Original Research Article

A HYBRID METHOD OF BACKPROPAGATION AND PARTICLE SWARM OPTIMIZATION FOR ENHANCING ACCURATION PERFORMANCE

ABSTRACT

Aims: Backpropagation is an algorithm for adjusting the weight of neural networks in training stage. The performance of backpropagation has proven to be superior in optimizing the weight of neural networks, however this method has a weakness in the initiation stage where the random process creates local optimal opportunities for the performance results. Applying an algorithm based on global search is a choice to solve the drawback of backpropagation. One of the global search methods that have superior performance is particle swarm optimization. In this research, hybridization of backpropagation and particle swarm optimization (BP-PSO) has been carried out to overcome the problem of backpropagation
Study design: Research Papers and Short Notes.
Place and Duration of Study: Department of Informatic Faculty of Mathematics and Natural Sciences Udayana University between June 2022 and November 2022.
Methodology: The dataset used in this study is handwriting image dataset of mathematical symbol. There are 240 symbols which is consist of 180 images for training and 60 for testing. The robustness of the PSO method in obtaining the optimum global solution is expected to help backpropagation out of optimal local solutions. The application of PSO is carried out at the initial weight initialization stage of the artificial neural network. The tunning parameters of the artificial neural network are number of neurons in the hidden layer and the value of the learning rate. There are three different combinations in the number of neurons in the hidden layer, namely 10, 20, and 30. Meanwhile, the learning rate values are five different combinations, namely 0.1 to 0.9. Then for the minimum error value is 0.01 and the maximum number of epochs is 1000 epochs. In each test scenario, 5 repetitions will be carried out
Results: The performance results showed that PSO has succeeded in optimizing backpropagation where the accuracy of the BP-PSO is higher than BP without optimization. The accuracy of BP-PSO is 97.2% while the BP is 94.4%. The optimal learning rate value is 0.1 and the optimal number of hidden layers is obtained at 30 neurons.
Conclusion: The performance results showed that PSO has succeeded in optimizing backpropagation where the accuracy of the BP-PSO is higher than BP without optimization. The optimization process of weighting the artificial neural network as the initial weight for later retraining shows a higher average accuracy and the average number of epochs is less than the artificial neural network that does not perform the optimization of initial weight.

Keywords: neural network, backpropagation, particle swarm optimization, classification, swarm intelligent

1. INTRODUCTION

One of the methods in artificial intelligence that inspired by the biological workings of the human neural system is artificial neural network (ANN) method. ANN has been developed as a generalization of mathematical models from the human or neurological understanding. Just like humans whose brains always learn from the environment so that they can manage the environment properly based on the experiences that have been gained. ANN learning is performed iteratively as the network is presented with training examples, similar to the way we learn from experience [7]. ANN has superior performance when compared to other methods [1]. This method requires a training process in order to be able to recognize class data. The training process is a stage of renewing the weight of the neural network through a series of learning iterations where the results of this training are the best weight values in recognizing patterns. Researchers have made various methods to optimize the weight value of this artificial neural network, one method that has superior performance is backpropagation [2]. The performance of backpropagation has been proven to have superior performance in optimizing the weight of the neural network. although this algorithm is better than some others, it also has a weakness where the resulting results can be local optimal values [3]. This weakness is influenced by the initial weight which is randomly selected [4]. This is what in some cases, backpropagation will experience poor performance when compared to other method schemes. To solve this problem, we can insert a global search based algorithm into backpropagation to find optimal initiation values of backpropagation weight.

There are various methods based on the search for global solutions, including genetic algorithms and various methods in intelligent swarms. If we compare the performance of genetic algorithms and swarm algorithms such as ant colony, bee colony, memetic, shuffled frog leap, particle swarm optimization (PSO), the results show that the PSO algorithm has the most superior performance compared to these other algorithms [5]. In model integer programming problem, PSO has been compared with Genetic Algorithm. The comparison results show that the PSO algorithm is superior in terms of complexity, accuracy, iteration and program simplicity in finding the optimal solution [6]. Seeing the superior performance of this PSO, this research applies the PSO method to optimize the backpropagation method to get out of the optimal local solution. At the testing stage, the performance comparison between the backpropagation method and PSO optimization and the backpropagation method without PSO application was performed. This is done to measure the effectiveness of the PSO in optimizing the weight of the training results.

