



A Blind Modulation Classification Method Based on Decision Tree and High Order Cumulants

Yulai He¹, Hua Wu^{1,3}(✉), Qinghe Zheng²(✉), Yang Liu^{1,3}, Abdussalam Elhanashi⁴, and Sergio Saponara⁴

¹ School of Information Science and Electrical Engineering (School of Artificial Intelligence), ShanDong JiaoTong University, Jinan 250357, China

wuhua@sdjtu.edu.cn

² School of Intelligence Engineering, Shandong Management University, Jinan 250357, China
15005414319@163.com

³ CAAC Key Laboratory of General Aviation Operation, Civil Aviation Management Institute of China, Beijing, China

⁴ Department of Information Engineering, University of Pisa, 56122 Pisa, Italy

Abstract. The classification of modulation schemes will be widely used in future communication systems, among which the classification accuracy and speed of modulation classification methods have always been the focus of the research. Our goal is to simplify the complexity of the classification model and improve the modulation recognition accuracy. In this paper, we propose the blind modulation classification method based on the decision tree and high order cumulants. During the experiments, the results are visualized to observe the effect. By comparing the probability of accurate classification at different signal-to-noise ratios at each stage, it is used as an overall measure of the performance of the entire classifier. Moreover, experiments show that the performance of the independent block level and the whole system has excellent accuracy and stability under the optimal threshold, signal-to-noise ratio, and the number of samples.

Keywords: Wireless communication · cognitive radio · modulation classification · feature learning

1 Introduction

Nowadays, wireless communication plays an important role in the communication systems in both civil and military fields. This requires us to make full use of channel resources and improve the effectiveness and accuracy of wireless communication [7]. In order to meet the different needs of various customers, the communication signals often use the method of matching modulation. In the development of wireless communication, the modulation scheme has also experienced many changes: from simple analog to complex digital. At the same time, the environment of wireless communication is becoming more and more complicated, and a variety of modulation schemes are included in a certain communication frequency band [8]. To effectively monitor and identify wireless

signals, the research on modulation classification has attracted numerous concerns by the academic community to make a breakthrough.

After repeated replication and modulation of traditional analog signals, long-distance transmission is more vulnerable to the interference of channel random noise [1]. In order to obtain more stable communication effects, digital communication application scenarios based on digital signals are gradually increasing [2]. In the digital communication process, the choice of modulation schemes will seriously affect the effect of information transmission and communication security, so it is important to choose scientific and secure signal modulation schemes. When the transmitter can freely set the modulation scheme of the signal, the receiver must get the information to demodulate the signal in order to transmit successfully.

Automatic modulation classification (AMC) methods are mainly divided into two categories: likelihood-based methods [3, 4] and feature-based classification methods. The former is based on the likelihood function of received signals and compares the likelihood ratio with the threshold value, which can be easy to implement; The latter typically uses multiple features and makes decisions based on the observations. Specifically, Zhang et al. [5] proposed a modulation classification method based on high-order cyclic cumulative feature which can realize the modulation classification of multi-channel signals.

This paper is organized as follows. In Sect. 2, we first introduce five kinds of modulation schemes and three sub blocks of the classification system. The proposed method focuses primarily on multi-carrier (MC), and single-carrier (SC) classification. Five modulation schemes include MC in orthogonal frequency division multiplexing (OFDM) signal and 4-quadrature amplitude modulation (4-QAM), 16-quadrature amplitude modulation (16-QAM), and 2-pulse amplitude modulation (2-PAM), 4-pulse amplitude modulation (4-PAM). In Sect. 3, extensive experimental results are intuitively reflected through confusion matrices. It can be found that performance is acceptable when the threshold was set to -0.2 and SNR equals to 20 dB.

