Automatic recognition of the digital modulation types using the artificial neural networks

Saad S. Hreshee
Electrical Engineering Department, College of Engineering, University of Babylon, Iraq

ABSTRACT
As digital communication technologies continue to grow and evolve, applications for this steady development are also growing. This growth has generated a growing need to look for automated methods for recognizing and classifying the digital modulation type used in the communication system, which has an important effect on many civil and military applications. This paper suggests a recognizing system capable of classifying multiple and different types of digital modulation methods (64QAM, 2PSK, 4PSK, 8PSK, 4ASK, 2FSK, 4FSK, 8FSK). This paper focuses on trying to recognize the type of digital modulation using the artificial neural network (ANN) with its complex algorithm to boost the performance and increase the noise immunity of the system. This system succeeded in recognizing all the digital modulation types under the current study without any prior information. The proposed system used 8 signal features that were used to classify these 8 modulation methods. The system succeeded in achieving a recognition ratio of at least 68% for experimental signals on a signal to noise ratio (SNR=5dB) and 89.1% for experimental signals at (SNR=10dB) and 91% for experimental signals at (SNR=15dB) for a channel with additive white gaussian noise (AWGN).

Keywords:
Artificial neural networks
Digital modulation
Smart receivers
Software defined radios

1. INTRODUCTION
The automatic recognition techniques (ART) of the modulation type of the received signal are an intermediate process between the initial processing of the received signal and the process of selecting and detecting the appropriate demodulator. These systems determine the modulation type of the signal used in the transmitting side, after extracting the features of that signal that help with the recognition process, which will enable us later to know the information in this signal by inputting it on the appropriate signal demodulator as shown in Figure 1 [1].

ART of the modulation type plays an important role in many civil and military applications [2]. ART can be used in the military field, in the electronic scanning and surveillance processes, where the recognition of the modulation type is very important in the national security by detecting a potential threat, which can then be concealed or jamming [2]. In the civil field, ART can be used in spectrum management processes, wireless network management, and frequency interference recognition. ART of the modulation type also plays a key role in the Software Defined Radios SDR and the intelligent receiver [1, 2].

Previously, the modulation type recognition processes were carried out with wide-beam receivers designed to be capable of scanning the entire spectrum. The modulation recognition process is made depending on the time format of the signal that appears on the output unit, based on this time format, the operator guessed
the type of modulation [3]. Obviously, this method of recognition requires a lot of effort versus poor performance and a high probability of error because it relies heavily on operator skills and experience [1]. The use of transmission recognition methods leads to the use and development of new ART like the input of the received modulated signal to all possible signal detectors and then recognition of the modulation type by comparing the output values of all detectors. The difficulty and poor performance of this method are also clear and are no longer used, especially after the use of digital data [4].

In the mid and late of the 1980s, two important ways of ART, which led to a significant development in this area. These two methods are the decision theoretic approach (DT) and the static pattern (SP) and both methods depend on guessing a separation algorithm and choose the appropriate thresholds for them [5, 6]. It is noted here that if these thresholds are not selected very precisely, this will result in a decrease in the algorithm performance, especially when the noise is present, so there is a need to find other methods to recognize the type of modulation that is better performing even with noise.

Since the beginning of the 1990s [7, 8], there has been oriented towards industrial intelligence in the general recognition and classification processes, which has allowed it to be used to recognize the types of modulation. This step has greatly improved the performance of the system than it was, especially in the presence of noise. Most of the previously proposed methods were designed to recognize analog modulation methods, but recent contributions and researches on this subject are more focused on digital communications because of the increasing use of digital modulation [8, 9].

Most of the existing ART of the modulation type can recognize and classify a few types of digital modulation methods [8]. Most of which can determine the type of digital modulation from signals with high SNR, but with increasing the noise, the results will be worst and the system then will be unable to determine the type of digital modulation. The importance of this research is to adapt artificial intelligence to propose an ART system capable of recognizing and classifying many different types of digital modulation methods. Using an Artificial Neural Network with its complex algorithm is to increase both the performance of the system and the immunity against noise even with signals of low SNR.

