Automatic Modulation Classification in Cognitive Radio Network

Shahrukh Munir
Reg.No:
297-FET/MSEE/F12

Supervisor/Co-Supervisor
Dr. Muhammad Amir
Dr. Sajjad Ahmed Ghauri

Department of Electronic Engineering
Faculty of Engineering and Technology, (FET)
International Islamic University, (IIU), Islamabad.

Cognitive radio networks.
Wireless communication systems.

Certificate of Approval

Title of Thesis: **"Automatic Modulation Classification in Cognitive Radio Networks"**

Name of Student: Mr. Shahrukh Munir

Registration No.: 297-FET/MSEE/F12

Accepted by the Faculty of Engineering and Technology, INTERNATIONAL ISLAMIC UNIVERSITY, in partial fulfillment of the requirement for the Master of Philosophy Degree in Electronic Engineering.

Viva Voce Committee

Supervisor
Prof. Dr Muhammad Amir
DEE, FET, IIUI

Co-Supervisor
Dr. Sajjad A. Ghauri
DEE, FET, IIUI

Internal Examiner
Dr. Ihsan ul Haq
DEE, FET, IIUI.

External Examiner
Dr. Hafiz Muhammad Faisal Zafar
Principal Scientist

M F. Zafar

Chairman
Prof. Dr Muhammad Amir
DEE, FET, IIUI.

Dean
Prof. Dr. Aqdas Naveed Malik
FET, IIUI.

بسم الله الرحمن الرحيم

رَبَّنَآ ءَاتِنَا فِى ٱلدُّنْيَا حَسَنَةً وَفِى ٱلْءَاخِرَةِ حَسَنَةً وَقِنَا عَذَابَ ٱلنَّارِ ﴿٢٠١﴾

Our Lord! Grant us good in this world

and good in the life to come

and keep us safe

from the torment

of the Fire.

Qur'an (2:201)

رَبَّنَا وَلَا تُحَمِّلْنَا مَا لَا طَاقَةَ لَنَا بِهِۦ ۖ وَٱعْفُ عَنَّا وَٱغْفِرْ لَنَا وَٱرْحَمْنَآ ۚ
أَنتَ مَوْلَىٰنَا فَٱنصُرْنَا عَلَى ٱلْقَوْمِ ٱلْكَٰفِرِينَ ﴿٢٨٦﴾

Our Lord! Impose not on us that which

we have not the strength to bear,

grant us forgiveness and have

mercy on us. You are our

protector. Help us

against those

who deny

the truth.

Qur'an (2:286)

This page is intentionally left blank

Abstract

The automatic modulation classification (AMC) is a hot research area in today communication system. AMC has vast application both in military and civilian application. In intelligent communication systems such as software defined radios networks and cognitive radio networks, AMC is the most important issue. In this thesis, AMC is performed by using pattern recognition approach. Higher order statistical features are selected for classification of different modulation schemes. Support vector machine (SVM) and feed forward back propagation neural network (FFBPNN) classifier are used for classification of the signals. Different QAM's modulation are classified in this research work. Some channel impairments are taken into consideration such as additive white guassian noise (AWGN) and flat fading such as Rayleigh and Rician. The comparison is made on classification accuracy on different channels for different number of samples at different SNR's. The proposed classifier, classification accuracy is optimized with one of the heuristics computational technique i.e. Genetic Algorithm. The simulation results are compared with and without optimization and also compared with the state of the art existing techniques.

DEDICATION

"I want to dedicate this work to Holy Prophet Muhammad (Peace be upon him) and his companions who laid the foundations of Modern civilization and paved the way for social, moral, political, economic, cultural and physical revolution"

I also thank my parents, brothers and sisters for their nerve ending moral support and prayers which always acted as a catalyst in my academic life.

Acknowledgement

The successful completion of this dissertation was made possible through the invaluable contribution of a number of people. To say “thank you “to all of you is not even enough to express our gratitude. You are all angel sent to us by ALLAH.

I would like to express our deep and sincere gratitude to our supervisor, Dr. Muhammad Amir, Chairman of Department of Electronics Engineering, International Islamic University, and Islamabad. His wide knowledge and his logical way of thinking have been of great value for me. His understanding, encouraging and personal guidance have provided me a good basis to present my thesis.

I deeply grateful to my co-supervisor Dr. Sajjad Ahmad Ghauri, for his detailed and constructive comments, and for his important support throughout this work.

I would like to say thank to all of our teachers, and friends, as they have been the source of guidance and inspiration throughout the period of this Project.

Above all, we thank our Allah, for giving me the intellect to understand the complexity of numbers, and giving me strength to complete this research.

Shahrukh Munir

Reg. No. 297-FET/MSEE/F12

List of Acronyms

AMC	Automatic Modulation Classification
FFBPNN	Feed forward Back Propagation neural network
SVM	Support Vector Machine
CR	Cognitive Radio
SDR	Software defined Radio
PDA	Personal Digital Assistant
GA	Genetic Algorithm
PSO	Particle swarm optimization
SA	Simulated annealing
FB	Feature based
DT	Decision Theoretic
HOM	Higher order moments
HOC	Higher order Cummulants
DE	Differential Evolution

List of Tables

List of Figures

Table of Contents

This page is intentionally left blank

Chapter 1

INTRODUCTION

1.1 Introduction

In Cognitive radio network, wireless communications environment is automatically sensed and radio spectrum band is efficiently used [1]. Awareness of Wireless radio spectrum, which is the adaptable proposal for spectrum access is dependent on, is a protuberant characteristic of cognitive radio networks [2]. In a conventional communications system, the receiver works cooperatively with the transmitter. That is, the receiver has deducible information of the modulation format of the transmitted signal. For an analog communication system, the modulation format includes modulation type, nominal carrier frequency, modulation index, etc. For a digital communication structure, the modulation schemes includes modulation type, symbol constellation, alphabet size, symbol rate, nominal carrier frequency, frequency deviation (for frequency modulated signals only), and a number of other parameters. Since both the transmitter and receiver are under the control of the system designers, the conventional communication studies generally focus on making communication systems more reliable, higher power and/or bandwidth efficient, and more secure [3].

As mentioned above, one of the essential requirements for a communication system is the security. The two users in communication system don't want their communication to be known to the third user. In contrast to this, the communication management authority might wish to detect these non-licensed transmitters. The essential step of doing so is to identifying or classifying the modulation scheme of intercepted signal, which is the signature of a transmitter. Such demands also arise in many other military and noncombatant applications such as surveillance, validation of signal, verification, identification of interference, selection of proper demodulation methods in software defined radio (SDR), electronic warfare and threat analysis [4].

Formerly, the communication system relied on manual classification of modulation scheme. In manual modulation classification, system uses a set of demodulators, each designed for unique modulation type. Modulation format of received signal can be decided by listening to the demodulator outputs. Automatic modulation classification (AMC) is dominant to the manual modulation classification due to integration of the automatic modulation classifier into the received signal [5].

AMC focus on modulation scheme identification of a given communication system with a high probability success rate and in a short span of time. Modern Communication systems are transforming themselves into crafty and intelligent, new technologies are adapted in order to increase user performance by reducing transmission power, maximization bandwidth utilization, and mounting connection reliability and system security [6].

Wireless communication technologies are expanding very quickly and the number of devices such as Mobile phones, personal digital assistant (PDA) and laptops which depends on wireless technologies are also increasing. That influence of wireless technologies is widen over wireless sensor networks for security purpose, for examples automation system for home, smart grid control, medicinal wearable, embedded wireless device, and for entertaining systems. By this boom, wireless technologies has upraised a huge demand for wireless spectrum band. Cognitive Radio (CR) is demanding technology for upcoming communications network which can exploit limited available network resources efficiently without creating interference with the primary users. It is different from traditional communication system in a way that the wireless devices can adapt their operational parameters by themselves. These parameters includes its operational frequency, transmission power and modulation format [7]. AMC originates in military, it was used for military intelligence communication applications such as surveillance of band spectrum, jamming of channel, assessment of threat risk and identification of interference [8].

In most communication systems, various licensed services are specified by a reserved spectrum bandwidth. Research shows that, maximum time reserved bandwidth remains unoccupied and existing utmost resources are wasted [9]. To solve this problem CR system offers a solution by giving the preference to the primary users (licensed services) by utilizing the reserved spectrum, and authorizing the secondary users in order to utilize the available spectrum when it is unused. This method improves the spectrum utilization, and permits primary and the secondary users together to utilize the available spectrum in an efficient way. The unlicensed users also known as secondary users are considered to be capable of recognizing the spectrum state (whether the spectrum is available or not). AMC is utilized to recognize the signal type in the spectrum and hence applied in the secondary transmitter [10].

In the wireless communication system, channels present are time variable and several hindrances occur in-between transmitter and receiver. In multipath fading channel, single carrier transmission method is used which results in corruption of signal, as in transmitted signal different frequencies have different gain of channel. This problem is solved by OFDM. The spectrum is divided into small sub bands then one sub carrier is used for every sub band. So, each of these small band of frequency is transmitted over the flat fading channel and inter symbol interference effect between theses mall frequency bands is minimized [11].

Furthermore, many different levels of modulation are being used which are dependent on information of channel condition for every sub band. Such kind of method can be identified as adaptive modulation. For instance IEEE 802.11a is the standard OFDM protocol, have throughput for 64 QAM in the range of 48 Mbps. But the probability of error rises with rise of modulation level. Therefore, high levels of modulation can be utilized by sub carriers having higher SNR values, also the lower levels of modulation can be utilized by low SNR value sub carriers, which result in considerable throughput improvement of a communication system.

The adaptive modulation system receivers needs to classify the modulation form for each sub carrier so as to choose the demodulation method for each modulation type.

This is possible by using a table called Bit allocation table (BAT), this table is transmitted with each frame transmitted to the receiver in order to share the modulation type information for each sub carrier. But, this bit allocation table creates an extra overhead, mainly for large numbers of sub carriers as well as small frames of OFDM. The pretty way out for this, is to use AMC on receiver end in order to classify the modulation format for respective sub carrier, thus overall system transmission rate is increased [12].

1.2 Applications:

AMC has many applications in cognitive radio network, software defined radio network (SDR), orthogonal frequency division multiplexing, and electronic welfare systems. AMC several civilian applications. Civilian application areas like monitoring of spectrum, identification of interference, monitoring and controlling of transmission, and SDR. Moreover, AMC also has military application. AMC military application areas are electronic warfare system, analysis of threat, acquisition of target and jamming of signals. For instance, one of major application in military is, it involve in classification of different modulation scheme of enemy intercepted signal when no information available about modulation [13-15].

1.3 Motivation and Problem Statement

The CR's have become a significant area for researcher in communication systems over the past few years. AMC is a key element which increases the overall cognitive radio networks performance. In current ages, several classification methods have been established. In the previous communication system, the receiver has information of type of modulation which is being used by the sender. The key aim of this dissertation is to empower the receiver in order to identify or classify the signal modulation automatically.

There are two methods for automatic modulation classification.

1. Decision theoretic approach
2. Pattern recognition approach.

To solve AMC problem, we have used feature based PR approach.

1.4 Contribution to the thesis

The thesis contribution is outlined below:-

- Proposed a modulation classification algorithms based on continuation of Support Vector Machine (SVM) and feed forward back propagation neural network (FFBPNN) Classifier using High Order Cumulants (HOCs) as a feature set. The proposed system has the following benefits:
  - i. It provides high accuracy of classification as compared to state of the art existing techniques in literature.
  - ii. Capable of to classify different forms of modulation even in the presence of AWGN noise as well as Rayleigh fading channel and Rician fading channel.
- Feature selection subsystem is based on moments and Cumulants is integrated with the proposed SVM Classifier which result in simplified model of classifier.
- Performance of classifier is optimized using evolutionary computing technique such as Genetic algorithm (GA) and Particle swarm optimization (PSO)

1.5 Thesis Outline

This thesis is organized as follows:

Chapter 1 presents the introduction of the thesis, and motivation to the research. This chapter also includes the problem statement for our research and our contribution towards the completion of our thesis.

Chapter 2 start with literature review of previous work done in this area. This chapter describes the modulation classification and different techniques for automatic modulation classification and features extraction method. Lastly, certain optimization algorithms have been discussed such as genetic algorithm, particle swarm optimization, simulated annealing and differential evolution.

Chapter 3 leads an overview to pattern recognition systems and feature extraction method. It also presents the structure and mechanism of support vector machine and feed forward back propagation neural networks as a classifier. Simulation results are incorporated at the end of this chapter which shows the supremacy of the classifier.

Chapter 4 shows the optimization of proposed classifier using GA. The simulation results are also compared with the result of chapter 3 and it is found that with optimization classification accuracy is much improved.

Chapter 5 conclude the thesis and give some future directions.

Chapter 2

LITERATURE REVIEW

In this chapter, different AMC techniques are discussed. It also presents various algorithms used for AMC. Different types of classifier are also discuss which are used for signal classification. The Figure 2.1 displays flow chart of literature review.

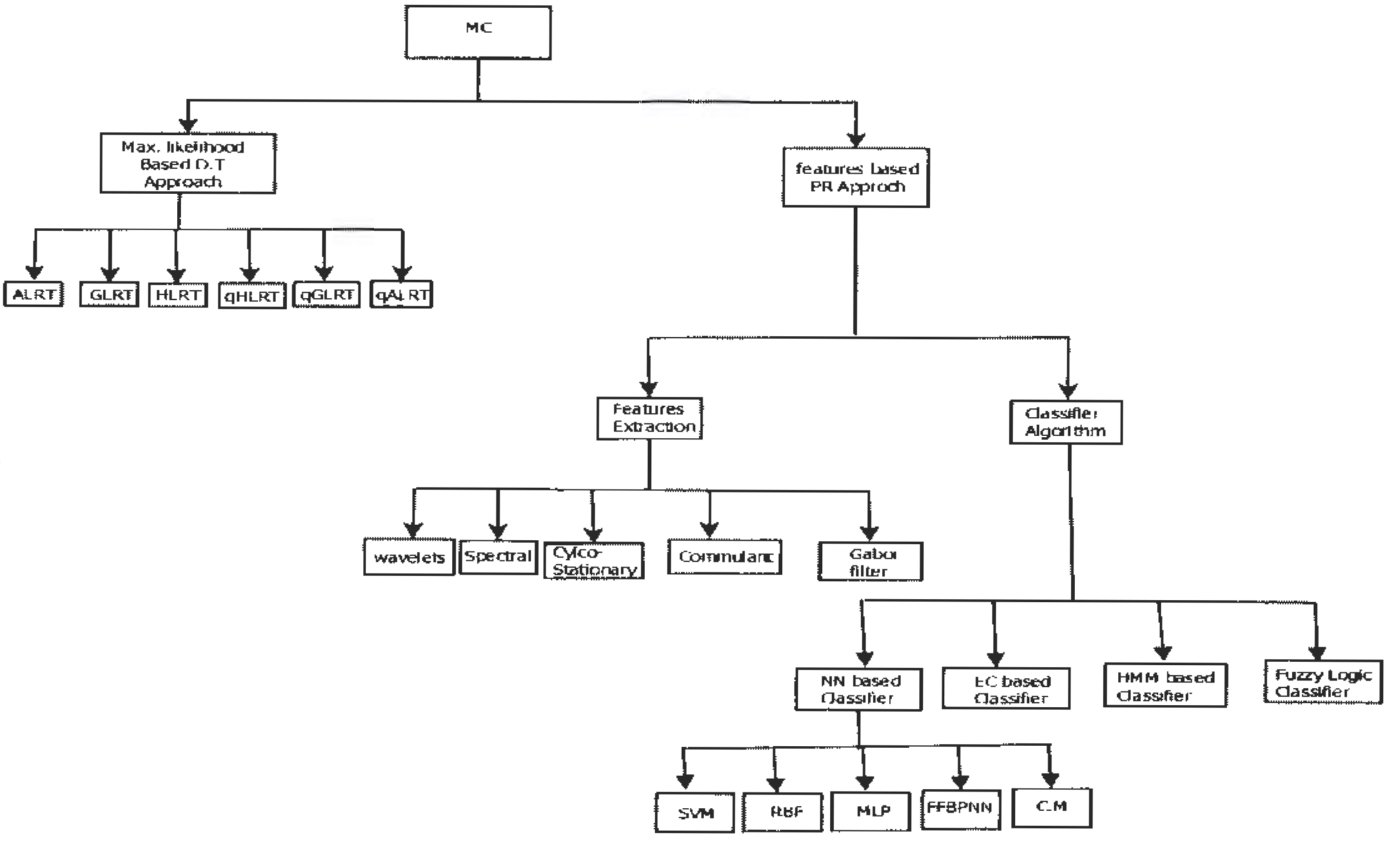

Fig. 2.1 Flow chart of Literature review

2.1 Automatic Modulation Classification (AMC)

AMC is a method of classifying different types of modulations from the received signal. AMC provides an in-between step between signal recognition and demodulation. In modern communication system AMC has various application. Due to its various application, AMC is a hot research area [16]. For a successful modulation classification different types of solutions are proposed. Normally Modulation classification is divided into two key approaches:-

- Maximum likelihood based decision theoretic approach.
- Feature-based pattern recognition approach.

