DOI: 10.17725/rensit.2020.12.319

Cumulative features for determining the type of signal manipulation

Vagid K. Kurbanaliev

A.I. Berg Central Research Institute of Radioengineering, http://www.cnirti.ru/
Moscow 107078, Russian Federation

E-mail: vagdag05@mail.ru

Received October 12, 2020; peer reviewed October 19, 2020; accepted October 26, 2020

Abstract: **This paper describes describes the apparatus of cumulant analysis in relation to the problem of recognizing the types of signal modulation. The article presents the results of using artificial neural networks in the task of automating the detection of intra-pulse modulation signs for the identification (classification) of signals. A mathematical model of a phase-shift keyed signal is developed, the main properties of this type of signals are described and a method is proposed that allows one to determine the type of signal manipulation based on the calculation of informative (cumulative) features. Simulation was carried out in Matlab/Simulink.**

Keywords: **phase-shift keying, artificial neural networks, cumulant analysis**

UDC 621.37

For citation: Vagid K. Kurbanaliev. Cumulative features for determining the type of signal manipulation. RENSIT, 2020, 12(3):319-328. DOI: 10.17725/rensit.2020.12.319

Contents

1. INTRODUCTION

Automatic recognition of the types of modulation of radio engineering and radio communication signals is an important function in modern electronic intelligence and electronic warfare systems. Automatic recognition of modulation types mainly consists of feature extraction and modulation classification. Modulation feature extraction consists of a series of transformations and analysis algorithms in the time domain, frequency domain, or time-frequency domain. Classification processing also consists of various pattern recognition and machine learning algorithms, deep learning and clustering algorithms. The problem of determining the parameters of signals is currently relevant for several reasons: determining the parameters will help to identify the transmitting device, in case of successful recognition of the type of modulation, the transmitted message can be restored, and it will also be possible to introduce controlled interference modulations and induce active interference to suppress radio location and communication channels. There are many algorithms for determining signal parameters. For example, in [1,2], a method is proposed for recognizing the type of modulation by the signal constellation. The reason for the shortcomings of this recognition method is low information content, the probability of correct recognition strongly depends on the signal-to-noise ratio (SNR).

This article discusses a method for recognizing phase-shift keyed signals based on cumulative analysis based on the use of more stable informative features. Cumulants (semi-invariants) are the coefficients of the Maclaurin series expansion of the characteristic function of a random variable. This method will make it possible to determine the type of modulation under conditions of a priori uncertainty. Neural networks were used to automate the identification process. The article discusses phase-shift keyed signals (BPSK, QPSK and 8-PSK), i.e. signals, during the manipulation of which a phase change occurs. Phase-shift keying signals are widely used in radio communication systems, since they contribute to an increase in the degree of noise immunity of the system and make it possible to effectively use the frequency range of the radio channel.

2. MATHEMATICAL MODEL OF A PHASE-SHIFT KEYED SIGNAL

Traditional analog modulation techniques change the high frequency signal in only one dimension. Modern quadrature (*I*/*Q*) modulators alter the RF carrier signal in two dimensions. They form a modulation signal in complex form, consisting of the sum of two baseband signals i(t) and q(t), where i(t) is the in-phase component and q(t) is the quadrature component. Orthogonally located on the quadrature plane, they form a vector with length A(t) and phase φ(t), see **Fig. 1**. Components i(t) and q(t) are normalized to a constant value ≤1.

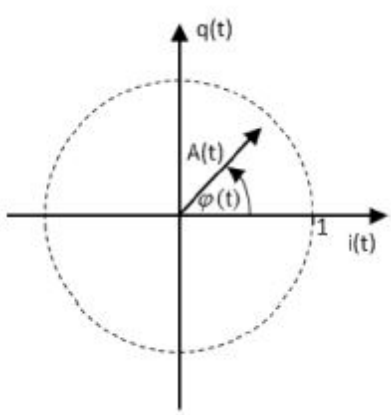

Fig. 1. *Modulation vector on the I/Q-plane.*

The constellation diagram provides a graphical representation of the *I* and *Q* components of a digitally modulated signal. In **Fig. 2** shows a constellation diagram for binary phase shift keying (BPSK), 4-PSK (QPSK) and for 8-PSK modulation at 20 dB SNR.

