Automatic Modulation Classifier Using Artificial Neural Network Trained by PSO Algorithm

Li Cheng[1] and Jin Liu[2]

[1] College of Information and Communication Engineering, Harbin Engineering University, Harbin, China
[2] Institution of China Ship Industry Corporation, Wuhan, China
Email: whtimesvip@sohu.com; instoshliu@sohu.com

Abstract—The back propagation (BP) neural networks have been commonly used for automatic modulation recognition since the late 1990s. However, the back propagation algorithm easily falls into local minimum and the network learning is sensitive to initial weight values which usually determined by experience. The particle swarm optimization (PSO) algorithm is a global heuristic searching technology. By combing these two techniques, this paper presents a novel automatic modulation classifier. By training BP neural network with PSO optimization, this classifier can identify seven digital signals with six input features. The experiment results show that this recognizer achieves about 92% recognition performance at the SNR of 0dB.

Index Terms—automatic modulation classifier, particle swarm optimization, back propagation neural network, parameter optimizations.

I. Introduction

An automatic modulation classification (AMC) is system that automatically identifies the modulation type of the received signal without the preventive knowledge of some parameters [1]. When the modulation types of an unknown signal are identified, it is possible to demodulate it properly with the right demodulator in order to recover the transmitted information [2]. This instrument of AMC could play a key role in various civilian and military applications. For civilian purposes, AMC can be used as interference identification, spectrum management and signal confirmation. Other applications are for military purposes such as emitter interception, electronic surveillance systems and electronic warfare.

There are two general classes of AMC algorithms, likelihood-based (LB) method [3]-[14] and feature-based (FB) method [15]-[24], respectively. The first method is optimal in the Bayesian Sense. It made the decision by comparing the likelihood ratio against a threshold. In this way, the probability of false recognition is minimized. But, this approach needs to know much prior knowledge, and suffers from computational complexity. The latter method made the decision by the extracted features' observed values. Although this method may not be optimal, it is usually simple to implement, with near-optimal performance, when designed properly [25]. The general process of signal modulation recognition which based on extracted features can be shown in Fig. 1. In the online phase, the properly designed classifier is responsible for classifying the incoming signals based on the features extracted [26].

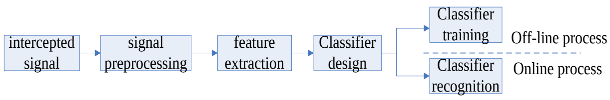

Figure 1. Process of AMC base d on pattern recognition

The techniques used in modulation recognizer are binary decision tree, Fuzzy clustering, K-nearest neighborhood, support vector machine, artificial neural network, etc. Among these techniques, artificial neural network (ANN) has been widely used in signal modulation classifier for its properties and capabilities of nonlinearity, adaptively and fault tolerance [27]. Presently the feed-forward neural network which includes multi-layer perceptron (MLP) neural network and radial basis function (RBF) neural network has widely been used in automatic modulation classifier. One particular structure called back propagation neural network is the most popular neural network architecture, which uses supervised learning to determine a complex input-output mapping in the field of AMC. A major drawback of the commonly used gradient-based BP algorithm, which is a local search method, is its easy entrapment at a local optimum point, especially for those non-linearly separable pattern classification problems or complex function approximation problem, so that BP may lead to failure in finding a global optimal solution [28]. Another drawback is that the convergence speed of the BP algorithm is too slow, and further, the convergence behavior of the BP algorithm depends very much on the choices of initial values of the network connection weights.

So, How to find a global optimization algorithm and a set of proper weights for BPNN is still a difficult problem in the designing of modulation recognizer. Due to quick convergence speed and better global exploration capability, PSO algorithm is applied to optimize the weights of the neural network so as to overcome the

Manuscript received April 20, 2013; revised May 26, 2013.
This work was supported by the National Defense Pre-Research Project under Grant No. 62201070301.
Corresponding author email: whtimesvip@sohu.com.
doi:10.12720/jcm.8.5.322-329

flaws of BP algorithm [29]. This paper outlines the performance of a new modulation recognizer with a variety of modulation types, e.g. 2ASK, 4ASK, BPSK, QPSK, 2FSK, 4FSK, OFDM, which can be recognized.

This paper is organized as follows: Section II states the signal model and feature extraction in our experiment. Section III briefly compares the multi-layer perceptron and the radial basis function network using in the automatic modulation classifier. Section IV presents the method of classifier using PSO-BP algorithm. Some simulation results are shown in experimental result section. Finally, conclusions and future plans are presented in section VI.