2.2 Maximum Likelihood (ML) based decision theoretic approach:

ML classifiers develop the modulation classification technique by means of multiple string of hypothesis test. In this test each modulation type represent one hypothesis of respective modulation type. ML classifier improves a classification process also it provide greatest possible rate of identification. It is achieved by determining a likelihood based function (which is generally a function of symbols transmitted) and the parameter of channel [17].

The contestant having a greater likelihood is selected as the kind of modulation.

$$\hat{m} = arg\ max_{s_m}(\rho(r|s_m)) \tag{2.1}$$

Where $\hat{m}$ symbolizes the receiver picked modulation type, $\rho(r|s_m)$ is the likelihood that the signal received is r and assumed that s_m is the transmitted signal, and r_m is being modulated by using m modulation type.

ML approach can be further divided into following six categories:

a. Generalized Likelihood Ratio Tests (GLRT)
b. Average Likelihood Ratio Tests (ALRT)
c. Hybrid Likelihood Ratio Tests (HLRT)
d. Quassi Average Likelihood Ratio Tests (q-ALRT)
e. Quassi Generalized Likelihood Ratio Tests (q-GLRT)
f. Quassi Hybrid Likelihood Ratio Tests (q-HLRT)

In the ALRT method signals unknown parameters such as noise power and signal constellation are used as random variables, and the received signal probability density function

(PDF) is computed by taking average of these random variable , assuming that their distributions are known. When this assumption is perfect then best classification accuracy is achieved by using this technique. On the other hand in GLRT, the received signal PDF is calculated by using the ML estimation of unknown paramcters, in which they worked by means of deterministic variable.

Although, ALRT needs complex mathematics specifically when there is an increase in random variables occur; in contrast, GLRT is less complicated but it has a disadvantage. It gives inappropriate classification accuracy when this one is used to classify the nested signal constellation such as 16 QAM and 64 QAM. This issue is resolved by HLRT. It contains the mutual benefits of ALRT and GLRT. So, HLRT is assumed in a way that some of its signal parameters are random and their PDF is also known and some of its parameters are deterministic [18].

HLRT variant is known as quasi HLRT (QHLRT) in which the unknown signal parameters are substituted by their moment based estimates. QHLRT have less complexity and have reasonably precise feature estimator.

Although, the ML approach can provide best solution for AMC but still features based approach is preferred for AMC due to its less complexity and suitable performance.

Table 2.1 : Likelihood based classifier summary

Author(s)	Classifier(s)	Modulations	Unknown Parameter(s)	Channel
(Kim et al., 1988 and Hsue et al., 1991)	Quasi-ALRT	BPSK, QPSK	Carrier Phase θ	AWGN
(Long et al., 1994)	Quasi-ALRT	16 PSK, 16QAM, V29	Carrier Phase θ	AWGN
(Beidas et al., 1995)	ALRT Quasi-ALRT	32FSK,64PSK	Phase jitter	AWGN
(Nandi et al., 1995)	Quasi-ALRT UW	BPSK, QPSK, 8PSK, 16PSK	Carrier Phase θ & timing offset	AWGN
(Chuggi et al., 1995)	HLRT	BPSK, QPSK, OQPSK	Carrier Phase θ, signal power S &	AWGN

			PSD N	
(Sapiano et al., 1996)	ALRT UW	BPSK, QPSK. 8PSK		AWGN
(Beidas et al., 1996 and Beidas et al., 1998)	ALRT Quasi-ALRT	32FSK,64FSK	Phase jitter and timing offset	AWGN
(Sills 1999)	ALRT	BPSK, QPSK, 16 QAM, V29, 32 QAM, 64QAM	Carrier Phase θ	AWGN
(Wei et al., 2000)	ALRT	16 QAM, V29		AWGN
(Panagiaotou et al., 2000)	GLRT HLRT	16PSK, 16 QAM, V29	Carrier Phase θ	AWGN
(Ho et al., 1995)	HLRT	QPSK, BPSK	Angle of arrival	AWGN
(Hong et al., 2003)	ALRT	BPSK, QPSK	Single level	AWGN
(Fahad Hameed et al., 2009)	ALRT	BPSK,QPSK	Single level	AWGN
(Jefresonl. Xu et al., 2011)	ALRT	BPSK, QPSK,QAM,	Single level	AWGN
(Francisco C.B.F Muller et al., 2011)	ALRT	BPSK,QPSK,QAM	Single level	AWGN
(Dongwoon et al.,2012)	ALRT	QAM	Single level	AWGN
(QinghuaShi et al., 2011)	ALRT	QAM	Carrier Phase θ	AWGN
(Fanggangwang et al., 2010)	GLRT	QAM	Carrier Phase θ	AWGN
(Jefreson et al., 2011)	GLRT	BPSK,QPSK,QAM	Single level	AWGN
(Zhechan Zhu et al., 2014)	GLRT	BPSK,QPSK,QAM	Carrier Phase θ	AWGN
(M.Derankht lan et al.,2011)	HLRT	BPSK,QPSK,QAM	Single level	AWGN
(Michael et al., 2013)	HLRT	BPSK,QPSK,QAM	Single level	AWGN
(Fahed et al., 2009)	QHLRT	BPSK,QPSK	Single level	AWGN
(Ramezai et al., 2013)	QHLRT	QAM	Single level	AWGN
(Sai et al., 2016)	ALRT	PSK, 16QAM and 32QAM	Single level	AWGN

2.3 Feature Extraction based Pattern Recognition Approach

In feature based (FB) approach, modulation format identification consist of two steps. The features extraction subsystem and the pattern recognizer. The features extraction subsystem is very significant and is associated to pattern recognizer. The typical pattern recognition system can reduce the amount of input data by extracting different attribute known as features. First of all, instead of dealing with stream of signal, received signal are represented by extracted features. There are different methods for features extraction such as wavelets based features extraction, spectral features, cyclostationary features, Gabor filter features extraction method and Cummulants bases features extraction method. Then, these features are fed into the classifier in terms of making decision about modulation format. There are various algorithms available used for classification of modulation schemes such as Fuzzy logic classifier, Evolutionary computing based classifier, hidden Markova model based classifier and neural network based classifier. Feature extraction approach is very simple to implement and it can bring optimum solution as compared to the likelihood based decision theoretic approach. This method is widely used for AMC. Several types of the features are explain below:-

2.3.1 Wavelet based features

A digitally modulated signals are cyclostationary signals and they shows their transient behavior when the data symbols varies. Each modulation type have its own transients, and this difference is used to classify the modulation type. For example, in frequency shift keying (FSK), the changes occurs in the frequency, whereas in phase shift keying (PSK) changes occurs in the phase. One more example for this is that the M-ary PSK signal M possible changes occurs in phase. The wavelet transform (WT) gives a constant Q analysis which is appropriate for detection of transient and its characterization. WT is computed by fast algorithms, thus permits real time identification [19].

2.3.2 Spectral based features

Instantaneous features are fit for signal which hold secreted info in a single domain, instantaneous frequency, instantaneous amplitude or instantaneous phase. Certain instantaneous features are described below: [20].

- **Instantaneous amplitude (σ_{aa})**

The instantaneous amplitude is defined as absolute value of standard deviation of a signal, and it is given by:

$$\sigma_{aa} = \sqrt{\frac{1}{N}\left(\sum_{A_n(i)>a_t} A_{cn}^2(i)\right) - \left(\frac{1}{N}\sum_{A_n(i)>a_t} |A_{cn}(i)|\right)^2} \; b \tag{2.2}$$

Where $A_{cn}(i)$ is the normalized centered instantaneous value at time t= i/f_s (i=1, 2 ...Ns), f_s is the sample rate, Ns is the no. of samples per signal fragment, a_t is value of threshold level for $A_n(i)$ below which estimation of the instantaneous phase is very noise sensitive. $A_{cn}(i) = A_{cn}(i)$ -1 where $A_n(i)= \frac{A_n(i)}{m_a}$ and $m_a =$ (1/Ns) $\sum_{i=1}^{N_s} A(i)$. It can be used to recognize the 2ASK and 4ASK. Since for 2ASK, absolute value of its instantaneous amplitude is constant [21].

- **Instantaneous Phase (σ_{ap})**

As the standard deviation of the absolute value of nonlinear centered components of the instantaneous phase evaluated over non weak section of received signal.

$$\sigma_{ap} = \sqrt{\frac{1}{c}\left(\sum_{A_n(i)>a_t} \phi_{NL}^2(i)\right) - \left(\frac{1}{c}\sum_{A_n(i)>a_t} |\phi_{NL}(i)|\right)^2} \tag{2.3}$$

Where c is the number of samples in $\{ \phi_{NL}(i) \}$ (at instant time t= i/f_s) for which $A_n(i) > a_t$ namely non-weak points. $\phi_{NL}(i) = \phi(i) - \phi_o$, where $\phi_o = \left(\frac{1}{N_s}\right)\sum_{i=1}^{N_s} \phi(i)$ is defined and we improve it to $\phi_o = \left(\frac{1}{C_t}\right)\sum_{a_n(i)>a_t}^{N_s} \phi(i)$.

- **Instantaneous frequency (σ_{af})**

Standard deviation of the absolute value of the normalized centered instantaneous frequency over non-weak segment of the intercepted signal:

$$\sigma_{af} = \sqrt{\frac{1}{c}\left(\sum_{A_n(i)>a_t} f_N^2(i)\right) - \left(\frac{1}{c}\sum_{A_n(i)>a_t} |f_N(i)|\right)^2} \tag{2.4}$$

Where $f_N(i) = \frac{f_c(i)}{r_s}$, $f_c(i) = f(i) - m_f$, $m_f = (1/N) \sum_{i=1}^{N} f(i)$. Where r_s is the symbol rate of digital sequence, C is the number of samples in { $f_N(i)$] (at instant time t= i/f_s)

For which $A_n(i) > a_t$ namely non weak points. This feature could differentiate between the modulation schemes without frequency information and the FSK modulation schemes and also between FSK2 and FSK4.

- **Power spectral density of the normalized-centered instantaneous amplitude γ_{max}**

γ_{max} is the maximum value of the power spectral density of the normalized-centered instantaneous amplitude of the intercepted signal segment, and is defined by

$$\gamma_{max} = \max|\mathrm{FFT}(A_{cn}(i))|^2 / N_s \tag{2.5}$$

This feature can express the character of signal's envelope and was added to differentiate between the modulation schemes that carry amplitude modulation and those that do not. For example, γ_{max} has a higher value for QAM 8 than for ASK4 because the former has amplitude levels 1 and 3 whereas the latter has amplitude levels 1 and 1/3. For frequency modulated signal, like FSK, there is no amplitude modulated information, so this parameter is very small [22].

- **Standard deviation of normalized instantaneous frequency σ_{fn}**

Standard deviation of the direct value of the normalized centered instantaneous frequency, evaluated over the over non weak segment of the intercepted signal

$$\sigma_{fn} = \sqrt{\frac{1}{c}\left(\sum_{A_n(i)>a_t} f_N^2(i)\right) - \left(\frac{1}{c}\sum_{A_n(i)>a_t} f_N(i)\right)^2} \tag{2.6}$$

σ_{fn} is used to discriminate between FSK2 and FSK4 signals.

2.3.3 Cyclostationary features

Cyclostationary is a pretty good feature for classification of modulation schemes. For example estimation of cyclostationary features can be magnificently done without demanding several preprocessing jobs that rely on unavailable channel information. Previous studies proposed a FB classifier that follows the decision tree algorithm and used cyclostationarity for testing, HO cyclostationarity at each node. The classifier controls BPSK, MPSK, AM, and M-QAM. The phase, frequency robustness, timing offset estimation errors are the highlights for this cyclostationary feature based algorithm. The firstly implemented technique get benefit of the fact that due to underlying periodicities, several time signals waveforms are exhibited as cyclostationary instead of stationary. Both means and autocorrelations are periodic for such kind of processes. Spectral correlation function (SCF) is obtained by taking Fourier transform (FT) of cyclic autocorrelation. Sometimes SCF is also known as spectral correlation density (SCD). The highest values of normalized SCF beyond whole frequencies cycles provide the cyclic frequency domain profile (CDP). Different types of modulation have their own unique pattern of CDP which helps in differentiating between different modulations format [23].

2.3.4 High Order Statistics

The moments of signals are obtained by taking the expectation of power signal, determined by moment order. The first order moment is obtained by taking the mean whereas the second moment order generally point to the signal power. Consider y_n, is the complex value stationary random process, and the p is the moment order, and is defined as:

$$M_{pq} = E[\, y^{p-q} (y^{*q})] \tag{2.7}$$

y^* is the complex conjugate of y, and q is power of the conjugate signal y.

Another way for calculation of signal statistics is higher order Cummulants.

To find the Cummulants of signal, firstly define the characteristics function of random variable x;

The first characteristics function of variable x is defined as: [24]:

$$\emptyset(s) = \int_{-\infty}^{\infty} f(x) e^{sx} d(x) \tag{2.8}$$

while 2nd characteristic function is given by:

$$\varphi(s) = \ln(\emptyset(s))$$

(2.9)

The Cummulants of signals are obtained by taking the derivatives of its second characteristic function about the origin.

$$\lambda_{p} = \frac{d^{p}\varphi(0)}{ds^{p}} \tag{2.10}$$

Where, p shows the Cumulants order. Mostly, HOCs are expressed in terms of signal function of HOMs. Usually, HOMs are calculated for signals with zero mean. So, the mean is subtracted from the received data symbols:

$$\widetilde{y_n} = y_n - \frac{1}{N}\sum_{n=1}^{N} y_n \tag{2.11}$$

Where N is the total number of symbols [25].

2.3.5 Gabor filter based features

Gabor atom is efficient mean for features extraction. Favored introduction and favored spatial recurrence best describe the Gabor filter. Normally, a 2-D Gabor channels goes about as a nearby band-pass channel with certain ideal joint confinement properties in the spatial area and in the spatial recurrence space. Typically, a picture is sifted with an arrangement of Gabor channels of various favored introductions and spatial frequencies that spread properly the spatial recurrence space, and the elements got structure an element vector field that is further utilized for examination, characterization, or division [26].

2.4 Classifier Algorithm

Classification is the process in which different modulation classes are identified and differentiated, on the basis of a training data. There are different classifier algorithms used for classification. Some of them are discussed below:-

2.4.1 Fuzzy logic based classifier

Fuzzy Logic can be conceptualized as a speculation of established rationale. Current Fuzzy Logic was created to demonstrate those issues in which uncertain information must be utilized or as a part of which the guidelines of derivation are defined in an exceptionally broad manner making utilization of diffuse classes. In Fuzzy Logic, which is additionally in some cases called diffuse rationale, there are two options as well in general continuum of truth qualities for coherent suggestions [27].

2.4.2 HMM Based classifier

In literature, HMM is used for classification of different modulations schemes. For detection of signal detection we assent that estimation of bandwidth is known. For signal recognition and extraction from the cyclic frequency domain profile (CDP), crest factor (CF) is used. CF is a dimensionless quantity. The CF is obtained by dividing waveform peak amplitude value to its root mean square value. When peak values are known, then this becomes a simple single cycle detector. For signal recognition, firstly its threshold values are calculated when there is no signal present, i.e. when AWGN is present then we have

$$C_{TH} = \frac{\max\left(I(\alpha)\right)}{\sqrt{\left(\sum_{N}^{a=0} \rho(\alpha)\right)/N}} \tag{2.12}$$

When the CF value is higher than CTH value then we announce the presence of signal. For the extraction in future all CDP peaks values are higber than the CTH and are encoded as 1 and all others are encoded as 0. This generate the binary feature vector and then it is fed into the HMM based classifier [28].

Discrete sequence or process is a Markov process. Where the future process gives the present and it is independent of past. Markov model is known as stochastic model of a system. It is being capable of finite states from 1, 2… S. from Markov property, we can determine the probability of arriving at the next stage by adding all the probabilities of arriving at that state [29].

2.4.3 Neural Network Based classifier

The modulation classifier which is based on the ANN approach is shown in fig.2.1. This classifier is considered to differentiate between different modulations format and it consists of the three main blocks. The preprocessing block wherein the key input characteristics are extracted from each and every part of signal frame. The training and learning blocking order to make way for classifier structure. The testing block in order to take decision about the signal modulation type.

