A new model based on improved ACA and BP to predict Silicon content in hot metal

Li Wang

Dong-qing Wang, Jia-jun Zhu, Xiang Zhao

Key Laboratory for Advanced Control of Iron and Steel Process (Ministry of Education)
School of Information Engineering
University of Science and Technology Beijing
Beijing, 100083, China
e-mail: wl3927@126.com

Key Laboratory for Advanced Control of Iron and Steel Process (Ministry of Education)
School of Information Engineering
University of Science and Technology Beijing
Beijing, 100083, China
e-mail: wangdongqing1984@163.com, zjy711@163.com
jazzkiss@gmail.com

Abstract—A new model based on improved ant colony algorithm (ACA) and Back-propagation (BP) is proposed to predict Silicon content of hot metal in blast furnace. BP algorithm has been widely used in training artificial neural network (ANN), which is an outstanding model to predict Silicon content. BP algorithm has many attractive features, such as adaptive learning, self-organism, and fault tolerant ability. All of them make BP one of the most successful algorithms in various fields. But BP suffers from relatively slow convergence speed, extensive computations and possible divergence for certain conditions. As a new bionic algorithm, the improved ACA has gained very good performance in solving traveling salesman problem (TSP) and other optimization problems. Its properties such as distributed computation, heuristic searching and robustness have well conquered the long convergence speed and premature problem, which are the main deficiencies of BP algorithm. Experiments show the model proposed has good performance in predicting Silicon content of hot metal.

Keywords-component; Silicon content prediction; BP; ACA; blast furnace

I. INTRODUCTION

Iron is one of the most important raw materials in the modern society whose production and quality have been the most influential remarks of economic power of a country. As the main upper procedure of metallurgical industry, blast furnace (BF) iron making plays a significant role in energy saving and technical development of the whole industry.

Maintaining the appropriate temperature in BF is the key to guarantee a good metallurgic performance. Unfortunately, due to highly complex characteristics of BF iron making such as nonlinearity, time lag, high dimension, big noise and distribution parameter etc, it is almost impossible to measure the exact temperature of hot metal in BF yet. As an index to evaluate the status of iron making process and quality of hot metal, Silicon content in hot metal also serves as the most important parameter to reflect temperature of hot metal in BF. Thus, the study of Silicon content of hot metal prediction models has great significance.

Recently, ANN has been widely used in many fields. ANN is designed to simulate human brain’s powerful information memorization and analysis ability to provide a solution to many tasks which are very difficult to solve by traditional computational methods. As a crucial factor to ANN’s performance, parameters (weights and thresholds) optimization has been studied for a long time. Among all the training algorithms, BP is one of the most famous methods in parameter determination. This algorithm has many attractive features, such as adaptive learning, self-organism, and fault tolerant ability which make the BP one of the most successful algorithms in various fields. Unfortunately, its properties of extensive computations and slow convergence speed often make it hard to gain desirable performance.

ACA was first proposed by M.Dorigo. Research suggests that ants can always find a shortest way between a food source and their nest. This interesting behavior inspired them to design a method to problem optimization. In the past decades, ACA has been an outstanding solution to TSP. It also provide a promising solution in many other problems such as vehicle routing problem (VRP), sequential ordering problem (SOP), graph coloring problem (GCP) and so on.

The main advantages of ACA are distributed computation, positive feedback and heuristic. Distributed computation can effectively prevents premature convergence; positive feedback guarantees a fast discovery of best solutions; heuristic enables ACA to find some potential good solutions in the early phase of the search process. These characteristics have well fixed the problems that exist in BP network. Therefore, the model proposed which combines BP with the improved ACA provides a promising approach to predict silicon content in hot metal.
II. METHODS DESCRIPTION

A. BP network model

BP network model has some connective layers which consist of many neurons. Each neuron in a former layer can be connected to any of the other neurons in a next layer. There is no feedback loops in the network.

Given X as input, Y as output, \{W_{ih}\} as input-hidden layer link weights, \{T_{ih}\} as hidden layer neuron thresholds, \{W_{oh}\} as hidden-output layer link weights, \{T_{om}\} as output layer neurons threshold.