II. Feature Extraction

A. Signal Model

In this paper, we take seven main schemes of digital signals into consideration, i.e., 2ASK, 4ASK, BPSK, QPSK, 2FSK, 4FSK, OFDM. The signal model that has been used in this paper was defined as in Eq. 1:

$$s(t) = x(t) + n(t) \tag{1}$$

where $n(t)$ is the additive white Gaussian noise. $x(t)$ represents the modulation type. The following models have been used for $x(t)$ [1]:

$$x_{ASK}(t) = A\mathrm{Re}\left[\sum_k A_k e^{j2\pi f_c t} g(t - kT_s)\right]$$
$$(A_k = 2i - M - 1;\ i = 0,1,\ldots,M-1) \tag{2}$$

$$x_{PSK}(t) = A\mathrm{Re}\left[\sum_k C_k e^{j2\pi f_c} g(t - KT_s)\right]$$
$$(C_k = e^{j\frac{2\pi i}{M}};\ i = 0,1,\ldots,M-1) \tag{3}$$

$$x_{FSK}(t) = A\mathrm{Re}\left[\sum_k e^{j2\pi(f_c + \Delta f_k)t} g(t - kT_s)\right]$$
$$(\Delta f_k = \left[i - \left(\frac{M-1}{2}\right)\right]\Delta f;\ i = 0,1,\ldots,M-1) \tag{4}$$

$$x_{OFDM}(t) = A\mathrm{Re}\left[\sum_k \sum_{n=0}^{N_p - 1} C_{n,k} e^{j2\pi n\Delta f t}\right]$$
$$C_{n,k} \in C;\ E\{C_{n,k}\} = 0 \tag{5}$$

The meanings of these symbols such as A, A_k, C_k, $C_{n,k}, \Delta f_k, T_s, \Delta f, f_c, N_p, M, g(t), E\{g\}$ and C will not described in detail.

B. Feature extraction

In order to discriminate above digital signal schemes, six extracted features can be chosen. Five key features that were used in [30] are also used in this paper. These features are derived from the instantaneous amplitude, phase and frequency of the intercepted signal. The other feature is referred to [31]. These features as the inputs of the classifier are:

1) γ_{max}, which is the maximum value of the spectral power density of the normalized-centered instantaneous amplitude. γ_{max} is used to discriminate between the signals that have amplitude information and those that have no amplitude information.

2) σ_{ap}, which is the standard deviation of the absolute value of the centered non-linear component of the instantaneous phase. This feature can be used to discriminate between the types that have absolute phase information and those that have no absolute phase information.

3) σ_{dp}, which is the standard deviation of the direct value of the centered non-linear component of the instantaneous phase. This feature can be used to discriminate between the types that have no direct phase information and those that have no direct phase information.

4) σ_{aa}, which is the standard deviation of the absolute value of the normalized-centered instantaneous amplitude. This feature is used to discriminate between 2ASK and 4ASK.

5) σ_{af}, which is the standard deviation of the absolute value of the normalized-centered instantaneous frequency. σ_{af} is used to discriminate between 2FSK and 4FSK.

6) K_{20}, which is the second-order moment and fourth moments value combination can be used to distinguish between QPSK and OFDM.

According to these kinds of features, seven digital signals can be identified by the following decision tree as in Fig.2.

According to the analysis of above features, Azzouz set the appropriate thresholds for different features and achieved the classification of signal modulation by the decision tree. How to set the suitable threshold is the difficulty of the method. Because the threshold is a fixed value, with the signal-to-noise ratio decreasing, the performance of this method will become worse. This paper take six features as the inputs of the neural network. After training the BP neural network by PSO, this method can identify seven modulation signals effectively.

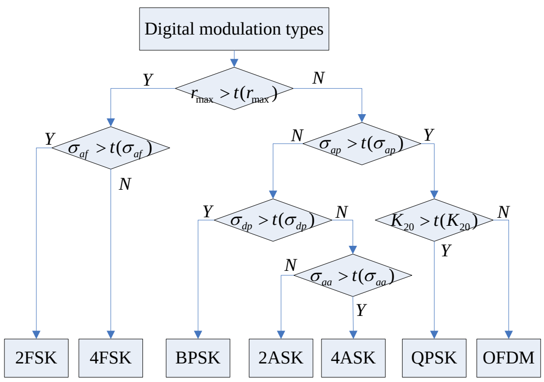

Figure 2. Flowchart for seven digital types recognition

III. Compared with Two Feed-Forward Neural Network Identifier

The feed-forward network which includes the Multi-layer perceptron and the radial basis function network is one of popular family of ANN. As two widely used feed-forward neural networks, BRF and MLP are often used as the design of automatic modulation classifier.