The digital signal fluctuates between two fixed values of the signal, and it is considered that this is not modulation, but a keying of the carrier wave. There are three options for manipulation:

- amplitude shift keying (AMn or ASK);
- frequency shift keying (FSK or FSK);
- phase shift keying (PSK or PSK).

Phase shift keying shifts the phase of the carrier waveform according to the digital bit sequence. An absolute phase angle is assigned to each transmitted symbol. For better separation, the phase states are usually evenly distributed over a 360° circle.

For an analytical description, the following expression is usually used:

$$V(t) = A\cos(\omega t+\theta), \quad (1)$$

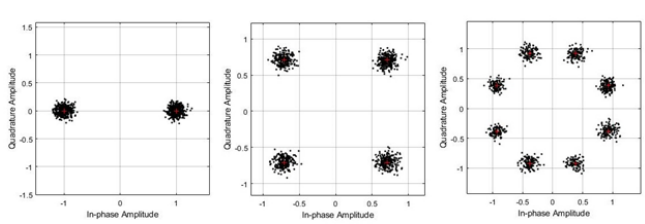

Fig. 2. *Constellation diagram for BPSK (left), QPSK (middle) and 8-PSK (right) modulation.*

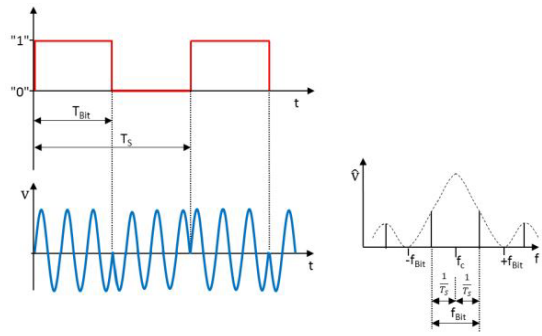

Fig. 3. *Phase Shift Keying in the frequency and time domains for a periodic bit sequence 1-0.*

where A is the amplitude, ω is the initial frequency, θ is the phase of the signal.

When manipulated with a square wave, the BPSK signal looks like this:

$$V(t) = \begin{cases} \hat{V}_c \cos \omega_c t & \text{for logical 1,} \\ -\hat{V}_c \cos \omega_c t & \text{for logical 0.} \end{cases} \quad (2)$$

where is the amplitude, ω_c is the angular frequency.

In **Fig. 3** shows a data signal and a modulation component with a step length Tbit and a period T_s in the time domain, as well as its spectrum, where

$f_{bit} = 1/T_{bit}$.

In **Fig. 4** shows the QPSK signal for the symbol sequence 10 11 01 10 00. For the carrier, the following expression can be written:

$$V_c(t) = \hat{V}_c \cos\left(\omega_c t + \frac{3\pi}{4} + \Delta\varphi\right). \quad (3)$$

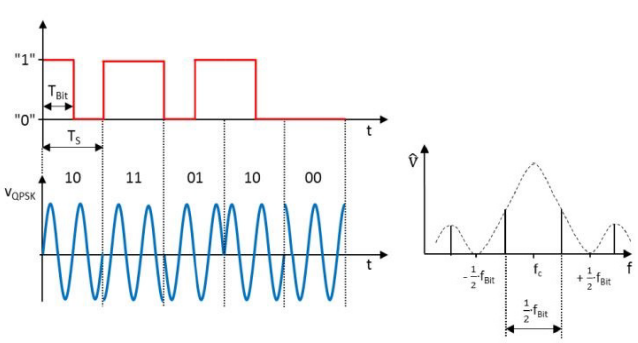

Fig. 4. *Time characteristics of the binary signal and the QPSK carrier waveform and the signal spectrum.*