2. RESEARCH AIM

This paper depends on the computational simulation method and mathematical modeling by using MATLAB program for its accuracy and simplicity to convert the models based on it to the real-time applications and followed the following methodology:

- Study and analyze the types of digital modulation under study.
- Investigation of the most important parameters of the studied digital modulation techniques.
- Analyze and compare these parameters at different values of SNR and examine their immune to this noise.
- The establishment of the artificial neural network (ANN), which will separate the types of digital modulation under study and then learning this network by two vectors (the inputs and targets vectors) that extracted from the key features with different noise levels. Through these two vectors, the network will be learned to be able to recognize the digital modulation type used even if the SNR has decreased and reached a low value.
- Testing the ANN.

3. MODULATION METHODS UNDER STUDY
Automatic recognition of the digital modulation types using the artificial neural networks (Saad S. Hreshee)

The modulation is one of the most important operations in the transmitter of the communications system. Since we cannot send the RF signal of the message wirelessly directly through the transmission channel, because it is usually of low frequencies, i.e. with large wavelengths and since the antenna length is usually equal to half or a quarter of the wavelength, its realization is practically impossible. The frequencies of various messages are also located within the same bandwidth. If the signals were sent directly, it will be overlapped making their separation in the receiver impossible. Therefore, these messages or signals are carried on high-frequency waves called (carriers) and supposed to have the corresponding frequency characteristics of the transmission channel. This carrying process is called modulation. In digital modulation techniques, an analog carrier signal is usually modulated by a binary message code, and this is carried out by changing one of the physical characteristics of the analog carrier (amplitude, frequency, phase or a combination of them) according to the change in the instantaneous value of the message signal containing the information [2, 10]. When the signal passes through a channel, it loses some of its energy and this is due to the impedance of the channel which causes the attenuation [11]. The shape of the signal also changes due to the distortion. The attenuation and distortion occur when more than one signal arrives through a lousy channel to the receiver side with different frequencies. Random noise is one of source of transmission losses, which generated from artificial sources and several natural. These losses in the transmitted signal make the identification of the modulation type used in the transmission of the signal is very difficult. In this paper, we will try to recognize the digital modulation type using artificial intelligence, so we'll take a quick look at the four types of digital modulation presented in this study. Figure 2 illustrate the difference between the theae modulation types (ASK, FSK, and PSK).

![Figure 2. Digital modulation signal with binary ASK, FSK, and PSK](image)

3.1. Amplitude shift keying (ASK)

This method of digital modulation corresponds to the amplitude modulation (AM) in the analog modulation. ASK is the amplitude modulation of the sinusoidal carrier wave, but with a digital signal, i.e. consists of ones and zeros. Thus, the carrier's amplitude changes as a result of the modulation between two values correspond to the Zeros and Ones as shown in Figure 2. Figure 3 show the digital modulation scheme of the binary ASK [11].

![Figure 3. Digital modulation scheme of the binary ASK](image)

3.2. Frequency shift keying (FSK)

Automatic recognition of the digital modulation types using the artificial neural networks (Saad S. Hreshee)
This method of digital modulation corresponds to the frequency modulation (FM) method of analog modulation. This frequency modulation is relatively simple and has low performance. The carrier frequency changes here between two values, when the digital modulation signal is a logical zero, the transmitted frequency is $f_1$, while when the digital modulation signal is a logical one, the transmitted frequency is $f_2$, as shown in Figure 2 while Figure 4 shows the digital modulation scheme of the binary FSK [11].

3.3. Phase shift keying (PSK)

This method in the digital modulation corresponds to the phase modulation method (PM) in the analog modulation. The PSK modulation is a form of frequency modulation, and the resulting signal after the modulation has a limited number of positions. When the number of phase positions of the output phase is two, the modulation method is called binary phase shift keying (BPSK) as shown in Figure 2 [11]. When phase positions of the output phase are four, the modulation called quadrature phase shift keying (QPSK). The phase positions of the output phase may become eight, which is called the eight-phase shift keying that is denoted by (8PSK), and the number of phase angle positions can be sixteen, the modulation being described as the hexadecimal phase shift keying and denoted by (16PSK) see Figure 5 for the digital modulation scheme of the binary PSK.