A. Automatic Modulation Classifier based on RBF

RBF neural network has simple network structure, fast learning rate and good generalization ability. So it has been extensively applied in the field of data classification, system modeling, control and fault diagnosis. RBF network is a three-layer feed-forward network. The transformation from the input layer to the hidden layer is non-linear, whereas the transformation from the hidden layer to the output layer is linear. The transfer function of the hidden layer unit using radial basis function is given in Eq. 6. The output layer value descried in Eq. 7:

$$\phi_i(x) = \exp\left[-\frac{\|X - C_i\|}{2\rho_i^2}\right] \tag{6}$$

$$y_i = \sum_{i=1}^{h} w_{ij}\phi_i(x) \tag{7}$$

In Eq. 6, $X = (x_1, x_2, \ldots, x_M)^T$ is input sample. $\phi_i(x)$ is the output of the i-th hidden layer. C_i is the center of the basis function. ρ_i represents the width. In Eq. 7, w_{ij} is the weight between hidden layer node *j* to the output node *i*. The training process of this neural network is divided into two stages. First, the input samples decide the center value and width of the Gaussian basis functions in each hidden layer node. After the parameters of the hidden layer determined, the weights between the hidden layer and output layer can be calculated by vary methods.

In the existing literatures, Ye Jian et al. [32] adopted fuzzy C-means clustering algorithm to obtain the parameters of radial basis function, while weights of the network are trained with gradient descent approach. Chen Jie et al. [33] introduced orthogonal least squares (OLS) radial basis function neural network which identified seven digital modulation types. Liu shu et al. [34] proposed an algorithm with which to implement the optimal classifier of RBF neural networks with genetic algorithm to classify the modulation types of CW, 2ASK, 2FSK, 2PSK and QAM.

The center value, the number of hidden layer nodes and width of the RBF network layer have great influences on the performance of the network. Several commonly used RBF network training algorithms are difficult to find the center and the width of the global optimum. And these training algorithms can not automatically determine the hidden layer nodes. These defects limit the RBF recognition of signal modulation type.

B. Automatic Modulation Classifier based on MLP

MLP neural networks compose of three parts: input layer, one or more hidden layer and output layer. Inputs are propagated through the network layer by layer and MLP gives an non-linear mapping of the inputs at the output layers [26]. This algorithm turns a set of input and output samples into a nonlinear optimization mapping problem, using gradient descent method to achieve the minimum mean square error (MSE) between the actual output and the desired output of the network. Azzouz presented two BPNN algorithms for digital modulation recognition. At the SNR of 10dB, the correct decision rate was more than 94.5% for the single hidden layer, while all the digital modulation types have been correctly classified with more than 96.% except 2FSK(=92.5%) for the double hidden layer. In order to avoid the problems of BPNN, M.L.D. Wong and A.K. Nandi [26] proposed a resilient back propagation algorithm (RPROP) MLP recognizer and compared with BP with momentum and adaptive learning rate MLP recognizer. Fu Wei-Hong et al. [35] proposed a novel scaled conjugate gratitude (SCG) BPNN classifier in order to accelerate the convergence speed. In [36], Li Jun-jun et al. imposed Levenberg-Marquardt algorithm to take the place of BP algorithm, then got faster convergence rate. In [37], Liu shu et al. enhanced a new BP recognizer which combined with genetic algorithms to overcome the drawbacks of back propagation algorithm. But the genetic algorithm consists of three essential operations: the selection operation, the crossover operation and mutation operation. These operations could tend to result in complicated calculation. In practice, the method which based on BP neural network training with genetic algorithm will cost a great deal of time to recognize the modulated signals. For this reason, we put forward a novel approach for the automatic modulation recognizer of communication signals based on PSO-BP neural network.

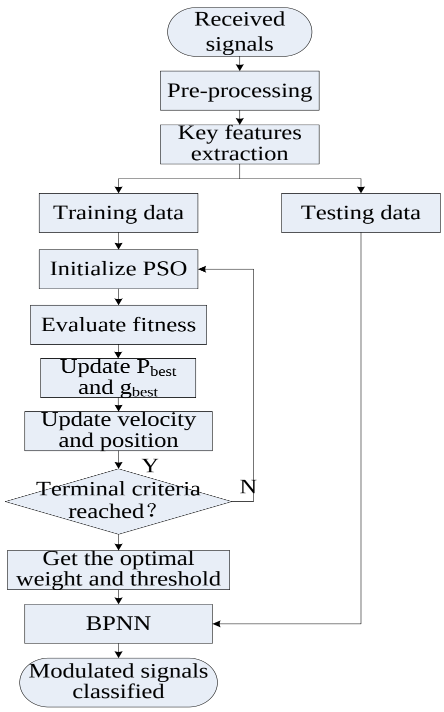

Figure 3. Procedures of modulation classification based on PSO-BP

IV. Classifier Based on Bio-Inspired Artificial Neural Network

To improve the performance of the original BP algorithm, two main factors have been taken into account: (1) selection of better energy function, (2) selection of dynamic learning rate and momentum. However, these improvements have not completely removed the disadvantages of the BP algorithm thoroughly. Because the PSO algorithm is able to find the global optimum with a large probability and high convergence rate, it is adopted to train the back propagation neural network (BPNN) in this study. The flow chart of AMC using BPNN based on PSO is shown in Fig. 3.

