Republic of Iraq
Ministry of Higher Education
and Scientific Research
University of Diyala
College of Science

Feature Extraction Of Modulated Signal Based on Swarm Optimization and Random Forest classifier

A Research
Submitted to the Department of Computer Science\ College of Sciences\ University of Diyala in a Partial Fulfillment of the Requirements for the Degree of Master in Computer Science

By

Batool Abd Al Hadi Sultan

Supervised By

Dr. Taha M. Hassan
Assistant Professor

Hadi A. Hamed
Assistant Professor

2020 A.D. **1441A.H.**

Chapter one

Introduction

1.1 Over view

Base-band signals might not be always appropriate to be used for the transmission over a given channel. These signals are modified to facilitate transmission. Parameters related to high-frequency carrier will be varied according to the variation of baseband signal. Shifting the range regarding the frequency spectrum of the signal is defined as modulation to gain several advantages: The long wavelength of baseband signal will need an impractically large antenna, and transmitting different modulated signals simultaneously as Frequency-Division Multiplexing (FDM) and Time Division Multiplexing (TDM). The digital and the analog modulations transform baseband signals into modulated band pass signals [1]. In conventional communication of digital communication system, the modulation formats include, modulation type, symbol rate, carrier frequency and so on. Technologies are rapidly moving toward secure and advanced communications. Rapid grow was achieved in terms of developing highly intelligent systems of communication. These systems could be vital for military and civilian applications, in which various systems of modulations are needed for securing communications. Automatic modulation classification (AMC) can be considered as in-between phase between detecting information that is carried through signals and its de-modulation provide simplicity in handling various standards of communications. Single receiver circuit could be used for the purpose of recognizing various systems of modulation, and after that de-modulate those in the coming signals that were transmitted through the use of various standards. AMC can also be used in spectrum management, interference identification, and signal confirmation.

Authorities related to civilians might want to look over their transmission so as to have a control over them in addition to monitoring and detecting the non-licensed transmitters [2] For military purposes its applications can be electronic warfare (EW), threat analysis and surveillance. Applications such as the systems of electronic support measures (ESM) are of high importance as an information source that is required for conducting target acquisition, warning, electronic counter measures (ECM), and the electronic counter-counter measures (ECCM) [3] . Automatic modulation classification (AMC) work on the basis of one the 2 methods: Pattern Recognition (PR) method or the Decision Theoretic (DT) method[4]. Probabilistic hypothesis testing arguments are used by the DT approach for the purpose of formulating recognition problem. The absence of robustness with regard to model mismatch as well as the complicated computations are the main disadvantages of DT methods. Also, because of the limitations of DT methods, they lack efficiency in the case when encountering various types related to the digital signals. With regard to PR methods, they have two major sub-systems: classifier sub-systems as well as the feature extraction sub-systems. The first one determines the signal's membership while the second will be extracting the features, also they do not require particular handling, thus they can be implemented easily. Many modulation recognition methods were stated in literature[5]. Another method was used in this thesis to identify the formation is (a Random forest). This approach is divided into four steps: signal generation and extraction features, phase improvement features, training phase for adjusting work-book structure (for determining biases and weights), and test phase, where the network performance is determined. In this thesis we used the second approach(PR) because it need less memory, simple and faster than the other techniques. More details about this approach will be discussed in later chapters.

1.2 Literature Survey

A comprehensive review of literature regarding automatic modulation techniques is presented. After improving the parameters of the signals included to increase the accuracy of the classification. In the presented chapter, brief review of different techniques for classification of modulation formats is presented.

