ELSEVIER   Science Press

Real estate appraisal system based on GIS and BP neural network

LIU Xiao-sheng[1], DENG Zhe[1], WANG Ting-li[2]

1. School of Architecture and Survey Engineering, Jiangxi University of Science and Technology,
Ganzhou 341000, China;
2. School of Applied Science, Jiangxi University of Science and Technology, Ganzhou 341000, China

Abstract: For the inefficiency and inaccuracy of appraisal method of traditional estate appraisal theory, the real estate appraisal system based on GIS and BP neural network was established. The structure of the system was designed which includes appraisal model, trade case, GIS database and query analysis module. With the help of the L−M algorithm in MATLAB software, BP neural network was improved and the trade cases were trained, then the BP neural network which has already been trained was tested. At the same time, the BP neural and GIS were put together to construct the hedonic price estimate model. The C# and ArcGIS9.3 were used to achieve the system in VS2008. City basic geographic data and real estate related information were used as the basic data in practice. The results show that the functions of querying, adding and editing the spatial data and attribute data are achieved and also the efficiency and accuracy of real estate are improved, so that the new method of real estate is provided by the system.
Key words: BP neural network; GIS; hedonic price; real estate appraisal

1 Introduction

The real estate appraisal is a very complex process. The accuracy of the appraisal is influenced by the accuracy of the obtained real estate information, the technology of appraisal staff and the size of errors caused by various factors in the process of appraisal [1−3]. These factors lead to imperfection of current real estate appraisal theory and inefficiency and inaccuracy of appraisal method. For this situation, it is necessary to construct a new way of appraisal. Firstly, GIS with its large database technology, super computer graphic processing and spatial analysis capability, provides a comprehensive and accurate information technology platform for real estate appraisal industry [4−6], and then neural network has the dynamic processing ability in nonlinear problem [7−9]. It is not necessary to preinstall the distribution of data, rules among variable or precise mathematical model. The prediction function like the "black box" principle, the input and output data for learning and training thereby avoiding the subjective factors in determining the impact of weight can be achieved and the appraisal workload can be reduced.

Thus the appraisal models were set up by using BP neural network and hedonic price theory, and integrating with GIS and establishing real estate appraisal system based on GIS and BP neural network.

2 System design

The system will combine BP neural network appraisal model with GIS technology, then the system construction idea is as follows. Firstly, collect a large number of real estate transaction cases information and geophysics space, attribute information, which are storaged and managed by GIS. Secondly, use mapping analysis function of Arc Map to quantify characteristic values of auxiliary housing. Then, construct BP neural network on MATLAB software with the quantified Eigen value which is already valued and construct the hedonic price valuation model through trained BP neural network. Finally, invoke this appraisal model for real estate price appraisal in the GIS-integrated hedonic price model [10−11]. The system overall design is shown in Fig. 1.

In Fig. 1, the transaction case data mainly storage information which is related with appraisal model, such as hedonic price model, the trained BP neural network
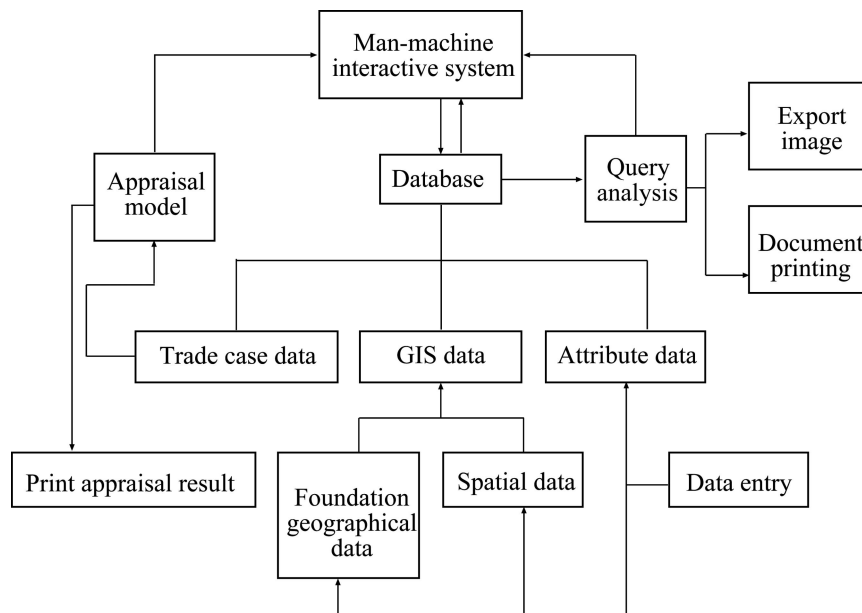
Fig. 1 Structure of real estate appraisal system based on GIS and BP neural network

