

Ultra wideband and Bluetooth detection based on energy features

Hossein Soleimani, Giuseppe Caso, Luca De Nardis, Maria-Gabriella Di Benedetto

Department of Information Engineering, Electronics and Telecommunications (DIET)

Sapienza University of Rome

Rome, Italy

{soleimani, caso, lucadn, dibenedetto}@newyork.ing.uniroma1.it

Abstract— Detection, classification, and recognition based on the detection of energy features of Ultra Wide Band (UWB) vs. signals emitted in the Industrial Scientific and Medical (ISM) radio bands, such as Bluetooth, is a challenging issue. This work addressed this issue by analyzing the behavior of UWB versus Bluetooth signals in various noisy environments. The focus was on identifying robust feature extraction algorithms, that would enable encoding UWB and Bluetooth signals with features such as, for example, short time energy, Fast Fourier transform energy, and derivatives of short time energy. Results of experimental analysis showed that with respect to other signals, short-time energy of UWB over small overlapping time windows had acceptable discriminative performance. The different feature selection algorithms were tested with the following classifiers; 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.

Index Terms— Ultra Wide Band, Bluetooth, Energy features, Noise, Machine Learning, Features Selection.

I. INTRODUCTION

Impulse radio (IR) Ultra Wide Band (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 have peculiar behavior [1], and this can be used to discover UWB in the radio environment. In particular, UWB has specific short time energy profiles. This paper extends the above previous investigation by implementing feature selection algorithms and by testing these algorithms with different classification methods.

Automatic wireless communication system recognition is fundamental in order to discover and recognize wireless networks over the RF spectrum. Different technologies have their own specific physical signals and the behavior of them may serve to discriminate technologies. The use of machine learning methods to detect signals operating in the ISM 2.4 GHz band may prove to be a successful way forward.

The aim of this paper was to investigate the behavior of UWB and Bluetooth in various noisy environments. We propose a solution including a simple energy detection receiver. Stating

that accuracy in classification can be obtained only with a preliminary accurate features extraction, we focused on introducing robust extraction algorithms, which enable to extract features and use them in the following classification step.

We choose short time energy as the basic feature for discriminating Bluetooth and UWB signals because it is independent respect to small variation of frequency, if the time window used to evaluate it is properly chosen. Following previous results [1], in this work, the time window was set to 5 ns, that is a very short interval respect to the standardized Bluetooth time slot (625 μ s). For this reason, the Bluetooth frequency hopping nature does not affect the evaluation of the chosen features. Once we extracted, we sort features with two algorithms and we use different classifier (Support Vector Machine with related kernel methods, Probabilistic Neural Networks, K-Nearest Neighborhood and Naïve Bayes) in order to find the best solution for improving the detection performance in different noisy conditions.

The rest of the paper is organized as follows: Section II introduces the characteristics of UWB, Bluetooth and noise signals. Section III shortly reviews machine learning methods and their application to the specific issue of wireless network recognition. An analysis of UWB energy profiles, compared against energy profiles of traditional technologies, and how it affects recognition and classification is taken into account in Section IV. Section V provides conclusions and introduces future work directions.

II. SYSTEM MODEL

Given their impulsive nature, IR-UWB signals differ from conventional radio signals [2]. Both impulsive and multi-carrier UWB transmissions was approved in the U.S by the Federal Communications Commission (FCC) in 2002, for both indoor and outdoor applications, and adopted in the IEEE 802.15.4a standard. In IR-UWB, there is no carrier signal. Carriers do not actually transmit information while carrying information used for modulating transmitted signals. Subsequently, carriers of traditional radio waste energy. Therefore, from the power consumption point of view, UWB seems to be a better choice. As a matter of fact, FCC stated that

Power Spectral Density (PSD) of UWB should not exceed specific emission masks, and specified quite low power emissions. Low energy levels allow transmission without license but since low energy signals are similar to noise, their identification is quite challenging.

Bluetooth is another quite known system in the Personal Area Network context and it can be used to connect computers, mobile phones, etc.

To the goal of this paper, we propose to study four different scenarios and, for this reason, the receiver is characterized by 4 different states: UWB plus White Gaussian (WG) noise, Bluetooth plus WG noise, UWB plus Bluetooth plus WG noise, and WG noise. Figure 1 shows the system model.

In the experimental setup, Bluetooth signals were synthetically generated using MATLAB. Following the Bluetooth standard (IEEE 802.15.1), time is divided into fixed slots of $625\mu s$.

We 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.