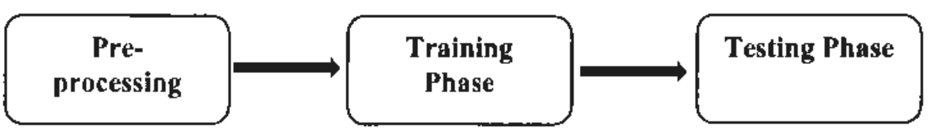

Fig. 2.2 ANN Classifier functional blocks

There are following classifier types which based on ANN classification algorithms.

- Support Vector Machine (SVM)
- Radial Basis function (RBF)
- Multilayer perceptron (MLP)
- Feed forward back propagation networks (FFBPN)

2.4.4 Evolutionary computing based classifier

In AMC, in order to select best features from a group of features the optimization methods is used sometimes it is also used to generate and test new features. Optimization techniques put great influence on classification accuracy. By using this technique the three is rise in percentage of correct classification. Different features are selected from the received signal for each type of modulation scheme. These features are used to determine its modulation kind.

For every classification problem, different types of classification features are selected from received signal in order to determine its type of modulation. By using optimization techniques the best features are selected which affect the percentage accuracy of classification and eliminate redundant features. The final outcome of optimization is to maintain a feature vector that has a less number of features as compared to original feature vector, as a result same accuracy or most of time better classification accuracy is achieved by using the original set of features[30].

2.5 Heuristics Techniques

The heuristic technique, sometimes simply called a heuristic, is a way to solve the problem by learning and discovery that apply the practical method which not guaranteed to become optimal or perfect, but it is enough for our immediate goals. Where outcome for an ideal solution is impossible or unpractical, in order to find the suitable solution a heuristic technique can be used. Heuristics provide the psychological shortcuts that comfort the mental load of making a decision. The examples of this technique are using thumb rule, an educated guess, a spontaneous judgment, outlining, stereotyping or common sense [31].

Overall, heuristic technique has many advantages:

- Insensitivity to the scope of problem.
- Provide global solution.
- Provide number of optimal solutions.

However, it has some disadvantages:

- It is more Complex
- Time consuming due to its iterative approach.

2.5.1 Genetic Algorithm (GA)

A genetic algorithm (GA) is also an optimization technique and it was invented in 1970 by John Holland. To solve optimization problem, GA optimization technique is used because it implement the idea of natural selection and genetics. GA is preferred in those problems where there are enormous number of solutions and the search space contains number of mountains and valleys. Using the previous possible solutions history and conducting a survey in parallel, GA is found to be adoptable for finding the best solution and it supervise the native best solution that cannot be simply ignored by the most of other optimization techniques. GA is most commonly apply on diverse optimization problems, while the goal is to select the best candidate from the population which can take to the best solution of problem. A criteria for selection is used by using different operators like recombination and mutation [32].

In AMC, many different types of features are proposed in order to differentiate between modulation schemes. Such features vary in power of classification. GA approach is used in order to pick the optimal features also to reject the identical features for avoiding and using a multiple redundant features which may effect in classification accuracy reduction [33].

The GA employment involves following points:

I. **Fitness function**:

It is an Objective function, which could be maximized or minimized by GA. For AMC, it is selected for the classification accuracy.

II. **Initial population:**

GA needs some starting values which are to be used for problem resolution (for AMC case, features classification). First of all, in order to cover all possible solutions, select an initial population randomly out of possible solutions set. To define an optimal solution among all possible solutions, fitness function evaluate all possible solution. Then this optimal solution is used in the next iteration and rest of all other solutions are washed out.

III. **Crossover and mutations:**

A living population, generally denoted as parent, and are utilized to produce a fresh set of probable solutions also called children. The Children are produced by their real parents with a process known as crossover. A crossover phase has, different types of functions like plus, minus, times, reciprocal, negator, abs, sqrt, sin, cos, tan, are to be used to in order to recombine different parents which produces a new child having common features as of their original parents. In addition, there is one more process known as mutation. It is used to bring a slight change in the children produced to retain their randomness behavior in the population of new generation.

IV. Stopping criteria:

As the new population is generated, the children solutions are examined with the fitness function (the fitness function is predefined), from among the new solutions the best children are retained for next iteration which symbolize the new parents. Generally, the fitness of solutions in second iteration is higher than the fitness of solution in the first iteration. This procedure for combing the features and solutions evaluation carry on for number of iterations till a particular condition is met. Either iteration touches to its extreme number or there is no significant addition in new generation fitness as compared to its former solutions [34].

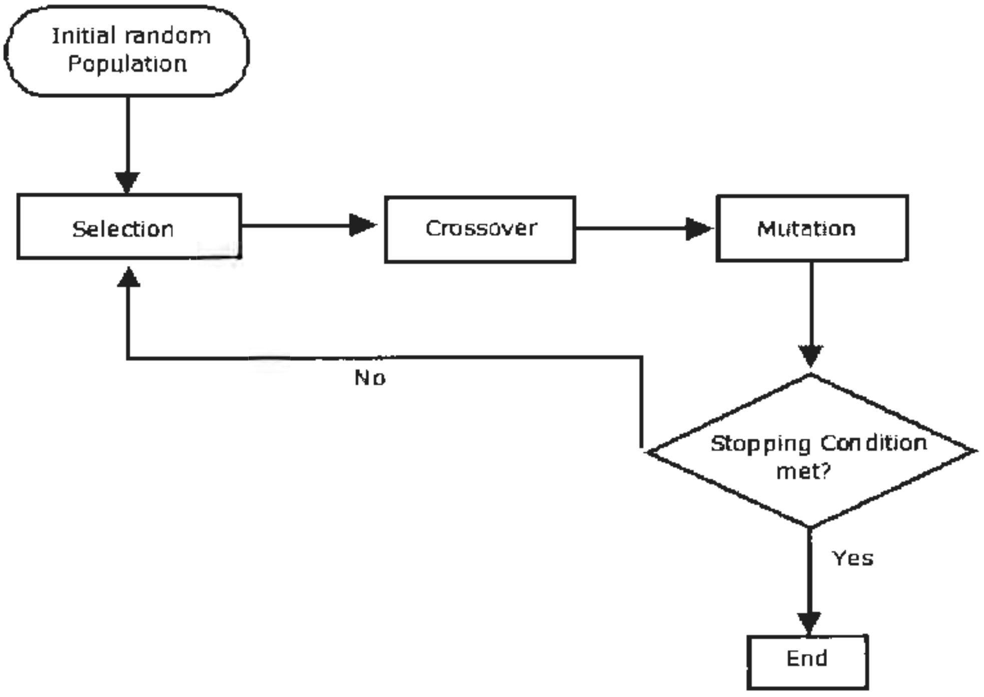

Fig. 2.3 GA Flow Chart

2.5.2 Particle Swarm Optimization

The Particle Swarm Optimization (PSO) was developed in 1995 and this optimization technique is a robust algorithm [35] . PSO copies the herd of bird's behavior, insect's swarms, or school of fish, when moving together in the form of group in order to catch their last destination. In PSO, the particles fly's over the multidimensional space in which all particle apt their position in accordance what they experience alongside the feedback from their neighboring particles. In PSO, the positions of each particle is originated randomly and assessed to compute the predefined fitness function [36].

For AMC, classifying features indicate the particles in the swarm, while using these features the classification accuracy achieved is known as fitness function. In first iteration, set of solution is randomly initiated, where each particle is a features vector. The particle having highest classification accuracy are stated as , local best, while those particles which takes to an optimal classification accuracy rate among rest iterations is known as global best [37].

Features of all particles are improved according to their earlier experience and also from the experience of total swarm in order to give rise in their fitness function in succeeding iteration. Near the completion of optimization, its final outcome is the set of particles, which can take to highest classification accuracy. Then such kind of features are being utilized in AMC problem rather than the original features set [38].

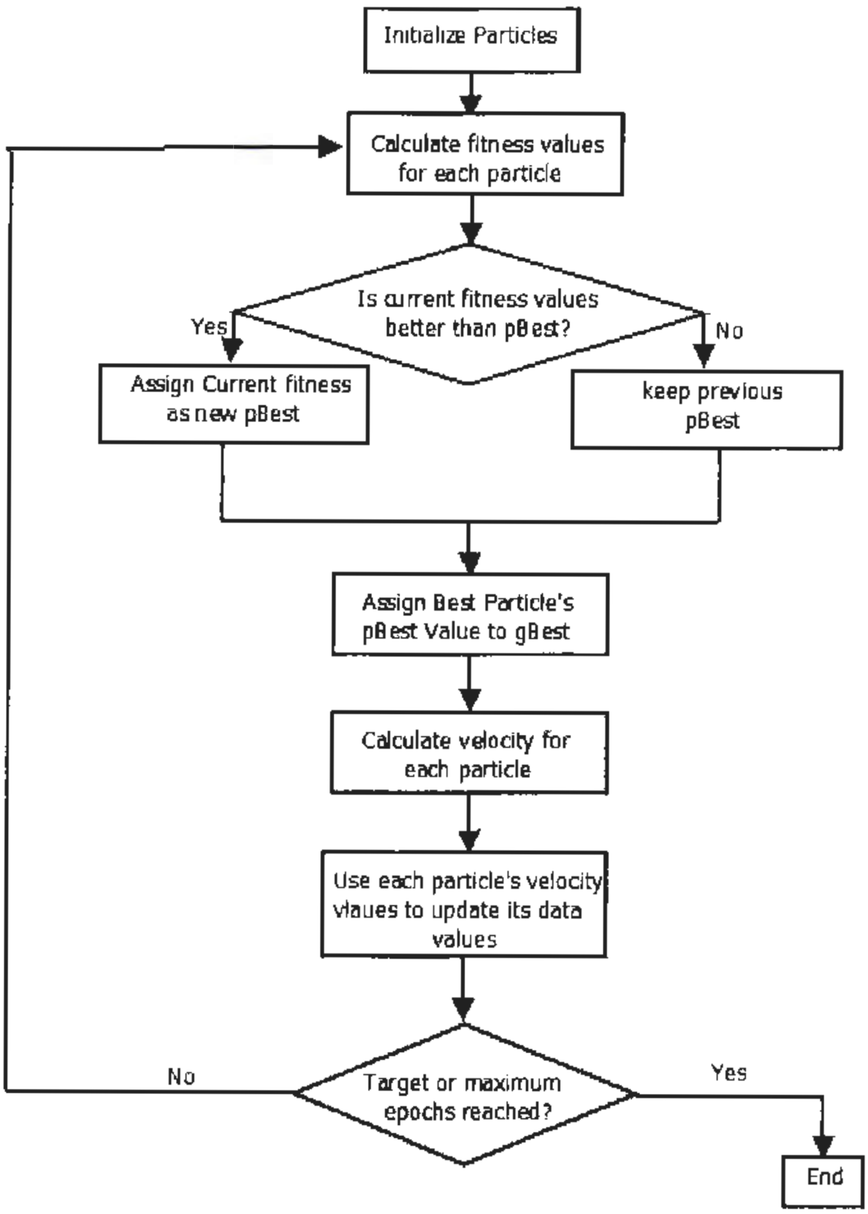

Fig. 2.4 PSO Flow Chart

2.5.3 Simulated annealing (SA)

Simulated annealing (SA) is a process for resolving unconstrained and constrained bounded optimization issues. The technique illustrate the physical heating procedure of a material and then gently dropping the temperature to reduce the flaws, thus reducing the system energy.

For each SA iteration, a new random point is initiated randomly. The distance between current point and the new point depend on the probability distribution with a scale related to temperature. The SA algorithm takes each new generated point that lessen the objective, besides, with a particular probability, the points which raise the objective. To accept those points which raise the objective, the algorithm avoids for being trapped in local minima, and for more possible solutions it is capable to explore globally. To lower its temperature systematically an annealing schedule is selected which decreases the temperature as the algorithm continues. With the decrease in temperature, the algorithm also decreases the magnitude of its search in order to come together to a minimum [39].

The SA algorithm inspiration was at first comes from annealing process in metal work. Annealing process consist of heating and chilling a material to modify its physical states due to the variation occurs in its internal conformation. As the metal absolutely cools then its new structure turn into fixed, as a result instigating the metal to hold its newly attained properties. In SA we have a temperature variahle which stimulate the heating process. At initial level we set it to high and then let it to become cool slowly as the algorithm procced. Although this temperature variable is very high and the algorithm will be acceptable with extra frequency, to agree on solutions which are worse than the current solution. This allows algorithm to jump out of any local bests to takes itself in early execution. Acceptance of worse solution increases as the temperature decrease, thus permitting the algorithm to slowly emphasis in on the search space area through which hopefully, a nearby optimal solution could be found. This process of cooling makes the SA algorithm extraordinarily effective in determining a close to optimal solution when dealing with large problems which have several local bests. The traveling salesman nature prohlem shows its perfect illustration [40].

Algorithm

Algorithm consist of following steps.

1. Random solution generation
2. Calculation of its cost by using cost function
3. Generation of neighboring random solution
4. Calculate the cost of new solution
5. Make comparison :
   - When cost $_{new}$ < cost $_{old}$: switch to the new solution
   - When cost $_{new}$ > cost $_{old}$: possibly switch to the new solution
6. Repeat steps from 3 to 5 until a suitable solution is found or reach its maximum level of iterations.

All these steps are explained in detailed as follow:

1. Random solution generation

We can generate random solution by our own choice. However, the main point is that it is random. It does not need to be our best guess at optimal solution.

2. Calculation of its cost using cost function

It also depends on problem, it can be as simple as counting the total number of miles traveled or it can be as complex as combining multiple factors. Cost Calculation for each solution is the most expensive part of algorithm, so it pays to keep it simple.

3. Random generation of neighboring solution

There is only one thing that differ from old solution and new solution. For new solution the two elements are switched and cost is recalculated. The main requirement for it is that it will be done randomly.

4. Calculation of cost of new solution

To calculate the new solution cost, the cost function is used. We can understand that is why it needs to execute well. It gets called with each iteration of algorithm.

5. **Comparison**

(i) If new solution cost < old solution cost

If the new solution cost is less than the previous cost of solution, then the new one is better. It makes the algorithm blissful and it is getting closer to an optimum solution. Then it will switch to new solution.

(ii) **If Cost of new solution is greater than old solution**

This is most interesting point of algorithm. Most of the time, the algorithm will avoid moving to a poorer solution. When it did this all of the time, then it get caught at local maxima. To avoid this problem, it sometimes vote for to keep the worse solution. To resolve this, algorithm calculates probability acceptance and then compares it with a random number [41].