3.4. Orthogonal QAM modulation

Orthogonal QAM modulation is a type of digital modulation, where the information is located in both amplitude and phase of the signal. It is difficult to distinguish between a quadrilateral and 4PSK, but it is easy to note the difference between QAM and PSK when the number of levels is higher than 4 or 16 (see Figure 6 that illustrates the digital modulation scheme of QAM) [11].

4. KEY FEATURES

The major challenge in the modulation recognition and classification systems is always in the process of finding a set of key features. These keys must be reduced in order to reduce the number of inputs on the recognizer or the classifier (artificial neural network in this paper) also its complexity must be low to be easy to calculate and suitable for real-time applications. At the same time, it must include all the necessary key
features to ensure a reliable recognition and classification process. So, the ART process can be representing a balancing process between these two considerations.

Many numerous properties are imposed by different digital signals. Therefore, the appropriate features should be selected to identify these signals. These features are selected depending on their sensitivity to the type of digital modulation and on being insensitive to noisy channels and variations of SNR.

The eight key features proposed and studied in this paper can be calculated from the instantaneous characteristics of the signal (instantaneous amplitude, instantaneous frequency, and non-linear instantaneous phase) as shown in Figure 7. ASK can be recognized depending based on their amplitude variations, the FSK may be classified by its fixed instantaneous amplitude, and the PSK information is defined by its phase [2, 12-14].

![Figure 7. Key features](image)

The eight features have been selected and used in this paper to separate the following basic digital modulation methods (64QAM, 2PSK, 4PSK, 8PSK, 4ASK, 2FSK, 4FSK, 8FSK). These key features, in turn, are based on three values of the received signal $y(i)$ which is:

4.1. **Instantaneous amplitude $a(i)$:**

Instantaneous amplitude represents the instantaneous value of the signal $y(i)$. This instantaneous value can produce two important signals: the modified instantaneous value $A_n(i)$ and the central modified instantaneous value $A_{cn}(i)$ [12].

Modified instantaneous value $A_n(i)$:

$$A_n(i) = \frac{A(i)}{m_A}$$

Central modified instantaneous value $A_{cn}(i)$:

$$A_{cn}(i) = \left[ \frac{A(i)}{m_A} \right] - 1$$

where $m_A$ represents the average value of the instantaneous terminal.

4.2. **Non-linear instantaneous phase $\phi_{NL}(i)$**

The instantaneous phase of the modulated signal consists of two components:

Linear component: This is caused by the carrier. A non-linear component: which is caused by the modulated signal (information). The non-linear component $\phi_{NL}(i)$ for the phase of the signal $y(i)$ calculated by subtracting the linear component of the phase from the overall phase $\phi(i)$ [12].

*Automatic recognition of the digital modulation types using the artificial neural networks (Saad S. Hreshee)*
4.3. Instantaneous frequency

The instantaneous frequency represented by:

$$f(i) = \frac{1}{2\pi} \frac{d\varphi(i)}{dt}$$  \hspace{1cm} (3)

This means that the instantaneous frequency represents the instantaneous changes in phase. This instantaneous frequency can produce an important signal, which is the instantaneous central frequency that is given by:

$$f_N(i) = \frac{f(i) - m_f}{f_d}$$  \hspace{1cm} (4)

where \(m_f\) represents the main value of instantaneous frequency and \(f_d\) represent the symbol rate.

4.4. Key features

The first key: (\(\sigma_p\)) the standard deviation of the nonlinear instantaneous phase: This key feature helps to separate between digital modulation methods that carry the information by phase from other modulation methods. In such methods, there will be a transition of the instantaneous phase with different phase-steps, depending on the level of the modulation, so the value of this key will be small for these modulations [14].

$$\sigma_p = \sqrt{\frac{1}{N_c} \left( \sum_{A_n(i) > A_t} \varphi_{NL}(i)^2 \right) - \frac{1}{N_c} \left( \sum_{A_n(i) > A_t} \varphi_{NL}(i) \right)^2}$$  \hspace{1cm} (5)

where \(N_c\) is the number of samples in \(\{\varphi_{NL}\}\) for \(A_n(i) > A_t\), where \(A_t\) is the threshold value of \(A_n(i)\) when the filter provides the minimum amplitude of the signal sample due to high noise sensitivity and \(\varphi_{NL}(i)\) is the nonlinear component of the \(n^{th}\) instantaneous phase of the sample.

