TARGET DETECTION USING LEARNING METHODS

by

Mohammad Moufeed Sahnoon

A Thesis presented to the Faculty of the
American University of Sharjah
College of Engineering
In Partial Fulfillment
of the Requirements
for the Degree of

Master of Science in
Electrical Engineering

Sharjah, United Arab Emirates

June 2017

Approval Signatures

We, the undersigned, approve the Master's Thesis of Mohammad Moufeed Sahnoon.

Thesis Title: Target Detection Using Learning Methods.

Signature	**Date of Signature** (dd/mm/yyyy)
___________________________ Dr. Khaled Assaleh Professor, Department of Electrical Engineering Thesis Advisor	_______________
___________________________ Dr. Usman Tariq Associate Professor, Department of Electrical Engineering Thesis Co-Advisor	_______________
___________________________ Dr. Hasan Mir Associate Professor, Department of Electrical Engineering Thesis Co-Advisor	_______________
___________________________ Dr. Mahmoud Ismail Associate Professor, Department of Electrical Engineering Thesis Committee Member	_______________
___________________________ Dr. Mamoun Abdel-Hafez Associate Professor, Department of Mechanical Engineering Thesis Committee Member	_______________
___________________________ Dr. Nasser Qaddoumi Head, Department of Electrical Engineering	_______________
___________________________ Dr. Ghaleb Husseini Associate Dean for Graduate Affairs and Research, College of Engineering	_______________
___________________________ Dr. Richard Schoephoerster Dean, College of Engineering	_______________
___________________________ Dr. Mohamed El-Tarhuni Vice Provost for Graduate Studies	_______________

Acknowledgement

I would like to thank my advisors: Dr. Khaled Assaleh, Dr. Usman Tariq and Dr. Hasan Mir for providing knowledge, guidance, support, and motivation throughout my research stages. I am greatly thankful for their noble assistance, worthy discussions, and suggestions. I would also like to express my very great appreciation to the American University of Sharjah as a whole and in particular the Electrical Engineering department for giving me the opportunity to complete my graduate studies and be a part of their graduate program. Finally, I wish to thank my family and friends for their tremendous support and encouragement.

To the soul

of my grandmother…

Abstract

Adaptive beamforming is an array processing method that can be used for target detection. In the absence of clutter signals, this method uses a one-dimensional adaptive filter called the space filter in the spatial dimension using a uniformly linear array as a receiver that is made of N-channels separated by a distance d. The N-receiver channels work on collecting target-free data that can be used as training data for the radar along with collecting the target signal with all types of interferences. The training data are then used to build the covariance matrix that is used in determining the adaptive beamformer filter weights. After that, the received data are projected onto these weights to null the jamming signals, minimize noise, and amplify the target signal. Finally, the output, after projection, is compared with a measured threshold value to decide upon the presence of the target. This conventional method suffers from several problems such as target cancellation when the training data collected are not target free. Furthermore, the amount of secondary data required is usually not available in such applications. Thus, different algorithms must be found or developed to overcome or improve the problems of the conventional method. In this report, a target detection system that involves direction of arrival estimation and learning based algorithms is proposed. The proposed system is assumed to overcome the problem of the jamming signal direction of arrival variations between the training and testing stages, signal-to-interference-plus-noise-ratio variations and the necessity for target free secondary data. Another target detection system is also proposed, i.e. the cascade system. This system uses the adaptive beamforming method as an unsupervised dimensionality reduction technique in line with the learning-based method for target detection, and it shows a comparable performance as compared to the original proposed system.

Keywords: Adaptive beamforming; MTI radar; DOA estimation; pattern classification.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

MTI	Moving Target Indicators
AMTI	Airborne Moving Target Indicators
GMTI	Ground Moving Target Indicators
STI	Stationary Target Indicators
SMTI	Stationary and Moving Target Indicators
SINR	Signal-to-Interference-plus-Noise-Ratio
DOA	Direction of Arrival
ULA	Uniformly Linear Array
RCS	Radar Cross Section
AWGN	Additive White Gaussian Noise
PCA	Principle Component Analysis
SVD	Singular Value Decomposition
LDA	Linear Discriminant Analysis
ESPRIT	Estimation of Signal Parameters via Rotational Invariance Tech.
MUSIC	Multiple Signal Classification
MSE	Mean-Squared Error
SVM	Support Machine Vector

Chapter 1. Introduction

In this chapter, a short introduction is provided about target detection and the encountered problems in this field. The chapter then moves to highlight the problem investigated in this study as well as the thesis contribution. Finally, the general organization of the thesis is presented.

1.1 Target Detection Overview

The need to know if a particular object or condition is present is sometimes significant for engineers and scientists. For instance, geophysicists explore the earth for oil or water; pilots need to locate other airplanes approaching their aircraft during a flight; astronomers search the universe for new planets or stars, etc. These situations usually make use of a threshold to be compared with the output of the processed received data. The decision of the presence of a target (the object or condition being sought) is taken if the threshold is exceeded [1].

In literature, different types of target indicators are developed and presented such as moving target indicators (MTI), airborne moving target indicators (AMTI), ground moving target indicators (GMTI), stationary target indicators (STI) and combined stationary and moving target indicators (SMTI) [2]. For the different types of MTI, the main goal is to differentiate a target against the clutter. In order to do so, the most common approaches take advantage of the Doppler's effect [2]. The moving target will change its distance from the radar system for a given sequence of radar pulses, hence, the reflection of the radar phase that returns from the target will be different for consequent pulses. On the other hand, STI mode of operation takes advantage of the fact that the target is moving against a stationary clutter [2].

In the process of detecting a target, the received signal is a combination of the target signal and all other interferences such as noise and jamming. The probability of detection P_D depends mainly on the probability of false alarm and the signal-to-interference-plus-noise-ratio (SINR). Jamming signals are considered the most powerful source of interference that may disrupt the received signal. These two interferences need to be suppressed in order to maximize the SINR, and as a consequence, this will maximize the probability of detection P_D. The clutter signals are not taken into consideration in this report since our radar platform is assumed to be the stationary.

1.2 Thesis Objectives

Due to the observed inefficient performance of the conventional space adaptive processing algorithm in detecting targets whenever the jamming signals direction of arrival (DOA) and the SINR are different than the training scenario, the interest of building a system that overcomes this problem has arisen. DOA estimation and learning based techniques were taken as an advantage to build the proposed system.

1.3 Research Contribution

The contributions of this research work can be summarized as follows:

- Propose a target detection algorithm that has greater immunity to noise and interference in severe environments compared to the conventional method. The proposed method is able to detect targets with a significant accuracy even when the noise and target signals are on the same power levels and when the jamming signals are much greater than the target signal, which the conventional method fails to handle.
- A new unsupervised method of dimensionality reduction that does not incorporate class information and helps in the target detection process is found by taking advantage of the adaptive beamforming output.

1.4 Thesis Organization

The rest of the thesis report is organized as follows: Chapter 2 states the problem definition together with illustrating the different signal types used. It also provides a literature review on array processing and pattern classification systems. Chapter 3 explains the proposed system for target detection and introduces another proposed cascade system that has a comparable performance with a significant dimensionality reduction in the training and testing data. Simulation results and performance evaluation are presented in Chapter 4. Finally, chapter 5 concludes by summarising all the findings presented in this thesis.

Chapter 2. Background and Literature Review

In this chapter, the problem architecture adopted in this thesis is presented along with all the associated signal types used. Then, the chapter provides a literature review on array processing techniques which includes adaptive beamforming, principle component analysis, fisher discriminant analysis and direction of arrival (DOA) estimation. Finally, the chapter concludes with an overview on pattern classification systems followed by two main classification systems which are the linear classifier and the second-order polynomial classifier.

2.1 Problem Definition

A uniformly linear array (ULA) receiver is used as the antenna pattern in this thesis report. This receiver consists of N channels that are separated by a distance d as shown in Figure 1. The distance d can be set depending on the expected wavelength that will be received by the ULA. The reference point of the ULA is assumed to be at the first element. The ULA elements are assumed to have the same radiation pattern. All the three algorithms which are the adaptive beamforming, linear classification and second-order polynomial classification that are discussed in this thesis report are assumed to have the previously mentioned antenna array geometry. These methods are considered as learning based methods since they are trained using a certain set of data before they are used.

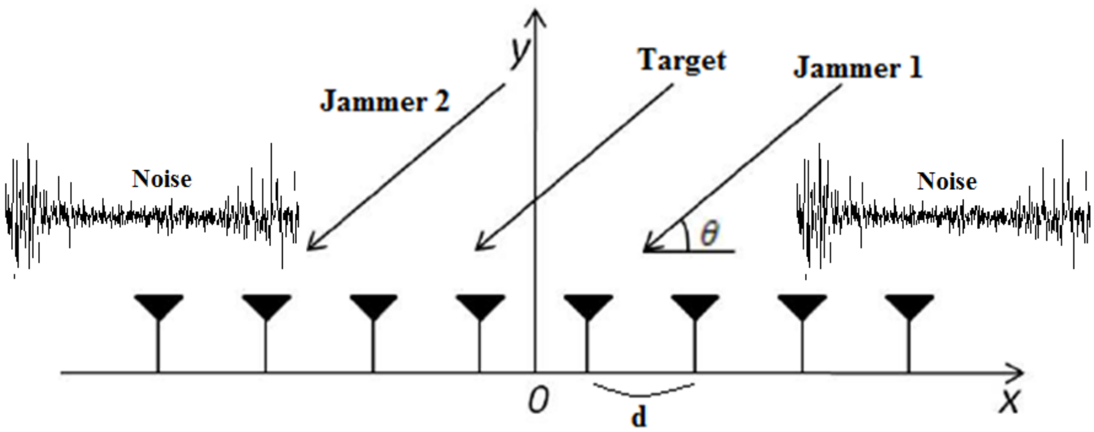

Figure 1: ULA receiver geometry.

2.2 Signal Types

This section clarifies all types of signals used throughout this thesis report. Such signals include the target, jamming and noise signals that are used to model the received signal to the radar system.

2.2.1 Target. In radar theory, the received target power depends on many variables such as the radar cross-section (RCS) of the target, the transmitted power, the array power gain, the target range gate, the radiation pattern, the receiver noise, and the radar's frequency of operation [3]. In our case, the target signal is characterized by two main parameters: the first parameter is the target DOA, θ_T, which is used to build the target steering vector $\boldsymbol{s}_T$. The steering vector definition is given by:

$$\boldsymbol{s} = \begin{bmatrix} 1 \\ e^{-j\pi sin\theta} \\ e^{-j2\pi sin\theta} \\ . \\ . \\ . \\ e^{-j(N-1)\pi sin\theta} \end{bmatrix} \tag{1}$$

where N is the number of the receiver channels. The spacing between the channels, d, is assumed to be half of the target wave length λ as shown in (2). This spacing value is chosen because it gives the best beam pattern with the number of ULA channels (N) chosen in this report as illustrated in [4].

$$d = \frac{\lambda}{2} \tag{2}$$

The steering vector definition is the same for all different types of signals used in this report; the only difference is in the signal DOA, θ, used to build the steering vector. The second parameter is the target relative power, ζ_T, which is seen at the receiver. This relative power is assumed to be constant, i.e. not random. The target signal model, $\boldsymbol{T}$, can then be given by:

$$\boldsymbol{T}(\theta_T, \zeta_T) = \zeta_T\, \boldsymbol{s}_T(\theta_T) \tag{3}$$

2.2.2 Jamming. Jamming signals appear centered at a certain azimuth angle and spread over all frequencies. These signals are correlated in the spatial dimension but uncorrelated in the temporal dimension. They can be suppressed by simply placing a null in the radar array pattern in the direction of the jammer [5]. Like the target signal, the jamming signal can be characterized by its angle of arrival and relative power seen at the receiver; however, the jamming signal power is not constant here but random. The jamming signal model, $\boldsymbol{J}$, is given by:

$$\boldsymbol{J}(\theta_J, \zeta_J) = \zeta_J \, \boldsymbol{s}_J(\theta_J) \tag{4}$$

where ζ_J is the jamming signal relative power, $\boldsymbol{s}_J$ is the jamming steering vector and θ_J is the jamming signal angle of arrival.