2 Modulation Classification Method Based on Decision Tree and High Order Cumulants

Through a series of modulation techniques, the higher communication rates and more efficient spectrum utilization can be obtained, thus spectrum shifting can be carried out, and the modulated symbols can be converted into a modulated signal suitable for channel

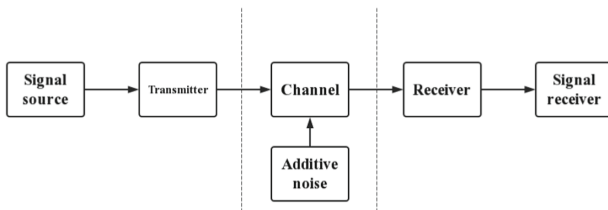


Fig. 1. Wireless communication system model.

transmission or channel multiplexing, which can improve the transmission effectiveness and transmission reliability of the system. The wireless communication system can be demonstrated by the following communication model in Fig. 1.

2.1 System Model Structure

The proposed system model is based on the decision tree, and the complete system can effectively identify the modulation scheme. As shown in Fig. 2, the modulation classification system is mainly divided into 3 blocks: (1) MC-SC block classifies between multi-carrier and single carrier signals; (2) MTC block distinguishes between M-QAM and M-PAM signals if the input to the block is a SC signal; (3) MLC block behaves as a modulation level classifier for classifying 4-QAM/16-QAM or 2-PAM/ 4-PAM.

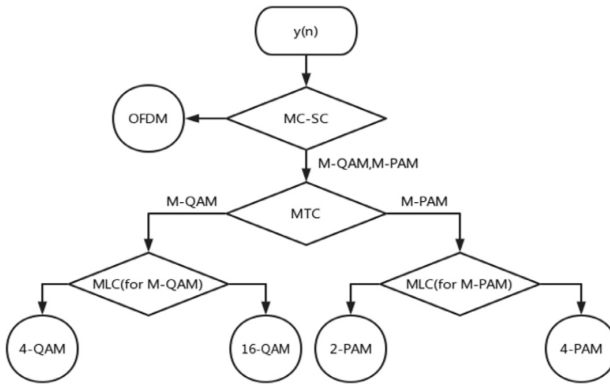


Fig. 2. Modulation classification system model.

2.2 MC-SC Block

The MC-SC block identifies MC and SC by judging the fourth-order cumulants of different signals by performing a Gaussian test. The cumulant C_{42} can be computed as follows:

$$C_{42} = \frac{1}{N_m} \sum_{n=1}^{N_m} |y[n]|^4 - |C_{20}|^2 - 2C_{21}^2 \tag{1}$$

where N_m is the number of samples and $y[.]$ is the received signal. Since OFDM is a mixture of many subcarrier waveforms, the input samples are close to a Gaussian distribution, its statistic C_{42} tends to be 0. For SC signals, C_{42} approaches to a non-zero value so that MC and SC can be classified. The C_{42} values for the modulation schemes with a threshold of 0.5 are shown in Table 1.

Table 1. The C42 values for the different modulation schemes.

Modulation scheme	C42
OFDM	0
4-QAM	-0.82
16-QAM	-0.52
2-PAM	-1.69
4-PAM	-1.1

2.3 MTC Block

For received signals, MTC block uses cyclic-auto-correlation (CAC) function to determine their cyclostationary features. The conjugate and non-conjugate CACs can be computed as

$$\begin{aligned}
 R_{x^*}(\alpha, \nu) &= \frac{1}{N} \sum_{n=0}^{N-1} x[n]x^*[n-\nu]e^{-j2\pi\alpha nT_s} \\
 R_x(\alpha, \nu) &= \frac{1}{N} \sum_{n=0}^{N-1} x[n]x[n-\nu]e^{-j2\pi\alpha nT_s}
 \end{aligned} \tag{2}$$

where α is the cyclic frequency, ν is the lag variable, N is a finite number of samples, T_s is the sampling period. Based on the CAC has peaks at different locations of cyclic frequencies α which is a function of symbol rate ($1/T$) and carrier frequency (fc), modulation schemes can be distinguished. The peaks of the CAC of M-QAM in terms of T and fc is $(1/T, 0)^*$, M-PAM are $(1/T, 0)^*$, $(2fc, 0)$ and $(2fc/T, 0)$. Combining these theoretical values, a feature vector of length 6 can be obtained as

