otimes R_{h_j h_j}^{auto}(t) = w(t) \otimes h_j^*(-t) \otimes h_j(t) = w(t) \otimes h_j(t) \odot h_j(t) \tag{6.4}$$

Let us consider a generic energy normalized channel realization $h(t)$ generated according to the IEEE 802.15.3a channel model from Equation 3.5 in 3.1.3:

$$h(t) = \frac{X}{\sqrt{E}} \sum_{l=0}^{L} \sum_{z=0}^{K(l)} \alpha_{z,l} \delta(t - T_l - \tau_{z,l}) \tag{6.5}$$

where the total energy of the channel is:

$$E = X^2 \sum_{l=0}^{L} \sum_{z=0}^{K(l)} |\alpha_{z,l}|^2 \tag{6.6}$$

Replacing the value of $h(t)$ from Equation 6.5 in Equation 6.4 we have:

$$\Psi_{jj} = w(t) \otimes \frac{X}{\sqrt{E}} \sum_{l=0}^{L} \sum_{z=0}^{K(l)} \alpha_{z,l} \delta(t - T_l - \tau_{z,l}) \odot \frac{X}{\sqrt{E}} \sum_{l=0}^{L} \sum_{z=0}^{K(l)} \alpha_{z,l} \delta(t - T_l - \tau_{z,l}) \tag{6.7}$$

This procedure provides a temporal energy focusing in the signal.

Let us now consider the proposed algorithm, and let us indicate with $h_i(t)$ be the selected CCH. The signal at the output of the channel is in this case given by:

$$\Psi_{ij}(t) = w(t) \otimes R^{cross}_{h_i h_j}(t) = w(t) \otimes h_i^*(-t) \otimes h_j(t) = w(t) \otimes h_i(t) \odot h_j(t) \tag{6.8}$$

similarly to the previous case one obtains:

$$\Psi_{ij} = w(t) \otimes \frac{X}{\sqrt{E}} \sum_{l=0}^{L} \sum_{z=0}^{K(l)} \alpha_{z,l} \delta(t - T_l - \tau_{z,l}) \odot \frac{X}{\sqrt{E}} \sum_{l=0}^{L} \sum_{z=0}^{K(l)} \alpha_{z,l} \delta(t - T_l - \tau_{z,l}) \tag{6.9}$$

Our goal is to find an optimal similarity function ensuring that the selected CCH is the best one. Ideally, the $h_i(t)$ with highest similarity to $h_j(t)$ should be the one that, when substituted for the $h_j(t)$, exhibits the most similar behavior. In other words, the signal $\Psi_{ij}(t)$ in Equation 6.8 between the candidate channel $h_i(t)$ to $h_j(t)$ should be the best match to the signal $\Psi_{jj}(t)$ corresponding to channel $h_j(t)$ in Equation 6.4.

Comparing $\Psi_{ij}(t)$ with $\Psi_{jj}(t)$ gives the possibility to identify the candidate channel that can produce maximum similarity. The goal is thus to determine $h_i(t)$ such that the best possible matching between $\Psi_{ij}(t)$ and $\Psi_{jj}(t)$ is obtained. Hence, different similarity criteria between $\Psi_{ij}(t)$ and $\Psi_{jj}(t)$ can be applied as input to the clustering algorithm in order to discover CCHs.

After an exhaustive analysis based on simulations, three different similarity metrics were selected based on Manhattan distance, Euclidean distance, and Cosine similarity, respectively.

These are described in the following, before introducing a new metric in Section 6.4.1.

The Minkowski distance between all pairs of channels is calculated as:

$$S(h_i(t), h_j(t)) = -\sqrt[p]{|\Psi_{ij}(t) - \Psi_{jj}(t)|^p} \tag{6.10}$$

where $p = 1$ and $p = 2$ gives the Manhattan (city-block) distance and Euclidean distance, respectively.

Cosine similarity is computed as follows:

$$S(h_i(t), h_j(t)) = \frac{\Psi_{ij}(t)\Psi_{jj}(t)}{|\Psi_{ij}(t)||\Psi_{jj}(t)|} \tag{6.11}$$

6.4.1 Matching similarity metric

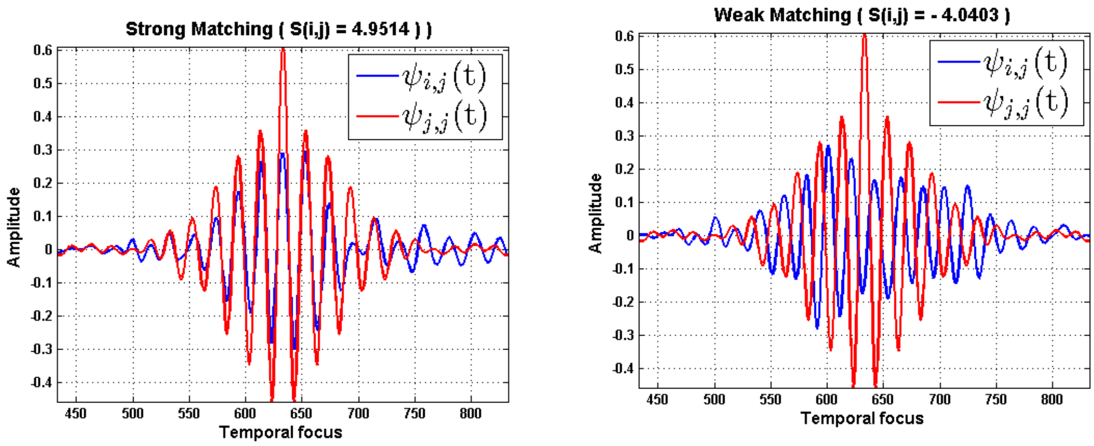
Moving from the observations on the definition of the signals to be matched presented in Section 6.4, in this section a novel metric is proposed, referred to as matching similarity metric.

The matching similarity metric computes the similarity of two vectors on the basis of the inner product of the signals defined in Equations 6.4 and 6.8:

$$S(h_i(t), h_j(t)) = \Psi_{ij}(t)\Psi_{jj}(t) \tag{6.12}$$

The matching similarity criterion compares channel $h_j(t)$ with all the $N-1$ candidate CCHs. The similarity between $h_i(t)$ and $h_j(t)$ shows the suitability of channel $h_i(t)$ to be as CCH for channel $h_j(t)$. The relation $S(h_i(t), h_j(t)) > S(h_x(t), h_j(t))$ denotes that channel i is a better representative of channel j than channel x.

The similarity function plays a significant role in achieving high performance for cluster-



(a) Strong Matching.

(b) Weak Matching.

Fig. 6.2 Comparison transmitted signal with TR and TR with two CCH candidates in optimal case(a) and non optimal case(b).

ing algorithms. We show in Section 6.6.1 that our proposed similarity function provides improvement over the other three mentioned distances.

A high similarity denotes that the substituted channel $h_i(t)$ as $\Psi_{ij}(t)$ leads to a focusing of the energy in the transmitted single pulse similar to the TR case using $\Psi_{jj}(t)$. In Equation 6.12 the matching between $\Psi_{ij}(t)$ and $\Psi_{jj}(t)$ is computed. We show in Figure 6.2 two examples of $\Psi_{ij}(t)$ and $\Psi_{jj}(t)$ along with their computed matching values. As can be seen in Figure 6.2a, a strong matching is observed in the case where the two waveforms are similar, i.e., channel $h_i(t)$ is a suitable candidate to be the cluster head of channel $h_j(t)$, whereas a weak matching is observed when the waveforms are dissimilar (Figure 6.2b).