2.2.3 Noise. In signal processing, noise is a general term for unwanted and unknown modifications that a signal may suffer during capture, storage, transmission, processing, or conversion [6]. Noise signals cannot be nulled because they are spread over all frequencies, and they come from every angle, however, in most cases, researchers do not care about the shape of the target signal as long as it can be differentiated from the interference signals. In this report, the term "noise" refers to the thermal noise seen at each receiver element. It is assumed to be an additive white Gaussian noise (AWGN) that is uncorrelated from one channel to another. The noise signal power, ζ_n, can be determined from the noise power spectral density, N_o, and the bandwidth of the receiver, B, and it is given by [6]:

$$\zeta_n = N_o \, B \tag{5}$$

2.3 Array Processing Techniques

In this section, different array processing techniques are presented. These techniques include adaptive beamforming, principle component analysis, fisher discriminant analysis and DOA estimation.

2.3.1 Adaptive beamforming. In the 1960s, adaptive beamforming was initially developed for the military use of sonar and radar [7]. Several modern applications exist for adaptive beamforming; one of the most significant applications is the commercial wireless networks such as 3GPP, LTE and IEEE 802.16 WiMax. These networks depend on adaptive beamforming to enable primary services within each standard [8]. In the literature, there are several approaches for designing a beamformer. The very first approach was implemented by Applebaum in 1965 where he used the maximization of the signal to noise ratio (SNR) in order to maximize the received signal power while minimize noise and all other interference signals such as jamming [7]. Another approach which uses the least mean squares (LMS) error was implemented by Widrow. The maximum likelihood method (MLM) was also developed by Capon in 1969. Both algorithms of Applebaum and Widrow are very similar and converge toward an optimal solution [9]. However, these techniques have implementation drawbacks. Reed proposed a technique known as Sample-Matrix Inversion (SMI) in 1974 which obtains the adaptive antenna array weights directly unlike the algorithms of Applebaum and Widrow [7].

Adaptive beamforming is a technique that performs adaptive spatial signal processing using an array of transmitters or receivers. In principle, it increases the target signal strength from a chosen direction while degrading all other signals from undesired directions. This technique is used mainly in radar systems to provide directional sensitivity without moving the array of receivers or transmitters physically. The way this method works is that it estimates the covariance matrix, $\boldsymbol{R_e}$, from a secondary set of data that is collected from the N-receiver channels. The secondary (training) data used in this process needs to be target-free to avoid what is known as target cancellation. In practice, this can be done by carrying out turning on the N-receiver channels when there is no target signal in the environment and observing all the interference signals that exist in the environment around the radar. The secondary data must also have the same statistical properties as the data under test which have target signal included. If any of these two conditions is not satisfied, the result will be either a poor estimate of the interference covariance matrix, hence a poor filter output, or the loss of the target as an outcome of target-cancellation [10]. After processing the received signal through the digital signal processing (DSP) unit, the discretized signal seen at the first channel, $\boldsymbol{X}_{d,1}$, is given by:

$$\boldsymbol{X}_{d,1} = [x_{11} \quad x_{12} \quad x_{13} \quad x_{14} \quad \cdots \quad x_{1K}] \tag{6}$$

where K refers to the number of samples used to discretize the received signal. Similarly, the discretized signal seen at the second channel, $\boldsymbol{X}_{d,2}$, is given by:

$$\boldsymbol{X}_{d,2} = [x_{21} \quad x_{22} \quad x_{23} \quad x_{24} \quad \cdots \quad x_{2K}] \tag{7}$$

Then we can write the discretized signal seen from the all N-receiver channels as follows:

$$\boldsymbol{X_d} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} & \cdots & x_{1K} \\ x_{21} & x_{22} & x_{23} & x_{24} & \cdots & x_{2K} \\ x_{31} & x_{32} & x_{33} & x_{34} & \cdots & x_{3K} \\ x_{41} & x_{42} & x_{43} & x_{44} & \cdots & x_{4K} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & x_{N3} & x_{N4} & \cdots & x_{NK} \end{bmatrix} \tag{8}$$

This discretized received signal form is held true for both training and testing data with or without a target signal. The notation used for data with no target in this report will be $\boldsymbol{X}_{d,notar,}$ As for data with no knowledge of whether there is a target or not, the notation will be $\boldsymbol{X_d}$. After obtaining the discretized received signal, the covariance matrix can be estimated using the following equation:

$$\boldsymbol{R_e} = \frac{\mathbf{1}}{\boldsymbol{K}} X_{d,notar}\, X_{d,notar}{}^{T} \tag{9}$$

where $\boldsymbol{X}_{d,notar}{}^{T}$ is the transpose of the no target discretized received signal matrix. Next, the filter weights vector $\boldsymbol{w}$ that is used to suppress the interference signals and amplify the target signal can be obtained using the estimated covariance matrix as shown below.

$$\boldsymbol{w} = \boldsymbol{R_e}^{-1}\, \boldsymbol{s} \tag{10}$$

Where $\boldsymbol{s}$ refers to the target space steering vector shown in (1). After building the estimated covariance matrix and getting the filter weights vector, the angle of the jamming signals and the angle of the target signal can easily be decided on. Our interest here is to null the jamming signals, suppress the noise signals, and amplify the target signal. This can be done by computing the filter output using the filter weight vector and the received signal. The filter output is given by:

$$\boldsymbol{y} = \boldsymbol{w}^{\boldsymbol{H}}\, \boldsymbol{X_d} \tag{11}$$

where $\boldsymbol{w}^{\boldsymbol{H}}$ is the Hermitian transpose of the filter weight vector. The magnitude of the output, $|\boldsymbol{y}|$, is then used to decide on the presence of a target. One way of deciding on the presence of a target is to compare the power of the output magnitude at the range gate[1] or samples that contain a target with those which do not include a target, and as this power difference increases, the performance of the system does also increase. Another way is to set a threshold value, Y_{th}, and compare the output magnitude with it. The decision will then be identified as:

$$D(|\boldsymbol{y}|) = \begin{cases} |\boldsymbol{y}| \geq Y_{th}, & H_0 \\ |\boldsymbol{y}| < Y_{th}, & H_1 \end{cases} \tag{12}$$

where H_0 refers to the hypothesis that indicates the presence of a target and H_1 is the hypothesis that indicates the non-presence of a target.

2.3.2 Adaptive beamforming recent algorithms. Several recently developed adaptive beamforming algorithms exist in the literature. One algorithm is the robust adaptive beamforming with precise main beam control which takes into account the steering vector uncertainties in the magnitude response of the adaptive beamformer and uses the semidefinite relaxation technique as approximate solver [11]. Another algorithm is the constant modulus reduced-rank beamforming which uses a generalized sidelobe canceller structure for interference suppression [12]. A third adaptive beamforming algorithm that is based on conjugate gradient algorithms is introduced in [13]. This algorithm offers two different methods of adaptive beamforming. The first method takes advantage of the diagonal loading technique [14], while the second method uses the regularization technique [15].

2.3.3 Wiener-Hopf filter. After building the estimated covariance matrix, $\boldsymbol{R_e}$, and getting the filter weights vector $\boldsymbol{w}$, the direction of arrival (DOA) of the jamming signals and the target signal can be determined. Our interest here is to null the jamming signals and amplify the target signal. This is done by using the wiener-hopf filter output, F, that is given by:

$$\mathrm{F} = |\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{s}| \tag{13}$$

The angle used to build the steering vector $\boldsymbol{s}$ is swept over all different possible angles, and for each one of these angles, the weight vector built using the adaptive beamforming method will determine whether this angle corresponds to a jamming

[1] Range gate refers to a certain distance on the ground.

signal or to a target signal. Based on this, nulling will occur on the direction of the jamming signals, and amplification will occur on the direction of the target signal. It is important to note here that noise signals cannot be nulled because they are spread over all frequencies, and they come from every angle; however, in most cases no much attention is given to the shape of the target signal as long as it can be differentiated from the interference signal.

2.3.4 Principle component analysis. Principal component analysis (PCA) is an unsupervised multivariate technique that analyzes a data set that consists of a number of observations described by several dependent features. The main goal of PCA is to project the data into a lower dimensional sub-space in the direction of their maximum variances regardless of their classes; thus, it is not always suitable for classification purposes as it may mix up the different classes and make the classification job harder. Mathematically, PCA depends upon the Eigen-decomposition of positive semidefinite matrices and upon the singular value decomposition (SVD) of rectangular matrices [16]. PCA is found to be unhelpful for the target detection process using the proposed systems mentioned in this thesis.

2.3.5 Fisher discriminant analysis. Fisher linear discriminant analysis (also called linear discriminant analysis (LDA)) is a method used in statistics, pattern recognition and machine learning to find a linear combination of features which separates two or more classes of events. The resulting combination can be used for dimensionality reduction purposes before the classification stage. LDA is closely related to PCA since both techniques apply linear transformations on a given data set. However, unlike PCA, LDA transformation is based on maximizing the ratio of "between-class variance" to "within-class variance" with the goal of reducing data variation in the same class and increasing the separation between classes [17].

2.3.6 Adaptive beamforming as a dimensionality reduction technique. In this thesis, it was found that adaptive beamforming can act as an unsupervised dimensionality reduction technique that is useful for target detection (classification) purposes. The researcher was able to apply adaptive beamforming on our training and testing data sets and reduce their dimensionality significantly along with getting a comparable classification output performance. This technique is used in the proposed cascade system that is illustrated in Chapter 3.

2.3.7 DOA estimation. Direction of arrival (DOA) estimation is a wide and significant research area in array signal processing where several engineering applications need sufficient algorithms for DOA estimation [18]. Spatial spectrum is a major concept in array signal processing theory. It estimates the signal's distribution in every direction in the space, hence, knowing the signal's spatial spectrum leads to the estimation of DOA of a signal. Consequently, spatial spectrum estimation can also be referred to as DOA estimation.

Various kinds of super resolution algorithms are present in the theory such as spectral estimation, Bartlett, Capon, ESPRIT, Min-norm and MUSIC [19]. One of the most popular subspace-based techniques to estimate the DOA of multiple signal sources is the MUSIC algorithm, an acronym that stands for multiple signal classification. Using the MUSIC algorithm involves large numbers of computations to search for the spectral angle. Therefore, in practical applications, its implementation can be challenging. This algorithm can only be used in uniform linear array (ULA) or non-uniform linear array whose arrays are restricted to a uniform grid [19].