$$F = \begin{bmatrix} |R_{x^*}(1/T, 0)|, |R_{x^*}(2fc - 1/T, 0)|, |R_{x^*}(2fc - 1/2T, 0)|, \\ |R_{x^*}(2fc, 0)|, |R_{x^*}(2fc + 1/2T, 0)|, |R_{x^*}(2fc + 1/T, 0)| \end{bmatrix} \tag{3}$$

where the element is proportional to the signal power and the feature vector has been normalized to unit power. The normalized asymptotic feature vector for M-QAM is and M-PAM are $V_1 = [1, 0, 0, 0, 0, 0]$ and $V_2 = [0.5, 0.5, 0, 0.5, 0, 0.5]$, respectively. By comparing the signal with feature vector, MTC can recognize M-QAM or M-PAM.

2.4 MLC Block

MLC block calculates the communication data field (CDF) of the received signals and compares them with theoretical CDFs of available modulation schemes and then chooses the one with the smallest distance from them. The CDF is calculated as

$$F_N(t) = \frac{1}{N} \sum_{n=1}^N I(z(n) \leq t) \tag{4}$$

where N is the number of samples, t is a threshold range, $z(n)$ is the received signal, and $I(\cdot)$ is 1 if the condition true, 0 otherwise. By calculating CDFs by replacing $z(n)$ with the corresponding input signal $x(n)$, MLC block further breaks down the signal into its modulation level: whether 4-QAM, 16-QAM or 2-PAM, 4-PAM.

3 Experimental Results and Analysis

3.1 Experiment Process

The performance of MC-SC block can be measured by evaluating the classification probability. By changing SNR, the classification probability of MLC for M-QAM and M-PAM is tested respectively.

3.2 Signal Waveform Plot

The initial signal provided was set before the experiments begin. Experiments are performed by setting up a transmitter that randomly transmits one signal in the MC and SC. The MC signal is an OFDM signal with 64 sub-carriers, 16 samples, and a sub-carrier spacing of 15.625 kHz. The SC signal applied by raised cosine filter is randomly composed of one of 4-QAM, 16-QAM, 2-PAM, and 4-PAM. The received signal is given as

$$y(n) = x(n)e^{j2\pi fc n Ts} + w(n) \tag{5}$$

where $x(n)$ is 1 MHz bandwidth transmitted signal, carrier frequency $fc = 2$ MHz, sampling duration $Ts = 0.25 \mu s$, the Gaussian noise $w(n) \sim CN(0, \sigma^2)$. The number of MC symbols and SC symbols are 25 and 2000 respectively. The frequency spectrum can then be generated by fast Fourier transforming (FFT) the signal. In Fig. 3, we show the amplitude spectrum and frequency spectrum of various modulation schemes.