6.5 Clustering

In this Section, the clustering algorithms and similarity metrics reviewed and introduced in Section 6.4 are evaluated in order to determine the best combination to be used in the system performance evaluation that will be carried out in Section 6.6.2. Next, we pick up the VSH, K log(K) EM and AP clustering algorithms described in Section 4.4.1 are compared using the metric selected in Section 6.6.1, leading to the selection of the best clustering algorithm. The random selection of the initial set of CCH also strong impact on performance. For this reason both K-log (K) EM and VSH are typically executed I_t times and the best partition is selected according to the network similarity performance indicator [36], [37]. AP algorithm can be carried out as Algorithm 2 to discover CCHs.

Algorithm 2 Discover CCHs using AP algorithm in learning step.

Input: Realize N channels from IEEE802.15.3a channel model [**?**]. $h_k(t)$, $k = \{1, \dots, N\}$
Output: Cluster labels of each channels $\Phi(h_j(t))$
Initialization:
$S(h_j(t), h_j(t))= \left\{ \begin{array}{ll} \text{use equations (6.10), (6.11) or (6.12)}, & \forall i, j \in k, i \neq j, \\ p, & \forall i, j \in k, i = j, \end{array} \right\}$
$A(h_i(t), h_j(t)) \leftarrow 0, R(h_i(t), h_j(t)) \leftarrow 0, \forall i, j \in k$
while $T \leq Threshold$ **do**
  updating responsibility and availability till converge
  $R(h_i(t), h_j(t)) \leftarrow S(h_i(t), h_j(t)) - \max_{j' \neq j} \{A(h_i(t), h_{j'}(t)) + S(h_i(t), h_{j'}(t))\}, \forall i, j \in k$
  $R^{new} = (\rho_d) R^{new} + (1 - \rho_d) R^{old}$
  $A(h_i(t), h_j(t)) \leftarrow \left\{ \begin{array}{ll} \min \left[0; R(h_j(t), h_j(t)) + \sum_{i' \notin (i,j)} \max \left[0; R(h_{i'}(t), h_j(t))\right]\right], & \text{for } i \neq j, \forall i, j \in k, \\ \sum_{i' \notin j} \max[0, R(h_j(t), h_j(t))], & \text{for } i = j, \forall i, j \in k, \end{array} \right\}$
  $A^{new} = (\rho_d) A^{new} + (1 - \rho_d) A^{old}$
end while
Calculate CCH's $h_j(t) \leftarrow [A(h_j(t), h_j(t)) + R(h_j(t), h_j(t))] > 0$,
Return selected labels $(\Phi(h_j(t)))$ for each channel based on maximum similarity from column of similarity

6.6 Experimentation on Channel Cluster Heads (CCHs) Pre-coding

This Section focuses on the analysis of the selected combination of clustering algorithm and similarity metric. The performance evaluation is based on Bit Error Rate (BER) as a function of received signal to noise ratio (SNR). With respect to the systems parameters defined previously the values in Table 6.1 will be adopted.

The analysis is organized in three steps. In Section 6.6.1 the metrics selected at the end of Section 6.6.2 and the novel matching metric introduced in Section 6.4.1 are compared using the AP algorithm as the reference clustering algorithm, and the metric leading to best performance is identified. In particular, note that the frame time T_f is set to value larger than the channel delay spread for all channel configurations considered in this work, thus ensuring that Inter-Sample Interference (ISI) is negligible in all cases. It is further more assumed that transmitter and receiver are perfectly synchronized.

6.6.1 Selection of similarity metric

In this section, the performance of the similarity metrics introduced in Sections 6.4 and 6.4.1 is evaluated by combining them with the AP algorithm.

Figure 6.3 shows the BER as a function of SNR for the matching metric (Equation (6.12)) compared with the Minkowski metric (Equation (6.10)) in the case $p = 2$ (Euclidean) and $p = 1$ (city block), and the Cosine metric (Equation (6.11)).

Affinity propagation takes as its input a collection of four different metric similarities between channels. The pair wise similarity between the channels is used for clustering channels with AP algorithms. Finally CCHs are discovered for different metrics and applied to the transmitted signals as pre-filter.

Figure 6.3 shows that the matching metric consistently leads to better results when compared to the other metrics and was thus selected as the reference metric in the proposed system. Note that results presented in Figure 6.3 were obtained by adopting the CM2 channel configuration, but similar results were observed for all configurations.

We can see from the figure that the BER average is lowest for the matching metric suggesting its appropriateness for cluster channel clustering algorithm. Thus, because of the optimal performance, in the following sections, the matching metric is used for simulations.

Table 6.1 Parameters settings used in performance evaluation.

Parameter	N_s	T_c	α	T_s	T_f
Value	1	2ns	0.4ns	$167ps$	$10ms$

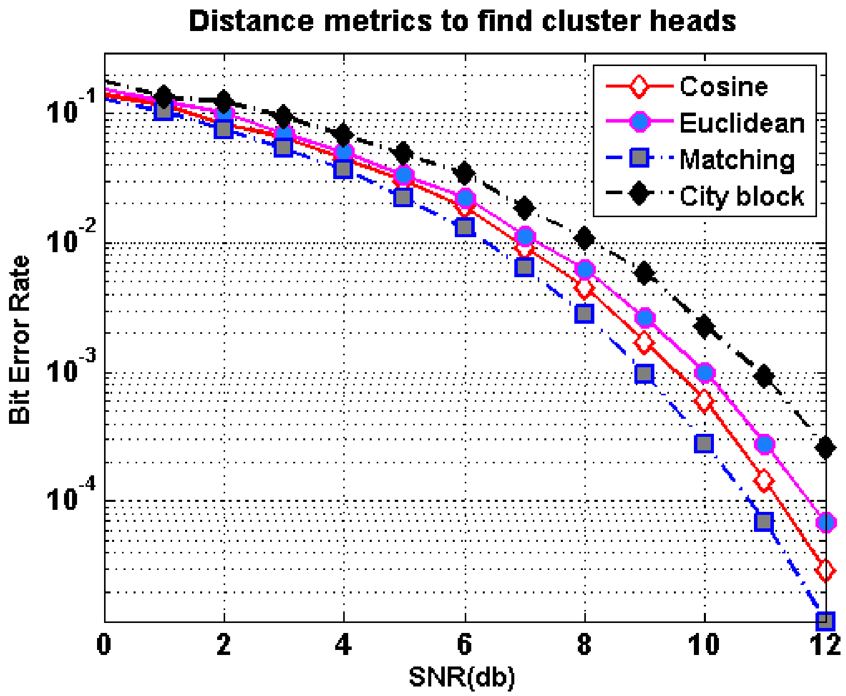


Fig. 6.3 Performance of the UWB system with different similarity metrics.

6.6.2 Selection of clustering algorithm

In this section the performance of the K-log(K) EM, VSH and AP clustering algorithms presented in Section 6.6.2 is compared in combination with the matching metric selected in Section 6.6.1. Using the settings presented in Table 6.2.

Figure 6.4 shows the BER as a function of the number of CIRs used for training in two cases; (a) feedback system using 4 bits (16 CCHs) and (b) 6 bits (64 CCHs).

Results demonstrate that when the number of CIRs used during the learning step is low the performance of the three algorithms is similar, but when the number of CIRs is large, AP and VSH perform better than EM. AP is however characterized by significantly faster learning process [37] suggesting its adoption as the reference clustering algorithm in the method proposed in Section 6.6.2.