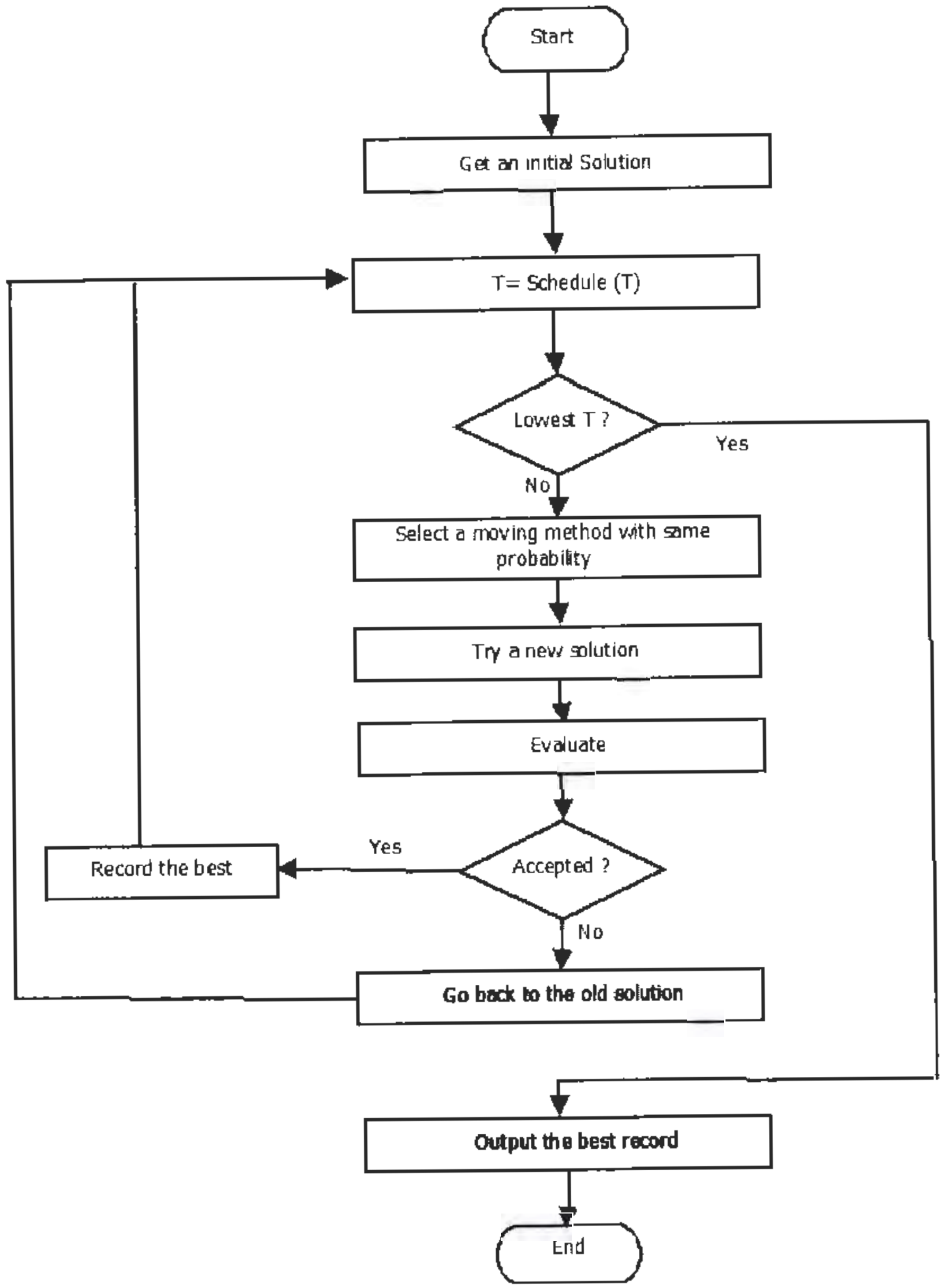

Fig. 2.5. Flow chart simulated annealing

2.5.4 Differential Evolution

Differential evolution was incepted by stone and price in 1996 by stone and price [42] and it has been showed to be a commanding and inspiring global optimization techniques as compared to the other evolutionary computing techniques. It has attracted the researchers due to its ease in implementation, fewer parameters are required to be tuned and they are highly random in nature. It can handle easily and efficiently, nonlinear, multimodal and non-differentiable cost function. It has also consistent and excellent convergence towards global minimum in successive independent runs [43]. Moreover, it has also been successfully applied to the solution of discrete, as well as, constrained problem and thus it has direct application in every field of science and engineering. Differential evolution is a kind of scheme which iteratively searches large spaces of candidate solution and tries for the improvement of candidate solution with respect to specified measurement of quality. In simple words, DE optimizes a problem in such a way that it maintains the candidate solution and then by using its simple formulae it creates a new solution by combining the existing ones. Now, it will keep only those candidate solution which have best score or fitness of the under consideration optimization problem. DE is basically have the combination of GA, Genetic programming and evolutionary programming. Like GA DE is also mainly based on the three operator's i.e. mutation, crossover and selection but its way of incorporating these operators is different from that of GA. From these three operators, role of mutation is significant in the DE algorithm performance and the DE strategies constituted are based on the mutation variants. Unlike GA, the average fitness function of DE monotonically decreases or increases without the requirement of elitism as the struggle between parents and children started after the cross over [44].The generic flow diagram of DE is show in the fig 2.5.

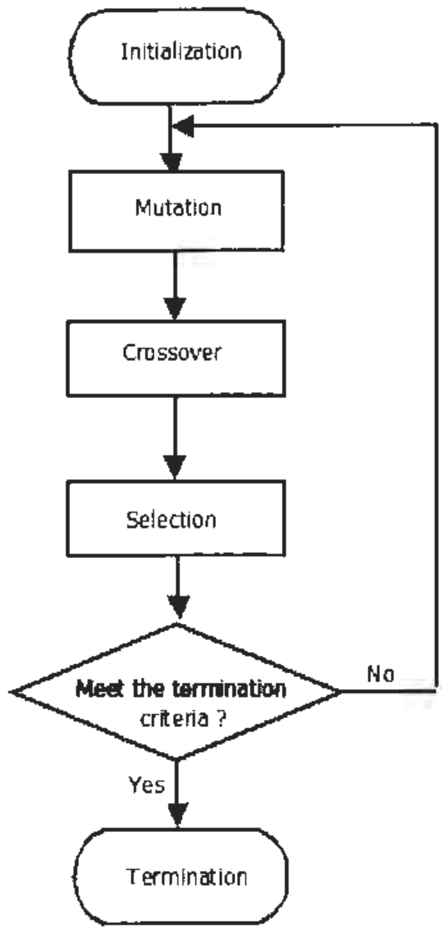

Fig. 2.6. Flow chart Differential Equation.

i. **Initialization:**

This step is exactly the same as PSO. Create random population of n chromosomes. Let, the entire population represented by matrix "C"

ii. **Updating**

In this step, update all the chromosomes of current "g_e". Now, select any chromosomes from randomly generated population e.g. choose $C_k^{i,ge}$ where " i" (i= 1,2 ……., N) represent the position of that particle chromosome in the population while " k" (k= any real number) is its respective length . The goal is to find the chromosome of the next generation i.e. $C^{i,ge+1}$ by using the following steps,

iii. Mutation

Due to this step, we have the name DE as it works on the difference of the vectors. Pick up any three different numbers (chromosomes) from 1 to N i.e. (n_1, n_2, n_3) under the following conditions ,

$$1 \leq n_1, n_2, n_3 \leq N$$

Where

$$n_i \neq n_k \ \forall i = 1,2,3$$

Now,

$$d^{i,ge} = C^{n_1,ge} + F\,(C^{n_2,ge} - C^{n_3,ge}) \tag{2.13}$$

Where 'F' is a constant whose values usually lie in the range 0.5 to 1.

iv. Crossover:

The crossover can be performed as,

$$o_k^{i,ge} = \begin{cases} d_k^{i,ge} & \text{If } rand() \leq CR \text{ or } k = k_{rand} \\ c_k^{i,ge} & 0/w \end{cases}$$

(2.14)

Where CR is cross over rate which is $0.5 \leq CR \leq 1$

v. Selection

The selection operation for the chromosome of next generation is performed as,

$$c^{i,ge+1} = \begin{cases} o^{i,ge} & \text{if } err(o^{i,ge}) \leq err(c^{i,ge}) \\ c^{i,ge} & 0/w \end{cases} \tag{2.15}$$

Repeat this for all chromosomes.

vi. Termination

The termination criterion of DE is based on the following results achieved,

i. If err $(c^{i,ge+1}) < \epsilon$, where ϵ is a very small positive number

ii. Total number of generation has reached,

Else go back to step 2.

vii. Storage

Store the entire result for later discussion and statistical analysis. [45]

Table 2.2: Features based classifier summary

Author(s)	Features	Modulations	Channel
(Azzouz et al., 1996)	Maximum power spectral density of normalized centered amplitude, standard deviations of normalized centered amplitude, phase and frequency	2ASK, 4ASK, BPSK, QPSK, 2FSK, 4FSK	AWGN
(Hsue et al., 1991 and Yang et al., 1991)	Variance of the zero- crossing interval sequence phase difference and zero-crossing interval histograms	UW, BPSK, QPSK, 8PSK, BPSK, 4FSK, 8FSK	AWGN
(Yang et al., 1997 and Yang et al., 1998)	PDF of phase	UW, BPSK, QPSK, 8PSK	AWGN
(Sapiano et al., 1995)	DFT of phase PDF	UW, BPSK, QPSK, 8PSK	AWGN
(Hong et al., 2002)	Variance of HWT magnitude, HWT magnitude and peak magnitude histograms	BPSK, QPSK,8PSK, 2FSK, 4FSK, BFSK,CP2FSK, CP4FSK, CP8FSK, MSK	AWGN
(Hong, 1999)	Variance of HWT magnitude and normalized HWT magnitude	QPSK. 4FSK, 16QAM	AWGN
(Wei et al., 2000)	Normalized fourth-order Cummulants of the received signal	BPSK,4ASK, 16QAM, BPSK, V29,V32, V29c	AWGN , impulsive noise, Co-channel interference
(Swami et al., 2000)	Normalized fourth-order Cumnulants of the received signal and AMA cost function	BPSK, 4ASK, QPSK, 16 QAM, V29, V32 64 QAM	Frequency Selective Channel
(Martr et al., 1997)	Fourth and second order moments of the received signal	QPSK, 16 QAM	AWGN
(Marchand et al., 1997 and Marchand et al. 1998)	Fourth and second order cyclic Cummulants of the received signal	QPSK, 16 QAM	AWGN
(Spooner 1995 and Spooner et al., 2000)	Sixth, fourth and second order Cyclic Cummulants of the received signal	MSK, QPSK, BPSK, 8QAM,OPSK, 16 QAM, 64QAM, V29	AWGN, Co-channel interference

(Dobre et al., 2003)	Eighth order cyclic Cummulants of the received signal	BPSK, QPSK, BPSK, ASK, 8ASK, 16QAM, 64QAM,256QAM	AWGN
(Dobre et al., 2004)	Eighth, sixth and fourth order cyclic Cummulants of the received signal	4QAM, 16 QAM	AWGN Impulsive noise
(Dobre et al., 2005)	Eighth order cyclic Cummulants of the signal at the output of a selection combiner	4ASK,8ASK,BPSK, OPSK, 16QAM, 32QAM, 64QAM	Rayleigh & Rician Fading Channels
(Yu et al., 2003)	DFT of the received signal	2FSK, 4FSK,8FSK,16FSK,32FSK	AWGN
(Fucai at el., 2007)	Wavelets based features	QPSK, 4FSK, 16QAM	AWGN
(Fucai et al., 2008)	Wavelets based features	QPSK, 16QAM	
(Me;gani et al., 2008)	Wavelets features	QPSK, 16QAM, 64QAM	AWGN
(Kamunlo et al., 2009)	Wavelets features	MSK, QPSK, BPSK, 8PSK,8QAM, QPSK,16QAM,64QAM, V29	
(Vladimire et al., 2009)	Sixth, fourth and second order cyclic Cummulants of the received signal	2FSK, 4FSK,8FSK,16FSK,32FSK	AWGN
(Vladimire et al., 2010)	Higher order Cummulants	QPSK, 4FSK, 16QAM	AWGN
(Fanggang et al., 2012)	Higher order Cummulants	QPSK, 4FSK, 16QAM	AWGN
(Michael et al., 2012)	Higher order Cummulants	QPSK, 16QAM	AWGN
(Sai et al., 2012)	Higher order Cummulants	QPSK, 16QAM, 64QAM	AWGN
(Zaerin et al., 2012)	Higher order Cummulants	QPSK, 4FSK, 16QAM	AWGN
(Jain et al., 2012)	Wavelets based features	QPSK, 16QAM	AWGN
(Subasi et al., 2013)	Wavelets based features	QPSK, 16QAM, 64QAM	AWGN
(Shadmand et al., 2013)	Wavelets based features	AWGN	AWGN
(Chang et al.,	Higher order moments	BASK,BPSK,QPSK, 16QAM,	AWGN

2014)		32QAM	
(Ghauri et al., 2014)	Eight , sixth , fourth and second order moments of received signal	2PSK to 64PSK, 2FSK to 64FSK, 2QAM t0 64QAM	AWGN
(Kawamoto et al., 2016)	Estimating the probability density of the input I/Q data using its input moments (GCA)	ASK, PSK, and QAM	AWGN
(Abdelmutalab et al., 2016)	Higher order Cummulants	BPSK , QPSK, 8PSK, 64QAM and 256QAM	AWGN

Chapter 3
AMC using SVM & FFBPNN

3.1 Introduction

In this chapter the system model for automatic model classification is discussed. The features which are extracted from the received signal are higher order statistical features which include moments and Cummulants. As Cumulants are made of moments, so moment are also used as feature. Then these features are fed into the classifier structure, which is based upon the combination of SVM and FFBPNN classifier. The simulation results are presented at three different SNRs values with different number of samples. The confusion matrix shows the classification accuracy increases as the number of samples of inputs signal are increased from 512 to 4096. The comparison of classification accuracy is also presented on different channels model such as with AWGN noise, Rician fading channel and Rayleigh fading channel.

3.2 System Model:

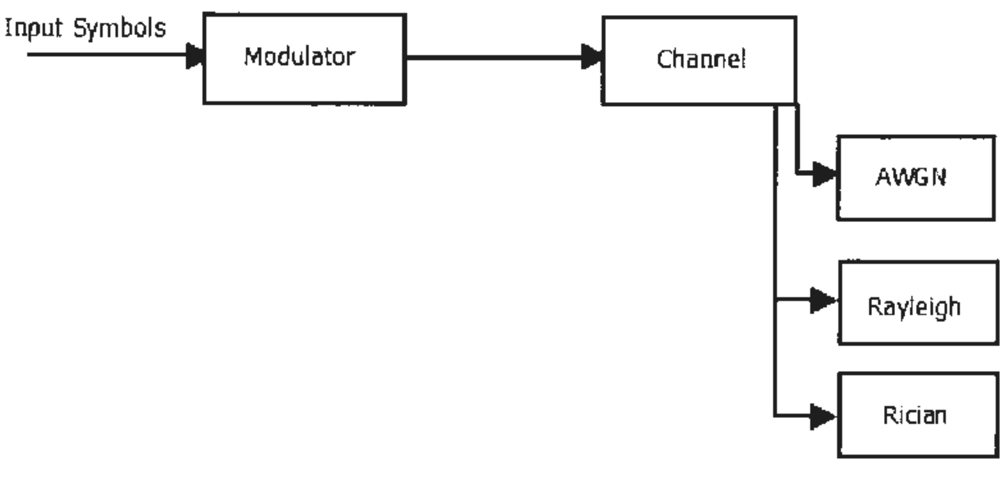

Fig.3.1 (a). Transmitter Side of Proposed System Model

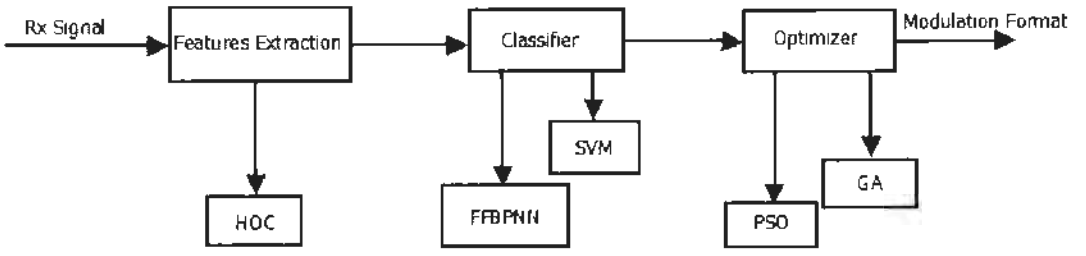

Fig. 3.1(b) Receiver Side for Proposed System Model

Figure 3.1(a) and 3.2 (b) depicts the generalized system model for an Automatic Modulation Classification (AMC). The signal is injected into the modulator for modulation subsequently, signal is transmitted over the channel. The noise is also added into the signal when passing through the communication channel. The noise considered in this research is AWGN, Rayleigh and flat fading. At the receiver side, the modulated signal is received, first of all preprocessing is executed in order to extract the features. The features taken, are higher order Cummulants. Once features extracted, then these features are fed into the classifier. The classifier used for determining AMC are SVM, FFBPNN, optimized SVM and FFBPNN. After the classification of different modulation format, the performance optimization is executed by using GA and PSO. The generalized expression for received signal is given as below:

$$r_n = s_n + g_n \tag{3.1}$$

where

r_n = received baseband signal, g_n = the additive white Gaussian Noise,

s_n = transmitted signal and is defined

$$s_n = K\, e^{-i(2\pi f_0\, n\, T + \theta_0)} \sum_{j=-\infty}^{j=\infty} S(l) \tag{3.2}$$

$$h(nT - jT + \epsilon_T T)$$

Where $S(l)$ is sequence of symbols at the input that is taken out from the set of M constellations of known symbols and the condition for symbols to be equiprobable is not necessary, K is the signal amplitude, f_0 is offset constant of frequency, T is the spacing of symbols, θ_n is phase jitter which differs from symbol to symbol, h is channel affects and ϵ_T is jitter timing.

3.3 Features Extraction

Usually, in pattern recognition approach, raw datasets are made from the received signals and this raw data is often compact into dimension prior it is being fed in the classifier. Such dimensionally decreases the data, also known as key features. These features contains the unique information of raw data. The benefit of feature extraction is, the work is to be done with fewer datasets. Application of pattern recognition approach most of the time face the dimensionality curse issue. The simplest way is to decrease the input dimension, by selecting some inputs and dispose of the remaining's. This process is known as feature selection therefore a smaller amount of resources are needed.