- ***Amudha, Karthik, & Sivakumari, 2015*** [6]- A hybrid algorithm has been suggested by the authors for integrating modified artificial bee colony (MABC) with an enhanced particle swarm (EPSO) for the purpose of predicting the infiltration detection. Such algorithms have been combined for the purpose of achieving improved results of optimization, also the classification precision will be achieved via 10-folds verification approach. The results show that when a dataset is categorized with all features, the average resolution of hybrid approaches is suggested. The ABC and MABC-EPSO accuracy rate were significantly increased to 94.36%.
- **HL & Shrinivasan, 2015** [7]- Suggested anovel algorithm for distinguishing six types of digital modulation approaches (ASK-2, FSK-2, PSK-2, ASK-4, FSK-4 and PSK-4). High resolution results for new algorithms have been demonstrated even when SNR = 4dB.
- **M Hamee & Wadi, 2015** [8]- In this study, the researchers investigated a method of classification of digital MFSK signals without prior information using the modified heterogeneity method. This method used for account features to form a FSK is sensitive with the FSK forming indicator although the signal-to-noise ratio (SNR) varies. Using a single-to-all SVM-OAA rating vector performance algorithm, a rating of 6 digitally modified signals has been achieved, with a rating accuracy of 85.85 at SNR =15Db.

- **Hakimi and Ebrahimzadeh, 2015[9] -** In this research, the researchers used a smart hybrid system to identify digital signal types. The system consists of three main units: the unit of extracting parameters, the classification using probabilistic neural networks (PNN), and the optimization using the bee algorithm. The proposed system gave high ratings accuracy results even when signal-to-noise ratio (SNRs) was low. Where the workbook achieved a success rate of about 91% for SNR> 0 dB.
- **Hassanpour, Pezeshk, & Behnia, 2016 [10] -** The researchers proposed a new algorithm to identify the configuration (AMR) based on pattern recognition for the purpose of determining the digital modulation systems. There are certain types of modulation such as: BASK, BFSK, BPSK, 4-ASK, 4-FSK, QPSK, 16-QAM, and AWGN that considered a channel. Features will be extracted from received signals that are considered in time, wavelets and frequency bands, To overcome the multi-layer problem, hierarchical structure will be examined on the basis of bilateral SVMs. Simulations will demonstrate the effectiveness related to the proposed features have been digitally separated. The modified signals in an extremely noisy environments with values of SNR are extremely low; Thus, it was indicated that the optimum selection is -5 dB, while the final accuracy ratio of 98.15 has been achieved at -10 dB.
- **Kanisha & Balarishnanan, 2016 [11]-** The researchers used the speech signal as input to the content recognition, where the signal processing involved three stages :pre-processing, extraction of features, and SVM. They used different optimization algorithms, including APSO. They extracted optimal features such as triple spectral, peak signal and separate waveform (DWT). These optimized features are used as an input to the SVM demonstrated results, that optimize

algorithm (APSO) has higher accuracy compared to the current SVM linear kernel technology.

- **Almaspour & Moniri, 2016 [12] .** In this study, the researchers selected suitable features for the modified input signal and used the neural network algorithm, choosing 10 modified signals (2ASK, 4ASK, 2FSK, 4FSK, 2PSK, 4PSK, 4QAM, 16QAM, 64QAM). The presence of Gaussian noise -5 dB to 20 dB .The results showed that the use of the NN algorithm led to a significant increase in the accuracy of recognition of the type of modification.
- **Liu, Proietti, Zhang, Lu, & Yoo, 2017 [13]-** The researchers proposed a method to determine the MFI format that provides high resolution even in a low-light signal system of less than 10 dB that uses peak detection and non-linear power conversion. The proposed method can specify the type of signal modulation format: BPSK, QPSK, 8-PSK, and QAM. The experimental results showed that the proposed MFI can achieve high accuracy even when SNR is 7 dB.
- **Wang, Guo, Dou, & Lin, 2018[14]** -The study aimed the complication to identify adjustments in SNR decline. According to entropy's information features and Dempster-Shafer manual theory, new automatic adjustment recognition approaches were suggested in this paper. First, Renee Interoperability and Single Entropy have been applied for obtaining adjustment features. Second, according to standard test theory, the new basic probability assignment function (BPAF) has been introduced. Finally, Dempster-Shafer has been applied as profiler. The results of the study suggest that the novel method has the ability of obtain high degree to recognition in the lower SNR rate .
- **Rajendran et al., 2018[15]-**They studied the classification problem of the distribution wireless spectrum sensor network. They proposed a data-based model to classify the AMC automatic adjustment