Table 1

Correspondence table for QPSK modulation

Character (two-bit)	Δφ
1 1	+45°
0 1	+135°
0 0	-135°
0 1	-45°

The manipulation is carried out in accordance with **Table 1.**

3. METHODS FOR RECOGNIZING THE TYPES OF SIGNAL MODULATION

There are several approaches to determining the type of modulation. The simplest is based on spectral analysis. Multiplying the current phase of the PSK signal by 2, 4, etc. allows you to consistently shoot one level of manipulation. A simple way to do this is to raise the signal to the appropriate power. Let, for example, the analyzed signal is described by expression (1), where in the case of BPSK signal processing θ = {0, π}. Let's square this signal

$$V^2(t) = A^2 \cos^2(\omega t + \theta) =$$

$$= A^2\left(\frac{1}{2} + \frac{1}{2}\cos(2\omega t + 2\theta)\right) =$$

$$= A^2\left(\frac{1}{2} + \frac{1}{2}\cos(2\omega t)\right).$$

.

Thus, we have obtained an expression that describes a signal with a doubled frequency, and does not contain information about the original phase of the signal. Obviously, a similar approach can be used to remove from QPSK, 8-PSK, etc. signals.

Each exponentiation operation leads to a corresponding transformation of the spectrum of the analyzed signal. Doubled, quadrupled, etc. harmonics appear in it. center frequency, which are signs that characterize a specific type of signal. But

the main disadvantage of this approach is its high sensitivity to signal interference during transmission. In practice, it may be easy to determine the frequency corresponding to the central harmonic, but it may not be possible to determine their number [3].

Of greatest interest is the approach to determining the type of signal modulation based on the threshold method, the classification of statistical features. Currently, informative features can be divided into spectral and cumulative.

Spectral features are calculated based on the instantaneous values of the parameters of the received signal: instantaneous amplitude, phase and frequency. The spectral features of the signal in the problem of recognizing the types of modulation were

Table 2

Spectral characteristics used to classify modulation.

The maximum value γ_{max} of the spectral power density for the normalized and centered instantaneous amplitude of the received signal	$\gamma_{\max} = \frac{\max \lvert DFT(A_{cn}(i)) \rvert^2}{N_s}$, where DFT is the discrete Fourier transform, Acn is the normalized and centered envelope of the input signal, N_s is the number of samples. Envelope normalization is performed as $A_{cn} = A_i/\mu_A - 1$, where μ_A is the average value of the instantaneous envelope values in a sample of samples of length N.
Standard deviation σ_{ap} of the absolute value of the centered nonlinear component of the instantaneous phase	$\sigma_{ap} = \sqrt{\frac{1}{N_c}\left(\sum_{A_n(i)>A_t} \varphi_{NL}^2(i)\right) - \frac{1}{N_c}\left(\sum_{A_n(i)>A_t} \lvert \varphi_{NL}(i) \rvert\right)^2}$, where Nc is the number of samples that satisfy the condition $A_n(i) > A_t$, where A_t is the threshold value that filters samples with a low signal amplitude due to their high sensitivity to noise, $\varphi_{NL}(i)$ is the nonlinear component of the instantaneous phase.
Standard deviation σ_a of the normalized and lumped instantaneous amplitude.	$\sigma_{ap} = \sqrt{\frac{1}{L}\left(\sum_{A_n(i)>t_{th}} a_{cn}^2(i)\right) - \frac{1}{L}\left(\sum_{A_n(i)>t_{th}} \varphi_{cn}(i)\right)^2}$, where L is the signal duration, t_{th} is the threshold value
Standard deviation σ_{dp} of the nonlinear component of the instantaneous phase	$\sigma_{ap} = \sqrt{\frac{1}{N_c}\left(\sum_{A_n(i)>A_t} \varphi_{NL}^2(i)\right) - \frac{1}{N_c}\left(\sum_{A_n(i)>A_t} \varphi_{NL}(i)\right)^2}$, The expression for σ_{dp} completely coincides with the expression for σ_{ap}, except for the instantaneous phase modulus
Standard deviation σ_{aa} of the absolute values of the normalized and centered instantaneous signal amplitude	$\sigma_{aa} = \sqrt{\frac{1}{N_c}\left(\sum_{i=1}^{N} A_{cn}^2(i)\right) - \frac{1}{N_c}\left(\sum_{i=1}^{N} \lvert A_{cn}^2(i) \rvert\right)^2}$, where A_{cn} is the normalized and centered instantaneous amplitude of the input signal
Standard deviation σ_{af} of the absolute values of the normalized and centered instantaneous frequency of the signal	$\sigma_{af} = \sqrt{\frac{1}{N_c}\left(\sum_{A_n(i)>A_t} f_N^2(i)\right) - \frac{1}{N_c}\left(\sum_{A_n(i)>A_t} \lvert f_N(i) \rvert\right)^2}$, where f_N is the normalized frequency
Kurtosis coefficient μ^a_{42} for normalized and centered instantaneous amplitude	$\mu_{42}^a = \frac{E\{A_{cn}^4[n]\}}{\{E\{A_{cn}^2[n]\}\}^2}$
Kurtosis coefficient μ^f_{42} for normalized and centered instantaneous frequency	$\mu_{42}^f = \frac{E\{f_N^4[n]\}}{\{E\{f_N^2[n]\}\}^2}$

