Construction Engineering Cost Evaluation Model and Application Based on RS-IPSO-BP Neural Network

Yuan Hong[1,2]
[1]Hunan University, Changsha, Hunan, China, 410079
[2]Ministry of Finance Fiscal Science Research Institute, Beijing, China, 100142
Email:hy6360298@163.com

Haibo Liao [1]
[1]Hunan University, Changsha, Hunan, China, 410079

Yazhi Jiang [1]
[1]Hunan University, Changsha, Hunan, China, 410079

***Abstract*—Aimed at coping with the complexity of construction engineering cost evaluation, the advantages of rough set theory, particle swarm algorithm and BP neural network are integrated to put forward a new model of construction engineering cost evaluation, namely, the model of construction engineering cost evaluation of optimized particle swarm and BP neural network on the basis of rough set theory. First, rough set theory was used to reduce the factors affecting construction engineering cost and optimize input variables of BP neural network. Then, the improved particle swarm algorithm with constriction factors is adopted to optimize the initial weights and thresholds. Through this method, BP neural network can be used in a better way to solve nonlinear problems and to improve the rate of convergence and the ability to search global optimum. An engineering project in a city of Hunan is selected to make empirical analysis. It shows that based on the features of engineering, this new model enjoys a high practical value as it can be applied to make scientific evaluation of costs of construction engineering.**

***Index Terms*—cost evaluation, Rough Sets, Particle Swarm Optimization, Artificial Neural Networks**

I. Introduction

There are many factors affecting building engineering cost, and the factors and costs show complicated nonlinear relationship[1] . In 1962, British Project Cost Information Service Department put forward BCIS model[2]. In 1974, Koehn and Kouskoulas[3] analyzed shortcomings of BCIS model, and came up with an improved model, namely $C=a_0+a_1v_1+a_2v_2+a_3v_3+a_4v_4+a_5v_5+a_6v_6$ on the basis of regression analysis. In the early 1980s, Monte Carlo random engineering estimation model was put forward. Since 1990s, artificial neural network[7]-[9] and other artificial intelligent algorithms have been widely used in project evaluation, risk assessment, pattern recognition, etc. Domestically, in 2007, Duan Xiaochen[10] proposed project cost evaluation methods on the basis of grey theory and fuzzy mathematics theory. However, the methods fail to make highly precise evaluation and subject to the influence of subjectivity. In 2009, Wang Chengjun[11] raised project cost evaluation method based on neural network. Compared with other methods, it enjoys high accuracy and objectivity. Yet, it only has slow convergence rate and tend to fall into local minimum point. In 2009, Gao Yanna[12]utilized genetic algorithm to optimize weight and threshold value of neural network, and hereby, established rural land expropriation evaluation model. In 2010, by taking advantage of genetic neural network, Jing Chenguang[13] set up highway engineering evaluation model. Optimizing neural network through genetic algorithm leads to better global optimization and solutions to complex nonlinear project evaluation problems.

In all the above mentioned models, sample data are not processed, and the factors affecting project cost are not uniformed. However, option of the factors and relevance of sample data will affect the speed and accuracy of algorithms, which in turn affects the precision of evaluation. The paper puts forward a model of construction engineering cost evaluation of optimized particle swarm and BP neural network on the basis of rough set theory. First, rough set theory, genetic algorithm and BP neural network are introduced. Then, a model of construction engineering cost evaluation on the basis of rough set theory and genetic neural network is established. Finally, engineering projects are taken to test the model.

II. Optimized Particle Swarm and BP Neural Network on the Basis of Rough Set Theory

A. Fundamental Principles of Rough Set (RS)[14]

A.1. Decision Table

Decision table is a kind of important and special knowledge representation system, which can show most of the decision problems. It can be defined as follows:

doi:10.4304/jcp.9.4.1020-1025

suppose information system $S=(U,A,V,f)$, $A=C\bigcup D$, $C\bigcap D\neq\phi$. C is condition attribute set and D, decision attribute set. Knowledge representation system with condition attribute set and decision attribute set is referred to as decision table. The factors affecting engineering project cost constitute the conditional attribute set, while the costs of unit area compose the decision attribute set. Table 1 is about influence index and decision, in which { S_1, S_2, ..., S_m} are similar completed commercial residential buildings; { C_1, C_2, ..., C_n} are factors or indexes affecting cost evaluation; { D_1, D_2, ..., D_m} are the unit area costs.