2.3.8 MUSIC DOA estimation algorithm. Let's assume a test signal of the form as in (8) is received by the ULA, and no information is available about the directions of arrival of each component (target and jamming) of this signal. Then, the job of the MUSIC estimator is to estimate the directions of arrival at which the energy is concentrated. To do so, the estimated covariance matrix, $\boldsymbol{R}_e$, of the test signal must be built using the same expression as in (9) but by using the $\boldsymbol{X}_d$ feature matrix instead of $\boldsymbol{X}_{d,notar}$. At this point, the eigenvalues of the estimated covariance matrix must be found and sorted in a descending order. In theory, there are several ways that may help in finding the maximum and the minimum eigenvalue of a system which will aid in the sorting job. One of these methods is the direct power method that helps in finding the maximum eigenvalue of a system [20]. Another method is the inverse power method which finds the minimum eigenvalue of a system [20]. However, once we have the eigenvalues of a system, the sorting job becomes easy, which helps in identifying the number of eigenvectors needed to build the interference subspace. Next, the eigenvectors of the system are sorted according to their associated eigenvalues. Once the eigenvalues stop changing and reach to a constant value (reach the elbow), all their associated eigenvectors are used to build the interference subspace. After that, the MUSIC DOA estimator can be determined by [21]:

$$\boldsymbol{MUSIC} = ((\|\boldsymbol{s^H}(\theta)\ \boldsymbol{E_i}\|_2)^{-1})^2 \tag{14}$$

where $\boldsymbol{s^H}(\theta)$ refers to the Hermitian transpose of the steering vector at a given angle θ and $\boldsymbol{E_i}$ refers to the interference subspace. Finally, to obtain the DOA of a signal at which the energy is concentrated, the angle, θ, used in the steering vector must be swept among all the possible angles, and the output of the MUSIC estimator must be observed. The higher the output at a certain angle, the more confidence is reached have about the DOA of a signal.

2.4 Pattern Classification

This section gives an overview on the theory of pattern classification. It describes the different stages needed to build a pattern classification system. Two types of classifiers are discussed at the end of this section: the linear classifier and the second-order polynomial classifier.

2.4.1 Pattern classification systems overview. Pattern recognition (classification) is a branch of machine learning that concentrates on the recognition of patterns and uniformities in data, although it is in some cases considered to be nearly synonymous with machine learning [22]. The main objective of a pattern classification system is to classify different objects into groups (classes) based on their different properties (features). The different stages needed to build a pattern classification system are illustrated in Figure 2.

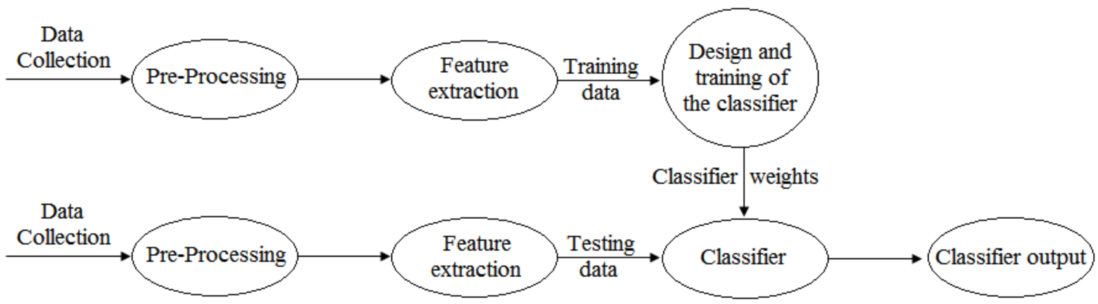

Figure 2: Pattern classification systems stages summary.

First, data is collected from the input of the system which could be voice, image, or data based input. Second, the pre-processing stage processes the useful data obtained through data collection. Then the different features are extracted from the useful data set, and according to these features the type of the classifier (linear or nonlinear) is selected. After that, the classifier is trained using a training data set and then tested using a testing data set. If the classifier succeeds in differentiating between the different classes, then it is said to be a suitable classifier for that specific application. However, if it fails, a different classifier type should be selected, and/or a different feature extraction technique should be used.

The feature extraction process is considered one of the most important stages in a pattern classification system. The obtained features are expected to include relevant information from the acquired received data so that the classification task can be done using this reduced representation instead of using the complete set of the data. To illustrate the importance of the feature extraction process, an example is shown in Figure 3. Figure 3(a) shows a combination of two features that is able to separate different classes efficiently. In Figure 3(b), on the other hand, the selected combination of features fails to discriminate between the two classes, hence, the classification task will be hard and maybe unobtainable.

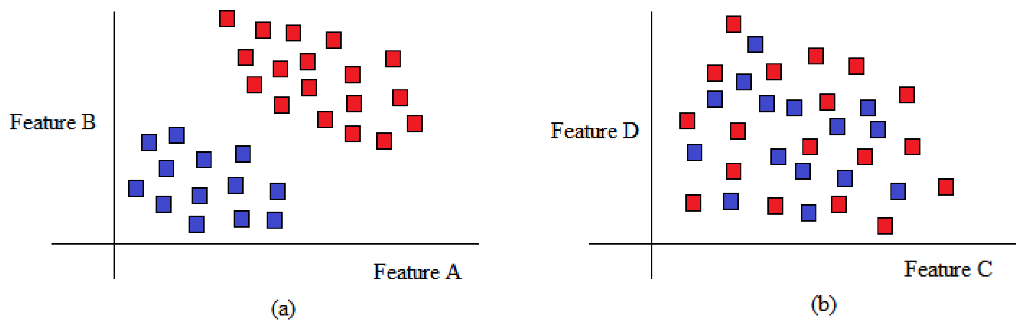

Figure 3: Example on effect of the feature extraction process on classification.

2.4.2 Linear Classifier. A linear classifier helps in reaching a classification decision based on the value of a linear combination of the characteristics. The characteristics of the object are also known as feature values and are typically presented to the machine in a vector called a feature vector. Such classifiers work well for practical problems such as document classification, and more generally for problems with many variables (features), reaching accuracy levels comparable to non-linear classifiers while taking less time to train and use [23]. Linear classifiers assume that the classes can be separated linearly. A linear classifier can be represented by a linear discriminant function of the form:

$$y(x) = \boldsymbol{w}^{\mathbf{T}}\boldsymbol{x} + w_o \tag{15}$$

where $\boldsymbol{x}$ represents a feature vector, $\boldsymbol{w}$ is the weight vector that will be projected onto the testing data to do the classification job and w_o is a constant bias value. A two-class problem with one, two, or three-dimensional feature space will have a decision boundary of a point, line, or plane respectively. Figure 4 shows a two-class problem with three-dimensional feature space. As the dimensionality of the feature space goes higher, the decision boundary shape gets more complex, but this will not necessarily make the classification job harder.

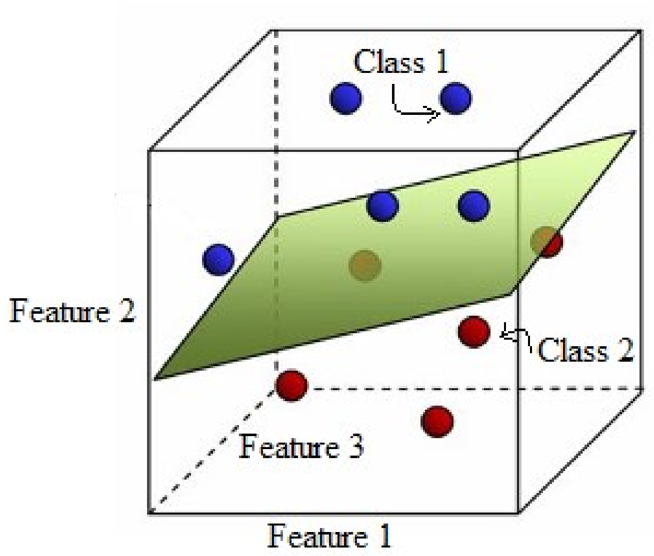

Figure 4: Two-class three-dimensional feature space problem example [24].

Many optimization methods can be used to train a linear classifier and determine its decision boundary. One of the methods uses a gradient descent algorithm to find the weight vector that minimizes the error between the actual output vector of the classifier and the target output vector. Another method depends on minimizing the mean-squared error (MSE) using the pseudoinverse of the matrix constructed from the feature vectors [25]. In this report, the second method is used to design the linear classifier.

To build the linear classifier, we need first to build the weight vector that will be used to map the input (target free or with a target) to the output (0 or 1 respectively).

This is done by using the linear classifier discriminant function shown in (15), however, to allow for the computation of the constant bias value, w_o, a column of ones is added to the feature vector $\boldsymbol{x}$. After building the weight vector, $\boldsymbol{w}$, we need to test the classifier for different target locations and with different noise and jamming power (different SINR).

2.4.3 Second-order polynomial classifier. The polynomial kernel looks not only at the given features of input samples to determine their similarity, but also at combinations of these [26]. In the context of regression analysis, such combinations are known as interaction features. The (implicit) feature space of a polynomial kernel is equivalent to that of polynomial regression, but without the combinatorial blowup in the number of parameters to be learned [27]. The discriminant function used in polynomial classifiers is the same as the one used for linear classifiers, shown in (15). The only difference here is that we apply polynomial expansion on the feature vector $\boldsymbol{x}$. To illustrate how polynomial expansion is done, assume the feature matrix, $\boldsymbol{x}$, is given by:

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x_{11}} & \boldsymbol{x_{12}} \\ \boldsymbol{x_{21}} & \boldsymbol{x_{22}} \end{bmatrix} \tag{16}$$

Then, the polynomial expansion of, $\boldsymbol{x}$, can be written as:

$$x_p = \begin{bmatrix} \boldsymbol{1} & \boldsymbol{x_{11}} & \boldsymbol{x_{12}} & \boldsymbol{x_{11}x_{12}} & \boldsymbol{x_{11}}^2 & \boldsymbol{x_{12}}^2 \\ \boldsymbol{1} & \boldsymbol{x_{21}} & \boldsymbol{x_{22}} & \boldsymbol{x_{21}x_{22}} & \boldsymbol{x_{21}}^2 & \boldsymbol{x_{22}}^2 \end{bmatrix} \tag{17}$$

This polynomial expansion is a second-order (quadratic) expansion. After setting the desired output vector $\boldsymbol{y}$, the filter weights vector, $\boldsymbol{w}$, can be obtained by:

$$\boldsymbol{w} = \boldsymbol{x_p}^{\dagger}\, \boldsymbol{y} \tag{18}$$

where $\boldsymbol{x_p}^{\dagger}$ is the pseudoinverse of the polynomial expansion of the data matrix $\boldsymbol{x}$.

Chapter 3. Proposed System

In this chapter, the proposed system for target detection is presented. First, the DOA estimation technique selection and the classifier type adopted are presented. Then, the method of operation of the proposed system is described. Finally, another proposed cascade system that involves a significant dimensionality reduction in the training and testing data with a comparable performance is introduced.

3.1 DOA Estimation Technique Selection

In this report, the MUSIC estimator is adopted as a part of the proposed system. It was found that this algorithm has a sufficient estimation of the DOA of the jamming signals since it mainly depends on the interference subspace to estimate the DOA of a given signal. Knowing an estimation of the DOA of the jamming signals will help us in choosing the appropriate classifier needed to detect the presence of a target.

3.2 Classifier Selection

Different types of classifiers such as the linear classifier, second-order polynomial classifier and support vector machine (SVM) were tested to be included in the proposed system. Our selection criteria are based on three different factors which are: (A) Computational Complexity, (B) Time Consumption, and (C) Classification Performance (accuracy). The linear classifier was found to achieve a significant performance as opposed to the other two classification techniques since it does not apply any feature processing before building the classifier weight vector. Linear classification also involves less computational complexity and consumes less time; therefore, it is selected to do the target classification task in both the proposed system and the proposed cascade system.