parameters; GIS data and attribute data; primary storaged geographical data; social data; economic data which are related with price appraisal, such as urban basic maps, facilities maps (education, business, transportation and roads, etc), trading cases and the project properties which have not been estimated, spatial data, etc. Query analysis is mainly aimed at querying information on trading case, assisting auxiliary appraisal, making analysis in unestimated surrounding circumstances property etc. After appraisal process, the users can archive and print the appraisal report and also update the estate price system information through inputting required data besides existing data.

3 Model construction

3.1 BP neural network structure

Back-propagation artificial network (BP network) is a multilayer feed forward neural network, which consists of input layer, hidden layer and output layer. BP network structure is shown in Fig. 2.

3.2 BP neural network design

BP neural network design includes input/output variable handling, network topology designing and network training parameters determining [12−13].
3.2.1 Input/output variable handling

Manipulation parameters of input layer and output layer which include input layer parameter normalization and output layer parameter de-normalization. It means that the parameter value should be controlled in the range of −1 to 1.

$$\overline{X} = \frac{2(X - X_{\min})}{X_{\max} - X_{\min}} - 1 \tag{1}$$

where X, $X_{\max}$, $X_{\min}$ and $\overline{X}$ are represented as value in handle sample, the max and min value in sample and the normalization data, respectively.
3.2.2 Network topology designing

1) Determine the variable of input layer and output layer. Input variables must follow two principles: the first one is input variable which has the largest effect on output variable and also can be detected; the second one is non-linear relation between input variables. Eleven price influencing factors are chosen as the network input layer variables. Each bargain price is the output variable.

2) Determine the number of network layer. The three-layer BP neural network structures include input layer, hidden layer and output layer because the function of reducing errors, raising network training efficiency and avoiding network size is too complicated.

3) Determine the number of nodes in hidden layer. The key of network training is to determine the number of nodes in hidden layer.

$$p = \sqrt{n+m} + \alpha \tag{2}$$

where p stands for the number of hidden nodes; n stands for the number of input layer nodes; m stands for the number of output layer nodes; α is the constant in the range from 1 to 10. Firstly, the formula was used to estimate the hidden nodes and then the number of nodes was chosen in small error of the corresponding network by training those nodes. The number of nodes is determined as 12 through this way.
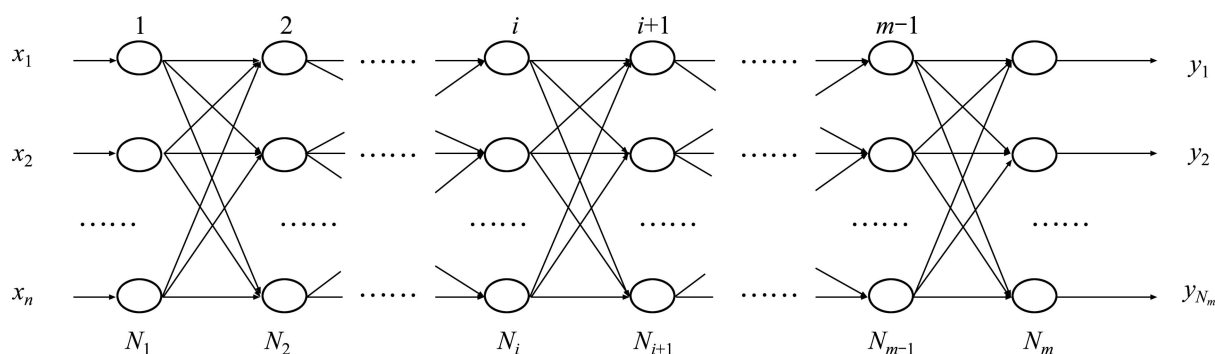
Fig. 2 BP neural network structure