A. Structure of Back Propagation Neural Network

The concept of artificial neural network is inspired from the view of human brain as a processor using interconnected neurons. BPNN is one of most popular neural network which based on backward error influence. The topological architecture of BPNN is shown in Fig. 4. BP learning algorithm comprises two processes of data flow, that is, forward propagation for information and backward propagation for error. The input pattern x_k produces the actual output y_k. During this course the information propagated from the input layer to the output layer. The error resulting from the difference between the target outputs d_k and the actual y_k back-propagated from the output layer to the previous layers so as to update the weights. The weights adjusted until the errors between the target and the predicted outputs were small enough, or a pre-determined number of epochs reached. [38] and [39] proved that BPNN have the ability to approximate any function to any degree of desired accuracy. But BPNN's inability lies in the limitations of the back propagation training algorithm.

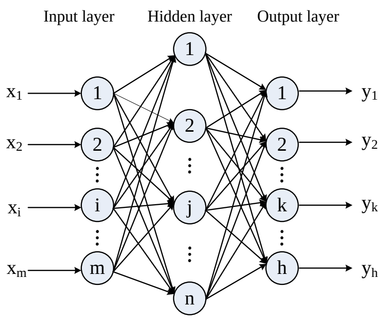

Figure 4. Topological architecture of the BPNN

B. The Standard Particle Swarm Optimization

Particle swarm optimization is a population based stochastic optimization technique proposed by Kennedy and Eberhart in 1995, which was inspired by the social behavior of animals, such as bird flocking and fish schooling [40]. The PSO system initially has a swarm of random solutions. The potential solutions, called particles, fly in a D-dimension search space with a velocity which is dynamically adjusted according to its own experience and that of its neighbors. The position of the i-th particle is expressed as $X_i = (x_{i1}, x_{i2}, \cdots, x_{iD})$; The velocity of the i-th particle is represented as $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$; Therefore the particle updates its velocity and position according to the following Eq. 8 and Eq. 9:

$$v_{id}^{k+1} = wv_{id}^{k} + c_1 r_1 \left(p_{id}^{k} - x_{id}^{k} \right) + c_2 r_2 \left(p_{gd}^{k} - x_{id}^{k} \right) \quad (8)$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \quad (9)$$

where k is the current step number, w is the inertia weight, c_2 and c_2 are the cognitive and social acceleration constants. r_2 and r_2 are random values between 0 and 1, x_{id}^{k} is the current position of the particle, p_{id} is the best one of the solution i-th particle has reached, p_{gd} is the best one of the solutions all the particles have reached. In addition, the velocity of the particle is limited to the range $[V_{\min}, V_{\max}]$.

In order to improve PSO algorithm searching capacity, linear inertial weight can be adopted. Inertia weight w was calculated a linearly decreasing equation with the iterative generations [41] in Eq.10:

$$w = w_{\max} - T \frac{w_{\max} - w_{\min}}{T_{\max}} \quad (10)$$

where T is the generation index representing the current number of evolutionary generations, and T_{max} is a pre-defined maximum number of generations. The maximal weights w_{max} and minimal weights w_{min} are usually set to 0.9 and 0.4, respectively.

C. Training the BPNN Classifier with PSO

1) The topology of BPNN

A single hidden layer BP feed-forward network was chosen as the recognizer. In Fig. 1, we set six nodes in the input layer and seven nodes in the output layer, while there are six typical features for inputs and seven classified digital signals for outputs. The number of nodes in the hidden layer could be computed out in Eq. 11:

$$n = \sqrt{m+h} + \alpha \qquad \alpha = 0, 1, \cdots 10 \quad (11)$$

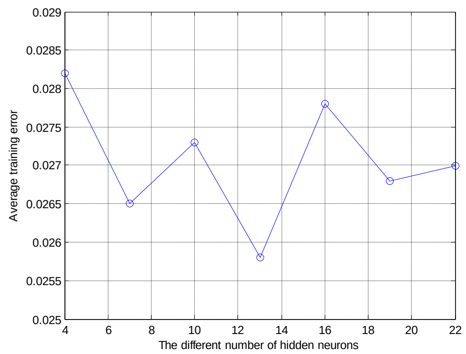

Figure 5. Average training error of different number of hidden neurons

The appropriate number of the neurons needs to be determined manually at this stage. Then we chose different numbers of neurons in the hidden layer to simulate the average training error. This is shown in Fig. 5. When the neurons set 13, the average training error is minimal. So, the BPNN with 13-node hidden layer is the optimum structure.