3.3 Proposed System Methodology

After studying the conventional adaptive beamforming method and applying it as a target detection technique, the following limitations are observed:

1. The amount of secondary data (that are target free) needed to train the weight vector online for different scenarios of SINR and DOA of the target and jamming signals is usually unavailable in practice.
2. The ability of handling variations in the DOA of the target and jamming signals between the training and testing stages is inefficient.

3. The ability to differentiate between the target and the jamming signal when that jamming DOA is close to the target DOA is somehow weak.

Therefore, a new system that can improve the previously mentioned weaknesses is introduced in this report. The proposed system consists of two stages: A) Training stage. B) Testing stage. In the training stage, first there is need for collecting different training data sets that represent the target at a certain DOA and an overlapping range of jamming DOAs. The maximum range of jamming DOAs that can be injected in each data set while keeping the system performance at its maximum was studied and found to be up to 6 different jamming DOAs. Data collected in this step is balanced, which means that 50% of the samples must contain information about the target, and the other 50% must be target free to allow the classifiers built using such data sets to differentiate between the target and no target cases. Since such data is complex by nature, the researcher preferred to split each feature vector into a real and imaginary component and concatenate them in one features' matrix, as shown in (19), where each features' matrix will hold information about a certain target DOA and a range of 6 jamming DOAs. This step takes place due to the need for reducing the computational complexity and time consumption as opposed to using the original complex received data without affecting the output performance of the system.

$$\boldsymbol{X_d} = \begin{bmatrix} x_{r11} & x_{r12} & x_{r13} & x_{r14} & \cdots & x_{r1K} \\ x_{r21} & x_{r22} & x_{r23} & x_{r24} & \cdots & x_{r2K} \\ x_{r31} & x_{r32} & x_{r33} & x_{r34} & \cdots & x_{r3K} \\ x_{r41} & x_{r42} & x_{r43} & x_{r44} & \cdots & x_{r4K} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{rN1} & x_{rN2} & x_{rN3} & x_{rN4} & \cdots & x_{rNK} \\ x_{i11} & x_{i12} & x_{i13} & x_{i14} & \cdots & x_{i1K} \\ x_{i21} & x_{i22} & x_{i23} & x_{i24} & \cdots & x_{i2K} \\ x_{i31} & x_{i32} & x_{i33} & x_{i34} & \cdots & x_{i3K} \\ x_{i41} & x_{i42} & x_{i43} & x_{i44} & \cdots & x_{i4K} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{iN1} & x_{iN2} & x_{iN3} & x_{iN4} & \cdots & x_{iNK} \end{bmatrix} \tag{19}$$

where x_{r43} represents a data point from the real component that belongs to the 4th feature and to the 3rd sample. While x_{i43} represents a data point from the imaginary component that belongs to the 4th feature and to the 3rd sample.

Second, linear classification is used to build different weight vectors (classifiers) using the different features' matrices mentioned before. Each classifier built will hold information about a certain target DOA and a range of jamming DOAs. Each range of jamming DOAs overlaps its adjacent ranges and consists of 6 different angles with a linear spacing of 2 degrees which allow up to 10 degrees DOA deviation. This is done to account for any inaccurate estimations taken by the MUSIC DOA estimator. If work is done in the range of angles in the first quadrature, i.e. angles between 0 and 90 degrees, there will be need for 17 different classifiers to account for the information of all the possible jamming DOAs. Luckily, the training stage can be done offline, and the 17 different classifiers can be saved for later use (testing stage). The training stage can be summarized as shown in Figure 5.

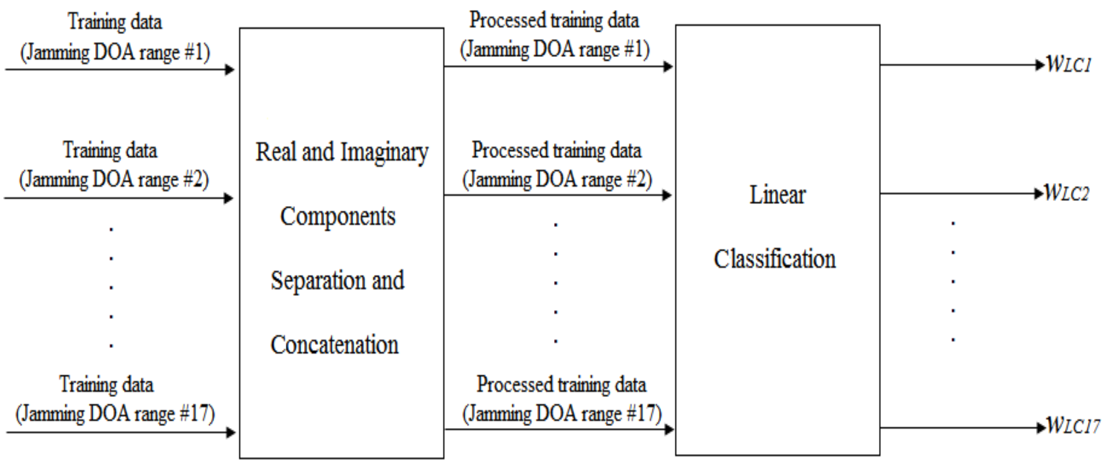

Figure 5: Proposed system training stage.

In the testing stage, the testing data must first be processed in the same manner as the training data; however, the restriction of balanced data mentioned in the training stage does not apply here. The testing data may have a target present at all samples; no target at any sample or target present at a certain range of samples (range gate). Second, the processed testing data will be passed to the MUSIC DOA estimator to get an estimation of the jamming DOA. Based on this estimation, the processed testing data will be directed to all suitable classifiers that have the estimated jamming DOA in their ranges. We can describe this step as a classifier selector step where all the suitable selected classifiers will be tested against the processed testing data, and the classifier that gives the best target detection accuracy is chosen. The testing stage can be summarized as shown in Figure 6.

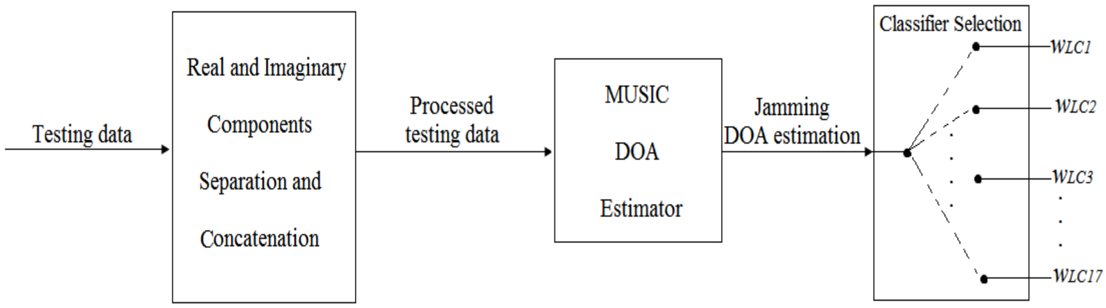

Figure 6: Proposed system testing stage.

3.4 Proposed Cascade System with Adaptive beamforming as a Dimensionality Reduction Technique

The proposed cascade system refers to the use of adaptive beamforming in line with the linear classification technique. The training stage in this system consists of collecting balanced training data sets that have the same specifications as in the previous proposed system. The only difference is that real components are used, and they are separate from the imaginary components of the training features' matrix to reduce the computational complexity and time consumption as mentioned earlier. Next, the processed training data will be projected onto the adaptive beamforming weights that have been trained on target free secondary data. After that, the real and imaginary components of the adaptive beamforming output will be used to train the different linear classifier weight vectors (classifiers). This step reduces the training data dimensionality for each component to be $1 \times K$ where K is the number of samples used. This means that the overall dimensionality used to train the linear classifier will be $3 \times K$, which represents one feature vector for the real components of the training features' matrix, one feature vector for the imaginary components and one feature vector for the constant bias value needed for the linear classification step. This method of dimensionality reduction is considered as an unsupervised method since the adaptive beamforming which is used to reduce the dimensionality does not include class labels in its weight calculation process. Figure 7 shows a summary of the proposed cascade system training stage.

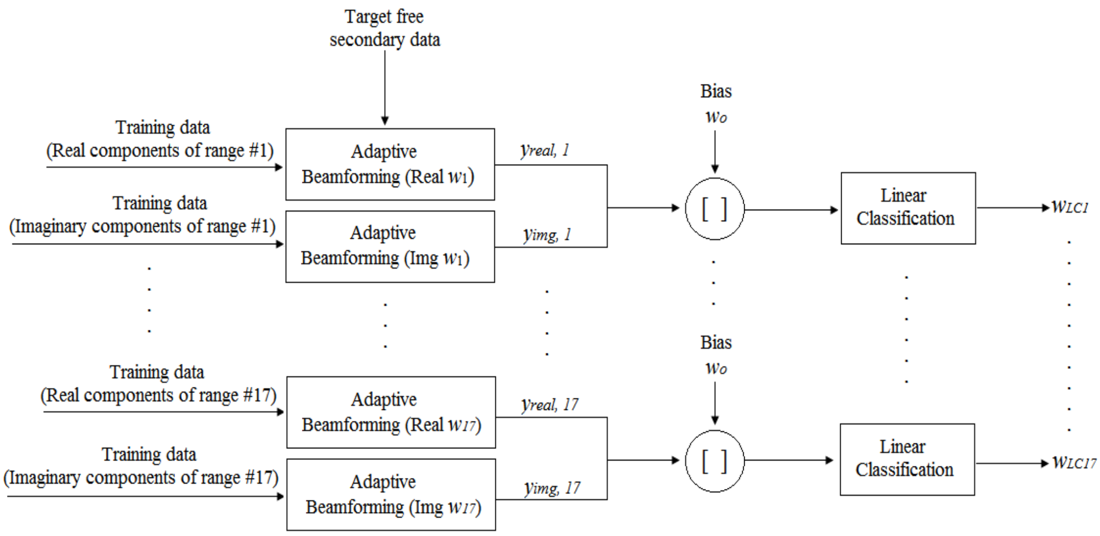

Figure 7: Proposed cascade system training stage.

In the testing stage, the same testing data specifications used in the previous proposed system are considered. Next, the processed testing data is passed to the MUSIC DOA estimator to get an estimation of the jamming signal DOA. Based on this estimation, an appropriate adaptive beamforming weight vector will be chosen to project the testing data real and imaginary separated components on. After that, the output of the adaptive beamforming step will be passed to all suitable linear classifiers, and the one that gives the best target detection accuracy will be chosen to contribute in the output of this system. Figure 8 shows a summary of the proposed cascade system testing stage.