first presented in the article by E.E. Azzouz, A.K. Nandi, published in 1995 and in a monograph by these authors, published a year later [4,5]. The physical basis for using this system of features is the features of the instantaneous values of the amplitude, phase and frequency of the signal for various types of modulation. Several spectral features are used to identify modulation that are suggested in the literature and are summarized in **Table 2**.

The use of spectral features has the following disadvantages: with low-precision synchronization of the carrier frequency, some features lose their information content; informative signs are sensitive to noise and depend on SNR [6].

Cumulative features make the algorithm less sensitive to unwanted noise and carrier frequency deviation. Cumulative features are calculated from low-frequency data based on the characteristic functions of stationary random processes [7, 8].

The cumulative features of two-dimensional random processes are expressed through their mixed moments, which are defined as follows: let there be a random stationary process - its conjugate, then the mixed moment is expressed by the formula where $k + n$ is the order of the mixed moment. In this case, the expression for the cumulant of order $k + n$. Distribution cumulants are in many ways much more informative distribution parameters than moments. This is mainly due to the fact that in many practical important cases, cumulants of distributions, in contrast to moments, can be neglected. Moreover, there are distributions of random variables, the cumulants of which, starting from a certain order, all vanish, while their moments are not equal to zero. For example, for a Gaussian distribution, only the first two cumulants are nonzero, and at the same time, none of the moments is zero [9].

Cumulants characterize the statistical relationship between the distributions of the instantaneous phase of the signal and can be expressed in terms of joint moments according to (4):

$$C_{k+n,n} = cum[\underbrace{s,...,s}_{k}\underbrace{\overline{s},...,\overline{s}}_{n}] = \\ = \sum_{\forall\Omega}(-1)^{p-1}(p-1)!E\left[\prod_{i\in\Omega_1}s_i\right]...E\left[\prod_{i\in\Omega_p}s_i\right], \quad (4)$$

where $p = k + n$ and the summation occurs over the set for $\Omega = (\Omega_1, \Omega_2, ..., \Omega_p)$ by $i = \overline{1, p}$.