The activation function associated with the hidden layer is log-sigmoid function and activation function associated with the output layer is the linear function. Log-sigmoid function and liner function are defined in Eq. 12 and Eq. 13:

$$y = \frac{1}{1-\exp(-x)} \tag{12}$$

$$y = x \tag{13}$$

2) The parameters of PSO

To evolve the network weights and thresholds with PSO, each particle will be represented as a group of weights and thresholds. So the particle consists of $m \times n + n \times h + n + h$ real numbers. The particle position vectors X contain the weights and bases of BPNN. It can be denoted as in Eq. 14:

$$X = \left[w_{ij}^1, \theta_i^1, w_{ij}^2, \varphi_j^2 \right] \tag{14}$$

where w_{ij}^1 $\left(i = 1, \cdots, m; j = 1, \cdots, n\right)$ and w_{ij}^2 $(i = 1, 2, \cdots, n;$ $j = 1, 2, \cdots, h)$ represent the connection weight between the input layer and the hidden layer, and that between the hidden layer and the output layer, respectively. θ_i^1 $\left(i = 1, \cdots, n\right)$ and φ_j^2 $\left(j = 1, \cdots, h\right)$ are the threshold values of hidden layer and output layer.

PSO fitness function which measures the performance of individual is defined as in Eq. 15:

$$Fitness(k) = \frac{1}{nh} \sum_{j=1}^{n} \sum_{i=1}^{h} \left(T_i - O_i \right) \tag{15}$$

where $Fitness(k)$ is the value of particle k's fitness, T_i is the expected output, O_i is the calculated output of individual network, n is the number of training set examples, and h is the number of output nodes. Thus $Fitness(k)$ is the normalized mean squared error (MSE) of the individual k on the training set. When the maximum number of generation comes to a climax, the training process comes to an end.

3) Training the PSO-BPNN

The steps of PSO-BPNN algorithm are as follows:

Step 1: Initialize population size, each particle's position and velocity randomly. Set the values of c_1, c_2, w_{max}, w_{min} and T_{max} . Fix the structure of BPNN.

Step 2: According to each particles position vectors which stand for the weight and threshold values calculate the actual output of BPNN. Evaluate current fitness in formula (15) for each particles position.

Step 3: For each particle, compare fitness and personal best of current agent. If fitness is better, then update the personal best.

Step 4: Compare each particles fitness and globe best. If fitness is better, then update the globe best.

Step 5: Update position, velocity and inertia weight of particles using formula (8), (9) and (10).

Step 6: If the maximum number of generation reaches, go to step 2, otherwise, the training ends. Then output a set of optimal weights and thresholds for BPNN.

V. Experimental Result

The PSO-BPNN classifier was trained using 600 samples from each of seven modulation schemes. These samples are divided into two datasets, 400 samples for training while the other 200 samples for testing. The carrier frequency, symbol rate and sampling frequency are respectively 10 kHz, 1.2kbit/s and 40 kHz. The noise environment is AWGN, and the SNR range is chosen from 0 to 20dB with 5dB interval. Let the particle population size be 40, the dimension of the particle be 189, the acceleration constants c_1 and c_2 be 2, and a linearly inertia weight starting at 0.9, ending at 0.4. The structure of the BPNN was selected as 6-13-7. The output neuron with the largest output is deemed to be the "wining neuron", that is, the detected modulation type. The whole process of AMC simulation are presented in Fig. 6.

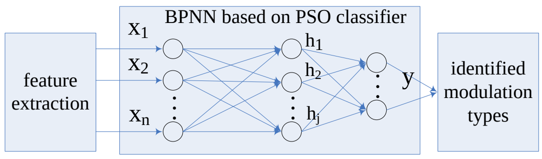

Figure 6. The whole process of AMC

TABLE I: MODULATION CLASSIFICATION ON SIMULATED SIGNALS WITH SNR=10 DB

Predicted Type	Actual Modulation Type						
	ASK2	ASK4	BPSK	QPSK	FSK2	FSK4	OFDM
ASK2	193	2					
ASK4	7	198					
BPSK			200	3			
QPSK				196			
FSK2				1	197		
FSK4					3	200	
OFDM							200

The details of classification of seven digital signals at 10dB were given in Table I. From this table, ASK2 and ASK4 is easy to confuse. This mainly results from the small difference of σ_{aa} between ASK2 and ASK4. In addition, the QPSK is easy to misjudge as BPSK and 2FSK.