The second key: (\(m_a\)) the average value of the instantaneous amplitude: The average value of the instantaneous amplitude for all modulation methods under study except (64QAM, 4ASK) is close to one. For the 4ASK method, since the absolute value of the signal changes between the two values \(1, 1/3\), the average value of the instantaneous amplitude is equal to \(2/3\), while this value will be greater than one for the orthogonal amplitude modulation method (64QAM) [14].

The third key: (\(\sigma_f\)) Standard deviation of instantaneous frequency: The digital modulation methods that carry the information by frequency are 8FSK, 4FSK, 2FSK and therefore its instantaneous frequency changes either between two values for the 2FSK method or between four value for the 4FSK method or between eight values for the method 8FSK. Although the intervals between these values are the same for all these methods, while for other methods, their instantaneous frequency is equal to zero. This key is therefore used to differentiate the digital modulation methods (8FSK, 4FSK, 2FSK) from the other modulation methods and this key can also be used to differentiate between these methods (the instantaneous frequency change area in 8FSK method is twice the area of frequency change in the 4FSK method, which is four times the area of frequency changes in the 2FSK method [14].

$$\sigma_f = \sqrt{\frac{1}{N_c} \left( \sum_{A_n(i) > A_t} f_N^2(i) \right) - \frac{1}{N_c} \left( \sum_{A_n(i) > A_t} f_N(i) \right)^2}$$  \hspace{1cm} (6)

The fourth key: (\(\sigma_a\)) The standard deviation of the instantaneous amplitude: It is another key feature that performs the same work of the key (\(m_a\)) and is used for the same features [14].

$$\sigma_a = \sqrt{\frac{1}{N_c} \left( \sum_{i=1}^{N} A_{cn}^2(i) \right) - \frac{1}{N_c} \left( \sum_{i=1}^{N} |A_{cn}(i)| \right)^2}$$  \hspace{1cm} (7)

The fifth key: (\(y_{max f}\)) The maximum value of the spectral power density of the central instantaneous frequency: This key helps to differentiate between the digital modulation methods (8FSK, 4FSK, 2FSK) and any other modulation methods (as (\(\sigma_f\)) works). As the instantaneous frequency change area of 8FSK is twice the area of instantaneous frequency change for 4FSK and four times the area of instantaneous frequency change for 2FSK so the 8FSK method has a higher value than they are in the two cases 4FSK and 2FSK which enables the differentiation between 8FSK, 4FSK, 2FSK methods. This key can be determined by [14]:
\[ \gamma_{max_f} = \max |DFT(f_{cn}(i))|^2 \]  

(8)

The symbol DFT represents the Discrete Fourier transform.

The sixth key: \( (\gamma_{max_a}) \) the maximum value of the spectral power density of the central instantaneous amplitude: This key helps to differentiate between the digital modulation methods that carry the information by amplitude i.e. (64QAM, 4ASK) and other methods of modification. This feature is given in the following relationship [14]:

\[ \gamma_{max_a} = \max |DFT(A_{cn}(i))|^2/N_s \]  

(9)

where \( N_s \) represents the number of samples in one frame of the signal.

The seventh key: \( (Kur_a) \) The fourth-degree torque of the central instantaneous amplitude: which given by:

\[ Kur_a = \frac{E\{A_{cn}^4(n)\}}{[E\{A_{cn}^2(n)\}]^2} \]  

(10)

where \( E \{ \} \) represents the expected value.

The fourth-degree torque represents the kurtosis around the average. It’s found that this key helps to differentiate between digital modulation methods that carry information in the amplitude, which has a small kurtosis than other methods [14].