TABLE 1

INFLUENCE INDEX AND DECISION

Sample	C_1	C_2	…	C_N	D
S_1	c_{11}	c_{12}	…	c_{1n}	d_1
S_2	c_{21}	c_{22}	…	c_{2n}	d_2
…			…	c_{3n}	d_3
S_m	c_{m1}	c_{m2}	…	c_{mn}	d_m

A.2.Attribute Reduction

According to RS, knowledge is the basic principle that objects follow to make assortment. Objects are represented by sets containing objects' characteristics and fundamental attribute of knowledge. In the process of classification, concepts form basic parts of knowledge, and knowledge forms modules of objects' domain of discourse. It can be defined through knowledge representation system[15].

$$S=(U,C,D,V,f) \tag{1}$$

In the equation, the non-empty finite set of U is called the domain of discourse. C is condition attribute set. D is resulting attribute set. $V=\bigcup_{a\in A}V_a$, V_a is threshold value of the attribute *a*. $f:U\times A\rightarrow V$ is an information function, defining attribute value of *S* in *U*. This data table to describe definition is referred to as knowledge representation system. The discernible matrix of S is $n\times n$, any element of which can be defined as follows.

$$a(x,y)=\{a\in A|f(x,a)\neq f(y,a)\} \tag{2}$$

$a(x,y)$ is a set of all the attributes of discernible objects *x* and *y*.

The discernibility function:

$$f(A)=\prod_{(x,y)\in u\times u}\sum a(x,y) \tag{3}$$

In the function, the conjunction expression of minimal disjunctive normal form is the reduction of attribute set *A*.

B. Improved Particle Swarm Optimization（IPSO）

B.1. Basic Particle Swarm Optimization （PSO）

In the international neural network academic conference of IEEE, 1995, American social psychologist J.Kennedy and electrical engineer R.C.Eberhart formally published the article Particle Swarm Optimization, marking the birth of POS. The basic idea is assuming a particle swam containing M particles can fly at a certain speed in D-dimensional search space. The attribute of particle i at t time is as follows.

Position: $x_i^t=(x_{i1}^t,x_{i2}^t,\cdots,x_{id}^t)^T$ (4)

$x_{id}^t\in[L_d,U_d]$, L_d and U_d are respectively upper limit and lower limit of the search space.

Velocity: $v_i^t=(v_{i1}^t,v_{i2}^t,\cdots,v_{id}^t)^T$ (5)

$v_{id}^t\in[v_{\min,d},v_{\max,d}]$, $v_{\min}$ and $v_{\max}$ are respectively the smallest and largest Velocity.

Optimal position of individuals: $p_i^t=(p_{i1}^t,p_{i2}^t,\cdots,p_{iD}^t)^T$; global Optimal position: $1\le d\le D,1\le i\le M$. $p_g^t=(p_{g1}^t,p_{g2}^t,\cdots,p_{gD}^t)^T$.

Then position of particle at time $t+1$ can be expressed in the following equation:

$$v_{id}^{t+1}=v_{id}^t+c_1r_1^t(p_{id}^t-x_{id}^t)+c_2r_2^t(p_{gd}^t-x_{id}^t) \tag{6}$$

$$x_{id}^{t+1}=x_{id}^t+v_{id}^{t+1} \tag{7}$$

c_1, c_2 are learning factors. Appropriate c_1 and c_2 can accelerate convergence and avoid falling into local optimum. r_1 and r_2 are any number between 0 and 1.

B.2. Improved Particle Swarm Optimization

With the increase of dimension of optimization problem, the basic PSO algorithm is prone to be premature, go into stagnation and even local extremum. On this basis, some people put forward many improved algorithms. Many scholars optimized the performance of particle swarm optimization by adding an appropriate inertia weight w, and gained some fruits[16].