Simulated each modulated signal at SNR, respectively, for 0dB, 5dB, 10dB, 15dB and 20dB, the recognition rate of PSO-BP classifier is shown in Table II. The results were the statistical means of 20 tests.

TABLE II: MODULATION CLASSIFICATION RATE ON SIMULATED SIGNALS WITH SNR=0, 5, 10, 15, 20 DB

SNR dB	**Predicted Modulation Type Correct Rate%**						
	ASK2	ASK4	BPSK	QPSK	FSK2	FSK4	OFDM
SNR=0	92	96.5	100	94.5	95	99	100
SNR=5	94.5	97.5	100	95.5	96.5	99.5	100
SNR=10	96.5	99	100	98	98.5	100	100
SNR=15	100	99.5	100	98.5	100	100	100
SNR=20	100	100	100	99.5	100	100	100

From Table II, through the above six kinds of features, PSO-BP classifier performed better in the recognition of the modulation signal. It can be seen that at 0dB the detection is still above 92%. Thus, in the case of low signal-to-noise ratio this algorithm can also obtain a high recognition rate.

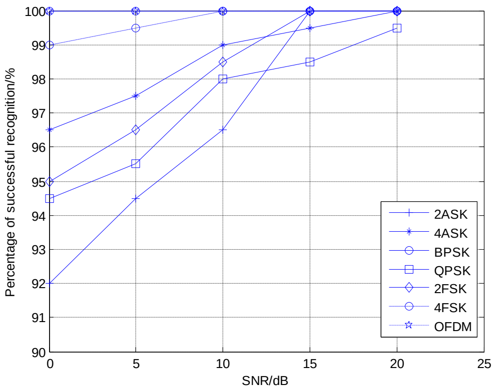

Figure 7. Percentage of successful recognition with PSO-BPNN classifier

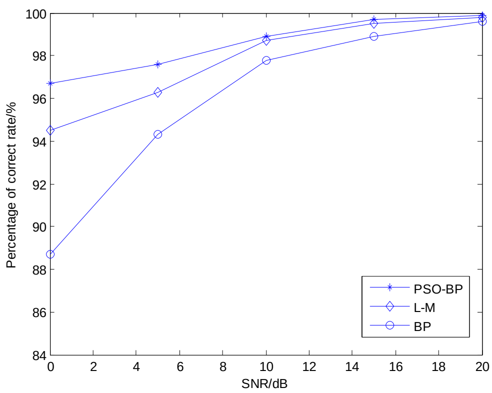

Figure 8. Recognition rate of PSO-BP, M-L and BP classifier

Fig. 7 shows the correct recognition rate of PSO-BP classification at SNR of 0, 5, 10, 15 and 20dB. This figure can visually see that the successful recognition rate has been improved accordingly, as SNR increasing.

Lenvenberg-Marquerdt (L-M) algorithm has fewer iterations and faster convergence. When the number of weights and thresholds is small, this algorithm can get better applied in practice. The recognition performance of PSO-BP classifier, L-M classifier and standard BP classifier were compared in Fig. 8. It is seen that the successful recognition rate of PSO-BP classifier is higher than M-L classifier and BP classifier in the same case of signal schemes and features extraction. In the conditions of low SNR, PSO-BP classifier can get higher recognition rate. Therefore, this algorithm has powerful advantage and stability.

This paper proposed distributed computing program to train the neural network based on PSO-BP algorithm and L-M algorithm. And compared the performance of this two training algorithm when achieving the same neural network error. The SNR is in the range of 0-20dB with steps of 5dB. From Table III, the recognition performance of PSO-BP classifier is slightly higher than the L-M classifier. But the L-M algorithm has more training times. So this PSO-BP classifier has better application prospect.

TABLE III: COMPARED CORRECT RATE AND ITERATIONS OF PSO-BP WITH L-M CLASSIFIER

SNR d/B	**L-M**		**PSO-BP**	
	Correct rate (%)	Iteration	Correct rate (%)	Iteration
SNR=0	94.5	436	96.7	384
SNR=5	96.3	452	97.6	384
SNR=10	98.7	475	98.9	384
SNR=15	99.5	475	99.7	384
SNR=20	99.8	475	99.9	384

VI. CONCLUSIONS

A new method for automatic classifier of the digital modulations has been presented in this paper. Making full use of the extensive mapping ability of the MLP neural network and the global convergence of the particle swarm optimization, this method overcomes the problems of basic BP algorithm by introducing new technique of PSO algorithm. The swarm intelligence method can decide the optimum weights and bases to minimize the error between the network output and the correct recognition. Extensive simulations and the experimental tests of these digital modulation types are carried out at different SNRs. Simulation results show that even at SNR of 0dB, this PSO-BP classifier has reached more than 92% recognition rate in the identification process of seven digital modulation types. This classifier can reduce calculation time and is simple to achieve.