To solve the problem of classifying phase-shift keyed signals, consider the values of cumulants C_{20}, C_{21}, C_{40}, C_{41}, C_{42}. The calculation takes place according to formulas 5-9 [10]:

$$C_{20} = \frac{1}{N}\sum_{n=1}^{N} s^2(t) \quad (5)$$

$$C_{21} = \frac{1}{N}\sum_{n=1}^{N} |s(t)|^2 \quad (6)$$

$$C_{40} = \frac{1}{N}\sum_{n=1}^{N} s^4(t) - 3C_{20} \quad (7)$$

$$C_{41} = \frac{1}{N}\sum_{n=1}^{N} s^3(t)\overline{s}(t) - 3C_{20}C_{21} \quad (8)$$

$$C_{42} = \frac{1}{N}\sum_{n=1}^{N} |s(t)|^4 - |C_{20}|^2 - 2C_{21}^2. \quad (9)$$

Cumulant values are complex numbers. The main differences in the values of cumulants for different types of digital modulation are manifested in the values of their real parts. Therefore, the values of the real parts of the cumulants are taken below as recognizing features.

Table 3

Moment values

	M_{20}	M_{21}	M_{40}	M_{41}	M_{42}
BPSK	1	1	1	1	1
QPSK	0	1	1	1	1
8-PSK	0	1	0	0	1

Table 4

Cumulant values

	C_{20}	C_{21}	C_{40}	C_{41}	C_{42}
BPSK	1	1	-2	-2	-2
QPSK	0	1	1	0	-1
8-PSK	0	1	0	0	-1

The values of cumulants and moments are presented in **Tables 3** and **4**, respectively, for three types of signal manipulation in the absence of noise.

As can be seen from Table 3, the moments for one type of manipulation have practically the same meaning and are not informative.

The main advantage of cumulants is their high accuracy and unambiguous classification.

Decision-making on classification can also be achieved by dividing into subgroups. The combined use of cumulants and moments allows you to reduce the likelihood of error both with respect to cumulative analysis only and relative application of moments only.

4. ARTIFICIAL NEURAL NETWORKS

In order to automate the process of recognizing the type of signal manipulation, artificial neural networks can be used [11-13].

In **Fig. 5** shows the structure of a two-layer neural network.

In the hidden first layer, each neuron has a vector of weights - the number of neurons in the hidden layer.

$$net^{(j,1)} = x^T w^{(j,1)} = w_0^{(j,1)} + \sum_{i=1}^{N} w_i^{(j,1)} x_i, \quad o^{(j,1)} = f(net^{(j,1)}), \tag{10}$$

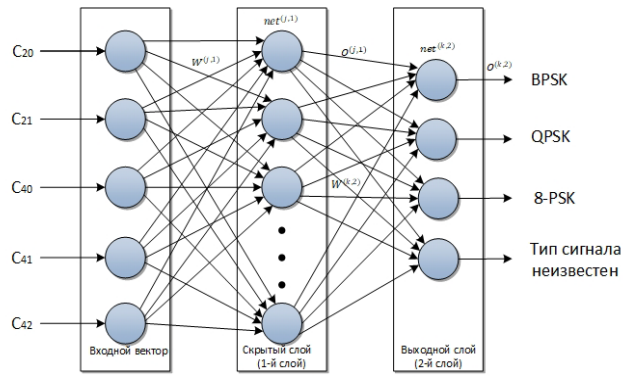

Fig. 5. *The structure of a two-layer neural network.*

Is the output of the j-th neuron of the hidden layer, is the input row vector, is the weight coefficient row vector, T is the transposition sign, is the nonlinear activation function.

In the output layer, each k neuron of the output layer has a vector of weight coefficients - the number of neurons in the output layer, which, in a particular case, is equal to the number of classes in the pattern recognition problem, o (1) is the vector of outputs of the hidden layer neurons, which are inputs to the output layer [14, 15].

$$net^{(k,2)} = w^{(j,1)T} o^{(1)} = w_0^{(k,2)} + \sum_{i=1}^{N} w_j^{(j,1)} o^{(j.1)}, \quad o^{(k,2)} = f(net^{(k,2)}),\ k = \overline{1, N_2}. \tag{11}$$

The output of the neural network reflects the type of manipulation of the received signal.

