Communication Modulation Recognition Method Based on Clustering Algorithm

Yuyang Cheng *, Shuying Shao
North China University of Technology, Beijing, Beijing, 100144, China
*Corresponding author’s e-mail: 18152010616@mail.ncut.edu.cn

Abstract: The modular identification of wireless communication signals is widely used in civil and military fields. With the rapid development of modern communication technology in China, signal modulation becomes complex and diverse. How to accurately identify communication signal modulation methods has become a research hotspot. Among them, multi band quadrature amplitude modulation (MQAM) is widely used in microwave communication, satellite communication and network communication. It has the advantages of high frequency band utilization and flexible modulation. Therefore, the recognition of MQAM modulation mode is of great significance. Traditional clustering algorithm has some limitations when it is used to identify MQAM signal modulation, such as too many iterations and long recognition time. For high-order modulated signals, a modulation recognition method based on Constellation Clustering is proposed. The constellation is used as the feature of modulation recognition, and the constellation of received signal is reconstructed in groups. The simulation results show that the collected characteristic parameters have strong interference ability to multi-channel, and the recognition rate reaches 97.6%. It has important engineering application value for signal modulation recognition in low SNR environment.

1. Introduction
With the rapid development of network information technology, wireless communication has become the mainstream of modern communication system. In order to meet the requirements of communication quality, it is necessary to select an effective modulation mode, so as to increase the appropriate signal modulation mode. Generally speaking, communication signal modulation method has a long history of development. In the early stage of the development of communication technology, there is only a simple analog modulation mode, which is simple, but its anti-interference ability is weak. With the rapid development of MIMO and OFDM technologies, the modulation method is not limited to single carrier, and multi carrier transmission modulation technology has become the mainstream method of modern system. As a result, it is more and more difficult to communicate with different kinds of signals.

In communication demodulation and modulation technology, constellation can be used for digital modulation analysis. It clearly shows the structure of the signal and the relationship between different modulation methods. In computer technology, cluster analysis is the basic data mining technology and intelligent recognition technology, while the modulation recognition of constellation cluster analysis combines the characteristics of communication algorithm and computer. Compared with other modulation recognition methods, it has many advantages, such as low complexity, easy implementation and strong scalability. Therefore, the research and design of MQAM automatic
recognition technology based on Constellation Clustering analysis has important significance and practical value in the field of digital communication [4-6].

In this paper, based on the idea of constellation reconstruction, eafcm grouping algorithm is proposed, which combines shape recognition and modulation recognition, and obtains some characteristics of modulation. At present, many scientists use c-means grouping algorithm (FCM) to reconstruct constellation. FCM algorithm can automatically classify data samples, but it is sensitive to the initial center, so it is necessary to predict the parameters of group C and convergence threshold. Based on FCM algorithm, this paper proposes an adaptive and efficient grouping algorithm (eafcm). This method uses subtractive clustering algorithm to determine the initial center of clustering, and then makes classification research based on the number of cluster centers. This algorithm reduces the randomness of the initial center and saves calculation time. For constellations with low signal-to-noise ratio or with certain time deviation, it can still be reconstructed and identified [7-10].

2. Communication Mode Recognition and Clustering Algorithm

2.1 Clustering Algorithm

Clustering algorithm, also known as group analysis, is a statistical analysis method to study (sample or index) classification problem, and is also an important mining algorithm. Clustering analysis is based on similarity. The similarity between models is higher in a group than in a group. The function of cluster analysis is to divide a data group into groups with the same attributes. Therefore, when the receiver intercepts the unknown signal, it only needs to use the specified database to extract the vector, and then use the appropriate model method to calculate the characteristic distance between the unknown signal vector and the signal in the database. Assuming that the recognition system has a functional database of $n$ different types of signals, the distance between the $j$-th function AAA of signal information and the $j$-th information feature $A_{ij}$ of class $i$ ($i = 1, 2, \ldots, n$) signal is

$$d_{ij} = |E_{ij} - A_{ij}|$$  \hspace{1cm} (1)

The information characteristic distance between the signal to be identified and the class $i$ communication signal is

$$D_i = \sqrt{d_{i1}^2 + d_{i2}^2 + \ldots + d_{in}^2}$$  \hspace{1cm} (2)