3.2.3 Setting of network training parameter

1) Initial weights setting

The initial value of the network has great influence on network training time, network convergence and local minimum. If the selected initial value of the network is too large, it will make weighted input into the saturated zone of activation function $f(x)$, which leads to small derivative values. And due to the weight adjustment formula, $f'(x)$ and δ are direct proportional, when the derivative $f'(x)$ tends to 0, δ also will tend to 0. It makes ω too close to 0, then the adjustment process could be stopped. Based on the above analysis, the $[-1, 1]$ random number is selected as network initial weights

2) Determination of training times

The training times of neural network directly affect the network training time. If the network is overtrained (over fitting), it will affect the generalization ability of the network, that is, after training, the network will make the error response to the sample that is out of the scope of training. Therefore, the sample data can be divided into two groups during network training, including the training samples and test samples. Let two group samples train alternatively, when the training times reach a certain number, the error of two samples would change from decrease at the same to one continues to decrease while the other increase, and this training times are required for the best training times. Trained sample data and test samples were trained and compared for many times, finally the largest training times of network is 1 000.

3) Selection of expected error

In the network training process, the expected error is an appropriate value after comparing and training for many times. Appropriate value refers to a smaller expected error and a larger number of nodes in the hidden layer and extend the training time. Usually two different expected error networks were trained and then various factors were considered to select one of them as the expected error. According to the actual training situation, 0.001 was selected for the expected error of network training

3.2.4 BP network promotion algorithm: L−M algorithm

Levenberg−Marquardt (L−M) algorithm is actually a combination-type algorithm of gradient descent method and the quasi-Newton method. Gradient descent method has fast descent speed in the first few steps, but is closer to the optimal value, and slowly declines in the error function. The quasi-Newton method is complicated. Therefore, L−M method combines the advantages of these two method and exclude their disadvantages.

3.3 BP neural network hedonic price appraisal model construction

Three-layer BP network of a single hidden layer was selected. The dimensions of input vector p and the output vector t are 11 and 1, respectively. Input vector was normalized by Premnmx function in order to improve network speed. With a random function initff, input layer to hidden layer weights, threshold −$w1$, $b1$, and the hidden layer to output layer weights, threshold −$w2$, $b2$ are initialized. Also, the following network parameters were set: the maximum expected error err_goal, the maximum number of iterations max_epoch and display frequency dis_freq. Trainlm function into network training and simuff function into network simulation were make use of. Finally, regression analyzes the network results by using postreg function, and then tests the network accuracy. In the training process, the initial matrix of the weight and threshold of each layer was saved as init.mat file, and the trained initial matrix of the weight and threshold of each layer was saved as netbp.mat file. Finally, the related functions of MATLAB neural network toolbox were used to carry on writing programming method, while the L−M algorithm was used to improve the standard BP network algorithm, for constructing the real estate appraisal model of BP network [14−15]. The system function module design is shown in Fig. 3.
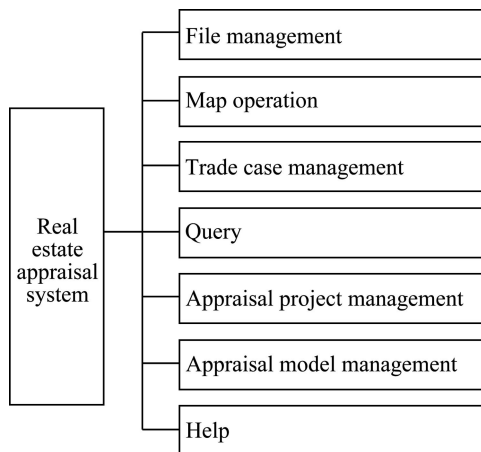
Fig. 3 System function module design

4 System development and implement

ArcGIS9.3 was chosen as the platform and C# language and geodatabase were selected to manage spatial data and attribute information [16]. The real estate appraisal system was developed based on GIS and BP neural network technology with VS2008. The main interface is shown in Fig. 4.