This article discusses three types of manipulation (BPSK, QPSK and 8-PSK), so the number of neurons in the output layer is 3 + 1, taking into account the unknown type. The number of neurons in the hidden layer 1 is set to 10. If the received signal is BPSK modulation, then the output of the neuron is "1", and the rest are equal to 0.

Preliminary training of the neural network is carried out using the Levenberg-Marquardt method (trainlm). To train a

neural network, a set of training samples is used, which consists of two components: input and target. The input component represents the values of the cumulants calculated for signals with the BPSK, QPSK and 8-PSK keying types, and the target is the desired state of the neural network outputs, corresponding to the output component of the sample. To assess the quality of training, the cross-validation method (cross-check) is used. The last action of the preparatory stage is the preservation of the weight connections of the neural network in long-term memory [16, 17].

In **Fig. 6** shows an algorithm for training a neural network.

At the stage of recognizing the type of manipulation of the signal received with the sampling frequency Fs in the frequency band ΔF, having a duration ΔT, we form a vector of features C20, C21, C40, C41, C42 in accordance with formulas (5) - (9). For this purpose, we form the in-phase and quadrature components from the received signal for sequential calculation of cumulants. The generated feature vector is fed to the input layer of the neural network. Then, one by one for the hidden and output layers of the neural network in each neuron, the value of the activation function is calculated based on the output values and the coefficients of the weight connections. Due to the fact that each neuron of the output layer corresponds to a certain type of manipulation, the choice of a neuron with the maximum value of Fout in a given interval will determine the desired type of manipulation. Otherwise, it is considered that the type of signal manipulation is unknown, and re-recording and analysis is required.

The generalized algorithm for recognizing the type of signal manipulation is shown in **Fig. 7.**

The training set consisted of 324 signals at different SNRs. The length of each signal was N = 1024 counts.