The formula for calculating the closeness between the signal to be identified and the sample of type $i$ is

$$N_i = 1 - \frac{D_i}{\sum_{j=1}^{n} D_j}$$  \hspace{1cm} (3)

2.2 Communication Modulation Recognition Mode

MPSK / MQAM module is a common modulation mode in communication. In recent years, some representative modulation recognition algorithms for these two modulation modes are proposed. At present, there are two kinds of algorithms for channel identification: one is based on the application range of the channel. The other is blind equalization, which eliminates the channel effect and uses AWGN channel recognition algorithm to identify. This method is suitable for many kinds of channels, and it is a promising modulation recognition method. The key of the second algorithm is blind equalization, which uses a single CMA equalizer to equalize MPSK and MQAM signals. Because CMA algorithm is based on the statistical characteristics of the signal, it has a weak effect on MQAM signal equalization. When CMA and AMA dual-mode blind equalization algorithm is used, each equalizer corresponds to only one modulation mode. The results show that the equalization effect and convergence speed of this algorithm are obviously better than the two algorithms in the previous paper. But its structure is complex, the selection of parameters in the algorithm is greatly affected by the
channel, and the amount of data required is also large. These two blind equalization methods are essentially cost minimization, and their convergence speed is slow and modulation mode is dependent on the input signal and converges to the local minimum.

3. Experimental Ideas and Design
The significance of modern signal communication is to transmit information safely, accurately, quickly and effectively to the destination through the communication channel. With the complexity and diversification of communication system, debugging mode and space transmission electromagnetic wave types, the channel environment density of transmission signal is higher and higher, so the frequency range set by the receiver is difficult to narrow, so that the received signal will not be invalid or distorted in the complex environment transmission process. Usually, in the long-distance and a lot of noise communication environment to achieve transmission without distortion, the transmitter communication signal needs to adopt different modulation methods, also can make the channel transmission capacity to achieve the best. Naturally, the research on the signal processing for the diversity of modulation methods must include the research on modulation recognition of modulation signals. In terms of military affairs, modern warfare is more complicated, and the well-known electronic warfare is one of them. In the military reconnaissance mission, sometimes it is necessary to use a reconnaissance plane with a specific frequency band to search within its set frequency range. By recognizing the enemy's communication signals, the modulated signals of the enemy's communication radio stations are monitored. The identified signal parameters are the so-called communication electronic warfare signal information. It is obvious that the most important measure to intercept and intercept the enemy signals is to intercept and intercept the signals of the enemy in many ways. In civil applications, signal modulation recognition is widely used in signal confirmation, interference identification and monitoring spectrum. Its main purpose is to prevent the illegal transmission of information when the radio station exceeds the specified limit of working parameters. In order to maintain the normal operation of different equipment, to improve the monitoring ability of different equipment. The basic task of this project is to be able to correctly and effectively identify the modulation mode of signals in the environment with a variety of signals and Gaussian noise interference, so as to lay a solid foundation for the application of signal modulation recognition. In this paper, 100 experiments on 4-256 order MQAM signals are carried out under the condition of Gaussian white noise and bit error rate of 15%. Among them, 2000 points of sample data are used for 4QAM, 8qam, 16QAM and 32qam, 4000 points of sample data are used for 64QAM and 128QAM signals, and 8000 points of sample data are used for 256qam.

4. Discussion
4.1. Modulation Recognition Algorithm Based on Signal Clustering Analysis
In this paper, the traditional clustering algorithm and eafcsm clustering algorithm are simulated for 100 times, and the operation points, cluster centers and running time of the two algorithms are counted. According to the experimental results, we investigated and analyzed the results, and the results are shown in Table 1. It can be seen from table 1 that traditional clustering algorithm updates 4QAM, 16QAM,..., 256qam with 2000, 4000 and 8000 sample points respectively, while eafcsm algorithm takes part in the operation of 1/3 of the traditional algorithm. In terms of the number of cluster centers, the improved algorithm does not reduce the accuracy of cluster centers without reducing the number of samples involved in the operation. From the operation time of the algorithm, the time of traditional clustering algorithm is approximately three times of that of the improved algorithm.