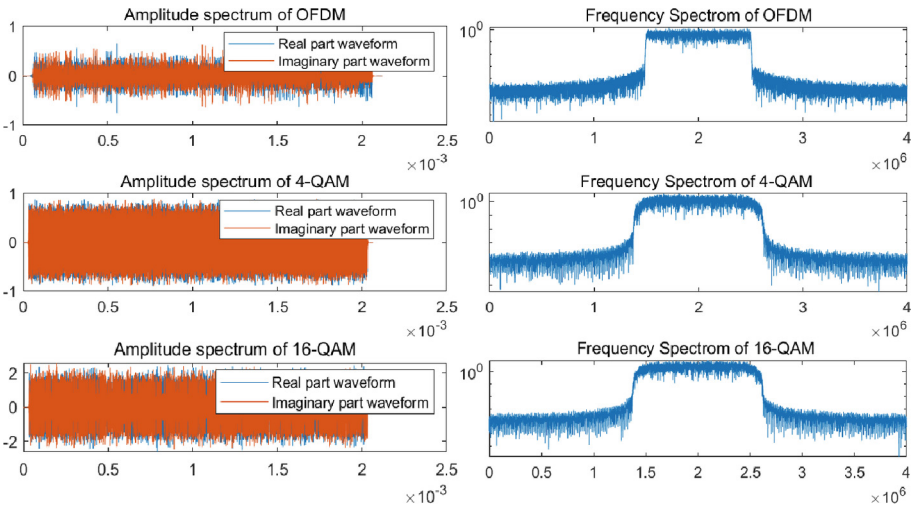


Fig. 3. Amplitude spectrum and frequency spectrum of various modulation schemes.

3.3 Experiment Results and Analysis

Probability of detection (P_d) is the probability that the received signal classified as MC. Probability of false alarm is the probability that the received signal classified as MC though it is not MC. As shown in Fig. 4, N_m determines the performance since a higher value of N_m can obviously improve the performance of MC-SC block. At the same time, SNR can make an impact on performance as well. We can obviously discover that SNR reaches the best state at 20 dB, which is better than 10 dB or 0 dB.

In Fig. 5, we depict the performance of MTC blocks in terms of classification probability. Here, the performance also increases as the number of samples to perform the calculation increases.

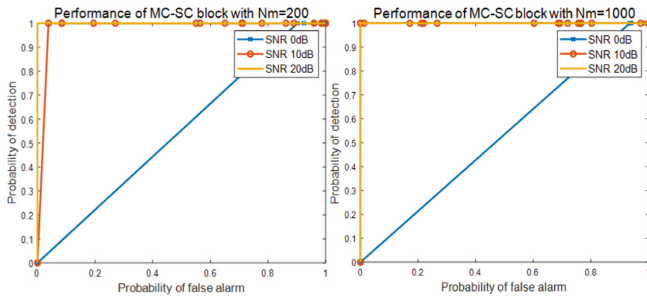


Fig. 4. Performance of MC-SC block ($N_m = 200$ and $N_m = 1000$).

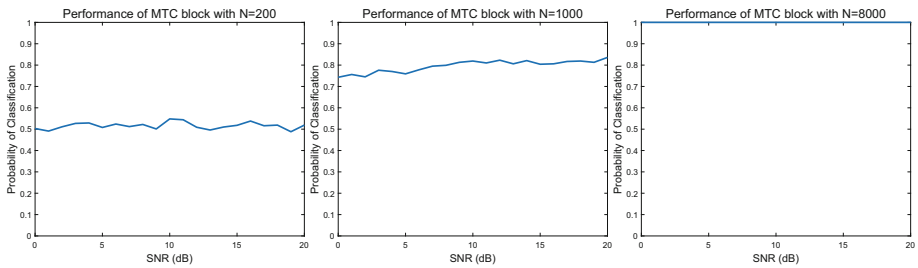


Fig. 5. Performance of MTC block ($N = 200$, $N = 1000$, and $N = 8000$).

Figure 6 shows the performance of MLC block. It can be seen that PAM classifier performs better than QAM classifier, which is due to that transmitted QAM signals' empirical CDF lies between the theoretical CDF of both level's signal. As shown in Fig. 7, the model achieves best performance (almost 100%) when the threshold γ was -0.2 and SNR = 20 dB.

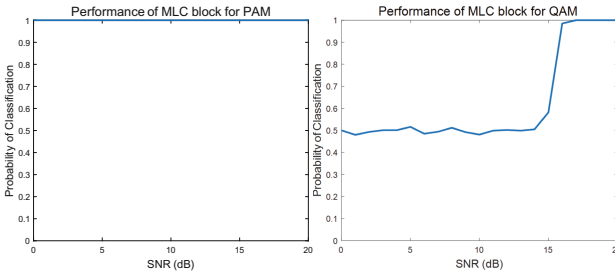


Fig. 6. Performance of MLC block for PAM and MLC block for QAM.