6.6.3 AP clustering algorithm setting

This section investigates how to configure the AP algorithm in order to achieve a satisfactory trade off between the three following competing requirements:

1. Maximize system performance by minimizing BER;
2. Minimize the amount of information on the feedback channel by minimizing the of CCHs;
3. Ensure the convergence of the clustering algorithm.

The analysis will focus on the impact of two parameters:

- The set of preference values that, as stated in [36], is the key parameter to control the behaviour of the AP algorithm.
- The number of CIRs used in the training phase.

Table 6.2 Parameters settings of clustering algorithm.

Parameter	ρ_d	$I_{t_{max}}$	$I_{t_{min}}$	I_t
AP	0.9	2000	200	1
VSH	-	-	-	100
EM	-	-	-	10000

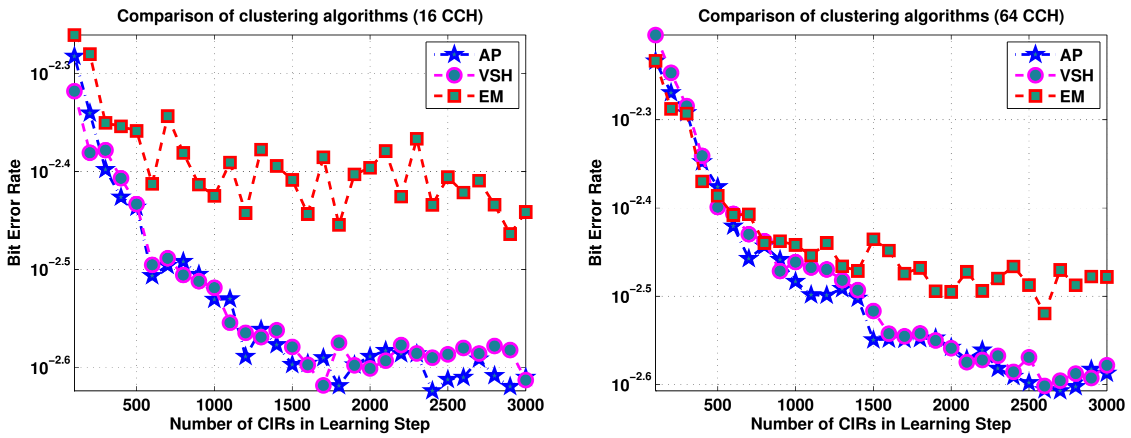


Fig. 6.4 Performance UWB system over clustering algorithms in 8 dB SNR.

It will determine and answer the mentioned issues in the following. At the core of our proposed approach lies the AP clustering algorithm. Hence, the convergence criteria and the parameter setting of AP are critical. Thus, we attempt to set the parameters in a way to obtain the optimal setting till improve the BER and decrease the number of cluster heads. As also discussed in previous sections, the UWB Transceivers based on CCHs require the number of cluster heads in order to properly function. This is a question: how many CCHs should be chosen while prior knowledge regarding the cluster quantity is unavailable.

The main parameter of AP is the preference value that controls the number of the cluster heads. Preference denotes the similarity of a channel $h_j(t)$ with itself, i.e., $S(h_j(t), h_j(t))$. The value of the preference influences the final number of cluster heads and consequently, it also affects BER. We start a simulation to find the optimal number of clusters without any prior knowledge.
Figure 6.5 presents the number of CCHs (a) and the BER (b) as a function of preference value. Figure 6.5 (a) shows that high preference values lead to a large numbers of CCHs.
In parallel Figure 6.5 (b) shows that the minimum BER is observed for a preference value of -30. The number of CCHs is also relatively low for this value which means that a low number of bits are needed to be sent back.
It is interesting to note the authors in [36] proposed the median values of similarities to be used for preference. However, results in Figure 6.5 show strategy not to lead to desirable results. For a median value preference of -0.01, AP found 208 CCHs with an average BER of 0.003016, higher than BER values observed for preference values lower than -10. Indeed for negative preference values, as the magnitude increases, the number of CCH and BER both decrease indicating that the preference value should be set to large negative value. A lower bound to this number is determined by the fact that when the value of preference becomes to small the algorithm does not converge.

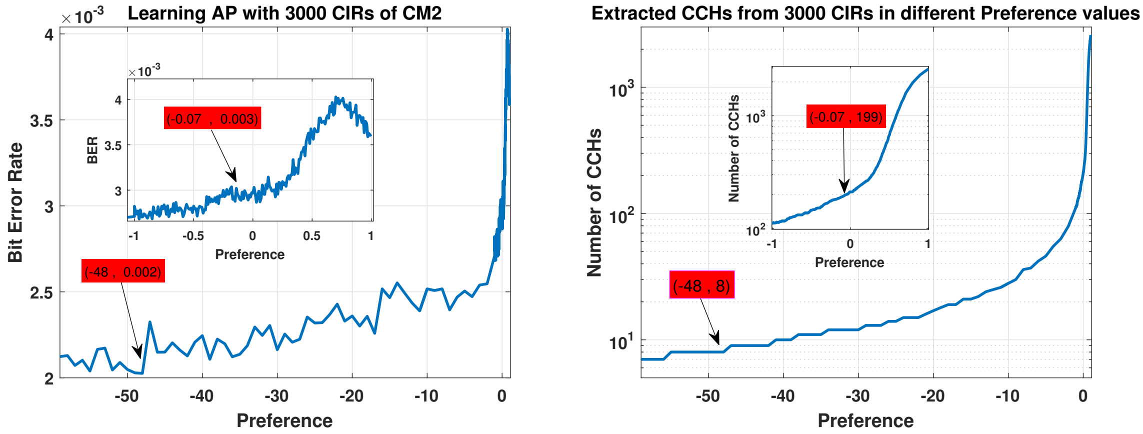


Fig. 6.5 The effect of (a) preference over the number of channels and (b) preference effect over the BER.

Moving to the number of CIRs in the learning step, Figure 6.6 presents the BER as a function of number of CIR shows that although the BER decreases as the number of CIR increases the performance gain become negligible above 10000 CIR.

In parallel, Table 6.3 shows that increasing the number of CIRs in the learning process leads to an increase, albeit slow, of the number of CCH. increasing CIRs in the learning process does not substantially increase the number of CCH as shown in Table 6.3. Thus, increasing amounts of CIR in the learning step, do not provides any noticeable gains in terms of BER performance. We can conclude that we do not need more CIRs in the learning process.

Result suggest does that the number of CIRs in the order of 10000 provide a good trade off between BER and number of CCHs leading to 3 or 4 bits to be sent back on the feedback channel, depending on the channel model and exact number of the CIRs.

We used the bisection method for AP [36] in order to adjust the preference values so to obtain CCH.

6.7 Performance Evaluation

All the transmitted signals in the proposed systems have the same power. The proposed TRP based on CCHs was compared with TRP as well as with ideal rake receiver (no TRP) and TRP with Random Channel (RC). All schemes are evaluated in terms of average BER as a function of SNR.

In the simulation of the case without time reversal a method of combining all resolvable paths in the detection of pulses, which is called all-Rake receiver, is utilized. In the TRP case,

Table 6.3 Amount of CCHs that are found in the learning step of the CRs (from 500 to 20000).