Then, \[y_i(k) = f(\sum_{j=1}^{m}w_{ij} f(\sum_{i=1}^{n}x_i(k)w_{ih} + T_{ih}) + T_{om})\]

Where \(H\) is the hidden layer neuron number; and

\[f = \frac{1}{1+\exp(-x)}\]

Given training sample matrix \(A = \{(x_i, y_i^*)|i=1, 2, \cdots d\}\) (\(x_i\) is the training sample input).

Root square error (\(E\)) is applied as the performance measurement for BP network training.

\[E = \frac{1}{d} \sum_{i=1}^{d} (y_i - y_i^*)^2\]

Where \(d\) is number of training samples; \(y_i\) is estimated output value; \(y_i^*\) is expected output value.

In the network training phase, weights and thresholds are adjusted to minimize \(E\). BP algorithm is the most commonly used ANN. The method proposed combines ACA with BP algorithm to train the neural network to gain desired performance.

B. Ant Colony Algorithm

According to the research, even if a single ant’s capability is limited and simple, the whole colony is highly structured. They cooperate with other members to accomplish difficult tasks. Among their behaviors, the most attractive one is the foraging behavior.

Ants can always find a shortest path between the food source and nest after a short period, even if an obstacle appears.

Ants use chemical pheromone to communicate with others. When ant meets a crossing, there are several paths to choose. If no one has selected these paths before, the ant will choose a random path, and then lays some pheromone (in varying quantity) on the ground. The shorter the path is, the more the pheromone there will be. Thus a pheromone trail is formed. When a following ant arrives, it detects and calculates the pheromone quantities of each path, and then chooses the path which has the most quantity. This new ant also lays pheromone on the ground. In this way, a positive feedback mechanism is provided. So, the pheromone on the shortest path grows faster, therefore more and more ants tend to choose the shortest path. The result is all ants will choose the shortest way quickly.

By simulating ants’ behavior, researchers proposed ACA, which has gained very good performance in solving TSP. However, ACA also suffers from the long convergence speed and premature problem. Many improved methods are proposed to solve these problems, in which Max-Min Ant System (MMAS) is the most famous and efficient one. MMAS provides pheromone control mechanism when updating pheromone, which can provide premature in searching, and also can decrease convergence speed greatly.

III. MODEL STRUCTURE

A. BP and improved ACA method

Given \(N\) parameters to be optimized (including all weights and thresholds of the BP neural network), every parameter has \(q\) effective digitals, each number’s value varies from 0 to 9, a number list with \(q\) elements was provide to represents these digitals.

To count the value of parameter: First, calculate the value corresponding to the numbers, then plus the amplificatory multiply \((AM, \text{required from a previous experiment using BP algorithm to train the network})\).

Given \(q=3\), number list=[6 8 9 5 1 2], \(AM = -0.0001\), then

\[\upsilon=(6 \times 1+8 \times 10+9 \times 100+5 \times 1000+1 \times 10000+2 \times 100000) \times (-0.0001)=-21.5986\]

For all parameters, each ant selects a value for every \(q\) effective digitals, save the record into a list, thus each ant represents a group of parameters and a neural network. Find out which ant has the best performance (means this network has smallest error with the training samples), record this ant’s choice as the final parameters for the BP neural network. In proposed the method, an effective digital for each parameter is represented as a “knot”. In this way, an ant’s record list can be mapped to an \(X-Y\) plane.

It takes the same long time for every ant to jump from a knot to another, and complete the selected task. Therefore, all ants will end their tourist at the same time. When an ant is jumping from knot \(A\) to knot \(B\), it will choose the digital from 10 values with a possibility given by function:

Then choose the digital using Roulette Wheel Select.

\[p_i(t) = \frac{[\tau_i(t)]^\alpha[\eta_i]^\beta}{\sum[\tau_i(t)]^\alpha[\eta_i]^\beta}\]

Parameters in formula (4):

\(\tau_i(t)\) : pheromone, quantity at \(i\)-th digital of knot \(B\)
\[ \eta_i; \text{ visibility, for } i-th \text{ digital.} \]
\[ \eta_i = \frac{10 - |y_i - y'_i|}{10} \quad (5) \]
\[ y_i; \text{the value of the } i-th \text{ digital} \]
\[ y'_i; \text{the expected value of } i-th \text{ digital (digital of the parameters trained by BP network)} \]
\[ \alpha; \text{the relative importance of trail, } \alpha \geq 0; \]
\[ \beta; \text{the relative importance of visibility, } \beta \geq 0; \]

