

# Research on Tacit Knowledge Acquisition Based on Gray Theory and RBF Neural Network

Xiaomin Guo, Zhiwei Zhou, Yamin Zhao, Mengmeng Zhang, Zhikang Lian, Lin Wei and Jianhua Zhang

School of Management Engineering, Zhengzhou University, Zhengzhou 450001 China

Corresponding Email: 530378118@qq.com

**Abstract.** In the era of knowledge economy, Knowledge has become the main means of production and the core source for value creation. However, especially tacit knowledge acquisition, knowledge acquisition is bottleneck in all kinds of knowledge system, cutting down the benefit of knowledge sharing, application and innovation. In view of this, in this paper, the main ideas and deficiencies of the existing methods of knowledge acquisition are generalized. And then, the basic principles, characteristics and the integrated advantage of gray theory and RBF neural networks are analyzed. On this basis, three kinds of models are designed and discussed, which is able to obtain tacit rule set by using gray theory integrated with RBF neural network. Simultaneously, an empirical analysis is carried out to analyze application results of tacit knowledge acquisition, which includes five kinds of models using independent and integrated strategies.

## 1 Literature Review and the starting point

### 1.1 Knowledge Acquisition connotation

The academic world on the "knowledge acquisition" of the understanding there are multiple versions, representative are: Breuker suggested that knowledge acquisition is to obtain useful knowledge from a specific source to solve the problem of knowledge or experience, it is the foundation for people to understand and transform the world [1]; Cassiman and Veugelers argue that knowledge acquisition is the fundamental activity that underpins the organization to achieve knowledge innovation [2]; In China, Wu Quan-yuan and Liu Jiang-ning pointed out that knowledge acquisition is the process of extracting knowledge from external knowledge sources into the computer system [3]. In view of the foregoing, we believe that knowledge acquisition refers to the analysis and extraction of problem-solving knowledge from knowledge sources, formats and codes them by appropriate knowledge representation methods then stored in the system knowledge base.

### 1.2 The Main Methods of Acquiring Knowledge and Its limits

Tacit knowledge and explicit knowledge as knowledge of the two basic forms, for its different characteristics of the commonly used knowledge acquisition method is different. There are two common methods of explicit knowledge acquisition: distributed search and data mining. The former is a search strategy, which can be used to query and retrieve knowledge of RDB, special documents, Internet sites, Web pages, and so on, and can gain explicit knowledge in related fields [4,5]. However, the distributed search algorithm is generally high complexity, low search efficiency. Data mining is to extract credible, effective and understandable new knowledge from a large amount of data. However, the knowledge gained by this method does not necessarily have good Interpretability.

Manual knowledge acquisition, semi-automatic knowledge acquisition and automatic knowledge acquisition are the main methods of tacit knowledge acquisition [6,7]. Manual knowledge acquisition through face-to-face communication with domain experts and experts, which is very inefficient. Compared with manual knowledge acquisition, Semi-automatic knowledge acquisition improves the efficiency of knowledge acquisition to a certain extent. However, relatively frequent I/O operations still restrict the efficiency of knowledge acquisition, and the production rules of manual access are often redundant or contradictory. The way of automatic knowledge acquisition is high efficiency, high degree of intelligence is the future direction of development; its main drawback is that algorithms maturity is not high, systematic training time-consuming data demand.

**1.3 Starting point**

Explicit knowledge is easy to code and transferable effectively. However, in the "knowledge iceberg" model, explicit knowledge is only part of the knowledge iceberg floating on the water surface. Tacit knowledge hidden beneath the water surface accounts for knowledge Iceberg about 90%. On the other hand, the famous SECI model shows that tacit knowledge plays an important role in ensuring and promoting the application, sharing and innovation of knowledge[8]. Tacit knowledge includes many types, such as informal experience and skills, difficult to express skills, recessive rule sets of data systems. Tacit knowledge in terms of the proportion of knowledge or the role of knowledge are far greater than the explicit knowledge. Therefore, how to do tacit knowledge acquisition well is the focus of knowledge management research.