Number of CIRs	500	2000	3000	5000	6000	7000	10000	20000
CM1	2	3	4	5	6	7	8	10
CM2	2	6	8	10	11	12	17	27
CM3	2	6	8	12	13	14	17	37
CM4	2	6	8	12	14	15	18	40

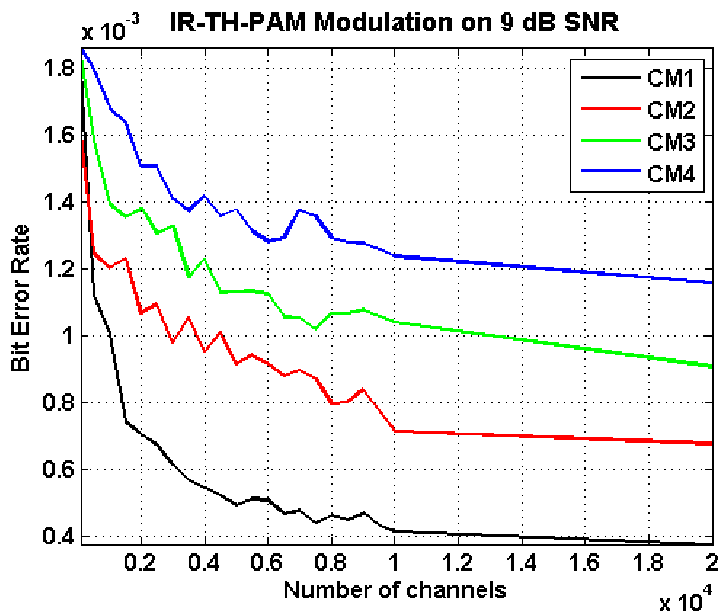


Fig. 6.6 CCHs extracted via the learning step from the range of 500 to 20000 CIRs. After the extraction of the CCHs are applied in TRP UWB systems as a pre-filter instead of the actual channel and the BER performances are evaluated.

the feedback channel is considered without noise for simulation. Thus the ideal channel, which does not have any disruptions or changes, is applied in the pre-filtering of transmitted pulses of TRP case. However, in the real world such a channel feedback does not exist and the real-world results of the TRP case will be inferior to our result. In the other words, in the results, ideal conditions are considered. While, the proposed pre-filter based on CCHs requires to send back CCHs codes.

The number of CCHs depends on the IEEE802.10.3.a environments. This means that the number of CCHs for CM1, CM2, CM3 and CM4 are 4, 8, 8 and 8, respectively. These are extracted from 3000 CIRs during the learning step with AP clustering algorithm and matching similarity metric. The proposed algorithm requires 3 ($2^3 = 8$) bits to be sent back from the feedback channel and reaches to the optimal BER performance of the CCHs as shown in Figure 6.7. With regards to only a three bit flow in the feedback channel, the proposed TR based on CCHs was able to send back a more predictable and protective channel coding making it more applicable in the real world.

In the testing step, matching similarity of channel estimations and CCHs is calculated and then the CCH label with the maximum similarity is returned back to the transmitter. The decoded label channel is then applied to the transmitting pulses.

Figure 6.7a shows the performance of different structures in line of sight (LOS) environment. It is shown that the CCHs were able to have the same performance of the TRP method in CM1. Figures 6.7b, 6.7c, 6.7d are the results over CM2, CM2 and CM3, respectively that illustrate the performance of proposed algorithm in terms of BER with respect to other structures.

The most important advantage of the proposed method is that it significantly reduces feedback load information with respect to the pure TR filters and one-bit quantization. For example, for pure TRP there is a need for the feedback of $\{K(l)+L+1\} * b$ bits. The sent back information in the feedback channel for the TRP over $167ps$ sampling times of CM1, CM2,

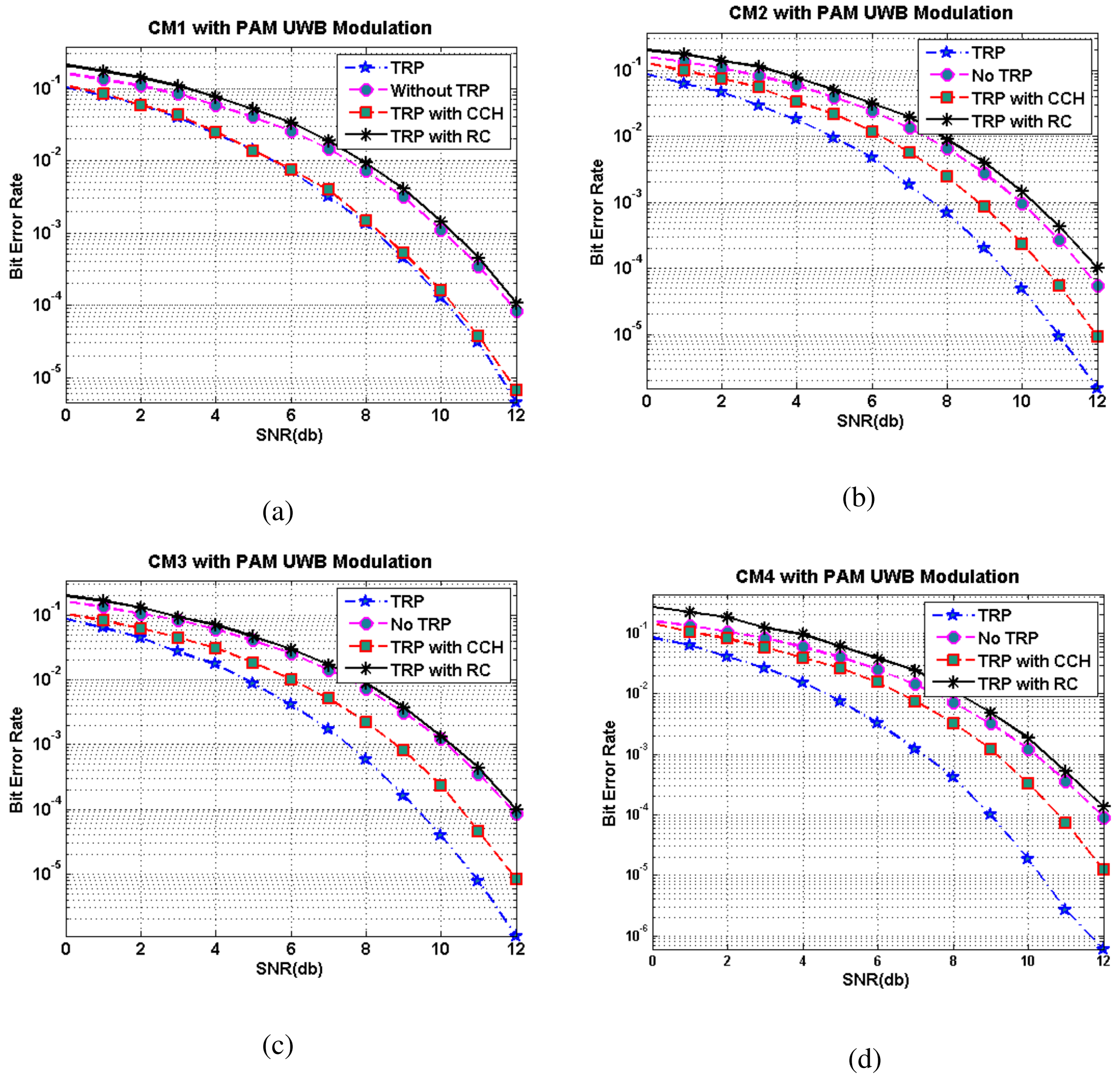


Fig. 6.7 Bit error rate performance for CM1 (a), CM2 (b), CM3 (c), and CM4 (d).

CM3 and CM4, which are the number of bytes allocated for each channel in Matlab software, are 7608, 8136, 85318 and 140313 bytes, respectively. Particularly the one-bit TRP methods need $\{K(l)+L+1\}*1$ bits to feedback. This amount is necessary for the pre-filter data on the transmitter side and the amount of data to be sent back depends on the channel model. Hence, they require a considerable amount of time to feedback.