Regarding UWB signals, they were generated adopting a Binary Pulse Position Modulation-Time Hopping (2PPM-TH) modulation technique. Other UWB parameters were: number of pulses per bit $NS = 1$, frame time $TS = 3$ ns, chip time $TC = 1$ ns, PPM time shift $\epsilon = 0.5$ ns, pulse duration $TM = 0.5$ ns, and pulse shaping factor $\tau = 0.25$ ns.

Received signals are illustrated in Fig. 2. The left-hand part of figure shows 10 ns long signals, with 40 dB SNR. Except for UWB signal, one can see that shapes of the other signals are similar to noise, even if a very high SNR is assumed. This leads to difficult signal identification. The right-hand part of Fig. 2 shows the short-term energy of the signals for overlapping time windows. It indicates that the energy variation of UWB is smooth, which is a helpful behavior to recognize it from other signals.

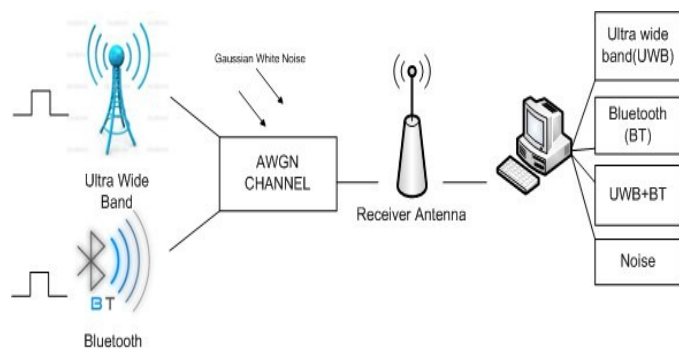


Figure 1- System model.

III. MACHINE LEARNING APPROACHES IN SIGNAL DETECTION

In this paper, the proposed approach is based on machine learning methods. The proposed solution is to identify signals

using their energy profiles and their derivatives. Depending on the scenario, we have 4 states that we need to identify. For this aim, we first extract the features and then sort them with feature selection methods. Finally, by classifying the features, we will show the effect of our features extraction on signal identification.

A. Feature Extraction

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 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.

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.

B. Feature Selection

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. In this Section, features that are more effective in the classification task are evaluated. To this end, two feature selection algorithms were applied: 1) minimum Redundancy Maximum Relevance (mRMR) (applied with three different methods) [3] and 2) Genetic Algorithm with Information Theory (GA) [4]. 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 relevance and redundancy using Mutual Information (MI). Following [3], this algorithm uses three different criteria to select features from the sorted subsets: Mutual Information Difference (MID (mRMR1)), defined as the difference between relevance and redundancy; Mutual Information Quotient (MIQ (mRMR3)), defined as the ratio between relevance and redundancy; and basic scheme (mRMR2), in which no particular combination of relevance and redundancy is defined. In any case, the output of each criteria is a vector with the indexes of the features that composes the optimal set of features.

C. Classification

The criteria used in mRMR and GA are the cross validated correct classification rate of 8 classifiers, where the class detection probabilities are shown using the confusion matrix and statistical graph. When a different classifier is employed, a correct classification rate is obtained for the mentioned classes. In the same experiment, various classifiers were applied to classify the data, generated using Matlab. From the several available classifiers, the following ones were selected: Support Vector Machine with different kernels (SVMs), K Nearest Neighborhood (KNN), Probabilistic Neural Network (PNN) and Naïve Bayes.

Most of previous works on classification have shown that SVMs are the most successful classification methods in machine learning [5]. 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 [6].

1) Support Vector Machine

- **Gaussian Radial Basis Function (RBF):**

$K(x_i, x_j) = \exp\{-\gamma|x_i - x_j|^2\}$ with $\gamma = 1$; we used RBF Kernel with Sequential Minimal Optimization (SMO) [7], that is a fast algorithm for training SVM by using pair-wise classification to break a multi-class problem into a set of 2-dimensional sub problems, eliminating the need for numerical optimization.

- **Multilayer Perception (MLP):** $K(x_i, x_j) = S(x_i^T x_j - 1)$

is a sigmoid function; we used MLP Kernel with the least-squares (LS) Method [8].

- **Quadratic kernel:** $K(x_i, x_j) = (x_i^T x_j - 1)^2$; We used quadratic Kernel with the least-squares (LS) Method.

- **Linear kernel:** $K(x_i, x_j) = x_i^T x_j$; We used linear Kernel with the SMO Method.

