Digital Modulation Classification Based On BAT Swarm Optimization and Random Forest

Batool abd alhadi sultan
College of Science, Department of computer
University of Diyala,
Diyala, Iraq
batoolabd1985@gmail.com

Taha Mohammed Hasan
College of Science, Department of computer
University of Diyala,
Diyala, Iraq
Dr.tahamh@sciences.uodiyala.edu.iq

Hadi Athab Hamed
Technical College-AL Mussaib
Al-Furat Al-Awsat Technical University
Babil, Iraq
Com had@atu.edu.iq

DOI :http://dx.doi.org/10.31642/JoKMC/2018/070104

Received Dec. 18, 2019. Accepted for publication Feb. 10, 2020

Abstract— *The applications of digitally modulated signals are still in progress and expansion. Automatic Modulation Identification (AMI) is important to classify the digitally modulated signals ..To get better results of the system suggested optimization the features to discard weak or irrelevant features in the system and keep only strong relevant features .In this work, present hybrid intelligent system for the recognition related to the digitally modulated signals where used . The proposed (AMI) had been built to classify ten most popular schemes of digitally modulated signals, namely (2ASK, , 2PSK, 4PSK, 8PSK, 8QAM,16QAM ,32QAM, 64 QAM, 128QAM, and 256QAM), with the signal to noise ratio ranging from (-2 to 13) dB. High-order cumulants (HOCs) as well as high-order moments (HOMs) were utilized. .In this thesis used , Bat Swarm Optimization (BA).The Random Forest (RF) classifier was introduced for the first time in this work. Simulation results of the System proposed , under additive white Gaussian noise channel, show that .While algorithm (BA Swarm Optimization) for the modulated signals we obtained a classification accuracy of around 92% for the SNR between (-2....12) dB.*

Keywords—Classification ,Bat Swarm Optimization Algorithm, Random Forests, Automatic Modulation Identification

I. INTRODUCTION

Automatic modulation recognition is middle step between signal detection and demodulation [1] .This step can recognize modulation type of received signal between the numbers of presupposed modulations. Automatic modulation recognition has important role in civil and military applications .Nowadays due to increasing digital modulations in civil and military applications, digital modulation recognition has particular importance. In general, for automatic modulation recognition, a few feature of received signal will be extracted and employed. Choice of appropriate features has important role in increasing efficiency of recognition [2] In this paper we have considered the following digital signals: 2ASK, 2PSK, 4PSK, 8PSK, 8 QAM,16QAM, 32QAM, 64QAM, 128QAM, 264QAM is optimizing the features of the modulated signals, Which leads to reducing the signal characteristics by increasing the accuracy of the system in detection and identification of the signal type, using of Bat Swarm optimization algorithm .In [3] Use new algorithm SVMs to identify are certain types of modulation such as: BASK, BFSK, BPSK, 4-ASK, 4-FSK, QPSK, 16-QAM .The simulation results proved to achieve the final accuracy of

98.15 at -10 dB. In [2] 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 increase in the accuracy of recognition of the type of modification .In [4] using the decision tree A dual-carrier bus system support has been trained on the features which have been extracted from high-order stacking tool. The results have indicated that the average accuracy of identifying formatting signals can be realized more than 94% when SNR is -5 dB . In [5] They proposed a data-based model to classify automatic settings without relying on expert features such as high-frequency moments. The accuracy of the results was 90% in the SNR variance ranging from 0dB to 20d. Also,The major aim of this study is optimizing the features of the modulated signals, Which leads to reducing the signal characteristics by increasing the accuracy of the system in detection and identification of the signal type, using of Bat Swarm optimization algorithm and Comparing the results obtained before and after the improvement and algorithm improve on classification accuracy by strengthening the parameters of the signals and increasing the accuracy .In this work, Random forest is as classified. Also, one of the main issues is selecting the suitable feature set. In previous paper , usually are numerous Features used to classify the composition leading to Improved efficiency. So many features are not effective enough one of the main reasons leads to restrictions Most of the techniques recognize a digital signal The features are related the general structure of this paper is as follows. After introduction, features extraction and algorithms of optimization are going to be reviewed in Section 2, the classified will be presented in Section 3. Section 4 Provide some simulation results, and lastly the conclusions of the study will be presented in section 5.