6.8 Performance Evaluation in ISI Condition

The robust channels similarity is proposed to find CCHs by using matching similarity in previous section. It simplified description of channels in the clustering process. Hence, robust discriminations of channels are significant to achieve high performance for UWB systems.

We define a new similarity criterion for high data rate. While the data rate is low, there are not any overlaps between sequence of pulses after apply the CCHs as per-filer . Therefore, we defined similarity criteria that adopt the nearest behaviour to the auto-correlation of channel which is presented in Section 6.4.

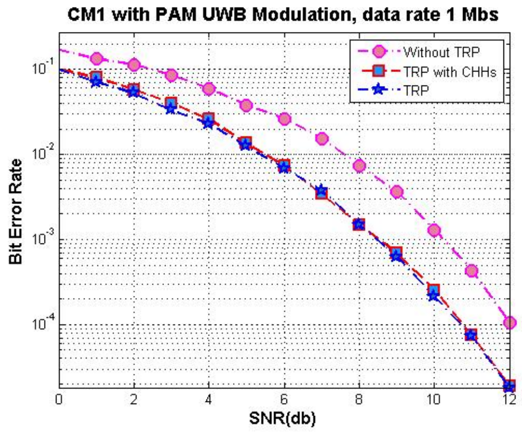
However, in high data rate, we have overlap betwwen the signals as Fig. 3.4 and then substitute channels should respect to its own channels with maximum concentration energy in centre.
We propose coherent energy factor that provide metric to evaluate the concentration of energy in the centre which is ratio of energy of main part of vector respect to all energy as follow:

$$CF = \frac{\int_{-nT_c}^{nT_c} \Psi_{ij}(t)dt}{\int \Psi_{ij}(t)dt} \tag{6.13}$$

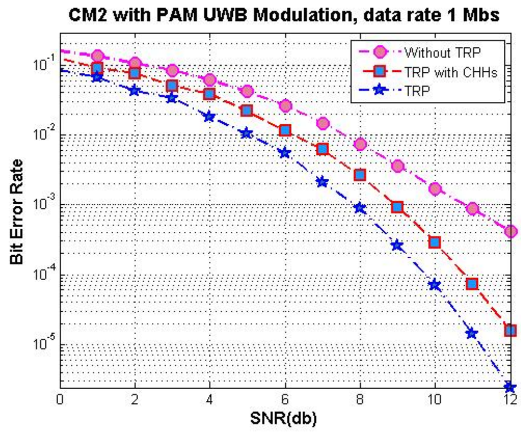
CF can be written interact of all channel is via computing energy of cross correlation of pairs channels. Our similarity criteria provide channels which can be more concentrate cross correlation in the centre that is useful for high data rate as follow:

$$S(h_i(t), h_j(t)) = CF - 1 \tag{6.14}$$

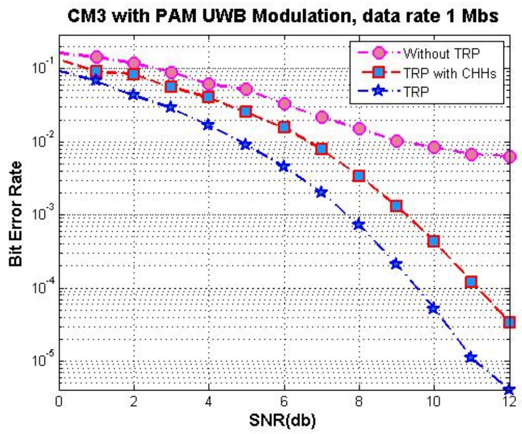
If we select the CCH with CF criteria, energy concentrate in the center. Thus, the similarity criteria selection should be improve the BER performance in high data rate in UWB systems.



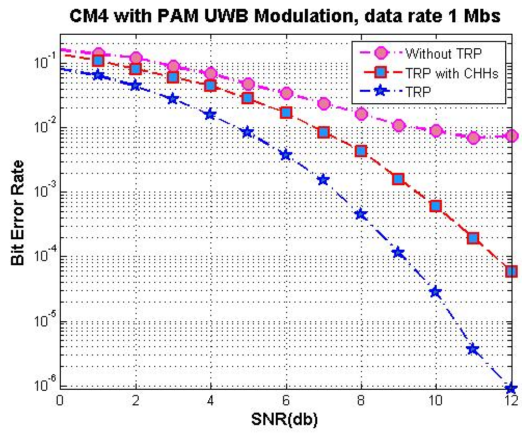
(a)



(b)



(c)



(d)

Fig. 6.8 Bit error rate performance at 1 Mbs for CM1 (a), CM2 (b), CM3 (c), and CM4 (d).

We introduced the CF criterion for similarity in high data rate. We also discover the CCHs with AP clustering algorithm. The result illustrates acceptable performance of new certain in high data rate. In the Fig. 6.8 show the performance of probability error for 1 Mbs. Fig. 6.9 demonstrate performance of different channel model in 25 Mbs and Fig. 6.10 is simulated in 50 Mbs.