Such features are used to define the PDF shape of a signal. The higher order moments and Cumulants behavior to several transformations is a key factor in determining how these quantities are useful in characterizing the signals in the systems. The one simple effect of translation is seen on the received signal is that the changes occur only in the mean. All the HOM's, HOC's and the variance are not affected. The rotation in the constellation of the received signals take place as a result of multipath or due to different distortions. Which affects the variances and HOM's or HOC's relatively. However, there are particular parameters which are invariant to rotation for instant eigenvalues and covariance matrix.

HOC's are the mathematical tool, which elaborates the characteristic of HO statistics of a random process. It not only eliminate the impact of Gauss noise, but it is also rotation's robust and it offers excursion to the constellation diagram [46].

3.3.1 Higher order statistics

The moment's Probability distribution is a generalized concept of expected value. When we recall the *ith* moment general expression of random variable, which is expressed in following equation

$$\mu_i \;=\; \int_{-\infty}^{\infty} (s-m)^i \; f(s)ds \tag{3.3}$$

Whereas m is random variable mean, and f (s) is the PDF of s random variable. The *ith* moment definition for discrete signal with finite levels is specified by following equation

$$\mu_i = \sum_{k=1}^{N} (s_k - m)^i \; f(s_k) \tag{3.4}$$

Whereas N represent the length of data. We have chosen zero mean signal for this research work. So

$$\mu_i = \sum_{k=1}^{N} (s_k)^i \; f(s_k) \tag{3.5}$$

Next, random variable auto moments can be defined as below

$$M_{pq}=\mathrm{E}\,[s^{p-q}(s^*)^q] \tag{3.6}$$

whereas p represent the order of moment and s^* is the complex conjugate of a signal s. suppose a zero mean discrete signal with the sequence of form as s= a+jb, by using the auto moments definition , the different orders expressions can easily be determined.

Suppose a scalar random variable "s" with characteristic function and with zero mean

$$\hat{f}\,(t) = \mathrm{E}\left\{e^{its}\right\} \tag{3.7}$$

Expand the logarithm of the characteristic function as a Taylor series, one can finds

$$\log \hat{f}\;(t) = \; k_1(jt) + \ldots\ldots\ldots + \frac{k_r\,(jt)^r}{r!} \tag{3.8}$$

$\dot{\kappa}_r$ is constants in equation (3.8) and are known as the Cumulants of s. The representation of p^{th} order of Cumulants is same as p^{th} order of moment.

$$C_{pq} = cum\left[\underbrace{s,\ldots\ldots,s}_{(p-q)terms}\quad,\underbrace{s^*,\ldots.s^*}_{(q)terms}\right] \tag{3.9}$$

The nth order Cumulants is the function of the moments order up to n

$$\text{cum}[s_1,\ldots\ldots,s_n] = \sum_{\forall v}(-1)^{q-1}\ (q-1)!\,E\left[\prod_{j\in v1} s_j\right]\ldots.E\left[\prod_{j\in v_q} s_j\right] \tag{3.10}$$

The summation index is over all partition = (1,……..,q) for indices set (1,2 …...,n) and q is total elements for the specified partition. Suppose n=3 . In this case the indices set available are (1,2 3) and four different kinds of partitions can be attained for these set : {(1,2,3)} leading to q=1 , {(1),(2,3)} leading q=2 , {2,(1,3)}leading q=2 , {3,(1,2)} , leading q=2, { (1),(2),(3)} leading q=3 . So

Cum$[s_1,\ s_2,\ s_3]$= $(-1)^{1-1}$ (1-1)! E $[s_1,\ s_2,\ s_3]$+ $(-1)^{2-1}$ (2-1)! E $[s_1]$,$E[\ s_2,\ s_3]$+ $(-1)^{2-1}$ (2-1)! E $[s_2]$, $E[s_1,\ s_3]$+ $(-1)^{2-1}$ (2-1)! E $[s_3]$,$E[\ s_1,\ s_2]$+ $(-1)^{3-1}$ (3-1)! E $[s_1]$.$E[\ s_2]$, $E[s_3]$= E $[s_1,\ s_2,\ s_3]$ – E $[s_1]$, $E[\ s_2,\ s_3]$- E $[s_2]$, $E\ [s_1,\ s_3]$- E $[s_3]$,$E[\ s_1,\ s_2]$-
2 E $[s_1]$, $E[s_2]$, $E[s_3]$ (3.11)

In the similar way, Cumulants expression up to eighth order can be calculated.

$$:C_{20} = \text{E}\ [y^2\ (n)] = \text{cumm}\{\text{y (n), y (n)}\} \tag{3.12}$$

$$C_{21} = \text{E}\ [|y\ (n)|^2] = \text{cumm}\{\text{y (n)},y^*(n)\} \tag{3.13}$$

$$C_{40} = M_{40} - 3M_{20}{}^2 = \text{cumm}\{\text{y (n), y (n), y (n), y (n)}\} \tag{3.14}$$

$$C_{41} = M_{40} - 3M_{20}M_{21} = \text{cumm}\{\text{y (n), y (n)},y^*(n)\} \tag{3.15}$$

$$C_{42} = M_{42} - |M_{20}|^2 - 2M_{21} = \text{Cumm}\ \{\text{y (n), y (n)},y^*(n),y^*(n)\} \tag{3.16}$$

$$C_{60} = M_{60} - 15\,M_{20}M_{40} + 30M_{20}{}^3 = \text{Cumm}\{\text{y (n), y (n), y (n), y (n),y(n),y(n)}\} \tag{3.17}$$

$$C_{61} = M_{61} - 5M_{21}M_{40} - 10M_{20}M_{41} + 30M_{20}{}^2M_{21} =$$
$$\text{cumm}\{\text{y(n),y(n),y(n),y(n),y(n)},y^*(n)\} \tag{3.18}$$

$C_{62} = M_{62} - 6M_{20}M_{42} - 8M_{21}M_{41} - M_{22}M_{40} + 6M_{20}{}^2M_{22} + 24M_{21}{}^2M_{22}$ = cumm {y (n), y (n),y(n),y(n),$y^*(n), y^*(n)$} (3.19)

$C_{63} = M_{63} - 9M_{21}M_{42} + 12M_{21}{}^3 - 3M_{20}M_{43} - 3M_{22}M_{41} + 18M_{20}M_{21}M_{22}$ = cumm {y (n), y (n),y(n),y(n),$y^*(n), y^*(n), y^*(n), y^*(n)$} (3.20)

$C_{80} = M_{80} - 35M_{40}{}^2 - 28M_{60}M_{20} + 420M_{40}M_{20}{}^2 - 630M_{20}{}^4$ = cumm{y (n), y (n), y(n), y (n), y (n), y (n), y (n),y(n)} (3.21)

$C_{84} = M_{84} - 16C_{63}C_{21} + |C_{40}|^2 - 18C_{42}{}^2 - 72\, C_{42}C_{21}{}^2 - 24\, C_{21}{}^4$= cumm{y (n), y (n), y (n), y (n), $y^*(n), y^*(n), y^*(n)$} (3.22)

We have calculated altogether the higher order features for the considered modulation formats. Table 3.1 displays the theoretical values of moments and Cummulants for the considered signal form. These results are obtained in the constraint of unit variance and free of noise also normalized by signal power, i.e., these results are attained by assuming that the signal is clean and has an infinite length. Although, practically, signals commonly subject to some kind of distortion, this distortion is mostly faced during the transmission or inside the transmitter and has a finite length [47].

Theoretical value of moments and Cummulants of considered digital signal types are shown in Table 3.1.

Table 3.1. Theoretical values of Higher order Moments & Cumulants

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
M20	1	0.1	3.8	0.36	0.9	0.73
M21	2	0.1	0.8	0.04	0.08	0.04
M40	2	0.9	0.9	0.73	0.2	0.62
M42	0	6.9	68.4	203	748.8	3535.26
M60	0.027	0.09	0.12	0.2	0.13	0.32
M63	1	2.8	16.8	36.3	96.1	312.5

M84	1	4	11.6	75.89	78.80	1108.5
C20	1	0.1	3.8	0.36	0.98	0.73
C21	1	2	5.8	10.1	19.3	42.04
C40	2	0.9	0.9	0.73	0.21	0.62
C42	2	1	0.9	0.66	0.67	0.61
C63	13	27.12	1157.8	6101.73	43319.7	445472.3

3.4. Classifier

The features extracted from received signal are fed into the classifier. The classifier classify the different modulations format on the basis of features. There are different types of classifier but in this thesis we consider SVM and FFBPNN classifier for the classification of different QAM format.

3.4.1 Support Vector Machine (SVM)

For classification of signal SVM classifier is selected. SVM is evolved from statistical learning theory, which indicates a lot of exceptional benefits in resolving insignificant learning problems, analysis of regression and high dimensional pattern recognition. SVM used its kernel functions for mapping the input data with the high dimensional feature space where nonlinear classification is treated as linear classification. When comparison is made between SVM and neural network, we came to know that SVM has a solid mathematical model, which can efficiently resolve the construction problem of high dimensional data model in the finite set of samples, and can converge to global best.

The SVM basic for solving the best linear hyper plane which could classify all the signals completely. Considered the training data as below:

$$\{(x_1, y_1), x_2, y_2), \ldots (x_i, y_i), x \in Rd, y \in \{+1, -1\}\} \tag{3.23}$$

Whereas x_i represent the feature space, $y_i = +1$ means that the signal belongs to the first class, $y_i = -1$ shows that the signal is member of second class. Such kind of data are separated through hyper plane w·x+b = 0. When training data are linearly distinguishable. Then the solution for optimal plane problem is the optimization problem.

Min*imize* $\frac{1}{2} \|w\|^2$

With reference to $y_i (w.x_i + b) \geq 1$

Lagrange multiplier is introduced for solving the quadratic programming problems and the best decision function is obtained by,

$$f_x = sign[\sum_{i=1}^{n} \alpha_i y_i (x_i, x) + b] \tag{3.24}$$

Whereas α_i is known as Lagrange multiplier.

For classification of nonlinear data, SVM make comparison by nonlinearly of training data with high dimensional feature space through its kernel function afterwards it is processed as linear classification. Decision function is given as,

$$f_x = sign[\sum_{i=1}^{n} \alpha_i y_i k(x_i, x) + b] \tag{3.25}$$

Whereas $k(x_i, x)$ indicates kernel function. The typical kernel functions consist of radial basis function

$$K(x, y = \exp(-\| x - y \|^2 / 2\sigma^2) \text{ and polynomial } K(x, y) = (1 + x \cdot y)^d. \tag{3.26}$$

In short, Modulation classification is based on SVM which includes followings steps,

i. **Feature Extraction**:

Some key features are extracted after which they are converted according to their SVM data format.

ii. **kernel function selection:**

Suitable kernel function is selected. Generally RBF kernel function is used.

iii. **kernel function parameter calculation:**

The best kernel function parameters with cross validation are determined.

iv. **Samples training:**

Sampled signals are trained and classifier model is obtained.

v. **Signals classification:**

Data are classified according to obtained model in the training phase [48].

Now a days SVM have got eminence in machine learning field also the classification of signal. Classification is done by recognizing the linear and nonlinear separation at the input.

By classification using SVM, separating function are specified by combining the kernels linearly which are associated by Support Vectors as

$$f_{(x)} = \sum_{x_j \in S} \alpha_j y_j K(x_j, x) + b \tag{3.27}$$

Where x_i indicates the training pattern, $y_i \in \{+1, -1\}$ represents the corresponding labels of class and S represents the Support Vectors set.

Dual formulation gives

$$\min_{0 \le \alpha_i \le C} W = \tfrac{1}{2} \sum_{i,j} \alpha_i Q_{ij} \alpha_j - \sum_i \alpha_i + b \sum_i y_i \alpha_i \tag{3.28}$$

Where α_i represent the corresponding coefficients and b denotes the offset, $Q_{ij} = y_i y_j$ K (x_i, x_j) shows the symmetric kernel matrix with positive definite and C indicates the used parameter to discipline error points for the case of inseparable. The dual conditions for Karush Kuhn Tucker (KKT) are stated as below

$$g_i = \frac{\partial W}{\partial \alpha_i} = \sum_i Q_{ij} \alpha_j + y_i b - 1 = y_i f(x_i) - 1 \tag{3.29}$$

and

$$\frac{\partial W}{\partial b} = \sum_j y_j \alpha_j \tag{3.30}$$

This divides the training data set into S set of Support Vector ($0 < \alpha_i < C$, $g_i = 0$), E is the set of errors ($\alpha_i = C$, $g_i < 0$) and R is the respectable classified set ($\alpha_i = 0$, $g_i > 0$).

When the error points are punished quadratically by C' penalty factor, then it showed that the problem decreases to a separable case with C = ∞. Afterward, modification in kernel function is done as

$$K'(x_i, x_j) = K(x_i, x_j) + 1/C' \delta_{ij} \tag{3.31}$$

Whereas $\delta_{ij} = 1$ when i = j and $\delta_{ij} = 0$ otherwise. The benefit for this formulation is that the SVM problem decreases in case of linearly separable.

It has been found that training of SVM includes the solution for quadratic optimization problem that contains the optimization routines usages from the numerical libraries. This stage is computationally demanding, which could be subjected to the problems of stability and it is not important to implement. Some appealing iterative systems for examples Nearest Point Algorithm (NPA), Sequential Minimal Optimization (SMO), etc. are being proposed to solve such kind of problem [49].

Because of machine learning based SVM has a structural risk that creates the hyper plane with N dimensions which separate the input data into different classes. The sigmoid kernel function of SVM is equal to two layer feed forward neural network. Furthermore, it can use polynomial function or radial basis function (RBF) in which network weights are obtained by solving the quadratic programming problem with linear constraints. Therefore, multiclass SVM classifier (MCSVM) is proposed. Figure 3.2 contains its hierarchical structure. The SVM classifier are presented on the basis of statistical learning theory. The SVM algorithms used in the middle of 1990's started with more accessibility of computing power, which build the strong path for several real time applications. The SVM basic deal with two class classification; however, it may perhaps work well for multiclass classification [50].

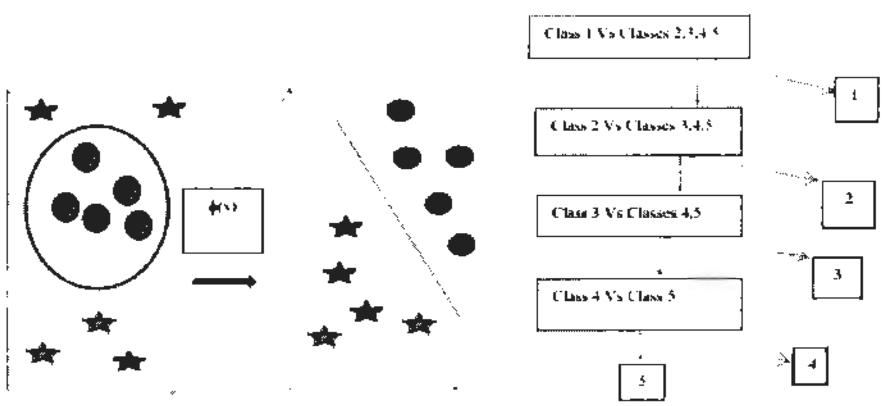

Fig. 3.2. The SVM Classifier

SVM algorithm has an empirical model and it include state of the art existing classification techniques. The SVM classifier is actually a classifier with two classes and it is based on the concepts of large margins and map- ping the data with a high dimensional space, and the SVM kernel functions. The first task of SVM is to maximize of the margin between the two adjacent data points that belongs to two different classes. The second aim is to make a constraint so that all data points be in the right class. This is solution for the two class that uses multi dimension features. Then two main goals of the SV classifier problem are then integrated in the optimization problem. SVC differentiate the points between two different set of linear classes by solving the problem of quadratic optimization for determining the most favorable separating hyper plane amongst two classes. This hyper plane increase the distance of the convex bodies for each class. These methods can be prolonged for nonlinear cases by inserting the input data in the nonlinear space by using their kernel functions. This SVC robustness initiates from solid statistical learning theory fundamental. SVC may be implemented to a separable data it can also be implemented to non-separable data points. In case of non-separable data points, one more design is added in the algorithm. This factor is the error weight which is affected by the points existing in the inappropriate class. In modulation classification, this problem arises at low SNR value. One more liberty of using this technique we have used this to classify the six different types of modulations, firstly we have classify the one modulation class from the rest and when the features of received signal do not belong to the one class and belongs to some other class, then we eliminate that class and select one class among other classes and classify that class from other classes and this process continues until the received signal is classify correctly [51].

3.4.2 Feed forward Back Propagation Neural Network Classifier

The key features of neural networks are their capability to determine the complex nonlinear input output relations that used sequential training processes, and adjust themselves to the data. Now a days for training of neural network back propagation is the most widely used search method. BPNN depend on the gradient algorithm in order to achieve the model weights and used the BP algorithm for minimizing of objective function. BPNN usually made up of three layers: fist is input layer, second is hidden layer and third is output layer [52]. Figure 3.3 illustrate the back propagation neural network architecture.