| Order | Sample points | Number of cluster | Time / S |
|-------|---------------|-------------------|----------|

Table 1. Comparison of the number of clustering centers and time before and after the improvement of the algorithm
| MQAM | tradition | improvement | Tradition and improvement | tradition | improvement |
|------|-----------|-------------|---------------------------|-----------|-------------|
| 4    | 2000      | 660         | 4                         | 2.06      | 0.99        |
| 8    | 2000      | 710         | 8                         | 3.65      | 1.56        |
| 16   | 2000      | 630         | 16                        | 7.71      | 2.60        |
| 32   | 2000      | 682         | 33                        | 15.23     | 5.53        |
| 64   | 4000      | 1365        | 66                        | 77.85     | 27.06       |
| 128  | 4000      | 1376        | 128                       | 211.96    | 69.98       |
| 256  | 8000      | 2760        | 260                       | 1085.42   | 393.19      |

In addition, the approximate model of the improved algorithm and the traditional algorithm is given, as shown in Figure 1. We can clearly see that the difference between each signal in the improved paste schedule model is very large. The similarity between the unknown signal and the known signal is close to 1, and the proximity to other types of samples is less than 0.8. The improved algorithm can judge the type of signal more clearly. In order to compare the advantages and disadvantages of the two methods, we compare the paste progress of the unknown signal and the known signal between the improved algorithm and the traditional algorithm, as shown in Figure 1. The improved model can judge the type of unknown signal more clearly and calculate the proximity between the signal to be identified and the known signal, so as to achieve the purpose of signal recognition.

![Figure 1. Comparison model of closeness between improved algorithm and traditional algorithm](image)

In this paper, the eafcm clustering algorithm is further simulated and analyzed, and the recognition rate comparison chart of QAM signal with different symbol length is obtained, as shown in Figure 2. When the fixed signal-to-noise ratio is 10dB, figure 2 shows the simulation results of the recognition accuracy rate with the change of the symbol length. It can be seen from the diagram that the recognition accuracy rate of the algorithm increases with the increase of the symbol length. When the
symbol length reaches 200, the performance of 16QAM, 32qam, 64QAM classification algorithm has reached the optimal performance; when the symbol length reaches 300, the recognition rate of 16QAM, 32qam, 64QAM classification algorithm reaches 97.6%. In practical application, the appropriate code length can be selected according to the actual situation.

![Figure 2. Recognition rate of QAM signal with different symbol length](image)

4.2. Suggestions on Communication Modulation Recognition Method Based on Clustering Algorithm

4.2.1. Modulation recognition process based on constellation

The key point of MQAM modulation recognition based on constellation is constellation recovery, which affects the final recognition rate of modulation signal. Before modulation recognition, it is necessary to know the modulation parameters such as carrier frequency and symbol rate to recover the signal constellation. However, in actual communication, the specific modulation parameters of the received signal are unknown to the receiver. Therefore, it is necessary to design an algorithm to estimate the carrier frequency, symbol rate and timing error in order to realize the carrier synchronization and symbol synchronization of the signal and recover the signal constellation Figure, and then identify. Firstly, the bandwidth and carrier frequency are estimated from the sampled signal, and then the down conversion and low-pass filtering are performed to eliminate the influence of the carrier and reduce the noise level. For the baseband sequence after down conversion, the symbol rate estimation algorithm can be used to estimate the signal symbol rate. After estimating the symbol rate, the best sampling time is selected for symbol rate sampling according to the code rate sampling principle and joint timing estimation, so as to recover the constellation of the signal. Finally, the subtractive clustering algorithm is used to cluster the constellation, and the modulation order can be obtained.

4.2.2. Signal parameter estimation

The difference sequence reflecting the signal symbol transformation can be obtained by the difference processing of single channel of received sequence. The existence of noise will change the signal waveform, and the sampling value inside the symbol is no longer a constant. When the SNR is low, the symbol transformation time point displayed in the corresponding difference sequence waveform is no longer clear. Some peak points caused by noise may affect the judgment of symbol transformation time point. Based on the above rules, most of the peaks caused by noise can be eliminated at the
expense of some normal peaks. Because only the digital modulation signal has the symbol, the analog signal does not have this definition. If the symbol of digital modulation signal is of equal length, that is, the number of sampling points is equal, then the position interval of single value at the time of symbol transformation must be equal. However, after the single value screening process described above, some correct single values will be filtered out, so the interval between adjacent single values may not be one symbol, but multiple symbols. Therefore, there must be a greatest common divisor greater than 1 in the sequence composed of intervals of all adjacent single values. Based on this conclusion, we extract feature parameters according to the above conditions.