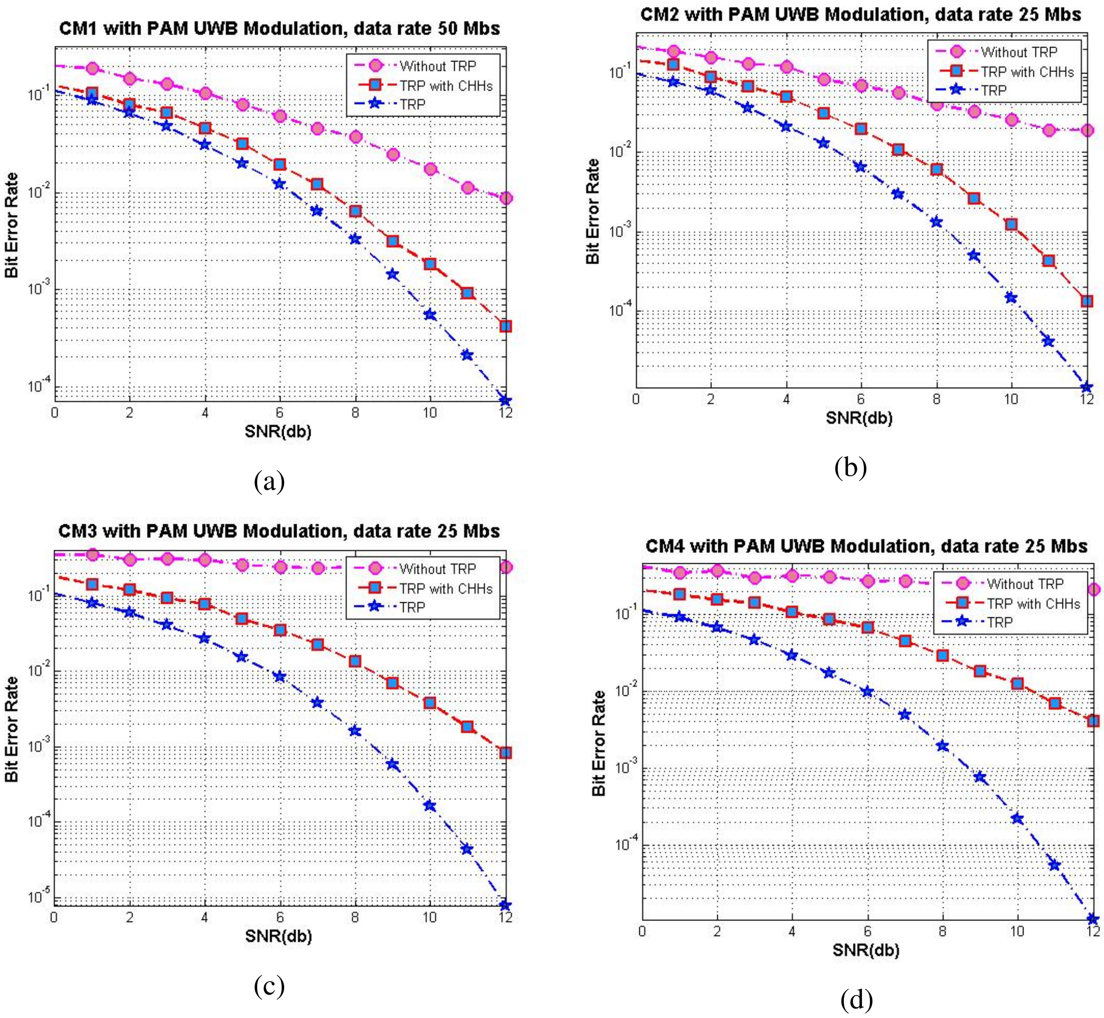


Fig. 6.9 Bit error rate performance at 25 Mbs for CM1 (a), CM2 (b), CM3 (c), and CM4 (d).

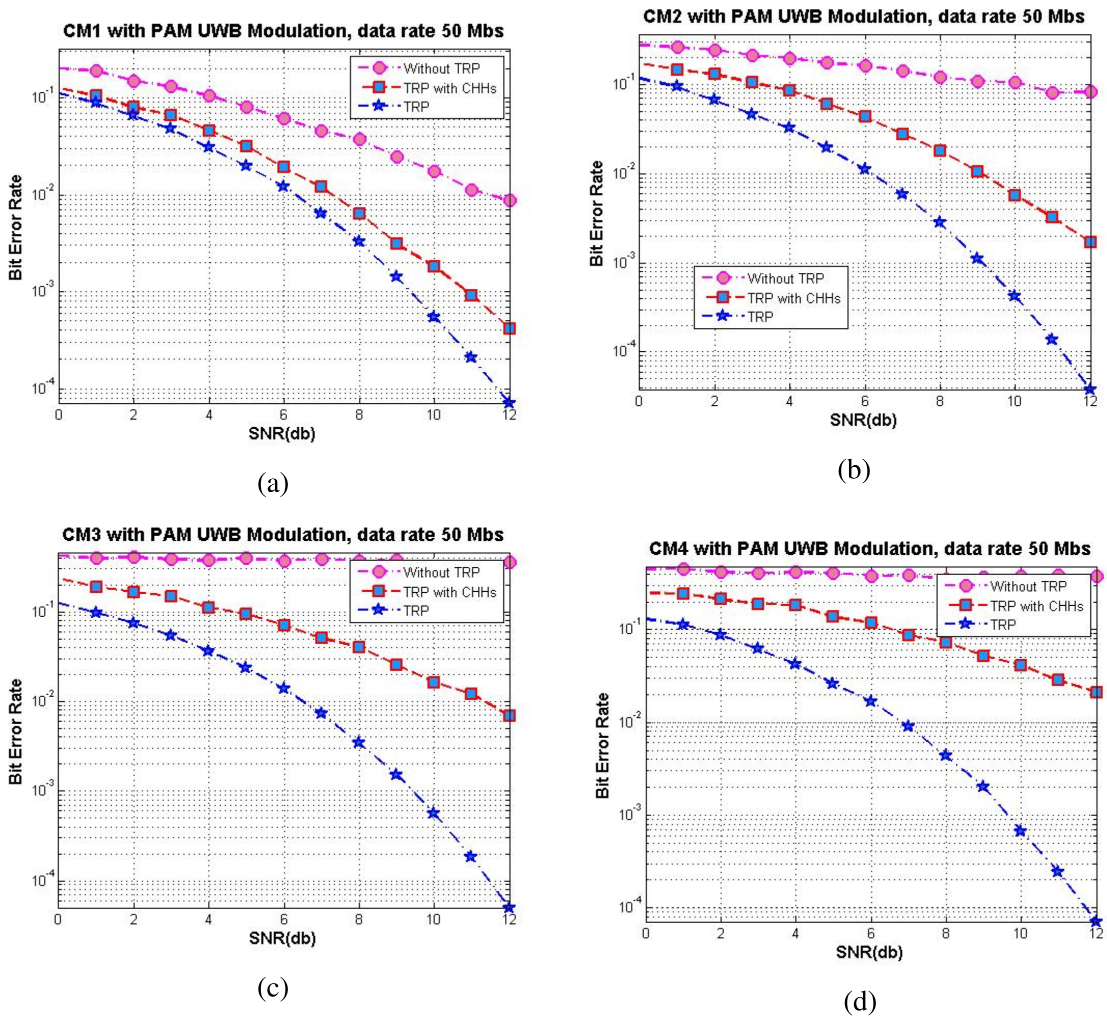


Fig. 6.10 Bit error rate performance at 50 Mbs for CM1 (a), CM2 (b), CM3 (c), and CM4 (d).

Chapter 7

Conclusion and Hint for the Future Research

In this work, we proposed a new detection algorithm based on energy features. Our target was to use energy profiles to discriminate different technologies. Since energy has a special behavior in UWB, the energy of signals in short time window was selected as a feature. We evaluated our algorithm for the detection of signals belonging to four different classes: UWB, Bluetooth, Bluetooth plus UWB and noise.
Based on the results, we provided several conclusions. We showed that the SVM with a Gaussian RBF kernel has demonstrated to yield the most accurate results for more than 10 dB SNR. For 0 dB Naïve Bayes has best accuracy and for less than 0 dB all of kernels have same performance. mRMR method had good performance to sort features. We plan to investigate different path loss in future work.

In this work also a new feedback design for the UWB systems was proposed. The proposed scheme relies on a learning process to determine the set of clusters and to select CCHs. For each channel estimation procedure the index of the best matching with CCH is sent back instead of the actual channel estimation which significantly reduces the amount of information to be sent to the feedback channel. We proposed matching criteria for the similarity of the channel and finally we discussed details of the design using different clustering algorithms. The simulation result over UWB systems in IEEE 802.15.3a multipath channel model shows that channel cluster heads are effective for the UWB-IR systems in order to achieve suitable BER performance. Also, in the proposed design the complexity of the transceiver is significantly reduced. CCHs have also optimized the performance of the UWB transceivers. This method requires a small amount of bits to be sent back in the feedback transceiver thus it needs a low bandwidth feedback, resulting in robustness with respect to noise. The only downside to this model is a slight increase in the bit error when compared to ideal time reversal pre-filtering.

In the case of high data rates, the transmitting pulses interfere with each other as the interference of pulses increase in channel with high delay spread. The proposed model will not be suitable without further operation. Therefore, it needs a new definition of similarity criteria and additional operations to reduce the Multi User Interference (MUI) and Inter Symbol Interference (ISI) for future work.

References

[1] Abbasi-Moghadam, D. and Vakili, V. T. (2012). A simo one-bit time reversal for uwb communication systems. *EURASIP Journal on Wireless Communications and Networking*, 2012(1):1–9.

[2] Abbasi-Moghadam, D. and Vakili, V. T. (2013). Effect of channel estimation error on time reversal uwb communication system. *Wireless personal communications*, 68(2):433–439.

[3] Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, pages 199–213. Springer.