**1.4 Model selection**

The gray RBF neural network model, which combines the RBF neural network model with the gray model, not only has the characteristics of small sample data modeling for gray system, but also has the advantages of neural network model with adaptive ability to nonlinear and inexact laws ; The model reduces the demand for raw data and the training time consumption of automatic knowledge acquisition to a certain extent.

**2 Gray RBF neural network model**

**2.1. The GM (1,1) model**

Gray GM (1,1) modeling process is as follows<sup>[9,10]</sup>, the original data sequence:

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$$

And a new sequence is obtained by accumulating it once

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$$

In this formula,

$$X^{(1)}(i) = \left\{ \sum_{j=1}^i x^{(0)}(j) \mid i = 1, 2, \dots, n \right\}$$

Order

$$D(i + 1) = -\frac{1}{2} [X^{(1)}(i) + X^{(1)}(i + 1)] \tag{1}$$

create the matrices **B** and **Y<sub>n</sub>**:

$$B = \begin{bmatrix} D(2) & 1 \\ D(3) & 1 \\ \dots & \dots \\ D(n) & 1 \end{bmatrix}, Y_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \dots \\ x^{(0)}(n) \end{bmatrix} \tag{2}$$

Let the series satisfy the first - order univariate ordinary differential equation:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b. \tag{3}$$

Among them, a is the development coefficient, b is the gray effect, both a and b are non-0 constant. The least squares method is used to solve a and b respectively, then:

$$p = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y_n \tag{4}$$

Thus, the cumulative series prediction model can be obtained:

$$\hat{x}^{(1)}(i + 1) = \left[ x^{(0)}(i) - \frac{b}{a} \right] e^{-ai} + \frac{b}{a}, (i = 0, 1, 2, \dots) \tag{5}$$

The predicted values for the series is:

$$\hat{x}^{(0)}(i + 1) = \hat{x}^{(1)}(i + 1) - x^{(1)}(i) \tag{6}$$

This is a function model based on time response and is the core formula of GM (1,1) model.

**2.2 RBF neural network**

RBF neural network, or radial basis function neural network, is a three-layer forward network. RBF neural network is superior to ordinary feed-forward network [11], because the convert of RBF can transform a nonlinearly separable

sample point in one space into a linearly separable sample point in another space. RBF network structure similar to the BP network. The first layer is the input layer, which is composed of the source nodes. The second layer is the hidden layer and the number of nodes depends on the requirements of described questions. The third layer is the output layer, which responds to the input mode. In this network, the transformation from the input space to the hidden layer space is nonlinear, and the transformation from the hidden layer space to the output layer space is linear; the transform function of the hidden layer unit is a radial basis function, which is a nonnegative nonlinear function of the locally distributed radial symmetric attenuation of the central point.

RBF network has a good nonlinear function approximation ability. It does not require the sample has a specific distribution and only uses the neural network itself reflects the non-linear change of things which directly estimate the non-linear change characteristics of things; At the same time, because it affects the network output with a few connection weights in the local area of the input space, it converges faster than the traditional neuron model, and is not easy to fall into the local minimum point and has better robustness.

**2.3 Gray RBF Neural Network Combined Model**

Gray GM (1,1) and RBF neural networks achieve integration in three way by series, parallel and embedded[12-14]. Among them, tandem Gray RBF neural network as shown in figure 1, the gain results of the rules knowledge of the multiple Gray GM (1,1) model as input of RBF network model, and then use RBF network 's non-linear mapping function to find out the Gray GM (1,1) model weight, so that the output get better accuracy of the rules knowledge output.

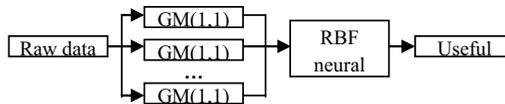


Fig.1 .Tandem Gray RBF neural network.

For initial datas, The parallel gray RBF neural network model use gray GM (1,1) model and the RBF neural network model to modeling of problem.. Then, the weighted average of the two models' output is obtained.

The embedded gray RBF neural network uses the gray GM (1,1) model to gray the sample data before inputting the data. the gray output of the RBF neural network is reduced, that is, the gray GM (1,1) model is used for the RBF output. And whitened to obtain a final output.