The system includes file management, map operation, trade case management, appraisal project management, appraisal model management and query function. The function of BP neural network appraisal is shown in Fig. 5.
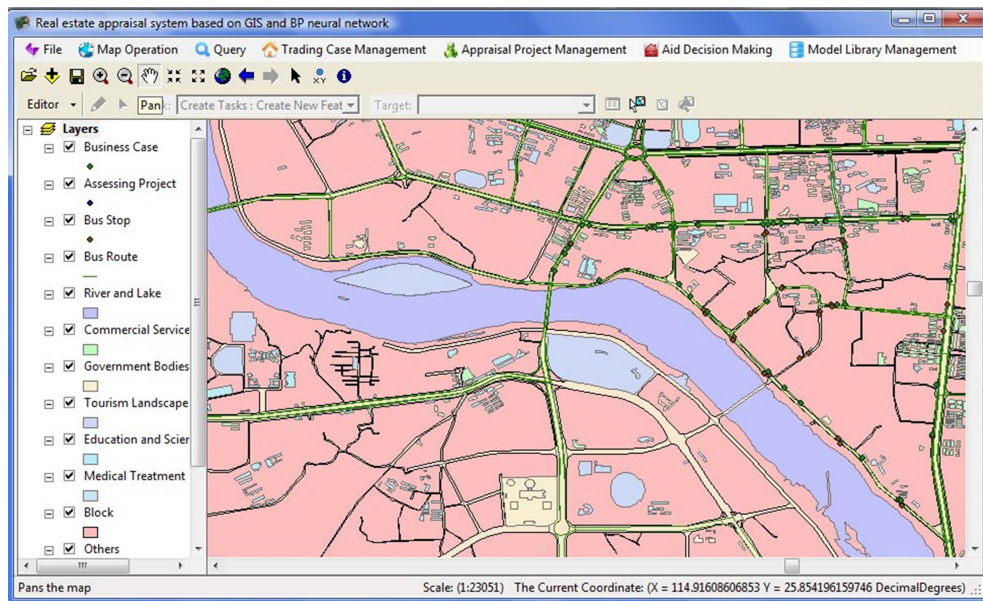