In 1999, M.Clerc suggested introducing constriction factor to PSO algorithm to replace traditional inertia weight. The velocity can be presented as:

$$v_{id}^{t+1}=K(v_{id}^t+c_1r_1^t(p_{id}^t-x_{id}^t)+c_2r_2^t(p_{gd}^t-x_{id}^t)) \tag{8}$$

Constriction factor K:

$$k = \frac{2}{\left|2-\varphi+\sqrt{\varphi^2-4\varphi}\right|}$$

$$\varphi = c_1 + c_2, \varphi > 4 \quad (9)$$

Compared with inertia weight, constriction factor K can control the velocity of particles more effectively, enhancing the local searching ability of the algorithm.

C.BP Neural Network

Back propagation artificial neural network (BPANN) is a multi-layer feedforward network mode， made up of input layer, hidden layer and output layer. Links exist among layers but not units of the same layer .The BP algorithm means, through back propagation of output errors, constantly adjusting and revising link weights and threshold value of neural algorithm so as to control the output errors. Usually, the node function of BPANN is "S" type function.

The usual action function f(x) is differentiable Sigmoid function:

$$f(x) = \frac{1}{1+e^{-x}} \quad (10)$$

Error Function R:

$$R = \frac{\sum (Y_{mj} - Y_j)^2}{2} \quad (j=1,2,\ldots,n) \quad (11)$$

In the equation, Y_j is output of expectation, Y_{mj} actual expectation, n length of sample.

The weight modification formula of BP algorithm can be expressed as:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_{pj} o_{pj} \quad (12)$$

W_{ij} is the link weight of neuron. η is learning rate of neural network. O_{pj} is the output of sample p. δ_{pj} is modification value of error.

D. Optimized Particle Swarm and BP Neural Network on the Basis of Rough Set Theory

(1) The following is the mathematical description of the optimization of RS-IPSO-BP neural network.

$$\begin{cases} \min E(w_i, w_o, \theta, r) = \frac{1}{2}\sum_{t=1}^{N_1}\sum_{k=1}^{n}(y_k(t) - \hat{y}_k(t))^2 = \min R \\ \text{s.t} \quad w_i \in R^{m\times p}, w_o \in R^{p\times n}, \theta \in R^p, r \in R^n \end{cases} \quad (13)$$

$E(w_i, w_o, \theta, r)$ is objective function. $\hat{y}_k(t)$ is output of neural network. $y_k(t)$ is the target value. w_i is the weight from input layer to hidden layer, and w_o the weight from hidden layer to output layer. Θ is the threshold value of hidden layer and r is the threshold value of output layer.

(2) The detailed steps of RS- IPSO-BP neural network algorithm are as follows.

Step1 Establish building engineering cost evaluation index set;

Step2 Discretizatize all attribute value;

Step3 Form decision table;

Step4 Reduce attributes to determine input index;

Step5 Determine topology structure of neural network, and randomly initialize weight and threshold value;

Step6 Initialize particle swarm, and set the size of particle swarm. Dimension of each particle is determined by the number of weights and threshold performing the linking function in neural network. Learning factors, the maximum and minimum of particles, and the largest and smallest velocity of particles are also determined. Position and velocity of particles are set randomly.

Step7 Define fitness function $f(i)$

Step8 Input the weighs and thresholds value of position vector s_i of each particle to neural network. Use learning sample to train neural network. Work out the actual output of neural network and, in the end, the Fitness value $f(i)$ of each particle.

Step9 Evaluate each particle, compare fitness value of new particles with individual extremum value of particles at the last moment, and update the optimal position and extremum of the individual particles.

Step10 Compare individual extremum of each particle with individual optimal value of global particles at the last moment, update the optimal position and extremum of the global individual particles.

Step11 Update velocity and position of particles as well as constriction factor k on the basis of equation (7), (8) and (9).

Step12 Reach the conditions to stop iteration. Utilize the optimal position of the population, that is G_b, to establish BP neural network, which means to take weights and threshold value optimized by the improved particle swarm as the initial weights and threshold value in BP neural network simulation.

Step13 Input test data set and output simulation value.

III. Application of the Evaluation Model in Building Engineering Cost Evaluation

A. Collection and processing of initial data of Building Engineering Cost Evaluation

A.1. Collection of Building Engineering Cost Evaluation Data

Sample data in this paper are from 20 cases of ordinary civil buildings engineering evaluation in Hunan Province. 16 of them are put into training set to train BP neural network. The other four form a test set to test the accuracy of the trained BP neural network model.