- **Polynomial kernel:** $K(x_i, x_j) = (x_i^T x_j + 1)^4$; A polynomial kernel of degree 4th is found to yield the same results with the cubic kernel. We used Polynomial Kernel with the least-squares (LS) Method.

2) Probabilistic Neural Network (PNN)

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 formula. PNN is based on biological neural network, processing the information and able to classify data [9]. PNN is an adapted version of the radial basis network and it 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. Table 1 shows, in detail, the confusion matrix measured for the PNN classifier on our data.

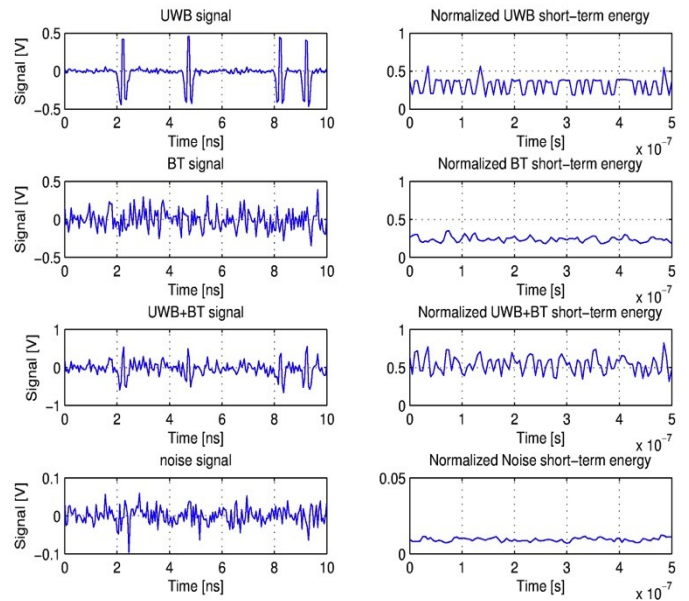


Figure 2- Different Signals and their energy with 40 dB SNR.

3) Naive Bayes classifier

Naive Bayes classifiers are discriminate and supervise learning methods that optimize conditional likelihood. The intuition of Naive Bayes is very simple and it is based on Bayes rules. In this classifier the decision is made based on conditional probability, where the likelihood of estimator is maximized. Results obtained with this classifier are shown in Table 1. Although this classifier is very simple, the performance is quite good.

4) K nearest neighbors

One of the simplest classification algorithms is the K Nearest Neighbors [10]. KNN takes the new point and classifies it according to the majority vote of the K nearest points 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 neighbors and the most suitable distance. New coming data look at and query from K neighborhoods then calculates

distance from these K neighborhoods and finally samples map to the nearest group and maps to the nearest one. Four K -NNs have been employed with different distance functions; however we use the Euclidean distance because it does not affect the classification accuracy. We choose $K=3$ neighborhoods to evaluate the total experiment. In this case, none of the results of the K -NN would be stable and thus valid for classification

D. Evaluation Criteria

Different metrics were used to measure the performance of our proposed algorithm. These metrics are useful to analyze data without prior assumptions about the data.

1) The 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.

2) 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 for each class.

The recall (correct classification rate) is the proportion of true positive in each class divided by true positive plus false negatives.

3) The precision is the proportion of the predicted true positive in each class divided by predicted true positive plus predicted false positives.

4) The F-measure is the harmonic mean of precision and recall and is calculated as:

$$2 \times \text{recall} \times \text{precision} / (\text{recall} + \text{precision}).$$

IV. ANALYSIS IN NOISY CHANNELS AND SELECTED FEATURE ANALYSIS

Data is analyzed to recognize and detect UWB and Bluetooth signals, as mentioned above. One of the most important issue in UWB is that there exists interference with other signals as well as noise. As previously shown in Fig. 2, energy in time windows of 5 ns with 50 percent overlapping was computed. Data were divided so that 70 percent were used to train and the remaining was used to classify.

A. Performance of different classifiers

Experiments were carried out with SVM based on five different kernels, K Nearest Neighborhoods, Probabilistic Neural Network, and Naïve Bayes classifier. In order to find the optimal separating hyperplane 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 hyperplane, 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 was 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 were selected randomly. The classifier was trained using the training set and the recall and accuracy were estimated on the test set.

B. Results with Confusion matrix

We investigated 4 classes in the database characterized by 20 dB SNR. We supposed that the distances between transmitters and receivers were constant. The performance of the predictive model is examined using the confusion matrix in Table 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 on the diagonal shows that test observation is correctly classified; on the other hand, any number in non-diagonal part means that it is not correctly classified. Table 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.