When all ants finish their tours, update the pheromone quantity for every digital by this function:
\[ \tau_i(t) = \rho \tau_i(t) + \Delta \tau_i(t, t + n) \quad (6) \]
\[ \Delta \tau_i(t, t + n) = \sum_{k=1}^{N_c} \Delta \tau^k_i(t, t + n) \quad (7) \]

Parameters in formula (6):
\[ \rho; \text{trail persistence }, 0 \leq \rho \leq 1, (1- \rho) \text{ represents trail evaporation} \]
\[ \Delta \tau^k_i(t, t + n) = \begin{cases} \frac{Q}{E^k}, & \text{if } k-th \text{ ant choose } i-th \text{ digital} \\ 0, & \text{otherwise} \end{cases} \quad (8) \]

Parameters in formula (8):
\[ Q; \text{a constant, it is the quantity of pheromone in a whole tour laid by an ant.} \]
\[ E^k; \text{the performance measurement of network using parameters calculated by } k-th \text{ ant's choice.} \]

Pheromone control:
If \( \tau_i(t) \) is greater upper threshold of pheromone, then set it to the upper threshold, if it is less than lower threshold of pheromone, set it to lower threshold.

B. Algorithm description

Step 1: train the neural network using BP algorithm, record the value of each parameter. Calculate and memorize the expected value of each digital for all knots.
Step 2: initialize:
Set: time = 0, all ants positioned on the first knot of the tour. Set loop number \( N_c = 0 \), initialize pheromone matrix, clear pheromone increment matrix, clear record matrix (each ant corresponds to a row vector).
Step 3: training phase:
Set \( i = 1 \), calculate the possibility for each digital of the first knot, and then use Roulette Wheel Select to choose a digital. Save the value into the record matrix.
Step 4: \( i = i + 1 \), if \( i < n^s q \), turn to Step 3, otherwise turn to Step 5;
Step 5: since then all ants have finished their tours. Calculate all the parameter values represented by ants and create \( m \) networks, calculate errors of each network with training samples. Find out the minimal error, save these parameters as the best one in this loop.
Step 6: if \( N_c = N_{c_{max}} \), end the program, print the best path; otherwise, \( N_c = N_c + 1 \), \( i = 1 \), update pheromone matrix using the given function and pheromone control mechanism, clear record matrix, turn to Step 2.

Step 7: calculate the last loop’s choice as the final parameters.
Step 8: testing phase:
Test performance of the neural network established with parameters required in above processes.

IV. EXPERIMENT AND ANALYSIS

A. Experiment data
In this paper, all experiment data are collected in an iron factory. After analysis, 7 parameters (Air flow, oxygen-rich volume, air temperature, ventilation, top pressure, silicon content in sinter, coke load) are chosen to be the input of the model.

BP neural network structure is designed as follows:
3 layers: Input layer has 7 nerves, middle layer has 7 nerves, and output layer has only 1 nerve. Therefore, there are 64 parameters needed to be optimized.
In the training phase, 60 samples are adopted as training samples; in the testing phase, 20 samples are adopted as testing samples. Set ants number \( M = 42 \) (numbers of parameters to be optimized according to the network model), ants number \( M = 42 \), maximum loop number \( N_{c_{max}} = 120 \), \( Q = 100 \), \( \alpha = 1 \), \( \beta = 5 \), \( \rho = 0.5 \).

B. Experiment results
The performance comparison of using ACA and improved ACA training BP network is showed in Fig.2 and Fig.3.
Therefore, according to the experiment results, with higher accuracy and lower error, the BP optimal method based on improved ACA has gained very good performance.
Figure 3. error comparison

V. PARAMETERS IN ACA OPTIMIZATION

In this part, the actual data which are all collected from the blast furnace are the input of the system. Experiments are conducted to optimize parameters in ACA in order to obtain the best prediction performance. This paper studies several parameters which have greatest impact on the algorithm, namely trail persistence $\rho$, relative importance of visibility $\beta$, relative importance of trail $\alpha$, quantity of pheromone $Q$.