Eighth key: \( (Kur_f) \) The fourth-degree torque of the central instantaneous frequency: which given by:

\[ Kur_f = \frac{E\{f_{N}^4(n)\}}{[E\{f_{N}^2(n)\}]^2} \]  

(11)

This feature helps to differentiate between digital modulation methods that carry information by frequency of the other methods of modification [14].

5. ARTIFICIAL NEURAL NETWORKS

For the recognition and classification processes, many techniques were found [3], one of these techniques was the artificial neural network (ANN), which is one of the most appropriate methods in the signal classification [3, 15]. Artificial neural networks are a collection of algorithms, working as the human brain neurons. These networks are designed to recognize patterns. The patterns identified by these networks are numerical, included in vectors, which must be translated into all real-world data, i.e. images, audio, text or time series. ANNs help to recognize and classify [16-18].

ANN merits in the ART are its high classification rate, reduced computational complexity, and improved accuracy. While its demerits are ineffective when a large number of input features are used and any small data set requires long training time for computation. In this paper, depending on neural network characteristics, we used the feedforward with multi-layered neural networks as a classifier and identifier and for the optimization of the network weights, we use backpropagation algorithm as the most common algorithms, using the controlled learning method, with the number of training pairs composed of patterns vector and targets vector as shown in Figure 8.

The aim of weights optimization is to get a stable state that makes the network properly responsive to the patterns vector used in the training process and to any new input vector and the training process continues until the error that exists between the output vector and the target to the smallest value (threshold value) which is predetermined (see Figure 8).

In this paper, the previous eight features used to build the patterns vector. The patterns vector was got from the concluding of the key features from a noisy modulated signal, i.e. a signal that crossed a communication channel, and a known modulation type and so we have another vector which is the targets vector. The neural network was learned through the patterns and targets vectors to get the ability to recognize and classify the type of digital modulation. This neural network was subjected to several experiments to change and modify its structure (changing the number of hidden layers or the number of nodes in these hidden layers or change the learning algorithm) until it finally got the best results appropriate for this application [19-25].
6. RESULTS AND DISCUSSION

The MATLAB environment was used to write system programs because it is a high-level language that contains libraries for communication and signal processing systems as well as a library of neural networks. So, we used parameters of this library to design, learn, and test the back-propagation, artificial neural network used in our system. Simulation programs to generate digital modulation signals proposed in this paper at a sampling frequency $f_s = 1200kHz$ and the coding rate $f_d = 12.5kHz$ to produce a specified number of frames for each digital modulation method and then partition the signal within a frame of length (4096) sample that determined to achieve the best results.

The eight key features of each frame were then extracted and grouped into one vector and for the best performance, a median filter was used with a window of seven samples, a sixth-degree non-recursive filter used to pass the signal through it before concluding the key features. Figure 9 to Figure 12 show the relationship between the key features of the digital modulation methods under study with the signal-to-noise ratio that changes between (0-20 dB) for an AWGN communication channel. We have included these curves to show how reliable these parameters are and how it is immune to Noise.

In order to give the learning to the artificial neural network, the key features were calculated at two values of SNR to eventually form the input vector (pattern vector) and the output vector (target vector) used to learn the network. The neural network proposed here consisting of an input layer with a number of nodes equal to the number of key features, two hidden layers and an output layer with a number of nodes equal to the number of modulation methods that must be recognized. Figure 9 to Figure 12 illustrate that with small values of SNR, most of the modulation methods are clear and can be detected but when the SNR increases, the value of key features begins to decrease clearly and converge with each other so that there is a difficulty to recognize between these methods while the methods that depend on the specified key feature will be distinguishable.
Figure 9. The effect of SNR for, (a) standard deviation of the non-linear instantaneous phase, (b) the average value of the instantaneous amplitude

Figure 10. The effect of SNR for, (a) standard deviation of the instantaneous frequency, (b) the standard deviation of the instantaneous amplitude

Figure 11. The effect of SNR for, (a) the maximum value of the spectral power density of the central instantaneous frequency, (b) the maximum value of the spectral power density of the central instantaneous amplitude
After completing the learning of the artificial neural network, we tested this network to recognize a testing modulated signals transmitted to the network, see Figure 13. As seen in Table 1, the modulation type with deferent recognition techniques and deferent SNRs used in some of the previous works are compared with the results of this paper.