4.2.3. Signal recognition based on signal features

The signal modulation information is reflected in the changes of amplitude, frequency and phase. Theoretically, all modulation modes can be identified by using the modified characteristics of these three variables. The search algorithm based on real-time signal features has been studied before. The advantages of this algorithm are simple and intuitive. It does not need a large amount of storage capacity in engineering application, and the calculation speed is fast. However, the recognition performance is poor under low SNR. Based on the analysis of 20 signals to be identified in this paper, we can deduce the most suitable recognition structure and recognition sequence, extract the corresponding feature parameters according to the respective characteristics of the signals, and determine the best decision threshold through a large amount of data testing, then the modulation recognition can be realized by the tree classification structure.

5. Conclusions

In this paper, a new method of signal modulation recognition is proposed. The algorithm uses eafcm clustering to reconstruct constellation. The algorithm has the advantages of simple structure, high efficiency and self-adaptive, and overcomes the shortcomings of c-means method. When the signal-to-noise ratio is low or the symbol time has a certain deviation, it can still achieve good recognition effect, which strongly proves the feasibility of the algorithm. The algorithm proposed in this paper is suitable for amplitude phase modulation signals which need carrier synchronization and bit synchronization, and its application range is limited. In this paper, the clustering algorithm itself is analyzed, and the collection and recognition efficiency of high-dimensional data needs to be further discussed. It is necessary to extract the features of high-dimensional data and achieve better clustering effect. The simulation results show that, even in the environment of 5dB SNR, the signals with different characteristics can still be recognized accurately. It provides an important theoretical basis for the wide application of clustering algorithm in signal recognition, video tracking, good working position and other fields.

References

[1] Li Q , Shen D , Wang F , et al. MQAM Modulation Recognition Based on AP Clustering Method[J]. Matec Web of Conferences, 2016, 44:01002.
[2] Li Y , Ni Z , Jin F , et al. Research on Clustering Method of Improved Glowworm Algorithm Based on Good-Point Set[J]. Mathematical Problems in Engineering,2018,(2018-3-5), 2018, 2018(PT.3):1-8.
[3] Liu T , Guan Y , Lin Y . Research on modulation recognition with ensemble learning[J]. Eurasip Journal on Wireless Communications & Networking, 2017, 2017(1):179.
[4] Storti G L , Paschero M , Rizzi A , et al. Comparison between time-constrained and time-unconstrained optimization for power losses minimization in Smart Grids using genetic algorithms[J]. Neurocomputing, 2015, 170(dec.25):353-367.
[5] Ulf, Petrausch. [Immunotherapy - immune-related adverse events and their management].[J]. Therapeutische Umschau. Revue therapeutique, 2019, 76(4):195-198.
[6] Fang F , Li X , Xing Y , et al. Research on Clustering Method of Airborne Threat Target Based on Density[J]. Journal of Physics: Conference Series, 2019, 1345(3):032026 (6pp).
[7] Wen-Bin Z, Yong M, Wei-Wei C, et al. Research on Shopping Basket Compression Based on Clustering Algorithm[J]. Computer Technology and Development, 2018.

[8] Zhang G, Zhao K, Yi Li. Research and Improvement Method Based on k-mean Clustering Algorithm[J]. Computer and Information Ence, 2019, 12(1):49-52.

[9] Yang F, Li Z, Luo Z. A New Specific Combination Method of Wireless Communication Modulation Recognition Based on Clustering and Neural Network[J]. Zhongshan Daxue Xuebao/Acta Scientiarum Naturalium Universitatis Sunyatseni, 2015, 54(2):24-29.

[10] Xi-Quan W U, Yong G. Novel Modulation Recognition Algorithm based on Genetic BP Neural Network[J]. Communications Technology, 2019.