**3 Empirical Analysis**

**3.1 Examples Background and Data Acquisition**

Stock prices are affected by many factors, with a high degree of complexity and nonlinearity. Stock opening and closing prices, the highest and lowest prices as the two extreme performance of stock prices, they and the stock of the actual price between the implied association rules. Changes in stock volume and turnover can reflect the level of stock activity. One of the most important indicators is the closing price, which is the benchmark for investors to make money and lose money. Due to the opacity of the market, investors can not get all the factors affecting the closing price.

In view of this, assuming that the government policy, the company's basic business conditions remain unchanged, we select Yutong Bus (600066) February 7, 2014 to March 18, 2014 the stock historical data as the original data 1) model, RBF neural network and three gray RBF neural network models were used to analyze and acquire the rule knowledge contained in the stock closing price data set. Then, Is applied to the closing price forecast for the stock, which is the forecast value of the closing price of the stock from March 19, 2014 to March 26, 2014; and finally, the actual value in the segment (see Table 1 Dashed line below) to test the accuracy of the predicted value, and then determine the level of validity of the acquired hidden rule knowledge.

**3.2 Implicit Rule Knowledge Obtainment of Five Different Models**

**3.2.1 Through the GM (1,1) model to obtain hidden rules**

The closing price of the part above the dotted line in Table 1 is taken as the original data, and the recessive rule is obtained by GM (1,1) model. First of all, the original data is accumulated to obtain the accumulated sequence  $X^{(1)}$ :

$$X^{(1)} = (18.01, 37.41, 56.7, 75.63, 95.08, 114.67, 134.45, 153.61, 172.44, 190.86, 209.32, 226.84, 243.93, 261.12, 278.71, 296.81, 315.5, 333.09, 350.48, 367.81, 385.47, 402.56, 419.65, 436.4, 453.2, 469.61, 486.41, 503.14)$$

On this basis, the matrix  $B$  and  $Y$  are established and  $P$  is obtained.

The simulation sequence of  $X^{(1)}$  is further calculated and the simulated sequence of  $s$  is generated by accumulation. The accuracy of the residual test  $P^{(0)}$  is:

$$P^{(0)} = (1 - \bar{\varepsilon}) \times 100\% = 97.2\% \tag{7}$$

$$B = \begin{bmatrix} D(2) & 1 \\ D(3) & 1 \\ \dots & \dots \\ D(n) & 1 \end{bmatrix} = \begin{bmatrix} -27.71 & 1 \\ -47.055 & 1 \\ \dots & \dots \\ -494.775 & 1 \end{bmatrix} \tag{8}$$

$$Y_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \dots \\ x^{(0)}(n) \end{bmatrix} = \begin{bmatrix} 19.4 \\ 19.29 \\ \dots \\ 16.73 \end{bmatrix} \tag{9}$$

$$p = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y_n = \begin{bmatrix} 0.0066 \\ 19.7283 \end{bmatrix} \tag{10}$$

**3.2.2 Obtain hidden rules through RBF neural network**

RBF neural network is calculated by using MATLAB software, RBF neural network is trained by newrb function, the simulation function is sim. The opening price, the highest price, the lowest price, the transaction volume and the transaction number of the part above the dotted line in Table 1 are taken as the output layer, and the number of input layer nodes is  $R = 5$ . The closing price as the output layer, the output layer node number  $S = 1$ . The newrb function has the function of automatically increasing the number of nodes in the hidden layer. Therefore, the hidden layer nodes are automatically adjusted by the program and are no longer set. All the sample data are normalized to remove the influence of different units on the RBF neural network model.

In this paper, the adopted normalization method is

$$\bar{x}(k) = \frac{2(x(k) - x_{\min})}{x_{\max} - x_{\min}} - 1 \tag{11}$$

in which  $k = 1, 2, \dots, 34$  represents the value of each of the different influencing factors, and  $x_{\min}$  and  $x_{\max}$  represent the minimum and maximum values of each influence factor, respectively. The normalized sample data is selected as the training sample, and the spread value is tested as the test data under the dotted line in Table 1 (numbers 29 to 31). Select the spread value from 37 to 45, and draw the simulated data and the actual data of different spread values, see Table 2. As can be seen from the table, when the spread value 43, the closest to the true value. The final RBF neural network is created as: net = newrb (P, T, 0.005, 43).