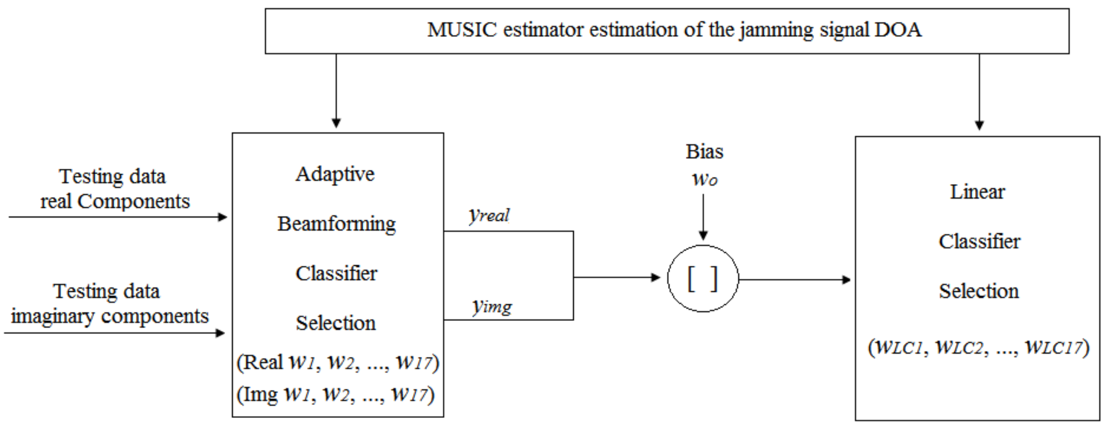

Figure 8: Proposed cascade system testing stage.

Chapter 4. Results and Analysis

In this chapter, the simulation results obtained for all the target detection systems mentioned in this thesis are presented. The first part is devoted to the decision making criteria used for target detection. The second part is dedicated to an output evaluation for the MUSIC DOA estimator. Next, a comparison between the conventional adaptive beamforming method and the proposed system is conducted. Finally, the proposed cascade system simulation results with adaptive beamforming as a dimensionality reduction technique are demonstrated. All the training and testing parameters used in this chapter are illustrated in Table 1 unless it is mentioned otherwise.

Table 1: Parameters values used to obtain the simulation results.

Parameter	Parameter definition	Value used
N	Number of the ULA channels.	16
K_{tr}	Number of samples used for training (offline).	300,000
K_{ts}	Number of samples used for testing (online).	200,000
Θ_T	Target DOA.	40 degrees
Θ_J	Jammer DOA.	All possible angles in the first quadrature with a linear spacing of 2 degrees.
SINR	Signal to interference plus noise ratio for both training and testing data.	-26 dB

4.1 Testing and decision Making Criteria

In the process of testing each proposed system, a threshold value is measured using the intersection point of the target and no target output of each technique. The output at each sample location of each target detection method is then compared to that threshold value. If the output exceeds the threshold value, the target is said to be present at that specific location, and it is given a new label equal to 1; otherwise, it is said to be not present and given a new label equal to 0. After that, the new predicted labels are compared with the actual labels (that includes information about the presence of the target and its range gate location), and an accuracy value is calculated. The threshold value measured for each of the adaptive beamforming technique and the linear classification technique is shown in Figure 9 and Figure 10 respectively. For the adaptive beamforming technique, the threshold value is found to be equal to 24.5 μVolts while for the linear classifier it is found to be 0.5. It is worth to mention that the linear classifier output is dimensionless, hence, its threshold value is dimensionless.

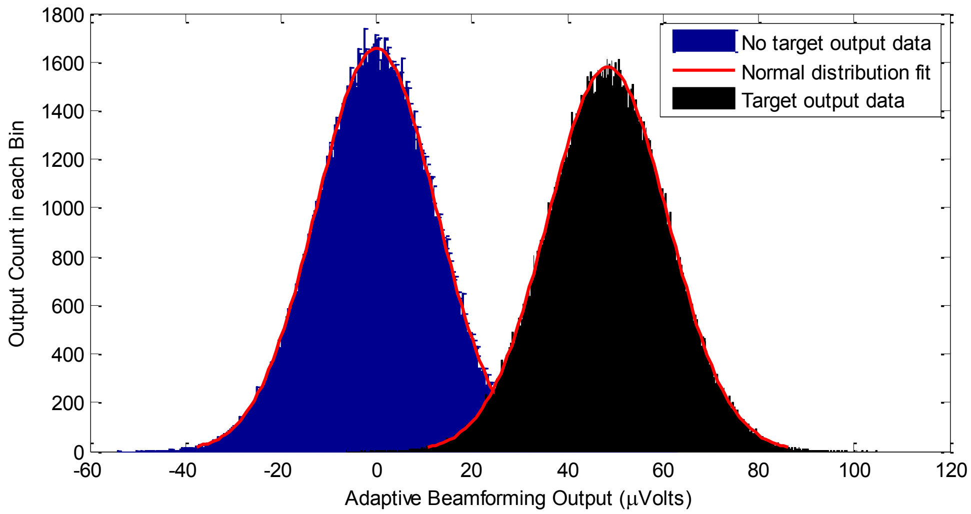

Figure 9: Threshold measurement for the adaptive beamforming technique output.

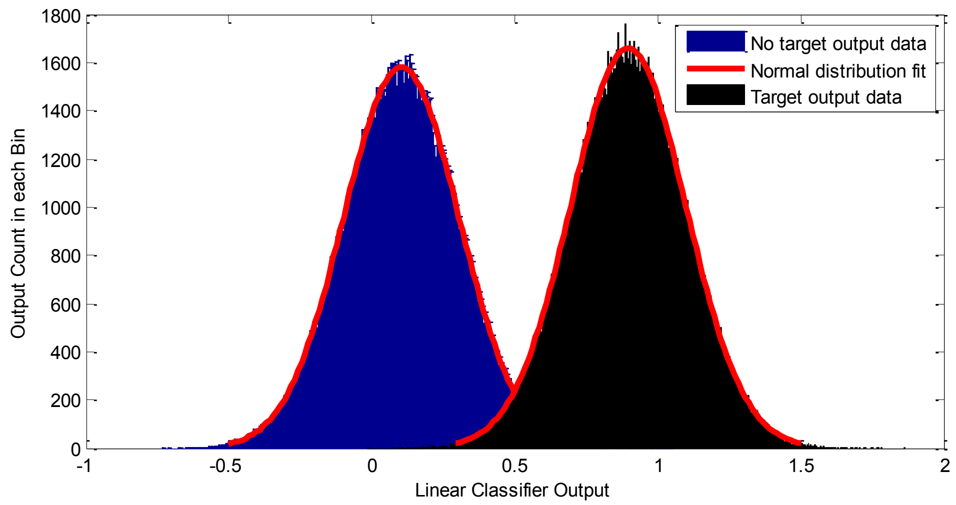

Figure 10: Threshold measurement for the linear classification technique output.

4.2 MUSIC DOA Estimator Output Evaluation

The jamming signal DOA estimation is considered as the first step in the testing stage after processing the received data. Therefore, steps must be taken to ensure that this estimation process leads to all the suitable classifiers that come in the second step of the testing process. Any wrong estimation of the jamming signal DOA will lead to a wrong classifier selection which will affect the target detection process negatively. From here, the idea of training different classifiers using an overlapping range of the jamming signal DOA was adopted to tolerate any inaccurate estimations. Figure 11 shows the jamming signal DOA estimation confusion of the MUSIC estimator.

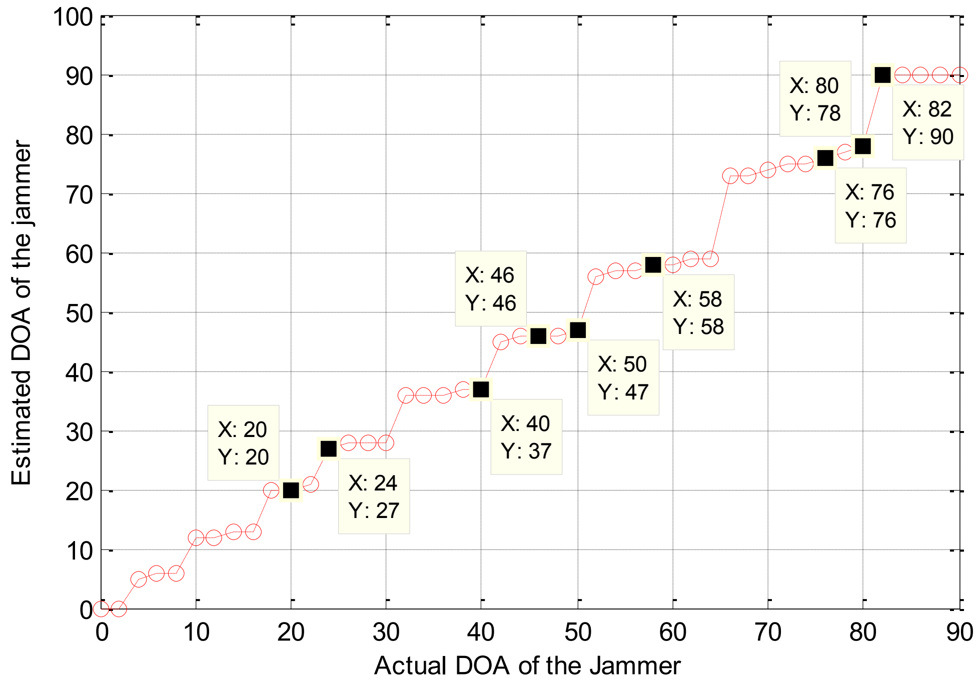

Figure 11: MUSIC jamming signal DOA estimation confusion.

As shown in Figure 11, the maximum inaccuracy in the jamming signal DOA estimation is detected at angle 82° where it is estimated to be equal to 90° which means that the estimation is 8 degrees apart from the actual jamming DOA. Hence, the range of jamming signal DOAs injected in each classifier is determined to be up to 6 angles, as mentioned before, that are apart from each other by 2 degrees. This means that these selected overlapping ranges can tolerate up to a 10 degrees' difference between the actual and estimated jamming signal DOA without affecting the overall performance of the target detection process. That accuracy of the MUSIC DOA estimator against the selected overlapping ranges for each of the jamming signal DOA when selecting the classifier that gives the best target detection accuracy among all the other suitable classifiers selected by the MUSIC DOA estimator is ensured to give a 100%

performance all the time which means that the correct classifier will be chosen everytime to contribute in the overall performance of the system.

4.3 Conventional Adaptive Beamforming vs. Proposed system

In this section, the proposed system for target detection is evaluated against the conventional adaptive beamforming technique using two different tests illustrated in the following sub-sections.

4.3.1 Range gate detection test. This test measures the ability of each target detection technique to detect the range gate bins which correspond to a certain distance on the ground at which the target signal appears. This test is applied in two different cases: the first case is when the jammer DOA is fairly apart from the target DOA, while in the second case, the jammer DOA is relatively close to the target DOA. For both cases, the target signal is present in the range gate bins between 100,000 and 150,000. Both the conventional adaptive beamforming and the proposed system are trained using the parameters values mentioned in Table 1 while the SINR is kept constant in both training and testing stages. For testing, Table 2 shows the DOA values chosen for the two different testing cases. Figure 12 and Figure 13 show the first case output of this test for both the adaptive beamforming technique and the proposed system respectively. As we can see, both techniques perform well when that jammer DOA is fairly apart from the target DOA.

Table 2: Range gate first and second testing cases DOA values.