II. FEATURE EXTRACTION AND OPTIMIZATION

A. *featuer extraction*

A typical pattern recognition system after doing some preprocessing operations, often reduce the size of a raw data set by extracting some distinct attributes called features .The need for feature extraction comes to scene due to the possible inability to use the raw data .In the signal recognition area, choosing the good features, not only enable the classifier to distinguish more and higher digital signals, but also help to reduce the complexity of the Classifier . Different types of the digital signal have different properties, therefore finding the proper features in order to identify them (especially in case of higher order moment ,cumlant)is a difficult task. Based on our researches, a the statistical features provide a fine way for discrimination of the considered digital signal types[6] .

II. *1. Moments*

Probability distribution moments can be defined as a model of expected value and defining the characteristic of probability density function. With regard to digital signals, the specification for i^{th} moment for finite length is specified via:

$$\mu_i = \sum_{K=1}^{N} (s_k - \mu)^i f(s_k) \tag{2.1}$$

Where N can be defined as the data length, s_k is the random variable, subscript (k) is an integer- valued variable, μ is mean value of random variable. Let the signal has a zero mean (μ= 0), thus equation (1) becomes:

$$\mu_i = \sum_{K=1}^{N} S_k^i \; f(s_k) \tag{2.2}$$

The auto-moment regarding random variable is:

$$E_{s,p+q,p} = E[S^P (S^*)^q] \tag{2.3}$$

The (p) represents the number of the non-conjugated terms, (q) represents the number of conjugated terms, (p+q) represent moment order, and (S) is discrete random variable[8]

A. *optimization*

In order to get better results of the system optimization is suggested to the features to discard weak or irrelevant features in the system and keep only strong relevant features, to check the accuracy of the system and comparing the results obtained from applying the feature optimization algorithm and without applying it.algorithm used for feature optimization (BAT swarm) .and not as an independent document. Please do not revise any of the current designations.

III. *2. Cumulants*

Cumulants, are also statistical features. If the characteristic equation of a random variable S with zero mean is :

$$\mathrm{f}(t) = E[e^{jts}] \tag{2.4}$$

Expanding the Logarithm of equation (2.4) by applying a Taylor series, we obtain:

$$g(t)\ log\left\{E[e^{jts}] \sum_{n=1}^{\infty} k_n \frac{(jt^n)}{n!}\right\}$$

Where(k_n) is called the cumulant, (t) is time. The nth order cumulant is comparable to nth order moment, thus

$$C_{s,p+q,p} = Cum\left[\underbrace{s(t),\dots,s(t),}_{pterms} \underbrace{S^*(t),\dots,S^*(t),}_{qterms}\right] \tag{2.5}$$

The cumulants can be derived from moment :

$$cum[s_1, \dots . s_n] = \sum_{\forall v} (-1)^{q-1} (q-1)!\, E\left[\prod_{j\epsilon v_1} s_j\right] \dots E\left[\prod_{j\epsilon v_q} s_j\right] \tag{2.6}$$

The summation will be implemented on partitions $v = (v_1, \dots \dots, v_q)$ for the indexes (1, 2, ..., n), (q) represents the number of elements in partition. [1]

III. METHODOLOGY

A. *Bat Swarm Algorithm*

This algorithm is simply follow the behavior of bat, initially each bat will assign a frequency (Random Frequency) these frequencies is known between ranges [f min, f max].Then the data set is bring to this algorithm the weight of the input-hidden and hidden-output is calculated and the transfer function is calculated and the rand value choose randomly. The three general equations that summarize the behavior of Bat is

$$fi = f\,min + (f\,max - f\,min)\beta \quad (3.1)$$

$$vt = vt-1 + (xt-1 - x^*)fi \quad (3.2)$$

$$xt = xt-1 + vt \quad (3.3)$$

where β [0, 1] is random vector drawn from uniform distribution[7].

.