**3.2.3 Obtain recessive rules through tandem gray RBF neural network**

GM (1,1) model was used to calculate the opening price, the highest price, the lowest price, the trading volume, the transaction number and the closing price of Yutong Company (600066), and the corresponding simulation sequence was obtained. As the posterior error ratio of turnover and the number of transactions are  $C_{\text{volume}} = 0.8526$  and  $C_{\text{turnover}} = 0.7842$  respectively, the accuracy of which can not meet the requirements. Therefore, only the opening price, highest price and the lowest simulation sequence of Yutong Bus (600066) are selected as input layer and the nodes number of input layer  $R$  is 3. The closing price as the output layer, the nodes number of output layer  $S$  is 1. Use the sample data of Nos. 1 to 28 in Table 1 as the training samples and the sample data of Nos. 29 and 30 as the test data to test the spread value. Debugging found that it is the closest to the true value when the spread value of 4. The final RBF network is: net = newrb (P, T, 0.005, 4). The results of the implicit rule application are shown in Table 3.

**Table 1** Yutong Bus (600066) stock data.

numb	Date	Opening quotation	highest	lowest	Volume	Turnover number	Closing
1	2014/2/7	17.78	18.08	17.7	22845	2802	18.01
2	2014/2/10	18.24	19.81	18.24	239398	17491	19.4
3	2014/2/11	19.38	20.17	19.18	195687	12882	19.29
4	2014/2/12	19.43	19.5	18.83	104345	11025	18.93
5	2014/2/13	19.15	19.77	19	180840	15579	19.45
6	2014/2/14	19.22	19.8	19.01	92542	8034	19.59

7	2014/2/17	19.87	19.95	19.41	82475	9618	19.78
8	2014/2/18	19.67	19.75	19.03	104130	10903	19.16
9	2014/2/19	19.2	19.38	18.8	99349	9690	18.83
10	2014/2/20	18.92	18.98	18.29	90619	10047	18.42
11	2014/2/21	18.4	18.6	18.26	54283	6321	18.46
12	2014/2/24	18.35	18.36	17.5	145797	14773	17.52
13	2014/2/25	17.52	17.66	16.47	209073	16751	17.09
14	2014/2/26	17.1	17.42	16.88	106920	9067	17.19
15	2014/2/27	17.23	17.9	17.06	180925	10830	17.59
16	2014/2/28	17.43	18.4	17.21	167603	11277	18.1
17	2014/3/3	18.11	19.2	18.11	199555	13515	18.69
18	2014/3/4	18.28	18.3	17.5	218370	17132	17.59
19	2014/3/5	17.75	17.85	17.35	119701	9734	17.39
20	2014/3/6	17.4	17.44	17.08	52471	5361	17.33
21	2014/3/7	17.4	17.82	17.35	68688	5642	17.66
22	2014/3/10	17.49	17.49	16.99	78765	6695	17.09
23	2014/3/11	17.09	17.15	16.85	65032	8971	17.09
24	2014/3/12	17.05	17.05	16.5	70396	6838	16.75
25	2014/3/13	16.66	16.98	16.66	35826	3584	16.8
26	2014/3/14	16.7	16.8	16.38	72118	6725	16.41
27	2014/3/17	16.61	16.83	16.4	82789	6623	16.8
28	2014/3/18	16.85	16.88	16.63	87685	8244	16.73
29	2014/3/19	16.72	16.72	15.52	257000	19285	15.98
30	2014/3/20	15.96	15.96	15.1	201594	18219	15.22
31	2014/3/21	15.17	15.8	14.86	172199	11498	15.57
32	2014/3/24	15.61	15.71	15.25	119907	9226	15.47
33	2014/3/25	15.72	16.2	15.68	137274	10072	15.86
34	2014/3/26	15.91	16.47	15.91	155977	14163	16.39