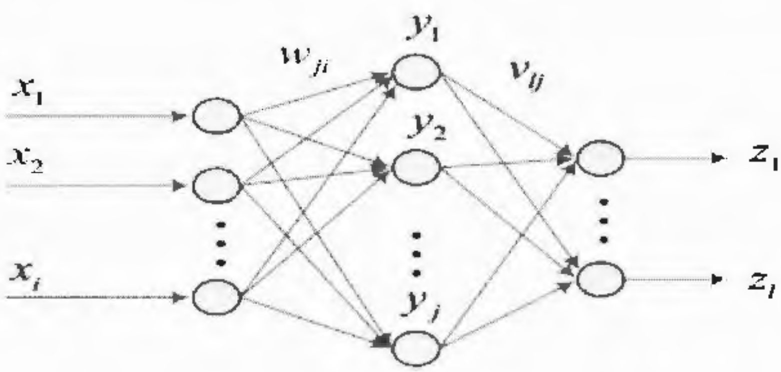

Fig.3.3. the Architecture of three layer of BPNN

The most fundamental processing elements called artificial neurons in BP network simulate the basic biotic neurons. The summation function of the neuron will be performed as in Equation (3.32).

$$b_j = T\left(\sum_i r_{ji} a_i + \delta_j\right) \quad (3.32)$$

$$y_i = f\left(\sum_i w_{ji} x_i + \theta_j\right) \quad (3.33)$$

$$z_l = g\left(\sum_j v_{lj} y_i + \phi_l\right) \quad (3.34)$$

$$E = \tfrac{1}{2} \sum_l (t_l - z_l)^2 \quad (3.35)$$

Where b_j is the j neuron activation level, T shows its transfer function, r_{ji} is the weighted value, δ_j is the bias. Therefore, the hidden layer outcome and the output layer defined by the equations (3.33) and (3.34). The output neuron error is specified in equation (3.35). In which x_i are the signal input and z_l are signals output, y_j are the hidden layer output, w_{ji} are the weights among input neuron j and hidden neuron i . v_{lj} is the weight among the hidden neuron i and output neuron j . ϑ_j is the hidden layer bias and $\emptyset_i$ is the output layer bias. f is the transfer function of hidden layer and g is the transfer functions of output layers. t_l is the projected output. E represent the error among the calculated and expected outputs. The main limitations of the BP algorithm are its slowness in convergence speed and its inability to escape local optima. Most frequently-used methods to improve upon the original BP focus on adding momentum values, changing learning rate and employing Levenberg Marquardt algorithm. Because of the gradient nature of BP neural network, those limitations of BP can be eliminated by adopting global search techniques, such as particle swarm optimization.

BPNN algorithm Explanation:-

The neural network structure consist of three layers with multiple input nodes which are determined by multiple features that multiple neurons outputs are equals to the classified modulation schemes. The neuron are present in the hidden layer and neuron numbers are arbitrary, it also depends on various situations of application and training algorithms. BPNNN consist of weights and neurons. The neurons contains the input nodes, output nodes and hidden nodes. The three layers of neurons are given below.

1. Input layer
2. Hidden layer
3. Output layer

BP algorithm can be summarized as follows:

Step 1 Initialization of weights for random values and setting of bias.

Step 2 Calculation of network output value.

Step 3 Error Calculation according to output targets.

Step 4 weight changes calculations from hidden layer to output layer.

Step 5 weight changes calculations from input layer to hidden layer.

Step 6 Repetitions from step 2 to step 5 till the cumulative error conform the stopping criteria [53].

3.5 Simulation Results

In this part, the under considered digital signals are simulated in the MATLAB environment. The complex baseband signal is consider. The modulated signal considered here are QAM2, QAM4, QAM 8, QAM 16, QAM 32, and QAM 64. The modulation classification is also performed in the presence of additive white Gaussian noise (AWGN). For every single trial, the random digital signal information is generated, to make sure that the result are independent of transmitted message. Estimation of features for all types of signal types is depend on the theoretical formula is explained in section 3.3. This section of thesis present the proposed classifier simulations result. The AWGN noise with different SNR values is added in the signals to check classification accuracy of proposed classifier with different number of samples i.e. 512, 1024, 2048 and 4096. Also classification accuracy performance is also checked on Rayleigh flat fading channel and Rician Channel. At the end comparison is made on classification accuracy for different channels.

3.5.1 Classification Accuracy on AWGN Channel

AMC of considered modulations schemes are firstly performed on AWGN Channels. As the number of samples taken are 512, 1024, 2048 and 4096. Then the classification accuracy is analyzed.

3.5.1.1 Classification accuracy with 512 number of Samples

The tables from 3.2 to 3.7 shows the classification accuracy of considered signal schemes with 512 number of samples on AWGN channel at three different SNR values of -10dB, 0dB and 10dB. As we see from results of both training and testing, the classification accuracy increase as SNR value increases. The tables shows that the average classification of training at SNR -

10dB is 93.15 and this average increase to 94.15 at 0dB SNR and finally reaches to 95.2 at 10 dB SNR. The average classification accuracy of testing at 0 dB SNR is 87.51, 90.83 at SNR 5dB and 91.98dB at 10dB SNR.

Table 3.2. Training data at SNR= 0dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	88.9					
QAM4		91.7				
QAM8			94.4			
QAM16				96.6		
QAM32					100	
QAM64						89.9

Table 3.3. Testing data at SNR= 0dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	83.3					
QAM4		84.4				
QAM8			80.6			
QAM16				92.8		
QAM32					100	
QAM64						84.1

Table 3.4. Training data at SNR = 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	89.9					
QAM4		92.4				
QAM8			94.4			
QAM16				96.6		
QAM32					99.9	
QAM64						91.7

Table 3.5. Testing data at SNR = 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	87.5					
QAM4		91.1				
QAM8			92.4			
QAM16				94.2		
QAM32					93.3	
QAM64						86.5

Table 3.6. Training data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	91.1					
QAM4		93.8				
QAM8			96.4			
QAM16				96.6		
QAM32					100	
QAM64						93.3

Table 3.7 Testing data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	88.9					
QAM4		91.8				
QAM8			93.1			
QAM16				94.6		
QAM32					93.6	
QAM64						89.9

3.5.1.2 Classification accuracy with 1024 number of samples

The tables from 3.8 to 3.13 shows the percentage of correct classification at different SNR values on AWGN channel with 1024 number of samples. The tables shows that classification accuracy increase as the SNR value increases also when we compare the result with the results of tables from 3.2 to 3.7 , we come to know that these result are better than with the result of table from 3.2 to 3.7. So, as the number samples increases the percentage of correct classification also increases. The tables from 3.8 to 3.13 shows that, the average classification accuracy at 0dB SNR is 94.95, 95.34 at 5dB SNR and 96.85 at 10dB SNR.

Table 3.8 Training data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	93.3					
QAM4		94.6				
QAM8			100			
QAM16				97.6		
QAM32					94.4	
QAM64						89.8

Table 3.9 Testing data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	87.2					
QAM4		91.4				
QAM8			100			
QAM16				94.6		
QAM32					94.4	
QAM64						86.2

Table 3.10 Training data at SNR = 5 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	90.8					
QAM4		94.6				
QAM8			95.6			
QAM16				96.6		
QAM32					100	
QAM64						94.4

Table 3.11 Testing data at SNR = 5 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	89.9					
QAM4		92.4				
QAM8			93.3			
QAM16				92.2		
QAM32					97.2	
QAM64						91.4

Table 3.12 Training data at SNR = 10dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	93.8					
QAM4		95.8				
QAM8			96.4			
QAM16				100		
QAM32					98.5	
QAM64						96.6

Table 3.13 Testing data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	91.6					
QAM4		93.2				
QAM8			93.8			
QAM16				95.8		
QAM32					96.6	
QAM64						92.2

3.5.1.3 Classification accuracy with 2048 number of samples

The tables from 3.14 to 3.19 shows the percentage of correct classification accuracy at SNR values of -10dB, 0dB and 10dB on AWGN channel with 2048 number of samples. Training and testing results are better than the results obtained with 512 and 1024 number of samples. The average classification with 2048 number of samples is 96.14 at 0dB SNR, 96.81 at 5dB SNR and 97.6 at 10dB SNR.

Table 3.14 Training data at SNR = 0dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	97					
QAM4		94.4				
QAM8			92.4			
QAM16				100		
QAM32					100	
QAM64						93

Table 3.15 Testing data at SNR = 0dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.4					
QAM4		93.3				
QAM8			91.5			
QAM16				100		
QAM32					100	
QAM64						92.2

Table 3.16 Training data at SNR= 5 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	94.4					
QAM4		96.6				
QAM8			96.6			
QAM16				97.6		
QAM32					100	
QAM64						95.7

Table 3.17 Testing data at SNR = 5 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	92.4					
QAM4		94.4				
QAM8			94.4			
QAM16				95.5		
QAM32					100	
QAM64						94.6

Table 3.18 Training data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.8					
QAM4		96.6				
QAM8			97.6			
QAM16				98.2		
QAM32					100	
QAM64						96.4

Table 3.19 Testing data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	95.8					
QAM4		95.6				
QAM8			96.6			
QAM16				96.6		
QAM32					100	
QAM64						94.5

3.5.1.3. Classification accuracy with 4096 number of Samples

The tables from 3.20 to 3.25 shows the percentage of correct classification with 4096 number of samples on AWGN channel. The training and testing data results shows that average classification accuracy with 4096 number of samples is higher than the classification accuracy with 512, 1024 and 2048 number of samples. Also the classification accuracy increase with the increase of SNR values.

Table 3.20. Training data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	98.8					
QAM4		94.6				
QAM8			94.6			
QAM16				94.4		
QAM32					99.4	
QAM64						100

Table 3.21. Testing data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.6					
QAM4		94.4				
QAM8			93.3			
QAM16				94.4		
QAM32					98.8	
QAM64						100

Table 3.22. Training data at SNR = 5 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	95					
QAM4		97.8				
QAM8			96.2			
QAM16				100		
QAM32					98.8	
QAM64						95.8

Table 3.23. Testing data at SNR= 5 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	94.8					
QAM4		97.2				
QAM8			96			
QAM16				100		
QAM32					98.8	
QAM64						95.2

Table 3.24. Training data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.7					
QAM4		98.8				
QAM8			98.2			
QAM16				100		
QAM32					100	
QAM64						98.8

Table 3.25. Testing data at SNR= 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.6					
QAM4		96.6				
QAM8			97.4			
QAM16				98.2		
QAM32					100	
QAM64						96.4

3.5.2 Classification accuracy on Rician Channel

The classification accuracy is also tested on Rician channel and compared with AWGN channel accuracy.

3.5.2.1 Classification accuracy with 512 number of Samples

The tables from 3.26 to 3.31 shows the percentage of correct classification of training and testing at different SNR values of 0dB, 5dB and 10dB respectively. The average classification accuracy at 0dB is 92.31, at 5dB is 93.93 and at 10dB is 94.13. Results shows that the classification accuracy rise with the rise of SNR. But Rician Channel has less classification accuracy than the AWGN channel with same number of samples and SNR values.

Table 3.26. Training data at SNR= 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	78.6					
QAM4		88.9				
QAM8			98.8			
QAM16				94.4		
QAM32					98.8	
QAM64						94.4

Table 3.27. Testing data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	61.1					
QAM4		86.9				
QAM8			94.4			
QAM16				92.9		
QAM32					94.4	
QAM64						88.1

Table 3.28. Training data at SNR = 5 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	88.7					
QAM4		90.5				
QAM8			95.2			
QAM16				96.4		
QAM32					95.2	
QAM64						97.6

Table 3.29. Testing data at SNR = 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	80.5					
QAM4		83.35				
QAM8			94.4			
QAM16				94.4		
QAM32					94.4	
QAM64						94.4

Table 3.30. Training data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	89.9					
QAM4		90.5				
QAM8			95.2			
QAM16				95.2		
QAM32					97.6	
QAM64						96.4

Table 3.31. Testing data at SNR= 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	88.9					
QAM4		88.9				
QAM8			94.4			
QAM16				94.4		
QAM32					94.4	
QAM64						88.9

3.5.2.2. Classification Accuracy with 1024 number of samples

The tables from 3.32 to 3.37 shows the training and testing classification accuracy with 1024 number of samples on Rician channel. The training and testing results shows that classification accuracy is higher than the 512 number of samples on Rician channel but has less classification accuracy as compared with the result on AWGN channel. However, the classification accuracy increase with the increase of SNR. The average classification accuracy of at 0 SNR is 94.6, 5dB SNR is 94.85 and 10 dB SNR is 95.7.

Table 3.32. Training data at SNR = 0dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	89.7					
QAM4		92.4				
QAM8			93.7			
QAM16				97.6		
QAM32					98.4	
QAM64						96.3

Table 3.33. Testing data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	88.9					
QAM4		87.05				
QAM8			92.6			
QAM16				96.3		
QAM32					92.6	
QAM64						92.6

Table 3.34. Training data at SNR = 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	93.7					
QAM4		95.2				
QAM8			93.7			
QAM16				87.3		
QAM32					99.2	
QAM64						100

Table 3.35. Testing data at SNR = 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	87.2					
QAM4		92.6				
QAM8			92.6			
QAM16				81.5		
QAM32					96.3	
QAM64						100

Table 3.36. Training data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	97.6					
QAM4		99.2				
QAM8			89.7			
QAM16				97.6		
QAM32					90.1	
QAM64						100

Table 3.37. Testing data at SNR = 10dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.3					
QAM4		92.6				
QAM8			88.9			
QAM16				96.3		
QAM32					88.9	
QAM64						96.3

3.5.2.3 Classification Accuracy with 2048 number of samples

The tables from 3.38 to 3.43 shows the percentages of correct classification on Rician channel with 2048 number of samples. The result has better classification accuracy compare to the results with 512 and 1024 number of samples. Moreover, the percentage of classification increases with the rise of SNR. These results have less classification accuracy than the AWGN channel results.

Table 3.38. Training data at SNR= 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	98.8					
QAM4		93.5				
QAM8			93.9			
QAM16				95.2		
QAM32					88.9	
QAM64						98.2

Table 3.39. Testing data at SNR= 0dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.4					
QAM4		92.7				
QAM8			91.2			
QAM16				93.6		
QAM32					83.3	
QAM64						96.4

Table 3.40. Training data at SNR= 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	99.4					
QAM4		89.9				
QAM8			95.7			
QAM16				97.2		
QAM32					93.06	
QAM64						99.2

Table 3.41. Testing data at SNR= 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	94.4					
QAM4		86.1				
QAM8			94.4			
QAM16				97		
QAM32					90.7	
QAM64						96.3

Table 3.42. Training data at SNR= 10dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	94					
QAM4		94.6				
QAM8			100			
QAM16				99.4		
QAM32					94	
QAM64						93.75

Table 3.43. Testing data at SNR= 10dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	91.7					
QAM4		94.4				
QAM8			97.2			
QAM16				97.2		
QAM32					88.9	
QAM64						93.05

3.5.2.4 Classification Accuracy with 4096 number of samples

The tables from 3.44 to 3.49 shows the classification accuracy with 4096 number of samples on Rician channel at different SNR values of 0dB, 5dB and 10 dB respectively. The training and testing results shows that classification accuracy increases with the increase of SNR. Also the classification with 4096 has higher accuracy percentage as compared with the 512, 1024 and 2048 number of samples. The classification on Rician channel has less accuracy when compared to AWGN channel accuracy.

Table 3.44. Training data at SNR = 0dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	98					
QAM4		93.1				
QAM8			93.3			
QAM16				96.6		
QAM32					93.3	
QAM64						98.2

Table 3.45. Testing data at SNR= 0dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.6					
QAM4		92.8				
QAM8			92.2			
QAM16				94.8		
QAM32					89.9	
QAM64						97

Table 3.46. Training data at SNR = 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	98.4					
QAM4		95.8				
QAM8			96.3			
QAM16				96.6		
QAM32					94.6	
QAM64						99.2

Table 3.47. Testing data at SNR = 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	97					
QAM4		93.1				
QAM8			92.8			
QAM16				95.1		
QAM32					90.1	
QAM64						97.3

Table 3.48. Training data at SNR= 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	99.1					
QAM4		96.6				
QAM8			97.4			
QAM16				96.6		
QAM32					96.6	
QAM64						99.8

Table 3.49. Testing data at SNR= 10dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	97.8					
QAM4		93.8				
QAM8			94.6			
QAM16				95.8		
QAM32					94.3	
QAM64						98.2

3.5.3 Classification Accuracy on Rayleigh Channel

The classification accuracy is also tested on Rayleigh channel and classification accuracy is also compared with AWGN channel and Rician channel accuracy.