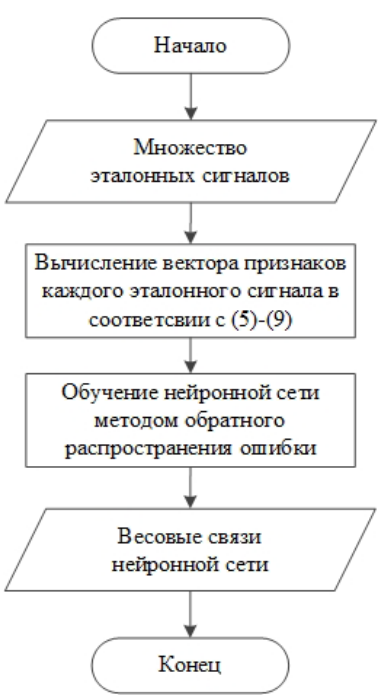

Fig. 6. *Neural network learning algorithm.*

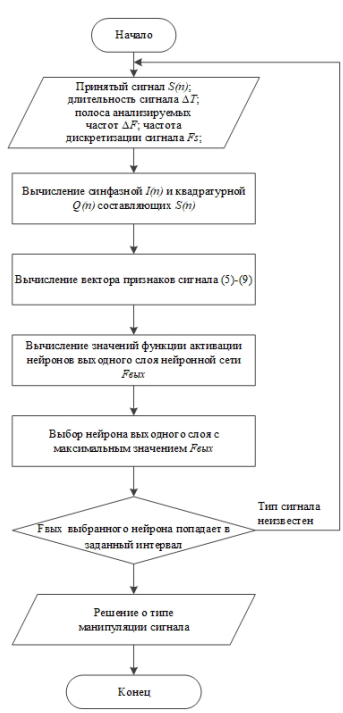

Fig. 7. *Generalized algorithm for recognizing the type of signal manipulation.*

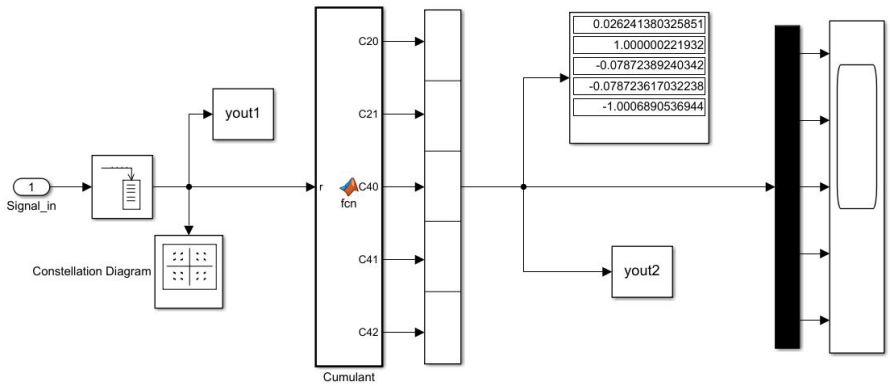

Fig. 9. Model for calculating cumulants.

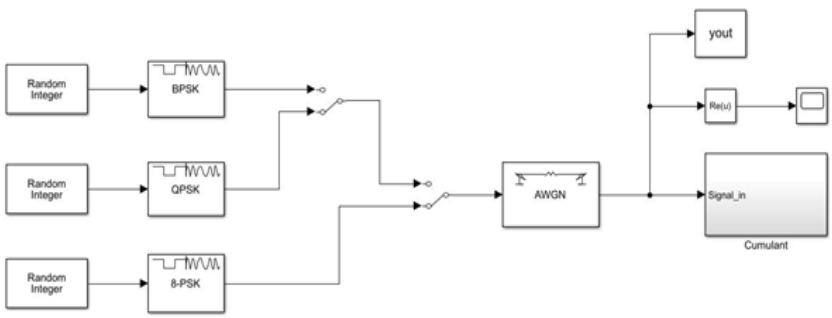

Fig. 8. *Mathematical model for generating initial signals.*

In **Fig. 8, 9** show the created mathematical models for generating initial signals and for calculating cumulant

5. RESULTS OF SIGNAL MODELING IN MATLAB/SIMULINK

The neural network parameters are shown in **Table 5**.

The estimation of the probabilities of recognition of the types of signal manipulation obtained in the MatLab/Simulink simulation environment is presented in **Table 6**.

As follows from the table, with an SNR of no worse than 4-5 dB, this network will work very well, but at low SNR values, the accuracy will deteriorate.

As a result of the experiment (a series of n = 1000 independent observations), the random variable obeys the binomial distribution law with a standard deviation (RMSD) where p is the probability of recognizing the type of manipulation. The error in measuring the recognition probabilities was ~ 1%. The conditional probability for a given SNR is selected as a criterion for evaluating the

Table 5

The main parameters of the neural network.

Number of layers	2
Number of inputs	5
Number of neurons in the hidden layer	10
Number of neurons in the output layer	3
Hidden Layer Activation Function	Tansig (hyperbolic tangent sigmoidal transfer function)
Activation function in the output layer	Purelin (linear transfer function)
The number of signal counts for calculating cumulants	1024

Table 6

Recognition results

SNR (dB)	Modulation types		
	BPSK	QPSK	8-PSK
4	0.96	0.95	0.95
10	0.98	0.97	0.96
15	0.99	0.99	0.98

effective decision-making for recognizing the type of signal manipulation.

6. CONCLUSION

A method for recognizing types of signal manipulation based on the calculation of cumulant features for training a neural network is considered. The article investigates cumulants up to the fourth order for recognition of three types of signal manipulation. This method allows recognizing types of manipulation with a probability of at least 0.95 at signal-to-noise ratios of 4 dB. The simulation results show the high efficiency of this approach. Analytical expressions for calculating cumulants are presented and the accuracy of the experiment is justified. A preliminary analysis of the method showed that the approach described in the work possesses the properties of nonparametry (robustness) and is similar to cepstral analysis. The implementation of this method will expand the range of tasks to be solved by remote methods and supplement the information obtained about the sources of radio emission.