A.2. Confirmation of condition attribute set and decision attribute set

In this paper, condition attribute set C contains 21 attributes, namely engineering type (C_1), structure type (C_2), height of each layer(C_3), number of layers (C_2),

base type (C_3), concrete supply(C_6)、type of pile foundation(C_7) exterior trim(C_8), interior trim(C_9), masonry engineering(C_{10}), building ground engineering(C_{11})、door and window engineering(C_{12})、ceiling trim(C_{13})、flat combination(C_{14})equipment engineering (C_{15}), water-power engineering (C_{16}), beam column engineering (C_{17}), the area of circulation (C_{18}), price index (C_{19}), and coefficient of structural area (C_{20}). Decision attribute set D is of single attribute, that is price of per square meter (D_1).

A.3. Data Discretization

When using rough set theory to deal with decision table, values in the table should be represented with discretized data. If codomain of condition attribute or decision attribute is discretized value, then it can be processed directly. If it is continuous value, then it must be discretized. Continuous data are discretized with the method of equidistance division in this paper. The results are shown in Table 2.

Reduction of condition attribute

After discretization, an improved heuristic algorithm[17] applied in rough set is adopted to make reduction. The minimum attribute reduction is {C_1、C_2、C_5、C_6、C_8、C_9、C_{11}、C_{12}、C_{15}、C_{20}}, all contained in attribute core.

B. Structure of RS-IPSO-BP Neural Network

(1) Adopt three-layer RS-IPSO-BP neural network. The number of input and output are 10 and 1 respectively. Set node number of hidden layer at 12 according to the equation $y = \sqrt{n+m} + a$.

(2) Process input and output data of neural network. Input data are the discretized value of the ten influence indexes in table 2. Output data are the value of per square meter after data discretization.

(3) Determination of parameters of neural network. The learning rate is 0.7 and factor of momentum 0.9. Size of particle swarm is 80. Learning factor $c_1 = c_2 =$ 2.0.Population n=40.The largest iteration is 1000.

TABLE2

DISCRETIZED SAMPLE DATA DECISION TABLE

sequence	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}	C_{17}	C_{18}	C_{19}	C_{20}	C_{21}	D_1
1	1	1	3	3	4	2	1	2	3	3	3	3	2	4	3	1	3	5	2	1	3	3
2	2	2	3	4	3	2	3	4	3	4	1	4	3	2	3	3	3	3	3	2	2	4
3	3	2	2	2	3	5	3	1	3	4	2	2	4	3	4	5	1	2	2	3	3	2
4	2	1	3	5	1	4	2	4	1	1	2	4	3	2	1	4	5	4	3	5	3	5
5	5	1	2	2	3	2	3	3	5	3	2	2	4	3	4	3	2	3	2	4	5	2
6	3	2	1	1	3	2	2	4	1	1	2	1	1	2	1	2	1	2	1	6	3	1
7	4	1	3	2	5	1	2	1	2	1	2	2	2	5	1	1	3	1	2	3	7	1
8	1	1	1	2	2	1	2	2	3	2	2	2	3	2	3	2	2	4	2	2	4	2
9	3	2	4	3	4	3	2	3	4	3	4	4	3	3	4	2	4	5	3	2	4	4
10	2	3	3	4	2	3	4	3	3	3	2	5	4	4	2	3	3	4	3	4	3	5
11	4	2	3	2	4	4	4	2	4	3	2	3	5	2	3	5	4	3	4	3	3	3
12	2	5	3	5	2	3	3	3	2	4	3	3	4	3	2	5	5	5	5	5	4	5
13	4	1	3	2	4	2	4	4	4	3	4	2	5	2	3	3	2	3	2	3	5	3
14	5	3	1	1	4	4	3	5	2	3	2	4	2	3	2	4	2	3	4	6	4	1
15	2	1	4	2	5	3	3	3	1	2	3	2	3	4	4	1	3	1	2	4	6	2
16	3	2	1	2	4	2	2	2	3	3	2	3	4	2	2	2	4	2	3	2	5	3
17	1	1	4	3	3	1	4	3	3	4	4	4	5	5	2	3	3	4	3	3	4	3
18	4	3	3	4	3	3	3	3	4	4	1	3	3	3	4	5	4	4	4	3	2	5
19	1	2	5	2	2	5	4	2	3	5	3	2	2	4	2	4	4	3	2	4	4	3
20	5	5	3	5	2	3	2	5	2	2	2	5	3	3	3	3	2	5	4	5	5	5