2. METHODOLOGY

2.1. Backpropagation

Backpropagation is an algorithm for training artificial neural networks to get optimal network weights. These set of weights determine the network's ability to recognize patterns. Backpropagation uses an error-output to fix the weights in the backward direction. In this study, a multi-layer backpropagation neural network scheme was used as shown in Figure 1 [11].

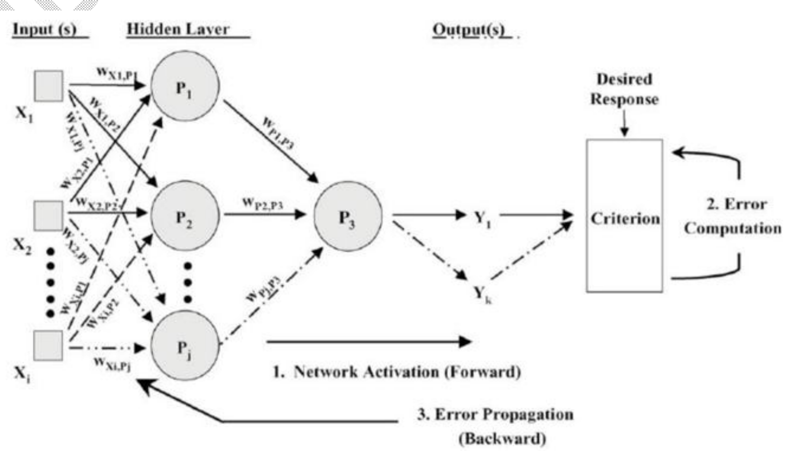

Fig 1. Back-propagation Network

The steps of Backpropagation algorithm are described bellow

1. Initialize the weights in the network with small random numbers
2. A pattern of data training is put into the network input layers.
3. Calculate all output unit in hidden layers and output layer

$$z\_net_j = v_{j0} + \sum_{i=1}^{n} x_i v_{ji} \quad (1)$$

$$z_j = f\left(z_{net_j}\right) = \frac{1}{1+e^{-z\_net_j}} \quad (2)$$

4. Calculate the unit error (δ) for every output layer and hidden layer. δ is the error unit that will be used in updating the weight. t_k is target output, Δw_{kj} is weight change, and α is learning rate.

$$\delta_k = (t_k - y_k) y_k (1 - y_k) \quad (3)$$

$$\Delta w_{kj} = \alpha \delta_k z_j \quad (4)$$

5. Modify all of the weight

$$w_{kj}(new) = w_{kj}(old) + \Delta w_{kj} \quad (5)$$

The next pattern from the training set is chosen and the step number 2 follows until all patterns processed. The algorithm ends when error is less then minimum criteria.

2.2. Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a population-based stochastic optimization algorithm motivated by intelligent collective behavior of some animals such as flocks of birds or schools of fish [13]. PSO as an optimization tool provides a population search-based procedure. PSO performs searching via a swarm of particles that updates from iteration to iteration. To seek the optimal solution, each particle moves in the direction to its previously best (pbest) position and the global best (gbest) position in the swarm [10]. The following are the stages of the particle swarm optimization algorithm:

1. Initialize particles with random speed and velocity.
2. Furthermore, the fitness value of each particle will be evaluated. Then the comparison process between particles and Pbest is carried out. If the particle fitness value is better than the Pbest value, the Pbest value will be converted into the particle fitness value.
3. comparison between the Gbest value and the best value for all particles. If the fitness value obtained is better than the Gbest value, the Gbest value will be converted into the particle fitness value.
4. Perform the process of changing the speed and rank of particles using formulas (6) and (7)
5. Repeat step 2 until it meets stopping criteria such as best fitness score or at maximum iteration

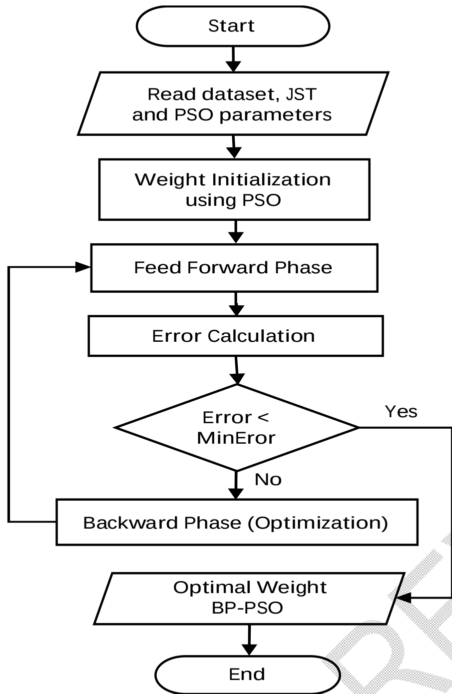

Figure 2. Hybridization Scheme

Some of the terms that are often used in PSO, namely:

- Swarm: population / herd of an algo-rithm.
- Particles: individuals in a swarm.
- Pbest (Personal best): the best rank of a particle.
- Gbest (Global best): the best position of the particle on a swarm.
- Velocity (v): a vector that determines the movement of particles as they move.
- Inertia weight (w): parameters controlling the speed of a particle.
- Learning speed (c_1 and c_2): constants for particles (c_1) and swarm (c_2).

In making changes to velocity, PSO has three parts, namely the social part, cognitive part and momentum part. These three parts are used in the velocity tracking process. The following is the PSO equation with the inertia weight [12].

$$v_{ij}(t+1) = w * v_{ij}(\mathrm{t}) + c_1 * \mathrm{rand}() * \left(p_{ij}(t) - x_{ij}(t)\right) + c_2 * rand() * \left(g_{ij} - x_{ij}(t)\right) \quad (6)$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(\quad (7)$$

Description :

v	=	Particel Velocity
x	=	Particel Position
w	=	Inertia Weight
c1	=	pBest Constanta
c2	=	gBest Constanta
p	=	*pBest*
g	=	*gBest*
rand()	=	Random Number

2.3. Hybridization Scheme

The major drawbacks of backpropagation learning algorithm is the problems of local minimal [8]. The robustness of the PSO method in obtaining the optimum global solution is expected to help backpropagation out of optimal local solutions. The application of PSO is carried out at the initial weight initialization stage of the artificial neural network. This can be seen from the results of previous studies where the optimal local cause of backpropagation is at the weight initialization stage which is carried out randomly [9].

3. RESULTS AND DISCUSSION

3.1. Hybridization

The dataset used in this study to measure the accuracy of PSO in optimizing backpropagation is a mathematical symbol image. There are 20 math symbols written by hand and using the image equation found in the Ms. Word Application. Each symbol has a number of data samples of 12 sample images. The total number of data amounted to 240 where the overall data was divided into 2, namely training data as much as 70% and test data as much as 30% of the entire dataset. Before this dataset enters the pattern recognition stage with the backpropagation and PSO methods, the data is first carried out in the preprocessing stage. The steps taken in the preprocessing stage are as follows:

1) Grayscaling: In this grayscaling process, the input image in the form of Color Image or RGB will be converted into grayscale or gray image. For each image pixel, the red, green, and blue component values will be taken. The method used for grayscaling is the weighted method where this method will produce grayscale values by adding up the values of 30% from red, 59% from green, and 11% from blue.
2) Binerization: In the binerization process, the grayscale image will be converted into a black and white / binary image using the thresholding method. This thresholding method uses a value to check whether the gray color of the image pixels will be converted to black or white. The threshold value used is 127 because it is the middle value of the value range 0 to 255.
3) Segmentation: This binaryization process uses a gray image where the color value of each pixel will be taken and will be compared with the threshold value, if the color value is greater than the threshold value (127) then the color value will be changed to 255 (white), otherwise the color value will be changed to 0 (black color). The result of this binary process is a binary image which only consists of white and black.
4) Normalization: The image resulting from the segmentation process has different sizes so that it is necessary to carry out the normalization process on the image size. In the normalization process, there are 2 stages, namely the padding and scaling stages. In the padding stage, the image that is not yet square (each side is the same) will be added to the sides to make it the same size. At the scaling stage, the size of the image will be changed to a size of 64x64 pixels. First, the value of the image height and width and the image to be normalized will be inserted. Then look for the largest value between the height and width of the image. If the height value is greater than the width, then the left and right values will be added according to the difference between the height and the width divided by 2. Whereas if the width value is greater than the height value, then the top and bottom values will be added according to the difference between width and height divided by 2. Then each the sides of the image will be added according to the top, bottom, left, and right values so that the image will now have the same size. This stage is called the padding stage. Furthermore, the image size will be converted to 64x64 pixels, this stage is called the scaling stage. The image produced from this normalization process is an image with a size of 64x64 pixels.