Further work of the research will be directed to apply other new bio-inspired theory with BPNN to improve the capacity of AMC. Another interesting aspect would be to adjust the number of nodes in the hidden layer of BPNN dynamically.

REFERENCES

[1] D. Grimaldi, S. Rapuano, and L. De Vito, “An automatic digital modulation classifier for measurement on telecommunication networks,” *IEEE Trans. on Instrumentation and Measurement*, vol. 56, pp. 1711-1720, October 2007.

[2] L. De Vito, S. Rapuano, and M. Villanacci, “An improved method for the automatic digital modulation classification,” in *Proc. IEEE IMTC Conf.*, Canada, 2008, pp. 1441-1446.

[3] J. A. Sills, “Maximum-likelihood modulation classification for PSK/QAM,” in *Proc. IEEE MILCOM Conf.*, Atlantic, 1999, pp. 217-220.

[4] C. Y. Huan and A. Polydoros, “Likelihood methods for MPSK modulation classification,” *IEEE Trans. on Communications*, vol. 43, pp. 1493-1504, Apr. 1995.

[5] L. Hong and K. C. Ho, “Classification of BPSK and QPSK signals with unknown signal level using Bayes technique,” in *Proc. ISCAS Conf.*, 2003, pp. IV.1-IV.4.

[6] B. F. Beidas and C. L. Weber, “Higher-order correlation-based approach to modulation classification of digitally frequency-modulated signals,” *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 89-101, Jan 1995.

[7] A. E. El-Mahdy and N. M. Namazi, “Classification of multiple M-ray frequency-shift keying signals over a Rayleigh fading channel,” *IEEE Trans. on Communications*, vol. 50, pp. 967-974, Jan 2002.

[8] P. Panagiotou, A. Anastasoupoulos, and A. Polydoros, “Likelihood ratio tests for modulation classification,” in *Proc. IEEE MILCOM Conf.*, Los Angeles, 2000, pp. 670-674.

[9] L. Hong and K. C. Ho, “Modulation classification of BPSK and QPSK signals using a two element antenna array receiver,” in *Proc. IEEE MILCOM Conf.*, 2001, pp. 118-122.

[10] N. E. Lay and A. Polydoros, “Modulation classification of signals in unknown ISI environments,” in *Proc. IEEE MILCOM Conf.*, San Diego, 1995, pp. 170-174.

[11] S. Lesage, J. Y. Tourneret, and P. M. Djuric, “Classification of digital modulations by MCMC sampling,” in *Proc. IEEE ICASSP Conf.*, Salt Lake City, 2001, pp. 2553-2556.

[12] A. Abdi, O. A. Dobre, R. Choudhry, Y. Bar-Ness, and W. Su, “Modulation classification in fading channels using antenna arrays,” in *Proc. IEEE MILCOM Conf.*, 2004, pp. 211-217.

[13] Y. C. Lin and C. C. J. Kuo, “Classification of quadrature amplitude modulated (QAM) signal via sequential probability ratio test (SPRT),” *Signal Process*, vol. 60, pp. 263-280, Aug. 1997.

[14] Y. C. Lin and C. C. J. Kuo, “Sequential modulation classification of dependent samples,” in *Proc. ICASSP Conf.*, Atlantic, 1996, pp. 2690-2693.

[15] E. E. Azzouz and A. K. Nandi, “Procedure for automation recognition of analogue and digital modulations,” *IEEE Proceedings Communications*, vol. 143, pp. 259-266, October 1996.

[16] A. K. Nandi and E. E. Azzouz, “Modulation recognition using artificial neural networks,” *Signal Processing*, vol. 56, pp. 165-175, Jan 1997.

[17] M. L. D. Wong and A. K. Nandi, “Automatic digital modulation recognition using spectral and statistical features with multi-layer perceptrons,” in *Proc. 6th Int. Symp. Signal Processing and its Applications Conf.*, Kuala Lumpur, 2001, pp. 390-393.

[18] A. K. Nandi and E. E. Azzouz, “Algorithms for automatic recognition of communication signals,” *IEEE Trans. on Communications*, vol. 46, pp. 431-436, Apr. 1998.

[19] S. Z. Hsue and S. S. Soloman, “Automatic modulation classification using zero crossing,” *IEE Proceedings F (Radar and Signal Processing)*, vol. 137, pp. 459-464, Dec 1990.

[20] K. Farrell and R. Mammone, “Modulation classification using a neural tree network,” in *Proc. IEEE MILCOM Conf.*, Boston, 1993, pp. 1028-1032.