**Table 2.** The simulated data and the actual data of different spread values.

Spread values	Output value of Sample 29	actual value of error	Output value of Sample 30	actual value of error	Output value of Sample 31	actual value of error			
37	-0.6688	0.0018	-0.806	-0.194	-0.8255	-0.0205			
38	-0.6704	0.0034	-0.8081	-0.1919	-0.8265	-0.0195			
39	-0.6718	0.0048	-0.8099	-0.1901	-0.8275	-0.0185			
40	-0.6731	0.0061	-0.8117	-0.1883	-0.8283	-0.0177			
41	-0.6743	-0.667	0.0073	-0.8133	-1	-0.1867	-0.8291	-0.846	-0.0169
42	-0.6755	0.0085	-0.8148	-0.1852	-0.8299	-0.0161			
43	-0.6765	0.0095	-0.8161	-0.1839	-0.8305	-0.0155			
44	-0.6775	0.0105	-0.8174	-0.1826	-0.8312	-0.0148			
45	-0.6784	0.0114	-0.8186	-0.1814	-0.8318	-0.0142			

3.2.4 Parallel gray RBF neural network model

Using the parallel gray RBF neural network model, the system output of the GM (1, 1) model and the RBF neural network based on the weight coefficient combination for the hidden rule acquisition is:

$$f_c = \frac{Var(e_2)}{Var(e_1)+Var(e_2)}f_1 + \frac{Var(e_1)}{Var(e_1)+Var(e_2)}f_2 = 0.088f_1 + 0.912f_2 \tag{12}$$

Correspondingly, The results of the implicit rule application are shown in Table 3.

**3.2.5 Embedded gray RBF neural network model**

Add up the sample data of the opening, the highest price, lowest price, volume, turnover and the closing price, and then use the previously described method to normalize the cumulative data. The normalized results of sample numbers 1 to 28 in Table 1 are selected as the training samples, and the sample data of numbers 29 to 31 are used as the test data to test the spread value. Debugging found that when the spread value of 9 when the closest to the true value. The final RBF neural network is: net = newrb (P, T, 0.005,9), which is the closest to the true value when the spread value is 9.

**3.2.6 Single input layer embedded gray RBF neural network model**

The embedded gray RBF(5) neural network model is adopted, but the sample data only adopts the closing price. Add up the sample data of the closing price, and then use the previously described method to normalize the cumulative data. The sample data to do the cumulative accumulation of the closing date, then the cumulative data using the methods described above normalized. The normalized results of sample numbers 1 to 28 in Table 1 are selected as the training samples, and the sample data of numbers 29 to 31 are used as the test data to test the spread value. Debugging found that when the spread value of 50 when the closest to the true value. The final RBF neural network is: net = newrb (P, T, 0.005, 50), which is the closest to the true value when spread is 50.

**3.3.Results analysis**

Table 3 describes the comparison of the output results from the five model implicit rule sets described above. It can be seen that the parallel RBF neural network model has better results than GM (1,1) model and RBF neural network model in the extraction and application of implicit rules of stock closing prices. And, in the implicit rule extraction and application of the stock closing price it also has more advantages than the RBF neural network model and the embedded RBF neural network model; On the other hand, the embedded RBF neural network model increases the influence factor of the closing price of single day after the completion of the ashing process, and the result of the application of its recessive rule set has a big error, while the single input layer RBF neural network model has been less influenced, the application result of the recessive rule set of it is less.

**Table 3.** Model rule extraction value.

Numb		29	30	31	32	33	34	Mean
actual data		15.98	15.22	15.57	15.47	15.86	16.39	15.748
GM(1,1)Model	Extract value	16.366	16.259	16.152	16.046	15.941	15.836	16.100
	Absolute error	0.386	1.039	0.582	0.576	0.081	-0.554	0.352
	Relative error/%	2.359	6.390	3.603	3.590	0.508	-3.498	2.159
	Extract value	15.66	15.776	16.165	16.122	16.91	17.013	16.274
RBF neural network model	Absolute error	-0.32	0.556	0.595	0.652	1.05	0.623	0.526
	Relative error/%	-2.043	3.524	3.681	4.044	6.209	3.662	3.180
	Extract value	15.969	15.723	15.76	15.798	16.001	16.16	15.902
tandem gray RBF neural network	Absolute error	-0.011	0.503	0.19	0.328	0.141	-0.23	0.154
	Relative error/%	-0.069	3.199	1.206	2.076	0.881	-1.423	0.978
Parallel gray RBF neural network model	Extract value	16.013	15.719	15.667	15.648	16.189	16.529	15.961
	Absolute error	0.033	0.499	0.097	0.178	0.329	0.139	0.213
	Relative error/%	0.206	3.175	0.619	1.138	2.032	0.841	1.335