After these four stages have been completed, it is continued with the feature extraction stage. This is done to get the features and their values for each dataset used in the pattern recognition stage. The digital image from preprocessing data will be used at the feature extraction stage to get the features of the image. The method used to perform the feature extraction process is a histogram of gradient. This method is a feature extraction method based on the slope or direction of the pixels in relation to the surrounding pixels. The result of this feature extraction is a vector which is a characteristic of the image.

3.2. Training and Testing Result

The first result of system testing in this study is the best artificial neural network architecture in the case of recognizing mathematical symbols. The parameters of the artificial neural network used are the number of neurons in the hidden layer and the value of the learning rate. There are three different combinations in the number of neurons in the hidden layer, namely 10, 20, and 30. Meanwhile, the learning rate values are five different combinations, namely 0.1 to 0.9. Then for the minimum error value is 0.01 and the maximum number of epochs is 1000 epochs. In each different test scenario 5 tests will be carried out.
The second system test result is the optimization effect of the initial weight value of the neural network using the particle swarm optimization algorithm. The results of this study will compare the accuracy value of the backpropagation method neural network with a random initial weight (ANN BP) and the backpropagation neural network with the initial weight generated using PSO.

Table 1. Test Results with Hidden Neurons 10

Learning	Epoch		Accuration	
Rate	BP	BP PSO	BP	BP PSO
0,1	1000	996	84,4	88,9
0,3	398,2	342,4	87,2	88,9
0,5	295,4	218,4	86,1	90,0
0,7	329,2	150,6	87,8	91,7
0,9	146	117,4	84,4	92,8
0,91	321,4	116,2	88,9	88,9
0,93	319,4	121,2	83,3	88,9
0,95	295,4	218,4	86,1	90,0
0,97	125,4	107,2	86,1	84,4
0,99	121,6	104	83,3	88,9

In this testing process, the level of program accuracy in recognizing mathematical symbols will be tested. In this study, several changes will be made to the learning rate to determine the effect on the resulting accuracy value. The learning rate value will be used to find the program's accuracy in recognizing mathematical symbols. The value of the accuracy rate can be calculated using equation (8).

$$P(N) = \frac{IN}{N} * 100\% \quad (8)$$

Description :
P (N) = level of accuracy
IN = The number of successfully recognized data
N = Total number of data

Tables 1-3 are the test results with Hidden Neurons 10, 20, and 30. The results of the test scenario in graphical form can be seen in Figure 3-5. From the results of the scenario testing that has been made, the optimal results of the back propagation weight parameter are in the learning rate value. If made in graphical form, changes in the performance of the backpropagation neural network with the scenario of the number of hidden neurons 10 can be seen in Figure 3. The highest accuracy value of 92.8% is achieved when learning is 0.9

on BP PSO and the highest value is 88, 9% when learning rate is 0.91 on backpropagation without PSO algorithm optimization. From the graph, it can be seen that almost all of the tested learning rate values, BP PSO performance has always superior accuracy rates compared to the backpropagation algorithm without PSO optimization. This clearly shows the effectiveness of PSO in optimizing the backpropagation neural network.