C. Training and Simulation of the Model

By taking the first 16 figures as training samples and the other 4 as test samples from the sample data Table 2, training and simulation of ordinary BP neural network, RS-BP neural network, RS-PSO-BP neural network, and RS-IPSO-BP neural network are made in MATLABR2007b.As is shown in figure 1, the rate of convergence and accuracy of RS-IPSO-BP (Line D) are superior to the other three models. Table 3 and figure 2 shows that the relative errors between evaluation results and factual results are below 5%, but RS-IPSO-BP model generates even smaller errors, which is within 2.5%,an acceptable rate. Thus, it can be concluded that compared with traditional BP neural network models, RS-IPSO-BP neural network model can work out better evaluation results.

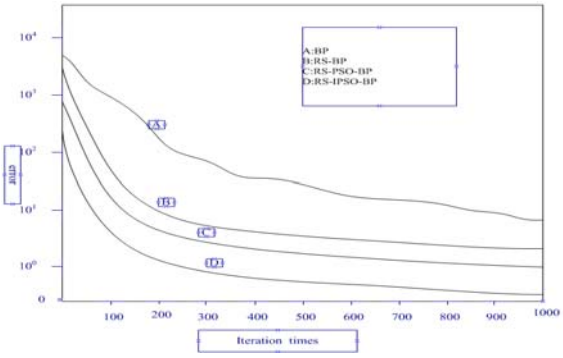

Figure1 Convergence performances of sample training of four algorithms

TABLE 3

EVALUATION SIMULATION RESULTS OF FOUR MODELS

Sample	Initial Price /m^2	BP	Relative Error	RS-BP	Relative Error	GA-BP	Relative Error	RS-GA -BP	Relative Error
17	4358	4586	0.0523	4580	0.0509	4260	0.0225	4319	0.0089
18	4685	4856	0.0365	4840	0.0331	4821	0.0290	4590	0.0203
19	4795	4965	0.0355	4960	0.0344	4921	0.0263	4850	0.0115
20	4657	4876	0.0470	4860	0.0436	4831	0.0374	4580	0.0165

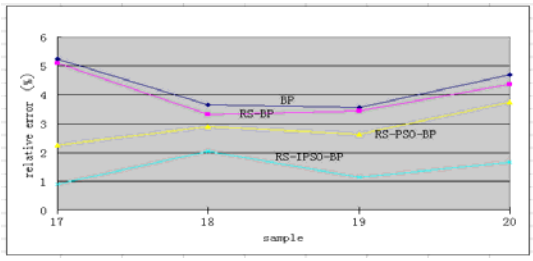

Figure 2 Relative errors of sample test simulation of four models

IV.CONCLUSIONS

(1) By adopting improved particle swarm algorithm to optimize the weight and threshold value of ordinary BP neural network, BP neural network is able to avoid falling into local minimum point.

(2)By taking rough set theory to simplify input data to the neural network, secondary factors influencing residential property evaluation are eliminated, which correspondingly reduces the data size and speed up convergence of neural network.

(3)Compared with ordinary BP model, RS-BP model, and RS-PSO-BP model, RS-IPSO-BP neural network, making full use of rough set theory and IPSO, is more scientific and able to make more accurate evaluation.

(4)The case study of the application of RS-IPSO-BP neural network in building engineering evaluation shows that the evaluation results are basically identical with the actual outcomes. The forecasting results are relatively stable, and the errors are relatively small (within 2.5%). Therefore, it serves as an effective new method for building engineering cost evaluation.

ACKNOWLEDGMENT

This work was supported by Project 71103060 of the China National Natural Science Funds and Project 09YBB079 of the Hunan Social Science Fund.