Parameter symbol	Value used
First case:	
Θ_T	40 degrees
Θ_J	20 degrees
Second case:	
Θ_T	40 degrees
Θ_J	32 degrees

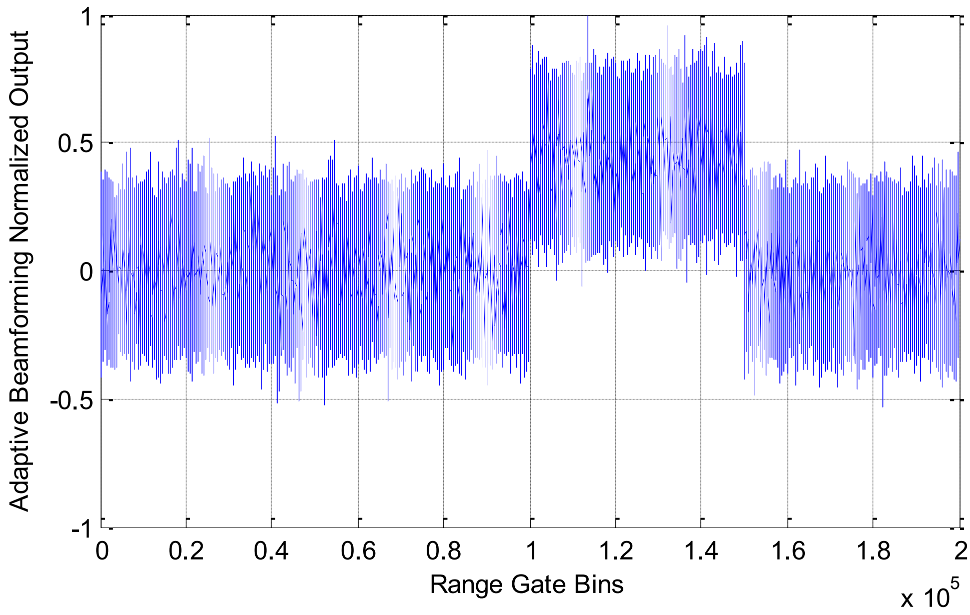

Figure 12: Adaptive beamforming range gate test (case 1).

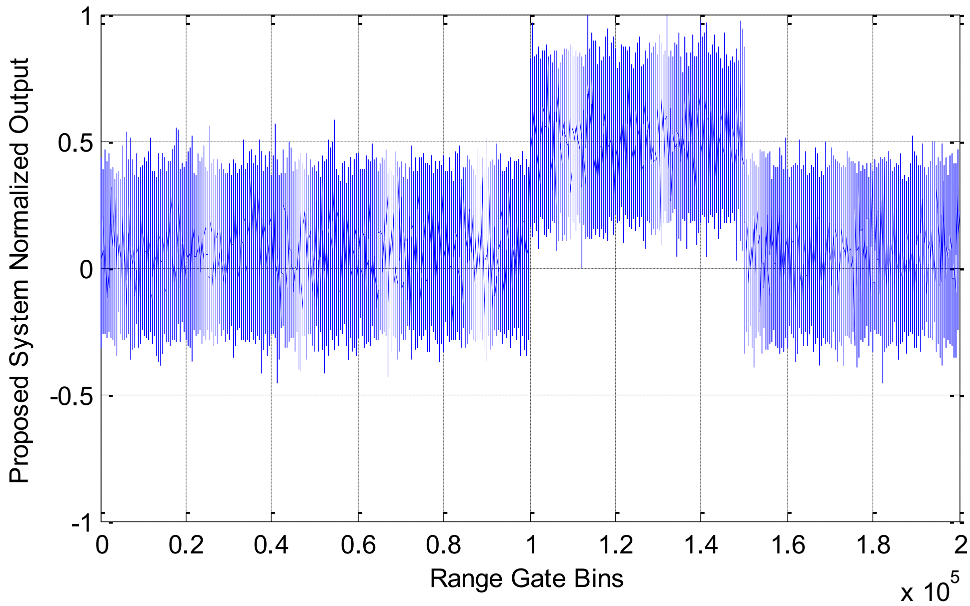

Figure 13: Linear classification range gate test (case 1).

On the other hand, Figures 14 and 15 show the second case output for both the adaptive beamforming technique and the proposed system respectively. Here, the adaptive beamforming technique performs worse and shows less immunity to interference as compared to the target detection proposed system. In the adaptive beamforming case, the target range gate bins can be barely differentiated from the other range gate bins; however, in the proposed system case, the target range gate bins can be fairly located.

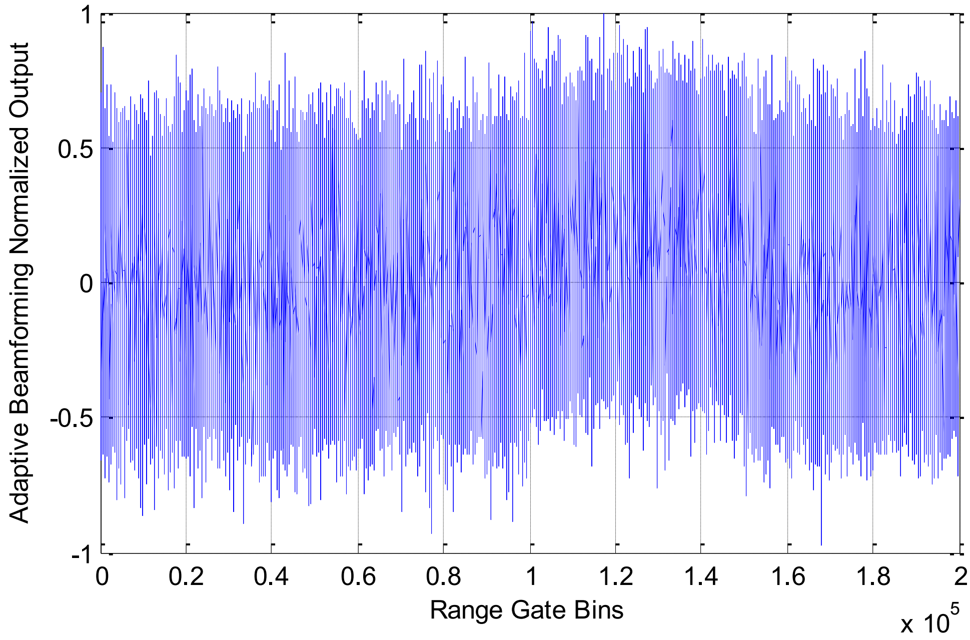

Figure 14: Adaptive beamforming range gate test (case 2).

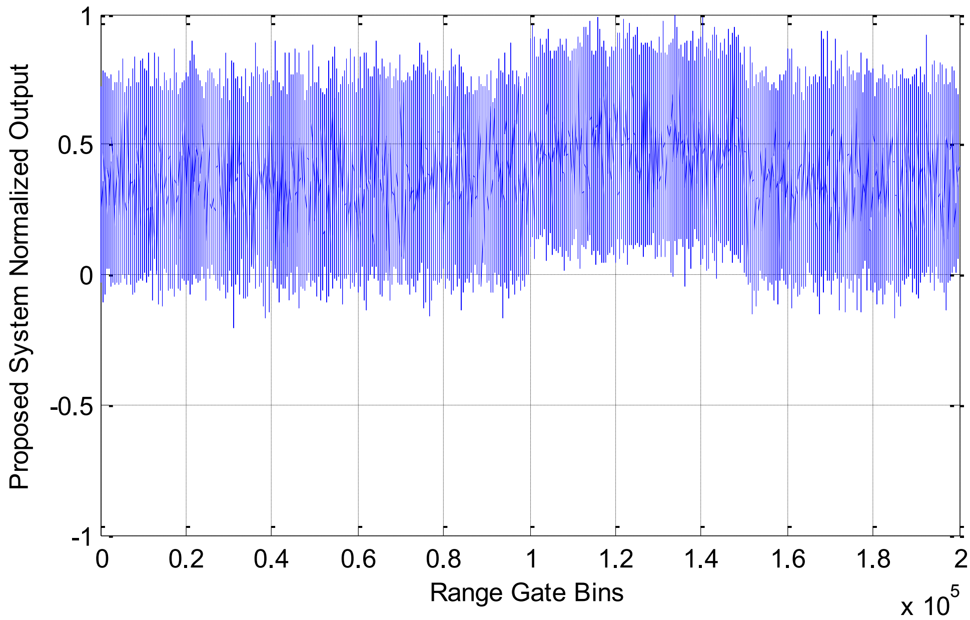

Figure 15: Linear classification range gate test (case 2).

4.3.2 Jammer DOA variations and target detection test. In some scenarios where the jamming signal location is not stationary, the jammer DOA keeps changing continuously. For such scenarios, the target detection process is affected by the jammer DOA depending on how close it is to the target DOA. Hence, this test measures the capability of each system to handle variations in the jammer DOA between the training and testing stages. To perform this test, the target DOA is chosen to be at 40 degrees for all data points while the jammer DOA is swept over all possible angles in the first quadrature as mentioned in Table 1.

Figure 16 shows the target detection accuracy of the adaptive beamforming technique that is trained on one jamming DOA scenario, which is at 20 degrees, and tested over all other jamming DOA scenarios with the existence of the target signal at all testing data points. As illustrated by Figure 16, whenever the jammer DOA used for testing is not the same as the one used for training, the target detection accuracy is worthless except for few other jamming DOAs. This result is considered as the base line from which the target detection output starts improving.

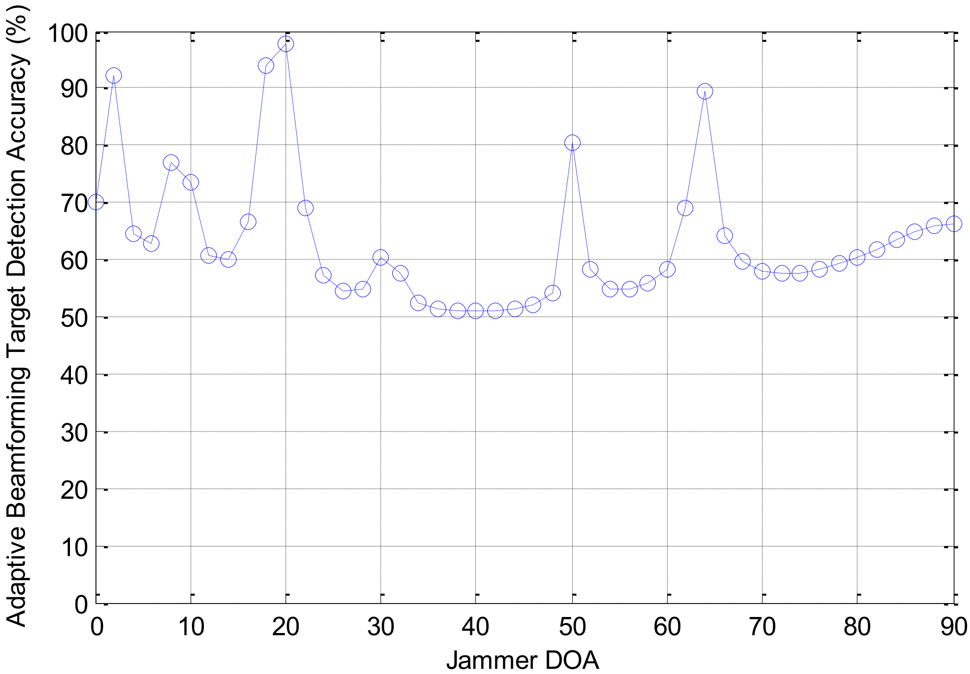

Figure 16: Adaptive beamforming target detection accuracy (data with target).

Figure 17 shows the same test with the same training and testing conditions except that the target signal does not exist in any of the testing data points. Similarly, the system is sure that no target exists whenever the training and testing conditions match; however, it appears uncertain whenever the conditions are different.