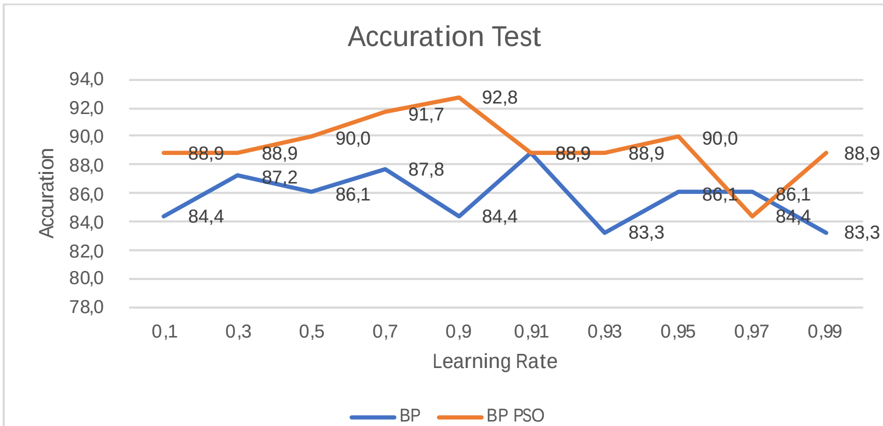

Fig 3. Graph of Hidden Neuron Test Results PSO = 10

Table 2. Test Results with Hidden Neurons 20

Learning Rate	Epoch BP	Epoch BP PSO	Accuration BP	Accuration BP PSO
0,1	564,4	560,8	93,3	95,0
0,3	202	187,2	93,3	95,6
0,5	159,6	116,2	91,7	93,9
0,7	86,2	82,8	91,1	95,6
0,9	70,2	64,6	94,4	96,1
0,91	70,8	64,2	89,4	95,0
0,93	77,6	63,4	90,6	95,6
0,95	188	112,4	84,4	90,6
0,97	64,2	61,6	91,1	94,4
0,99	259,6	59,2	92,2	94,4

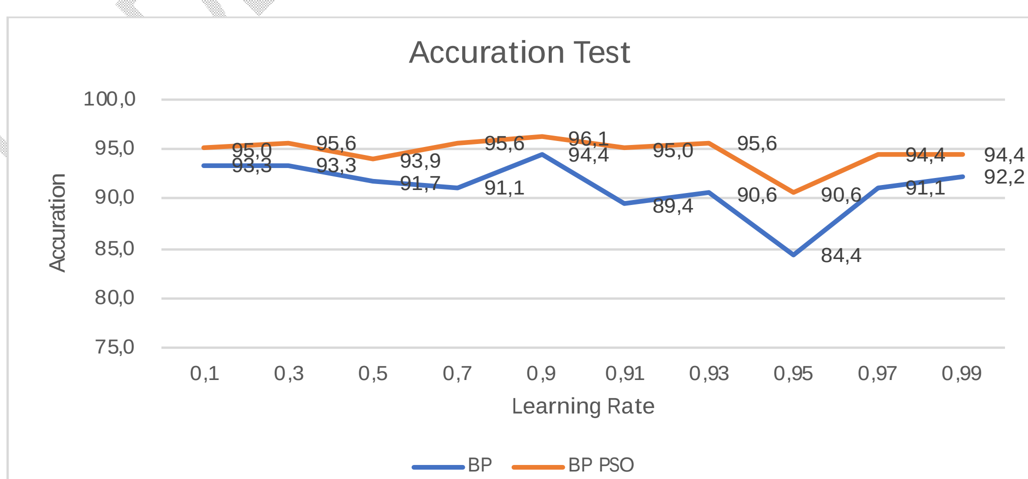

Fig 4. Graph of Hidden Neuron Test Results PSO = 20

Table 3. Test Results with Hidden Neurons 30

Learning	Epoch		Accuration	
Rate	BP	BP PSO	BP	BP PSO
0,01	1000	1000	78,3	87,2
0,03	70,8	64,2	89,4	95,0
0,05	920,2	895,6	91,7	94,4
0,07	618,8	623,8	92,2	95,6
0,09	596,4	487,8	92,8	96,7
0,1	579,8	444	91,7	97,2
0,3	161,8	150,6	93,3	96,1
0,5	96,2	91,6	92,2	96,1
0,7	103,2	65,8	92,2	96,1
0,9	54,4	52,2	92,8	95,6