![Figure 12](image1.png)

Figure 12. The effect of SNR for, (a) the fourth-degree torque of the central instantaneous amplitude, (b) the fourth-degree torque of instantaneous central frequency attributed

![Figure 13](image2.png)

Figure 13. Recognition percentage ratios for each modulation type at different values SNR

Table 1. A comparison of the results of this study with some of the classification methods of digital modulation used previously

<table>
<thead>
<tr>
<th>References</th>
<th>The used modulation types</th>
<th>The used model and technique</th>
<th>The obtained success rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Almohamad et al.</td>
<td>2ASK, QPSK, 16QAM</td>
<td>PCA + SVM (SNR: 0 - 30 dB)</td>
<td>99.83 (on average)</td>
</tr>
<tr>
<td>[26]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ali, and Yang Yu</td>
<td>BPSK, 4QAM, 16QAM, 64QAM</td>
<td>2-Layered deep neural network model</td>
<td>95.40</td>
</tr>
<tr>
<td>[9]</td>
<td></td>
<td>(SNR: 10 dB)</td>
<td></td>
</tr>
<tr>
<td>Dai et al. [27]</td>
<td>2PSK, 4PSK, 8PSK, 16QAM, 64QAM</td>
<td>A fully connected 2-layer feed-forward deep neural network (DNN)</td>
<td>94.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(SNR: 15 dB)</td>
<td></td>
</tr>
<tr>
<td>Our work</td>
<td>64QAM, 2PSK, 4PSK, 8PSK, 4ASK, 2FSK, 4FSK, 8FSK</td>
<td>Feedforward + Backpropagation with 2 hidden layers ANN (SNR: 5, 10, and 15 dB)</td>
<td>68%, 89.1%, 91% respectively but when we exclude the PSK, the result will be 75.2%, 98.8%, 99.6% respectively</td>
</tr>
</tbody>
</table>

7. CONCLUSION

In this study, an artificial neural network with an input-layer of eight nodes, two hidden-layers and an output-layer of eight nodes too are used to get the best results. The proposed system proves its ability to achieve a recognition ratio of 68% for a testing signal with SNR=5dB, and 89.1% for a testing signal with SNR=10dB and 91% is known for a testing signal with SNR=15dB. The channel proposed in this work was the AWGN and hence the relatively low recognition rate for the SNR=5dB is due to the high sensitivity of some of the selected determinants of high noise. Some types of modulation understudy got the best recognition ratio, where the 64QAM was the best one, while another modulation technique which is the 4PSK got the lowest recognition. Truly, when we exclude all the PSK modulation method studied in this paper, the result will be 75.2%, 98.8%, 99.6% for the SNR= (5dB, 10dB, and 15dB) respectively and that means, the PSK is very sensitive to the AWGN channel. The proposed ANN recognition algorithm is very efficient in the 64QAM and 4ASK in any SNR ratio. For the M-FSK modulation method, the proposed algorithm can be representing a good recognition method for the high values of SNR. The study also proved that the ANN algorithm is not efficient for M-PSK modulation signals, especially for large M.

REFERENCES


BIOGRAPHY OF AUTHOR

Asst. Prof. Dr. Saad Saffah Hasson Hreshie was born in Baghdad, Iraq, in 1974. He received his BSc degree in Electrical Engineering from the Electrical Engineering Department, College of Engineering, University of Babylon, Iraq, in 1997. He obtained his MSc in Electronic Engineering from the Electrical and Electronic Engineering Department, University of Technology, Iraq, in 2000, while his PhD. in Electronic and Communication Engineering from the Electrical Engineering Department, College of Engineering, University of Basrah, Iraq, in 2007. Since 2001, he has been with the staff of the Department of Electrical Engineering, College of Engineering, University of Babylon, where he is now the head of the electrical engineering department. His main research interests are antenna, signal processing, encryption and communication security.