and based on LSTM long-term memory. They used the time-based capacitance model and stage information from the configuration diagrams contained in the training data without relying on expert features such as periodic top order moments. The results was averaged 90% accuracy in varying SNR ranging from 0dB to 20d.

- **Kurniansyah, Wijanto, & Suratman, 2018 [16]-** In this research, extracting the executed parameters is a statistical advantage of high ranking in time domain. Order 4 is the applied statistical order. The signals information have been transmitted on transmission channel with the existence of AWGN noise interference with variable signal quality that ranged from zero to forty dB. ANN has been utilized for classifying the modification. AMC has ability of distinguishing between AM, LSB, USB, BPSK and QPSK, and 8PSK modulation. The accuracy rate related to the system in modulation classification process without using non-linear conversions is 65.5 percent on signal quality of ten dB. After that, AMC accuracy through the use of non-linear transformations on incoming signal reached 88.8 percent on signal quality ten dB.
- ***Sun et al., 2019 [17]*** This paper proposed a method for determining the format of blind modification using the decision tree dual-carrier bus system support has been trained on features extracted from the high-order stacking tool. The results suggested that the average Accuracy of the identification of optical format coordination signals can be achieved more than 94% when SNR is -5 dB.

in fact, Paper submitted by Hakimi and Ebrahimzadeh, 2015 **[9]**under title "Digital modulation classification using the bees algorithm and probabilistic neural network based on higher order statistics" The closest research with our study , have used the same type of signal and same kind of features, But in this research were used two algorithms to improve system performance (CSO & BA) and Random Forest for

classifying the signal. .the researcher depended on bee algorithm to improve system performance and the classification using probabilistic neural networks (PNN) where better results were obtained.

1.3 Thesis Contribution

The main contributions of this thesis are:

1-Utilization of Random Forest (RF) as a classifier for identification digitally modulate signal was introduced at the first time by this Thesis.

2-The improved the performance the proposed classifier Random Forest (RF)is presented based on Statistical features. by Chicken swarm optimization and Bat swarm optimization was introduced at the This technique first time by this Thesis.

1.4. The Aim of Thesis

The purpose of this study is to design and implementation an intelligent system able to get better results of Via optimization the features to discard weak or irrelevant features in the system and keep only strong relevant features using developed optimization techniques this reduces the complications of time and cost , thus increasing the accuracy of the system in identifying the modified and detect digitally modulated signals, without prior knowledge of the transmitted signal thus investing it within its applications each type according to its use, using MATLAB and JAVA programs.

The objectives this thesis are as follows:

1- Develop optimization methods to extract the main features of the received signals that could be used to distinguish the different between communication signals.

2- Develop algorithms to improve the characteristics of the signal extracted from the vertical to strength the parameters and increase the accuracy of their classification

3 - Develop algorithms to define the configuration based on the main features extracted from the signal after improve them using Random forest.

1.5. Thesis Layout

Furthermore, the other parts of the presented thesis consist :

Chapter Two: Theoretical and Background

Digital modulation signals and feature extraction present an extensive overview optimization feature extraction of signal moduled and classification two features extraction techniques are given.

Chapter Three: The Proposed modulation Recognition System

Chapter three introduces the steps of the suggested system, describes the developed algorithms to execute the system

Chapter Four: The Experimental Result and Test

This chapter presents the experiments and the results which are obtained from the system running and performance measures related to tested results

Chapter Five: Conclusions and Suggestions for Future Work

Chapter five provide conclusions list that is derived from the results related to the current work and few future works with suggestions.