3.5.3.1. Classification Accuracy with 512 number of samples

The tables from 3.50 to 3.55 shows the percentage of correct classification accuracy on Rayleigh channel with 512 number of samples at SNR values of 0dB, 5dB, 10dB respectively. The tables shows that classification accuracy increases with increase of SNR. The classification results on Rayleigh channel has less classification accuracy than the results on rician and awgn channels. However, the average classifications for each SNR values are 90.55, 92.83 and 94.11 respectively.

Table 3.50. Training data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	81.5					
QAM4		82.7				
QAM8			88.9			
QAM16				97.6		
QAM32					96.4	
QAM64						96.2

Table 3.51. Testing data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	77.8					
QAM4		77.8				
QAM8			88.8			
QAM16				89.9		
QAM32					90.2	
QAM64						91.4

Table 3.52. Training data at SNR = 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	84.5					
QAM4		96.4				
QAM8			95.2			
QAM16				94		
QAM32					95.2	
QAM64						91.7

Table 3.53. Testing data at SNR= 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	83.3					
QAM4		94.4				
QAM8			88.9			
QAM16				83.3		
QAM32					94.2	
QAM64						83.3

Table 3.54. Training data at SNR= 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	83.3					
QAM4		97.8				
QAM8			97.2			
QAM16				96.8		
QAM32					95.2	
QAM64						94.4

Table 3.55. Testing data at SNR= 10dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	81					
QAM4		97				
QAM8			93.5			
QAM16				94.4		
QAM32					94.4	
QAM64						90

3.5.3.2 Classification Accuracy with 1024 number of samples

The table from 3.56 to 3.61 shows considered modulation schemes classification accuracy on Rayleigh channel with 1024 number of samples with three different SNR values. The results include the training and testing for each modulation schemes. The average classification of both training and testing increases with the increase of SNR values. Also the classification accuracy result for this number of samples are higher than the classification accuracy then the 512 number of samples on the same channel. However, the classification accuracy has less accuracy than the classification accuracy on rician and awgn channels.

Table 3.56. Training data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	87.5					
QAM4		83.7				
QAM8			96.2			
QAM16				91.3		
QAM32					94.2	
QAM64						93.3

Table 3.57. Testing data at SNR= 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	87					
QAM4		82.6				
QAM8			91.3			
QAM16				91.3		
QAM32					87	
QAM64						82.6

Table 3.58. Training data at SNR = 5 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	86.9					
QAM4		97.6				
QAM8			90.5			
QAM16				96.4		
QAM32					95.2	
QAM64						91.7

Table 3.59. Testing data at SNR= 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	83.3					
QAM4		94.4				
QAM8			88.9			
QAM16				88.9		
QAM32					94.4	
QAM64						77.8

Table 3.60. Training data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	87.6					
QAM4		98.2				
QAM8			97.6			
QAM16				96.7		
QAM32					95.4	
QAM64						94.4

Table 3.61. Testing data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	87.5					
QAM4		97.8				
QAM8			93.3			
QAM16				94.4		
QAM32					94.4	
QAM64						89.9

3.5.3.3 Classification Accuracy with 2048 number of samples

The tables from 3.62 to 3.67 shows the training and testing classification accuracy with 2048 number of samples on Rayleigh channel. The table results shows that the classification accuracy increases with increase of SNR. Also, classification accuracy is higher than the classification accuracy of 512 and 1024 numbers of samples on Rician channel. When we compared the classification accuracy on AWGN and Rician channel. The result shows that it has less accuracy when compared to these two channels accuracy.

Table 3.62. Training data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	83.3					
QAM4		98.8				
QAM8			91.7			
QAM16				94		
QAM32					94	
QAM64						86.9

Table 3.63. Testing data at SNR = 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	80.9					
QAM4		96.6				
QAM8			83.4			
QAM16				91.9		
QAM32					92.2	
QAM64						77.8

Table 3.64. Training data at SNR = 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	95.7					
QAM4		96.2				
QAM8			83.7			
QAM16				95.2		
QAM32					96.2	
QAM64						96.2

Table 3.65 Testing data at SNR = 5dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	95.2					
QAM4		95.7				
QAM8			82.6			
QAM16				82.6		
QAM32					95.7	
QAM64						95.7

Table 3.66 Training data at SNR = 10dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.1					
QAM4		93.3				
QAM8			93.9			
QAM16				93.8		
QAM32					97.6	
QAM64						98.1

Table 3.67. Testing data at SNR= 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	95.7					
QAM4		88.3				
QAM8			91.3			
QAM16				95.7		
QAM32					94.4	
QAM64						95.7

3.5.3.4 Classification Accuracy with 4096 number of samples

The table from 3.68 to 3.73 shows the percentage of correct classification with 4096 number of samples on Rayleigh channel at different SNR values. The result shows that the classification accuracy rises with the rise of SNR. Also, the classification accuracy is higher than the classification accuracy when compared with the result of tables from 3.50 to 3.67. Classification accuracy with 4096 number of samples on Rayleigh channel has less as compared with the classification accuracy on Rician and AWGN channels.

Table 3.68 Training data at SNR= 0 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	95.7					
QAM4		88.6				
QAM8			91.3			
QAM16				95.7		
QAM32					94.4	
QAM64						95.7

Table 3.69 Testing data at SNR= 0dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	89.8					
QAM4		83.3				
QAM8			88.4			
QAM16				94.4		
QAM32					92.6	
QAM64						87.1

Table 3.70 Training data at SNR = 5 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.4					
QAM4		95.2				
QAM8			90.2			
QAM16				92.6		
QAM32					97.6	
QAM64						95.2

Table 3.71 Testing data at SNR= 5 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	93.3					
QAM4		94.6				
QAM8			88			
QAM16				91.4		
QAM32					92.4	
QAM64						94.8

Table 3.72 Training data at SNR = 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	97					
QAM4		94				
QAM8			94.2			
QAM16				97.6		
QAM32					96.5	
QAM64						98.1

Table 3.73 Testing data at SNR= 10 dB

	QAM2	QAM4	QAM8	QAM16	QAM32	QAM64
QAM2	96.8					
QAM4		93.5				
QAM8			92.8			
QAM16				94.4		
QAM32					92.9	
QAM64						96.6

3.5.4 Comparison of Classification Accuracy

The tables 3.74 and 3.75 shows respectively training and testing average classification accuracy of QAM 16 modulation format on different channels at three different SNR values.

Table 3.74. Training Performance comparison of recognizer for different channel at different SNR values

Training

Channel	No. of Samples	10 dB	5 dB	0 dB
AWGN	512	95.2	94.15	93.15
	1024	96.85	95.34	94.95
	2048	97.6	96.81	96.14
	4096	98.75	97.26	96.96
Rician	512	94.13	93.93	92.31
	1024	95.7	94.85	94.6
	2048	95.95	95.73	94.57
	4096	96.75	96.83	95.42
Rayleigh	512	94.11	92.83	90.55
	1024	94.98	93.05	91.03
	2048	95.46	93.86	91.45
	4096	96.24	94.54	93.56

The table 3.74 shows the comparison of training result of considered modulation schemes with different numbers of samples at different SNR values and on different channels. The tables shows that the classification accuracy increases with the increase of SNR and also increase as the number of samples increases. The table also show that the classification accuracy on AWGN channel is higher than the Rician and Rayleigh channels. The Rician channel has higher accuracy rate than the Rayleigh channel. The Rayleigh channel has least percentage of classification accuracy.

Table 3.75. Testing Performance comparison of recognizer for different channel at different SNR values

Channel	No. of Samples	10 dB	5dB	0 dB
AWGN	512	91.98	90.83	87.51
	1024	93.86	92.74	92.3
	2048	96.52	96.14	95.56
	4096	97.54	97	96.25
Rician	512	91.65	90.24	86.3
	1024	93.21	91.7	91.67
	2048	93.74	93.15	92.26
	4096	95.75	94.24	93.89
Rayleigh	512	91.25	87.9	84.25
	1024	92.89	87.95	86.96
	2048	93.52	91.25	87.13
	4096	94.50	92.42	89.27

The table 3.75 shows the testing classification accuracy for considered modulation schemes on different channels at different SNR values with different numbers of samples. The table shows that the training classification accuracy of considered modulation schemes is increase with the increase of SNR values, also increases as the number of samples increases. The table also shows that the percentages of correct classification also varies on different channels. It shows that AWGN channel has most classification accuracy than Rician and Rayleigh channel. The least classification accuracy is on Rayleigh channel.

It is also clear from results that the testing classification accuracy of considered modulation schemes has less classification accuracy than training classification accuracy.

3.6 Summary

Automatic modulation classification is performed by using feature extraction method. Higher order Cumulants and moments based features of received signal are extracted. Then these features are fed into the classifier for classification of modulation scheme. The proposed classifier are based on Support vector Machine and feed forward back propagation neural network classifier. The automatic modulation is performed in the presence of additive white Gaussian noise (AWGN). Classification accuracy is compared on different number of samples at different SNR's. Comparison of classification accuracy is made on AWGN, Rician and Rayleigh channels. By simulation results it can be concluded that as number of samples increases the classification accuracy also increased. Moreover, classification accuracy increase with the increase of SNR. From simulation results we can also conclude that classification accuracy of AWGN channel is greater than other channel. Classification accuracy of Rician channel is greater than Rayleigh channel but less than that of AWGN channel. Rayleigh channel has least classification accuracy than AWGN and Rician channels.

Chapter 4
AMC using Optimized SVM

4.1 Introduction

In this chapter, results obtained in chapter 3 are optimized using evolutionary computing technique. The optimization is done by using genetic algorithm (GA). The proposed classifier performance is also compared with the classifier proposed in chapter 3 and existing state of the art techniques.

4.2 Features used for Automatic Modulation Classification.

For classification of modulation schemes pattern recognition approach has been used. A higher order statistical feature extraction method is used for AMC. The second, fourth and sixth and eighth order moments as well as Cummulants are used to extract the features of received signal [54].

4.3 Optimization

The result obtained in the previous chapter are optimized. There are different optimization techniques available but in our research Genetic Algorithm optimization technique is selected for optimization of result.

4.3.1 Optimization by using GA

SVM performance can be improved by selecting optimal values of the hyper parameters: though, it is very challenging problem. GA has high efficiency characteristics due to which it used as global optimization in several areas. Genetic Algorithm adopts the survival of the fittest theory of Darwin because GA has stochastic optimization algorithm. In order to apply GA in selection model, followings issue are considered:

First, scheme for encoding, secondly, initial population production methodology, thirdly fitness function, and fourthly, genetic operators like imitation, mutation and crossover.

The encoding of problem solution into the chromosome is the basic issue for GA. Selection of SVM parameters with constraint is another issue for optimization problem. From experiment it has been clear that variables in GA, with real coded schemes have better performance than the binary encoded scheme in signal classification. Therefore, the real encoded schemes are chosen to represent the parameters. The parameters search space or selected intervals are C ε [2:4:50], $\sigma\epsilon$ [0.1:0.1:2]. In order to avoid population convergence difficulties, the population size (POP_Size) is selected as 16. For initial population production, designed parameters initial values are distributed over the solution space consistently. The probabilistic selection function also known as normalized geometric ranking had been used in such a way that individuals which are based on the fitness function have greater chance of selection. The non-uniformly mutation function which uses the random number for mutation is based on existing number of generation and maximum number of generation , amongst the many others parameters, is accepted. The crossover which produce the linear extrapolation of two individuals is based on the fitness function information. The maximum generation number decide the termination criteria of solution process. The testing data classification success rate is used as fitness criteria in the process of function evaluation [55].

4.4 Simulation Results

Table 4.1. Training Performance comparison of recognizer for different channel at different SNR values after optimization

Channel	No. of Samples	10 dB	5 dB	0 dB
AWGN	512	97.4	96.3	95.1
	1024	99.1	97.4	96.2
	2048	99.3	97.9	98.8
	4096	99.9	99.5	98.9
Rician	512	96.5	95.1	94.12
	1024	97.2	96.8	96.71
	2048	98.3	97.9	96.8
	4096	98.7	98.1	97.9
Rayleigh	512	96.5	95.1	92
	1024	97.8	96.7	93.5
	2048	98.4	97.3	94.3
	4096	98.5	97.9	95.1

The table 4.1 shows the comparison of training classification accuracy of considered modulation schemes at different SNR values after optimization. Theses result shows that classification accuracy is increased after optimization.

Table 4.2. Testing Performance comparison of recognizer for different channel at different SNR values after optimization

Channel	No. of Samples	10 dB	5 dB	0 dB
AWGN	512	94.5	92.1	88.1
	1024	97.7	93.6	93.1
	2048	98.6	98.1	97.2
	4096	99.9	99.2	98.2
Rician	512	96.7	93.2	91.1
	1024	97.8	96.5	92.3
	2048	98.1	97.0	94.5
	4096	98.6	97.2	95.7
Rayleigh	512	93.5	92.1	90.0
	1024	96.3	93.4	91.2
	2048	97.9	94.15	92.3
	4096	98.7	95.3	93.4

The table 4.2 shows the comparison of testing classification accuracy on different channels and at different SNR values with different numbers of samples after optimization. The tables shows that the classification accuracy has improved after optimization.

4.5. Comparison with the Chapter 3 Results.

Table 4.3 comparison of classification accuracy with and without optimization at SNR 10dB

Channel	No. of Samples	10 dB With optimization	10 dB Without optimization
AWGN	512	97.4	95.2
	1024	99.1	96.85
	2048	99.3	97.6
	4096	99.9	98.75
Rician	512	96.5	94.13
	1024	97.2	95.7
	2048	98.3	95.95
	4096	98.7	96.75
Rayleigh	512	96.5	94.11
	1024	97.8	94.98
	2048	98.4	95.46
	4096	98.5	96.24

The table 4.3 shows the comparison of results without optimization and with optimization at SNR 10dB on each channel with different number of samples. The table shows clearly that the classification accuracy has increase after optimization on each channel with different number of samples respectively.

Table 4.4 comparison of classification accuracy with and without optimization at SNR 0 dB

Channel	No. of Samples	5 dB With optimization	5 dB Without optimization
AWGN	512	92.1	94.15
	1024	93.6	95.34
	2048	98.1	96.81
	4096	99.2	97.26
Rician	512	93.2	93.93
	1024	96.5	94.85
	2048	97.0	95.73
	4096	97.2	96.83
Rayleigh	512	92.1	92.83
	1024	93.4	93.05
	2048	94.15	93.86
	4096	95.3	94.54

The table 4.4 shows the comparison of results without optimization and with optimization at SNR 0dB on each channel with different number of samples. The table shows clearly that the classification accuracy has increase after optimization on each channel with different number of samples respectively.

Table 4.5 Comparison of classification accuracy with and without optimization at SNR 0 dB

Channel	No. of Samples	0 dB with optimization	0 dB without optimization
AWGN	512	95.1	87.51
	1024	96.2	92.3
	2048	98.8	95.56
	4096	98.9	96.25
Rician	512	94.12	86.3
	1024	96.71	91.67
	2048	96.8	92.26
	4096	97.9	93.89
Rayleigh	512	92	84.25
	1024	93.5	86.96
	2048	94.3	87.13
	4096	95.1	89.27

The table 4.5 shows the comparison of results without optimization and with optimization at SNR 0dB on each channel with different number of samples. The table shows clearly that the classification accuracy has increase after optimization on each channel with different number of samples respectively.

4.6 Performance comparison of proposed classifier with state of art existing techniques

Table 4.6 Performance comparison with existing techniques

Reference	Channel	Features & Classifier	No. of samples	SNR Value	Previous classification accuracy(CA)	Proposed classification accuracy (CA)
67	AWGN	4^{th} and 8^{th} order Cumulants, ANN Classifier	1024	0 dB	57.2	96.2
				10 dB	75.36	99.1
2	AWGN	discrete wavelet transform	512	0 dB	95.4	95.1
	Rayleigh				92.7	94.12
68	AWGN	Standard deviation of Phase, difference of Phase, absolute difference of Phase, Instantaneous Amplitude, Envelop	512	10 dB	88	97.4
69	AWGN	Spectral Features SVM Classifier	1024	0	89.3	96.2
				5	96.1	97.4
70	AWGN	Spectral features hierarchical classifier	2048	10 dB	97.7	99.3
71	AWGN	Clustering Algorithm, NN Classifier	512	0dB	78.4	95.1
				5 dB	93.3	96.3
				10 dB	96.4	97.4
7	AWGN	Spectral features used and ANN classifier is used	512	5 dB	80.9	96.3
				10 dB	84.6	97.4
64	AWGN	6^{th} , 8^{th} order Cummulants, SVM classifier	1024	0 dB	94.92	96.2
				5	96.85	97.4
46	AWGN	Cummulants based features, SVM Classifier	4096	0 dB	96.3	99.21
				5 dB	98.95	99.93
52	AWGN	4^{th}, $6^{th}$and 8^{th} order cummulants, SVM classifier	4096	0	98.48	99.21
				5	99.86	99.93

The table 4.6 show the classification accuracy performance of proposed classifier with the existing state of the art techniques. It is clear from above table that our proposed classifier has highest classification accuracy as compared with the other existing techniques.