[4] Banfield, J. D. and Raftery, A. E. (1993). Model-based gaussian and non-gaussian clustering. *Biometrics*, pages 803–821.

[5] Benedetto, M.-G. D., Boldrini, S., Martin, C. J. M., and Diaz, J. R. (2010). Automatic network recognition by feature extraction: a case study in the ism band. In *Cognitive Radio Oriented Wireless Networks & Communications (CROWNCOM), 2010 Proceedings of the Fifth International Conference on*, pages 1–5. IEEE.

[6] Bkassiny, M., Li, Y., and Jayaweera, S. K. (2013). A survey on machine-learning techniques in cognitive radios. *Communications Surveys & Tutorials, IEEE*, 15(3):1136–1159.

[7] Bluetooth (2003). *Bluetooth SIG. Assigned Numbers - Bluetooth Baseband*. https://www.bluetooth.org/foundry/assignnumb/document/baseband.

[8] Bluetooth (July , 2004). *In-Stat/MDR. Bluetooth 2004: Poised for the Mainstream*. http://www.instat.com/r/nrep/2004/IN0401211MI.htm.

[9] Blum, C. and Roli, A. (2003). Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys (CSUR)*, 35(3):268–308.

[10] Boldrini, S., Ferrante, G. C., and Benedetto, M.-G. D. (2011). Uwb network recognition based on impulsiveness of energy profiles. In *Ultra-Wideband (ICUWB), 2011 IEEE International Conference on*, pages 327–330. IEEE.

[11] Cassioli, D., Win, M. Z., Vatalaro, F., and Molisch, A. F. (2007). Low complexity rake receivers in ultra-wideband channels. *Wireless Communications, IEEE Transactions on*, 6(4):1265–1275.

[12] Chang, Y.-H., Tsai, S.-H., Yu, X., and Kuo, C. J. (2007). Ultrawideband transceiver design using channel phase precoding. *Signal Processing, IEEE Transactions on*, 55(7):3807–3822.

[13] Chuang, J. C. (1987). The effects of time delay spread on portable radio communications channels with digital modulation. *Selected Areas in Communications, IEEE Journal on*, 5(5):879–889.

[14] Clancy, C., Hecker, J., Stuntebeck, E., and Shea, T. O. (2007). Applications of machine learning to cognitive radio networks. *Wireless Communications, IEEE*, 14(4):47–52.

[15] Cristianini, N. and Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.

[16] Danev, B., Luecken, H., Capkun, S., and El Defrawy, K. (2010). Attacks on physical-layer identification. In *Proceedings of the third ACM conference on Wireless network security*, pages 89–98. ACM.

[17] Danev, B., Zanetti, D., and Capkun, S. (2012). On physical-layer identification of wireless devices. *ACM Computing Surveys (CSUR)*, 45(1):6.

[18] Dasgupta, S. and Schulman, L. (2007). A probabilistic analysis of em for mixtures of separated, spherical gaussians. *The Journal of Machine Learning Research*, 8:203–226.

[19] Dasgupta, S. and Schulman, L. J. (2000). A two-round variant of em for gaussian mixtures. In *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, pages 152–159. Morgan Kaufmann Publishers Inc.

[20] Davies, D. L. and Bouldin, D. W. (1979). A cluster separation measure. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2):224–227.

[21] De Nardis, L., Fiorina, J., Panaitopol, D., and Di Benedetto, M.-G. (2013). Combining uwb with time reversal for improved communication and positioning. *Telecommunication Systems*, 52(2):1145–1158.

[22] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38.

[23] Derode, A., Tourin, A., de Rosny, J., Tanter, M., Yon, S., and Fink, M. (2003). Taking advantage of multiple scattering to communicate with time-reversal antennas. *Physical Review Letters*, 90(1):014301.

[24] Derode, A., Tourin, A., and Fink, M. (1999). Ultrasonic pulse compression with one-bit time reversal through multiple scattering. *Journal of applied physics*, 85(9):6343–6352.

[25] Dezfooliyan, A. and Weiner, A. (2015). Spatio-temporal focusing of phase compensation and time reversal in ultra-wideband systems with limited rate feedback. *Communications, IET*.

[26] Dezfooliyan, A. and Weiner, A. M. (2013). Phase compensation communication technique against time reversal for ultra-wideband channels. *Communications, IET*, 7(12):1287–1295.

[27] Di Benedetto, M.-G. and Giancola, G. (2004). *Understanding ultra wide band radio fundamentals*. Prentice Hall PTR.

[28] Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

[29] Edelmann, G. F., Akal, T., Hodgkiss, W. S., Kim, S., Kuperman, W., Song, H. C., et al. (2002). An initial demonstration of underwater acoustic communication using time reversal. *Oceanic Engineering, IEEE Journal of*, 27(3):602–609.

[30] Emami, S., Hansen, J., Kim, A., Papanicolaou, G., Paulraj, A., Cheung, D., and Prettie, C. (2004). Predicted time reversal performance in wireless communications using channel measurements. *IEEE Communications Letters*.

[31] Esmailzadeh, R. and Nakagawa, M. (1993). Pre-rake diversity combination for direct sequence spread spectrum mobile communications systems. *IEICE Transactions on Communications*, 76(8):1008–1015.

[32] (FCC), F. C. C. (2002). Spectrum policy task force report. *ET Cocket 02-155*.

[33] Ferrante, G. C., Fiorina, J., and Di Benedetto, M.-G. (2012). Complexity reduction by combining time reversal and ir-uwb. In *WCNC*, pages 28–31.

[34] Fink, M. (1992). Time reversal of ultrasonic fields. i. basic principles. *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on*, 39(5):555–566.

[35] Foerster, J. R., Pendergrass, M., and Molisch, A. F. (2004). A channel model for ultrawideband indoor communication. *MERL (Mitsubishi Electric Research Laboratory) Report*.

[36] Frey, B. J. and Dueck, D. (2007). Clustering by passing messages between data points. *science*, 315(5814):972–976.

[37] Frey, B. J. and Dueck, D. (2008). Response to comment on" clustering by passing messages between data points". *science*, 319(5864):726–726.

[38] Gandetto, M. and Regazzoni, C. (2007). Spectrum sensing: A distributed approach for cognitive terminals. *Selected Areas in Communications, IEEE Journal on*, 25(3):546–557.

[39] Giancola, G. and Di Benedetto, M.-G. (2006). A novel approach for estimating multi-user interference in impulse radio uwb networks: The pulse collision model. *Signal processing*, 86(9):2185–2197.

[40] Goldsmith, A. (2005). *Wireless communications*. Cambridge university press.

[41] Guo, N., Sadler, B. M., and Qiu, R. C. (2007). Reduced-complexity uwb time-reversal techniques and experimental results. *Wireless Communications, IEEE Transactions on*, 6(12):4221–4226.

[42] Guyon, I., Gunn, S., Nikravesh, M., and Zadeh, L. A. (2008). *Feature extraction: foundations and applications*, volume 207. Springer.

[43] Hämäläinen, M., Saloranta, J., Mäkelä, J.-P., Oppermann, I., and Patana, T. (2003). Ultra wideband signal impact on ieee802. 11b and bluetooth performances. In *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*, volume 1, pages 280–284. IEEE.

[44] Han, J., Kamber, M., and Pei, J. (2011). *Data mining: concepts and techniques: concepts and techniques*. Elsevier.

[45] Hansen, P. and Mladenović, N. (1997). Variable neighborhood search for the p-median. *Location Science*, 5(4):207–226.