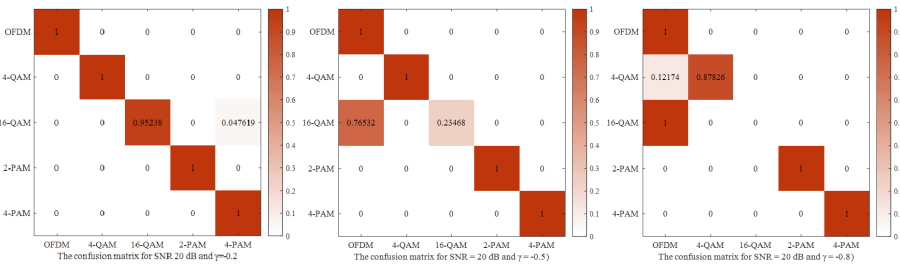


Fig. 7. Confusion matrices of different hyper-parameters setting.

4 Conclusion

A blind modulation classification method based on decision tree and high order cumulants is presented in the paper. The techniques for classifying various types of classification schemes at every stage viz. MC-SC, MTC, and MLC are implemented. The MC-SC block is accomplished by high-order cumulant, then the MTC block is accomplished by feature vector and the MLC block is accomplished by CDFs. The simulation results show that this method of direct classification can obtain a high classification accuracy under the condition of high SNR. In the future, our goal is to improve the classification accuracy under low SNR conditions by pre-estimating the necessary parameters such as carrier frequency, code rate, and modulation index.

Acknowledgements. This research was supported by Shandong Provincial Natural Science Foundation (Grant ZR2020MF151), CAAC Key Laboratory of General Aviation Operation (Civil Aviation Management Institute of China) (Grant CAMICKFJJ-2020-2), National Natural Science Foundation of China (Grant U1933130 and 71731001), and Research and Demonstration of Key Technologies for the Air-Ground Collaborative and Smart Operation of General Aviation (Grant 2022C01055).

References

1. Kim, D.H., Choi, J.H.: Analysis of the transmission performance of control pilot signal lines for charging communication in electric vehicles. *Int. J. Automot. Technol.* **21**(2), 519–525 (2020)

2. Al-Salehi, A.R., Qureshi, I.M., Malik, A.N., et al.: Dual-function radar-communications: information transmission during FDA radar listening mode. *Int. J. Microwave Wirel. Technol.* **12**(1), 1–12 (2020)
3. Dobre, O., Zarzoso, J., Bar-Ness, Y., Su, W.: On the classification of linearly modulated signals in fading channel. In: *Proceedings CISS Conference*, Princeton, March 2004
4. Zheng, Q., Zhao, P., Wang, H., Elhanashi, A., Saponara, S.: Fine-grained modulation classification using multi-scale radio transformer with dual-channel representation. *IEEE Commun. Lett.* **26**(6), 1298–1302 (2022)
5. Zhang, C.Y., Yue, H.T., Wang, B.W., et al.: Deep learning pattern recognition method of partial discharge ultrasonic signal based on similar matrix blind source separation and convolutional neural network. *Power Syst. Technol.* **43**(6), 7 (2019)
6. Zheng, Q., Zhao, P., Zhang, D., Wang, H.: MR-DCAE: Manifold regularization-based deep convolutional autoencoder for unauthorized broadcasting identification. *Int. J. Intell. Syst.* **36**(12), 7204–7238 (2021)
7. Zheng, Q., Yang, M., Wang, D., et al.: An intelligent wireless communication model based on multi-feature fusion and quantile regression neural network. *J. Intell. Fuzzy Syst.* (2021). Early Access
8. Zheng, Q., Zhao, P., Li, Y., Wang, H., Yang, Y.: Spectrum interference-based two-level data augmentation method in deep learning for automatic modulation classification. *Neural Comput. Appl.* **33**(13), 7723–7745 (2020). <https://doi.org/10.1007/s00521-020-05514-1>