In the development of the research topic, it is supposed to expand the number of recognized types of modulation and the selection of significant cumulants of higher orders and the implementation of this method on the target FPGA platform of Altera, Xilinx. Also, thanks to a new set of essential features (combination of cumulants, moments and spectral characteristics), due to the expansion of the vector of features of classified signals, it is possible to increase the probability of correct recognition of the types of signal modulation.

REFERENCES

1. Ajemov SS, Klenov NV, Tereshonok MV, Chirov DS. Methods for recognizing types of digital signal modulation in cognitive radio systems. Moscow University Bulletin. Series 3. Physics. Astronomy, 2015, 6: 19-27.
2. Velampalli C. Hierarchical blind modulation classification in the presence of carrier frequency offset. Master's Thesis. Communications Research Center, 2010: 1-39.
3. Stepanov AV, Matveev SA. Methods of computer processing of signals from radio communication systems. Moscow, SOLON-press, 2003, 145 p.
4. Azzouz EE, Nandi AK: Automatic identification of digital modulation types. Signal Processing, 1995, 47 (1): 55-69. DOI: 10.1016 / 0165-1684 (95) 00099-2.
5. Azzouz EE, Nandi AK. Automatic Modulation Recognition of Communication Signals. Kluwer Academic Publishers, Dordrecht, Netherlands, 1996.
6. Dhamyaa H. Al-Nuaimi, Ivan A. Hashim, Intan S. Zainal Abidin, Laith B. Salman, Nor Ashidi Mat Isa. Performance of Feature-Based Techniques for Automatic

Digital Modulation Recognition and Classification. Electronics, 2019, 8 (12): 1407. DOI: 10.3390 / electronics8121407.

7. Avedyan ED, Dam VN. On the choice of cumulant features in the problem of recognizing types of digital modulation of radio signals Informatization and communication, 2015, 4: 11-15.
8. Dam Van Nyit. Neural network technologies in the problem of automatic recognition of digital modulation types. Dissertation. Moscow, MIPT, 2018, 159 p.
9. Malakhov AN. Cumulant analysis of random non-Gaussian processes and their transformations. M., Sov. radio, 1978, 376 p.
10. Zhechen Zhu, Asoke K. Nandi. Automatic modulation classification principles, algorithms and applications. London, John Wiley & Sons, Ltd, 2015, 175 p.
11. Ajemov SS, Tereshonok MV, Chirnov DS. Method and device for automatic recognition of types of manipulation of radio signals. Patent RU 2510077, IPC G06N 3/02, 10/27/2013 Bull. No. 30.
12. Elizarov VV, Kasatkin AS, Nalivaev AV, Smirnov PL, Shepilov AM. Method for automatic recognition of the type of manipulation of radio signals. Patent RU 2622846, IPC G06N 3/02, 20.06.2017 Bull. No. 17.
13. Kolbasko IV, Kvasov AV, Yuriev IA, Fesenko MV. A method for recognizing the types of manipulation of radio signals. Patent RU 2682304, IPC G06N 3/02, 03/18/2019 Bull. No. 8.
14. Khaikin S. Neural networks: a complete course. M., Ed. house "Williams", 2006, 1104 p.
15. Galushkin AI. Neural networks: basic theories. M., Hotline-Telecom, 2010, 480 p.
16. Vorontsov KV. A combinatorial approach to assessing the quality of learning algorithms. In the book: Mathematical problems of cybernetics. Ed. ABOUT. Lupanov. Moscow, Fizmatlit, 2004, Volume 13, p. 5-36.
17. Lupanov OB (ed.) Mathematical problems of cybernetics. Moscow, Fizmatlit, 2004.
18. MathWorks [Electronic resource]. - Access mode: https://www.mathworks.com/ (date of access: 09/01/2020).