[21] K. C. Ho, W. Prokopiw, and Y. T. Chan, “Modulation identification of digital signals by the wavelet transform,” *IEE Proceedings Radar,Sonar and Navigation*, vol. 147, pp. 169-176, Aug 2000.

[22] L. Hong and K. C. Ho, “Identification of digital modulation types using the wavelet transform,” in *Proc. IEEE MILCOM Conf.*, Atlantic, 1999, pp. 427-431.

[23] A. Swami and B. M. Sadler, “Hierarchical digital modulation classification using cumulants,” *IEEE Trans. on Communications*, vol. 48, pp. 416-429, Mar 2000.

[24] Z. Zhijin and S. Junna, “A new method for modulation types recognition based on the time frequency representation,” in *Proc. 6th International Conference on Signal Processing*, 2002, pp. 26-30.

[25] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, “Survey of automatic modulation classification techniques: Classical approaches and new trends,” *IET Communications*, vol. 1, pp. 137-156, Apr 2007.

[26] M. L. D. Wong and A. K. Nandi, “Automatic digital modulation recognition using artificial neural network and genetic algorithm,” *Signal Processing*, vol. 84, pp. 351-365, Feb 2004.

[27] Simon Haykin, “Neural networks expand SP’s horizons,” *IEEE Signal Processing Magazine*, vol. 13, pp. 24-49, Mar 1996.

[28] J. Upendar, C. P. Gupta, G. K. Singh, and G. Ramakrishna, “PSO and ANN-based fault classification for protective relaying,” *IET Generation, Transmission & Distribution*, vol. 4, pp. 1197-1212, Oct 2010.

[29] Z. Youpeng and S. Hongsheng, “Turbo-Generator vibration fault diagnosis based on PSO-BP neural networks,” *Wseas Transations on Systems and Control*, vol. 5, pp. 37-47, Jan 2010.

[30] E. E. Azzouz and A. K. Nandi, “Automatic identification of digital modulation types” *Signal Processing*, vol. 47, pp. 55-69, Nov 1995.

[31] Chen Mei and Zhou Qi, “Research on automatic modulation recognition based on improved neural network,” *Journal of Chongqing University of Posts and Telecommunication*, vol. 21, pp. 764-770, 2009.

[32] J. Ye, L. D. Ge, and Y. X. Wu, “An application of improved RBF network in modulation recognition,” *Acta Automatica Sinica*, vol. 33, pp. 652-654, 2007.

[33] J. Chen and H. Chen, “Automatic digital modulation recognition based on OLS radial basis function neural network,” *Journal of Shanghai Jiaotong University*, vol. 38, pp. 26-29, 2004.

[34] S. Liu and H. Y. Wang, “Modulation recognition of communication signals based on RBF neural networks,” *Systems Engineering and Electronics*, vol. 30, pp. 1241-1244, 2008.

[35] W. H. Fu, X. N. Yang, X. W. Zeng, and N. A. Liu, “Novel method for blind recognition of communication signal based on time-frequency analysis and neural network,” *Signal Processing*, vol. 23, pp. 775-778, 2007.

[36] J. J. Lin, M. Q. Lu, and Z. M. Feng, “Modified automatic digital modulation recognition algorithm,” *Systems Engineering and Electronics*, vol. 27, pp. 2023-2024, 2005.

[37] S. Liu and H. Y. Wang, “Modulation recognition of communication signals using MLP neural networks trained with hybrid genetic algorithms,” *J. Wuhan Univ.(Nat. Sci. Ed.)*, vol. 54, pp. 104-108, 2008.

[38] K. I. Funahashi, “On the approximate realization of continuous mappings by neural networks,” *Neural Networks*, vol. 2, pp. 183–192, 1989.

[39] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feed-forward networks are universal approximators,” *Neural Networks*, vol. 2, pp. 395-366, 1989.

[40] J. Kennedy and R. Eberhart, “Partical swarm optimization,” in *Proc. IEEE International Conference on Neural Networks*, Perth, 1995, pp. 1942-1948.

[41] Y. Shi and R. Eberhart, "A modified particle swarm optimiser," in *Proc. IEEE International Conference on Evolutionary Computation*, Anchorage, 1998, pp. 69-73.

Jin Liu received the B.S. Degree from Xidian University, in 1984. He is currently a professor at No.722 Institution of China Ship Industry Corporation, Wuhan, China. His current research is concentrated on wireless digital communication, cognitive radio and military communications.

Li Cheng received the B.S. Degree and M.S. Degree in 2004 and 2008, respectively, from Wuhan University of technology and China Ship Research and Development Academy, China. She is currently working towards the Ph.D. Degree in signal processing for communications at Harbin Engineering University, China. Her research interests include digital signal processing, automatic classification of modulations and intelligence algorithms.