REFERENCES

[1] X.C. Duan, “Study of Theory and Method of Overall Investment Control on Government Investment Projects”, Beijing Science Press, pp87-223, 2007.

[2] F. Liu, Z. Huang, “A Study of Fixed Asset Investment Evaluation in China and Abroad”, Infrastructure Optimization, vol.32, No.8, pp56-58, 2003.

[3] Kousloulas Vasily, Koehn Edward, “Predesign Cos-Estimation Function for Buildings”, Journ-al of the Contraction Division ASCE., vol.34, No.7, pp589~604, December, 1974.

[4] Wall DM. “Distributions and correlations in Mon-Carlo simulation”, Construct Manage Economic, 1997, No.15, pp241-258.

[5] W.C. Wang, “Simulation-facilitated model for assessing cost correlations”, Journal of Company Aided Civil infrastructure Engineer-ing, 2002, vol.18 No.5, pp. 368-380.

[6] Tummala VMR,Burchett JF. “Applying a risk management process (RMP), to manage cost risk for an EHV transmission line Project”, International Journal of Project Management, 1999, vol.7No.4, pp223-235.

[7] Dybowski R,Roberts S J.“ Confidence intervals and Prediction Intervals for feed-forward neural networks”, In Clinical applications of artificial neural networks .Cambridge Universi-ty Press.200l, vol.8 No.7, pp298-327.

[8] HegaZy T,Ayed A.“Neural network model for Parametric cost estimation of highway Projects”, Journal of Construction Engineering and Management, vol. 9, No.3, pp210-219, 1998.

[9] G.H.Kim ,D.S.Seo, K.I.Kang,“ Hybrid Models of Neural Networks and Genetic Algorithms for Predicting Preliminary Cost Estimates”, Journal of Computing in Civil Engineering,vol.15, No.8,pp208-211,2005.

[10] X.C. Duan, J.W.Zhang, “Study of Theory and Method of Overall Investment Control on Government Investment Projects”, Beij-ing Science Press, pp12-156, 2007.

[11] C.J.Wang,X.H.Zuo,“ Building Engineering Cost Evaluation Forecasting Model Based on BP Neural Network ”,Journal of Shanxi Uni-versity of Finance and Economics,vol.14,No.11,pp327-330,2009.

[12] Y.N.Gao, D.L.Zhu, Y.Q.Chen and X.F.Zhang, “ Rural Land Expropriation Price Evaluation Model Based on Neural Network”, Theory and Practice of Systematic Engineering,vol.12,No.04,pp103-110,2009.

[13] C.G.Jin , X.C. Duan, “ Study of Engineering Cost Evaluation Method Based on Neural Network ”, Journal of Shijiazhuang Railway University(Social Science,,vol.15,No.04, pp11-17,2010.

[14] W.X. Zhang. “Rough Set Theory”, Beijing Science Press, pp24-36, 2008.

[15] B.Shi, Y.X.Li, X.H.Yu, et al. “A modified particle swarm optimization and radial ba-sisfunction neural network hybrid algorithm model and its application”, WRI Global Con-gres on Intelligent Systems, Xiameng, China, pp134-138, January, 2009.

[16] L.P. Zhang and H. J. Yu,"Analysis and Improvement of Particle Swarm Algorithm" Information and Control, vol.33,No.5 , pp513-517,2004.
[17] L. Zhang and X.Y.Lu. "An improved algorithm for attribute reduction and rule generation",Journal of Information and Computational Science, vol.5,No.5, pp2611-2618,2008.

Yuan Hong Hunan Province, China. Birthdate: November, 1981. is Applied Economics Ph.D., graduated from school of Finance and Taxation, Zhongnan University of Economics and Law. And research interests on intelligent control and project management.

He is an associate professor of Economy and Trade College, Hunan University. Meanwhile, he is a post doctor of Ministry of Finance Fiscal Science Research Institute.

Yazhi Jiang Hunan Province, China. Birthdate: December, 1987. is Master of Taxation Administration ,graduated from Economy and Trade College , Hunan University. And research interests on performance.

Haibo Liao Hunan Province, China. Birthdate: December, 1989. is Master of Taxation Administration ,graduated from Economy and Trade College , Hunan University. And research interests on performance.