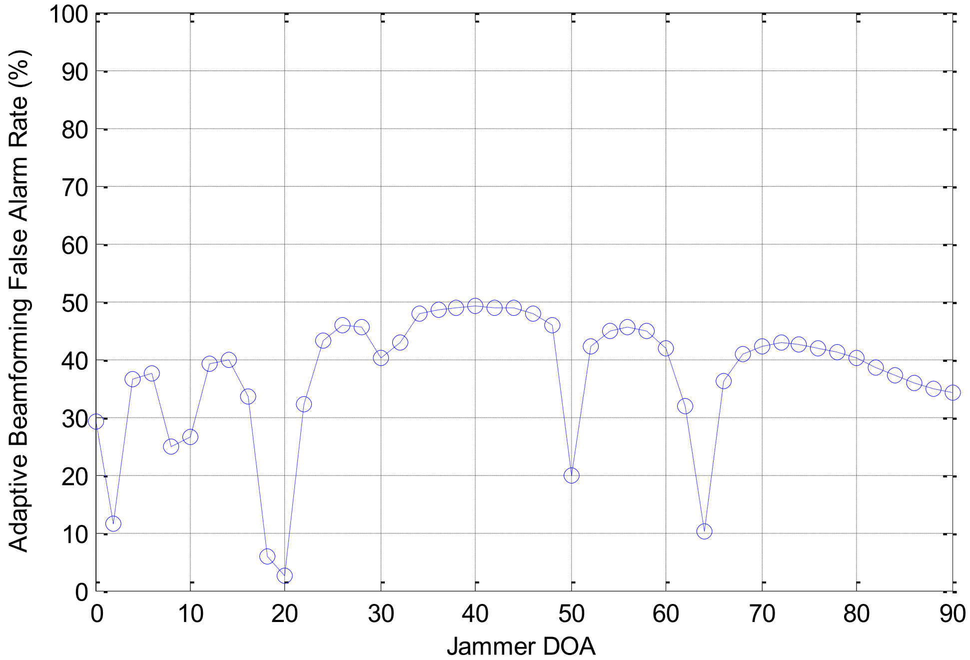

Figure 17: Adaptive beamforming false alarm rate (data with no target).

The proposed system improved the target detection accuracy for this test significantly, since it uses multiple classifiers trained on all possible scenarios of jamming signal DOA. The same exact testing conditions are applied on the proposed system. The first is achieved by using a linear classifier to do the classification job, as mentioned in section 3.3. The second is maintained by using an adaptive beamformer instead of the linear classifier. The result of this test for both systems, when the target signal exists at all testing data points, is shown in Figure 18.

As shown in Figure 18, both systems perform in the same way whenever the jammer DOA is approximately 10 degrees or more apart from the target DOA; however, when they are less than 10 degrees apart, the proposed system with linear classification outperforms the proposed system with adaptive beamforming by a maximum accuracy difference of 18%, which occurs at 38 degrees of jamming DOA, and a minimum accuracy difference of 4% which occurs at 44 degrees of jamming DOA. The minimum target accuracy dip, in the proposed system with linear classification, happens when the target and the jammer DOAs are exactly the same which makes the target classification job nearly impossible, yet, it gives us the information about the target DOA which is in this case equal to 40 degrees.

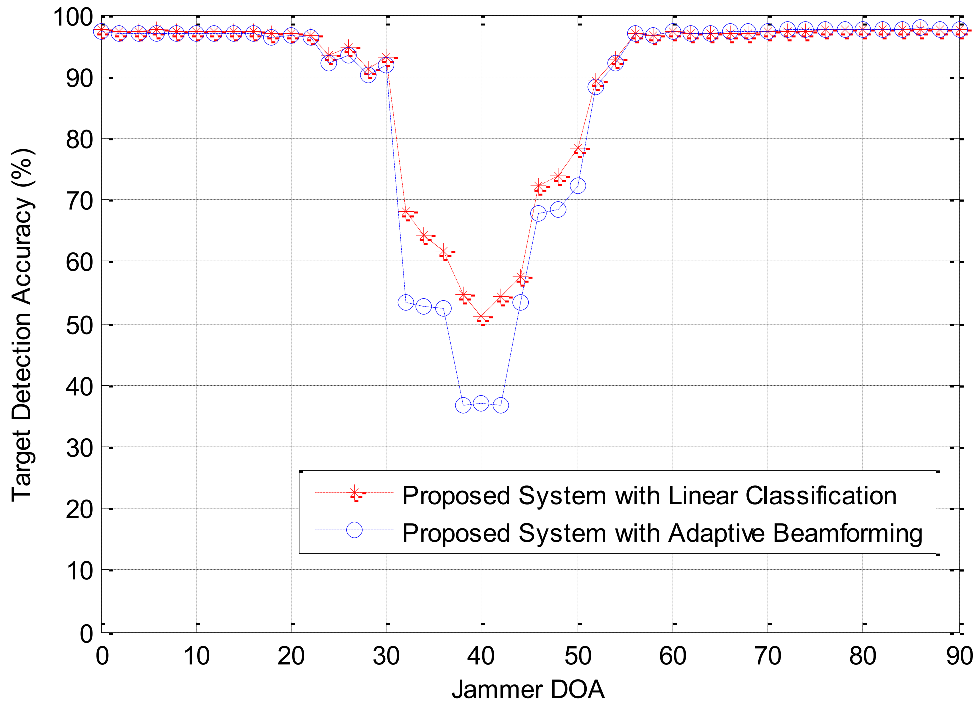

Figure 18: Proposed systems target detection accuracy for the jamming DOA variations test (data with target).

In contrast, the proposed system with adaptive beamforming minimum target detection accuracy dip is not clear. Therefore, it cannot be used to specify the target DOA. This comes from the fact that the adaptive beamforming beam pattern is wider than the linear classification beam pattern. This fact can be proven by training both methods on a single target DOA and testing them on all possible target DOAs as shown in Figure 19 which illustrates that the adaptive beamforming method doesn't only suffer from a wider beam pattern, as compared to the linear classification method, but also from undesirable side lobes.

Figure 20 shows the other side the jamming DOA variations test, which is when all the testing data points do not include a target signal. We can notice from Figure 20 that the proposed system with adaptive beamforming is doing slightly better than the proposed system with linear classification. However, since this test is for the no target case, the false alarms here are insignificant as compared to the false alarms in the target existence case. Generally, from the test results obtained in this section, the proposed system with linear classification is considered the best proposed system for target detection among all the other systems discussed in this thesis. The assumption that any target detection accuracy that is above 50% is an indication of the presence of a target and anything below it is an indication of the absence of the target leads to the conclusion that the proposed system is superior for jamming DOA variations.

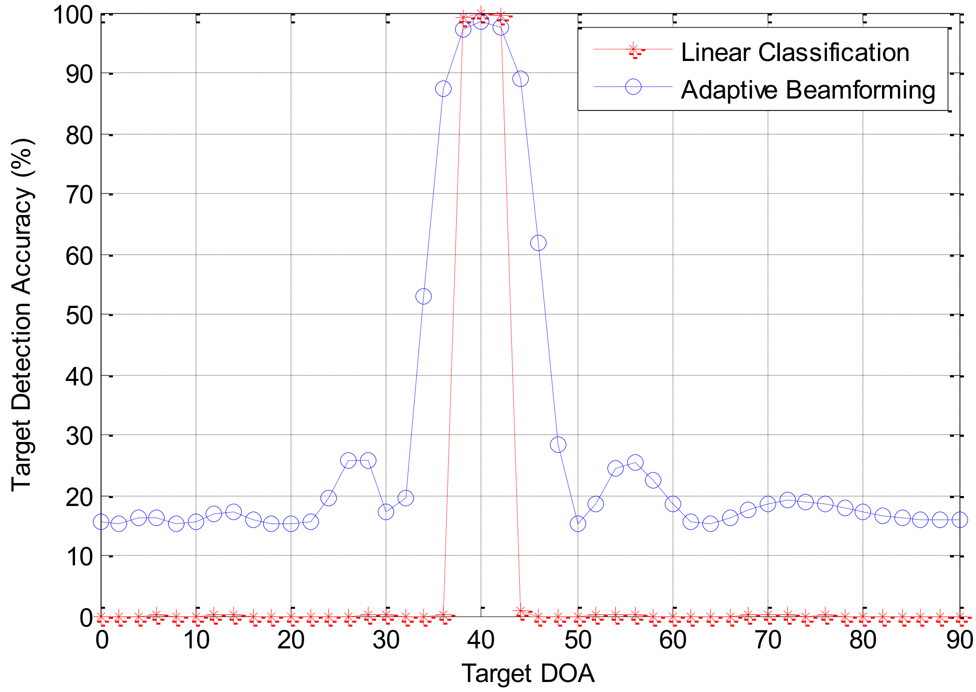

Figure 19: Beam pattern width of the adaptive beamforming and linear classification techniques.

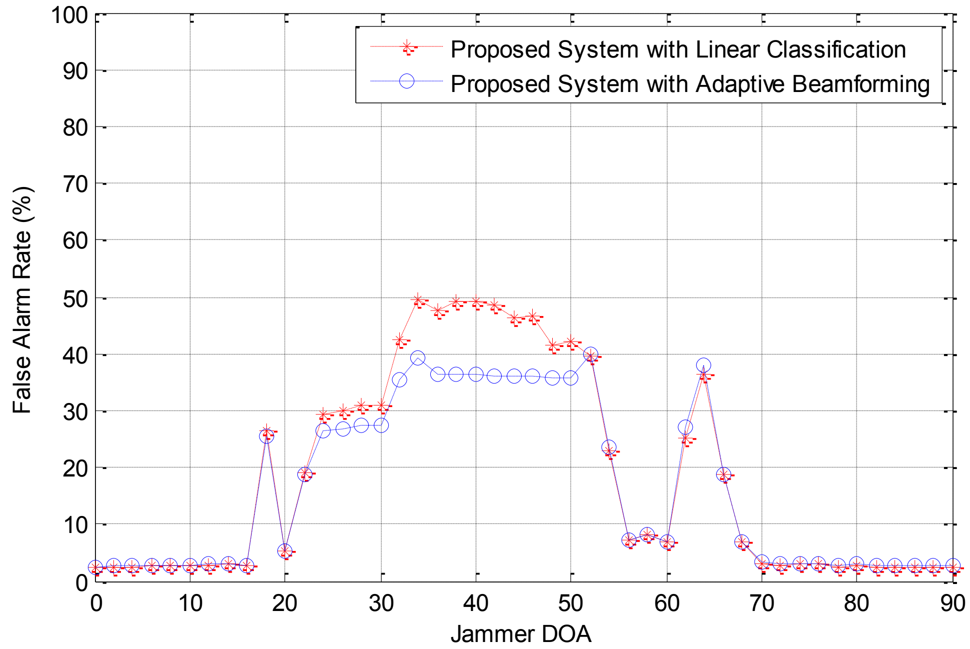

Figure 20: Proposed systems false alarm rate for the jamming DOA variations test (data with no target).

4.4 Proposed Cascade System Evaluation

Although the dimensionality dealt with in this paper is not relatively high, it is always preferable to deal with a lower dimensional data as long as the performance is not affected. The result of using PCA, for dimensionality reduction, in line with the proposed system for both target and no target cases is shown in Figure 21. As we can see, PCA projected data are useless since they put the classifier into a confusion mode between the target and other interferences. This result is expected since PCA does not incorporate class information as mentioned earlier.