Table 4.7, shows the training and testing accuracy of proposed and existing classifier, and it is found that proposed classifier performs better in both scenarios also at lower SNR's.

Table 4.7: Average Classification Accuracy Compared with [64]

Channel/No of Samples	0 dB without Optimization			
	Training [64]	Training [Proposed]	Testing [64]	Testing [Proposed]
AWGN 1024	93.26	**95.34**	91.45	**92.74**
	5 dB without Optimization			
	96.75	**96.85**	96.54	**96.68**
	0 dB with Optimization			
	94.92	**97.40**	93.64	**95.6**
	5 dB with Optimization			
	98.95	**99.12**	98.75	**98.71**

Table 4.8, shows the comparison of classification accuracy with the existing techniques at 0 dB and 5 dB of SNR with 4096 number of samples on AWGN channel with optimization.

Table 4.8: Classification Accuracy Comparison with existing techniques

Channel/No of Samples	0dB [52]	0dB Proposed	5dB [52]	5dB Proposed
AWGN 4096	98.48	**99.21**	99.86	**99.93**
	0dB [46]	0dB Proposed	5dB [46]	5dB Proposed
	96.30	**99.21**	98.95	**99.93**

4.7 Summary

The result of automatic modulation classification obtained in the previous chapter are optimized using evolutionary computing techniques. Optimization is done by using Genetic Algorithms (GA). The optimized results are compared to the result without optimization. The simulation results shows the better classifier accuracy as compared to existing techniques and results of chapter 3.

Chapter 5

Conclusion and future direction:

This chapter present the conclusion of work and achievement of our work in this thesis. It also describe the future direction work in this field.

AMC has vast significance in enhancing the consumption of the available band and enhancing the communication systems throughput. Besides, it has several applications in civilian areas as well as in military. There are two different techniques for AMC. One is likelihood based AMC and other is features based AMC. The first approach is difficult to implement while features based approach is simple to implement. Therefore, features based approach is used in this thesis. For an automatic modulation classification (AMC) higher order Cumulants based features are used and AMC is executed by combining SVM and FFBPNN classifier. Different QAM modulations types are considered for automatic modulation classification of signals. Simulations is demonstrated under AWGN, Rayleigh and Rician noise of different values of SNR's. It has been found that percentage accuracy of classification is higher at high SNR also percentage accuracy of classification rises when optimization is performed on the results of chapter 3. The proposed classifier gives high classification accuracy as compared than the simple classifier.

In this thesis higher order statistics features are used. Different features base approach with different modulation schemes can be used for AMC. Different classifier can also be used for modulation classification.

Biblography

[1] Zhao Jianli , Wang Tingting, "Identification of Cognitive Radio Modulation," International Conference on Mechatronic Science, Electric Engineering and Computer, pp. 1773-1776, 19-22 August 2011.

[2] Jian Liu,Qiang Luo, "A Novel Modulation Classification Algorithm Based on Daubechies5 Wavelet and Fractional Fourier Transform in Cognitive Radio," IEEE, pp. 115-120, 2012.

[3] Zaihe Yu, "Automatic Modulation Classification of Communication Signals", Department of Electrical and computer Engineering New Jersey Institute of Technology, August, 2006.

[4] Ijaz Mansoor Qureshi, Sajjad Ahmed Ghauri, "Automatic Classification of Digital Modulated Signals using Linear Discriminant Analysis on AWGN Channel," in 1st International Conference on Information and Communication Technology Trends (ICICTT), 2-5 September, 2013.

[5] E. Erdem, Master Thesis on digital modulation recognition, The Graduate School of natural and applied sciences of middle east techinical university, December 2009.

[6] Adalbery R. Castro, Lilian C. Freitas, Claudomir C. Cardoso, Joao C. W. A. Costa and Aldebaro B. R. Klautau, "Modulation Classification in Cognitive Radio, Foundation of Cognitive Radio System, ISBN: 978-953-51-0268-7, pages 43-60, 2012.

[7] Rada A. El-Khoribi, Mahmound Ahmed Ismail Shoman and Ahmed Galal Ahmed Mohammed, "Automatic Digital Modulation Recognition using Artificial Neural Network in Cognitive Radio," International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Giza, Cairo,Egypt, May - June 2014.

[8] M.L.D. Wong and Ashoke Kumar Nandi "Automatic digital modulation recognition using artificial neural network and genetic algorithm," Signal Processing , Elsevier, pp. 351-365, 2004.

[9] Richard Engelman, Kwaku Abrokwah, George Dillon, Gardner Foster, Gordon Godfrey, Trey Hanbury, Charlene Lagerwerff, Wayne Leighton, Mike Marcus, Roger Noel, Jerilyn Payton, Jennifer Tomchin, John Williams, Allen Yang," Federal Communications Commission Spectrum Policy Task Force", Report of the spectrum efficiency working group, pp. ET-Docket 02-135, Nov, 2002.

[10] Barathrm Ramkumar, "Automatic modulation classification for cognitive radios using cyclic feature detection," IEEE Circuits and Systems Magazine, vol. 9, no. 2, pp. 27-45, June ,2009.

[11] A. N. M.Wong, "Efficacies of selected blind modulation type detection methods for adaptive OFDM systems," 1st International Conference on Signal Processing and Communication Systems (ICSPCS), p. 243–246, December, 2007.

[12] L. Haring, Y. Chen, A.Czylwik "Automatic modulation classification methods for wireless OFDM systems in TDD mode," IEEE Transactions on Communications,, vol. 58, no. 9, p. 2480–2485, September 2010.

[13] Asas Hussain, Sajjad Ahmed Ghauri, M. Farhan Sohail, Sheraz Alam Khan and Ijaz Mansoor Qureshi "KNN based Classification of Digital Modulated Signals," IIUM Engineering Journal, 2016.

[14] Muhammad Waqar Aslam, Pattern recognition using Genetic Programming for classification of diabetes and modulation data, Liverpool: Ph.D. dissertation, the University of Liverpool, 2013.

[15] Barathrm Ramkumar, Automatic Modulation Classification and Blind Equalization for Cognitive Radios, Virginia,: Ph.D. dissertation, Virginia Polytechnic Institute and State University , 2011.

[16] Ameen Abdelmutalab, Khaled Assaleh and Mohamed El-Tarhuni, "Automatic Modulation Classification using polynomial classifier," 25th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, p. 785–789, 2014.

[17] Ruoyu Li, Modulation classification and parameter estimation in wireless networks, New York: Master's thesis, Syracuse University, 2012

[18] Fahed Hameed, Octavia A. Dobre and Dimitrie C.Popescu "On the likelihood-based approach to modulation classification," IEEE Transactions on Wireless Communications, vol. 8, no. 12, p. 5884–5892, December 2009.

[19] K.C.Ho, W.Prokopiw and Y.T.Chan, "Modulation identification of digital signals by the wavelet transform," IEE Proceedings, vol. 147, no. 4, pp. 169-176, 9th March 2000.

[20] Sajjad Ahmed Ghauri and Ijaz Mansoor Qureshi, "M-PAM Signals Classification using Modified Gabor Filter Network," Mathematical Problems in Engineering, 2015.

[21] Changyi Yin, Bringbring Li, Yanling Li"Modulation classification of MQAM signals from their constellation using clustering," in 10th Second International Conference on Communication Software and Networks (ICCSN), February 2010.

[22] A. K. a. D. Kubanek, "Algorithms of digital modulation classification and their verification," WSEAS Transactions on Communications, vol. 9, no. 9, p. 563–572, September, 2010.

[23] Sajjad Ahmed Ghauri, Ijaz Mansoor Qureshi, Basir, S., Din, H., "Modulation Classification using Spectral Features on Fading Channels," Science International, vol. 27(1), pp. 147-153, 2014.

[24] N. Geisinger, "Classification of digital modulation schemes using linear and non linear classifiers," Master's thesis, Naval Postgraduate School, California, 2010.

[25] Q. I. S. I. K. N. Sajjad Ahmed Ghauri, "Modulation Classification using Cyclostationary Features on Fading Channels," Research Journal of Applied Sciences Engineering & Technology (RJASET), vol. 7(24), pp. 5331-5339, 2014.

[26] Q. I. M. A. C. T. Sajjad Ahmed Ghauri, "Higher Order Cummulants based Digital Modulation Classification Scheme," Research Journal of Applied Sciences Engineering & Technology (RJASET), vol. 6(20), pp. 3910-3915, 2013.

[27] Q. I. Sajjad Ahmed Ghauri, "M-PAM Signals Classification using Modified Gabor Filter Network," Mathematical Problems in Engineering, 2015.

[28] M. Harbaji, "Classification of common partial discharge types in oil-paper insulation using acoustic signals," Master's thesis, American University of Sharjah, Sharijah , 2014.

[29] Q. I. M. A. C. T. Sajjad Ahmed Ghauri, "Automatic Digital Modulation Classification Technique using Higher Order Cummulants on Faded Channels," J. Basic. Appl. Sci. , vol. 4(3), pp. 1-12, 2014.

[30] Q. I. Sajjad Ahmed Ghauri, "Automatic Classification of Digital Modulated Signals using Linear Discriminant Analysis on AWGN Channel," 1st International Conference on Information and Communication Technology Trends (ICICTT), 2-5 September, 2013.

[31] Shi Qinghua and Y. Karasawa, "Non coherent Maximum Likelihood Classification Quadrature Amplitude Modulation Constellations: Simplification, Analysis, and

Extension", IEEE transactions on wireless communications, VOL. 10, NO. 4 pp. 1312-1322, 2011.

[32] G. X. a. G. Q. W. Dan, "A new scheme of automatic modulation classification using wavelet and WSVM," in 2nd International Conference on Mobile Technology, Applications and , November, 2005.

[33] H. H. S. H. Sajjad Ahmed Ghauri, "Comparison of different Population Strategies for Multiuser Detection using Genetic Algorithm," World Congress on Internet Security (WorldCIS), 2013.

[34] R. G. A. Ebrahimzadeh, "Classification of Communication signals using an optimized classifier and efficient features," The Arabian journal for Science and Engineering, vol. 35, no. 1B, pp. 225-235, 2010.

[35] J. K. a. R. Eberhart, "Particle Swarm optimization," 1995.

[36] M. Zubair, "Application of Particle swarm optimization to digital communication," 2009.

[37] O. K. a. M. El-Sharakawi, "Economic dispatch using Particle Swarm Opti-mization for combined cycle generators," IEEE/PES Power Systems Conference and Exposition, PSCE, pp. 1-9, 09, March 2009.

[38] P. M. J. M. A. B. Sajjad Ahmed Ghauri, "Channel Estimation using Continuous Particle Swarm Optimization," in IEEE conference of Information Society, UK, 2012.

[39] "mathwork," [Online]. Available: http://www.mathworks.com/help/gads/what-is-simulated-annealing.html. [Accessed 1 July 2016].

[40] L. Jacobson, "theprojectspot.com," 11 04 2013. [Online]. Available: http://www.theprojectspot.com/tutorial-post/simulated-annealing-algorithm-for-beginners/6. [Accessed 13 7 2016].

[41] k. E. Geltman, "katrinaeg.com," Kenton Hamaluik., 20 2 2014. [Online]. Available: http://katrinaeg.com/simulated-annealing.html. [Accessed 13 7 2016].

[42] K. P. RAINER STORN, "Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," Journal of Global optimization , pp. 341-359, 1996

[43] K. P. J. A. L. Rainer M. Storn, "Differential Evolution: A practicle approach to Global Optimization.," Springer , 2005.

[44] K. P. R. STORN, "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," Journal of Global Optimization, vol. 11, pp. 341-359, 1996.

[45] Fawas Zaman, Estimation of Direction of arrival for adaptive beamforming, PHD Dissertation , Department of Electronics Engineering IIUI, 2013

[46] R. G. Ata Ebrahimzadeh, "Blind digital modulation classification in software radio using the optimized classifier abd feature subset selection," ELSEVIER , pp. 50-59, 2010.

[47] I. M. Q. N. M. A. C. Sajjad Ahmed Ghauri, "Higher Order Cummulants based Digital Modulation Recognition Scheme," Research Journal of Applied Sciences, Engineering and Technology 6(20), pp. 3910-3915, 2013.

[48] J. L. Dan Liu, "A Novel Signal Recognition Algorithm Based on SVM in Cognitive Networks," National Natural Science Foundation of China (No. 60932002), Fundamental Research Funds for the Central Universities , 2010.

[49] M. N. M. S.V.N. Vishwanathan, "SSVM : A Simple SVM Algorithm," Indian Institute of Science, Bangalore 560 012, India.

[50] Digital Modulation Recognition Using Support Vector Machine Classifier," IEEE Transaction, pp. 2238-2242, 2004.

[51] J. W. Hanan M. Hamee, "Automatic Modulation Recognition for MFSK Using Modified Covariance Method," International Journal of Electrical and Computer Engineering (IJECE), vol. 5, no. 3, pp. 429-435, June 2015.

[52] Ataollah Ebrhimzadeh Sherme, "A Novel method for automatic modulation recognition", Elsevier Journal of Applied Soft Computing, pp. 453-461, 2011.

[53] Z. Q. CHEN Mei, "Cooperative automatic modulation recognition in cognitive radio," Elsevier, no. 2, pp. 46-52, 2010.

[54] Z. Z. A. K. N. Muhammad waqar Aslam, "Automatic Modulation Classification Using Combination of Genetic Programming and KNN," IEEE Transaction on wireless communicatios, vol. 11, no. 8, pp. 2742-2750, August , 2012.

[55] H. H. S. H. Sajjad Ahmed Ghauri, "Comparison of different Population Strategies for Multiuser Detection using Genetic Algorithm," in World Congress on Internet Security (WorldCIS), 2013.

[56] B. L. a. Y. L. C. Yin, "Modulation classification of MQAM signals from their constellation using clustering," 10th Second International Conference on Communication Software and Networks (ICCSN), p. 303–306, February 2010.

[57] A. K. a. D. Kubanek, "Algorithms of digital modulation classification and their verification," WSEAS Transactions on Communications, vol. 9, no. 9, p. 563–572, September, 2010.

[58] M. Harbaji, "Classification of common partial discharge types in oil-paper insulation using acoustic signals," Master's thesis, American University of Sharjah, Sharjiah, 2014.

[59] G. X. a. G. Q. W. Dan, "A new scheme of automatic modulation classification using wavelet and WSVM," 2nd International Conference on Mobile Technology, Applications and Systems,2005, pp. 1-5, November, 2005.

[60] A. Swedan, "Acoustic detection of partial discharge using signal processing and pattern recognition techniques," Master's thesis, American University of Sharjah, Sharjiah, 2010.

[61] B. L. a. Y. L. C. Yin, "Modulation classification of MQAM signals from their constellation using clustering," Second International Conference on Communication Software and Networks, ICCSN, p. 303–306, 10, February 2010.

[62] M. Zubair, "Application of Particle swarm optimization to digital communication," 2009.

[63] K. P. RAINER STORN, "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," Journal of Global Optimization 11, pp. 341-359, 1996.

[64] R. G. Ata Ebrahimzadeh, "Classification of Communication signals using an optimized classifier and efficient features," The Arabian journal for Science and Engineering, vol. 35, no. 1B, pp. 225-235, 2010.

[65] A. E. Sherme, "A novel method for automatic modulation recognition," Elsevier, pp. 453-461, 2011.

[66] Digital Modulation Recognition Using Support Vector Machine Classifier," IEEE Transaction , pp. 2238-2242, 2004.

[67] Ali Ozen and Celal Ozturk, "A Novel Modulation Recognition Technique Based on Artificial Bee Colony Algorithm in the Presence of Multipath Fading Channels" IEEE transaction , 2013.

[68] Bin lee et al., "Modulation Identification Using Neural Network for Cognitive radios", Software Defined Radio Forum Technical Conference, 2005.

[69] Xiaojie Gong, Li Bian, Qi Zhu, "Collaborative Modulation Recognition Based on SVM" Sixth International Conference on Natural Computation (ICNC 2010), 2010.

[70] Zhuan Ye, Gokhan Memik, John Grosspietsch, "Digital Modulation Classification using temporal waveform features for cognitive radios", The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07), 2007.

[71] LIU Ai-sheng, ZHU Qi, Automatic modulation classification based on the combination of clustering and neural network, The Journal of China Universities of Posts and Telecommunications, 2010.