[46] He, A., Bae, K. K., Newman, T. R., Gaeddert, J., Kim, K., Menon, R., Morales-Tirado, L., Neel, J. J., Zhao, Y., Reed, J. H., et al. (2010). A survey of artificial intelligence for cognitive radios. *IEEE Transactions on Vehicular Technology*, 59(1-4):1578–1592.

[47] Hu, N. and Yao, Y.-D. (2012). Identification of legacy radios in a cognitive radio network using a radio frequency fingerprinting based method. In *Communications (ICC), 2012 IEEE International Conference on*, pages 1597–1602. IEEE.

[48] Imada, S. and Ohtsuki, T. (2005). Pre-rake diversity combining for uwb systems in ieee 802.15 uwb multipath channel. *IEICE transactions on communications*, 88(5):2193–2199.

[49] Kyritsi, P. and Papanicolaou, G. (2005). One-bit time reversal for wlan applications. In *Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005. IEEE 16th International Symposium on*, volume 1, pages 532–536. IEEE.

[50] Lampe, L. and Witrisal, K. (2010). Challenges and recent advances in ir-uwb system design. In *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pages 3288–3291. IEEE.

[51] Love, D. J., Heath Jr, R. W., Lau, V. K., Gesbert, D., Rao, B. D., and Andrews, M. (2008). An overview of limited feedback in wireless communication systems. *Selected Areas in Communications, IEEE Journal on*, 26(8):1341–1365.

[52] Ludwig, O., Nunes, U., Araújo, R., Schnitman, L., and Lepikson, H. A. (2009). Applications of information theory, genetic algorithms, and neural models to predict oil flow. *Communications in Nonlinear Science and Numerical Simulation*, 14(7):2870–2885.

[53] MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA.

[54] Mitchell, T. M. (1997). Machine learning. *Computer Science Series (McGraw-Hill, Burr Ridge, 1997) MATH*.

[55] Molisch, A. F. (2005). Ultrawideband propagation channels-theory, measurement, and modeling. *Vehicular Technology, IEEE Transactions on*, 54(5):1528–1545.

[56] Molisch, A. F., Balakrishnan, K., Cassioli, D., Chong, C.-C., Emami, S., Fort, A., Karedal, J., Kunisch, J., Schantz, H., Schuster, U., et al. (2004). Ieee 802.15. 4a channel model-final report. *IEEE P802*, 15(04):0662.

[57] Nguyen, H. T. (2007). On the performance of one bit time reversal for multi-user wireless communications. In *Wireless Communication Systems, 2007. ISWCS 2007. 4th International Symposium on*, pages 672–676. IEEE.

[58] Nguyen, H. T. (2008). Optimal one bit time reversal for uwb impulse radio in multi-user wireless communications. In *World Academy of Science, Engineering and Technology. Conference*, pages 298–304.

[59] Nguyen, H. T., Kovacs, I. Z., and Eggers, P. C. F. (2006). A time reversal transmission approach for multiuser uwb communications. *Antennas and Propagation, IEEE Transactions on*, 54(11):3216–3224.

[60] Peng, H., Long, F., and Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(8):1226–1238.

[61] Pilehvar, M. T., Jurgens, D., and Navigli, R. (2013). Align, disambiguate and walk: A unified approach for measuring semantic similarity. In *ACL (1)*, pages 1341–1351.

[62] Platt, J. et al. (1998). Sequential minimal optimization: A fast algorithm for training support vector machines.

[63] Qiu, R. C. (2006). Generalized time domain multipath channel and its application in ultra-wideband (uwb) wireless optimal receiver-part iii: system performance analysis. *Wireless Communications, IEEE Transactions on*, 5(10):2685–2695.

[64] Qiu, R. C., Zhang, J. Q., and Guo, N. (2006). Detection of physics-based ultra-wideband signals using generalized rake with multiuser detection (mud) and time-reversal mirror. *Selected Areas in Communications, IEEE Journal on*, 24(4):724–730.

[65] Saleh, A. A., Valenzuela, R., et al. (1987). A statistical model for indoor multipath propagation. *Selected Areas in Communications, IEEE Journal on*, 5(2):128–137.

[66] Scanlon, P., Kennedy, I. O., and Liu, Y. (2010). Feature extraction approaches to rf fingerprinting for device identification in femtocells. *Bell Labs Technical Journal*, 15(3):141–151.

[67] Selim, S. Z. and Alsultan, K. (1991). A simulated annealing algorithm for the clustering problem. *Pattern recognition*, 24(10):1003–1008.

[68] Selim, S. Z. and Ismail, M. A. (1984). K-means-type algorithms: a generalized convergence theorem and characterization of local optimality. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (1):81–87.

[69] Smola, A. J. and Schölkopf, B. (1998). *Learning with kernels*. Citeseer.

[70] Soleimani, H., Caso, G., De Nardis, L., and Di Benedetto, M.-G. (2014). Ultra wideband and bluetooth detection based on energy features. In *Ultra-WideBand (ICUWB), 2014 IEEE International Conference on*, pages 96–101. IEEE.

[71] Strohmer, T., Emami, M., Hansen, J., Papanicolaou, G., and Paulraj, A. J. (2004). Application of time-reversal with mmse equalizer to uwb communications. In *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE*, volume 5, pages 3123–3127. IEEE.

[72] Suski II, W. C., Temple, M. A., Mendenhall, M. J., and Mills, R. F. (2008). Radio frequency fingerprinting commercial communication devices to enhance electronic security. *International Journal of Electronic Security and Digital Forensics*, 1(3):301–322.

[73] Suykens, J. A., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J., Suykens, J., and Van Gestel, T. (2002). *Least squares support vector machines*, volume 4. World Scientific.

[74] Teitz, M. B. and Bart, P. (1968). Heuristic methods for estimating the generalized vertex median of a weighted graph. *Operations research*, 16(5):955–961.

[75] Toonstra, J. and Kinsner, W. (1996). A radio transmitter fingerprinting system odo-1. In *Electrical and Computer Engineering, 1996. Canadian Conference on*, volume 1, pages 60–63. IEEE.

[76] Wasserman, P. D. (1993). *Advanced methods in neural computing*. John Wiley & Sons, Inc.

[77] Win, M. Z., Scholtz, R., et al. (2002). Characterization of ultra-wide bandwidth wireless indoor channels: a communication-theoretic view. *Selected Areas in Communications, IEEE Journal on*, 20(9):1613–1627.

[78] Win, M. Z. and Scholtz, R. A. (1998). Impulse radio: How it works. *IEEE communications letters*, 2(2):36–38.

[79] Win, M. Z., Scholtz, R. A., et al. (1998). On the energy capture of ultrawide bandwidth signals in dense multipath environments. *IEEE Communications Letters*, 2(9):245–247.

[80] Win, M. Z., Scholtz, R. A., et al. (2000). Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications. *IEEE Transactions on communications*, 48(4):679–689.

[81] Woelfle, M., Temple, M., Mullins, M., and Mendenhall, M. (2009). Detecting, identifying and locating bluetooth devices using rf fingerprints. In *2009 Military Communications Conference (MILCOM 2009)*.

[82] Xu, R., Wunsch, D., et al. (2005). Survey of clustering algorithms. *Neural Networks, IEEE Transactions on*, 16(3):645–678.

[83] Zanetti, D., Danev, B., et al. (2010). Physical-layer identification of uhf rfid tags. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, pages 353–364. ACM.

[84] Zia, U., Uppal, M., and Naqvi, I. H. (2015). Robust feedback design for time reversal uwb communication systems under csit imperfections. *Communications Letters, IEEE*, 19(1):102–105.