B. *Random forest*

IV. Random forest can be considered as ensemble classifier which contain a lot of decision trees and output the class which is mode of class's output via separate trees. The term was derived from the random decision forests which has been initially suggested via [12]. The approach combine feature's random selection and the Breiman's "bagging" approach, presented. For the purpose of constructing set of decision trees with controlled variation. The random forests combine the tree predictors in a way that every one of the trees depend on the random vector values that are sampled independently and with same distribution for every forest tree. The generalization error regarding forest of the tree classifiers depend on strength regarding distinct trees in forest and association between them. Random forests operate effectively on huge databases that can handle huge volumes of input variables with no deletion. It offers assessment regarding the significant variables in classification. Random forest in considered to be un-biased toward the assessment regarding the generalized error throughout forest formation. The algorithm also effectively calculate the missing data as well as preserving precision with approaches used to balance the errors in un-balanced class population datasets. The Resultant forests could be managed as inputs to future datasets. It provides information regarding the association between classification and variables. It operates extremely well in outlier detection, labeling of un-supervised clustering and data views Streaming Random Forest learning Algorithm .Random forest algorithm consist of the following steps*:*

V.

- Step1: Assuming that S is the number of training samples, whereas P is the number of variables in the classifier.
- Step2: assume that p is the number of input variables that are utilized for determining decision at a node of the tree where p must be considerably smaller than P.
- Step3: Choose a training set for certain tree via choosing S times with replacement from every S training sample available. Via predicting the classes, the remaining samples are utilized for estimating the error of the tree.
- Step4: in order to make a decision at one of the nodes, randomly choose p variables for every one of the tree nodes. Calculate the optimal split in the training dataset according to the p variables.
- Step5: every one of the trees will be grown to its maximum possible level so that there isn't any more pruning. [3] " Fig. 1" shows The structure of Random Forest .

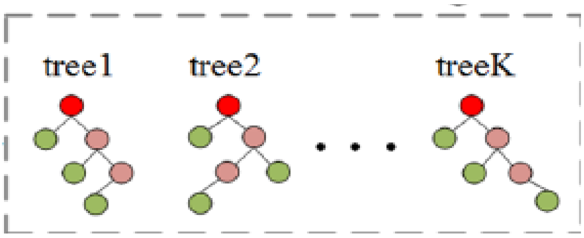


Fig1. The structure of Random Forest

IV. V. IMPLEMENTATION AND RESULT

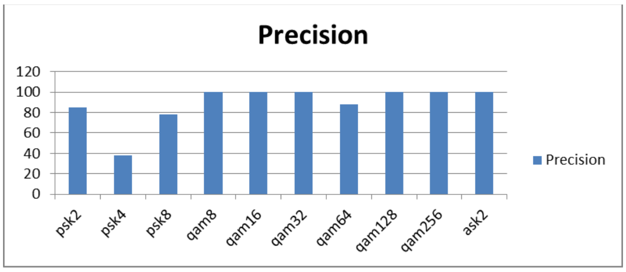
This section discussed the results when we used the algorithm Bat Swarm Optimization and Random Forests to classify 10 types of The modified signal (8 QAM, 16 QAM, 32 QAM, 64 QAM, 128 QAM, 256 QAM, 2 PSK, 4 PSK, 8 PSK, 2 ASK Within the level of SNR (-2, -1,0,1,2,3,4,5,6,7, 8,910,11 and 12) dB and comparing these results when classifying without optimization the feature of signals in an algorithm A random forest is for each type of signal and the accuracy of the rating as shown in Table 1. Each type of signal .The success rate in identifying signals after using the BA algorithm is higher as indicated in Table 3 , which means more system efficiency ."Fig. 2" shows the proposed systemAuthors and Affiliations

Results of classification accuracy using Random forest without optimization where the accuracy ratio was 90% to the proposed methodology.