Accuracy is an advantageous metric in many applications, and is calculated from the confusion matrix. Fig. 3 shows accuracy performance of signals detection with different classifiers. Accuracy illustrates global performance, which means the capability of classifiers to discriminate classes' samples. High accuracy demonstrates power of a classifier to separate data of the 4 classes. Signal evaluation, presented in Fig. 3, is the global performance of the approaches with regards to the all 4 classes. Fig. 3 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 leads to worst results independently of SNR. If we look at all of SNR values, SVM RBF has acceptable efficiency. Other classifiers have very similar results; hence, classifier selection with different SNR is imposed according to the maximum accuracy. For example, in SNR= -20 dB we should choose the SVM quadratic with 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.

Table 1- Confusion matrix on the 7 features when 30% of the samples of 20 dB SNR data are used for testing.

| | | mRMR on 20 dB with 7 features | | | | | | | | |
|-------------------|----------------|-------------------------------|-----|-----|-----|----------------|-----|-----|-----|-----|
| SVM Polynomial | Linear | 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 | RBF | 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 | KNN | 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 | Naive Bayes | 102 | 39 | 1 | 1 | Naive Bayes | 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 |

C. Feature Selection Analysis

We show in Fig. 4 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. 4, 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 4 demonstrates the effect of feature selection showing that with fewer data we have the same performance. The best 4 features were chosen by the different feature selection algorithms as follows.

With mRMR1 selected feature were Short window time energy, Gaussian normalized window time energy, Short Fourier transform of short time energy, and Fast Fourier transform short time energy.

mRMR2 sorted features such as 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.

mRMR3 selected features as short window time energy, short Fourier transform of short time energy, fast Fourier transform short time energy, and Gaussian normalize window time energy. Stability of methods is shown to be good with compared to the genetic algorithm. Feature selection is some kind of statistical study that if you have good feature selection you can do better storage. MID is more stable than MIQ.

Finally, the Genetic algorithm choose 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.

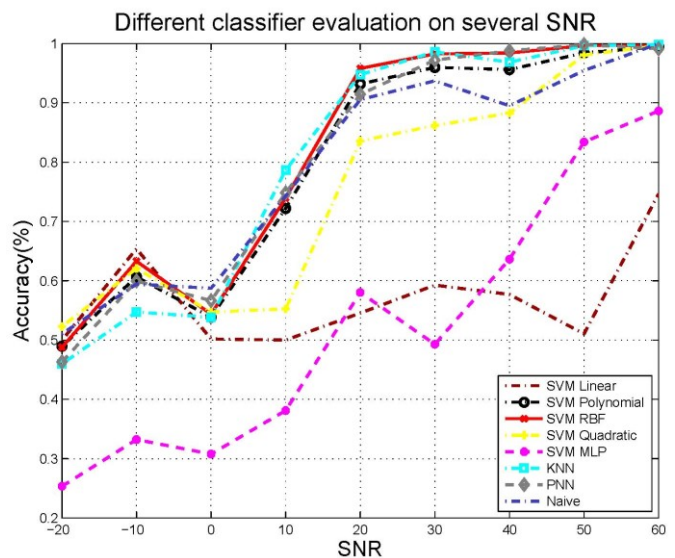


Figure 3- Evaluation classifier in different SNR.

D. 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 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, for 10 dB SNR, UWB+BT with RBF kernel provides the best correct classification rate of 99.3% , 98.61%, 98.95% respectively. We have seen precision Noise has the second best performance of 96.7%, whereas UWB have the recall with SVM RBF of 93%. The recall of noise signals for RBF is 82.5%. Bluetooth (BT) provides performance of 33.5% according to F-measure when RBF is used.

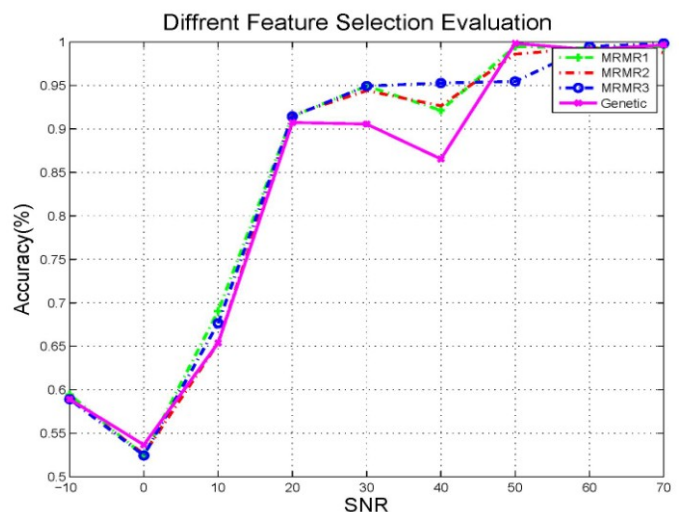


Figure 4- Evaluation feature selection in different SNR.