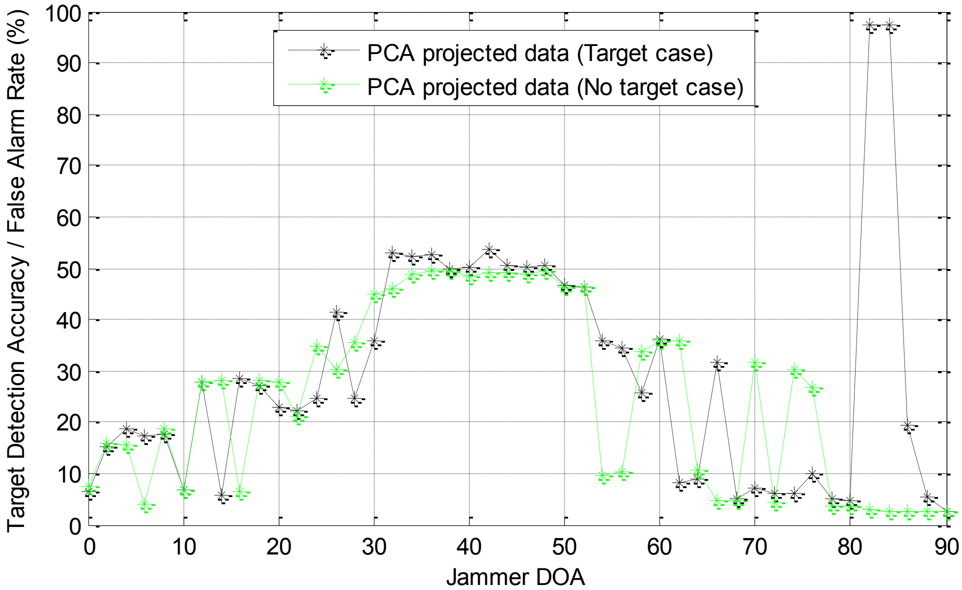

Figure 21: Target detection accuracy / False alarm rate for the jamming DOA variations when using PCA projected data in line with the proposed system.

Figure 22 shows the target detection accuracy for target and no target cases of the proposed cascade system, discussed in section 3.4, which uses the adaptive beamforming output to reduce the dimensionality of the original data. This will make the followed linear classification step faster and reduce its computational complexity. As shown in Figure 22, the performance of the proposed cascade system (target and no target case) is approximately the same as the performance of the proposed system with linear classification, shown previously in Figures 18 and 20, with relatively insignificant accuracy drops (in the target case) and an increase (in the no target case) at the jamming angles around 40 degrees where the target is located.

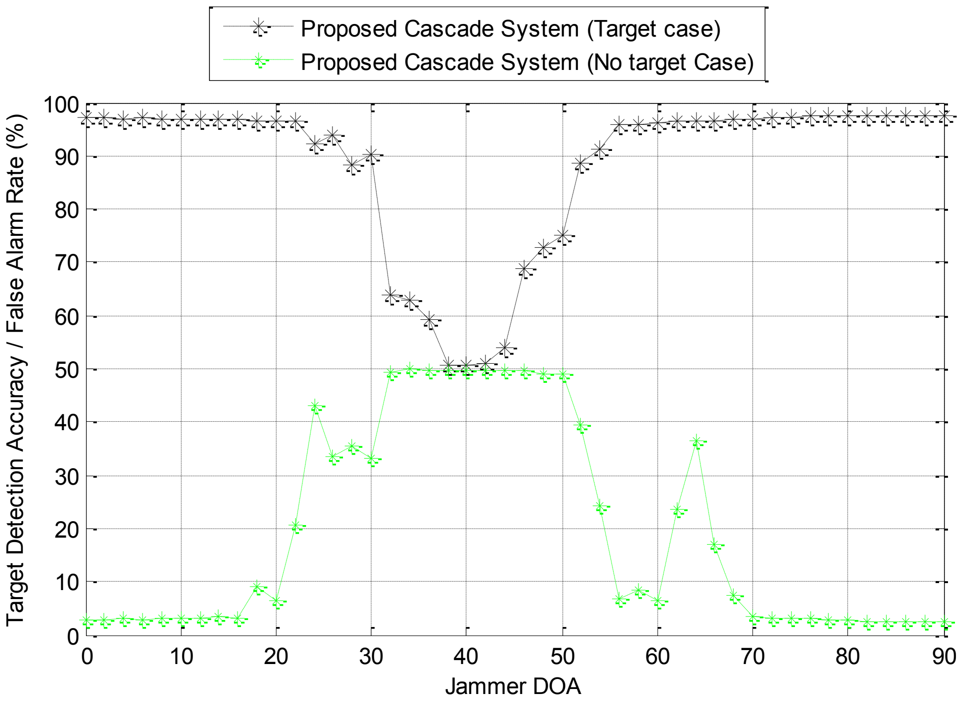

Figure 22: Target detection accuracy / False alarm rate for the jamming DOA variations in the proposed cascade system.

4.5 Computational Complexity Analysis

To measure the computational complexity of each system, the running time is measured and recorded in Table 3. The training stage running time is not taken into consideration for the measurements recorded in Table 3 since this stage is done offline and do not have to be repeated everytime a test is done. The software used to perform all the simulation results is MATLAB R2014a and the machine specifications are listed in Table 4.

Table 3: Running time of each proposed algorithm.

Algorithm used	Running time (sec)
Proposed system with adaptive beamforming.	1.4374
Proposed system with linear classification.	0.8511
Proposed cascade system.	45.0944

Table 4: Machine specifications.

Specification	**Value**
Processor	Intel® Core™ i7-4500U @ 1.80 GHz – 2.40 GHz
RAM	8 GB
Operating system	Windows 8.1 Single Language
System type	64-bit operating system

Chapter 5. Conclusion

A new target detection algorithm based on jamming signal DOA estimation and learning based techniques was introduced in this thesis. It was shown that this proposed algorithm was able to improve several problems that the conventional method suffers from. These problems include the jammer DOA variations between the training and testing stages and the necessity for target-free secondary data for training.

Also, a new unsupervised method of dimensionality reduction through the use of the adaptive beamforming output was introduced. Like PCA, this method does not incorporate the different class information (labels). However, it helps in classification purposes as opposed to PCA which is not always good for such applications. This technique of dimensionality reduction was used in the proposed cascade system which showed approximately an equivalent output performance to the proposed system with linear classification.

It must be said that the proposed system is not considered superior in all situations. As shown using the simulation results, the target DOA is assumed to be fixed due to the assumption of a stationary radar platform that is always directed to the expected target DOA. However, in many practical cases, the target DOA is not fixed but is rather changing continuously. The proposed system can overcome this problem if the training stage is done online, yet, this will increase the computational complexity of the system.

References

[1] S. W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*, 2nd ed. San Diego, CA: California Technical, 1997.

[2] J. C. Whitaker, *The Electronics Handbook*. Boca Raton, FL: CRC Press, 1996.

[3] M. A. Richards, J. A. Scheer and W. A. Holm, *Principles of Modern Radar: Basic Principles*. Raleigh, NC: Scitech, 2010.

[4] V. Rabinovich and N. Alexandrov, "Typical Array Geometries and Basic Beam Steering Methods," in *Antenna Arrays and Automotive Applications*, New York: Springer, 2013, pp. 23-54.

[5] J. Ward, "Space-time adaptive processing for airborne radar," Lincoln Laboratory, MIT, Massachusetts, Tech. Rep. 1015, Dec. 1994.

[6] V. P. Tuzlukov, *Signal Processing Noise*. Boca Raton, FL: CRC Press, 2002.

[7] J. S. Blogh and L. Hanzo, *Third-generation Systems and Intelligent Wireless Networking: Smart Antennas and Adaptive Modulation*. New York: Wiley, 2002.

[8] Q. Li, G. Li, W. Lee, M. i. Lee, D. Mazzarese, B. Clerckx and Z. Li, "MIMO techniques in WiMAX and LTE: A feature overview," *IEEE Commun. Mag.,* vol. 48, no. 5, pp. 86-92, May 2010.

[9] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. Wiley, 1980.

[10] A. E. Khatib, K. Assaleh and H. Mir, "Space-Time Adaptive Processing Using Pattern Classification," *IEEE Trans. Signal Process.,* vol. 63, no. 3, pp. 766-779, Feb. 2015.

[11] B. Liao, C. Guo, L. Huang, Q. Li and H. C. So, "Robust Adaptive Beamforming With Precise Main Beam Control," *IEEE Trans. Aerosp. Electron. Syst.,* vol. 53, no. 1, pp. 345-356, Feb. 2017.

[12] X. Wu, Y. Cai, M. Zhao, R. C. d. Lamare and B. Champagne, "Adaptive Widely Linear Constrained Constant Modulus Reduced-Rank Beamforming," *IEEE Trans. Aerosp. Electron. Syst.,* vol. 53, no. 1, pp. 477-492, Feb. 2017.

[13] M. Zhang, A. Zhang and Q. Yang, "Robust Adaptive Beamforming Based on Conjugate Gradient Algorithms," *IEEE Trans. Signal Process.,* vol. 64, no. 22, pp. 6046 - 6057, Nov. 2016.

[14] J. R. Guerci, *Space-Time Adaptive Processing for Radar*. Norwood, MA: Artech House, 2003.

[15] H. L. V. Trees, *Optimum Array Processing*. New York: Wiley, 2002.

[16] H. Abdi and L. J. Williams, "Principle Component Analysis," *Wiley Interdiscipl. Rev. Comput. Statist.,* vol. 2, no. 4, pp. 433-459, June 2010.

[17] M. Welling, *Fisher linear discriminant analysis*, 2005.

[18] Z. I. Khan, "Performance Evaluation and Analysis of Direction of Arrival Estimation Using MUSIC, TLS ESPRIT and Pro ESPRIT Algorithms," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 4, no. 6, pp. 4948-4958, June 2015.

[19] T. Lavate, V. Kokate and A. Sapkal, "Performance Analysis of MUSIC and ESPRIT DOA Estimation Algorithms for Adaptive Array Smart Antenna in Mobile Communication," in *Second International Conference on Computer and Network Technology (ICCNT)*, Bangkok, 2010, pp. 308-311.

[20] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. New York: Oxford Univ. Press, 1965.

[21] M. Devendra and K. Manjunathachari, "DOA estimation of a system using MUSIC method," in *2015 International Conference on Signal Processing and Communication Engineering Systems*, Guntur, 2015, pp. 309-313.

[22] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[23] G.-X. Yuan, C.-H. Ho and C.-J. Lin, "Recent advances of large-scale linear classification," *Proc. IEEE,* vol. 100, no. 9, pp. 2584-2603, Sept. 2012.

[24] E. Alpaydin, *Introduction to Machine Learning*. Cambridge, MA: MIT Press, 2004.

[25] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*. New York: Wiley, 2001.

[26] C. Robert, "Machine Learning: a Probabilistic Perspective," *CHANCE,* vol. 27, no. 2, pp. 62-63, 2014.

[27] Y. Goldberg and M. Elhadad, "splitSVM: fast space-efficient non-heuristic polynomial kernel computation for NLP applications," *Proc. 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers,* June 2008, pp. 237-240.

Vita

Mohammad Moufeed Sahnoon was born in 1993 and received his primary and secondary education in Sharjah, UAE. He received his B.Sc. degree in Electrical and Electronic Engineering from the University of Sharjah in 2015.

In September 2015, he joined the Electrical Engineering graduate program in the American University of Sharjah as a graduate teaching assistant and completed his M.S. in Electrical Engineering degree requirements in 2017. During his master's study, he participated in the graduate student research conference (GSRC UAE’17) held in Khalifa University in Abu Dhabi, UAE with his thesis findings. His research interests include machine learning, pattern classification and signal processing.