Table 1. Represents results Accuracy of the proposed algorithm's without optimization

Type of signal	8 QAM	16 QAM	QAM 32	64 QAM	128 QAM
Accuracy of signal	100	100	100	88	100
Type of signal	**QAM 256**	**PSK 2**	**PSK 2**	**PSK 8**	**2ASK**
Accuracy	100	85	38	78	100

of signal					



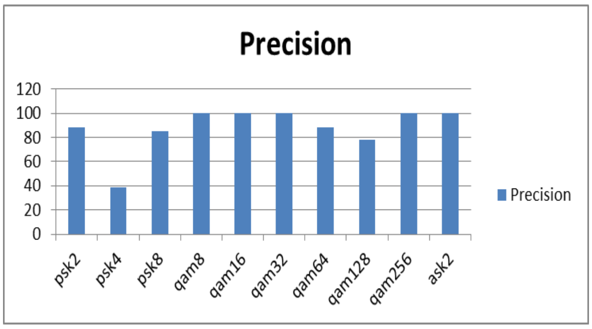
- Fig. 2. The classification Accuracy without optimization

I. Represents results Accuracy using Random Forest with BA where the accuracy ratio was 91 % to the proposed methodology.

II.

III. Table 2. Represents results Accuracy of the proposed algorithm's with optimization

Type of signal	**8 QAM**	**16 QAM**	**QAM 32**	**64 QAM**	**128 QAM**
Accuracy of signal	100	100	100	88	100
Type of signal	**QAM 256**	**PSK 2**	**PSK 2**	**PSK 8**	**2ASK**
Accuracy of signal	100	85	38	78	100



- Fig 3. The classification Accuracy with optimization

V. VI. CONCLUSION

In this paper, Ten types of electromagnetic signals embedded in the MATLAB program were created within an SNR level ranging from (-2,-1,0,1,2,3,4,5,6,7,8,9,10,11,12) dB . Then extract the statistical characteristics (moment, cumulant) of the signals, above. improve the features, by Bat swarm optimization algorithm . Using the Random Forest as a classifier When optimizing features, Bat Swarm's algorithm resulted in the highest rating accuracy even at a low SNR level of about 92%.

Future Work More work is needed to separate more sets of digital and analog signals with the use of other features and other optimization Technic .

REFERENCES

1] M. Azarbad, S. Hakimi, and A. Ebrahimzadeh, "Automatic recognition of digital communication signal," *Int. J. energy, Inf. Commun.*, vol. 3, no. 4, pp. 21–34, 2012.

[2] W. Su, J. L. Xu, and M. Zhou, "Real-time modulation classification based on maximum likelihood," *IEEE Commun. Lett.*, vol. 12, no. 11, pp. 801–803, 2008.

[3] P. K. HL and L. Shrinivasan, "Automatic digital modulation recognition using minimum feature extraction," in *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2015, pp. 772–775.

[4] S. Almaspour and M. R. Moniri, "Automatic modulation recognition and classification for digital modulated signals based on ANN algorithms," *nine*, vol. 3, no. 12, 2016.

[5] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Trans. Cogn. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, 2018.

[6] M. Bellanger and M. Bellanger, "Adaptive digital filters," 2001.

[7] J. Bagga and N. Tripathi, "Automatic modulation classification using statistical features in fading environment," *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, vol. 2, no. 8, pp. 3701–3709, 2013.

[8] S. B. Sadkhan-Smieee, A. Q. Hameed, and H. A. Hamed, "Digitally

modulated signals recognition based on adaptive neural-fuzzy inference system (ANFIS)," *Int. J. Adv. Comput. Technol.*, vol. 7, no. 5, p. 57, 2015.

[9] C. Qu, S. Zhao, Y. Fu, and W. He, "Chicken swarm optimization based on elite opposition-based learning," *Math. Probl. Eng.*, vol. 2017, 2017.

[10] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R news*, vol. 2, no. 3, pp. 18–22, 2002.

[11] A. Liaw and M. Wiener, "Classification and regression by randomForest. R News. 2002; 2: 18–22." 2016.

[12] S. Revathi and A. Malathi, "Detecting Denial of Service Attack Using Principal Component Analysis with Random Forest Classifier," *Int. J. Comput. Sci. Eng. Technol.*, vol. 5, no. 3.