Embedded gray RBF neural network model	Extract value	23.034	20.622	17.538	15.321	16.027	17.751	18.382
	Absolute error	7.054	5.402	1.968	-0.149	0.167	1.361	2.634
	Relative error/%	30.624	26.195	11.221	-0.973	1.042	7.667	12.630
Single input layer embedded gray RBF neural network model	Extract value	15.846	15.058	15.376	15.24	15.59	16.074	15.531
	Absolute error	-0.134	-0.162	-0.194	-0.23	-0.27	-0.316	-0.218
	Relative error/%	-0.85	-1.08	-1.26	-1.51	-1.73	-1.97	-1.40

## 4 Conclusion

Automatic knowledge acquisition can effectively save the cost of knowledge acquisition, reduce the time consumption, improve the access efficiency, and become the improvement direction of KM knowledge acquisition "bottleneck". Gray RBF neural network model combines the characteristics of gray system "weakening data randomness, better revealing the data state" and RBF neural network has higher self-learning and self-adaptability; Different tacit knowledge acquisition environments provide flexible support. Knowledge explicitization is the basic activity of SECI helix, which is helpful to deepen knowledge cognition and benefit knowledge sharing, application and innovation benefit. Follow-up research direction.

## Acknowledgement

This work is supported by the Excellent Young Teacher Development Fund Project of Zhengzhou University (Foundation No. 2015SKYQ15).

## References

1. Joost Breuker. A Cognitive Science Perspective on Knowledge Acquisition. *International Journal of Human-Computer Studies*, 2013, 71(2): 177-183.
2. Bruno Cassiman, Reinhilde Veugelaers. In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition. *Management Science*, 2006, 52(1): 68-82.
3. Wu Quan-yuan, Liu Jiang-ning. *Artificial Intelligence and Expert System*. Beijing: China Railway Publishing House, 1995.
4. WU Bao-gui, DING Zhen-guo. Research on Distributed Search Engine Based on Map / Reduce. *New Technology of Library and Information Service*, 2007 (8): 52-55.
5. JIANG Jian-hong. Research on the Technology of Distributed Distributed Search Engines. *Science Technology and Engineering*, 2007, 7 (10): 2418-2424.
6. Kaj U. Koskinena, Pekka Pihlanto, Hannu Vanharanta. Tacit knowledge acquisition and sharing in a project work context. *International Journal of Project Management*, 2003, 21(4): 281-290.
7. Zhang Jianhua, *E-government knowledge management*. Beijing: Science Press, 2010.
8. Ikujiro Nonaka, Ryoko Toyama, Noboru Konno. SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation. *Long Range Planning*, 2000, 33(1): 5-34.
9. LIU Si-feng, DENG Ju-long. Application of GM (1,1) model. *Systems Engineering -Theory & Practice*, 2000 (05): 121-124.
10. Yang Hualong, Liu Jinxia, Zheng Bin. Improvement and application of grey prediction GM (1,1) model. *mathematics practice and understanding*, 2011, 41 (23): 39-46.
11. CHANG Yu-qing, ZHANG Hong-yan, WANG Shu. Injection Speed Control Based on RBF Neural Network Inverse System. *Journal of Northeastern University (Natural Science)*, 2013, 34 (2): 174-177.
12. YUAN Jing-ling, ZHONG Luo, LI Xiao-yan. Study and development of gray neural network. *Journal of Wuhan University of Technology*, 2009, 31 (3): 91-93.
13. Xing X. Gray neural network improved algorithm and its application. *Huazhong University of Science and Technology Master's thesis*, 2011.
14. ZHANG Yi, YANG Jianguo. Thermal error modeling of machine tools based on gray neural network. *Journal of Shanghai Jiaotong University*, 2011, 45 (11): 1581-1586.