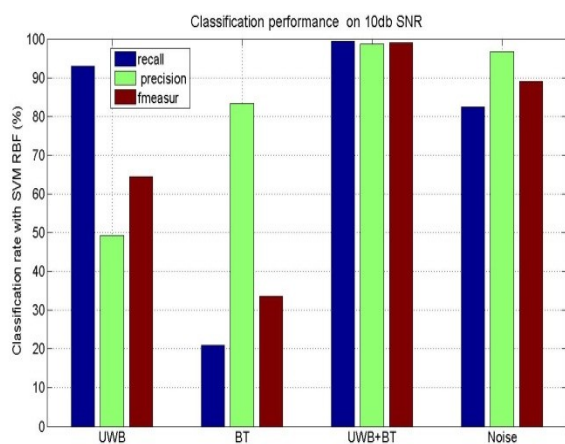


Figure 5- Classification performance on 10 dB with SVM RBF on different classes using recall, precision and f-measure criteria.

V. CONCLUSIONS

In this paper, we have examined the presence of UWB versus Bluetooth, and noise. Furthermore we concentrated on extracting robust features that make possible to discriminate the mentioned classes. 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 are the features extracted from receiver. Afterwards, we evaluated worthiness of each feature with feature selection methods such as minimum Redundancy Maximum Relevance and feature selection based on Genetic Algorithm (GA) and Information Theory. Receiving signals were UWB, Bluetooth, Bluetooth plus UWB and noise on which we tested the performance of our algorithm to understand receiving signals. Support Vector Machine classifier with related kernel methods, Probabilistic Neural Networks, K Nearest Neighborhood and Naïve Bayes, were used to investigate behavior features extracted from receiving signals. We evaluated the algorithms for the detection of signals belonging to four different classes: Based on the results, we provided several conclusions; we showed that the SVM with Gaussian RBF kernel gives the most accurate results for more than 10 dB SNR. For 0 dB, Naïve Bayes has the best accuracy, and for less than 0 dB all kernels have same performance.

In this paper we supposed that the distance between transmitters and receivers are fixed but in realistic environment propagation of signals have different distances. We plan to investigate different path losses and technologies in future work.

ACKNOWLEDGMENT

This work was carried out in the framework of the joint Telecom Italia - DIET Dept. Labs "AWESAM". Part of this work was supported by the "Ricerca Scientifica 2013" project by Sapienza University of Rome, COST Action IC0902, ICT ACROPOLIS NoE FP7 project n.257626, and funded by Telecom Italia in the contract 2014 with Sapienza University of Rome "Internet of Things".

REFERENCES

- [1] S. Boldrini, G.C Ferrante, M.G Di Benedetto, "UWB network recognition based on impulsiveness of energy profiles" IEEE international conference on ultra-wideband (ICUWB), 14-16 Sept. 2011, pp. 327-330, 2011.
- [2] M.-G. Di Benedetto, and G. Giancola, "Understanding Ultra Wide Band Radio Fundamentals", 1st Ed., Prentice Hall PTR, 2004.
- [3] H.C. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy" IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 8, pp. 1226-1238, 2005.
- [4] O. Ludwig, U. Nunes, R. Araujo, L. Schnitman, "Application of information theory, genetic algorithm, and neural models to predict oil flow" Communications in Nonlinear Science and Numerical Simulation, Vol. 14, Issue 7, p. 2870-2885, 2008.
- [5] B. Scholkopf, and A.J. Smola, "Learning with Kernels" MIT Press, Cambridge, MA. 2002.
- [6] N. Cristianini and J. Shawe-Taylor, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods" Cambridge University Press, 2000.
- [7] J. C. Platt, "Sequential minimal optimization: A fast algorithm for training support vector machines" Technical Report MSR-TR-98-14, Microsoft Research, April 1998.
- [8] J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, "Least Squares Support Vector Machines" World Scientific, Singapore, ISBN 981-238-151-1, 2002.
- [9] T. Mitchell, "Machine Learning" McGraw-Hill, 1997.
- [10] P.D. Wasserman, "Advanced Methods in Neural Computing" New York, Van Nostrand Reinhold, pp. 35-55, 1993.