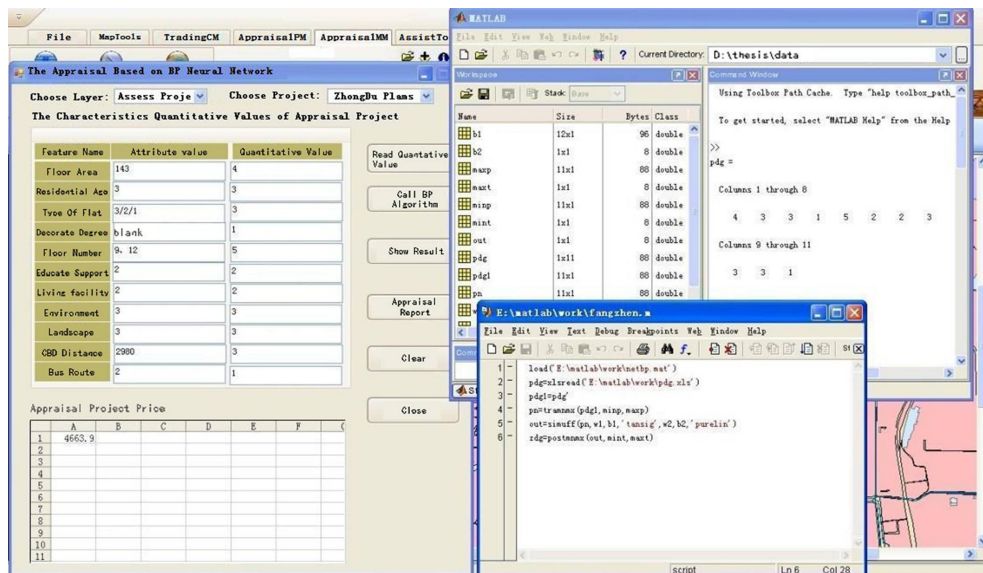**Fig. 4** Main interface of system

Fig. 5 BP neural network appraisals

6 Conclusions

For the question of traditional estate appraisal theory, inefficiency and inaccuracy of appraisal method and so on, first, the system quantifies the real estate characteristic factors by the specific quantitative criteria. The BP neural network is designed through the methods of set the network topology, the network training parameters are determined, the network hidden nodes are selected, the BP network algorithm is improved, and Matlab software is used to achieve the network training, verification and simulation functions to construct the valuation model base on BP neural network.

2) The system design is introduced. The system uses a city basic geographic data and real estate related information as the basic data to practice. The system achieves the functions of document management, map operation, trading case management, valuation model management, support tools, and query and also improves the efficiency and accuracy of real estate, so that the system provides the new method of real estate.

References

[1] GAO Bing-hua. The real estate appraisal [M]. Wuhan: Huazhong University of Science and Technology Press, 2006: 20−49. (in Chinese)

[2] GUO Bing. Theory and practice of real estate appraise [M]. Beijing: Chemical Industry Press, 2008: 3−4. (in Chinese)

[3] ZHU Ping-heng. Theory and practice of real estate appraise [M]. Dalian: Dongbei University of Finance and Economics Press, 2007: 134−145. (in Chinese)

[4] GUO Da-zhi. Principle and application of GIS [M]. Xuzhou: China University of Mining and Technology, 2002: 34−90. (in Chinese)

[5] WU Xin-cai. Principles and methods of GIS [M]. Beijing: Electronic Industry Press, 2002: 50−69. (in Chinese)

[6] WANG Da-seng, ZHU Da-ming, TONG Chun-ya. Real estate appraisal system based on GIS [J]. Surveying and Mapping, 2006, 22(1): 16−18. (in Chinese)

[7] HOU Yuan-bing, DU Jing-yi, WANG Mei. Neural network [M]. Xi'an: Xidian University Press, 2007: 88−107. (in Chinese)

[8] YAN Ping-fan, ZHANG Chang-shui. The man-made neural network and simulate evolution calculate [M]. Beijing: Tsinghua University Press, 2005: 20−56. (in Chinese)

[9] HAN Li-qun. Artificial neural network theory, design and application [M]. Beijing: Chemical Industry Press, 2002: 77−98. (in Chinese)

[10] GU Xiao-yu, WANG Wei-an. The real estate appraisal system based on GIS [J]. Microcomputer applications, 2003, 19(10): 20−30. (in Chinese)

[11] TIAN Yu, ZHENG Wen-hua. GIS in Real estate management information system study [J]. Construction Management Modernization, 2007, 19(2): 48−50. (in Chinese)

[12] YANG Li-meng, LIU Kai-di. BP neural network application in the real estate appraisal [J]. Hebei Architectural Science and Technology, 2004, 21(2): 105−108. (in Chinese)

[13] WANG Xi, LI Wei. Artificial neural networks in the application of real estate appraisal [J]. Shanxi Construction, 2010, 36(21): 23−24. (in Chinese)

[14] JIA Sheng-hua, WEN Hai-zheng. Hedonic price model of real estate development and application of the theory [J]. Foreign Economics and Management, 2004, 26(5): 42−44. (in Chinese)

[15] ZHOU Li-ping, LI Hui-ming. GIS-based hedonic price model of real estate [J]. Xi'an University of Architecture Technology, 2008, 27(2): 21−34. (in Chinese)

[16] ZHOU Chang-fa. Programming C# numerical algorithms [M]. Beijing: Electronic Industry Press, 2007: 10−89. (in Chinese)

(Edited by HE Yun-bin)