A. Trail persistence $\rho$

$\rho$ is trail persistence, whose range is $[0,1]$. Experimental results show that when $\rho$ is valued $[0.5, 1]$, these results are closer to the actual condition and the algorithm has faster convergence speed.

According to experience, set: $M=42$, maximum loop number $N_c_{max}=120$, $Q=0.6$, $\alpha=1$, $\beta=5$. Then set $\rho=0.5$, 0.75, and 0.9. Compare the errors of the system output based on each $\rho$ and choose the better one. The experimental results showed 0.5 and 0.75 are better. Then set it as 0.5, 0.6, and 0.75. The performance comparison experimental result showed that 0.75 and 0.6 are better. Then set it as 0.75, 0.7, 0.65, 0.6. Performance comparison is showed in Fig.4. As figure 4 shown, 0.7 is the best one.

B. Relative importance of visibility $\beta$

$\beta$ affects the transfer probability when the ant selects the path. If its value is too large, the algorithm will fall into a local optimal solution. Lots of experiments show when $\beta$ is valued $[4.5, 5]$, the algorithm has better convergence.

Set: $M=42$, maximum loop number $N_c_{max}=120$, $Q=0.6$, $\alpha=1$, $\rho=0.7$. Then set $\beta=4.5$, 4.75, and 5.0. Compare the errors of system output based on each $\beta$ and choose the better one. The comparison experimental result showed that 4.75 and 5.0 are better. Then set it as 4.7, 4.8, 4.9, and 5.0. Performance comparison experimental result showed that 4.8 and 4.9 are better. Then set it as 4.85, 4.875, 4.9, 4.925, and 4.95. Performance comparison is showed in Fig.5 As figure 5 shown, 4.9 is the best one.

Relative importance of trail $\alpha$

$\alpha$ indicates the relative importance of path. The larger its value is the more cooperation ants there will be. It is valued around 1.

Set: $M=42$, maximum loop number $N_c_{max}=120$, $Q=0.6$, $\rho=0.7$. Then set $\beta=4.9$. Set $\alpha=1$, 0.8, and 0.6. Compare the errors of system output and choose the better one. Performance comparison experimental result showed that 0.8 and 1.0 are better. Then set it as 0.85, 0.9, 0.95, and 1.0. Performance comparison experimental result showed 0.95 and 1.0 are better. Then set it as 0.96, 0.98, and 1.0. Performance comparison is showed in Fig.6. As figure 6 shown, 1.0 is the best one.
C. Quantity of pheromone $Q$

$Q$ reflects the convergence speed of the algorithm. Experiments show algorithm is close to the global optimal solution when $Q$ is valued (0, 1).

Set: $M=42$, maximum loop number $N_{\text{max}}=120$, $\rho=0.7$, $\beta=4.9$, $\alpha=1$. Set $Q=0.4$, 0.6, and 0.8. Compare the errors of system output and choose the better one. Performance comparison result showed that 0.4 and 0.6 are better; then set it as 0.4, 0.45, and 0.5. Performance comparison result showed that 0.5 is better; then set it as 0.5, 0.6. Performance comparison is showed in Fig.7.

As figure 7 shown, 0.5 is the best one.

![Figure 6 performance comparison](image)

Figure 6 performance comparison

![Figure 7 performance comparison](image)

Figure 7 performance comparison

Thus, through experience, the best parameters are $M=42$, $Q=0.5$, $\rho=0.7$, $\beta=4.9$. Set $\alpha=1$.

VI. CONCLUSION

A model based on improved ACA and BP is proposed to predict in Silicon content in hot metal. Also it makes prediction for the silicon content using the serialized Si data collected online from a certain domestic metal company.

ACKNOWLEDGMENT

This paper is supported by Beijing Key Discipline Development Program (No. XK100080537).

This model use BP to set up the control and prediction relationship between control variables with state variables and Silicon sequence data, and use improved ACA to train BP. In this way, it has higher hit rate than that using normal ACA. With the application of “black box” capability of neural network and the prediction of silicon content using clear relationship of input and output, the model is of more practical value than existing ones that is rather complicated. It can provide a good guidance and future in industrial automation.

REFERENCES