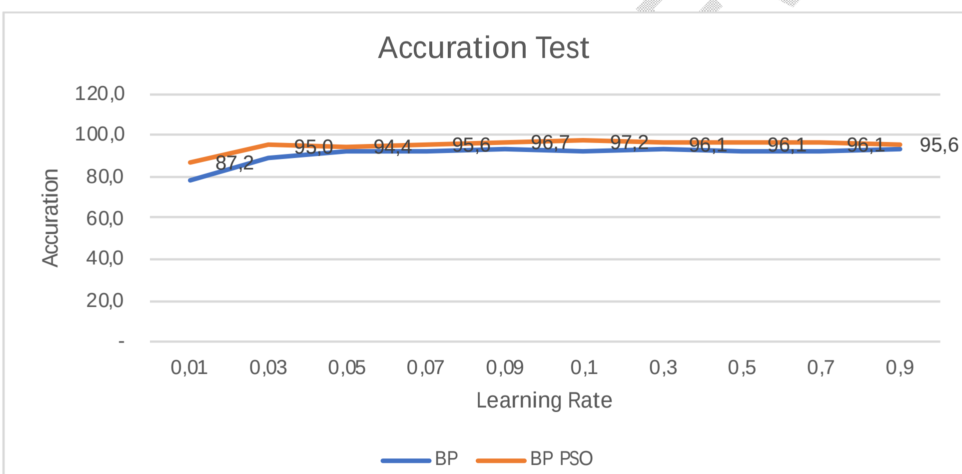

Fig 5. Graph of Hidden Neuron Test Results PSO = 30

4. CONCLUSION

From the results of trials and evaluations of the research that has been done, the following conclusions can be drawn.

1. The test scenario with the highest average accuracy is obtained by using a learning rate of 0.1 and the number of neurons in hidden layer 30 with an average accuracy of 94.4% in artificial neural networks without optimization and by using initial weight optimization using particle swarm optimization, it reached 97.2 % accuracy.
2. Learning rate has an effect on neural network training where a learning rate of 0.1 will result in a better average accuracy but a greater number of epochs and in the other hand a learning rate of 0.9 will result in a poor average accuracy but a smaller number of epoch in a scenario using the same number of hidden layer neurons.

From the testing process that has been carried out, it shows that the optimization process of weighting the artificial neural network as the initial weight for later retraining shows a higher average accuracy and the average number of epochs is less than the artificial neural network that does not perform the optimization of initial weight.

REFERENCES

Mahdiyah, U., Irawan, M.I., Imah, EM. Study Comparison Backpropagation, Support Vector Machine, and Extreme Learning Machine For Bioinformatics Data. *Journal of Computer Science and Information*. 2015;8(1): 53-59.

Amalia, O. A., Putri, A. H. M. Comparison of backpropagation artificial neural network and SARIMA in predicting the number of railway passengers. *Journal of Physics: Conference Series*. 2020;1663(2020):1-14.

Du, W. Research on Evaluation of Enterprise Performance Based on BP Neural Network Improved by Levenberg-Marquardt Algorithm. *Proceedings of the 2015 International Conference on Automation, Mechanical Control and Computational Engineering*. 24-26 April 2015, Jinan, Shandong. 2015;169-173.

Lee, A., Geem, Z.W., Suh, K.D. Determination of Optimal Initial Weights of An Artificial Neural Network by Using Harmony Search Algorithm. *Applied Sciences*. 2016;6(6):1-17.

Garg, A., Juneja, D. A Comparison and Analysis of various extended Techniques of Query Optimization. *International Journal of Advancements in Technology*. 2012;3(3):184-194.

Wihartiko, F.D., Wijayanti, H., Virgantari, F. Performance comparison of genetic algorithms and particle swarm optimization for model integer programming bus timetabling problem. *IOP Conference Series: Materials Science and Engineering*. 12 October 2017, Tangerang, Indonesia. 2017;1-6.

Basheera, I.A., Hajmeer, M. Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods. 2000;43(1):3-31.

Burse, K., Manoria, M., Kirar, V. P. S. Improved Back Propagation Algorithm to Avoid Local Minima in Multiplicative Neuron Model. International Journal of Electrical and Computer Engineering. 2010;4(12):1776-1779.

Chang, Y.T., Lin, J., Shieh, J.S., Abbod, M.F. Optimization the Initial Weights of Artificial Neural Networks via Genetic Algorithm Applied to Hip Bone Fracture Prediction. *Hindawi Publishing Corporation Advances in Fuzzy Systems* 2012:;1-9.

Zhang, Y., Balochian, S., Agarwal, P. Bhatnagar, V., Housheya, O. J. Artificial intelligence and its applications. *Mathematical Problems in Engineering*. 2016;1-12.

Kaur, J., Kaur, P. A Review: Artificial Neural Network. *International Journal of Current Engineering and Technology*. 2018;8(4):904-910.

Kong, F., Jiang J., Huang, Y. An Adaptive Multi-Swarm Competition Particle Swarm Optimizer for Large-Scale Optimization. *Mathematics*. 2019;7(521):1-13.

Wang, D., Tan, D., Liu, L. Particle swarm optimization algorithm: an overview. *Springer-Verlag Berlin Heidelberg*. (2018);22:387-408